International Journal on Advances in Internet Technology

2020 vol. 13 nr. 1&2

The *International Journal on Advances in Internet Technology* is published by IARIA.
ISSN: 1942-2652
journals site: http://www.iariajournals.org
contact: petre@iaria.org

Sponsored by IARIA
www.iaria.org

International Journal on Advances in Internet Technology
Volume 13, Number 1 & 2, 2020

Ruay-Shiung Chang, National Dong Hwa University, Taiwan
Tzung-Shi Chen, National University of Tainan, Taiwan
Xi Chen, University of Washington, USA
IlKwon Cho, National Information Society Agency, South Korea
Andrzej Chydzinski, Silesian University of Technology, Poland
Noël Crespi, Telecom SudParis, France
Antonio Cuadra-Sanchez, Indra, Spain
Javier Cubo, University of Malaga, Spain
Sagarmay Deb, Central Queensland University, Australia
Javier Del Ser, Tecnalia Research & Innovation, Spain
Philipe Devienne, LIFL - Université Lille 1 - CNRS, France
Kamil Dimililer, Near East Universiy, Cyprus
Martin Dobler, Vorarlberg University of Applied Sciences, Austria
Jean-Michel Dricot, Université Libre de Bruxelles, Belgium
Matthias Ehmann, Universität Bayreuth, Germany
Tarek El-Bawab, Jackson State University, USA
Nashwa Mamdouh El-Bendary, Arab Academy for Science, Technology, and Maritime Transport, Egypt
Mohamed Dafir El Kettani, ENSIAS - Université Mohammed V-Souissi, Morocco
Armando Ferro, University of the Basque Country (UPV/EHU), Spain
Anders Fongen, Norwegian Defence Research Establishment, Norway
Giancarlo Fortino, University of Calabria, Italy
Kary Främling, Aalto University, Finland
Steffen Fries, Siemens AG, Corporate Technology - Munich, Germany
Ivan Ganchev, University of Limerick, Ireland / University of Plovdiv “Paisii Hilendarski”, Bulgaria
Shang Gao, Zhongnan University of Economics and Law, China
Emiliano Garcia-Palacios, ECIT Institute at Queens University Belfast - Belfast, UK
Kamini Garg, University of Applied Sciences Southern Switzerland, Lugano, Switzerland
Rosario Giuseppe Garroppo, Dipartimento Ingegneria dell'informazione - Università di Pisa, Italy
Thierry Gayraud, LAAS-CNRS / Université de Toulouse / Université Paul Sabatier, France
Christos K. Georgiadis, University of Macedonia, Greece
Katja Gilly, Universidad Miguel Hernandez, Spain
Mariusz Głąbowski, Poznan University of Technology, Poland
Feliz Gouveia, Universidade Fernando Pessoa - Porto, Portugal
Kannan Govindan, Crash Avoidance Metrics Partnership (CAMP), USA
Bill Grosky, University of Michigan-Dearborn, USA
Jason Gu, Singapore University of Technology and Design, Singapore
Christophe Guéret, Vrije Universiteit Amsterdam, Nederlands
Frederic Guidec, IRISA-UBS, Université de Bretagne-Sud, France
Bin Guo, Northwestern Polytechnical University, China
Gerhard Hancke, Royal Holloway / University of London, UK
Arthur Herzog, Technische Universität Darmstadt, Germany
Rattikorn Hewett, Whitacre College of Engineering, Texas Tech University, USA
Quang Hieu Vu, EBTIC, Khalifa University, Arab Emirates
Hiroaki Higaki, Tokyo Denki University, Japan
Dong Ho Cho, Korea Advanced Institute of Science and Technology (KAIST), Korea
Anna Hristoskova, Ghent University - IBBT, Belgium
Ching-Hsien (Robert) Hsu, Chung Hua University, Taiwan
Chi Hung, Tsinghua University, China
Edward Hung, Hong Kong Polytechnic University, Hong Kong
Raj Jain, Washington University in St. Louis , USA
Edward Jaser, Princess Sumaya University for Technology - Amman, Jordan
Terje Jensen, Telenor Group Industrial Development / Norwegian University of Science and Technology, Norway
Yasushi Kambayashi, Nippon Institute of Technology, Japan

Georgios Kambourakis, University of the Aegean, Greece
Atsushi Kanai, Hosei University, Japan
Henrik Karstoft , Aarhus University, Denmark
Dimitrios Katsaros, University of Thessaly, Greece
Ayad ali Keshlaf, Newcastle University, UK
Reinhard Klemm, Avaya Labs Research, USA
Samad Kolahi, Unitec Institute Of Technology, New Zealand
Dmitry Korzun, Petrozavodsk State University, Russia / Aalto University, Finland
Slawomir Kuklinski, Warsaw University of Technology, Poland
Andrew Kusiak, The University of Iowa, USA
Mikel Larrea, University of the Basque Country UPV/EHU, Spain
Frédéric Le Mouël, University of Lyon, INSA Lyon / INRIA, France
Juong-Sik Lee, Nokia Research Center, USA
Wolfgang Leister, Norsk Regnesentral (Norwegian Computing Center), Norway
Clement Leung, Hong Kong Baptist University, Hong Kong
Longzhuang Li, Texas A&M University-Corpus Christi, USA
Yaohang Li, Old Dominion University, USA
Jong Chern Lim, University College Dublin, Ireland
Lu Liu, University of Derby, UK
Damon Shing-Min Liu, National Chung Cheng University, Taiwan
Michael D. Logothetis, University of Patras, Greece
Malamati Louta, University of Western Macedonia, Greece
Maode Ma, Nanyang Technological University, Singapore
Elsa María Macías López, University of Las Palmas de Gran Canaria, Spain
Olaf Maennel, Loughborough University, UK
Zoubir Mammeri, IRIT - Paul Sabatier University - Toulouse, France
Yong Man, KAIST (Korea advanced Institute of Science and Technology), South Korea
Sathiamoorthy Manoharan, University of Auckland, New Zealand
Chengying Mao, Jiangxi University of Finance and Economics, China
Brandeis H. Marshall, Purdue University, USA
Constandinos Mavromoustakis, University of Nicosia, Cyprus
Shawn McKee, University of Michigan, USA
Stephanie Meerkamm, Siemens AG in Erlangen, Germany
Kalogiannakis Michail, University of Crete, Greece
Peter Mikulecky, University of Hradec Kralove, Czech Republic
Moeiz Miraoui, Université du Québec/École de Technologie Supérieure - Montréal, Canada
Shahab Mokarizadeh, Royal Institute of Technology (KTH) - Stockholm, Sweden
Mario Montagud Climent, Polytechnic University of Valencia (UPV), Spain
Stefano Montanelli, Università degli Studi di Milano, Italy
Julius Müller, TU- Berlin, Germany
Juan Pedro Muñoz-Gea, Universidad Politécnica de Cartagena, Spain
Krishna Murthy, Global IT Solutions at Quintiles - Raleigh, USA
Alex Ng, University of Ballarat, Australia
Christopher Nguyen, Intel Corp, USA
Petros Nicopolitidis, Aristotle University of Thessaloniki, Greece
Carlo Nocentini, Università degli Studi di Firenze, Italy
Federica Paganelli, CNIT - Unit of Research at the University of Florence, Italy
Carlos E. Palau, Universidad Politecnica de Valencia, Spain
Matteo Palmonari, University of Milan-Bicocca, Italy
Ignazio Passero, University of Salerno, Italy
Serena Pastore, INAF - Astronomical Observatory of Padova, Italy
Fredrik Paulsson, Umeå University, Sweden
Rubem Pereira, Liverpool John Moores University, UK

Yulia Ponomarchuk, Far Eastern State Transport University, Russia
Jari Porras, Lappeenranta University of Technology, Finland
Neeli R. Prasad, Aalborg University, Denmark
Drogkaris Prokopios, University of the Aegean, Greece
Emanuel Puschita, Technical University of Cluj-Napoca, Romania
Lucia Rapanotti, The Open University, UK
Gianluca Reali, Università degli Studi di Perugia, Italy
Jelena Revzina, Transport and Telecommunication Institute, Latvia
Karim Mohammed Rezaul, Glyndwr University, UK
Leon Reznik, Rochester Institute of Technology, USA
Simon Pietro Romano, University of Napoli Federico II, Italy
Michele Ruta, Technical University of Bari, Italy
Jorge Sá Silva, University of Coimbra, Portugal
Sébastien Salva, University of Auvergne, France
Ahmad Tajuddin Samsudin, Telekom Malaysia Research & Development, Malaysia
Josemaria Malgosa Sanahuja, Polytechnic University of Cartagena, Spain
Luis Enrique Sánchez Crespo, Sicaman Nuevas Tecnologías / University of Castilla-La Mancha, Spain
Paul Sant, University of Bedfordshire, UK
Brahmananda Sapkota, University of Twente, The Netherlands
Alberto Schaeffer-Filho, Lancaster University, UK
Peter Schartner, Klagenfurt University, System Security Group, Austria
Rainer Schmidt, Aalen University, Germany
Thomas C. Schmidt, HAW Hamburg, Germany
Zary Segall, Chair Professor, Royal Institute of Technology, Sweden
Dimitrios Serpanos, University of Patras and ISI/RC ATHENA, Greece
Jawwad A. Shamsi, FAST-National University of Computer and Emerging Sciences, Karachi, Pakistan
Michael Sheng, The University of Adelaide, Australia
Kazuhiko Shibuya, The Institute of Statistical Mathematics, Japan
Roman Y. Shtykh, Rakuten, Inc., Japan
Patrick Siarry, Université Paris 12 (LiSSi), France
Jose-Luis Sierra-Rodriguez, Complutense University of Madrid, Spain
Simone Silvestri, Sapienza University of Rome, Italy
Vasco N. G. J. Soares, Instituto de Telecomunicações / University of Beira Interior / Polytechnic Institute of Castelo Branco, Portugal
Radosveta Sokullu, Ege University, Turkey
José Soler, Technical University of Denmark, Denmark
Victor J. Sosa-Sosa, CINVESTAV-Tamaulipas, Mexico
Dora Souliou, National Technical University of Athens, Greece
João Paulo Sousa, Instituto Politécnico de Bragança, Portugal
Kostas Stamos, Computer Technology Institute & Press "Diophantus" / Technological Educational Institute of Patras, Greece
Cristian Stanciu, University Politehnica of Bucharest, Romania
Vladimir Stantchev, SRH University Berlin, Germany
Tim Strayer, Raytheon BBN Technologies, USA
Masashi Sugano, School of Knowledge and Information Systems, Osaka Prefecture University, Japan
Tae-Eung Sung, Korea Institute of Science and Technology Information (KISTI), Korea
Sayed Gholam Hassan Tabatabaei, Isfahan University of Technology, Iran
Yutaka Takahashi, Kyoto University, Japan
Yoshiaki Taniguchi, Kindai University, Japan
Nazif Cihan Tas, Siemens Corporation, Corporate Research and Technology, USA
Alessandro Testa, University of Naples "Federico II" / Institute of High Performance Computing and Networking (ICAR) of National Research Council (CNR), Italy
Stephanie Teufel, University of Fribourg, Switzerland

International Journal on Advances in Internet Technology
Volume 13, Numbers 1 & 2, 2020

CONTENTS

James Okolica, Air Force Institute of Technology, UnitedStates
Dillon Pettit, Air Force Institute of Technology, Unites States
Marvin Newlin, Air Force Institute of Technology, United States

pages: 83 - 96

On Heterogeneity of Management and Orchestration Functional Architectures in 5G Slicing

Eugen Borcoci, University POLITEHNICA of Bucharest - UPB, Romania
Cosmin Contu, University POLITEHNICA of Bucharest - UPB, Romania
Andra Ciobanu, University POLITEHNICA of Bucharest - UPB, Romania

Identifying and Analyzing Obscure Venues Using Obscure Words in User-provided Reviews

Masaharu Hirota
Faculty of Informatics
Okayama University of Science
Okayama-shi, Okayama
Email: hirota@mis.ous.ac.jp

Jhih-Yu Lin
Graduate school of System Design
Tokyo Metropolitan University
Hino-shi, Tokyo
Email: lin-jhihyu@ed.tmu.ac.jp

Masaki Endo
Division of Core Manufacturing
Polytechnic University
Kodaira-shi, Tokyo
Email: endou@uitec.ac.jp

Hiroshi Ishikawa
Graduate school of System Design
Tokyo Metropolitan University
Hino-shi, Tokyo
Email: ishikawa-hiroshi@tmu.ac.jp

***Abstract*—When sightseeing, many people visit different places such as restaurants, hotels, and tourist spots. Some of these venues, while worthwhile, are considered obscure, secret, not well-known, or having little popularity. Their extraction and recommendation are vital to improving the satisfaction of tourists. This research proposes a method for discovering obscure venues using classifiers for identifying reviews, including obscure impressions. To achieve this goal, in this research, a model was developed to classify venues as obscure or not obscure using reviews with language indicating their obscurity. In addition, we compare various methods for generating feature vectors and the models for classification. This research also analyzes the differences among venues perceived by reviewers as being obscure. We demonstrate the performance of the proposed approach by indicating that the posting destination of obscure reviews differs for each user.**

Keywords–Tourism information; Text classification; Support Vector Machine; Review Analysis.

I. INTRODUCTION

A considerably shorter pre-version of this paper has already been published in [1].

In recent years, it has become commonplace for many people to give their opinions and impressions regarding several spots as tourist spots, hotels, and restaurants, on review websites such as Yelp [2], Expedia [3], and TripAdvisor [4]. In this paper, we call such spots venues. Reviews written about venues describe information regarding the venues themselves and the impressions to them and behaviors of the users. Such reviews are useful for travel planning, obtaining information on travel destinations, tourist behavior, and visitor impressions of popular tourist spots. Therefore, many studies have extracted tourism information from user-provided reviews [5][6].

Some venues are obscure, secret, not well-known, or having little popularity. Despite not being popular, such venues may be well-regarded by visitors. In this paper, these are collectively called "obscure". Because some obscure venues can lead to improved tourist satisfaction and the acquisition of repeat visitors, some methods for describing obscure venues and recommending them to tourists have been proposed [7] [8]. Definitions regarding obscure venues have been proposed in such studies. Studies on this subject commonly define an obscure venue as one in which the visibility for tourists is low, but the value is high. For example, the authors in [7] defined obscure spots as less known, but still worth visiting, and extracted such spots. Also, [9] extracted hidden tourist spots with low popularity but a high level of satisfaction. However, precisely identifying obscure venues is difficult because the places that people feel are obscure depends on their own personality.

In this research, we identify obscure venues from review sites, and the proposed approach focuses on words in the text of the venue reviews. This research then extracts obscure reviews without directly defining obscure to accommodate the fact that the impression of a venue differs among different people. For this research, we regard a venue with many reviews written about the impression of its obscurity as an obscure venue (hereinafter referred to as "obscure review". Also, we call other reviews "non-obscure review").

This research extracted such reviews from all reviews on a particular venue. In this paper, a review is defined as an obscure review if its text contains terms related to "obscure" (hereinafter referred to as "obscure words"). If the ratio of reviews of a venue that includes obscure words accounts for the majority, the venue is defined as obscure.

The aim of this research is the identification of obscure venues using user-provided reviews that include obscure words. However, in most cases, the number of reviews on a venue is small. Because an obscure venue might be less well-known by people even if worthwhile, there will be few reviews for such venues. Also, few reviews obtain obscure words. As a result, the number of reviews to be classified as obscure is insufficient for identification of obscure venues. Moreover, it is unrealistic to define all expressions related to the word obscure. Therefore, to extract obscure reviews that do not include obscure words but rather the description of an obscure venue, this research applies the classification model of the representation of contents of a review as obscure or not, regardless of whether a review contains an obscure word. Reviews that do not contain obscure words were classified using the model, and the classifier was evaluated using a dataset of reviews submitted by users.

Moreover, different reviewers have posted various reviews

on different venues, and the criteria by which a venue is considered obscure differs according to the reviewer. Therefore, this research revealed that the reviewer who posts an obscure review for each venue is different. As a result, this research examined the efficiency of the proposed approach in identifying obscure venues using the obscure-word based classifier without a direct definition of the term obscure.

A summary of contributions from this research is as follows.

- We design a new approach for identifying obscure venues using user-provided reviews.
- We propose a classifier for identifying obscure reviews without the obscure words.
- We analyze the posting destination of obscure reviews differently for each user.

The remainder of this paper is organized as follows. Section II presents previous studies related to this topic. Section III describes our proposed method for the development of a classifier for discovering obscure reviews by using obscure words and the identification of obscure venues. Section IV describes the experiments evaluating our proposed method using the Yelp dataset. Section V describes an analysis of the hypothesis that an obscure venue is perceived differently for each user and discuss the extracted obscure reviews and venues. Section VI provides some concluding remarks along with a discussion of results and areas of future work.

II. Related Works

The main aim of our research was to find obscure venues for tourism analysis using user-provided reviews posted to social media sites. This section introduces the related studies published in the area of analysis of tourism information using reviews and extracting obscure venues. Also, vectorizing documents is an essential procedure for review analysis, because the performance of vectorization has massive effects on the classification of them. Therefore, we describe the related studies of document vectorization.

A. Analysis for tourism using reviews

Research has been conducted on the extraction of tourism information through user-generated content on social media sites. Also, extracting helpful or useful information from text data like reviews and blogs is one of the research tools used to analyze reviews. Our proposed research on extracting obscure venues from reviews is related to the analysis of reviews for recommendations and the analysis of tourism information.

[10] analyzed factors affecting the perceived usefulness of reviews to findings contributing to tourism marketers. [11] predicted where memorable is the travel destination using the user-generated photographs in blogs. [12] proposed a method for identifying dimensions of satisfaction using an unsupervised learning algorithm with numerical and textual information from user-generated online reviews, and analyzed the multiple factors contributing to consumer satisfaction. [13] predicted how helpful a review is and presented a list of ranked reviews based on an evaluation. [14] proposed a method for detecting reviews that reliably predict foodborne illnesses using review classification. [15] analyzed online review to identify insights through a case study, and found them. For example, overall review star rating correlates well with the sentiment scores for both the title and the full content of the online reviews. [16] proposed a method for detecting the topic of phrases in helpful recommending reviews. [17] proposed a method for aspect-based opinion mining of tourism reviews to classify them into negative or positive aspects. [18] proposed an approach for sentiment classification of online hotel booking opinions using a dependency tree structure. [19] investigated the valence of online reviews and modeled them with hotel attributes and performance. [20] analyzed the online reviewer profile, and exposed its image can significantly enhance consumers evaluation of review helpfulness. [21] concluded that sentiment analysis plays an important role in the analysis of tourism reviews and summarize their studies.

These studies analyzed user-provided reviews on social media sites for improving sightseeing satisfaction. This paper tackles the analysis of user perception of obscure venues based on reviews.

B. Extracting obscure venues from social media sites

Studies have been conducted on extracting obscure venues and tourist spots from social media sites. Because obscure spots are expected to spread tourists to other tourist spots and improve the satisfaction of the tourism experience, some studies extracting posts on such spots have been conducted.

[7] proposed a method for evaluating sightseeing spots that are less well-known but are worth visiting. [8] defined the term obscure to indicate spots that are not famous but have high evaluations, and extracted such spots based on name recognition and user evaluations. [9] proposed a method for providing tourism information on hidden spots for increasing tourism satisfaction. [22] extracted hot and cold spots based on a spatial analysis of user-generated content to extract knowledge of tourist behaviors. [23] proposed a method for less-known tourist attractions by using a clustering algorithm from geo-tagged photographs on Flickr.

This study used a classifier to extract obscure venues using reviews that include the word obscure to comprehensively deal with familiarity, popularity, and attractiveness. The main characteristic of this research is the extraction of sightseeing spots recognized by reviewers as obscure venues by using the classifier.

C. Document vectorization

Various methods have been proposed to vectorize documents. The traditional approach for vectorizing documents is some hand-craft features such as Term Frequency-Inverse Document Frequency (TFIDF), bag-of-words, and n-grams. Also, unsupervised representation learning have been used [24], [25], [26]. However, in recent years, pre-trained deep language representation model has been highly successful in the domain of Natural Language Processing (NLP) [27], [28], [29], [30]. Especially, Bidirectional Encoder Representations from Transformers (BERT) [27] achieved state-of-the-art performance on various NLP tasks.

To generate a document vector using BERT, the most commonly used approach is to average the BERT output layer or by using the output of the first token (the [CLS] token). However, this approach has been pointed out as unsuitable by [31]. Sentence-Bert [31] (SBERT) is a model that BERT [27] to derive semantically meaningful sentence embeddings that can be compared using cosine-similarity. Therefore, in this paper, we use SBERT for generating a document vector.

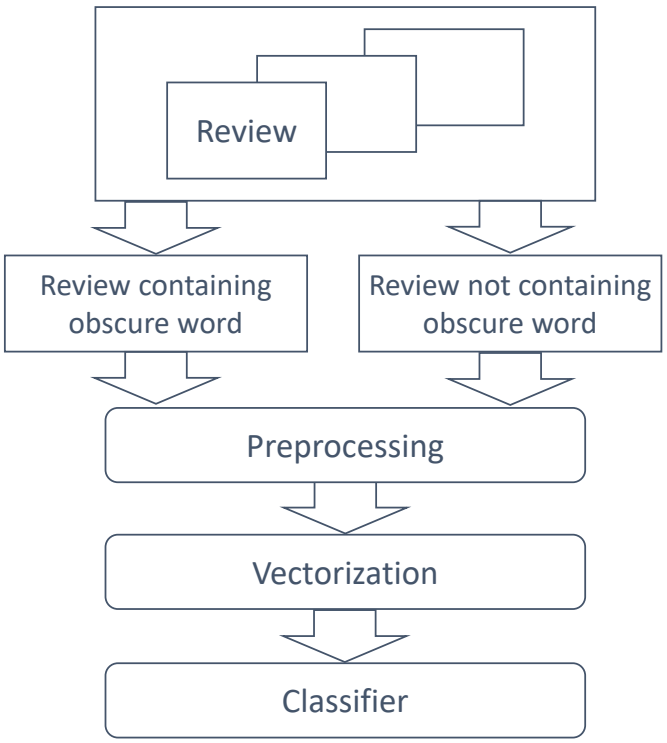


Figure 1. Overview of classifier for extracting obscure reviews using obscure words.

TABLE I. OBSCURE WORDS.

secret spot	secret place
best kept secret	best-kept secret
well-kept secret	well kept secret
local secret	obscure spot
hidden spot	hidden place
little known	little-known
good out of the way	

III. Proposed Method

In this section, we describe our proposed method for discovering obscure venues using obscure words and classification algorithms from user-provided reviews.

This research extracted reviews including obscure words and generated a classifier for both obscure and non-obscure reviews. We indicate an overview of our proposed classifier in Figure 1. First, we extract obscure and non-obscure reviews from reviews. Next, we apply the preprocessing method for the reviews. Next, we apply a vectorization method to generate a document vector. Finally, we create a model of the classifier using a vector to classify a review as obscure or not.

After this process, the classifier is applied to all reviews on a venue, and the venue is classified as obscure or non-obscure based on the reviews classified as obscure.

A. Obscure words

This section explains obscure words that we used for extracting obscure reviews.

In this research, obscure words are used to identify obscure venues from all reviews in a venue. This research defined 13 obscure words, as shown in Table I. The criterion for selecting obscure words is to select an English phrase manually that seems to represent a word indicating obscurity, and an expression that has no meaning other than obscurity.

However, these words do not cover all words expressing user perceptions of obscurity. Also, for example, phrases such as "little well known" can assume word choices and various spelling variations. Preparing all those phrases or words included in reviews is not realistic. However, it is desirable to extract obscure reviews from all of them that contain unknown obscure phrases or do not include those phrases. Therefore, we conduct supervised learning using obscure reviews including these words to discover obscure reviews not including them.

B. Preprocessing

This section describes the preprocessing applied to vectorize the reviews for machine learning.

First, reviews written in English were extracted from all reviews. In this paper, to detect the language of the texts we applied langdetect [32] to them.

Also, we extract reviews where the text has more than 30 words. This reason is because the classification is difficult when the number of words is small.

The texts from the extracted reviews were converted into lower-case texts. Next, we apply stop-word elimination and stemming to each word. This research defined 319 stop words, such as "the" and "and," which are commonly used in sentences.

C. Vectorization

Next, the preprocessed reviews were vectorized for determining what words in reviews might be more efficient for extracting obscure reviews. In this paper, we tried two vectorization methods. First, is TFIDF, which is one of the major hand-craft features. The other is SBERT, which is a pre-trained deep language representation model.

1) TFIDF: First, TFIDF were applied to the texts.

In this paper, we calculated the TFIDF of each word t in review r. The term frequency $tf(t,d)$ and inverse document frequency $idf(t,D)$ are calculated using the follow equations:

$$tf(t,r) = \frac{f_{t,r}}{\sum_{t \in r} f_{t,r}} \tag{1}$$

$$idf(t,R) = \log \frac{|R|}{|\{r \in R : t \in r\}|} \tag{2}$$

where the number of reviews is $|R|$, and $f_{t,r}$ is the number of occurrences of word t in review r.

Then, the TFIDF of each word t in review r in reviews R is calculated through the following equation:

$$tfidf(t,r,R) = tf(t,r) \times idf(t,r) \tag{3}$$

2) SBERT: SBERT is a modification of the pre-trained BERT network that uses siamese and triplet network structures to derive semantically meaningful sentence embeddings that can be compared using a similarity function.

SBERT uses the output of CLS-token or all output vectors from BERT. First, this model applies the pooling operation to the vector. In this paper, we adopt all output vectors and mean pooling. Also, to fine-tune BERT, SBERT used siamese networks to update the weights and the objective function is the following equation:

$$\omega = softmax(W_t(u, v, \|u - v\|)) \tag{4}$$

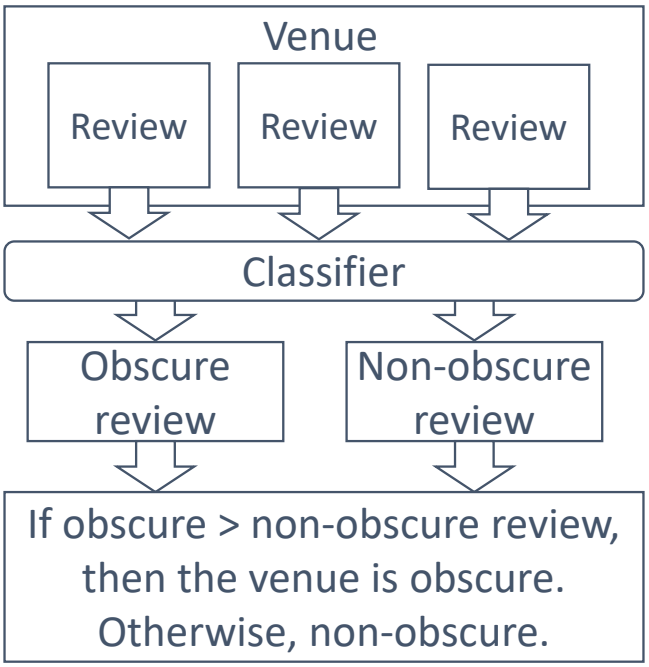


Figure 2. Overview of procedure for identification of obscure venues using obscure and non-obscure reviews.

Here, n is the dimension of the sentence embeddings and k is the number of labels, W is the weights of siamese network.

In this paper, we used the pre-trained model of SBERT using NLI models [31], in which this model was generated using the combination of two datasets [33] [34]. Next, the preprocessed reviews were vectorized for determining what words in reviews might be more efficient for extracting obscure reviews.

D. Classification of obscure reviews

In this section, we describe the procedure for generating a classification model of reviews regardless of whether they are obscure reviews. Our method proposed in this research identifies obscure venues using obscure reviews even if the review does not include obscure words. Therefore, our proposed method creates a classifier for identifying such reviews that do not include obscure words but when their content represents an obscure venue.

A method is proposed to classify the reviews into obscure or non-obscure reviews. In this research, we apply a binary classification method using vectors generated as described in Section III-C. The first class is thus obscure reviews, which consists of reviews that contain an obscure word. The other class is non-obscure reviews, which consists of reviews that do not contain an obscure word.

This research used three binary classification methods to classify reviews as obscure or not obscure: Support Vector Machine (SVM) [35], Random Forests (RF) [36] and LightGBM [37].

E. Identification of obscure venue

Herein, we describe how to find obscure venues using a classifier. Figure 2 shows an overview of the procedure for the identification of an obscure venue. We collect all review texts of a venue and apply the classifier described in Section III-D to the reviews. Finally, we count the reviews classified as obscure or non-obscure reviews of a venue. As a result, this research regards an obscure venue as one in which the percentage of obscure venues is greater than the threshold. In this paper, when the ratio of reviews classified as obscure among all reviews on a venue is larger than half, the venue is considered obscure, otherwise, it is non-obscure.

IV. Experiments of classification performance

In this paper, we evaluate the performance of our proposed method through an evaluation experiment based on classification. We describe the experimental conditions of the dataset and the evaluation criteria. Also, we describe our experiments conducted for the evaluation of obscure review discovery.

A. Dataset

Herein, we describe the dataset used for this experiment. We used the Yelp Dataset Challenge (round 13) [38], which includes 192,609 venues and 6,685,900 reviews which were written by 1,637,138 users. After we applied the preprocessing procedure as described in III-B, the number of reviews, venue, and users is 518,8614, 165,060 and 602,988, respectively.

This research comprises 1,780 reviews that mention an obscure word at least once. Table II shows the number reviews containing each obscure words. Here, we replaced the name of the venue into "@" to anonymize it. About 45% of reviews in the table contain the word "best kept secret". Therefore, this word is a general phrase for representing obscure venues. However, this table shows various words representing obscure venues are used. We used these reviews to generate a classifier for identifying a review as obscure or not. Also, we prepared the same number of randomly selected reviews from reviews which do not contain the obscure words.

The reviews with and without obscure words were randomly split into a ratio of 4:1 for training and testing data. As a result, the number of training data is 2,848 and testing data is 712.

B. Evaluation criteria

We used the following widely used performance measures for classification: Accuracy, Recall, Precision, and F-measure. To calculate them, we exploited the concepts of True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN), which shown in Table III. TP is the number of obscure reviews that are predicted as obscure. TN is the number of non-obscure reviews that are predicted as non-obscure. FP is the number of non-obscure reviews that are predicted as obscure. FN is the number of obscure reviews that are predicted as non-obscure. Using them, Accuracy, Recall, Precision, and F-measure are calculated as the following equations.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + TN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$F - measure = \frac{2 * Precision * Recall}{Precision + Recall} \tag{8}$$

TABLE II. THE NUMBER OF REVIEWS FOR EACH OBSCURE WORD AND EXAMPLES OF A PART OF REVIEWS

Obscure word	The number of reviews	Example
secret spot	106	I only gave 3 stars because I don't want you blowing up my secret spot!
secret place	112	It is a secret place that not even all locals know about, and the pizza is great.
best kept secret	789	As others have said, this place is one of the best kept secrets.
best-kept secret	71	@ it's a best-kept secret you only share with your close friends.
well kept secret	42	Overall, @ has been a well kept secret amongst those in the know for a long time.
well-kept secret	96	We enjoy coming here nonetheless. Maybe it's just a well-kept secret.
local secret	39	Even though many people knew about it, it still seemed like a local secret.
obscure spot	4	One of our favorite relaxed but obscure spots with the decor of an opium den slash western saloon.
hidden spot	155	It's in such a hidden spot you wouldn't know it was there unless you looked it up or saw people walking out of a hallway with a pizza box!
hidden place	124	I've been wanting to try this little hidden place for over a year now and finally found the time.
little known	188	If it's not on someone's list of high-quality, yet little known local-area sports bar destinations it should be.
little-known	43	I'd say it's the best little-known hard dip ice cream place in town.
good out of the way	11	First, let me get the good out of the way. The kids who got my order were nice, and the restaurant was clean.
sum	1,780	

TABLE III. CONFUSION MATRIX.

		Predict	
		Positive	Negative
Correct	Positive	TN	FN
	Negative	FP	TP

TABLE IV. EVALUATION RESULT: ACCURACY.

	RF	SVM	LightGBM
TFIDF	0.74	0.74	0.74
SBERT	0.75	0.77	0.73
Ave.	0.75	0.76	0.74

TABLE V. EVALUATION RESULT: PRECISION.

		RF	SVM	LightGBM
TFIDF	Obscure	0.76	0.76	0.76
	Non-obscure	0.73	0.74	0.74
	Ave.	0.75	0.75	0.75
SBERT	Obscure	0.78	0.82	0.75
	Non-obscure	0.73	0.74	0.73
	Ave.	0.76	0.78	0.74

TABLE VI. EVALUATION RESULT: RECALL.

		RF	SVM	LightGBM
TFIDF	Obscure	0.72	0.72	0.73
	Non-obscure	0.76	0.77	0.77
	Ave.	0.74	0.75	0.75
SBERT	Obscure	0.71	0.72	0.72
	Non-obscure	0.80	0.84	0.76
	Ave.	0.76	0.78	0.78

C. Experimental conditions

This section describes the procedure used for the creation of classifiers for obscure reviews.

This experiment used a Gaussian kernel for the SVM kernel function and entropy and Gini impurity for a split of nodes in Random Forest. In addition, the hyperparameters of those methods were searched using Optuna [39] with five cross-validations, which is a software for automatically optimizing hyperparameters. We used the parameters with the highest accuracy measured through this experiment. In addition, we used the Python software scikit-learn [40] for the implementation of the SVM, RF, TFIDF, and evaluation criteria in the following experiments. Also, we used [41] for the implementation of the LightGBM.

D. Evaluation results

In this section, we describe and discuss the evaluation results of classifying reviews into obscure or non-obscure reviews.

Table IV shows the evaluation results of the classification of obscure reviews through the procedure described above using accuracy. Also, Tables V, VI, and VII show the evaluation results of the classification of obscure reviews through the procedure described above using f-measure, precision, and recall. In those tables, "Obscure review" shows the reviews that include an obscure word, whereas "Non-obscure review" shows reviews that do not include an obscure word.

Also, in Tables IV, V, VI, and VII, comparing TFIDF and SBERT used for document vectorization, the evaluation scores of the SBERT is better than TFIDF in most cases. This reason is that the procedure for generating TFIDF is a simple way and does not consider the context and meaning of sentences, but SBERT uses a complex model considering them and can generate better feature vector. Also, the evaluation score of the combination of RF and LightGBM with TFIDF has often better performance than SBERT. As described in Section III-C1, a dimension in the vector generated by TFIDF shows the degree of appearance of one word in one document. On the other hand, the vector of SBERT is generated using neural networks and one vector does not have a specific role. Also, RF and LightGBM have functions for feature engineering such as feature bagging and exclusive feature bundling. Therefore, we think that those methods choice better dimensions from document vectors by themselves and showed better performance. However, SBERT generated a better feature vector in the overall evaluation.

Comparing the results shown in Tables VII, V, and VI for obscure and non-obscure reviews, the evaluation scores of the non-obscure reviews are lower than those of the obscure reviews. In particular, there is a vast difference between both scores regarding the recall rate. The evaluation score is achieved because reviews with an obscure word are misclassified as non-obscure in certain cases because the number of reviews in the training dataset is unbalanced. However, the purpose of this research is to identify obscure venues using extracted obscure reviews. As shown in Table V, the precision of the obscure reviews was 0.82 (the combination of SBERT and SVM), which shows that it is rare for a classifier to misclassify the content of reviews unrelated to obscurity.

In Table IV, the best combination of feature and classification methods is SVM and SBERT in almost cases. However,

TABLE VII. EVALUATION RESULT: F-MEASURE.

		RF	SVM	LightGBM
TFIDF	Obscure	0.74	0.75	0.75
	Non-obscure	0.75	0.74	0.75
	Ave.	0.75	0.75	0.75
SBERT	Obscure	0.75	0.76	0.74
	Non-obscure	0.76	0.79	0.74
	Ave.	0.76	0.78	0.74

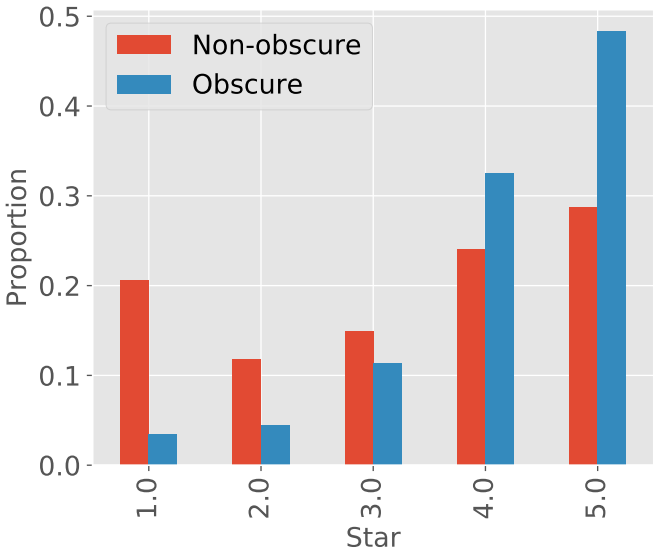


Figure 3. Distributions of proportion of stars on obscure and non-obscure reviews.

the difference in performance with other combinations is small. These results show that various document vectorization and classification methods can classify reviews into obscure and non-obscure. Therefore, our approach which uses the reviews containing obscure words to discover obscure reviews is effective.

E. Analysis of stars in obscure and non-obscure reviews

This section describes and discusses the difference of the stars in Yelp between obscure and non-obscure reviews. Here, the star represents the evaluation score, in which a reviewer evaluates a venue on a scale of 1 to 5 (1 = bad and 5 = good).

In general, we think that reviewers who wrote obscure reviews are considered to have a positive evaluation to the venue. Therefore, we can assume that the stars of obscure reviews are high. On the other hand, in the case of non-obscure, the reviewer wrote not only positive ratings to venues such as popular restaurants but also negative ratings, because they also wrote about those with a bad impression. Therefore, we can assume that the stars of non-obscure reviews are varied values. As a result, we believe that if distributions on stars of obscure and non-obscure are different and are similar to the above explanation, our classifier may classify reviews correctly.

Figure 3 shows the distributions of proportions of stars on obscure and non-obscure reviews. Here, blue bar shows the proportions of obscure reviews in each star value. Also, red bar shows the proportions of non-obscure reviews in each star value.

In Figure 3, the distributions of stars on obscure and non-obscure reviews are clearly different. The stars of obscure reviews are biased toward higher values. On the other hand, the stars of non-obscure reviews are evenly distributed. Therefore, Figure 3 shows that our classifier could classify reviews into obscure and non-obscure appropriately.

V. Analysis of obscure reviews and venues

In this section, we analyze obscure reviews and obscure reviews by using our classification methods. First, we discuss the obscure reviews and venues extracted by our method. Next, we discuss the categories of obscure venues. Finally, evaluate and discuss the differences in which each reviewer evaluates a venue as obscure or not.

A. Analysis of obscure review

This section describes and discusses obscure reviews extracted by using our proposed classifier.

In this experiment, we apply the classifier to all reviews. We used the document vectorization is SBERT and the classification algorithm is SVM, because this combination indicated the best performance in Section IV-D.

The number of reviews classified as obscure reviews is 312,151 (this is approximately 15% in all reviews). Table VIII shows some example of obscure reviews. Here, we replaced the name of the venue into "@" to anonymize it. In terms of review No. 1 of Table VIII, the reviewer wrote the location of the venue is negative but the food is positive. There are such texts in reviews classified as obscure reviews. The review No. 2 was written about a restaurant and the text contains the phrase "hidden gem". This phrase is a metaphorical expression for representing a place not very well known or unexpected find. Also, the review No. 3 and No. 4 contains the phrase representing obscure venues, but our obscure words in Table I does not include them. Therefore, their result shows that our classifier can find obscure reviews even if the texts do not directly contain the obscure words or phrases we have not prepared.

Also, we confirmed more obscure reviews manually. As a result, those reviews include many phrases of "I knew for the first time," "It was hard to access, but the service was good," and so. These phrases seem to be related to obscurity. Therefore, we believe that our method discovers venues that people have evaluated as obscure.

B. Classification results of obscure venue

This section describes and discusses the evaluation results of discovering an obscure venue using a classifier. In this experiment, we apply the classifier to all reviews of a venue and calculate the percentage of reviews classified as obscure.

The number of venues which were classified as the obscure venue is 10,915 (this is approximately 6% in all venues). Figure 4 shows the histogram of proportion of obscure reviews which were classified by our methods. This figure uses bins that are separated from 1.0 to 0.0 in 0.05 units. Here, for 1.0 and 0.5 in the figure, due to the small number of reviews included in the venue, the value is large. Venues without obscure reviews dominant in Figure 4. Also, most venues have a small percentage of obscure reviews. However, some venues have a high percentage of obscure reviews, and we regard such venues as obscure venues. Therefore, we believe that our approach can discover obscure venues.

TABLE VIII. Examples of reviews classified as obscure review.

No.	Text
1	Awesome place to eat! It may not look like much on the outside, but trust me...this place has some of the tastiest food in town.
2	This is a hidden gem. The decor is mixed but the food is excellent.
3	This is a great west side secret and I will be sure to refer the many people I encounter with in my position and let them know where I got my nails and toes done!
4	Located at the less well known spot of the @, the food court is less busy in comparison, thus it's never a hassle to find an empty seat.

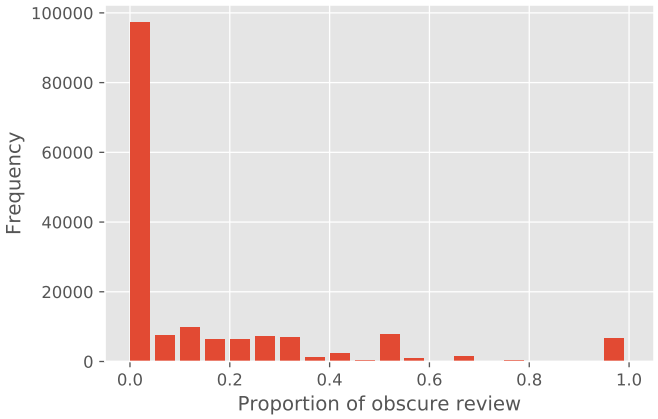


Figure 4. Histogram of proportion of obscure reviews in each venue.

C. Analysis of obscurity in each category

In this section, we analyze the obscure venues in each category. We denote the venue where the percentage of obscure reviews is 50% or more, according to the description in Section III-E, and find the proportion of venues classified as obscure within the same category.

We calculate the proportion of venues classified as obscure within a category. Here, the dataset from Yelp has 1,300 categories. Also, the venue in Yelp has at least one category. We used 62 categories whose number of reviews in a category is 1,000 or more.

We show the top 30 categories with the percentage of obscure venues in each category, as indicated in Figure 5. In this figure, the vertical axis shows the proportion of venues classified as obscure within the same category, and the horizontal axis shows the category names in Yelp. The highest percentage of obscure venues is for "Arts & Entertainment" at approximately 25%. This category has 3,886 venues in Yelp and 986 venues were classified as obscure venues. "Arts & Entertainment" has various subcategories in Yelp such as "Museums", "Stadiums & Arenas", and "Planetarium". However, these subcategories are not included in the ranking.

In Figure 5, the 2nd and 3rd categories are "Active Life" and "Shopping" at approximately 25%, respectively. "Active Life" has subcategories such as "Fitness & Instruction", "Baseball Fields", and "Parks". "Fitness & Instruction" is ranked at 18th. The obscure venue of this category occupies 25% of "Active Life". Also, "Shopping" has various various subcategories such as "Women's Clothing", "Fashion", and "Home & Garden". These subcategories are included in the top 30. Therefore, reviewers are likely to think of these subcategories as obscure venues.

In addition, according to Figure 5, the top categories with a high percentage of obscure venues contain many categories used in daily life such as "Shopping", "Education", and "Bakeries". In contrast, the subcategories of "Restaurants "such as "Steakhouses" , "Seafood", and "Breakfast & Brunch" where many people go to popular venues ranked the low. In these categories, popular venues are sometimes a type of sightseeing spot. However, as described in Table VIII, some venues were classified as obscure venues by our classifier. Therefore, we believe that such a result is correct as an analysis of obscure venues by categories.

D. Differences between venues evaluated as obscure for each reviewer

This section analyzes the differences among venues considered by reviewers as obscure.

Herein, we show the difficulty of providing a unique definition for obscure venues using our proposed method for obscure venue extraction. Using the classifier described in Section III-D, we classify whether a user review on a venue is obscure or not. Then, if the types of reviews on the venue are different, the venue that the user feels is obscure is different.

This research focused on cases in which two different reviewers posted similar reviews on two venue pairs. Two patterns of venues whose reviews refer to obscurity were considered, as shown in Figure 6. Pattern ① is a case in which two reviewers posted an obscure review and a non-obscure review to different venues. This pattern represents a case in which the reviewer felt that the referred venue was different. Pattern ② is a case in which the reviews posted by two different reviewers are the same for the referred venues. This pattern is one in which the venues the reviewers felt as obscure are the same. Therefore, if there is a certain number of reviews considered as pattern ①, it can be said that the venue perceived as obscure is different for each reviewer; the classification of obscure reviews reveals the contribution of the identification of obscure venues.

The procedure of this experiment is as follows. First, obscure venues to which two users posted similar reviews were extracted. During this experiment, 10,915 obscure venues that had obscure reviews were extracted, comprising more than 50% of all reviews. The classifier was then applied to the written reviews as described in Section IV. The numbers of the two patterns were calculated based on the classification results.

Table IX shows the experimental results. From Table IX, pattern ① comprised approximately 74% of the total. In other words, the combination of 74% of reviewers differs from the venue that was perceived as obscure. This result shows that the venues perceived as an obscure venue are not necessarily the same for all reviewers. Therefore, the approach of abstractly treating as obscure a review that includes an obscure word without criteria on the obscure venue used to extract the venue has the potential to be effective.

VI. Conclusion

In this research, we proposed a method for identifying obscure venues by extracting reviews that include descriptions

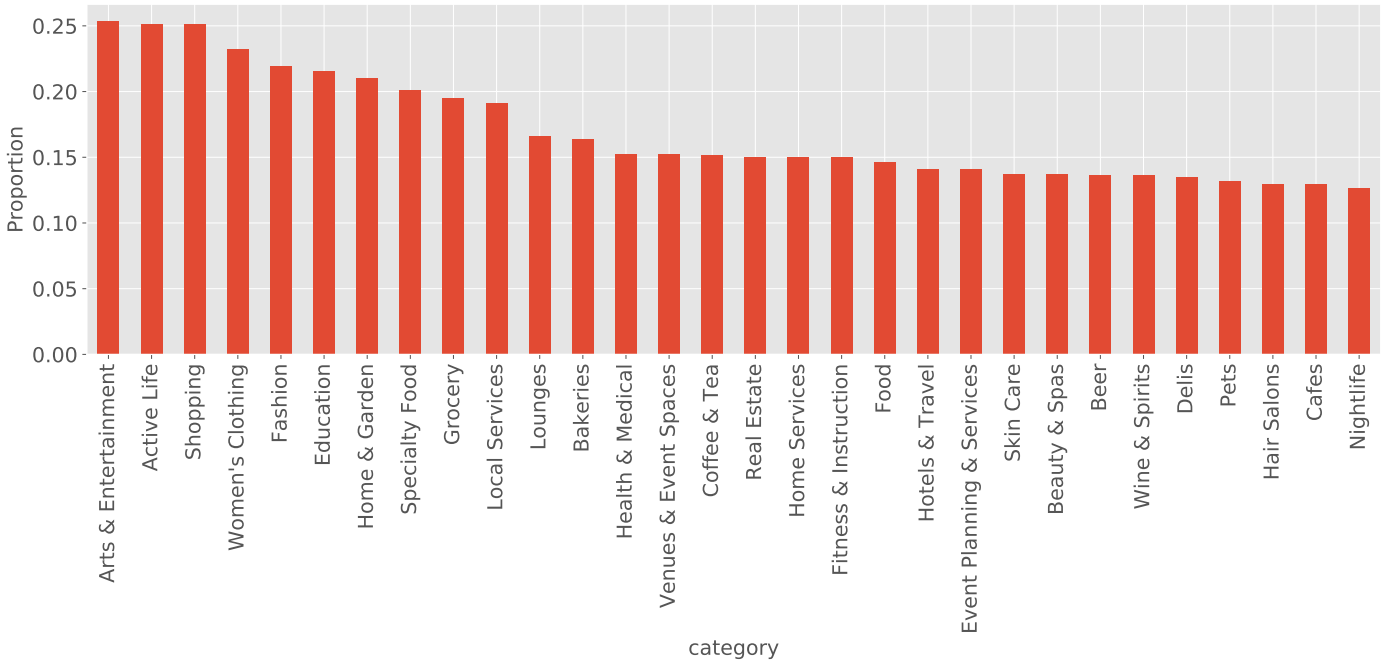


Figure 5. The top 30 categories with a high percentage of obscure venues in each category.

TABLE IX. PERCENTAGE OF DIFFERENCES IN REVIEWERS FEELING A VENUE AS BEING OBSCURE.

Pattern ①	17,206
Pattern ②	23,234
① / (① + ②)	0.74

regarding obscure posts on Yelp. Through reviews that include obscure words, a classifier was created to differentiate the reviews describing obscurity from those that do not, based on reviews in which the reviewers recognize the venues as being obscure. Evaluation results showed that the classifier is useful for extracting obscure reviews. Also, we discussed the differences of stars of obscure and non-obscure reviews to evaluate our method qualitatively. Furthermore, this research formulated and verified the hypothesis that venues perceived as obscure by reviewers are different. As a result, the venues perceived as being obscure are not necessarily the same for all reviewers, and our hypothesis is useful for discovering obscure venues.

Future studies will include a more detailed experiment and analyze obscure venues and the various categories present in each city. This paper is limited to analyzing obscure venues extracted using our proposed method in a qualitative manner. For a discovered venue, it is necessary to analyze whether it is obscure or not and to evaluate how useful or helpful the information is. For this purpose, we will conduct questionnaires by evaluators on the obscure venues by our proposed method. Further studies may apply our classifier to more various reviews such as another review site to discover obscure venues.

Also, there is necessary to examine the validity of obscure words. In this research, we used 13 obscure words, as Table shown in I. The experimental results represented that these obscure words are effective. However, we do not confirm that these words account for the majority of this meaning. The future work about obscure words investigates the validity of those words by questionnaires.

Also, the performance improvement of our obscure classifier is other future research. Although we used the LightGBM, SVM, and RF for this study, various methods for classification of texts such as graph convolutional network [42] and recurrent neural network [43] have been proposed. Also, because the number of reviews within obscure words is few, semi-supervised learning methods such as self-training [44] and label propagation [45] are suitable approaches for the situation. Those approaches might improve the obscure classifier, and we can extract more obscure reviews.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Numbers 16K00157, 16K16158 and 19K20418, and Tokyo Metropolitan University Grant-in-Aid for Research on Priority Areas Research on social big data.

REFERENCES

[1] M. Hirota, M. Endo, and I. Hiroshi, "Identifying obscure venues using classification of user reviews," in Proceedings of The Eleventh International Conference on Advances in Multimedia, ser. MMEIDA 2019, Mar 2019, pp. 7–12, ISBN:978-1-61208-697-2, URL:http://ns2.thinkmind.org/index.php?view=article&articleid=mmedia_2019_1_20_58003.

[2] "Yelp," URL: https://www.yelp.com/ [accessed: 2019-02-27].

[3] "Expedia," URL: https://www.expedia.com/ [accessed: 2019-02-27].

[4] "Tripadvisor," URL: https://www.tripadvisor.com/ [accessed: 2019-02-27].

[5] D. Ukpabi, S. Olaleye, E. Mogaji, and H. Karjaluoto, "Insights into online reviews of hotel service attributes: A cross-national study of selected countries in africa," in Information and Communication Technologies in Tourism 2018, B. Stangl and J. Pesonen, Eds. Cham: Springer International Publishing, 2018, pp. 243–256.

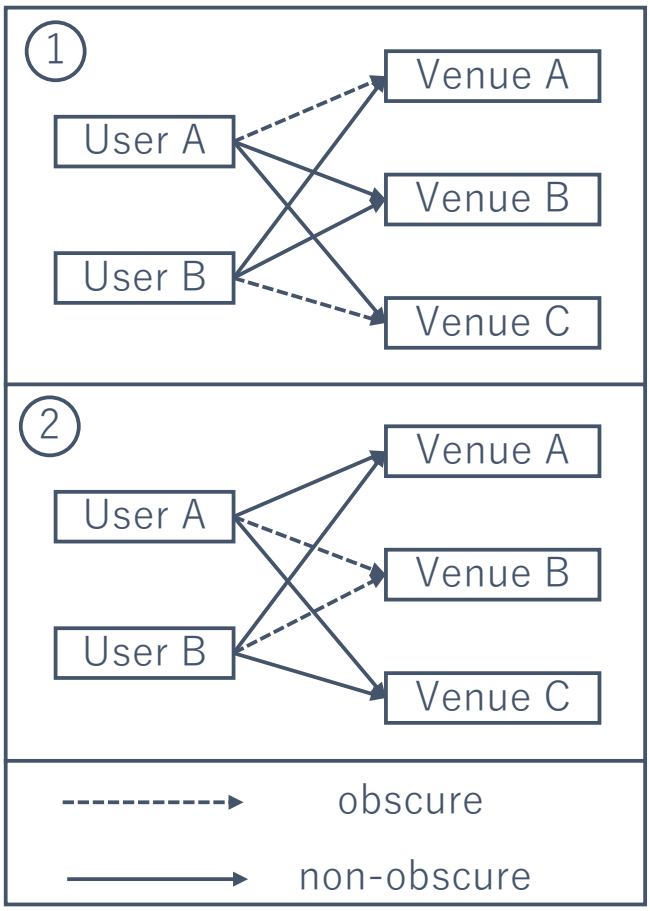


Figure 6. Pattern in which two reviewers evaluate venues as obscure.

[6] V. Browning, K. K. F. So, and B. Sparks, "The influence of online reviews on consumers' attributions of service quality and control for service standards in hotels," Journal of Travel & Tourism Marketing, vol. 30, no. 1-2, 2013, pp. 23–40.

[7] C. Zhuang, Q. Ma, X. Liang, and M. Yoshikawa, "Anaba: An obscure sightseeing spots discovering system," in 2014 IEEE International Conference on Multimedia and Expo, 2014, pp. 1–6.

[8] D. Kitayama, "Extraction method for anaba spots based on name recognition and user's evaluation," in Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services, ser. iiWAS '16. ACM, 2016, pp. 12–15.

[9] S. Katayama, M. Obuchi, T. Okoshi, and J. Nakazawa, "Providing information of hidden spot for tourists to increase tourism satisfaction," in Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, ser. UbiComp '18. ACM, 2018, pp. 377–380.

[10] Z. Liu and S. Park, "What makes a useful online review? implication for travel product websites," Tourism Management, vol. 47, 2015, pp. 140 – 151.

[11] M. Toyoshima, M. Hirota, D. Kato, T. Araki, and H. Ishikawa, "Where is the memorable travel destinations?" in Social Informatics. Cham: Springer International Publishing, 2018, pp. 291–298.

[12] Y. Guo, S. J. Barnes, and Q. Jia, "Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation," Tourism Management, vol. 59, 2017, pp. 467 – 483.

[13] C. Vo, D. Duong, D. Nguyen, and T. Cao, "From helpfulness prediction to helpful review retrieval for online product reviews," in Proceedings of the Ninth International Symposium on Information and Communication Technology, ser. SoICT 2018. ACM, 2018, pp. 38–45.

[14] Z. Wang, B. S. Balasubramani, and I. F. Cruz, "Predictive analytics using text classification for restaurant inspections," in Proceedings of the 3rd ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics, ser. UrbanGIS'17. ACM, 2017, pp. 14:1–14:4.

[15] W. He, X. Tian, R. Tao, W. Zhang, G. Yan, and V. Akula, "Application of social media analytics: a case of analyzing online hotel reviews," Online Information Review, 2017.

[16] R. Dong, M. Schaal, M. P. O'Mahony, and B. Smyth, "Topic extraction from online reviews for classification and recommendation," in Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, ser. IJCAI '13. AAAI Press, 2013, pp. 1310–1316. [Online]. Available: http://dl.acm.org/citation.cfm?id=2540128.2540317

[17] M. Afzaal, M. Usman, A. C. M. Fong, S. Fong, and Y. Zhuang, "Fuzzy aspect based opinion classification system for mining tourist reviews," Advances in Fuzzy Systems, 2016, pp. 1–14.

[18] T. S. Bang and V. Sornlertlamvanich, "Sentiment classification for hotel booking review based on sentence dependency structure and sub-opinion analysis," IEICE Transactions on Information and Systems, vol. E101.D, no. 4, 2018, pp. 909–916.

[19] P. Phillips, S. Barnes, K. Zigan, and R. Schegg, "Understanding the impact of online reviews on hotel performance: An empirical analysis," Journal of Travel Research, vol. 56, no. 2, 2017, pp. 235–249.

[20] S. Karimi and F. Wang, "Online review helpfulness: Impact of reviewer profile image," Decision Support Systems, vol. 96, 2017, pp. 39 – 48.

[21] A. R. Alaei, S. Becken, and B. Stantic, "Sentiment analysis in tourism: Capitalizing on big data," Journal of Travel Research, vol. 58, no. 2, 2019, pp. 175–191.

[22] E. van der Zee, D. Bertocchi, and D. Vanneste, "Distribution of tourists within urban heritage destinations: a hot spot/cold spot analysis of tripadvisor data as support for destination management," Current Issues in Tourism, vol. 23, no. 2, 2020, pp. 175–196.

[23] L. Jhih-Yu, W. Shu-Mei, H. Masaharu, A. Tetsuya, and I. Hiroshi, "Less-known tourist attraction analysis using clustering geo-tagged photographs via x-means," International Journal on Advances in Systems and Measurements, vol. 12, no. 3&4, 2019, pp. 215–224.

[24] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), 2014, pp. 1532–1543.

[25] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in Advances in Neural Information Processing Systems 26, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 3111–3119. [Online]. Available: http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf

[26] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," Transactions of the Association for Computational Linguistics, vol. 5, 2017, pp. 135–146.

[27] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," 2018.

[28] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le, "Xlnet: Generalized autoregressive pretraining for language understanding," 2019.

[29] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep contextualized word representations," 2018.

[30] A. Adhikari, A. Ram, R. Tang, and J. Lin, "Docbert: Bert for document classification," 2019.

[31] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks," in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 11 2019. [Online]. Available: http://arxiv.org/abs/1908.10084

[32] "langdetect," URL: https://pypi.org/project/langdetect/ [accessed: 2020-02-20].

[33] S. R. Bowman, G. Angeli, C. Potts, and C. D. Manning, "A large annotated corpus for learning natural language inference," in Proceedings of the 2015 Conference on Empirical Methods in Natural Language

Processing (EMNLP). Association for Computational Linguistics, 2015.

[34] A. Williams, N. Nangia, and S. Bowman, "A broad-coverage challenge corpus for sentence understanding through inference," in Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). Association for Computational Linguistics, 2018, pp. 1112–1122. [Online]. Available: http://aclweb.org/anthology/N18-1101

[35] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, 1995, pp. 273–297.

[36] L. Breiman, "Random forests," Machine learning, vol. 45, no. 1, 2001, pp. 5–32.

[37] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," in Advances in Neural Information Processing Systems 30, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Curran Associates, Inc., 2017, pp. 3146–3154.

[38] "Yelp dataset challenge (round 13)," URL: https://www.yelp.com/dataset/challenge [accessed: 2020-02-20].

[39] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," 2019.

[40] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, 2011, pp. 2825–2830.

[41] "Lightgbm, light gradient boosting machine," URL: https://github.com/microsoft/LightGBM [accessed: 2019-02-27].

[42] L. Yao, C. Mao, and Y. Luo, "Graph convolutional networks for text classification," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 7370–7377.

[43] P. Liu, X. Qiu, and X. Huang, "Recurrent neural network for text classification with multi-task learning," 2016.

[44] M. Pavlinek and V. Podgorelec, "Text classification method based on self-training and lda topic models," Expert Systems with Applications, vol. 80, 2017, pp. 83–93.

[45] Z.-W. Zhang, X.-Y. Jing, and T.-J. Wang, "Label propagation based semi-supervised learning for software defect prediction," Automated Software Engineering, vol. 24, no. 1, 2017, pp. 47–69.

Security Risk Analysis of the Cloud Infrastructure of Smart Grid and IoT - 4-Level-Trust-Model as a Security Solution

Katrin Neubauer
Dept. Computer Science and Mathematics
Ostbayerische Technische Hochschule
Regensburg, Germany
email:
katrin1.neubauer@oth-regensburg.de

Sebastian Fischer
Secure Systems Engineering
Fraunhofer AISEC
Berlin, Germany
email:
sebastian.fischer@aisec.fraunhofer.de

Rudolf Hackenberg
Dept. Computer Science and Mathematics
Ostbayerische Technische Hochschule
Regensburg, Germany
email:
rudolf.hackenberg@oth-regensburg.de

***Abstract*—The digital transformation has found its way into business and private life. It consists of digitization and digitalization. Digitization means the technical process and digitalization is the socio-technological process. Technologies of digitization are Cloud Computing (CC), Internet of Things (IoT) and Smart Grid (SG), which are separate technologies. The increasing digitalization in the private sector and of the energy industry connect these technologies. Actually, there is no connection between the CC infrastructure and the SG infrastructure at the moment, because in Germany the SG is currently under construction. If one looks at the CC and IoT, it must be stated there is an connection between the IoT infrastructure and the CC infrastructure as a service provider. To connect the technologies CC, IoT and SG and also build an SG cloud for innovative services, the new laws for privacy must be implemented. For privacy and security analyses it is important to know which data can be stored and distributed on a cloud. To illustrate this analysis, we connect the SG infrastructure with the IoT. An IoT device (car charging station) should be able to transfer data to and from the SG. SG is a critical infrastructure and the IoT device a potential insecure device and network. We show the communication between the smart meter switching box and the IoT device and the data transferred between their clouds. The charging station is connected to the SG to get the current amount of renewable energy in the grid. This is necessary to create a new smart service. But this service also generates private data (e.g., name, address, payment details). The private data should not be transferred to the IoT cloud. For the connection of SG and IoT, availability, confidentiality and integrity must be ensured. A risk analysis over all the cloud connections, including the vulnerability and the ability of an attacker, the resulting risk and the 4-Level-Trust-Model for security assessment are developed. Furthermore, we show the application of the 4-Level-Trust-Model in this paper.**

Keywords*—*Smart Grid; Internet of Things; security analysis; safety-critical infrastructure; cloud computing; 4-Level-Trust-Model

I. Introduction

This paper extends the already published paper "Risk Analysis of the Cloud Infrastructure of Smart Grid and Internet of Things" [1] with more detailed information of the risk analysis and the 4-Level-Trust-Model as a security solution for the main problem with the different data.

With the increasing digitalization in our world, new technologies, like the Internet of Things (IoT), have a great influence on our future way of life. Non-technical user are using connected technologies to improve their comfort without knowing about the possible risks.

But not only private technologies are increasing their digitalization, the future Smart Grid (SG) is also a highly networked system. In order to use these innovative services, which emerge from the digitalization, a third technology, Cloud Computing (CC) is necessary. With all three technologies combined, new services can be offered and the transformation of the energy system can be successfully implemented.

In Germany, the integration of the intelligent energy supply system (SG) is creating a new IT infrastructure. The intelligent measuring system (iMSys), containing a basic meter (smart meter) and the smart meter gateway (SMGW) [2] are currently installed in many households and companies in Germany. But other countries like Italy or Sweden are already further ahead with the development of the SG infrastructure.

The digitalization of the electricity grid brings new dangers and challenges in the area of IT-Safety and -Security. These can even allow attacks from the internet where no physical access to the network is necessary. Besides the SG, all kind of devices are getting a connection to the internet. These devices can range from smart refrigerators to connected cars and are called IoT. Most of the time, existing devices are getting a communication interface and are connected to the internet over a gateway or directly.

IoT, just like SG, also brings new IT-Security and -Safety dangers. For consumer devices, the damage is normally not high, but the lack of IT-Security in consumer devices, which are connected to other networks, can lead to serious damage in other (critical) infrastructures. One big challenge is the high number of newly connected devices. Services for only a few devices, are getting new ones on a large scale, which are not necessary persistent. They are very flexible and appear and disappear quickly in their lifetime. This volatility is a big challenge for the security of existing and new services.

Beside of all dangers, new technologies are emerging and the new challenges must be solved. The smart services are required for future applications and the connection between

SG and IoT is necessary, to regulate the amount of energy in the grid. For these services, the cloud platform is needed as a connection between both technologies. It can be described as a data hub, for data storage, analysis and the services.

Both technologies, SG and IoT, are implementing their own cloud platform with the corresponding infrastructure. These independent clouds must be connected in order to offer new services with the desired added value. The potential insecure device and infrastructure of IoT should be able to communicate in both directions with the critical infrastructure of the SG. The security objectives availability, confidentiality, integrity and privacy must nevertheless still be ensured. Therefore, new risks and attack vectors emerge and new requirements for authentication and authorization are needed.

In this paper, we connect a IoT and SG cloud and perform a risk analysis over our example architecture to see the new problems and dangers of this connection. In the next step, a security model, based on IoT security standards and a new 4-Level-Trust-Model is developed. Finally, the model is applied to the example, to show the benefits of our security model.

The paper is structured as follows. Section II covers the related work and existing publications. In Section III, we describe our architecture and the corresponding challenges for the connection between the two technologies. In the next section, the security analysis is performed and Section V, describes the security model, which is applied in Section VI to our example. Finally, the conclusion is given.

II. Related Work

IoT devices are potential insecure devices. The security gaps in IoT devices can be protected with known principles. The problem is that they are not used by the manufacturers. One reason for this could be problem of costs. It is important for research to respond to new challenges in this field.

One challenge is the scarce resources of IoT devices. Already known encryption algorithms need to be adapted or changed to work more effectively and operate acceptably with low-performance hardware (e.g., PRINCE [3]). As an alternative, the new development of suitable algorithms can be considered (e.g., Secure IoT - SIT [4]).

Some publications cover the details about the necessary encryption and communication protocols, but do not classify the different data, e.g., [5] and [6]. The publication "Enabling privacy and security in Cloud of Things: Architecture, applications, security & privacy challenges" [7] describes an architecture for the combination CC and IoT, the "Cloud of Things". With the 4-Level-Trust-Model, a solution is developed to deal with the the mentioned security and privacy threats.

Currently, insecure devices are in use. For this situation, solutions must be found to continue the operation. The "Quad9 DNS Privacy and Security Service" is an example for this problem. Several companies (including IBM) have developed a special DNS server (Quad9 DNS Privacy and Security Service), which should ensure the security as well as privacy of the IoT devices. Quad9 automatically blocks requests to infected sites. As a last challenge, manufacturers must be "forced" to improve IT security. This can be accomplished by guidelines and certifications.

For SG exist an European architecture model so called Smart Grid Architecture Model (SGAM). This model was developed in the context of the European standardization mandate M/490. The SGAM includes the visualization, validation and structuring of SG projects from the beginning of the project as well as for the standardization of SG. The model was also used for the SG architecture development at different organizational levels. In this model, security is not explicitly considered. This publication describes security as a cross-cutting topic [8]. The architectural models of the countries differ in principle, but they are mostly based on the SGAM. In Germany, the SG itself is regulated by the specifications of the Federal Office for Information Security (BSI) and is regarded as the state of the art (communication) [9]. The BSI was commissioned by the legislator to develop specifications for a SMGW in order to guarantee a secure infrastructure for intelligent measuring systems [10]. The intelligent measuring systems will be integrated into a communication network with the central element SMGW as a communication unit [11]–[13].

Security and privacy considerations for IoT application on SG with a focus on survey and research challenges presented are shown in [14] and [15]. The publication gives a brief insight SG and IoT application on SG. Furthermore, the publication identifies some of the remaining challenges and vulnerabilities related to security and privacy. A security and communication analysis of SG, IoT and CC in Germany are shown in [16]–[18]

Classical models for IT security assessment are the BSI-Standards (BSI-Standards 200-1, 200-2 and 200-3 [19]–[21]) or ISO/IEC 27000:2018 [22], which classically consider the IT processes within a company. With highly scalable and distributed systems (such as CC, IoT and SG), the entire IT process must be considered. In [23]–[26] security is considered during the development process of software. The security evaluation of data is based on a 2-level trust model shown in [27]. These known models for security modeling as well as the 2-level trust model are not suitable for cyber physical systems (CPS).

The handling of data when they leave the "SG", requirements for authentication and authorisation in future SG-IoT-cloud application and how to deal with service provider who access data (service charging station) in critical infrastructures are open questions. For this open question there is no related work.

III. Architecture Challenges for Smart Grid and IoT

First, we describe the challenges in SG and IoT individually, then we present our example architecture with the communication and the corresponding data.

TABLE I. Energy-Supply: Today - Future

today	future
central supply	decentralized supply
bilateral and wholesale trade (local markets)	centralization (regional market)
transfer energy	transfer energy and data
reading of the meter content: once a year (manual)	smart metering: transfer data all 15 minutes

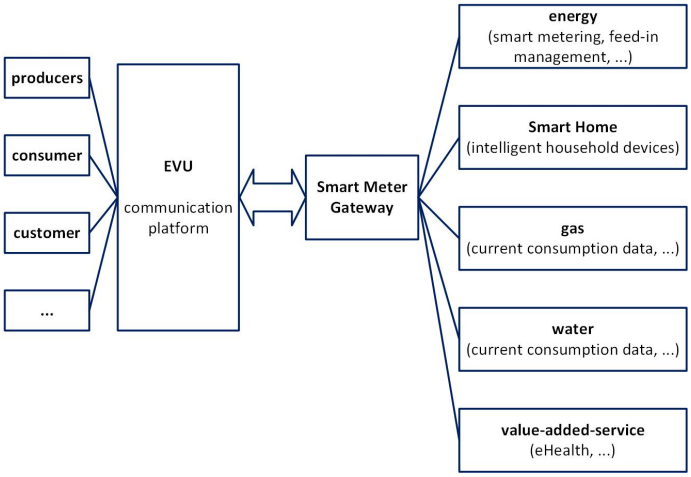


Figure 1. Application example Smart Grid

A. Smart Grid

In the future SG (compare Table I), large amounts of data are generated daily when meter data are read out on a regular basis. These data have to be stored, archived and analyzed. The conversion of the entire energy supply system from a centralized supply to a decentralized supply is progressing continuously in Germany [28]. The transformation of the energy supply to an intelligent energy supply system creates numerous new opportunities and challenges. The changeover to renewable energies alone, such as wind power or solar energy, creates new challenges for future systems in the SG.

The SG infrastructure is not only used for the use case "energy" like smart metering (see Figure 1). Further filed of application are smart home, gas, water and value-added service. The energy supplier (EVU) operates a data platform to connect the users. In this case, user are producer, consumer and customer. The SMGW is the secure interface and communication unit between the household and the EVU.

The conversion is not only taking place in Germany, but also in other European countries. Pioneers are countries like Italy and Sweden [28]. However, these rollouts have already been carried out the dangers from a security and safety perspective. With regard to security of supply, attacks on control systems of the power grid via the Internet represent a growing threat, because on the one hand, the power grid can be controlled or manipulated over it. On the other hand, data requiring protection about the consumer and their behaviour can be accessed. This is because data of varying origin and quality is processed and analysed in real time. As a result, access to the systems must be guaranteed for different groups of people.

B. Internet of Things

The Internet of Things is defined in the ISO/IEC 20924:2018 standard as a infrastructure, which connects entities with services which react to information from the physical and virtual world [29]. This includes all connected devices nowadays, regardless if they are connected to the Internet or not. In our paper, we restrict this definition to common IoT devices, which benefit from a connection with the SG. This mainly includes smart home devices with a high energy consumption like a smart charging station.

Especially smart home devices are currently highly insecure, because of the increasing amount of devices [30] and the cheap price. Nearly every home appliance devices needs a connection to some smartphone application and the internet to control them remotely. This leads to a fast development of new features and connection points without enough time to care about the security. The second security issue is price, because no customer is willing to pay more for a device just because it was designed to be secure. The cheapest device with the most features is usually always bought.

Botnets like Mirai [31] and other malware are using insecure IoT devices, to attack other networks. The security problems of IoT are not new and the majority of them can be solved with common IT-Security methods. This shows the OWASP IoT Project. The top vulnerabilities in IoT devices, like default or weak passwords, are simple to fix [32].

Because of this, we consider IoT devices as insecure. There are too much insecure devices in operation and there is no prove of security of new devices. Nevertheless, we connect an insecure IoT network to a probably insecure cloud and this cloud finally to the SG.

C. Architecture Smart Grid and IoT

The SG reference architecture consist of the Local Metrological Network (LMN), the Wide Area Network (WAN) and the Home Area Network (HAN). The connection between these networks takes place through the SMGW. The LMN consists of all the gas meter, electricity meter, etc. The WAN is outside of the building and describes the connection over a wider range. The Gateway-Administrator and the energy supplier (Energieversorgungsunternehmen in German, short EVU) are located in the WAN. The last network, the HAN, is the local network in the building with the connected (smart home) devices. The SG cloud extends the common SG reference architecture, as shown in Figure 2.

The IoT network is located in the HAN with all connected IoT devices. Some devices are connected over a gateway to the Internet. The IoT Cloud and the user can access theses devices over these HAN connections.

New services can use the central stored data of the cloud platforms and make it available to the user. As shown in Figure 2, the connection between IoT and SG can be established

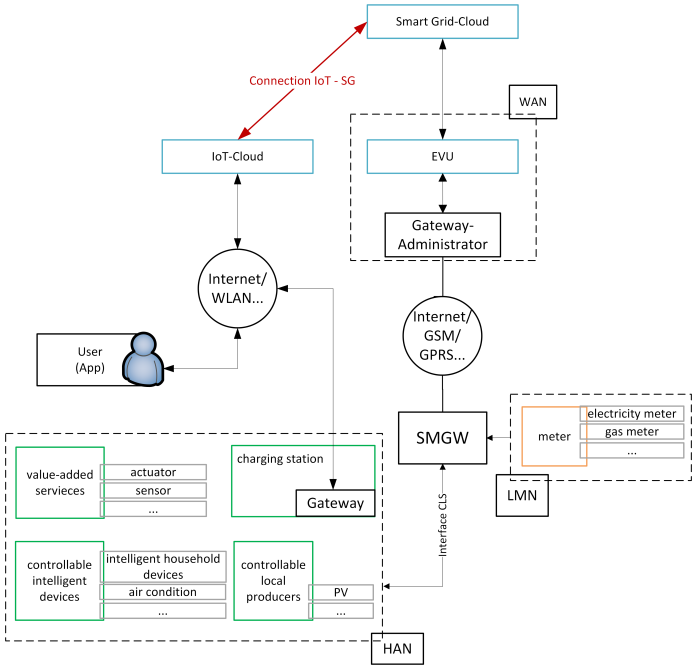


Figure 2. Architecture Cloud Application on Smart Grid with IoT

between the two clouds (IoT Cloud and SG Cloud). It is the main focus of this paper, as this is where the two technologies are combined and data exchange takes place.

D. Application Example

For a realistic and useful example, we use a smart car charging station with a cloud connection. The charging station is considered as insecure, as well as the whole IoT network (Gateway, Cloud, Applications). The SG network is as extensive as described in the previous section with all the components. As meter, a electrical meter is used, because the smart service should connect the charging station cloud and the SG cloud. The service can load the electric car, connected to the charging station, at the most suitable times. The grid can charge the car, when a lot of energy is produced and therefore in the grid. If the grid is low on energy (can be detected by the current frequency), the car can supply the grid with stored energy to stabilize it. The smart service connects the two clouds, because at this point it is possible to get all necessary data from both technologies.

E. Communication between Smart Grid and IoT

By connecting the two clouds, data is exchanged. To determine the risk of the connection, it is necessary to know which data is transferred. There are only the communication data from IoT and SG listed, because not all stored data is exchanged.

1) Communications data Internet of Thing: The following data is stored in the IoT cloud and transmitted to the smart service as needed:

- ID Connected car
- ID gateway (charging station (CS))
- IP-Address gateway (CS)
- Sum of energy consumption (CS)
- Current energy consumption / supply (CS)
- History of energy consumption / supply (CS)
- Time to load the car
- User data (CS)
  - Name
  - E-Mail

The connected car and the history can be used to create a profile of the user. This includes the times, the user is normally at home or at work. This data is private data and should be protected.

2) Communications data Smart Grid: The following data is generated and stored in the SG cloud, as well as transmitted to the smart service:

- Information about the smart meter (ID, IP-Address)
- Current energy consumption
- Current price for electricity
- Information about the customer
  - Name
  - Address
  - Payment details

The information about the smart meter or the current energy consumption can be used to create a profile of the household (user). This is partly equal to the profile of the connected car, but can be extended to the whole household an therefore other people. In conclusion, like the connected car data, this data is also private data and should be protected. Special data protection precautions must be taken, as this may allow conclusions to be drawn about third parties (other people in the household).

IV. Security Analysis for Safety-Critical Systems

The security analysis starts with the description of the attack vectors. From these vectors, the threads are derived. In the next step, the risk is shown for every thread, based on the ability of the attacker and the possible damage. Finally, practical examples show the potential danger in our example architecture.

A. Attack vector Smart Grid and Internet of Thing

There are four kind of attack vector categories: hardware manipulation attacks (physical attacks), software manipulation attacks, network-based attacks and privacy related attacks [33] [34]. Each attack tries to get unauthorized access to the infrastructure or inflict some damage to it [35].

Hardware manipulation or physical attacks are performed locally on the device. It is possible to change the hardware and the software. Mostly malware is installed, which leads to data manipulation and sniffing. In context SG, a complete shutdown of the grid would be possible in the worst case. However, sensitive (private) data can also be tapped and modified. In

the case of IoT devices, for example, the software can be modified so that the device acts as a spy and forwards all data to the attacker. Hardware attacks thus open up all possibilities for an attacker, but are very difficult to execute.

With software manipulation attacks, it is possible to change the software (or firmware) of the device. These attacks can be done remotely over the internet or any other network. The attacker uses a weakness in the running software (e.g., buffer overflow, code injection) to execute his own code or tries to manipulate the administrator of the device to install the malicious software. As with hardware manipulation attacks, in the worst case the SG can be shut down or sensitive (private) data can be modified or tapped.

Network-based attacks like identity theft, denial of service, cascading malware propagation (Business IT & Plant Control) and monitor, traffic analysis (passive attacks) are using the network to inflict damage. They can be used to get data or to disable the service. These attacks are difficult to protect from, because the hole network (internet) is not controlled.

The last category are privacy related attacks. With these attacks, user-specific data are collected and used to inflict personal damage to the customers or the energy supplier. They can be combined with other attacks or used to trick the administrator to install malicious software (social engineering).

According to IoT and SG, the following risks are possible: manipulation of measured values and time, manipulation of the communication between IoT cloud and SG cloud, misuse of energy data and/or sensitive data, sabotage of the power grid and sabotage of mobility (example: charging station).

B. Security threats: Infrastructure Smart Grid and Internet of Things

The risk analysis for both, the IoT cloud and the SG cloud, are including the ability of an attacker and the potential damage, which are leading to a risk for the associated attack. With a lower ability, it is more likely for an attacker, to use this kind of attack [36]. The potential damage of an attack is related to the real damage (destroy some parts of the grid or the unavailability of services) and the personal damage, caused by stolen private information. For example, an attacker gets private data from the SG, the ability needs to be high, but the damage is high, too. This lead to a high risk overall [37].

Because of strict specifications and regulations of the SG in Germany, the ability of an attacker must be high in the most cases.

1) DoS and DDoS: A (distributed) denial of service (DDoS or DoS) attack tries to flood the device or network with too much data, so the service becomes unavailable. This kind of attack can be performed distributed with a lot devices from a botnet at low costs.

For IoT devices there is low damage, because most of them are just for comfort features. Necessary devices, like electric cars, are not available in high amounts at the moment, so not many of them are affected. For the SG, such an attack can lead to a shutdown of the grid, because the SG is unable to broadcast the current amount of energy in the grind and all connected cars start charging. The medium and high damage, combined with the low ability needed, are leading to medium to high risks for both technologies.

Ability of an attacker
IoT: low SG: low
Damage
IoT: medium SG: high
Risk
IoT: medium SG: medium / high

2) Malware: For using a malware, the attacker needs to know or find a vulnerability in the software. This can be very easy in IoT devices, because of the bad security situation. For eamxple, the mirai botnet started by using easily guessable login credentials to compromise the devices [31].

The SG is strictly regulated in Germany by the Federal Office for Information Security with the technical regulations TR-03109 [8]. This certification is needed to operate the devices, so they can be declared as secure and the ability of an attacker has to be high to attack them.

The damage for IoT is similar to the one for the DoS and DDoS attacks. But the SG can be compromised and the attacker can shutdown the hole grid or even damage hardware components.

The derived risk of a malware attack is therefore medium for IoT and medium to high for SG.

Ability of an attacker
IoT: low SG: high
Damage
IoT: medium SG: high
Risk
IoT: medium SG: medium / high

3) Broken Authentication: Like shown at malware attacks above, the broken authentication is very similar. The IoT devices are not secure and the SG is considered as secure, because of the certification.

The damage and the risk were also assessed as in the malware section. A broken authentication can lead to a full compromisation of the device or the network.

Ability of an attacker
IoT: low SG: high
Damage
IoT: medium SG: high
Risk
IoT: medium SG: medium / high

4) Broken Encryption: The broken encryption is also very similar to the malware attacks. The IoT devices are not secure and the SG is considered as secure, again because of the certification.

The damage is not so high as malware or broken authentication, because only the data send over the network can be attacked. Depending on the content of the data, personal information may be included, but the confidentiality of the

data is not necessary for the operation. The damage at IoT can be low to medium, because of the different device types. The SG can expose more personal information, so the damage is medium.

Because of the low ability and the low to medium damage, the risk of IoT is medium. In the SG a high ability is needed, which leads to medium damage, the risk is declared as medium.

Ability of an attacker
IoT: low SG: high
Damage
IoT: low / medium SG: medium
Risk
IoT: medium SG: medium

5) Data leakage: When a part of the data or all data are exposed, the damage and the risks are the same as by broken encryption. The ability is also rated the same, but can be a bit lower, because sometimes no encryption at all is used for IoT devices.

Ability of an attacker
IoT: low SG: high
Damage
IoT: low / medium SG: medium
Risk
IoT: medium SG: medium

6) Data manipulation: As mentioned before, data manipulation can be easily performed in IoT environments, because of missing regulations. For example, the IoT Cloud can be attacked and adopted, because easy to guess passwords are used. As the last sections, the SG network is secure and no data can be manipulated.

In the most cases, the manipulation of data for IoT devices is only possible for one kind of device or one manufacturer. This limits the damage and therefore has no great effect on the SG. If SG data are manipulated, it can lead to some damage, but not for all user, just for the affected ones. Therefore, the damage for IoT is low and for SG medium.

The risk was assessed as before. A low ability and a low damage are leading to a low risk. A high ability and a medium damage to a medium risk.

Ability of an attacker
IoT: low SG: high
Damage
IoT: low SG: medium
Risk
IoT: low SG: medium

7) Hardware manipulation: It is very difficult to get access to the hardware. The cloud server are most of the times under good protection, especially in the SG and the devices are installed in the house. If an attacker gets access to one house, it is only one device affected and not the whole network. These points are leading to a medium and a high ability for the attacker.

If it is possible to get hardware access to the cloud, the damage can be medium to high. For an IoT device, the attacker only gets access to one or some manufacturer. But the SG cloud can be used to shutdown the whole grid.

The risk is straight forward for IoT, because the ability and the damage are both medium. For the SG, the risk is medium because it is difficult to attack the server infrastructure.

Ability of an attacker
IoT: medium SG: high
Damage
IoT: medium SG: high
Risk
IoT: medium SG: medium

	IoT risk	SG risk
DoS und DDoS	medium	medium / high
Malware	medium	medium / high
Broken Authentication	medium	medium / high
Broken Encryption	medium	medium
Data leakage	medium	medium
Data manipulation	low	medium
Hardware manipulation	medium	medium

Figure 3. Summary of the risks

C. Summary of the security analysis

As shown in Figure 3, the summary of the risks shows that the SG is always exposed to at least medium risk (sometimes medium to high), while for IoT, the maximum is medium. This shows a need for action, especially for SG.

D. Examples

In the following, we show a few examples of how the problems by connecting IoT and SG can be recognized. The first two examples are from [1]. As an IoT device and the according infrastructure are currently highly insecure [38], all problems are realistic and the data from IoT can be considered easily accessible.

Example 1: The user can register his IoT device in the IoT cloud only with a valid E-Mail address and a username. No further information is needed. The IoT provider only knows that this username has loaded his car 20 times per month. By exchanging data with the smart meter, detailed information(name, address) about the user can be transferred. Now it is possible to identify the user.

Example 2: The energy service provider does not need any information of the connected car of the user. But with additional information from the IoT charging station, it is possible to tell when the user is at home or if he gets visited by another person with an electric car. This part is very important. A third user can be tracked with his car, without knowing it.

Example 3: The SG customer does not wish to disclose any personal information about his purchasing behaviour or financial situation. If, however, data of the car (cheap or expensive

car) and the charging points (e.g., at which supermarket the car is charged) are exchanged, an exact profile of the user can be created with the additional personal information from the SG.

Example 4: A hacked charging station can be made by software to charge at times when the electricity price is high. This can also result in financial damage for the user.

All four examples are showing the importance of a security and privacy orientated connection. As default no data should be transferred between the clouds. The user should have to confirm each data exchange.

V. Security model

We are presenting two parts, to improve the security of our example. The first part consists of security standards for IoT, which are currently under development and the second part shows a 4-Level-Trust-Model. The security standards are just a overview, but the 4-Level-Trust-Model is a development by our own.

A. IoT Security Standards

In order to increase the security of IoT devices, security and privacy must be taken into account during the development phase (Security- and Privacy-by-Default). Since most manufacturers are currently foregoing such measures because of the costs, guidelines and standards must be developed to implement a minimum level of security.

In Germany, DIN SPEC 27072 was published in 2019 [39], which sets minimum security requirements for consumer devices. These include a secure password, encryption, updates, etc. This standard is currently not mandatory and manufacturers of IoT devices can voluntarily develop their products according to it.

In 2020, the UK Government has published a guideline, which is also aimed at consumer devices and will be binding. It focuses on three aspects [40]:

- IoT device passwords must be unique and not resettable to any universal factory setting.
- Manufacturers of IoT products provide a public point of contact as part of a vulnerability disclosure policy.
- Manufacturers of IoT products explicitly state the minimum length of time for which the device will receive security updates.

Besides these local standards, the European Telecommunications Standards Institute (ETSI) is working on EN 303 645 - Cyber Security for Consumer Internet of Things. This standard is currently available as a draft and has similar requirements according to the DIN standard. The main focus are also consumer devices, but with less restrictions.

The current EN 303 645 draft consists of requirements, grouped into the following thirteen topics [41]:

- No universal default passwords
- Implement a means to manage reports of vulnerabilities
- Keep software updated
- Securely store sensitive security parameters
- Communicate securely
- Minimize exposed attack surfaces
- Ensure software integrity
- Ensure that personal data is protected
- Make systems resilient to outages
- Examine system telemetry data
- Make it easy for consumers to delete personal data
- Make installation and maintenance of devices easy
- Validate input data

TABLE II. New ability of an attacker and new risk by applying standards

	ability	ability new	risk	risk new
DoS und DDoS	low	low	medium	medium
Malware	low	medium	medium	medium
Broken Authentication	low	high	medium	medium
Broken Encryption	low	high	medium	low
Data leakage	low	medium	medium	low
Data manipulation	low	medium	low	low
Hardware manipulation	medium	medium	medium	medium

In the future, when IoT devices will be developed with the help of security standards, the risk analysis will no longer have to assume that the attacker needs little effort to attack the device and the risk will be significantly reduced. Furthermore, for most use cases, considerably less personal information can be collected and stored, making it more difficult to obtain sensitive information. For example, a networked refrigerator does not have to identify its user exactly (with name, address, etc.). Authentication without further user details or a social media account is sufficient.

The risk analysis for IoT changes with the improvements of the standards. The requirements about passwords, updates, encryption, authentication and data minimization are leading to higher abilities and therefore to a lower risk. The new values for the ability of an attacker and the risk can be seen in Table II. These new values are just assumptions, because some requirements are not mandatory and can improve the security even more, depending on whether they are implemented.

B. 4-Level-Trust-Model for safety-critical systems

Data are to be regarded as endangered property. The 4-level trust model is intended to protect the data (depending on its specification). For example, all smart meter data are data worthy of protection [42]. This is a statement of the state conference of data protection officers. It means that all data (IP adress, frequency, customer data, etc.) must be specially treated during processing, transmitting and storing. Data generator and user are human and machine also. The processing of data is in real-time not now but in future. SG is a variant of CPS. The requirements of future Systems like SG are:

- High scalable:
  The use case data logging electricity shows us the Data flaw from final consumers to the energy supplier. This

means 2 million participants and 192 million consumption values per day.
- Volatile:
  If we have a look inside the communication. There are data transfer every 15 minutes.
- High data volume:
  For example, 2 million households generating 22 gigabyte data per day.
- Different types of data:
  Customer data, power consumption, IP address, etc.

Security assessment must be adapted with regard to these additional requirements. The 4-Level-Trust-Model for safety-critical systems was developed based on the requirements. The 4-Level-Trust-Model for safety-critical systems is one option of the role-based trust model for safety-critical systems [43]. This is a model for security assessment for CPS. Classically, data are divided into two categories - secure and insecure. This is described as the classical security model. In the new 4-Level-Trust-Model for safety-critical systems the data are categorized in 4 categories. The categorization depends on the requirements analysis for CPS. The 4-Level-Trust-Model for safety-critical systems is defined as follows.

1) Category: non sensitive data
   - All data that do not contain any personal reference or have been made anonymous.
   - There are no effects of damage or damage that has occurred for the affected person.
   - The security level is low.
2) Category: high sensitive data I
   - All data which, through the combination of several data in category 2 and 3, have a personal reference, but do not have a direct reference themselves (e.g., network status data).
   - The damage effects are limited and manageable. Any damage that has occurred is relatively easy to heal for the affected person.
   - The security level is minimal.
3) Category: high sensitive data II
   - All data which, through the combination of a further date in categories 2 and 3, have a personal reference, but do not have a direct reference themselves (e.g., status data of a meter).
   - The impact of the damage can be assessed as significant by one person. Damage that has occurred for the person affected can be healed with increased effort.
   - The security level is intermediate.
4) Category: high sensitive data III (personal data)
   - All data that are personal data or data worth protecting according to the Federal Data Protection Act (e.g., name, address).
   - The effects of the damage have reached an existentially threatening, catastrophic extent. Damage that has occurred to the affected person cannot be healed.
   - The security level is high.

Table III shows the 4-Level-Trust-Model for safety-critical systems with the coding and the security level. The 4-Level-Trust-Model for safety-critical systems permits to consider the security assessment of data.

TABLE III. Evaluation criteria data security

category	description	security level	coding
1. Category	non sensitive data	low	0
2. Category	high sensitive data I	minimal	1
3. Category	high sensitive data II	intermediate	2
4. Category	high sensitive data III	high	3

With the 4-Level-Trust-Model it is possible to evaluate data and information of a use case in CPS with regard to security. By subdividing the data worthy of protection, a further gradation between personal data and sensitive data is made. With this model, appropriate security measures can be selected. The security measures for SG must be taken from the respective standards of the BSI. Security measures for IoT must be taken from the corresponding standards (see above).

The proposed model is an extension of the 3-Level-Model (such as security evaluation according to the BSI standards) and is a possibility to perform security evaluation in CSP. The 4-level model has proven itself in application.

VI. Application example: 4-Level-Trust-Model for safety-critical system

In the following section, we present the security assessment based on the 4-Level-Trust-Model for safety-critical systems. The application example is SG and IoT: charging station (see Section III, part D). Table IV shows the security assessment in detail. We categorized the data and matched the security level.

For example, the “ID connected car” is a data type for the third category. The security level is “intermediate” and the coding is “2”. In combination with one data of the second category is an personal reference possible. Another example is the assessment of the history of energy consumption / supply CS. This is a data type for the second category. The security level is “minimal” and the coding is “1”. In combination with several data (e.g., ID Gateway CS, IP-address) of the third category is an personal reference possible. For example, the information about the customer are a data type from the fourth category. The security level is “high” and the coding is “3”. The data are personal data like name or street.

This security analysis enables the selection of appropriate security measures (e.g., authentication). For the authentication of devices which transmit data such as ID Gateway CS (category 3), a procedure that guarantees a high level of security can be selected. On the other hand, a minimum level of security can be ensured for the authentication of devices that transmit data assigned to the category 2 (e.g., history of energy consumption / supply CS).

TABLE IV. Evaluation of the data security: use case charging station (SG and IoT)

data	category	security level	coding
ID connected car	3. Category	intermediate	2
IP-Address Gateway (CS)	3. Category	intermediate	2
ID Gateway (CS)	3. Category	intermediate	2
IP-Address smart meter	3. Category	intermediate	2
IP-Address smart meter	3. Category	intermediate	2
Sum of energy consumption CS	2. Category	minimal	1
Current energy consumption / supply CS	2. Category	minimal	1
History of energy consumption / supply CS	2. Category	minimal	1
Time to load the car	2. Category	minimal	1
User data CS	4. Category	high	3
smart meter ID	3. Category	intermediate	2
IP-Address smart meter	3. Category	intermediate	2
SMGW ID	3. Category	intermediate	2
IP-Address SMGW	3. Category	intermediate	2
Current energy consumption (SG) smart meter ID	2. Category	minimal	1
IP-Address smart meter	3. Category	intermediate	2
Current price for electricity smart meter ID	2. Category	minimal	1
IP-Address smart meter	3. Category	intermediate	2
smart meter ID	3. Category	intermediate	2
IP-Address smart meter	3. Category	intermediate	2
Information about the customer	4. Category	high	3
IP-Address smart meter	3. Category	intermediate	2

VII. Conclusion

In this paper, we show different challenges for the digitization and digitalization. New connected technologies, like the SG and IoT are getting connected. This can result in some serious security issues, because the SG is a critical infrastructure and current IoT devices are insecure. In our application example, a car charging station with its corresponding cloud (IoT) is connected to the SG infrastructure. For this use case example, we carried out a security analysis for safty-critical infrastructures. We show the attack vectors of SG an IoT and the security threats. The SG is always exposed to at least medium risk (sometimes medium to high), while for IoT, the maximum is medium. Due to these high risks, security by connecting the two technologies must be significantly improved. Four examples were presented of why lack of security is a problem.

For improvement, IoT devices can be secured by applying standards, like the DIN 27072 or the European version from ETSI - EN 303 645. As an advanced solution, we introduced the 4-Level-Trust-Model for safety-critical systems. The 4-Level-Trust-Model is one option of role-based trust model. With this model, data and information of a system can be evaluated. A distinction is made between personal data and sensitive data. With this security assessment, CPS can be evaluated. This model offers assistance in the selection of appropriate security measures. We have shown the application of the 4-Level-Trust-Model using the application example "connect a charging station with a cloud to SG infrastructure".

The security standards and our trust model can only help to decrease the risks. To establish a highly secure connection between IoT and SG, more considerations are needed. The interfaces must be clearly defined and communication must be restricted accordingly. A detailed risk analysis on the concrete architecture is as necessary as extensive penetration tests.

The 4-Level-Trust-Model provides a good basis and the next step is to implement the model to demonstrate its functionality in practice. It will be some time before the two technologies (IoT and SG) are connected in Germany and by then, a complete secure infrastructure model can be developed.

References

[1] K. Neubauer, S. Fischer, and R. Hackenberg, Risk Analysis of the Cloud Infrastructure of Smart Grid and Internet of Things, The Tenth International Conference on Cloud Computing, GRIDs, and Virtualization, pp. 82-87, 2019.

[2] M. Irlbeck, Digitalisierung und Energie 4.0 Wie schaffen wir die digitale Energiewende?, Springer Fachmedien Wiesbaden GmbH, pp. 135-148, 2017.

[3] H. Kim and K. Kim, Toward an Inverse-free Lightweight Encryption Scheme for IoT, Conference on Information Security and Cryptography, 2014.

[4] M. Usman, I. Ahmed, M. I. Aslam, S. Khan, and U. A. Shah, SIT: A Lightweight Encryption Algorithm for Secure Internet of Things, CoRR abs/1704.08688, 2017.

[5] I. Ben Dhaou, A. Kondoro, A. Kelati, D. S. Rwegasira, S. Naiman, N. H. Mvungi, and H. Tenhunen, Communication and Security Technologies for Smart Grid, International Journal of Embedded and Real-Time Communication Systems (IJERTCS), 8(2), pp. 40-65, 2017.

[6] D. G. Korzun, I. Nikolaevskiy, and A. Gurtov, Service Intelligence and Communication Security for Ambient Assisted Living, International Journal of Embedded and Real-Time Communication Systems (IJERTCS), 6(1), pp. 76-100, 2015.

[7] A. A. A. Ari, O. K. Ngangmo, C. Titouna, O. Thiare, A. Mohamadou, and A. M. Gueroui, Enabling Privacy and Security in Cloud of Things: architecture, applications, security & privacy challenges, Applied Computing and Informatics, 2019.

[8] M. Uslar, C. Rosinger, and S. Schlegel, Application of the NISTIR 7628 for Information Security in the Smart Grid Architecture Model (SGAM), VDE Kongress, 2014.

[9] Bundesamt fuer Sicherheit in der Informationstechnik, Technische Richtlinie, BSI TR-03109, 2015.

[10] P. Peters and N. Mohr, Digitalisierung im Energiemarkt: Neue Chancen, neue Herausforderungen, Energiewirtschaftliche Tagesfragen, pp. 8-12, 2015.

[11] V. C. Gungor et al., A Survey on Smart Grid Potential Applications and Communication Requirements, IEEE Trans. Ind. Inf. 9 (1), pp. 28-42, 2013.

[12] X. Li et al., Securing smart grid. Cyber attacks, countermeasures, and challenges, IEEE Commun. Mag. 50 (8), pp. 38-45, 2012.

[13] C. Wietfeld, C. Muller, J. Schmutzler, S. Fries, and A. Heidenreich, ICT Reference Architecture Design Based on Requirements for Future Energy Marketplaces, 1st IEEE International Conference on Smart Grid Communications, pp. 315-320, 2010.

[14] F. Dalipi and S. Y. Yayilgan, Security and Privacy Considerations for IoT Application on Smart Grids. Survey and Research Challenges, IEEE 4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), pp. 63-68, 2016.

[15] M. Yun and B. Yuxin, Research on the architecture and key technology of Internet of Things (IoT) applied on smart grid, International Conference on Advances in Energy Engineering, pp. 69-72, 2010.

[16] B. Genge, A. Beres, and P. Haller, A survey on cloud-based software platforms to implement secure smart grids, 49th International Universities Power Engineering Conference (UPEC), pp. 1-6, 2014.

[17] S. Bera, S. Misra, and J. Rodrigues, J.P.C: Cloud Computing Applications for Smart Grid. A Survey, IEEE Trans. Parallel Distrib. Syst. 26 (5), pp. 1477-1494, 2015.

[18] Y. Simmhan, A. G. Kumbhare, B. Cao, and V. Prasanna, An Analysis of Security and Privacy Issues in Smart Grid Software Architectures on Clouds, IEEE 4th USENIX International Conference on Cloud Computing (CLOUD), pp. 582-589, 2011.

[19] Bundesamt fuer Sicherheit in der Informationstechnik, BSI-Standard 100-1 Managementsysteme fuer Informationssicherheit (ISMS), 2008, [Online]. Available from: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/ITGrundschutzstandards/BSI-Standard_1001.pdf?__blob=publicationFile&v=2 [retrieved: 02, 2020].

[20] Bundesamt fuer Sicherheit in der Informationstechnik, BSI-Standard 200-2 IT-Grundschutz Methodology, 2017. Available from: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/International/bsi-standard-2002_en_pdf.pdf?__blob=publicationFile&v=1 [retrieved: 02, 2020].

[21] Bundesamt fuer Sicherheit in der Informationstechnik, BSI Standard 200-3: Risk Analysis based on IT Grundschutz, 2017. Available from: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/International/bsi-standard-2003_en_pdf.pdf?__blob=publicationFile&v=2 [retrieved: 02, 2020].

[22] ISO/IEC Information Technology Task Force, ISO/IEC 27000:2018 Information technology Security techniques Information security management systems Overview and vocabulary, 2018.

[23] R. Matuleviius, N. Mayer, H. Mouratidis, E. Dubois, P. Heymans, and N. Genon, Adapting Secure Tropos for Security Risk Management in the Early Phases of Information Systems Development, International Conference on Advanced Information Systems Engineering, pp. 541-555, 2008.

[24] P. Bresciani et al., Tropos: An Agent-Oriented Software Development Methodology, Autonomous Agents and Multi-Agent Systems 8, pp. 203236, 2004.

[25] D. Mellado, C. Blanco, and L. Sanchez, A systematic review of security requirements engineering, Computer and Standards & Interfaces, Volume 32, Issue 4, pp. 153-165, 2010.

[26] L. Compagna et al, How to integrate legal requirements into a requirements engineering methodology for the development of security and privacy patterns, Artif Intell Law 17, pp. 130, 2008.

[27] K. Boroojeni, M. Amini, and S. Iyengar, Smart Grids: Security and Privacy Issues, Springer International Publishing, 2017.

[28] Ernst u. Young GmbH, Kosten-Nutzen-Analyse fuer einen flaechendeckenden Einsatz intelligenter Zaehler, 2013.

[29] International Organization for Standardization, ISO/IEC 20924:2018 Information technology - Internet of Things (IoT) - Vocabulary, 2018.

[30] Heise Medien, Wachsende Bedrohung durch unautorisierte IoT-Geraete, 2020. [Online]. Available from: https://www.heise.de/ix/meldung/Wachsende-Bedrohung-durch-unautorisierte-IoT-Geraete-4668472.html [retrieved: 02, 2020]

[31] B. Herzberg, I. Zeifman, and D.Bekerman, Breaking Down Mirai: An IoT DDoS Botnet Analysis, 2016. [Online]. Available from: https://www.imperva.com/blog/malware-analysis-mirai-ddos-botnet [retrieved: 01, 2020].

[32] The OWASP Foundation, OWASP Internet of Things, 2018. [Online]. Available from: https://owasp.org/www-project-internet-of-things/ [retrieved: 02, 2020].

[33] R. Sichler, Smart und sicher geht das?, Springer Fachmedien Wiesbaden, pp. 463-494, 2014.

[34] S. Babar, A. Stango, N. Prasad, J. Sen, and R. Prasad, Proposed embedded security framework for Internet of Things (IoT), Electronic Systems Technology (Wireless VITAE), 2010.

[35] C. Eckert, IT-Sicherheit, Konzepte - Verfahren - Protokolle, Boston De Gruyter, 2012.

[36] L. ben Othmane, H. Weffers, and M. Klabbers, Using Attacker Capabilities and Motivations in Estimating Security Risk, Symposium On Usable Privacy and Security, 2013.

[37] The OWASP Foundation, OWASP Risk Rating Methodology, 2019.

[38] The OWASP Foundation, Internet of Things Project, IoT Vulnerabilities, 2019.

[39] Deutsches Institut fuer Normung, DIN SPEC 27072: Informationstechnik - IoT-fhige Gerte - Mindestanforderungen zur Informationssicherheit, pp. 1-16, 2019.

[40] United Kingdom Department for Digital, Culture, Media & Sport, Government response to the Regulatory proposals for consumer Internet of Things (IoT) security consultation, 2020.

[41] European Telecommunications Standards Institute, Draft ETSI EN 303 645 V2.0.0, pp. 1-30, 2019.

[42] Konferenz der Datenschutzbeauftragten des Bundes und der Lnder und Dsseldorfer Kreis, Orientierungshilfe datenschutzgerechtes Smart Metering, 2012. [Online]. Available from: https://www.datenschutz.rlp.de/fileadmin/lfdi/Dokumente/Orientierungshilfen/oh_smartmeter.pdf [retrieved: 02, 2020].

[43] K. Neubauer, S. Fischer, and R. Hackenberg, Work in Progress: Security Analysis for Safety-critical Systems: Smart Grid and IoT, ARCS Workshop, 32nd International Conference on Architecture of Computing Systems, pp. 1-6, 2019.

An Improved Adaptive Beamforming-based Machine Learning Method for Positioning in Massive MIMO Systems

Chong Liu, Hermann J. Helgert

School of Engineering and Applied Science, The George Washington University,
Washington, DC 20052
Email: cliu15@gwu.edu, hhelgert@gwu.edu

***Abstract*—Outdoor localization will become very essential in the development of 5G applications. Current localization techniques mainly relying on GPS and sensors can mostly overcome problems caused by path loss, background noise and Doppler effects, but multiple paths in complex indoor or outdoor environments present additional challenges. In this paper, we propose an improved adaptive *BeamMaP* that can instantaneously locate users in dynamic environment urban after training input data and steer the beams efficiently in a distributed massive Multiple-Input Multiple-Output (MIMO) system. We also design an adaptive algorithm to improve the performance of the model under the dynamic weather. To simulate a realistic environment, we evaluate the positioning accuracy with multiple channel fingerprints collected from uplink Received Signal Strength (RSS) data, including Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS), in the training data sets. Based on the adaptive beamforming, we employ the Rice distribution to sample the current mobile users locations in the testing data sets. Our simulation results achieve Reduced Root-Mean-Squared Estimation Error (RMSE) performances with increasing volume of training data, and the performances of RMSE are very close to the Bayesian Cramer-Rao bounds. We prove that our proposed positioning model is more efficiency and steadier compared with kNN and SVM in the dynamic weather conditions, and also demonstrate the effectiveness of the adaptive beamforming model in the online testing process.**

Keywords–outdoor localization; machine learning; adaptive; beamforming.

I. INTRODUCTION

The future developing technologies, such as autonomous vehicles, Virtual Reality (VR), high-speed data center networking and the Internet of Things (IoT), are relying on more efficient bandwidth distribution and higher speed transmission [2] [3] [4] [5] [6]. Meanwhile, the next generation of wireless networks 5G should provide more accurate localization of the connected mobile devices and distribute the limited bandwidth in a more efficient way. Some new technologies employed in localization, especially including the massive MIMO and beamforming technologies, are explored in the 5G system [7]. The innovative design of massive MIMO disclosed in some publications utilizes a large number of upgraded array antennas (more than one hundred) to multiplex messages for several devices simultaneously. This component, implemented in future Base Stations (BSs), has been shown to play an essential role in positioning of Mobile Users (MUs) in cellular networks, including increased spectral efficiency, improved spatial diversity, and low complexity [8]. More importantly, a distributed design for massive MIMO is beneficial for positioning due to the better spatial diversity, which will be employed in this paper. Some proposed solutions applying the MIMO positioning techniques are mainly focused on the received signal information from the users, such as the Angle-of-Arrival (AoA), Time-of-Arrival (ToA), and Received Signal Strength (RSS) [9] [10] [11]. These features, singly or in combination, can be used in the localization of mobile users in indoor or outdoor environments.

Even though positioning in cellular networks widely uses the Global Positioning System (GPS) in urban or rural areas, the method becomes unreliable when the LoS and NLoS signals are difficult to distinguish, such as in highly cluttered multipath scenarios (tens meters error) [12]. In addition, it consumes the phone battery quickly on GPS. In some conventional method using the two-step localization techniques, the received LoS signals are processed at different base stations and AoA and/or ToA of each user can be obtained. Then the position of the user can be found by triangulation calculation [10]. However, the LoS path may be damped or obstructed, leading to large positioning errors, as is often the case in complex scenarios. Also, [10] is exploiting channel properties to distinguish LoS from NLoS signal paths, resulting in an improvement of performance. However, a large data gain with a combination of LoS and NLoS signal paths will require high computational complexity.

Our solution is to employ a machine learning regression technique based on the efficient beamforming transmission patterns to estimate the location of MUs after collecting amounts of LoS and NLoS data. Our model can instantaneously predict the locations of MUs after generating the Machine Learning (ML) regression network model and help the base stations to distribute beams in an efficient way. Moreover, the proposed design with improved adaptive algorithm can implement the real-time detection to update the input data sets including LoS and NLoS multipath channels. The main contributions of our work are as follows:

- We employ a supervised machine learning regression approach to accurately locate the MUs in a single cellular system.
- We present extensive performance results from simulations exploring the effects of various componential parameters.
- We prove our proposed machine learning method is more efficiency and steadier in the positioning system compared with kNN and SVM.
- We build different testing users models to compare

adaptive and switched beamforming in proving the adaption of our ML model.

The new contributions over [1] includes:

- We add the adaptive algorithm to better initialized before training the input data in the machine learning process.
- We compare the performance of RMSE under different size antennas with Bayesian Cramer-Rao bounds and prove the correction.
- We add the extensive experiments in different weather environments to prove the better performance in our adaptive model.
- We compare our machine learning method with different regression models in the dynamic environments and indicate the better performance as tradeoff between localization bias and response time.

The rest of this paper is organized as follows: Section II discusses different kinds of machine learning localization methods in the wireless network system. Section III presents the proposed positioning system design, including the input data sets collected for training, the machine learning model and testing process. In Section IV, we design our adaptive algorithm in improving the performance of whole system. In Section V, we present performance evaluation results to analyze the impact factors and implement the comparison in different schemes. Section VI presents our conclusions and future works.

II. Related work

In this section, we illustrate the applications of machine learning used for wireless localization and express our proposed method and contributions.

A. Machine Learning methods used for localization

Big data collections combined with machine learning methods have been widely mentioned in solving the mobile users indoor or outdoor localization in some proposed literature [11] [13] [14]. Similar to our proposed [1], some existing innovative ML methods are also commonly based on collecting some signal information as location fingerprints including RSS, ToA or AoA, and through training, modeling and testing to implement the localization.

Hossain et al. [15] and Xie et al. [16] introduce an unsupervised machine learning technique k-Nearest Neighbors (kNN) or an improved kNN scheme to solve the indoor localization problem. Through collecting RSS as fingerprint using Bluetooth and Wi-Fi signals from multiple access points (APs) in [15], a designed regression method is introduced to reduce the training time and facilitate under-trained location systems. The principle behind nearest neighbor methods is to find a predefined number of training samples closest in distance to the new point, and predict the label from these samples. Hossain et al. [15] employ kNN and Bayesian probabilistic model as the regression algorithms for localization in a lecture theater environment. Also, an improved kNN as Spearman-distance-based indoor location system is mentioned in [16], the spearman rank correlation coefficient being as a label metric is calculated after obtaining the unknown position fingerprints (RSS). The spearman distance is acquired based on the spearman rank correlation coefficient and used to combine with the original kNN approach, which proves an improvement performance compared with original kNN.

Tran et al. [17] and Kim et al. [19] proposed a supervised machine learning technique, that is, support vector machines (SVM) to estimate the geographic locations of users in a wireless sensor network where most sensors are without owning an effective self-positioning functionality. Even though SVM is a classification method, it is proved that the localization error can be decreased after given by an appropriate training data size and kernel functions in [17] [19]. Tran et al. [17] assumed that each node is repeatedly positioned as the centroid of its neighbors until convergence. The training data sets are collected through beacon nodes information where two nodes can communicate with each other if no signal blocking entity exists between them. The kernel function used for training is defined based on hop counts only. Kim et al. [19] build the training model based on the raw RSS data sets measured from each sensor. Then a least-square SVM mechanism is explored and implemented on a designed kernel function. Both of them confirmed the estimation performance more accurate and robust than the conventional method.

The supervised deep learning techniques are also employed in the positioning systems due to the higher performance compared with traditional methods in [21] [22]. Rizk et al. [21] introduce the data augmentation method to generate synthetic data with pairs of CID (represents the cell tower unique ID) and RSS fingerprints and utilize the deep learning approach to train the received generated data. A neural network including three hidden layers is designed and processed the training step. The proposed system can receive the improved performance in the evaluation of indoor and outdoor scenarios. Also, another novel deep learning indoor localization system termed as DeepFi is presented in [22]. DeepFi system architecture composed of an offline training phase and an online localization phase utilizes the deep learning method to train all the weights of a deep neural network. The input training data as fingerprints are the channel state information (CSI) collected from some Wi-Fi network interface, which calculated from many subcarriers in an orthogonal frequency division multiplexing (OFDM) system. DeepFi scheme was validated in the representative indoor environments.

However, those efficient supervised or unsupervised techniques still have some limitations in localization of wireless networks. For example, kNN employed in [15] [16] are able to provide good performance in uniformly distributed references, but we have to choose a better regression depending on the different k dimension. The changing k process will generate the large number of input training data and cause higher computational complexity. Also, supported machine learning methods, such as Support Vector Machines (SVM) [17] [19] are easy to cause over-fitting in the regression when the number of features is much greater than the number of samples, so it relies on large numbers of sensors to acquire the data in the wireless sensor network. Thus, when the number of MUs in the outdoor increases, it will increase the time computational complexity to distinguish the LoS and NLoS

signals from multiple different sensors and need more cross-validation iterations to avoid over-fitting. Additionally, deep learning method [21], is explored to predict the coordinates of MUs after collecting amounts of RSSs and/or AoAs through different base stations. It will cause the estimation to be degraded when the number of MUs increases and interference between cellular areas becomes dramatically higher. Even though DeepFi scheme [22] was validated in two representative indoor environments, it ignores the complexity of the dynamic environment if implemented in outdoor network.

In addition, current research in exploiting the machine learning techniques points out that offline optimization can also dramatically improve the speed of test processes and the accuracy of estimation through collecting a considerable amount of multiple channel paths parameters. However, the impact of realistic aspects such as multiple channels in different paths sent from MUs are not all considered. It means that only LoS channels in the cellular networks are considered and some strong NLoS signals in the urban areas are ignored. Moreover, [11] considers the magnitude of a channel snapshot represented in a sparse domain and translates it into a convolutional neural networks (CNNs) image identification problem, which is constrained on the fixed data array such as delay and angles in the static LoS channels. It ignores the real-time channel variations that are not presented in the training data sets.

Based on the features of raw data sets, such as RSS information that is easier to be collected and run in our system, we employ a Gaussian Process Regression (GPR) model to estimate the locations of MUs, discussed in [9]. GPR is a generic supervised learning method designed to solve regression and probabilistic classification problems. Under this method, an unknown nonlinear kernel function is assumed to be random, and to follow a Gaussian Process (GP). In contrast to kNN and SVM, GPR is able to provide probabilistic output, for example, the posterior distribution of the MU position, after given an online measurement and a set of fingerprints with RSS vectors. Besides, without LoS and NLoS identification, this machine learning approximation method can efficiently identify MUs positions after training with limited reference users, and it significantly decreases the computational complexity as well.

B. Our Approach and Contributions

In this paper, we propose a novel positioning technique, called Beamforming-based Machine Learning for Positioning (*BeamMaP*) to meet the above challenges. *BeamMaP* employs a machine learning regression technique based on the efficient beamforming transmission patterns in order to estimate the location of MUs. *BeamMaP* can instantaneously predict the locations of MUs after generating the Machine Learning (ML) regression network model and help the base stations to distribute beams in an efficient way. Moreover, *BeamMaP* can implement the real-time detection to update the input data sets including LoS and NLoS multipath channels, and also an improved adaptive *BeamMaP* can adequately satisfy with dynamic atmosphere in the 5G system.

The *BeamMaP* design is illustrated in Figure 1. The beamforming system in each BS installed massive MIMO antennas serves more than one MU. When a MU transmits on the uplink, we can obtain a vector of RSS (or a fingerprint)

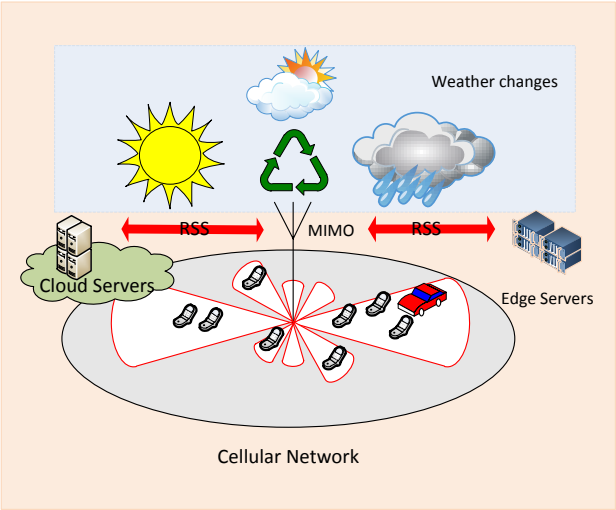


Figure 1. BeamMaP positioning system in cellular networks.

comprising LoS and NLoS multipath measured by the massive antennas array in the BS. The detected uplink signals or RSS information are collected and submitted to the edge servers or cloud servers for calculation. Then the adaptive array systems can formulate a single or more beams with different weights to different directions according to the demands of MUs. Furthermore, MUs can process signals from a single MIMO base station, provided the BS and users were synchronized, which can be easily implemented by a two-way protocol by adding some additional overheads [23]. Besides, in order to avoid the pilot contamination occurred in massive MIMO system between cells, some reuse pilot schemes and particular modulation technology, such as Orthogonal Frequency-Division Multiplexing (OFDM) or Code-Division Multiple Access (CDMA) should be applied in our system [24]. Furthermore, massive MIMO systems combined with beamforming antenna technologies are considered to play a key role in the next generation wireless communication systems [25]. Optimal beamforming techniques, such as adaptive beamforming, are mentioned to be employed in localization and provide energy saving of the MIMO systems. *BeamMaP* employs adaptive beamforming as a candidate in building the testing process. Compared with switched beamforming, adaptive beamforming can cover a larger area of MUs when the number of beams and bandwidths range shared are the same, and it also offers more comprehensive interference rejection [25]. Therefore, *BeamMaP* not only can improve the efficiency of coverage for users, but can also result in significant reduction in energy consumption of base stations.

III. BeamMaP Positioning System and Algorithm Design

Driven by the above motivations, the *BeamMaP* framework is illustrated in Figure 2.

We firstly need to collect the fingerprints (RSS vectors) to generate the training data sets. Due to the unknown directions of MUs, we assume the beams weights in a uniform distribution trying to cover more MUs in comparison with the random distribution in the beginning status. Then, *BeamMaP*

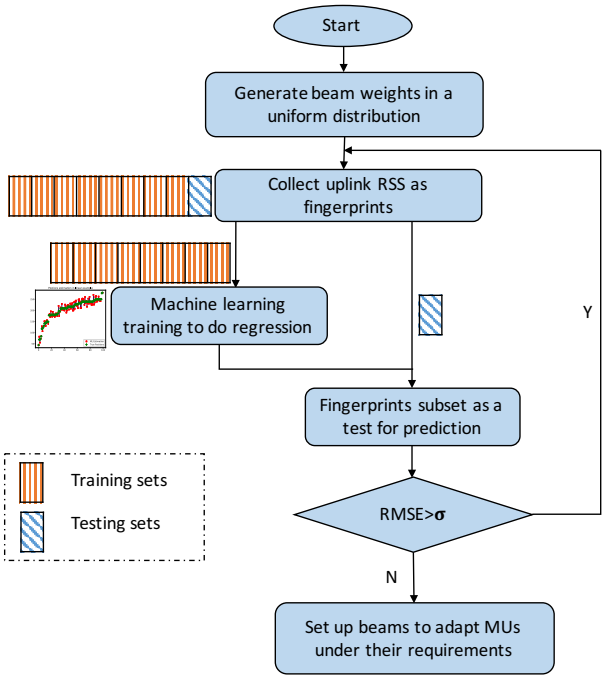


Figure 2. BeamMaP's positioning system framework (adaptive σ chosen)

TABLE I. BASIC NOTATIONS REPRESENTATIVE.

Notation	Corresponding meaning
K, k	the number of antennas in BS, antenna index
M, m	the number of MUs, MU index
ρ	the transmission power of each mobile user
S	the number of training reference MUs
s_m $\mathbf{s}$	the symbol vector transmitted by the mth mobile user, the sum symbol vectors transmitted by all MUs
r_k $\mathbf{r}$	the received symbol vector at the kth antenna in BS, the sum signal strength vectors in BS
$h_{k,m}$ $\mathbf{H}$	fading uplink channel between mth MU and kth antenna , the uplink channel matrix between all MUs and BS antennas
$\alpha_{k,m}$ $q_{k,m}$	small-scale fading coefficient between mth MU and kth antenna , large-scale fading coefficient between mth MU and kth antenna
n_k $\mathbf{n}$	the additive white Gaussian noise vector received in the antenna k the sum additive white Gaussian noise vectors in the BS
$p_{k,m}$ $\mathbf{p}_m$	RSS of mth MU at kth antenna in BS uplink RSS vectors of MU in all antennas of BS
n	the Path Loss Exponent (PLE) for LoS or NLoS channel
δ_s	the shadow fading in dB
$\widetilde{\mathbf{p}}_a$ $\widetilde{\mathbf{P}}$	the uplink RSS vector for the ath training MU the training data matrix for S coordinates of MUs chosen
$\widehat{\mathbf{p}}_m$	the uplink RSS vector of the mth testing MU
$(\widehat{\mathrm{x}}_m, \widehat{\mathrm{y}}_m)$ $(\widetilde{\mathbf{x}}_m, \widetilde{\mathbf{y}}_m)$	the coordinate of the mth testing user in vector $(\widehat{\mathrm{x}}, \widehat{\mathrm{y}})$ the coordinate of the mth training user in vector $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{y}})$
$[\mu^x]_m$ $[\sigma^x]_m$	the estimation value of the mth testing user $\widehat{\mathrm{x}}_m$-coordinate the variance for errors of user $\widehat{\mathrm{x}}_m$-coordinate

starts to explore the GPR method to train the collected raw data arrays, which include the RSSs of LoS and NLoS in the scenario. Some parameters set up in the ML regression model are able to be estimated in the training process. Furthermore, in order to avoid the overfitting in the training process, we follow the K^*-fold cross-validation to partition a sample of input data sets into complementary subsets, performing one subset as the training set (the orange blocks in the figure), and validating the analysis on the other subset as the testing set (the blue blocks in the figure). Multiple rounds of cross-validation are performed using different partitions, and the validation results are combined (e.g., averaged) over the rounds to give an estimate of the model's predictive performance. Moreover, we choose the Root-Mean-Square Estimation Error (RMSE) as the metric, which will be introduced in the experiment section. We set up a threshold σ to analyze the training process of the ML model. If the RMSE in the model is larger than σ, it will back up to the beginning of the ML process, requiring that the ML process continue the training process. If the RMSE is less than or equal to σ, the parameters in the model have been generated successfully in the estimation, and we should adjust the system to set up beams to cover the mobile users under their requirements. The detailed model is designed in the following part.

A. Input Data Sets for Training –Uplink Transmission in 5G MIMO Model

In this section, we build a wireless network model to locate Mobile Users (MUs) in a single cellular 5G network system. We assume one Base Station (BS) with K ($K \geqslant M$) antennas to serve M single-antenna MUs in the cell. We consider MUs simultaneously transmit M symbols, $\mathbf{s} = (s_1, ..., s_M)^T$, the massive MIMO antennas array in the base station can receive the sum signal strength vectors $\mathbf{r} = (r_1, ..., r_K)^T$:

$$\mathbf{r} = \sqrt{\rho}\mathbf{H}\mathbf{s} + \mathbf{n} \tag{1}$$

Here ρ is a constant denoting the transmission power of each mobile user; $\mathbf{H}$ is the $K \times M$ channel matrix, with $h_{k,m} = \alpha_{k,m}\sqrt{q_{k,m}}, \forall k = 1, ..., K$ and $m = 1, ..., M$ as the transmission channel element for mth mobile user uplink to the kth antenna in the base station. $\alpha_{k,m}$ and $q_{k,m}$ are respectively the small-scale and large-scale fading coefficients. The large-scale fading $q_{k,m}$ (related to shadowing noise variance) is assumed to be a constant in the urban or suburban environment, and the small-scale fading $\alpha_{k,m}$ is considered to be an independent and identically distributed complex Gaussian distribution (Rayleigh distribution), with $\alpha_{k,m} \sim \mathcal{CN}(0, 1)$. In addition, $\mathbf{n} = (n_1, ..., n_K)^T$ represents the additive white Gaussian noise vector given by $n_k \sim \mathcal{N}(0, 1)$. We list the basic notations in Table I.

From (1), we are considering the sum signal strength vectors from all users to antennas. In order to separate the multiple users RSS in $\mathbf{r}$, we have different schemes to extract the kth user RSS r_k. In order to capture the effective signals, the pilot signal vector s_k should be modulated as mutually orthogonal during transmission so that it can satisfy $s_i^H \cdot s_j = 0$ $(i \neq j)$ [24]. Particular modulation techniques, such as OFDM or orthogonal CDMA employed as the coded schemes in the transmission systems. Minimum MSE (MMSE) being an appropriate solution, we can simply extract each user signal strength from the combination signals of all MUs and then distinguish the signals and noise by setting a threshold in the receiving part.

$$\mathbf{s}^H\mathbf{r} = \sqrt{\rho}\mathbf{H} + \mathbf{s}^H\mathbf{n} \tag{2}$$

Taken all assumptions into account, we can acquire the single user's RSS as $p_{k,m}$ in:

$$p_{k,m} = \left\| s_m^H r_k \right\|^2 = \rho \left| h_{k,m} \right|^2 = \rho \alpha_{k,m}^2 \left| q_{k,m} \right| \quad (3)$$

Also, we accumulate all MU uplink power vectors from all antennas in BS: $\mathbf{p}_m = [p_{1,m}^{\text{dB}}, p_{2,m}^{\text{dB}}, ..., p_{K,m}^{\text{dB}}]$. Established on the received power model, we can acquire the power data sets by converting (3) to the log distance path-loss model but they are limited in the lower frequency and small cellular environment [26]. Additionally, through our experiment, we observe the COST Hata model (COST is a radio propagation model that extends the urban Hata model to cover a more elaborate range of frequencies, which is developed by a European Union Forum for cooperative scientific research.) also cannot adapt the different higher frequency 5G network system, even though it is popularly employed in the urban cellular network [27]. Also, the path loss models currently employed in the 3GPP 3D model is the ABG model form but without a frequency dependent parameter and additional dependencies on base station or terminal height, and only used in LoS scenario [28]. Therefore, we are considering to employ the Close-in (CI) free space reference distance Path Loss (PL) model, which is noted multi-frequency and covers the 0.5-100 GHz band [28] [29]. The CI-PL model is also transferred from (3) to adapt LoS and NLoS realistic scenarios through adding the free space path loss and optimizing the parameters:

$$\text{P}_{loss}(f_c, d)[\text{dB}] = \text{FS}(f_c, 1\text{m}) + 10n\log_{10}(\frac{d}{1\text{m}}) + \delta_s \quad (4)$$

Here f_c is the carrier frequency in Hz, n is the Path Loss Exponent (PLE) describing the attenuation of a signal passing through a channel, d is the distance between MU and each antenna in BS and δ_s is the shadow fading in dB. The Free Space Path Loss (FS) in (4) is standardized to a reference distance of 1 m. FS with frequency f_c is given by:

$$\text{FS}(f_c, 1\text{m}) = 20\log_{10}(\frac{4\pi f_c}{\nu}) \quad (5)$$

In (5), ν denotes the speed of light. The CI-PL model is represented as the relationship between propagation path loss and TX-RX distance based on a straight line drawn on a two-Dimensional (2D) map, passing through obstructions, and used in both LoS and NLoS environment. While we are considering CI-PL in the urban cellular network of 5G system model, the parameters are measured as $n = 2.0, \delta_s = 4.1$ dB in LoS and $n = 3.0, \delta_s = 6.8$ dB in NLoS using omnidirectional antennas [28]. Due to the same transmission power assumed for each MU, we can use the CI-PL model as the RSS parameters to acquire the training data sets. Additionally, for each MU's uplink transmission, multiple antennas can receive multipath signals, some of them are LoS and the others are NLoS responses. So we consider the LoS probability model in the current 3GPP/ITU model in the MIMO receiving part when setting up the training data. It means the uplink response array of MIMO antenna includes LoS and NLoS components for each MU. From [28], in terms of Mean Squared Error (MSE) between the LoS probability from the data and the models, we choose the d_1/d_2 model as follows:

$$p(d) = \min(\frac{d_1}{d_2}, 1)(1 - e^{-\frac{d}{d_2}}) + e^{-\frac{d}{d_2}} \quad (6)$$

Where d is the 2D distance between MU and antennas in meters and d_1, d_2 can be optimized to fit a scenario of parameters (we choose $d_1 = 20$, $d_2 = 39$ because it acquires minimum MSE in adapting the urban scenario).

B. Machine Learning Model

Given the RSS vector $\mathbf{p}_m = [p_{1,m}^{\text{dB}}, p_{2,m}^{\text{dB}}, ..., p_{K,m}^{\text{dB}}]$, our goal is to find the position of the mth MU in the two dimensional plane, denoted by (x_m, y_m). We build the functions $f_x(.)$ and $f_y(.)$, which take the uplink RSS vector $\mathbf{p}_m$ of a given user m as input and provide the user's location coordinates (x_m, y_m) as output, and try to learn as follows:

$$x_m = f_x(\mathbf{p}_m) \quad \text{and} \quad y_m = f_y(\mathbf{p}_m), \forall x_m, y_m \quad (7)$$

Derived from CI-PL model for the input training model, the learning functions can be classified as a nonlinear regression problem. We follow GPR as a supervised machine learning approach, with a training phase and a test phase, to learn $f_x(\mathbf{p}_m)$ and $f_y(\mathbf{p}_m)$. In the training level, we consider RSS vector $\mathbf{p}_m$ derived from the CI-PL model in both LoS and NLoS conditions. Prior to it, we need to acquire the antennas coordinates, the training users coordinates, and some other parameters. In the testing phase, the testing users are chosen in a Rice distribution to satisfy the adaptive beamforming pattern, whose location coordinates are unknown.

C. Training and Beamforming-based Prediction Phase

GPR uses the kernel function to define the covariance over the objective functions and uses the observed training data to define a likelihood function. Gaussian processes are parameterized by a mean function μ_x and covariance function $\mathbf{K}(\mathbf{p}_i, \mathbf{p}_j)$, which means $f_x(.), f_y(.) \sim \mathcal{N}(\mu, \sigma^2)$. Usually the mean matrix function is equal to 0, and the covariance matrix function, also known as kernel matrix function, is used to model the correlation between output samples as a function of the input samples. The kernel matrix function $\mathbf{K}(.,.)$ contains $k(\mathbf{p}_i, \mathbf{p}_j), \forall i, j = 1, ..., M$ as the entries to define the relationship between the RSS of the users. We usually use a weighted-sum of squared exponential and linear functions, which servers the stationary component and non-stationary component respectively, to generate the regression function:

$$k(\mathbf{p}_i, \mathbf{p}_j) = \upsilon_0 e^{-\frac{1}{2}\mathbf{A}\|\mathbf{p}_i - \mathbf{p}_j\|^2} + \nu_1 \mathbf{p}_i^T \mathbf{p}_j \quad (8)$$

Here $\mathbf{A} = \mathbf{diag}(\eta_k), \forall k = 1, ...K$. It will cover the LoS and NLoS matching with each MU. So the parameters vector $\Lambda = [\upsilon_0, \mathbf{A}, \upsilon_1] = [\upsilon_0, \eta_1 ..., \eta_K, \upsilon_1]$ can be estimated from the training data. In order to learn the target vector $\overline{\Lambda}$, we choose S coordinates of MUs as the training data matrix $\widetilde{\mathbf{P}}$ denoted $\widetilde{\mathbf{P}} = [\widetilde{\mathbf{p}_1}, \widetilde{\mathbf{p}_2} ... \widetilde{\mathbf{p}_S}]$ and use the maximum-likelihood method to predict the $(\widetilde{x}, \widetilde{y})$-coordinates. According to the property of a Gaussian process, we can acquire the learned vector $\overline{\Lambda}$

by employing the maximum-likelihood of the $S \times 1$ training $\widetilde{\mathrm{x}}$-coordinate vector:

$$\overline{\Lambda} = \underset{\Lambda}{\operatorname{argmax}} \log(p(\widetilde{\mathbf{x}}|\widetilde{\mathbf{P}}, \Lambda)) \sim N(\widetilde{\mathbf{x}}; 0, \widetilde{\mathbf{K}}) \tag{9}$$

The parameter vector follows as GP, which is a non-convex function as shown in the [9], and can not be solved well in the training process. Several methods introduced in [30], such as stochastic gradient descent, mini-batching or momentum, can help to solve the non-convex problem. Established on the ML method in the training problem, we decided to employ stochastic gradient descent method [30] to obtain the optimum vector $\overline{\Lambda}$ in convergence to a local maximum.

In the prediction phase, the predictive distribution $p(\widehat{\mathrm{x}}_m|\widetilde{\mathbf{P}}, \widetilde{\mathbf{x}}, \widehat{\mathbf{p}}_m)$ in terms of posteriori density function, is applied as estimation of the testing user $\widehat{\mathrm{x}}_m$-coordinate, which also follows the Gaussian distribution with mean $[\mu^x]_m$ and variance $[\sigma^x]_m$, $\widehat{\mathrm{x}}_m|\widetilde{\mathbf{P}}, \widetilde{\mathbf{x}}, \widehat{\mathbf{p}}_m \sim \mathcal{N}([\mu^x]_m, [\sigma^x]_m)$:

$$\begin{aligned}
[\mu^x]_m &= \sum_{a=1}^{S} k(\widehat{\mathbf{p}}_m, \widetilde{\mathbf{p}}_a)[\widetilde{\mathbf{K}}^{-1}\widetilde{\mathbf{x}}]_a, \\
[\sigma^x]_m &= k(\widehat{\mathbf{p}}_m, \widehat{\mathbf{p}}_m) - \sum_{a=1}^{S}\sum_{b=1}^{S} k(\widehat{\mathbf{p}}_m, \widetilde{\mathbf{p}}_a)[\widetilde{\mathbf{K}}^{-1}]_{ab} \cdot k(\widetilde{\mathbf{p}}_b, \widehat{\mathbf{p}}_m)
\end{aligned} \tag{10}$$

Where the mean $[\mu^x]_m$ indicates the estimation value of test user $\widehat{\mathrm{x}}_m$-coordinate and the variance $[\sigma^x]_m$ represents the variance for errors of user $\widehat{\mathrm{x}}_m$-coordinate. $\widehat{\mathbf{p}}_m$ denotes the received power vector of the mth testing MU, and $\widetilde{\mathbf{p}}_a$ denotes the ath power vector in the received training power matrix $\widetilde{\mathbf{P}}$. For the computational complexity of GPR, we observe from (10), $[\mu^x]_m$ needs to sum up S operations for $\widetilde{\mathbf{K}}^{-1}\widetilde{\mathbf{x}}$, which requires $\mathcal{O}(S^2)$. In total, $[\mu^x]_m$ incurs a time complexity of $\mathcal{O}(S^3)$.

Subsequently, we choose the locations of test MUs based on the beamforming pattern. Beams can be optimized to distribute and spread with the demand users. In the real scenarios, some hot spot areas need large bandwidth and some other areas only need small bandwidth to satisfy with few mobile users. The locations of MUs always follow a Rice distribution. Therefore, the coordinates of test users in positions prediction can be chosen from input fingerprints following a Rice distribution, which will satisfy with the beams distribution in an adaptive way. *BeamMaP* being as a prediction assistant, it will cooperate with a better beamforming scheme to distribute the bandwidths in efficiency. During the experiments, we will compare with switched beamforming patterns, which beams are distributed uniformly in the system. Furthermore, we employ the same proposed regression method to estimate the $\widehat{\mathrm{y}}_m$-coordinate of test user. Also, we can acquire the mean $[\mu^y]_m$ and variance $[\sigma^y]_m$ as the predictive parameters.

IV. Adaptive Algorithm Design for the Selection of the Input Data Sets

In the machine learning design process, both initialization and momentum are known to be crucial since poorly initialized network can not be trained well [31]. For the training phase in the machine learning process, the selection of input data sets should be of importance in training the ML model. According to some proposed papers, wireless communications suffers a RSS loss or degrade in the network quality during bad weather or climatic change, which can affect the regional communication. The effects of atmosphere in RSS need to be considered in the analysis of dynamic outdoor conditions [32] [33]. So we realize that the rain volume will affect the signal attenuation in some range especially in the crowd cities. If the environment of testing data sets is different from the training sets, it will definitely cause the increase of the estimation error rate in the testing. Therefore, before starting the training process, we learn that the selection or classification of input data sets can better improve the performance.

Even though our previous chapter *BeamMaP* [1] is considered in the relatively stable outdoor condition, the effects of atmosphere in RSS will be considered in the designing the adaptive algorithm for selection of the training data. In order to adapt to the different environments in the outdoor urban, we adopt the different training data sets. In the transmission of wireless signals, attenuation is due to the scattering and absorption of electro magnetic waves by drops of liquid water, temperature and humidity [33]. However, we collect the data and do the training in the day time cycle, and find that temperature and humidity have not much fluctuation in hours. Then rain is shown as a major source of attenuation for microwave propagation above 5 GHz especially in 5G system [34] [35] [41]. The signal attenuation increases as its wavelength approaches the size of a typical raindrop (1.5 mm). Thus we will employ the different regular weather conditions such as sunny, drizzle (including cloudy) and rainy (including showers) in the dynamic environments. In the initialization process, we will manually choose the different data groups to do the training process in these different conditions. In order to realize the dynamic model, the status of weather conditions will be classified into $S[0]$, $S[1]$ and $S[2]$ depending on the volume of rain in the time slots. In the practice, we usually add the rain volume sensor in the antennas to help and decide the status of weather. For example, when the sensor finds that rain volume is zero (Sunny status) at that time, we will employ $S[0]$ data sets in the training process. The selection process is the initialization step in our ML model. It will help to implement the localization estimation in adapting with different weather. Then, we can start the machine learning algorithm and testing in the following steps.

The pseudocode of the algorithm is shown in Algorithm 1.

V. Performance Evaluation

In this section, we conduct simulations to evaluate the performance of *BeamMaP* as the machine learning method in estimating the locations of testing MUs. In order to simulate a realistic environment, we set up the fundamental parameters of path loss model based on the 5G 3GPP/ITU Micro-Urban model [28].

A. Parameters Set Up

The parameters used in the simulation are shown in Table II. According to the analysis of different environment in

Algorithm 1 ADAPTIVE ALGORITHM FOR POSITIONING

1: Initialize: Initial positions and set up the beams in a uniform distribution.
2: **for** $i = 1...k(k = 3)$ **do**
3: **if** Rain volume is zero **then**
4: Choose status $S[0]$
5: **if** Rain volume is small or medium **then**
6: Choose status $S[1]$
7: **if** Rain volume is large **then**
8: Choose status $S[2]$
9: **Input:** Measurement data sets in the $S[i]$ condition.
10: Compute $\mathbf{K}(\mathbf{p}_i, \mathbf{p}_j), \forall i, j = 1...M$,
$[\mu^x]_m = \sum_{a=1}^{S} k(\widehat{\mathbf{p}}_m, \widetilde{\mathbf{p}}_a)[\widetilde{\mathbf{K}}^{-1}\widetilde{\mathbf{x}}]_a$
11: **Until** | RMSE| $\leq \sigma$
12: **Output:** Estimated target position $\widehat{\mathrm{x}}_m = [\mu^x]_m$, set up beams in a specific directions according to the location distribution.

TABLE II. PARAMETERS FOR SIMULATION.

Description	Value
Path loss parameters (5G 3GPP/ITU Micro-Urban model [28])	$n = 2.0, \delta_s = 4.1$ dB for LoS, $n = 3.0, \delta_s = 6.8$ dB for NLoS, $d_1 = 20$, $d_2 = 39$
Modulation Scheme	OFDM (Orthogonal CDMA)
MU transmit power	23 dBm (200 mW)
Minimum SNR for channel estimation	1 dBm
Number of antennas in BS	64 (8×8), 100 (10×10), 144 (12×12)
Maximum number of training fingerprints	90000
Number of testing MUs	100
The space between antennas	0.12, 0.3, 0.48 m
The space between training MUs	1 m
Threshold to control the training process (σ)	[5, 35] m

Section III-A, the path loss parameters n and δ_s are chosen for adapting the crowded urban area. The MU transmit power is chosen as per LTE standards to be 23 dBm [36]. In practice testing, the minimum SNR required is determined by the normalized MSE of the channel estimates [28]. For our simulations, we set the minimum required SNR to 1 dB. Considering that currently the number of MIMO antennas of the BS can be designed from 64 to 156, we assume $K = 64, 100, 144$ antennas uniformly distributed as a 8×8, 10×10 and 12×12 squares. We assume that the MIMO antennas are installed at the center of a cellular network, which can distribute the beams in each direction with the same maximum reach. Figure 3 shows an example of the deployment of the base station antennas and the surrounding reference MUs consisting of a squared antennas array with 16 antennas covering $x \in [5, 30]$ and $y \in [10, 70]$ area (meters in unit). The fingerprints for MUs are distributed in a grid covering dimensions $x \in [50, 130]$ and $y \in [20, 140]$. We split the fingerprints into a training part and a testing part, then follow the K^*-fold cross-validation method (i.e., $K^* = 10$) to do the regression and average the result over several runs.

The coordinates of MUs and antennas are selected as positive values in the simulation. In order to reduce the interference between the uplink received signals in the massive MIMO, spatial separation for antennas is on the minimum order of 2 to 3 wavelengths and usually in 5 to 8 wavelengths (or more)

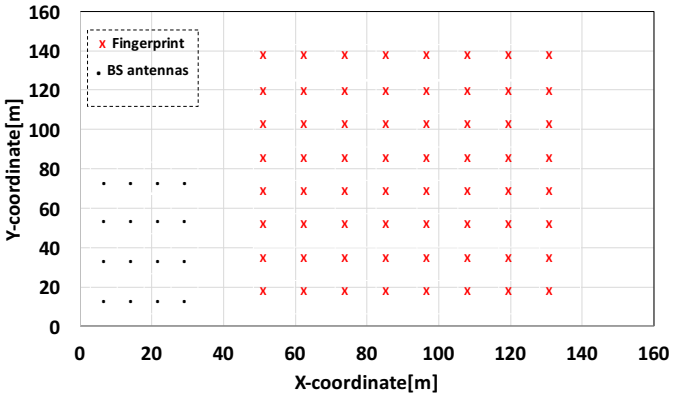


Figure 3. The deployment example of MIMO antennas (BS antennas) and reference MUs (Fingerprint)

[37] [38]. In our simulation, the spacing between antennas can be selected from 0.12 to 0.48 m, which is based on the times of 5 GHz OFDM signal wavelength (5 GHz = 6 cm). If without considering the influence of the other parameters, we assume the space between antennas be 0.48 m to better differentiate the RSS vectors in the simulation. In addition, we choose $S = 90000$ as the maximum number of fingerprints with 1 meter spacing between MUs in a grid covering about 300 m $\times$ 300 m, which covers 95% of LoS components in the single cellular system. In practice, for example, we can install a cellular BS with a 12 $\times$ 12 square antennas on the top roof of our engineering building located in Washington DC of United States. Each antenna equipped with one transceiver can receive and/or send the signals from and/or to each MU. The coordinates of references MUs will be chosen in a grid around the building, the spaces between MUs are set up as 1 meter. We can use a moving MU in each chosen locations to send the signals to all the receivers in BS each time. The computers as a RSS reader in BS will calculate each RSS vector from the signals of the reference MUs and accumulate all the uplink RSSs as the training data sets. Due to lack of hardware support, the RSS vector $\mathbf{p}_m$ for each MU in antennas is generated from the CI-PL model in (4) and (5), which has been proved in the Aalborg, Denmark environment [28].

Meanwhile, each antenna in MIMO can receive LoS or NLoS from the different direction. In order to model the real-life scenario including LoS and NLoS, the RSS matrix $\widetilde{\mathbf{P}}$ as the fingerprints collected from all antennas follows the LoS and NLoS distribution in (6). We calculate them through generating a probability function in the simulation. During the training phase, while we are learning the parameter vector $\overline{\Lambda}$, we run the training locations on randomly choosing the start points (numbers of training references vectors can be chosen in the different order), so as to avoid the convergence to a bad optimal solution. We assume the threshold $\sigma \in [5, 35]$ m, which needs to be feasibly chosen depending on the different training data sets to fit in the experiment. In the testing phase, we choose the Rice distribution of test users from the fingerprint RSS vectors to efficiently steer beams in a flexible way. We follows that $R \sim \mathrm{Rice}\left(|\nu|, \sigma\right)$ has a Rice distribution if $R = \sqrt{X^2 + Y^2}$ where $X \sim N\left(\nu \cos\theta, \sigma^2\right)$ and $Y \sim N\left(\nu \sin\theta, \sigma^2\right)$ are statistically independent normal random variables and θ is any real number. The testing mobile users can be distributed in any

direction. Considering the testing MUs are around the antennas array, we assume that the maximum distance between central of antenna array and test mobile users set up as $\nu = 150$ meters, and the variance distance between adjacent testing mobile users set up as $\sigma = 1$ meter. The Rice distribution is selected as $R \sim \text{Rice}(150, 1)$ through experiments because of the maximum 150 meters coverage of a single cell network and variance of spacing in 1 m.

B. Performance on Metrics in Static Environment

After RMSE is reaching less than σ, we test the accuracy of the simulation model in using the linear sampling coordinates, which are convenient to observe. For example, we use a 12 $\times$ 12 antenna array located in $x \in [40, 46]$ and $y \in [100, 106]$ area as reference locations. In order to observe the tracking locations in a 'linear' status, we initialize to employ a linear log-function ($y = 50\log x$) to sample the positions of 100 testing mobile users from fingerprints within [0, 100]. The X coordinates keep the same in the comparison results. We can then track the MUs and compare with their true positions, as shown in Figure 4(a). It is simple to find the estimated position of testing users not far from the linear true positions 'line', where the interval between them can not exceed 8.5 m. Due to the limitation of test users and sampling, we are not able to decide other impact factors for the accuracy of estimation. Additionally, we choose the testing target users in random route distributed within $x \in [0, 100]$,$y \in [0, 210]$ area and distributed in sparse distance to predict the X-coordinate and Y-coordinate at the same time. The red dots represent the ML estimation position, and the green dots are the true users position. It is shown in Figure 4(b) that the proposed ensemble method receives the expected results, which the average location error is around 5 meters much less than the conventional methods results.

Furthermore, we use the RMSE as the metric to analyze the performance of the estimation methods. RMSE is formulated as:

$$\text{RMSE} = \sqrt{\frac{\sum_{m=1}^{\widehat{M}} (\hat{\text{x}}_m - [\mu^x]_m)^2 + (\hat{\text{y}}_m - [\mu^y]_m)^2}{\widehat{M}}} \quad (11)$$

Where $[\mu^x]_m$ and $[\mu^y]_m$ are the estimation of test user's coordinates $\hat{\text{x}}_m$ and $\hat{\text{y}}_m$, respectively. $\widehat{M}$ is the number of testing MUs.

In Figure 5, we are trying to determine the influence of training samples for different number of antennas in the base station. As the antennas are installed in a fixed space, some of them will receive the LoS signals and others will receive the NLoS signals. The distribution between LoS and NLoS follows the probability function of LoS in (6), as assumed previously. We show 95% confidence intervals from 30 trials for each data point. As observed from Figure 5, we know when the sampling in training locations increases, the RMSE keeps decreasing with fixed antennas size, which means acquiring the higher the accuracy of estimation. When the sampling is the same, more LoS signals will be received in the large size antenna array, which will help to decrease the interference, while fewer NLoS signals will be identified as LoS in the receiver. For example,

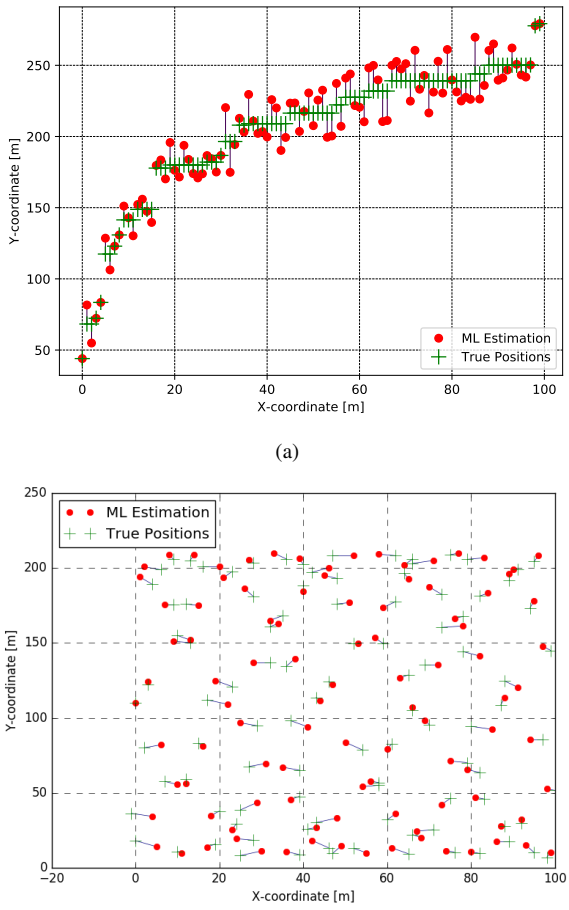


Figure 4. Position estimation of Testing MUs (a) in a linear distribution, (b) in random distribution.

RMSE in 12×12 antennas is almost half of 8 $\times$ 8 in the same sampling condition. Also, the higher dimension of fingerprints for training will acquire more accuracy estimation in the terms of the increase number of antennas.

In order to know the effect of antenna size in a MIMO system, we change the spacing between antennas as in Figure 6. The RMSE for different spacing but the same number of antennas shows no significant change. When the space is changed from 0.12 m to 0.30 m, the differential in RMSE for 8 $\times$ 8, 10 $\times$ 10 and 12 $\times$ 12 antennas is 5 m on average. However, comparing the spacing in 0.12 m and 0.48 m, the RMSE is dramatically decreased, caused by the ability of identification between LoS and NLoS, and the size of sampling.

C. Adaptive Algorithm Implementation in the Dynamic Environment

It is well known that Bayesian Cramer-Rao bound (BCRB) is an optimistic bound in a non-linear estimation problem where the outliers effect generally appears, leading to a quick increase of the MSE. This threshold effect is not predicted by BCRB. The particular value for which the threshold effect appears is a necessary feature to define the estimator

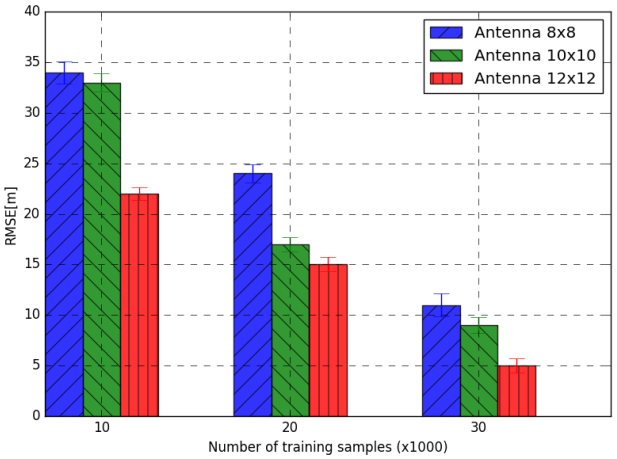


Figure 5. RMSE vs. number of training samples for different size of antennas array

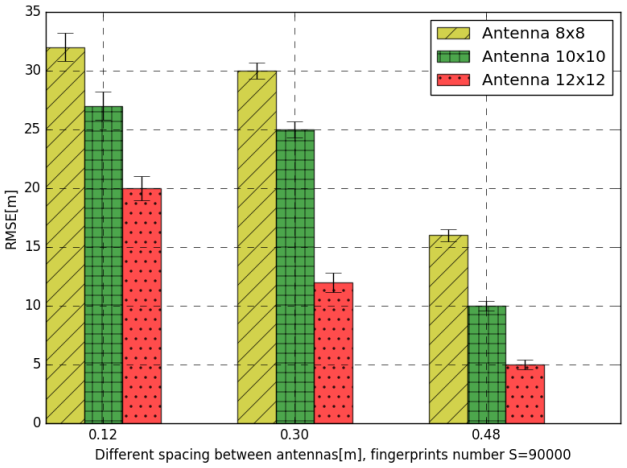


Figure 6. RMSE vs. different spacing between antennas for different size of antennas array

optimal operating area and tightness becomes a prominent quality looked in non-linear estimation problems proved in [39]. Similarly, in order to analyze the RMSE performance of Gaussian Process method, we employ the BCRB to observe the tightness with the different noise level for LoS and NLoS signals:

$$\text{BCRB} = \sqrt{\frac{1}{\widehat{M}}(\text{Tr}([\sigma^x]_m + [\sigma^y]_m))} \qquad (12)$$

Where $[\sigma^x]_m$ and $[\sigma^y]_m$ are the variances for errors of users in $(\widehat{\text{x}}_m, \widehat{\text{y}}_m)$-coordinates and $\widehat{M}$ is the number of testing MUs. We assume LoS and NLoS signals with the same shadowing noise but different Path Loss Exponent (PLE) to distinguish.

In Figure 7, we plot the BCRBs on the RMSE performance of the GP methods under study, setting the shadowing noise level δ_s for LoS and NLoS to change from 1 dB to 6 dB, which can be regarded as different scenarios in practice. We employ the two different antennas sizes $K = 8 \times 8$, $K = 12 \times 12$ to observe. After through the relative large training process in $S = 90000$ and testing, the achieved RMSE are very close to the theoretical BCRBs for $K = 64$ and $K = 144$. With the increase of noise, the RMSE will become larger but in the accepted range. We also find that the BCRBs are tighter for a larger K. It is expected in the reason of the receiver sensitivity. When the number of antennas in the BS becomes larger, the receiving experience of test RSS values will keep in the high sensitive level. At the same time, the receiving matrix in RSS will be generated in higher efficiency. Otherwise, with the smaller K, a smaller fraction of the total number of antennas in the base station would experience receiving RSS below the receiver sensitive level and will cause a small amount of information loss in the training process.

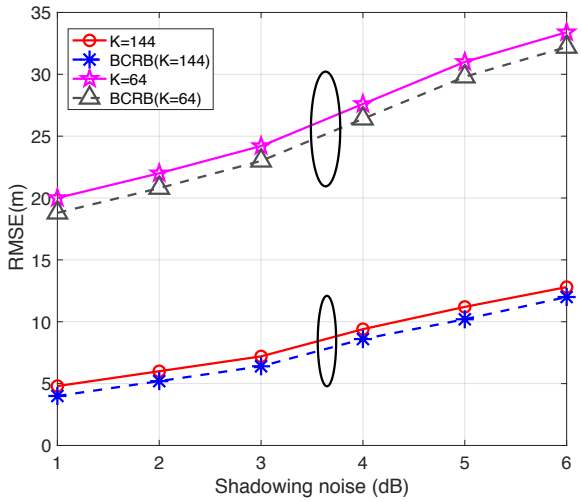


Figure 7. RMSE along with their BCRBs in the different number of antennas, for the different shadowing noise levels

We next evaluate the performance of our improved adaptive algorithm and compare with proposed *BeamMaP* in the different environments. According to our analysis in the previous section, the attenuation of the signal strength is mainly affected by the rainfall in the outdoor of 5G system. We set up three different scenarios and parameters according to the rain volume. We assume the LoS and NLoS signal strength average attenuation is δ_{rain} for 0.5 dB (in drizzle) and for 1.4 dB (in heavy rain) over 10 GHz link in the 5G wireless network [34] [35]. The path loss function in (4) becomes $\text{P}_{loss}[\text{dB}] + \delta_{rain}$ for LoS and NLoS channel in the simulation. The other parameters are the same following in the Table II and the number of antennas is chosen as $K = 12 \times 12$. We assume the hourly training data sets are relatively invariability within 24-48 hours for *BeamMaP* and set as the benchmark in the comparison.

Figure 8(a), 8(b) and 8(c) show that our adaptive algorithm works in the best performance with the increased number of training samples. For example, in the Figure 8(a), when the current weather is sunny or without rain drop, the Adaptive-BeamMaP presents the initiative training input data are chosen in the sunny measurement as status $S[0]$. At the same time, the testing data sets in RSS vectors are also chosen from the fingerprints in the sunny condition. The benchmark in BeamMaP(Drizzle) and BeamMaP(Rainy) represents the training input data sets are collected in the drizzle and rainy

environments while the testing fingerprints are chosen from the fingerprints in the sunny status. And we also set up BeamMaP(All) as baseline scheme by collecting and training the input fingerprints which covers all the weather conditions (including Sunny, Drizzle and Rainy) without classification. The selection probability of the three conditions in the input data sets is assumed the same (1/3), while we set up the training data sets in BeamMaP(All) scheme. Similarly, Figure 8(b) and 8(c) show the testing environments are in the Drizzle and Rainy status respectively, and the training fingerprints are chosen different under the different schemes.

Firstly, we observe the RMSE of all conditions decreases and gradually becomes steady with training samples increasing from 10 to 90 ($\times$ 1000) in Figure 8(a), 8(b) and 8(c). This is expected because we train the models with more and more fingerprints, the process will tend to project the coordinates of testing users onto the output reference location coordinates space in regression. Secondly, we observe that the BeamMaP(Drizzle), BeamMaP(Rainy) and BeamMaP(All) provide the higher values than the Adaptive-BeamMaP scheme. This is because the first two methods do not utilize the original testing RSS vectors, whereas the adaptive scheme utilizes the same environment RSS vectors for both training and testing. While the bias introduced by the different shadowing noise levels in the training, the testing data not belongs to the same weather conditions will degrade RMSE performance in the different levels. In addition, even though the BeamMaP(All) tends to close the curve of Adaptive-BeamMaP, the input RSS vectors will generate the overlap in the same coordinates within the different weather conditions and it will cause the increase of the variances for RMSE. Thirdly, BeamMaP(Drizzle) and BeamMaP(All) in RMSE performance are more closer to the optimize scheme, it depends on the less differences on the levels of shadowing noise combined in input fingerprints. BeamMaP(Rainy) being the worst case demonstrates the rainfall largely affects RSS receiving in the higher frequency wireless system and causes the differences between input and testing data sets. Similar to the condition in Figure 8(a), 8(b) and 8(c) also reflect the better performance in RMSE for Adaptive BeamMaP, compared with the other schemes. The gap in the curves between the different schemes shown in the Figure 8(b) is minimal, because the propagation loss generated in Drizzle is close to the Sunny status and also the Rainy status. In other words, the smaller bias of the training RSS vectors between Drizzle and other status will cause the close performances. In addition, even though BeamMaP(All) shows the relatively good performance in the comparison, it has to increase the time complexity in the training process because of the diversity sampling.

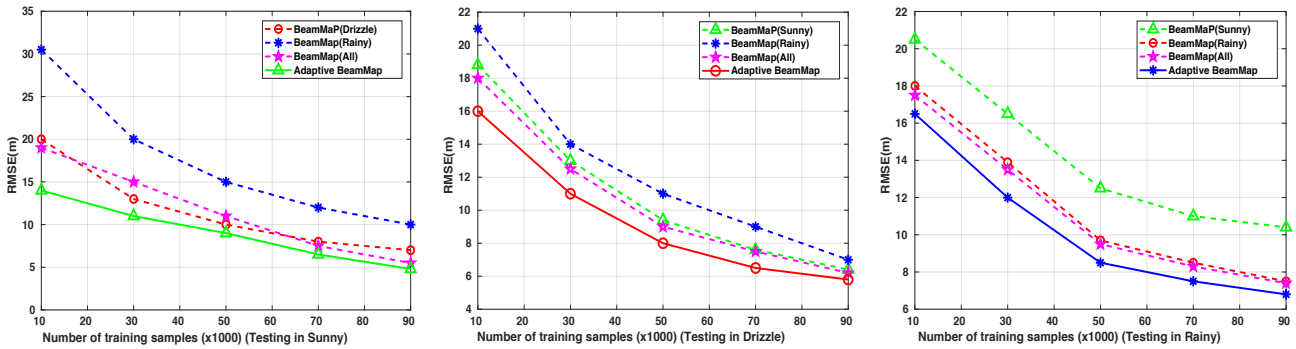
D. Comparison with Other Machine Learning Algorithms

We compare performance of the algorithms based on accuracy in RMSE and running time of machine learning between different machine learning approaches (BeamMaP, kNN and SVM) in the dynamic environments. kNN and SVM algorithms based on RSS fingerprints introduced in some indoor or outdoor localization techniques [18] [19] are acquired some improvements in the coordinates estimation. These fingerprinting-based approaches are all based on the matching of the online data to the existing database. The online data with RSSI values are gathered from each WIFI or beacon in the building or outdoor environment [18] [19], which can represent the features of a specific location. The existing database represents the testing data selected from the fingerprints. However, in order to keep the fairness of the experiments, we employ the same input training data sets (outdoor model) in these three models to study the advantages, disadvantage and effectiveness between them. In general, the localization with fingerprints can be viewed as a simple nonlinear equation, in which the values of each parameter are entered and the outputs are the coordinates of the target locations. We run the simulations simultaneously on the three same workstations (Ubuntu 16.04 LTS system on 3.6 GHZ Intel Core i7-4790 CPU with eight cores). The shadowing noises for LoS and NLoS are set up to change from 1 dB to 6 dB. The same training data sets are generated through CI-PI model. The details of models for kNN and SVM are designed below.

The k-nearest neighbor algorithm is a simple and effective classification and regression method in machine learning applications. [18] introduces kNN scheme to solve the localization problem. The proposed designed regression method is to find a predefined number of training samples closest in distance to the testing point and predict the label from the samples. In our comparison experiments, the input training samples are assumed the same. The basic procedure is to initialize k to the chosen number of neighbors (RSS vectors as fingerprints) in the beginning. For calculating the similarity between a training and testing fingerprint we use the Euclidean distance between RSS vectors, which is a well-established and extensively used procedure in kNN regression. Our objective is to minimize the Euclidean distance function between the training RSS vectors and testing vector $\sum_k(\mathbf{p}_k - \widehat{\mathbf{p}}_k)^2$. We sort the ordered collections of distances and indices from smallest to largest (in ascending order) by the distances. Then the first k entries from the sorted collection will be collected and calculate the RMSE of model. Depending on the training mobile users under outdoor instead of indoor environment [18], we need to choose a different k to optimize the ML process in the simulation. The indoor experiment chooses k=4 in the optimized prediction model [18], but in our outdoor, the RMSE can become much smaller when k is chosen a larger number in the simulation below. In the simulation, we start from $k = 1$ to observe changes of the RMSE metric. When k is chosen larger, the accuracy of localization becomes more precision until $k = 10$, and then behaves worse after 10. Our mission is to compare these algorithms in the best optimized model, so we only select $k = 1$, $k = 4$, and $k = 10$ shown in the results below.

A support vector machine (SVM) is able to analyze existing data and learn the relations between the input data and predicted outputs. In the model design, a non-linear kernel function is used to maximize the margin between classes by transforming the space into a higher dimension, where the problem can be solved in a linear way. There are three most popular kernel functions: polynomial, Gaussian radial basis function (RBF) and hyperbolic tangent. Since RBF is one of most popular and proven empirical effective kernel function in [19], it is adopted in our simulation model. Based on the same input training samples, the kernel function is showed below.

$$K(\mathbf{P}_i, \mathbf{P}_j) = \exp(-\lambda \|\mathbf{P}_i - \mathbf{P}_j\|^2) \tag{13}$$



(a) (b) (c)

Figure 8. Adaptive BeamMaP VS. BeamMaP: (a) Testing in Sunny. (b) Testing in Drizzle. (c) Testing in Rainy.

Where λ is a free parameter. The training of SVM is to minimize the structural risk. The most significant parameters required when working with the RBF kernel of the SVR model are C and γ. A list of values to choose from should be given to each hyper parameter of the SVM model. We need to change these values and experiment more to see which value ranges give better performance. Grid Search method is employed in the process of performing hyper parameter tuning, in order to determine the optimal values for the given model [20]. Also a K^*-fold cross-validation process (K^* =10) is performed in order to determine the hyper parameter value set which provides the best accuracy levels. Intuitively, the γ parameter defines how far the influence of a single training example reaches, with low values meaning 'far' and high values meaning 'close'. The C parameter trades off correct classification of training examples against maximization of the decision function's margin. Established on the features of RSSI vectors, we set up γ in [10^{-4}, 10^{-3}, 0.01, 0.1, 0.2, 0.5, 0.6, 0.9], C in [0.001, 0.01, 0.1, 1, 2, 4, 6, 8, 10] and do the estimation of RMSE to optimize the model. We plot one RMSE accuracy heatmap as an example in Figure 9 to observe the optimization process when the shadowing noise is chosen 2 dB. The blue area in the figure represents the accuracy of localization is higher when RMSE is smaller. It is obvious the best parameters are C=1.0, γ= 0.1 with a smallest RMSE 15.5 m. For larger values of C, a smaller margin will be accepted if the decision function in Equation 7 of ML model is better at classifying all training points correctly. A lower C will encourage a larger margin, therefore a simpler decision function, at the cost of training accuracy. Also the behavior of the model is very sensitive to the γ parameter. If γ is too large, i.e.,$\gamma = 0.6$, RMSE is larger, the radius of the area of influence of the support vectors only includes the support vector itself and no amount of regularization with C will be able to prevent overfitting. Finally we can also observe that for some intermediate values of $\gamma = 0.1$ and $C = 1$, we get best perform model, while it is not necessary to regularize by enforcing a larger margin. When the shadowing noise becomes larger in the simulation, we continue to optimize the model and acquire the smallest RMSE as the best result.

In the comparison experiments, we employ K^*-fold cross-validation method in the experiments, K^* is set as 10 in all the three models. Since the number of fingerprints is $S = 90000$, the testing time in running each model is calculated

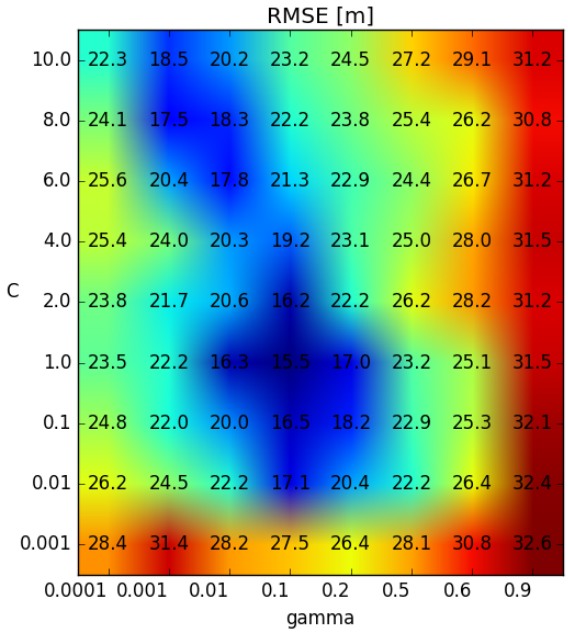


Figure 9. Heatmap of the RMSE [m] in SVM scheme as a function of γ and C when shadowing noise is 2 dB.

after estimating the coordinates of 9000 testing samples each time. The training time for each model is calculated until the model is optimized. For example, the training time of kNN is calculated after receiving the optimization of k value and RMSE. For SVM, we also need to tune a better γ and C in the process. All results are shown in Table III. In general, with the increase of shadowing noise, the RMSE (in meters) for all approaches gradually becomes larger. Compared with kNN and SVM, RMSE for the proposed *BeamMaP* is obviously smaller and has better accuracy. SVM takes too much time (about 16.2 hours) to train a model, which renders it a poor training candidate. Although the training time for kNN is much less than *BeamMaP* and SVM, the testing time for our proposed is averaged as 0.35 s which is less than the others.

kNN being as a unsupervised method, is served as positioning the target MU through collecting and analyzing the closest k reference neighbors. The time complexity known as

TABLE III. Comparison between different machine learning models.

Shadowing Noise	RMSE[m]				
	BeamMaP	kNN (k=1)	kNN (k=4)	kNN (k=10)	SVM
1 dB	3.5	12.1	10.5	8.5	10.2
2 dB	8.4	14.5	12.1	10.2	15.5
3 dB	15.6	16.8	16.2	20.2	20.4
4 dB	20.3	25.2	24.8	23.5	24.7
5 dB	24.3	29.3	28.2	27.5	28.8
6 dB	29.3	34.2	32.3	30.4	32.5
Phase	**Running time**				
Training	7.5 hours	58 mins	1.2 hours	2.1 hours	16.2 hours
Testing	0.35 s	0.45s	0.45 s	0.45 s	0.80 s

$\mathcal{O}(KS + kS)$ is depended on the S cardinality of the training data set and the K (the number of antennas) dimension of each sample. In particular, the optimized kNN (k=10) regression algorithm on average reduces the prediction error by roughly 20% and 40% compared to the benchmarks for k=4 and k=1. k=4 in [18] plays a good performance in the indoor environment, because the compromise is that the distinct boundaries within the feature space are blurred. However, a large k value is more precise as it reduces the overall noise in the outdoor environment.

Despite SVM is mostly used in the linear condition, our nonlinear problem needs to be transferred into the quadratic problem directly, which involves inverting the kernel matrix. It has complexity on the order of $\mathcal{O}(S^3)$ same with our proposed model. But in order to tune the parameters in optimizing the model, it will spend much more time in the training process. In addition, these two models kNN and SVM employed in [18] [19] only choose LoS signals in the RSS vectors of training data sets, the NLoS elements have to be removed and become 0 in the experiments. It will increase the removal algorithm (dispatch NLoS signals) before the training process and cause the large increase of running time. In general, the shortest testing time spent and smallest RMSE in the simulation will prove that our proposed model is steadier and better optimized in the much noisy or highly cluttered multipath scenarios, also the gap of the training time between them can be shortened if the future advanced hardware employed.

E. Comparison between the Different Beamform Patterns in the Testing Phase

In this section, we aim to compare the different beamforming employed in the localization. In the analysis of characteristics of beamforming techniques, we realize to model the different distribution of mobile users in the testing phase to meet the different beamform patterns, which could decide the direction of antennas transmission and bandwidth distribution. Beamforming schemes are generally classified as either switched-beam systems or adaptive array systems. A switched-beam system depends on a fixed beamforming network that yields established predefined beams [25]. In the adaptive beamforming, perfect adaptive beams attempt to reduce the interference between users and achieve considerably improved offered power resources [25]. In our model, it can be expressed that switched beam pattern represents the selection of actual mobile users follows the uniform distribution and the testing mobile users are selected in the same distribution from the fingerprints. Also adaptive beam pattern represents the selection of actual mobile users follows the Rice distribution (power consumption in the antennas is fixed but need to distribute non-uniform in any direction) and the testing mobile users are selected in the Rice distribution. We choose the testing mobile users from the input fingerprints due to the cross-validation process.

In order to compare adaptive beamforming with switched beamforming, the number of antennas is set up as 12×12 to maximize the sampling ratio in the fingerprints collection. The other parameters set up is the same with above experiments. Similarly, we assume that the maximum distance between central of antenna array and test mobile users set up as $\nu = 150$ meters, and the variance distance between adjacent testing mobile users set up as $\sigma = 1$ meter. The Rice distribution is selected as $R \sim \text{Rice}(150, 1)$ through experiments to cover a single cell network and variance of spacing in 1 m. During the testing phase, we model the switched beamforming as a uniform distribution with the same mean and variance as the Rice distribution in adaptive beamforming. It is shown in the Figure 10, the estimation errors of localization decrease with the number of training becoming more. We also observe that adaptive beamforming or Rice distribution in the *BeamMaP* system plays a better role, it can reach the 72.8%, 85.3% and 92.4% of RMSE of uniform distribution with the same training index (10, 20, 30 $\times$ 1000). However, with the increase of training fingerprints, the gap between them will become smaller easily, which proves the adaption of our localization system.

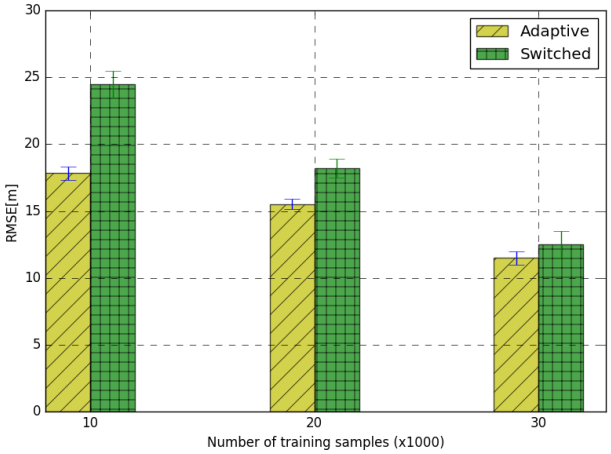


Figure 10. RMSE vs. number of samples for different beamforming patterns

More efficiency for adaptive beamforming is achieved by randomly selecting the testing users similar to Monte-Carlo sampling. The reason is that more testing users are gathered together in one direction for the adaptive pattern, but testing users in uniform distribution (switched pattern) are separately localized, which will accumulate the estimation errors and lead to the increase of RMSE. Adaptive beamforming system in the base station being as a better candidate in the future wireless network can be better assisted by our localization method.

VI. Conclusion and Future Works

In this paper, we present an improved adaptive *BeamMaP* positioning method in massive MIMO systems. It consists of an adaptive algorithm for the selection of input fingerprints,

a supervised machine learning approach and an online adaptive beamforming testing process, to estimate the position of mobile users. *BeamMaP* can estimate the location of the MUs within 5 meters deviation in milliseconds, which is much better than some conventional methods like GPS. Numerical results show the accuracy of positioning, as determined by the size of training samples, the dimension of antennas and the spacing of antennas. The achieved RMSE performances are proved to close to Bayesian Cramer-Rao bounds. In addition, our improved adaptive *BeamMaP* exhibits the better performance than original *BeamMaP* in different weather during the hourly time, while achieving comparable performance as other machine learning schemes such as kNN and SVM in the dynamic environments. Moreover, we conclude that our *BeamMaP* localization method can serve in the different beamforming systems and performs better in the adaptive beamforming wireless system. However, the RSS as input data seems more sensitive established on the limited fingerprints collected, some steadier features such as channel states information (CSI) and Time-of-Arrival (ToA), can become the next alternatives in the future works. In addition, some deep learning or hybrid machine learning methods can be explored and make more improvements.

REFERENCES

[1] C. Liu and H. J. Helgert, "BeamMaP: Beamforming-based Machine Learning for Positioning in Massive MIMO Systems," In Proceedings The Eleventh International Conference on Evolving Internet, INTERNET 2019, June 30 - July 4, 2019, Rome, Italy, ISSN: 2308-443X, ISBN: 978-1-61208-721-4, pp. 16-23. [Online]. Available: https://www.thinkmind.org/.

[2] S. K. Routray, P. Mishra, S. Sarkar, A. Javali, and S. Ramnath, "Communication bandwidth for emerging networks: Trends and prospects," arXiv preprint arXiv:1903.04811, 2019.

[3] M. S. Leeson, "Introductory chapter: The future of mobile communications," In The Fifth Generation (5G) of Wireless Communication. IntechOpen, 2019.

[4] C. Liu, M. Xu, and S. Subramaniam, "A reconfigurable high-performance optical data center architecture," in 2016 IEEE Global Communications Conference (GLOBECOM), pages 1–6. IEEE, 2016.

[5] M. Xu, C. Liu, and S. Subramaniam, "PODCA: A passive optical data cen- ter network architecture," Journal of Optical Communications and Networking, 10(4):409420, 2018.

[6] C. Tang, S. Xia, C. Liu, X. Wei, Y. Bao, and W. Chen, "Fog-enabled smart campus: Architecture and challenges," In International Conference on Security and Privacy in New Computing Environments, pages 605–614. Springer, 2019.

[7] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," IEEE Communications Magazine, vol. 52, no. 2, 2014, pp. 74–80.

[8] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," IEEE journal of selected topics in signal processing, vol. 8, no. 5, 2014, pp. 742–758.

[9] S. Kumar, R. M. Hegde, and N. Trigoni, "Gaussian process regression for fingerprinting based localization," Ad Hoc Networks, vol. 51, 2016, pp. 1–10.

[10] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, "Direct localization for massive MIMO," IEEE Transactions on Signal Processing, vol. 65, no. 10, 2017, pp. 2475–2487.

[11] J. Vieira, E. Leitinger, M. Sarajlic, X. Li, and F. Tufvesson, "Deep convolutional neural networks for massive MIMO fingerprint-based positioning," arXiv preprint arXiv:1708.06235, 2017.

[12] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, "Global positioning system: theory and practice," Springer Science & Business Media, 2012.

[13] M. Brunato and R. Battiti, "Statistical learning theory for location fingerprinting in wireless LANs," Computer Networks, vol. 47, no. 6, 2005, pp. 825–845.

[14] M. A. Alsheikh, S. Lin, D. Niyato, and H.-P. Tan, "Machine learning in wireless sensor networks: Algorithms, strategies, and applications," IEEE Communications Surveys & Tutorials, vol. 16, no. 4, 2014, pp. 1996–2018.

[15] A. M. Hossain, H. N. Van, Y. Jin, and W.-S. Soh, "Indoor localization using multiple wireless technologies," in 2007 IEEE International Conference on Mobile Adhoc and Sensor Systems. IEEE, 2007, pp. 1–8.

[16] Y. Xie, Y. Wang, A. Nallanathan, and L. Wang, "An improved k-nearest-neighbor indoor localization method based on spearman distance," IEEE signal processing letters, vol. 23, no. 3, 2016, pp. 351–355.

[17] D. A. Tran and T. Nguyen, "Localization in wireless sensor networks based on support vector machines," IEEE Transactions on Parallel and Distributed Systems, vol. 19, no. 7, 2008, pp. 981–994.

[18] F. Lemic et al., "Regression-based estimation of individual errors in fingerprinting localization," IEEE Access, 7:3365233664, 2019.

[19] W. Kim, J. Park, J. Yoo, H. J. Kim, and C. G. Park, "Target localization using ensemble support vector regression in wireless sensor networks," IEEE transactions on cybernetics, vol. 43, no. 4, 2012, pp. 1189–1198.

[20] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, 12:28252830, 2011.

[21] H. Rizk, A. Shokry, and M. Youssef, "Effectiveness of data augmentation in cellular-based localization using deep learning," arXiv preprint arXiv:1906.08171, 2019.

[22] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," IEEE Transactions on Vehicular Technology, vol. 66, no. 1, 2016, pp. 763–776.

[23] H. Wymeersch et al., "5G mm wave downlink vehicular positioning," in 2018 IEEE Global Communications Conference (GLOBECOM), 2018, pp. 206–212.

[24] A. Gorokhov, A. F. Naguib, A. Sutivong, D. A. Gore, and J. Tingfang, "Pilot signal transmission for an orthogonal frequency division wireless communication system," Oct. 4 2016, US Patent 9,461,859.

[25] E. Ali, M. Ismail, R. Nordin, and N. F. Abdulah, "Beamforming techniques for massive MIMO systems in 5G: overview, classification, and trends for future research," Frontiers of Information Technology & Electronic Engineering, vol. 18, no. 6, 2017, pp. 753–772.

[26] S. Jung, C. Lee, and D. Han, "Wi-Fi fingerprint-based approaches following log-distance path loss model for indoor positioning," in Intelligent Radio for Future Personal Terminals (IMWS-IRFPT), 2011 IEEE MTT-S International Microwave Workshop Series on. IEEE, 2011, pp. 1–2.

[27] M. Hata, "Empirical formula for propagation loss in land mobile radio services," IEEE transactions on Vehicular Technology, vol. 29, no. 3, 1980, pp. 317–325.

[28] K. Haneda et al., "5G 3GPP-like channel models for outdoor urban microcellular and macrocellular environments," in Vehicular Technology Conference (VTC Spring), 2016 IEEE 83rd. IEEE, 2016, pp. 1–7.

[29] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networkswith a focus on propagation models," IEEE Transactions on Antennas and Propagation, vol. 65, no. 12, 2017, pp. 6213–6230.

[30] P. Jain and P. Kar, "Non-convex optimization for machine learning," Foundations and Trends® in Machine Learning, vol. 10, no. 3-4, 2017, pp. 142–336.

[31] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning," in International conference on machine learning, 2013, pp. 1139–1147.

[32] Recommendation ITU-R P. 676–10, attenuation by atmospheric gases, International Telecommunications Union, 2013. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.676-11-201609-I!!PDF-E.pdf [retrieved: June 2020]

[33] J. Luomala and I. Hakala, "Effects of temperature and humidity on radio signal strength in outdoor wireless sensor networks," in 2015 Federated Conference on Computer Science and Information Systems (FedCSIS). IEEE, 2015, pp. 1247–1255.

[34] Z. Qingling and J. Li, "Rain attenuation in millimeter wave ranges," in 2006 7th International Symposium on Antennas, Propagation & EM Theory. IEEE, 2006, pp. 1–4.

[35] I. Shayea, T. Abd. Rahman, M. Hadri Azmi, and A. Arsad, "Rain attenuation of millimetre wave above 10 GHz for terrestrial links in tropical regions," Transactions on Emerging Telecommunications Technologies, vol. 29, no. 8, 2018, p. e3450.

[36] P. Joshi, D. Colombi, B. Thors, L.-E. Larsson, and C. Törnevik, "Output power levels of 4g user equipment and implications on realistic RF EMF exposure assessments," IEEE Access, vol. 5, 2017, pp. 4545–4550.

[37] "Cisco 1250 dipole antenna spacing," 2019, [Online]. Available: https://community.cisco.com/t5/other-wireless-mobility-subjects/specs-on-distance-between-antennas/td-p/1030478 [retrieved: June. 2020]

[38] "TE Connectivity," 2020, [Online]. Available: https://www.electronicspecifier.com/products/communications/te-connectivity-antenna-separation-in-mimo [retrieved: June. 2020]

[39] Z. Ben-Haim and Y. C. Eldar, "A lower bound on the bayesian mse based on the optimal bias function," IEEE Transactions on Information Theory, vol. 55, no. 11, 2009, pp. 5179–5196.

[40] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, "LANDMARC: indoor location sensing using active RFID," Wireless networks, vol. 10, no. 6, 2004, pp.701–710.

[41] M. Tamosiunaite, S. Tamosiunas, M. Zilinskas, and G. Valusis. "Atmospheric attenuation of the terahertz wireless networks," In Broadband Communications Networks-Recent Advances and Lessons from Practice. InTech, 2017.

"Objection, Your Honor!": False Positive Detection in Sender Domain Authentication by Utilizing the DMARC Reports

Kanako Konno
Amazon Web Services Japan K.K.
Tokyo, 141-0021, Japan
Email: kankon@amazon.co.jp

Naoya Kitagawa
Research and Development Center for Academic Networks,
National Institute of Informatics
Tokyo, 101-8430, Japan
Email: kitagawa@nii.ac.jp

Nariyoshi Yamai
Division of Advanced Information Technology and Computer Science Institute of Engineering,
Tokyo University of Agriculture and Technology
Tokyo, 184-8588, Japan
Email: nyamai@cc.tuat.ac.jp

***Abstract*—Information leakage and phishing scams caused by spoofed e-mails have become serious problems, particularly in the fields of business and e-commerce. Sender domain authentications, such as Sender Policy Framework (SPF), DomainKeys Identified Mail (DKIM), and Domain-based Message Authentication, Reporting, and Conformance (DMARC), are effective countermeasures against spoofed e-mails. In particular, DMARC is one of the most effective methods of sender domain authentication. However, sender domain authentication methods erroneously classify legitimate e-mails, such as forwarded e-mails, as malicious e-mails. Because sender domain authentication is usually processed prior to content filtering, the fact that sender domain authentications generate a large number of false positives is a serious problem. In this paper, we propose a method to detect false positive deliveries in sender domain authentications based on the legitimacy of the senders' IP addresses by adapting X-means clustering to the reports generated by the reporting function of DMARC. Our approach consists of three phases: DMARC report summarization, X-means clustering, and legitimate sender detection. Applied to actual DMARC reports, we found that our method detected 214,153 e-mails on average sent from 347 legitimate senders' IP addresses on average as legitimate e-mails per day. We evaluate our results focusing on the legitimate deliveries sent from the detected legitimate senders and the detected false positives generated by existing sender domain authentications. The evaluation results confirmed that our method can detect large numbers of legitimate e-mails, including the false positive e-mails, such as forwarded e-mails, which cannot be correctly identified using existing sender domain authentication technologies.**

***Keywords*–Spoofed e-mail; SPF; DKIM; DMARC; Clustering.**

I. INTRODUCTION

This paper is an extended version of our previous study presented at the Eleventh International Conference on Evolving Internet [1]. In our previous study, we proposed a mechanism to detect e-mail forwarding servers, which are a type of legitimate e-mail sending server, via clustering. In this paper, as an enhancement to our previous study, we propose a method to detect many types of legitimate e-mail sending servers, in addition to forwarding servers. By utilizing our method proposed in this paper, e-mail system administrators can detect a variety of legitimate deliveries that have been false positives with conventional sender domain authentications.

E-mail is one of the most utilized communication services worldwide. However, especially in business, e-mail has a serious problem due to the rapid increase in information leakage and phishing scams enabled by spoofing e-mail. According to the statistics report of the Federal Bureau of Investigation, the total financial damage due to spoofed e-mails was 26.2 billion US dollars from June 2016 to July 2019 [2]. Spoofed e-mails are used by spammers to steal sensitive information or send malicious programs, such as computer viruses.

Sender domain authentication has been proposed as an effective countermeasure to spoofed e-mails. Sender Policy Framework (SPF) [3] and DomainKeys Identified Mail (DKIM) [4] are methods that are widely used. SPF is a method, in which the receiver confirms whether the e-mail sender's IP address is legitimate by checking the original sender's SPF record, which is a list of IP addresses that the sender may use to send e-mails. However, SPF cannot verify forwarded e-mails correctly because the sender's IP address is changed to the forwarder's IP address, which is not included in the sender's SPF record when the e-mails are forwarded. In DKIM, the receivers verify the digital signatures generated from the header and body of the e-mail and confirm whether the e-mail has been rewritten by spammers. DKIM allows a third-party's domain to sign e-mails; therefore, DKIM has the problem that spoofed e-mails signed by a spammer's own malicious domain will incorrectly pass its verification.

Domain-based Message Authentication, Reporting, and Conformance (DMARC) [5] is one of the most effective sender domain authentication frameworks and includes reporting and policy controlling mechanisms. DMARC utilizes both the SPF and DKIM authentication mechanisms to verify e-mails. DMARC has a reporting function that enables an e-mail sender to receive a "DMARC aggregate report" (hereafter, called the DMARC report). This report provides information, such as the header of the e-mail and the authentication results. In general, DMARC reports are used to confirm the effectiveness of sender domain authentications by e-mail senders. However, we can also observe the transmission behavior for each e-mail sending server by analyzing the information in the DMARC reports.

Anti-spam methods are generally operated in three phases: Transmission Control Protocol (TCP) and Simple Mail Transfer Protocol (SMTP) session monitoring and blacklist, sender domain authentication, and content filtering. In such an anti-

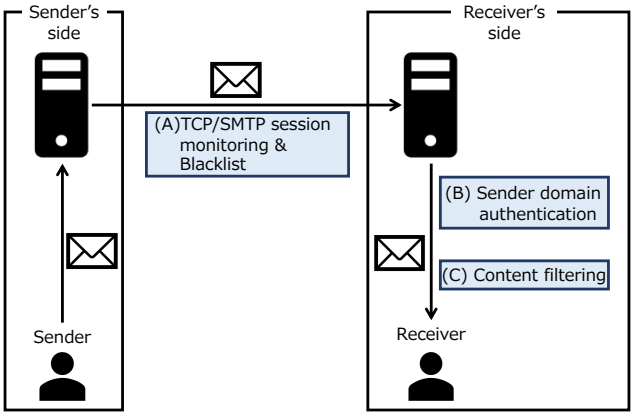


Figure 1. General flow of anti-spam measures.

spam operation, e-mail servers process sender domain authentications before implementing the content filtering method. Therefore, it is essential to reduce the number of false positives in the sender domain authentication. Conversely, the increase in false negatives in the sender domain authentication caused by reducing the number of false positives is not a critical problem.

In this paper, we propose a method to detect legitimate e-mail senders in order to reduce false positives in the sender domain authentication via X-means clustering [6] using the massive amounts of available DMARC report data. Our approach consists of three phases: DMARC report summarization, X-means clustering, and legitimate sender detection.

To test our approach, we apply it to actual DMARC report data. To evaluate our results, we investigate the details of the detected legitimate e-mails sent from the legitimate senders and the false positive deliveries in the sender domain authentications. Our evaluation results indicate that our method detects false positive e-mails, such as forwarded e-mails, which cannot be correctly determined by existing sender domain authentication technologies.

This paper organized as follows. In Section II, we explain several existing anti-spam methods. In Section III, we describe the design of our approach. Then, we describe the dataset that we use in our experiment in Section IV. Section V shows the results generated when applying our method to the dataset. In Section VI, we evaluate our results focusing on the number and ratio of false positive deliveries of the different sender domain authentication technologies. Finally, we present our concluding remarks in Section VII.

II. Related Work

Anti-spam measures are generally processed in phases. In this section, we show several approaches for each of the phases shown in Figure 1.

A. TCP/SMTP session monitoring and blacklists

Greylisting [7] is a method that checks the retry function for establishing an SMTP session. In general, legitimate e-mail senders try to resend an e-mail after a period of time when an e-mail is temporally rejected. Conversely, spammers who use massive e-mail sending tools do not try to resend e-mails. This technique, which takes advantage of such differences in sending behavior, is effective as a countermeasure against spammers sending large amounts of e-mail.

SMTP tarpitting [8] detects spam e-mails by delaying a response to the sender's server. Spammers generally try to send as many spam e-mails as possible in a short period of time. Therefore, they tend to ignore a response from a receiver's server or abandon sending the spam e-mails altogether. Even though SMTP tarpitting can eliminate such transmissions with priority on delivery efficiency, it also delays transmissions by legitimate senders. Therefore, legitimate e-mails may not be delivered correctly.

Kitagawa et al.'s method [9] inspects the SYN packet retry function for establishing a TCP session between a sending host and a receiving host. This method is effective for spam delivery that gives priority to e-mail delivery efficiencies such as greylisting and SMTP tarpitting.

Even though these methods are highly effective against conventional spam transmission, it is expected that the reduction effect for the cleverly spoofed e-mails that have become a social problem in recent years is not sufficient.

A blacklist mechanism checks whether the sender's IP address and/or domain name is registered in an attacker IP address and/or domain name list, i.e., a blacklist. Blacklists provided by MxToolBox [10], Spamcop Blocking List [11], Barracuda Reputation Blocklist [12], and Spamhaus blocklist are popular. The Spamhaus blocklist, provided by The Spamhaus Project, an international non-profit organization, is the most famous and widely used IP blacklist. The Spamhaus blocklist is managed by dedicated teams in 10 countries and maintains its by tracking cyber threats such as spam, phishing, and malware, worldwide. However, blacklists have a disadvantage in that it takes time for both the removal of legitimate IP addresses from the list and the registration of malicious IP addresses to the list to be reflected in the service. For example, when an Internet service provider (ISP) sends an IP address removal request to the Spamhaus blocklist, the Spamhaus blocklist evaluates the legitimacy of the IP address, and the removal request is processed within 24 hours according to the Spamhaus blocklist policy.

B. Sender domain authentication

SPF, DKIM, and DMARC are popular methods of sender domain authentication. We describe these three methods in Sections II-B1, II-B2, and II-B3, respectively.

1) SPF: SPF is a method of checking the SPF record of the sender domain to make sure that the IP address of the sender's SMTP server is legitimate. Senders indicate a list of IP addresses of SMTP servers that may send e-mails from their domain as the SPF record on the Domain Name System (DNS) content server in advance. To verify an e-mail with SPF, the recipient queries the sender's Envelope-From domain's DNS content server for the SPF record to check if the SPF record contains the IP address of the sender's SMTP server. However, the verification of forwarded e-mails with this method will not be successful even for legitimate mail. Figure 2 shows an example in which SPF authentication fails for a forwarded e-mail.

As shown in this figure, when a message is forwarded, the original IP address of the SMTP server is changed to the relay

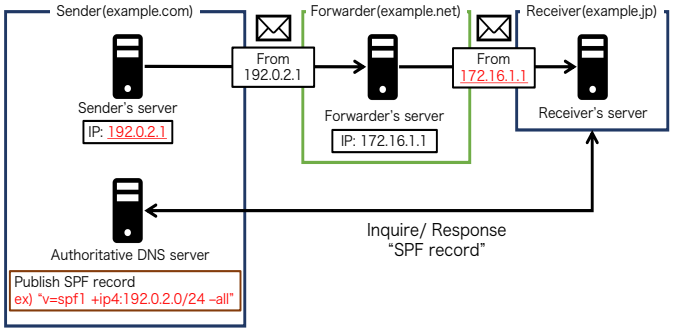


Figure 2. An example of an SPF authentication failure for a forwarded e-mail.

```
Return-Path: <sender@example.com>
(snip)
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed;
    d=signer.example; s=20191225;
    bh=Za3JDErJrJPrpL+bXkLoOcl2gQi1jwTNEIAraa8oTDU=;
    b=yRKl7uiICDa7nBw2I0yQECGgnWWwNX+H42tMm2T4/MI/S
    6fgRL/XoOyYyNb14BtR5H7I0O8mXQKUB78cyFJj75Wy0w2RBb
    SnHTbOYM3KmEnzqu4lrFLlovRoI=
(snip)
From: <sender@example.com>
To: Receiver@example.net
```

Figure 3. An example of an e-mail header.

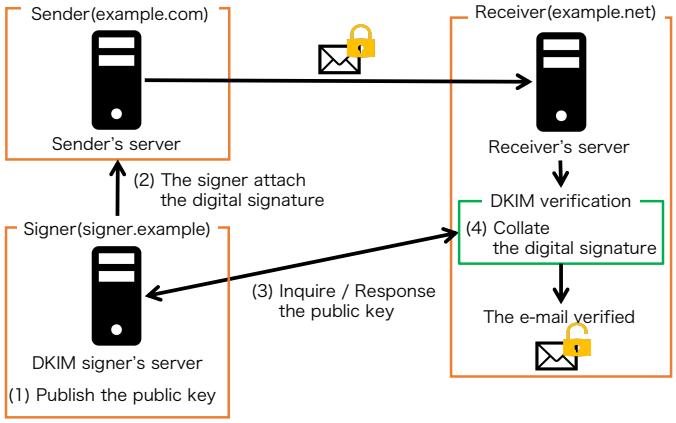


Figure 4. An example of the DKIM verification flow.

server's IP address which is not include the SPF record. As a result, there are many cases in which a valid mail fails this verification.

2) DKIM: DKIM is an authentication method that uses the digital signature generated from the body and header of an e-mail. Figure 3 shows an example of an e-mail header and Figure 4 shows an example of the DKIM verification flow.

First, to use the DKIM mechanism, the sender domain ("example.com" in Figure 4) prepares a private key and public key pair in advance and publishes the public key on their authoritative DNS server for DKIM verification ("signer.example" in Figure 4). Then, the sender domain ("example.com") generates the DKIM signature from the body and header of the e-mail

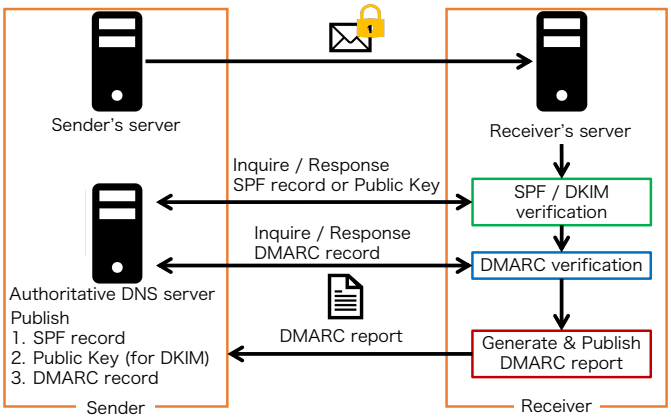


Figure 5. Flow of the DMARC verification.

using the private key and attaches it to the e-mail header as the DKIM signature, as shown by the "b=" tag in Figure 3.

Next, the receiver ("example.net" in Figure 4) requests the public key from the sender specified domain authoritative DNS server, as is shown in the "d=" tag of the DKIM signature ("signer.example" in Figure 3 and Figure 4). Then, the receiver obtains the hash value from the digital signature using the public key and compares it to the value of the "bh=" tag of the DKIM signature. If these values are the same, the e-mail passes the DKIM verification. With this mechanism, DKIM can even correctly verify forwarded e-mail, unlike SPF.

As shown in Figure 4, the DKIM signature domain does not need to match with the name of the sender's domain. Our observations confirmed that approximately 75% of DKIM-compatible domains use a third-party signature. However, the receiver cannot distinguish whether the third-party signer is legitimate. As a result, spammers can send spoofed e-mails with a DKIM signature using their own malicious domain that will pass the verification.

Additionally, in DKIM authentication, an administrator must change the key periodically, but the key may be expired or the key information may be misdescribed. In such cases, the validation will be failed even if the e-mail delivery is legitimate.

3) DMARC: DMARC is a reporting and policy controlling framework using both the SPF and DKIM mechanisms to authenticate e-mails. Although DMARC is a relatively new technology, the adoption rate of DMARC has been increasing in recent years. One of the reasons for this is that in addition to the UK and Australian governments, the US government has also required government agencies to support DMARC [13] [14] [15]. In addition, many mail service providers (MSP), ISP, financial institutions around the world have also adopted DMARC.

Figure 5 shows the flow of the DMARC verification. To use DMARC, the sender domain administrator must publish the SPF record for SPF verification and the public key for DKIM verification on an authoritative DNS server in advance to correspond to at least one of the two authentication mechanisms. Moreover, the sender domain needs to publish the DMARC record on their DNS server. For example, when the sender domain is "example.com," a DMARC record is published as a

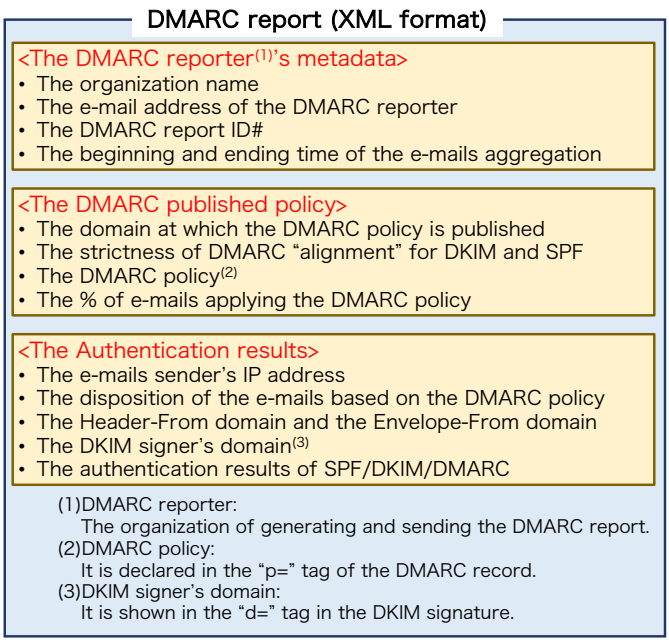


Figure 6. Example of a DMARC report.

TXT record "_dmarc.example.com" under the following rules:

v=DMARC1; p=reject; rua=mailto:rua@example.net.

In the policy controlling function, DMARC provides a mechanism for the administrator of the sender domain to declare the policy for how the receiver handles an e-mail that fails sender domain authentication in the "p=" tag of the DMARC record. The value of the "p=" tag has three variations: "none" (nothing even in the case of authentication failure), "quarantine" (quarantine the authentication failure e-mail), and "reject" (reject the authentication failure e-mail).

In the reporting function, an e-mail receiver sends the DMARC report to the e-mail address of the administrator of the sender domain shown in the "rua=" tag of the DMARC record. The DMARC report provides information, such as e-mail domains, authentication results, and the effectiveness of the DMARC policy. Examples of information included in DMARC reports are shown in Figure 6. With this, the administrator of the sender domain can determine the performance of the DMARC authentication and they can take measures to prevent spoofed e-mails from abusing their domain.

According to the concept of "alignment," DMARC verification fails when the domains for SPF and DKIM verification are different from the sender's Header-From domain. The sender's Header-From domain need not be the same as the Envelope-From domain or the DKIM signature domain. However, spammers can easily imitate the Header-From domain. As a countermeasure, using alignment, the receiver can check whether the Header-From domain is correct. The sender domain can choose from two mode of alignment strictness, "strict" and "relaxed," using the DMARC record.

When the administrator of the sender domain uses the "strict" mode, DMARC verification passes only when the Header-From address and the domain for SPF or DKIM verification match completely. Conversely, when the alignment mode is "relaxed," DMARC verification will succeed if subdomains of the Header-From address and subdomains of the domain for SPF or DKIM verification match.

DMARC is one of the most effective countermeasures to spoofed e-mail. However, DMARC cannot solve the issue that SPF cannot properly verify forwarded e-mails. SPF cannot properly authenticate forwarded e-mails because the sender's IP address changes to the forwarder's IP address when the e-mails are forwarded. Moreover, because DKIM allows third-party signatures, which are commonly used worldwide, as described in Section II-B2, e-mails signed by a third-party signer will fail the DMARC verification due to alignment.

Therefore, there are cases in which legitimate forwarded e-mails will fail the DMARC authentication, e.g., when e-mails use a third-party signature or the e-mail domains are not compatible with DKIM.

C. Content filtering

A large number of content filtering methods have been proposed over the years. Content filtering is an effective and widely used anti-spam method. This method adapts classifiers to the content or the attached files of the e-mail. The Bayesian filter [16] [17] [18] is a well-known content filtering method using the Bayes theorem to classify the e-mail content. In addition, natural language processing [19], support vector machines [20] [21], and machine learning [22] [23] are widely used as classifiers in content filtering methods.

In actual operation, as shown in Figure 1, content filtering has a high calculation cost and is therefore used after reducing the number of e-mails to be inspected by other anti-spam methods. SpamAssassin [24] [25], for example, scores e-mails based on keywords, the public database, and a Bayesian filter to detect spam e-mails. This method uses several anti-spam methods, such as blacklist [26] [27] and sender domain authentication methods, when the e-mails are received, prior to applying the Bayesian filter.

III. Design of our method

As described in Section II, e-mail servers are generally operated using a combination of multiple anti-spam measures. In addition, sender domain authentication is processed prior to e-mail content filtering. Therefore, it is important to reduce false positives in the sender domain authentication to achieve reliable e-mail server operation. However, as described in Section II-B, sender domain authentication may mistakenly determine legitimate e-mails as spoofed e-mails in the case of forwarded e-mail, misdescription of DKIM key information, and DKIM third-party signatures, etc.

To overcome this issue, we propose a method to detect false positives generated by the existing sender domain authentications by analyzing large-scale DMARC report data using an X-means clustering analysis.

As shown in Figure 7, our method consists of the following three phases: (A) DMARC reports summarization, (B) Summarized DMARC report clustering, and (C) Legitimate senders detection.

A. The DMARC reports summarization

First, we describe the DMARC report aggregation ((A) in Figure 7). As described in Section II-B3, the DMARC reports

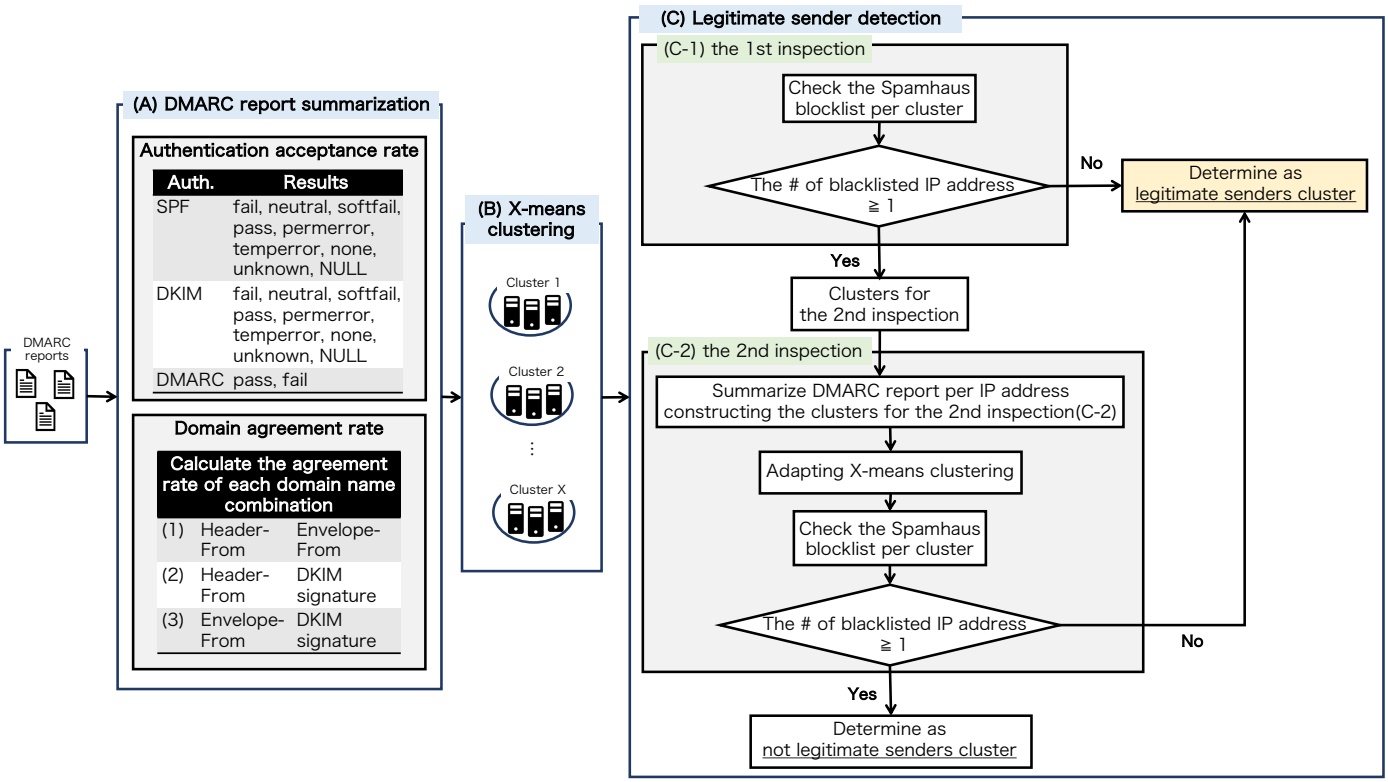


Figure 7. Overview of our method.

are provided in the XML format; therefore, it is necessary to convert the original DMARC report data into numeric data for the clustering process. Additionally, our method should summarize the DMARC reports by the sender's IP address to identify legitimate senders.

We summarize DMARC report data to allow the adaptation of a clustering analysis focusing on the results of the sender domain authentication and the e-mail domain names. As a summarization of the sender domain authentication results, we calculate the acceptance rates of SPF, DKIM, and DMARC for each IP address. Each of these three authentication methods has several authentication results, as shown in Figure 7. Our method calculates the percentage of e-mails for each authentication result per IP address.

Next, to summarize the e-mail domain names, we calculate the agreement rate for the three domain name combinations. The DMARC mechanism compares the Header-From to the Envelope-From domain ((1) in Figure 7) and the DKIM signature domain ((2) in Figure 7) for the DMARC alignment inspection. Conversely, the Envelope-From domain is not compared to the DKIM signature domain ((3) in Figure 7) in the sender domain authentication verification process. However, because we consider the combination (3), which is not for sender domain authentication, as having a relationship, it can be used to improve the accuracy of our approach.

B. Clustering the summarized DMARC reports

Second, we cluster of the aggregated DMARC reports ((B) in Figure 7). Our method adapts a clustering algorithm to the summarized DMARC report data. This clustering phase is used to classify the sender's IP addresses exhibiting similar e-mail transmission behavior trends, including those with respect to the authentication results and the consistency between the domain names related to sending e-mails with respect to the clusters. Actually, as presented in our previous study [1], we confirmed that our previous method can classify a plenty of legitimate forwarders in one cluster. Based on these results, we consider that our method can classify the sender's IP address according to the similarity associated with the e-mail sending operation in addition to the classification of the forwarding servers in our clustering phase.

We assume that the dataset is a large number of DMARC reports. Therefore, non-hierarchical clustering is better than hierarchical clustering for our method. In addition, when using a non-hierarchical clustering algorithm in our approach, we assume that it is difficult to determine the number of clusters because the scale of the DMARC reports is not constant depending on the DMARC report receiving domain. Several algorithms have been proposed that can automatically estimate the number of clusters in non-hierarchical clustering; such algorithms include affinity propagation [28], the Bayesian Gaussian mixture model [29], and X-means clustering [6]. However, to appropriately estimate the number of clusters for affinity propagation, we need to set the number of the "preference," which is the preference value for each point, depending on the dataset. Meanwhile, to estimate the appropriate number of clusters for Bayesian Gaussian mixture model, we need to adjust the parameter "reg_covar," which is a regularization added to the diagonal of the covariance, depending on the dataset. Therefore, because our method is to

applied to actual DMARC reports, the affinity propagation and Bayesian Gaussian mixture model approaches are not suitable as clustering algorithms for our approach. X-means, which is also a non-hierarchical clustering approach, is a K-means extended algorithm proposed by Pelleg and Moore [6]. K-means is one of the most popular clustering methods but has a shortcoming, in which the number of clusters, K, needs to be provided by users in advance.

Conversely, X-means can determine the number of clusters, X, via iterations of K-means and splitting decisions based on the Bayesian information criterion (BIC) without complicated parameter adjustments. Accordingly, our method uses an X-means clustering analysis to classify the sender's IP address.

In the X-means clustering flow, the senders' IP addresses are divided into clusters according to their e-mail transmission behavior trends, such as the consistency between the domain names related to the e-mail sending and its authentication results.

C. Legitimate sender detection

Third, we detect the legitimate senders in our proposed approach ((C) in Figure 7). We determine legitimate senders clusters based on the Spamhaus blocklist that is the most famous IP blacklist in the world. This detection flow consists of two inspections as (C-1) and (C-2) in Figure 7.

1) The first inspection:

The first inspection ((C-1) in Figure 7) checks all of the IP addresses in all of the clusters to determine if they are listed in the Spamhaus blocklist. Then, our method classifies the clusters that do not include any IP addresses registered in the Spamhaus blocklist as legitimate senders clusters. Other clusters that have one or more IP addresses in the Spamhaus blocklist are passed to the second inspection ((C-2) in Figure 7).

2) The second inspection:

As mentioned in Section III-C1, the clusters to be checked in the second inspection ((C-2) in Figure 7) have not been determined to be clusters of legitimate senders in the first inspection (C-1) because one or more of their IP addresses are registered in the Spamhaus blocklist. In other words, these clusters consist of both blacklisted IP addresses and non-blacklisted IP addresses. As an example, let us consider the cluster that consists of one blacklisted IP address and 99 white IP addresses. As described in Section II-A, the registration, and deregistration of IP addresses on the blacklist may be delayed. Therefore, even if the e-mail sending operation of this cluster is legitimate, 99 non-blacklisted IP addresses may be affected by the one blacklisted IP address, whose deregistration from the blacklist has been delayed. Accordingly, our method performs a second inspection (C-2) to further improve the false positive detection performance. However, because these clusters actually contain at least one or more blacklisted IP addresses, our method cannot use these clusters to detect legitimate servers. Therefore, as the second inspection (C-2), our method performs the following clustering to detect additional legitimate servers.

As the first step in the second inspection (C-2), our method adapts the DMARC report summarization phase and X-means clustering phase to the IP addresses constructing clusters in the same way as in (A) and (B) in Figure 7. Then, as in the first inspection (C-1), our method checks the Spamhaus blocklist to determine whether one or more IP addresses are listed for each cluster. Further, as with the first inspection (C-1), if no IP addresses are listed in the Spamhaus blocklist, our method classifies the clusters as legitimate senders clusters. Otherwise, our method determines these clusters to be non-legitimate senders clusters.

TABLE I. LIST OF ABBREVIATIONS USED IN THIS PAPER.

Abbreviations	Details
Day	Day the DMARC report received
All_IP	The total number of sender server IP addresses in the DMARC reports
All_mail	The total number of e-mails constructing the DMARC reports
All_rep	The total number of DMARC reports
Tgt_IP	The IP addresses adapting to our method *These IP addresses send 90% of all e-mails constructing the DMARC reports.
1st_Tgt_IP	The IP addresses for X-means in the first inspection ((C-1) in Figure 7) *These IP addresses are same as those in Tgt_IP.
2nd_Tgt_IP	The IP addresses for X-means in the second inspection (C-2 in Figure 7) *These IP addresses are not detected as legitimate senders in the first inspection ((C-1) in Figure 7).
1st_C	The number of clusters as the clustering result of the first inspection ((C-1) in Figure 7)
2nd_C	The number of clusters as the clustering result of the second inspection ((C-2) in Figure 7)
1st_Leg_C	The legitimate sender clusters detected in the first inspection ((C-1) in Figure 7)
2nd_Leg_C	The legitimate sender clusters detected in the second inspection ((C-2) in Figure 7)
1st_Leg_IP	The legitimate IP addresses in *Leg_C* detected in the first inspection ((C-1) in Figure 7)
2nd_Leg_IP	The legitimate IP addresses in *Leg_C* detected in the second inspection ((C-2) in Figure 7)
Leg_IP	All legitimate IP addresses detected by our method *(the # of *Leg_IP*) = (the # of *1st_Leg_IP*) + (the # of *2nd_Leg_IP*)

IV. DATASET

In this section, we describe the dataset we used to test our method. We use the actual DMARC reports received from November 1 to November 30, 2019, at one of the most famous ISP domains in Japan.

The abbreviations that we use in the following discussion and in the results are shown in Table I. Table II shows the number of sender IP addresses ("*All_IP*"), DMARC reports ("*All_rep*"), and e-mails ("*All_mail*") in the DMARC report dataset used in this experiment. As shown in the bottom-most row of the "*All_rep*" column in Table II, we observed 74,199 DMARC reports on average (45,884–100,536). These DMARC reports are constructed by 501,927 e-mails on average (385,115–637,727) that sent from 11,418 sender IP addresses on average (7,390–19,330), as shown in the bottom-most row of the "*All_mail*" and the "*All_IP*" column in Table II, respectively. The "*# of Tgt_IP*" and the "*Tgt_IP/All_IP (%)*" columns in Table II show the number and the ratio to the "*All_IP*" of the sender IP addresses that we apply to our method.

TABLE II. UTILIZED DATASET.

Day	*All_IP*	*All_mail*	*All_rep*	*# of Tgt_IP*	*Tgt_IP / All_IP (%)*
Day 1	12,614	438,216	59,794	1,804	14.3
Day 2	10,314	420,850	51,527	1,346	13.1
Day 3	9,445	494,184	53,033	1,164	12.3
Day 4	7,390	436,984	45,884	1,074	14.5
Day 5	7,641	447,544	46,655	1,100	14.4
Day 6	10,839	592,334	59,038	1,242	11.5
Day 7	11,617	495,553	65,319	1,703	14.7
Day 8	10,996	495,411	67,184	1,812	16.5
Day 9	11,624	491,806	75,665	2,080	17.9
Day 10	9,757	486,201	71,167	2,030	20.8
Day 11	8,013	402,857	65,537	1,972	24.6
Day 12	11,297	510,453	79,405	2,228	19.7
Day 13	12,789	561,485	86,690	2,469	19.3
Day 14	12,584	588,425	92,324	2,537	20.2
Day 15	12,014	626,296	85,930	2,399	20.0
Day 16	11,835	554,598	83,702	2,468	20.9
Day 17	9,796	428,524	78,384	2,428	24.8
Day 18	8,381	385,115	73,551	2,389	28.5
Day 19	11,323	520,894	79,461	2,319	20.5
Day 20	12,179	456,908	76,983	2,478	20.3
Day 21	19,330	637,727	100,536	3,149	16.3
Day 22	13,891	543,432	90,515	2,897	20.9
Day 23	12,027	488,000	81,262	2,714	22.6
Day 24	11,372	506,560	74,459	2,318	20.4
Day 25	8,773	386,306	65,246	2,315	26.4
Day 26	11,523	507,158	77,065	2,517	21.8
Day 27	12,520	493,725	78,698	2,667	21.3
Day 28	15,950	567,088	98,018	3,007	18.9
Day 29	12,467	518,344	81,857	2,717	21.8
Day 30	12,250	574,834	81,073	2,562	20.9
Minimum	7,390	385,115	45,884	1,074	11.5
Maximum	19,330	637,727	100,536	3,149	28.5
Average	11,418	501,927	74,199	2,197	19.3

As shown in the bottom-most row of Table II, the number of *Tgt_IP* accounts for 19.3% on average (11.5–28.5%) of *All_IP*. According to our observations in Table II, *Tgt_IP* sends more than 90% of the e-mails of *All_mail*. By contrast, the remaining IP addresses, which are not *Tgt_IP*, send less than 10% of the e-mails of *All_mail*. Because these remaining IP addresses, which send only a few e-mails, will constitute noise for the X-means clustering algorithm, we utilize only the DMARC reports, for which the senders' IP addresses are included in *Tgt_IP*.

V. RESULTS

In this section, we explain the results obtained by applying our method to the dataset described in Section IV. Table III shows the results of the X-means clustering and legitimate IP address detection.

First, the average number of IP addresses for the first inspection (*1st_Tgt_IP*) is 2,197 per day (1,074–3,149). In the first inspection ((C-1) in Figure 7), our method divided *1st_Tgt_IP* into 20 clusters on all days, as shown in the "*1st_C*" column in Table III. As the result of the first inspection (C-1), the number of legitimate sender clusters (*1st_Leg_C*) was 15 per day on average (12–17), as shown in the "# of *1st_Leg_C*." Moreover, 324 IP addresses per day on average (164–493) were contained within *1st_Leg_C*, as shown in the "# of *1st_Leg_IP*" column in Table III. Then, as described in Section III, the second inspection ((C-2) in Figure 7) applied DMARC report summarization, X-means clustering, and legitimate sender detection to the clusters for the second inspection (*2nd_Tgt_IP*, which was generated by the first inspection (C-1). As shown in the "# of *2nd_Tgt_IP*" column in Table III, the number of IP addresses subject to the second inspection (*2nd_Tgt_IP*) was 1,873 per day on average (910–2,753). In the second inspection (C-2), our method classified *2nd_Tgt_IP* into 20 clusters on all days, as shown in the "*2nd_C*" column in Table III. The second inspection determined 5 clusters on average (1–11) to be legitimate sender clusters (the "# of *2nd_Leg_C*" column in Table III). In addition, *2nd_Leg_C* consisted of 23 IP addresses on average (1–133), as shown in the "# of *2nd_Leg_IP*" column in Table III.

As a result of applying our method to the dataset, our method detected 347 legitimate senders' IP addresses per day on average (178–732), as shown in the "# of *Leg_IP*" column in Table III.

VI. EVALUATION

In this section, we evaluate the results of applying our method, as described in Section V, to the dataset. As described in Section III, none of the legitimate IP addresses detected in our method are included in the Spamhaus blocklist. This means that none of the legitimate IP addresses detected using our approach were the known spammer IP addresses.

Forwarded e-mails are prone to false positives with sender domain authentications, as described in Section II-B. To determine if our method successfully detected forwarded e-mails as legitimate senders, we confirmed the classification results of five IP addresses that were known forwarding servers in the domain of the ISP that received the DMARC reports used as the dataset. We confirmed that these five IP addresses were successfully classified in the same cluster, which was detected as a legitimate sender cluster by our method.

Then, we evaluated the results focusing on the following two points: the detected legitimate e-mails (A) and the detected false positive deliveries with respect to the sender domain authentications (B).

A. The detected legitimate e-mails

First, we checked the number of e-mails sent from the IP addresses (*Leg_IP*, which consisted of 347 servers on average, as shown in Table III) of the servers that our method detected as legitimate senders.

Figure 8 shows the number of legitimate e-mails sent from the legitimate IP addresses detected by our approach. As shown in this figure, combining the first and second inspections, our method detected 214,153 legitimate e-mails per day on average (110,484–340,473). From this result, we confirmed that our method can detect a large number of legitimate e-mails in the sender authentication.

As mentioned in Section II-A, blacklist techniques have an issue in that both the registration and deregistration of IP addresses is delayed. This delay can cause many non-blacklisted IP addresses to be incorrectly classified into the same cluster as a few blacklisted IP addresses, as we described in Section III-C2. To counter this problem, our method performs a second inspection after the first inspection, as described in Section III-C2. As we can see from Figure 8, which shows the number of legitimate e-mails detected by our method, the second inspection in our method was able to detect 20,141 additional legitimate e-mails per day on average (146–116,888). In particular, for example on *Day* 22

TABLE III. THE RESULTS OF APPLYING OUR METHOD TO THE DATASET.

Day	# of *1st_Tgt_IP*	*1st_C*	# of *1st_Leg_C*	# of *1st_Leg_IP*	# of *2nd_Tgt_IP*	*2nd_C*	# of *2nd_Leg_C*	# of *2nd_Leg_IP*	# of *Leg_IP*
Day 1	1,804	20	17	694	1,110	20	6	38	732
Day 2	1,346	20	16	431	915	20	2	4	435
Day 3	1,164	20	15	199	965	20	3	7	206
Day 4	1,074	20	16	164	910	20	10	14	178
Day 5	1,100	20	17	175	925	20	3	7	182
Day 6	1,242	20	16	277	965	20	5	8	285
Day 7	1,703	20	17	370	1,333	20	4	13	383
Day 8	1,812	20	17	365	1,447	20	4	15	380
Day 9	2,080	20	12	208	1,872	20	8	99	307
Day 10	2,030	20	14	260	1,770	20	1	1	261
Day 11	1,972	20	15	192	1,780	20	7	28	220
Day 12	2,228	20	17	369	1,859	20	1	2	371
Day 13	2,469	20	17	378	2,091	20	5	17	395
Day 14	2,537	20	14	327	2,210	20	2	2	329
Day 15	2,399	20	16	326	2,073	20	3	26	352
Day 16	2,468	20	14	316	2,152	20	6	24	340
Day 17	2,428	20	13	215	2,213	20	7	35	250
Day 18	2,389	20	16	244	2,145	20	2	9	253
Day 19	2,319	20	17	336	1,983	20	11	45	381
Day 20	2,478	20	17	493	1,985	20	1	7	500
Day 21	3,149	20	15	396	2,753	20	4	10	406
Day 22	2,897	20	14	166	2,731	20	9	133	299
Day 23	2,714	20	16	371	2,343	20	3	9	380
Day 24	2,318	20	14	318	2,000	20	6	21	339
Day 25	2,315	20	14	279	2,036	20	1	4	283
Day 26	2,517	20	16	327	2,190	20	5	11	338
Day 27	2,667	20	15	348	2,319	20	5	29	377
Day 28	3,007	20	14	396	2,611	20	5	41	437
Day 29	2,717	20	15	392	2,325	20	4	12	404
Day 30	2,562	20	15	379	2,183	20	4	18	397
Minimum	1,074	20	12	164	910	20	1	1	178
Maximum	3,149	20	17	694	2,753	20	11	133	732
Average	2,197	20	15	324	1,873	20	5	23	347

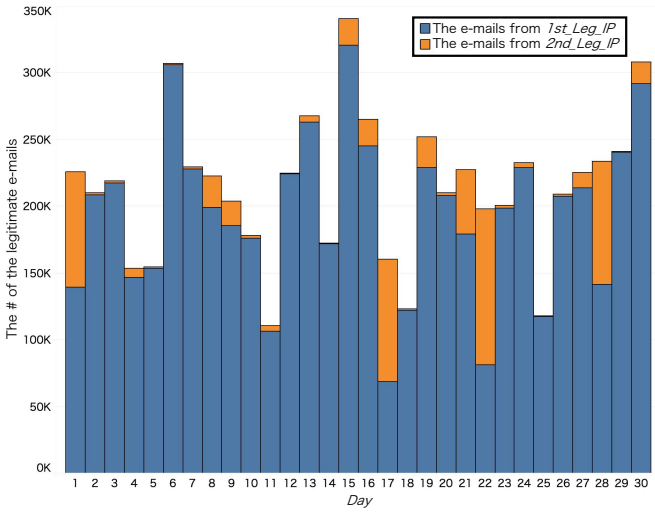


Figure 8. The number of e-mails sent from the legitimate IP addresses detected by our method.

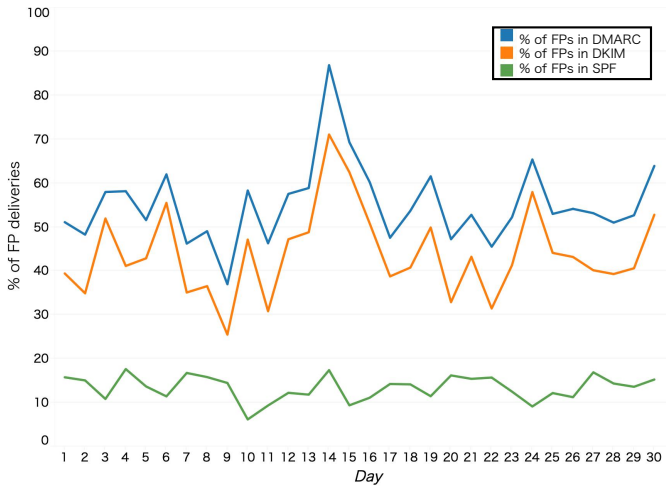
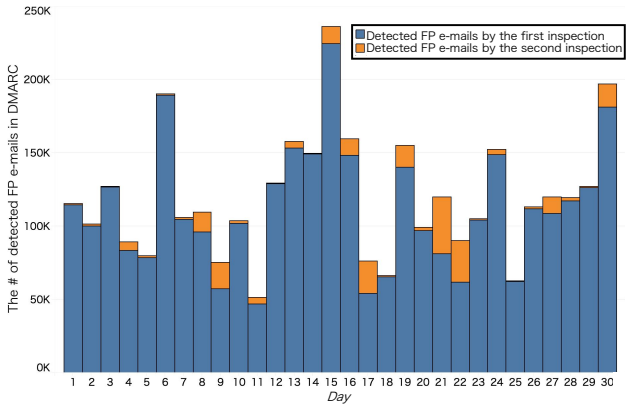


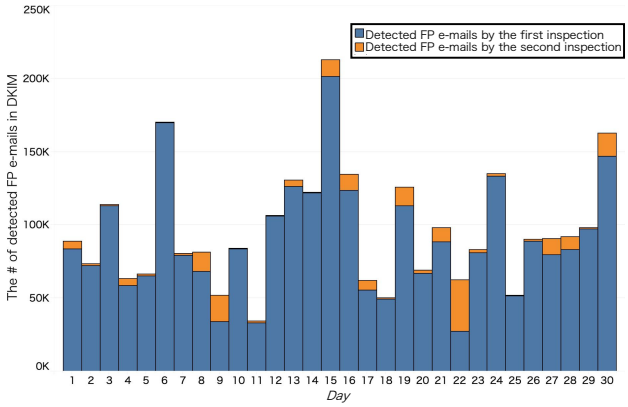
Figure 9. The percentage of false positive deliveries in the sender domain authentication for e-mail deliveries from the legitimate senders detected by our method.

in Figure 8, the first inspection found 81,375 legitimate e-mails, while the second inspection found an additional 116,888 legitimate e-mails. In other words, approximately 59.0% of the legitimate e-mails detected on that day were detected by the second inspection. These results show that the second inspection was able to detect many legitimate senders that were incorrectly classified as non-legitimate sender clusters during the first inspection. Therefore, we confirmed that our method can improve the detection performance by performing a second inspection in addition to the first inspection.

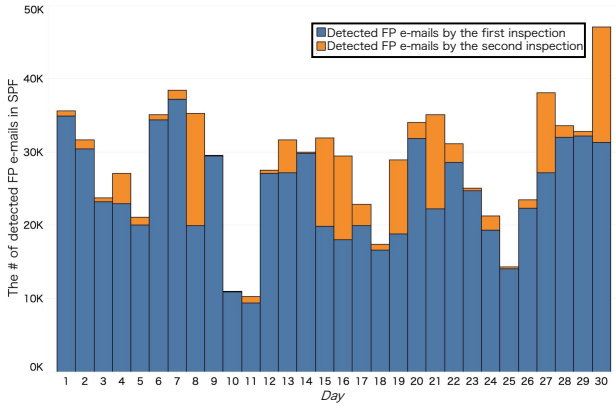
The detected legitimate e-mails contain e-mails that both failed and passed sender domain authentications. Figure 9 shows the ratio of e-mails that failed the SPF, DKIM, and DMARC authentications to the total legitimate e-mails shown in Figure 8 for each day. As shown by the blue line in Figure 9, the ratio of DMARC failed e-mails was 55.1% on average (37.0–86.8%). The orange line in Figure 9 shows the ratio



(a) False positive deliveries in DMARC.



(b) False positive deliveries in DKIM.



(c) False positive deliveries in SPF.

Figure 10. The number of false positive deliveries detected by our approach.

of DKIM failed e-mails, which was 43.9% on average (25.5–71.0%). Meanwhile, the green line in Figure 9 indicates that the ratio of SPF failed e-mails to legitimate e-mails was 13.4% on average (6.2–17.6%).

According to these results, a large number of sender domain authentication failure e-mails are contained in the detected legitimate e-mails. We consider these sender domain authentication failure e-mails in detail in Section VI-B.

B. The detected false positive deliveries in the sender domain authentications

As mentioned in Section VI-A, because the legitimate e-mails detected by our method were sent from legitimate IP addresses, the legitimate e-mails that failed SPF, DKIM, and DMARC are false positives in the sender domain authentications.

In this section, we investigate the number of sender domain authentication failure deliveries, that is, the false positive deliveries, for each method of sender domain authentication.

Figure 10 shows the number of detected false positives for the SPF, DKIM, and DMARC authentications. As shown in Figure 10(a), 119,405 legitimate e-mails on average (51,132–235,908) were detected as false positives in the DMARC authentication by our method, including both the first and the second inspections. In addition, as shown in Figure 10(b), our method detects 96,129 legitimate e-mails on average (34,051–212,661) as false positives in the DKIM authentication. Meanwhile, Figure 10(c) indicates that our method detected 28,466 legitimate e-mails on average (10,322–47,010) as false positives in the SPF authentication.

From these results, we confirmed that our method is able to detect various types of deliveries that are false positives in sender domain authentication without using e-mail contents. By utilizing the proposed method, e-mail system administrators can significantly reduce the false positives that occur with conventional sender domain authentication.

VII. Conclusions

In general anti-spam operation, e-mails are inspected by sender domain authentications prior to content filtering. Therefore, it is critical to reduce the false positives in the sender domain authentication, as opposed to the false negatives, to enable reliable e-mail server operation.

In this paper, we proposed a method to detect false positives generated by existing sender domain authentications by analyzing massive amounts of DMARC report data using an X-means clustering analysis.

Our approach consisted of three phases: DMARC report summarization, X-means clustering, and legitimate sender detection.

In the DMARC report summarization, our approach summarized the DMARC reports for each e-mail sender's IP address focusing on the results of the sender domain authentications and combinations of the Header-From domain, the Envelope-From domain, and the DKIM signature domain.

Then, our approach adapted X-means clustering to the summarized DMARC reports to classify the e-mail sender's IP address based on transmission behavior, such as the consistency between the domain names related to e-mail sending and its authentication.

Next, our approach detected the legitimate sender clusters by processing two inspections. The first inspection checked whether the IP addresses in the clusters were included in the Spamhaus blocklist. If no IP addresses in the cluster were included in the Spamhaus blocklist, the first inspection determined the cluster to be a legitimate sender cluster. The other clusters consisted of both blacklisted IP addresses and non-blacklisted IP addresses. However, non-blacklisted IP addresses may be incorrectly classified into the same cluster as a few blacklisted IP addresses because both the registration and deregistration of IP addresses in the blacklist are not processed immediately. Therefore, to improve the performance of the legitimate sender detection, a second inspection checked the clusters that were not determined to be legitimate sender clusters in the first inspection.

In the second inspection, as in the first inspection, our method aggregated the DMARC reports for the IP addresses that were subject to the second inspection, performed X-means clustering, and determined the validity of the clusters using the Spamhaus blocklist.

We applied our method to actual DMARC report data and detected 214,153 e-mails on average (110,484–340,473) sent from 347 legitimate senders' IP addresses on average (178–732) as legitimate e-mails per day.

In addition, to evaluate the effect of reducing the false positives that occur in the sender domain authentication when using our method, we investigated the percentage of e-mails sent from the legitimate sender addresses that failed sender domain authentications using the DMARC reports. As a result, we confirmed that, on average, 13.4% (6.2–17.6%), 43.9% (25.5–71.0%), and 55.1% (37.0–86.8%) false positives occurred when using SPF, DKIM, and DMARC, respectively. This result shows that our method detects false positive e-mails in conventional anti-spam systems by detecting e-mails, such as forwarded e-mails, which cannot be correctly classified by existing sender authentication technologies.

Our method does not use e-mail contents, only DMARC report data, and can effectively detect deliveries that would be false positives with conventional sender domain authentications. In addition, since this method can be operated independently of the sending and receiving resources of the e-mail system, it can be installed without increasing the load on the entire e-mail system.

The evaluation for accuracy of our method when our method is operated continuously in actual large scale e-mail system is our future work.

Acknowledgments

This paper was written based on the research while Kanako Konno and Naoya Kitagawa, who the authors of this paper, were with Tokyo University of Agriculture and Technology, Japan. The authors would like to thank TwoFive, Inc. for supporting this study. The authors would also like to thank Mr. Shuji Sakuraba (Internet Initiative Japan Inc.) for constructive discussion.

References

[1] K. Konno, N. Kitagawa, S. Sakuraba, and N. Yamai, "Legitimate E-mail Forwarding Server Detection Method by X-means Clustering Utilizing DMARC Reports," in the Eleventh International Conference on Evolving Internet (INTERNET 2019), 2019, pp. 24–29.

[2] FBI (Federal Bureau of Investigation), "Business email compromise the $26 billion scam," 2019, [Online]. Available: https://www.ic3.gov/media/2019/190910.aspx [Accessed: 1st Jun. 2020].

[3] M. Wong and W. Schlitt, "Sender Policy Framework (SPF) for authorizing use of domains in e-mail," 2006.

[4] D. Crocker, T. Hansen, and M. Kucherawy, "DomainKeys Identified Mail (DKIM) signatures," sep 2011.

[5] M. Kucherawy and E. Zwicky, "Domain-based message authentication, reporting, and conformance (DMARC)," 2015.

[6] D. Pelleg and A. Moore, "X-means: Extending K-means with Efficient Estimation of the Number of Clusters," in Proceedings of the 17th International Conference on Machine Learning, 2000, pp. 727–734.

[7] E. Harris, "The Next Step in the Spam Control War: Greylisting," 2019, [Online]. Available: http://projects.puremagic.com/greylisting/whitepaper.html [Accessed: 1st Jun. 2019].

[8] T. Hunter, P. Terry, and A. Judge, "Distributed tarpitting: Impeding spam across multiple servers," in Proceedings of the 17th Large Installation Systems Administration, vol. 3, 2003, pp. 223–236.

[9] N. Kitagawa, H. Takakura, and T. Suzuki, "An anti-spam method via real-time retransmission detection," in Proceedings of 18th IEEE International Conference on Networks (ICON), 2012, pp. 382–388.

[10] MxToolbox, Inc., "MxToolbox," [Online]. Available: https://mxtoolbox.com/SuperTool.aspx [Accessed: 1st Jun. 2020].

[11] Cisco Systems, Inc, "SpamCop Blocking List," [Online]. Available: https://www.spamcop.net/bl.shtml [Accessed: 1st Jun. 2020].

[12] Barracuda Networks., "Barracuda Reputation Block List (BRBL)," [Online]. Available: http://www.barracudacentral.org/rbl [Accessed: 1st Jun. 2020].

[13] S. M. Jones, "DMARC Required For UK Government Services By October 1st," 2016, [Online]. Available: https://dmarc.org/2016/06/dmarc-required-for-uk-government-services-by-october-1st/ [Accessed: 1st Jun 2020].

[14] S. M. Jones, "Australian Government Agency Recommends DMARC, DKIM, and SPF," 2016, [Online]. Available: https://dmarc.org/2016/08/australian-government-agency-recommends-dmarc-dkim-and-spf/ [Accessed: 1st Jun. 2020].

[15] U.S. Department of Homeland Security, "Binding Operational Directive 18-01 – Enhance Email and Web Security," 2017, [Online]. Available: https://cyber.dhs.gov/bod/18-01/ [Accessed: 1st Jun. 2020].

[16] P. Graham, "A plan for spam," 2002, [Online]. Available: http://www.paulgraham.com/spam.html [Accessed: 22th Feb. 2020].

[17] P. Graham, "Better bayesian filtering," 2003, [Online]. Available: http://www.paulgraham.com/better.html [Accessed: 22th Feb. 2020].

[18] I. Androutsopoulos, J. Koutsias, K. V. Chandrinos, and C. D. Spyropoulos, "An experimental comparison of naive Bayesian and keyword-based anti-spam filtering with personal e-mail messages," in Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval, 2000, pp. 160–167.

[19] S. Aggarwal, V. Kumar, and S. D. Sudarsan, "Identification and detection of phishing emails using natural language processing techniques," in Proceedings of the 7th International Conference on Security of Information and Networks, 2014, pp. 217–222.

[20] H. Ducker, D. Wy, and V. N. Vapnik, "Support vector machines for spam categorization," IEEE Transactions on Neural networks, vol. 10, no. 5, pp. 1048–1054, 1999.

[21] W. Feng, J. Sun, L. Zhang, C. Cao, and Q. Yang, "A support vector machine based naive Bayes algorithm for spam filtering," in 2016 IEEE 35th International Performance Computing and Communications Conference (IPCCC). IEEE, 2016, pp. 1–8.

[22] T. S. Guzella and W. M. Caminhas, "A review of machine learning approaches to spam filtering," Expert Systems with Applications, vol. 36, no. 7, pp. 10 206–10 222, 2009.

[23] M. Crawford, T. M. Khoshgoftaar, J. D. Prusa, A. N. Richter, and H. A. Najada, "Survey of review spam detection using machine learning techniques," Journal of Big Data, vol. 2, no. 1, p. 23, 2015.

[24] "The Apache SpamAssassin Project," [Online]. Available: http://spamassassin.apache.org/ [Accessed: 1st Jun. 2020].

[25] J. Mason, "Filtering spam with spamassassin," 2002.

[26] S. Sinha, M. Bailey, and F. Jahanian, "Shades of Grey: On the effectiveness of reputation-based "blacklists"," in 2008 3rd International Conference on Malicious and Unwanted Software (MALWARE), 2008, pp. 57–64.

[27] C. J. Dietrich and C. Rossow, "Empirical research of ip blacklists," in ISSE 2008 Securing Electronic Business Processes. Springer, 2009, pp. 163–171.

[28] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," Science, vol. 315, no. 5814, pp. 972–976, 2007.

[29] H. Attias, "A variational baysian framework for graphical models," in Advances in Neural Information Processing Systems 12, S. A. Solla, T. K. Leen, and K. Müller, Eds. MIT Press, 2000, pp. 209–215.

A Review on IoT Frameworks Supporting Multi-Level Interoperability – The Semantic Social Network of Things Framework

Antonios Pliatsios, Christos Goumopoulos
Information and Communication Systems Engineering Dept., University of the Aegean
Samos, Greece
e-mail: icsdd18007@icsd.aegean.gr, goumop@aegean.gr

Konstantinos Kotis
Dept. of Cultural Technology and Communication
University of the Aegean
Mytilene, Greece
e-mail: kotis@aegean.gr

Abstract— The Internet of Things (IoT) paradigm advocates the massive use of sensing and communication technologies embedded in the physical world, which provides the potential to collect huge volumes of data and connect them to intelligent systems. As the number of IoT devices is increasing with geometric progress, ensuring interoperability and handling of the big heterogeneous data they generate is of major importance for the development of smart applications and services. In this context, a systematic review of contemporary IoT frameworks based on a multi-level interoperability consideration is performed and findings are critically discussed. Challenges and open issues that emerge in this research area are pointed out, and research opportunities and insights are suggested. Motivated by the shortcomings of the current solutions to support open, interoperable, intelligent and collaborative IoT environments, the concept of Semantic Social Network of Things (SSNT) is introduced. SSNT specifies the integration of device-to-device collaborative services which semantically enable heterogeneous objects to (socially) interact and participate in communities of smart objects. By establishing social relationships and taking collaborative actions, such communities can support users to achieve their goals. A middleware-based framework architecture is presented to enact the SSNT abstraction, and a proof-of-concept application in the smart agriculture domain is outlined to demonstrate important features of this approach.

Keywords- Review; Internet of Things; Interoperability; IoT frameworks; Ontologies; Semantic Social Network of Things

I. INTRODUCTION

The Internet of Things (IoT) is the up-and-coming big step in the field of technology. The IoT concept, initially utilized as an umbrella term for a range of various emerging technologies such as “embedded internet” and “pervasive computing”, is currently paving its way for being the key driver for digital transformation in several application domains among which manufacturing, automotive, health, smart cities and smart farming.

IoT growth is explosive and there are already billions of connected smart objects, embedded systems, sensors and microcontrollers that have penetrated our world connecting home users, businesses, public facilities and enterprise systems. New technologies are being developed to meet the continuous incremental requirements of a new digital world where heterogeneous devices have been connected, forming a part of the IoT ecosystem. Since the density of IoT systems and technologies is becoming increasingly high, ensuring interoperability and handling of large-scale heterogeneous data is turning into a vital key factor in the development of successful smart applications [1].

Undoubtedly, there are still many challenges to overcome in order to fully realize the IoT vision [2][3]. The vast number of interconnected devices gives rise to scalability, heterogeneity and several interoperability issues [4]. One of the crucial issues is that IoT landscape is made up of proprietary devices and platforms that were created to provide a single service and act as "vertical silos" [5]. These silos require the creation of cross-domain, cross-platform and cross-organizational services due to their lack of interoperability and openness. Thus, there is an important need to revise the philosophy of IoT platforms and focus on trying to build synergies between different IoT platforms. This will lay the foundation for interconnecting IoT devices and services that collaborate together to achieve a common goal defined implicitly or explicitly by people.

In this paper, a review of contemporary IoT frameworks is performed to analyze and evaluate relevant contributions related to the establishment of open, interoperable, intelligent and collaborative IoT environments. Accordingly, a classification scheme is proposed to effectively represent the results of the related literature review analysis. The classification is based on the four interoperability levels, i.e., technical, syntactic, semantic and organizational, explored in our previous work [1], extended by the broader scope of the systematic literature review performed. The comparative analysis of the explored IoT frameworks allows us to identify important limitations, challenges and open issues that future research needs to address. Our investigation on the topic can be framed by the following research questions:

RQ1: Do the current IoT frameworks provide solutions supporting multi-level interoperability?

RQ2: What is missing from current IoT frameworks in order to fully support open environments/spaces of heterogeneous but collaborative smart objects?
RQ3: What are the open issues that future researchers should focus on in terms of smart objects interoperability?
RQ4: How collaboration and social interaction mechanisms can at a conceptual level address multi-level interoperability issues in open IoT environments?

The literature review follows a systematic approach consisting of three phases as suggested by [6]:

1) Review planning: specification of research questions and classification scheme; and development of the review protocol which includes the research strategy (literature databases, research keywords) and the definition of inclusion/exclusion criteria.
2) Review running gathering of scientific publications according to the research strategy; and selection of relevant work by applying the selection criteria.
3) Review reporting: overview of the selected work; and comparative analysis of the explored solutions based on the specified classification scheme.

Regarding the review protocol, several academic bibliography sources were used such as Web of Science, Google Scholar, IEEE Xplore Digital Library, Elsevier Scopus, ACM digital library, Citeseer library, Science Direct, and arXiv.org in order to search for relevant scientific contributions of the last 10 years. Search keywords were limited to the following terms: Internet of things, Web of Things, Interoperability, Ontologies, Semantics, and Social IoT. In addition, the following search expressions were used: IoT Frameworks addressing interoperability OR Interoperability OR Internet of Things OR Semantic Web of Things AND Semantic Web Technologies OR Interoperability OR ontology.

Besides the chronological filtering, other selection criteria for the bibliography collection included the publication language (studies had to be written in English) and the pertinence to the research agenda of the review. Selected studies had to present initiatives related to interoperability in the IoT domain, as well as current IoT frameworks that provide solutions improving interoperability, covering at least one of the research questions stated. Both conference and journal papers were eligible but not short studies.

Motivated by the identified shortcomings of the reviewed solutions to support open, interoperable, intelligent and collaborative IoT environments the concept of Semantic Social Network of Things (SSNT) is introduced. SSNT specifies the integration of device-to-device collaborative services which semantically enable heterogeneous objects to (socially) interact and participate in communities of smart objects. By establishing social relationships and taking collaborative actions, such communities can support users to achieve their common goals. In a sense, the interoperable societies of things, services and people are forming an SSNT structure that allows scalable object/service discovery as in the case of social networks of humans.

Towards realizing the SSNT concept, a framework is proposed for the establishment and exploitation of social relationships among heterogeneous but interoperable smart things. A high-level architecture is presented specifying the main components that enable things/objects to be identified as potentially able to participate in communities of smart things/objects, creating groups of common interest and working collaboratively towards achieving common goals. Furthermore, a proof-of-concept application in the smart agriculture domain is outlined to demonstrate important features of this approach. In this example scenario, summaries of sensor data are translated to the RDF modeling language based on the Semantic Sensor Network (SSN) ontology. When the sensor data streams are semantically annotated, semantic techniques (e.g., SPARQL queries and reasoning) can be used for efficient processing. Then, social groups of objects (generating and consuming the annotated data) are created that aim to achieve common goals, and new knowledge is produced from their interaction.

The contributions of this paper can be summarized as follows:

- We provide an extensive review of the up-to-date research progress on contemporary solutions regarding interoperability in the IoT domain.
- We propose a classification which contributes to representing a deep analysis of a comprehensive literature review, as well as comparing IoT frameworks with a view to providing solutions supporting multi-level interoperability.
- We identify a number of limitations, challenges and open issues that future studies in this research area of IoT need to focus on.
- We introduce the concept of Semantic Social Network of Things (SSNT) to describe a network of things that 'speak', 'behave', 'collaborate' and 'co-exist' just like a 'social network' of people, establishing social relationships and taking collaborative actions to support users to achieve their common goals.
- We propose an architectural design for the SSNT framework that specifies the main software components to seamlessly confront the problem of multi-level interoperability tackling also the constraints of devices with limited resources. An evaluation scenario of the SSNT framework in the agricultural domain, representing an instantiation of the SSNT framework, is also provided.

The structure of the paper is as follows. Section II presents background knowledge and motivation. In Section III state of the art approaches confronting interoperability in the IoT domain are reviewed and reported. Section IV outlines essential design requirements to develop a novel interoperable IoT framework and highlights open research

challenges, as it also discusses the architecture and main modules of the proposed SSNT framework, with an aim to enhance interoperability and collaboration in IoT environments. Finally, Section V concludes the paper.

II. Background and Motivation

This section outlines the evolution of existing approaches in the direction of establishing interoperable and cooperative IoT environments. In addition, the multi-level interoperability taxonomy that is used in the systematic review of IoT frameworks is presented.

A. From the Internet of Things to the Semantic Social Network of Things

The IoT concept implies that all things are harmoniously connected so they can communicate, and they are also easily accessible from the Internet to deliver services to end-users [7]. Presently, one of the biggest problems which the IoT is facing, concerns the lack of interoperability, arising from the heterogeneity of devices, systems, protocols and platforms. Consequently, it is necessary to focus on an interoperable and collaborative IoT. A first step in this direction is provided by *Web of Things* (*WoT*) [8]. WoT provides an Application Layer that simplifies the development of IoT applications composed of multiple devices across different platforms and application domains. WoT develops IoT with a common stack based on web services. Unlike IoT that focuses on the Network Layer, WoT assumes that connectivity between devices is achieved and focuses on how to build IoT applications. But even if the problem of interconnection with the help of web protocols such as HTTP (Hypertext Transfer Protocol Secure) and CoAP (Constrained Application Protocol) has been resolved, the problem of perception and context awareness in IoT ecosystems remains.

For this reason, the *Semantic Web of Things* (*SWoT*) is proposed [9]. SWoT is a current exploration area targeting to assimilate Semantic Web-based technologies with the IoT. It can also be considered as a transformation of the WoT by incorporating semantics. SWoT targets the ability to exchange and use information among data and ontologies. However, the challenges to move from the IoT and WoT towards the SWoT are numerous; some of these are to define a common description that allows data, and device description to be universally understandable, create extensible annotations, i.e., from minimal semantic descriptions towards more elaborate ones.

Currently, there are significant ongoing efforts for the definition of common semantics to collaborate on different data modelling approaches. Cross-domain interoperability is expected to be one of the main drivers for the realization of the next state of IoT computing paradigm which is already getting shape under the term of *Internet of Everything* (*IoE*). The IoE "*is bringing together people, processes, data, and things to make network connections more relevant and valuable than ever before-turning information into actions that create new capabilities, richer experiences, and unprecedented economic opportunity for businesses, individuals, and countries*" [10]. Figure 1 depicts the IoE data management model.

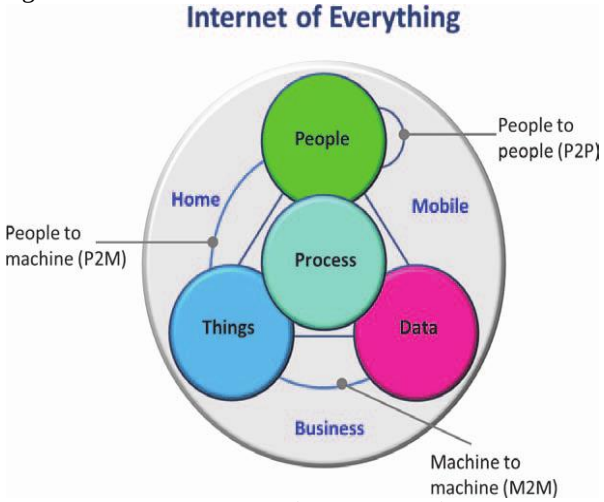


Figure 1. Data Management Model for the Internet of Everything [10].

Another approach towards a collaborative and interoperable IoT is the *Social Internet of Things* (*SIoT*). In social IoT, different devices work together to create social relationships with each other (such as social relationships on social network of people) [11][12]. The basic idea is to utilize human social networks (e.g., Twitter) as service discovery and provisioning infrastructure. However, the proposed notion does not align with the fundamental concept of IoT in which the ubiquitous connectivity of objects is envisioned to provide services to humans. Another attempt is made in a related work where authors discussed the integration of IoT with social networks [13]. An important step in laying down the vision of SIoT is taken in [14][15]. Therein, the various policies to determine the establishment and management of relationships among IoT objects are discussed. Different perspectives between human and IoT social networks are outlined in Figure 2.

SIoT defines several forms of socialization between objects. Firstly, the parent-object relationship is defined between objects manufactured by the same company. In addition, between objects there are relationships of people who share experiences, for example in a discussion or in their work or in any interaction. Another type of relationship is defined for objects owned by the same user such as smartphones, computers, smart TVs, etc. This relationship is called the object-ownership relationship. Finally, social-object relationship is defined when devices come in contact with their owners, such as smartphones belonging to friends. To manage the resulting network and relationships, a foreseen SIoT architecture is made of four major components [15] among others. Relationship management enables SIoT to begin updating and terminating relationships between objects. The service discovery identifies which items can provide the required service in the same way that people search for friendships and information. The composition of

services allows for interaction between objects. Reliability management aims at understanding how information is processed by other members.

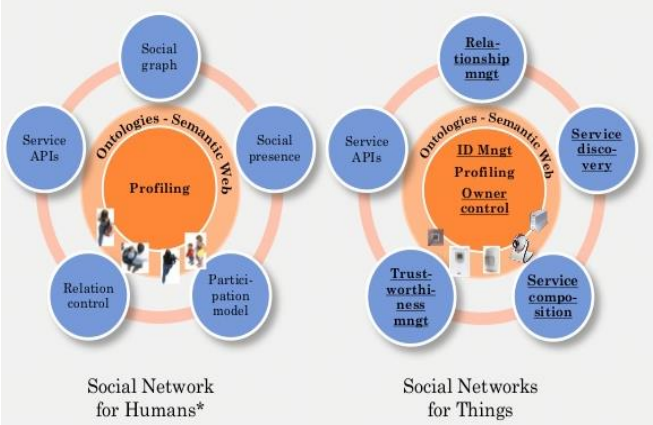


Figure 2. Comparison between Social Network of Humans and SIoT [14].

In this paper an approach that seeks to provide mechanisms to create interoperable, collaborative and open IoT environments is proposed and captured in the concept of *Semantic Social Network of Things* (*SSNT*). SSNT entails a network of things that 'speak', 'behave', 'collaborate' and 'co-exist' just like a 'social network' of people. For example, different intelligent objects that are able to interconnect and make decisions in an interoperable way, without human intervention, constitute an SSNT. This should not be misinterpreted as smart objects in a social network. Even a more elaborated case is that these devices may not only inform each other but negotiate a result. For example, smart home sensors working together to adjust the power consumption to suit the user preferences and concurrently optimize cost based on electricity provider charging policies.

An everyday life application scenario is described next to make the SSNT more understandable. Let us consider a businesswoman named Rafaela, who lives in Athens and employs an SSNT network and an SSNT recommendation system. Rafaela plans to invite her colleagues, who are based in other countries, to Athens for a critical meeting next week. She wants to make an appointment that should be accessible to all of her partners based on their availability. For this purpose, she initiates an appointment using her system which is based on an SSNT network containing all the information available to Rafaela and her associates. It is important to note that by the time Rafaela uses the SSNT network, IoT devices have already maintained social relationships with other IoT devices using the SSNT perception layer. The system coordinates with the affiliate scheduling systems and proposes an appointment time for her and other affiliates based on their availability, and the availability of airline flights. This is done by overcoming problems of lack of data interoperability as her colleagues are located in countries where the date and time system is different (e.g., USA, China).

In addition, Rafaela has some health problems, the most important of which is asthma, so there is information on her health in a system based on an SSNT network via wearables and sensors mounted in indoor and outdoor environments. In this way, an SSNT network can recommend that at the meeting location the air quality is acceptable or not. Again, data interoperability problems are overcome as city/building sensors measure their parameters, for example, in different units of measurement. Based on these recommendations from the SSNT networks, the proper recommendations of the meeting place and time can be provided. When all users confirm this appointment, the system sets it and sends an invitation to everyone.

Table I summarizes the evolution route from IoT to the proposed SSNT. It shows the key features and objectives of the approaches discussed such as WoT, SWoT, IoE, SIoT up to the SSNT. As shown in Table I, WoT attempts to reuse and adapt existing web technologies to build new applications and services [9]. SWoT focuses on machine-understandable data and in the description of data with common vocabularies, concentrating on the reuse of domain knowledge. SIoT targets to utilize human social networks as service discovery and provisioning infrastructure [11][12]. Finally, the proposed SSNT framework focuses on a network where different intelligent objects are able to interconnect and make decisions without human intervention leveraging on semantically annotated information.

B. IoT Interoperability Levels

There are numerous definitions in the literature for interoperability. The IEEE defines interoperability as "*the ability of two or more systems or components to exchange information and use the information exchanged*" [16]. Moreover, interoperability can be defined as a measure of the degree to which diverse systems, organizations, and/or individuals are able to work together to achieve a common goal [17]. IoT interoperability is a multifaceted issue and the solutions to be addressed must be in line with many factors that are also referred to in the literature as interoperability levels. A taxonomy of interoperability for IoT is based on four levels: technical, syntactic, semantic and organizational interoperability [18][19]. In the following each level is analyzed explicitly.

1) Technical Interoperability

Technical Interoperability includes three sublevels of classification, namely, the interoperability of devices, the interoperability of networks and the interoperability of platforms.

a) Device Interoperability

Typically, an IoT system or IoT Sensor Network communication is designed using one of the popular low-level standard technologies like Zigbee, Bluetooth Mesh, Z-wave, WiFi, etc. for devices to collaborate with each other. One of the current IoT challenges is to add a new device in an existing network that is having a different communication protocol to collaborate compared to the existing device network.

TABLE I. FROM SOCIAL NETWORK AND IOT TO SSNT.

	Source	Nodes	Connection	Enabling Technologies/Services	Target
World Wide Web	[2]	Web pages	Hyperlinks	HTML, XML	Share resources
Social Network	[11][12]	Persons	Social Relations	Network analysis, Community detection	Analysis relation principles and evolution
IoT	[2][3][4][7]	Devices/ Objects/ Things	Wireless signals	RFID, LoRaWan, Bluetooth, GPS, IPv6, …	Remote detection and control
WoT	[7][8]	Web-enabled objects	Web, Smart Gateways	REST, HTTP, CoAP, JSON, Web sockets	WoT attempts to reuse and adapt existing web technologies to build new applications and services
SWoT	[7][8][9]	Machine-Understandable Objects	Semantic Web, Smart Gateways	JSON-LD, linked data, Ontologies, Linked Open Vocabularies, Reasoners	Machine-understandable data - Describe data with common vocabularies - Reuse domain knowledge - Link to other data - Ease the reasoning
IoE	[10][11]	People, Things, Data	Internet, TCP/IP	IPv6 extensions (MIPv6, GLoWBAL IPv6)	Intelligent connection. Machines will become more intelligent and cognitive by having more access to data and expanded network opportunities.
SIoT	[12][13][14][15]	Objects, Humans, Data	Social Relations of Things' owners	Relationship management, Service discovery, Service composition, Trust management	Utilize human social networks (e.g. Twitter) as service discovery and provisioning infrastructure
SSNT	this work	Objects, Humans, Data	Social Relations and semantic links of Things, Platforms, Networks	SSNT Architecture Layers and Framework Modules	A social network where heterogeneous intelligent objects are able to interconnect and make decisions without human intervention

Interoperability of the IoT devices is hence becoming more and more important to build a scalable, adaptable and a seamless IoT device network [20]. The IoT ecosystem needs interoperability to create seamless programmability or configurability of the various products or devices or sensors to connect and collaborate. There is a need for a consolidated common standard that makes devices communicable, operable, and programmable, regardless of make, model, manufacturer, or industry. For example, consider a smart home scenario where the light bulbs and thermostats use ZigBee, speakers communicate with Bluetooth, and switches communicate through WiFi. Interoperability in this example enables different devices to understand and translate between these disparate communications technologies. An ideal IoT platform would offer a pool of standardized communication protocols where the device manufacturers may select the appropriate protocols [20] (e.g., CoAP for constrained devices). In the literature, device level interoperability relies either on a gateway solution (sometimes called protocol converters) that can be extended using plug-ins, to support new communication protocols or by instructing the device vendors to only use the protocols that are supported (such as Fosstrak). For example, the Apple HomeKit, If-This-Then-That (IFTTT) Eclipse Ponte and Light-Weight M2M (LWM2M) are some of the gateway solutions in the literature [21].

Devices that are integrated into the world of IoT are becoming more and more ubiquitous. These smart devices / things are either devices with a lot of computing power like smartphones and Raspberry Pi, or devices with built-in microswitches and low-power actuators, such as Arduino, Wispmote, Libelium, and others [22]. The problem of interoperability at this level is due to the inability of all these devices with different architectures and power levels to interact properly.

b) Network Interoperability

Moreover, due to the variety and heterogeneity of IoT devices, many communication protocols have been developed to cover all requirements in the IoT market. Home appliances, such as smart air conditioners, refrigerators, televisions, etc., use WiFi and 2G / 3G / 4G cellular communications. Other mobile devices use more low-power and short-range wireless technologies, such as Bluetooth, ZigBee, Beacons, RFID belonging to the WBAN IEEE 802.15.6 family. While a new category created for sensor applications is that of long-range and Low-Power Wide-Area Networks (LPWAN). Some of them are the wireless technologies LoRaWan, SigFox and NB-IoT [23]. This level of interoperability refers to the difficulty of communication of the IoT devices using different communication protocols.

At this level of interoperability, mechanisms are used that allow the continuous exchange of messages between systems across different heterogeneous networks. These include issues such as addressing, routing, resource optimization, security, QoS and mobility support.

c) Platform Interoperability

The IoT platform is a comprehensive suite of services that facilitates services, such as development, maintenance, analysis, visualization and intelligent decision-making capabilities in an IoT application. Interoperability issues of IoT platforms appear because many of these systems are tailored for specific IoT applications. Some of the most popular platforms are Google Cloud Platform, IBM Watson IoT, ThingWorx, oneM2M, Microsoft Azure Cloud, ThingSpeak [24]. Each of the above platforms follows its data sharing policy, it has its operating system, and this has the effect of creating heterogeneous IoT systems and increasing the problem of interoperability.

Today, the IoT environment comprises vertically oriented platforms for things. Developers who want to use them need to negotiate access individually and adapt to the platform-specific API and information models. Having to perform these actions for each platform often outweighs the possible gains from adapting applications to multiple platforms. This fragmentation of the IoT and the missing interoperability result in high entry barriers for developers and prevent the emergence of broadly accepted IoT ecosystems.

Today, we are dealing with various vertically oriented and mostly closed systems. Architectures for IoT are built on heterogeneous standards [25][26][27] (e.g., IETF CoAP, OASIS MQTT, OMA LWM2M, OGC SWE, or OneM2M) or even proprietary interfaces. As a result, most existing and emerging IoT platforms offer heterogeneous ways of accessing things and their data. This causes interoperability problems when overarching, cross-platform, and cross-domain applications are to be built, and eventually prevents the emergence of vibrant IoT ecosystems

For example, the Apple HomeKit supports its own open source language Swift, Google Brillo uses Weave, and Amazon AWS IoT offers SDKs for embedded C and NodeJS [24]. This non-uniformity causes hindrance for application developers to develop cross-platform and cross-domain IoT applications. Developers need to obtain extensive knowledge of the platform specific APIs and information models of each different platform to be able to adapt their applications from one platform to another. A cross-platform IoT application can access different IoT platforms and integrate data from various platforms. After cross-platform interoperability is enabled, cross-domain interoperability can be achieved in which different platforms within heterogeneous domains are federated to build horizontal IoT applications. For example, a smart home platform can provide domain-specific enablers such as air temperature and lighting conditions. These enablers can then be exploited by other IoT platforms, such as smart healthcare, to provide more innovative applications and scenarios.

2) Syntactic Interoperability

Syntactic interoperability refers to the interoperability of data formats and encodings used in any exchange of information or services between heterogeneous systems and IoT entities. Such forms of standardization are, for example, XML (Extensible Markup Language), JSON (JavaScript Object Notation) and RDF (Resource Description Framework) [28]. The encoding and decoding of messages are done using editorial rules, defined by a grammar. The problem of syntactic interoperability arises due to the great variety of grammars that each architecture employs and consequently, the IoT devices could not communicate properly.

Syntactic interoperability, provided, for instance, by XML or the SQL (Structured Query Language) standards [29], is a prerequisite to semantic definitions. It involves a common data format and common protocol to structure any data so that the manner of information processing will be interpretable from the structure. It also allows detection of syntactic errors, thus allowing receiving systems to request resending of any message that appears to be garbled or incomplete. No semantic communication is possible if the syntax is garbled or unable to represent the data. However, the information represented in one syntax may in some cases be accurately translated into a different syntax. Where accurate translation of syntaxes is possible, systems using different syntaxes may also be interpreted accurately. In some cases, the ability to accurately translate information among systems using different syntaxes may be limited to one direction, when the formalisms used have different levels of expressivity (ability to express information).

3) Semantic Interoperability

Semantic interoperability is characterized as the ability to transmit information, data and knowledge among agents, services and applications in a meaningful way, inside and outside the Semantic Web [30][31]. It is the description of smart devices according to their data, services, and capabilities in mechanically comprehensible form using a common vocabulary. Semantic interoperability is achieved

when the exchange of data is made harmoniously independent of the structure of the original data giving a common meaning [32]. This can be done either by existing standards or agreements on the form and importance of data or can be done using a common vocabulary either in a schema and/or in an ontological approach [33].

The use of an ontology is the most common way of adding semantics to the IoT data. It is a way of modelling information that extends the concept of the Semantic Web into the IoT. The most important Semantic Web technologies have been standardized by the World Wide Web Consortium and are: Resource Description Framework RDF - a lightweight data metadata model for describing ontology properties, SPARQL, and the RDF Query Language.

Existing solutions [35][36] suggest the use of unified ontologies to address semantic interoperability issues and automation related to the heterogeneity of data. However, the multiple possible consolidations developed by field experts pose many challenges as each consolidated ontology proposes its autonomous classification. It is therefore imperative to improve ontology matching and ontology alignment [37] to discover the most appropriate strategies that can overcome the heterogeneity problem in the IoT and bridge the semantic gap between IoT entities at the level of Information / Applications.

4) Organizational Interoperability

Organizational interoperability refers to the successful organization of a system to communicate effectively and to transmit the information in a harmonious manner [37]. To do this, the other three levels of interoperability, i.e., technical, syntactic and semantic interoperability, must be ensured. High organizational interoperability means that information has been properly transmitted irrespective of the heterogeneity of devices, networks, types of compilation and modelling of information [38].

Organizational interoperability is concerned with the coordination and alignment of business processes and information architectures that span both intra- and inter-organizational boundaries. Coordination of business processes across organizational boundaries is essential if a single, aggregated view of a service from the customers' perspective is to be achieved. It is suggested that administrations could develop an exemplar scheme that would define standard approaches to each of the main requirements of any public service and use this exemplar to benchmark all other services; that common functionality could be provided on a shared basis through a broker service to reduce development, deployment and operational costs to the public administration and to each service fulfilment agency. Furthermore, it ensures consistency of experience for users of services across all agencies in the public sector through the use of agreed standards across all services; that expenditure reviews could be undertaken to ensure that financial priority is given to those schemes that comply with the structured customer support services set out above and with interoperability standards; and that each administration could develop a central programmed of organization development assistance and funding to bring this change about.

Table II provides a summary of the aforementioned interoperability levels analysis.

TABLE II. SUMMARY OF THE ANALYSIS OF THE INTEROPERABILITY LEVELS.

Interoperability Level	Source	Aim	Objects	Solutions	State of Knowledge
Technical	[20][21][22][23][24][25][26][27]	Technically secure data transfer	Signals	Protocols of data transfer	Almost developed
Syntactic	[28][29]	Processing of received data	Data	Standardized data exchange formats, e.g. XML	Almost developed
Semantic	[30][31][32][33][34][35][36][37]	Processing and interpretation of received data	Information/ Knowledge	Common directories, data keys, ontologies	Theoretically developed, but practical implementation problems
Organizational	[37][38]	Automatic linkage of processes among different systems	Processes (workflow)	Architectural models, standardized process elements	Conceptual clarity still lacking, vague concepts with large scope of interpretation

III. REVIEW OF IOT FRAMEWORKS ADDRESSING INTEROPERABILITY

This section presents our comprehensive review on existing IoT frameworks. The research was launched at a previous conference paper and is enriched with more information. The section concludes with a discussion on the technologies described and summarized their limitations and challenges.

A. Examined Solutions

A significant research effort has been devoted to providing solutions in the direction of increasing interoperability at all four levels presented in Section II. In this section, we examine solutions provided by eight related research efforts: BiG-IoT, INTER-IoT, VICINITY, AGILE, Open-IoT, Machine-to-Machine Measurement (M3) Framework, FIESTA IoT and SymbIoTe. These projects are developing interoperability solutions at different interoperability levels and for this purpose were chosen to be analyzed in this work.

1) BIG-IoT

BiG-IoT [38][39] focuses on addressing the semantic and organizational levels of IoT interoperability issues by creating the BiG-IoT API. It is about a generic web platform that unifies multiple platforms and different middleware. The Web API and semantic information representation models are defined in cooperation with the Web of Things Interest Group at W3C, expanding the standards of this community. The project has chosen schema.org as a basic vocabulary of concepts.

Through the API, which has a defined architecture, it is easier to create applications and services for heterogeneous platforms. To increase the level of interoperability at semantic, but especially at the organizational level the IoT API is framed by the following functions [40]:

- Identity management for registering resources.
- Discover resources according to user-defined search criteria.
- Access metadata, and data (download data as well as publish / record feeds).
- Vocabulary management for semantic descriptions of concepts.
- Security, including identity management, authorization and key management.
- Billing that allows you to make money through payment and billing mechanisms.

2) INTER-IoT

The INTER-IoT project aims to comprehensively address the lack of interoperability in the IoT realm by proposing a full-fledged approach facilitating "voluntary interoperability" at any level of IoT platforms and across any IoT application domain, thus guaranteeing a seamless integration of heterogeneous IoT technology [41].

INTER-IoT is based on the following main functionalities to address technical and syntactic interoperability:

- Methods and tools for providing interoperability among and across each layer of IoT platforms.
- A global framework called INTER-FW for programming and managing interoperable IoT platforms, including INTER-API and several interoperability tools for every layer.
- Engineering Methodology based on the CASE tool for IoT platforms integration/interconnection.

Three main types of interoperability (i.e., technical, syntactic and semantic) are enabled by INTER-IoT [24][42]. Universal syntactic and semantic interoperability among any platform with different data formats and ontologies is possible through the INTER-IoT DS2DS (Data & Semantics-to-Data & Semantics) solution. Moreover, other INTER-IoT layers like D2D (Device-to-Device) and N2N (Networking-to-Networking), can provide organizational interoperability among smart elements, enabling connectivity to the network.

3) VICINITY

The VICINITY project aims at interfacing cloud-based platforms from various application domains by providing "interoperability as a service" for the IoT [43]. The proposed interoperable platform is presented as a virtual neighborhood, a “social network” where users can share access to their smart objects without losing control. The project team has thoroughly reviewed all existing standards and platforms, selecting those needed to build a service or increase interoperability.

The project is not so concerned with technical interoperability. For communication between devices, wireless networks like WiFi and ZigBee are mainly used. Main goal of the VICINITY project is to increase semantic interoperability. Using the standard W3C Web Language Ontology, specific ontologies are developed in a variety of areas, such as ontologies for energy and building, extending the SAREF reference ontology [44] interoperability.

The VICINITY ontology network is composed of cross-domain ontologies, addressing the modelling of general concepts like time, space, Web of Things. It will represent the information for exchanging IoT descriptor data between peers. Domain-oriented ontologies aim to cover vertical domains, such as Health, Transport, Buildings, etc.

4) AGILE

The AGILE project builds a modular open-source interoperable Gateway solution (hardware and software gateway) for the IoT focusing on the physical, network communication, processing, storage, and application layers [24][45]. The AGILE software modules are addressing

functions, such as device management, communication networks like area and sensor networks and solutions for distributed storage. Moreover, the AGILE approach includes security features that allow users to share data in a trusted way.

The AGILE project focuses on technical interoperability both at hardware and software levels. Within the project, various popular and low-cost technologies, such as Raspberry Pi are being developed and expanded. This creates the "Gateway Maker", a proposal to create interoperable gateways that will be used for multi-purpose and heterogeneous purposes. At the same time, the project provides open-source code and a web-based environment (Node-Red) for developers to develop new, innovative applications. The project does not address any approach to the semantic and organizational level of interoperability. The architecture comprises four layered domains.

5) Open-IoT

Open-IoT focuses on increasing semantic interoperability [46][47]. In the framework of the project, a middleware platform was created that allows semantic integration of applications on the cloud. For information modelling, the ontology of W3C sensor networks (SSN) are used as a common standard for the semantic integration of various IoT systems. Appropriate infrastructures collect and semantically comment on the data of the different sensors. Also, another semantic technique called Linked Data is used to enrich the data and interface it.

Open-IoT innovates with other programs as it implements a platform with modules for collecting data and applications in cloud computing infrastructures, modules for creating semantically interoperable applications, and applications for mobile sensors. The implementation of semantic techniques in the cloud is something that adds value to the project and makes it stand out from other similar solutions. These functionalities provide a basis for the development of novel applications in the areas of smart cities and mobile crowdsensing, while also enabling large scale IoT experimentation and increase the level of organizational interoperability. The project does not address any approach to the technical and syntactic level of interoperability.

6) Machine-to-Machine Measurement (M3) Framework

The M3 Framework project focuses on addressing the lack of semantic interoperability in IoT. The framework of the project assists the developers in semantically annotating M2M data and in building innovative applications by reasoning on M2M data originating from heterogeneous IoT systems and domains. To increase the level of interoperability at syntactic, but especially at the semantic level the M3 Framework is framed by the following layers [48][49]:

- Perception layer, which consists of physical IoT devices, such as sensors, actuators and RFID tags.
- Data acquisition layer, which focuses on collecting raw data from IoT devices/sensors and converting them in a unified way, such as RDF/XML compliant with the M3 ontology.
- Persistence layer, which takes over to store M3 in a database to store semantic sensor data which is called the triple store.
- Knowledge management layer, which is responsible for finding, indexing, designing, reusing and combining domain-specific knowledge, such as ontologies and datasets to update M3 domain ontologies, datasets and rules.
- Reasoning layer, which infers new knowledge using reasoning engines and M3 rules extracted from Sensor-based Linked Open Rules (S-LOR) [49].
- Knowledge query layer executes SPARQL (an SQL-like language) queries on inferred sensor data.
- Application layer, which employs an application (running on smart devices) to parse and display the results to end-users.

7) FIESTA IoT

The FIESTA-IoT project is a Research and Innovation Action under the European Horizon 2020 Programme addressing the topic 'Future Internet Research and Experimentation'. The project focuses on large-scale experiments in the IoT domain that will utilize data and resources from heterogeneous IoT platforms [50]. These experiments provide a variety of tools and good practices to increase the interoperability of IoT heterogeneous platforms. FIESTA project promotes researchers and experimenters to share and reuse data from diverse IoT testbeds using semantic technologies seamlessly and flexibly.

The FIESTA-IoT architecture is a set of functional blocks allowing [51]:

- Testbed data streams and resources to be plugged into FIESTA-IoT; be discoverable using FIESTA-IoT and be accessible via FIESTA-IoT services.
- Semantic querying of both linked data sets (of collected testbed data) and IoT service APIs.
- Secure access to testbed resources by authenticated and authorized experimenters.

8) SymbIoTe

The SymbIoTe project (symbiosis of smart objects across IoT environments) focuses on the implementation of a flexible and secure interoperability middleware across IoT platforms. The main goal of the project is to create IoT applications on IoT platforms as well as dynamic and adaptive smart spaces that they can collaborate [51][52]. This is accomplished by:

- A semantic IoT search engine for connected (virtualized) smart objects (i.e., IoT resources) registered by platform providers;
- An abstraction layer for unified and secure usage of those resources across platforms;
- High-level, domain-specific APIs ("Enablers") for rapid cross-platform application development;
- IoT platform federations, i.e., associations between two platforms facilitating their secure interaction, collaboration and bartering of resources;
- Dynamic and self-configurable smart spaces offering interoperability for collocated devices and gateways;
- A secure interworking protocol between the IoT platforms, gateways and smart devices.

The SymbIoTe is built around the concept of virtual IoT environments provisioned over various cloud based IoT platforms. Virtual IoT environments are an abstraction composed of virtual representations of actual sensors and actuators being exposed by their host platforms to third parties. The symbIoTe framework is built around a hierarchical IoT stack and spans over different IoT platforms. Smart objects are expected to be connected to IoT gateways within the smart spaces which also host various computing and storage resources. The local infrastructure shares the available local resources (connectivity, computing and storage) and is connected to platform services running in the cloud. The architecture comprises four layered domains.

B. Discussion

The existing solutions are dealing with the heterogeneity of devices, data and services. Some of them integrate semantic web technologies to enhance interoperability [41][42][46][47][48][49]. The absence of standardized activities, life cycles and methodologies as well as a set of techniques and tools hinder an interoperable IoT. To all existing solutions interoperability challenges remain still present. For instance, they neither use the same model to structure the data produced by objects/things nor the same reasoning approach to deduce new knowledge from data produced by objects/things. To assess the degree of interoperability maturity and answer research question RQ1, Table III summarizes the results of the state-of-art IoT frameworks that were analyzed in this review.

At technical and syntactic level AGILE, VICINITY and INTER-IoT attempt to provide solutions by creating Generic Gateways and device-to-device modules that integrate several wireless and wired technologies. All of these need to be incorporated into supported technologies like families of Low Power and Wide Area wireless networks (LoRaWan, SigFox, etc.), as well as other short-range wireless indoor technologies, such as Beacons.

A recurring aspect is that most efforts are focused on addressing the semantic interoperability challenge. The VICINITY platform uses the standard W3C Web Language Ontology and implements cross-domain ontologies, whereas Open-IoT extends SSN ontology, and uses semantic tools such as Linked Data. BiG-IoT expands the standards of WoT and uses vocabulary management for handling semantics tools. Moreover, INTER-IoT increases semantic interoperability compared to the rest of the platforms by introducing different data formats and ontologies through the INTER-IoT DS2DS solution. In addition, the M3 Framework project addresses the semantic interoperability by the use of innovative semantic tools, such as M3 ontology tools, reasoning engines and M3 rules extracted from S-LOR. In addition, FIESTA-IoT project provides a blueprint of experimental infrastructure, software tools, semantic techniques, certification processes and best practices enabling IoT testbeds/platforms to interconnect their facility resources in an interoperable semantic way. Finally, symbIoTe, support fair and trustworthy interactions between platforms without a centralized mediator, so that IoT platform owners can engage in direct partnering relationships by use of symbIoTe platform federations.

At organizational level, BiG-IoT creates a common and generic API (Application Programming Interface) between the different IoT middleware platforms. Open-IoT implements a cloud-based middleware platform with innovative tools and functionalities. Also, VICINITY project creates a framework that follows the philosophy of interoperability as a service for "IoT Neighborhood" with many modules and tools. Moreover, the INTER-IoT platform increases the levels of organization interoperability with INTER-API, which includes several interoperability tools for every layer. Moreover, M3 Framework project with innovative semantic engines and solutions at the application layer, which parses and displays the results to end-users, increases the organizational interoperability level. Furthermore, FIESTA-IoT enables execution of experiments across multiple IoT testbeds, based on a single API for submitting the experiment and a single set of credentials for the researcher and the portability of IoT experiments. The focus is on resource sharing in the form of mutual registration, resource announcement, and subscriptions to information about resources offered by different platforms. However, features for the management of platform federations and collaboration mechanisms for fair and social interactions are not defined in most of the projects. Only VICINITY, and SymbIote have moved clearly in the philosophy of collaborative and open IoT Environments. Thus, by adopting this approach, organizational interoperability is increased, which, as we have argued, is not largely addressed by existing solutions. However, the tools that they proposed are still at an early stage and need to be evaluated in the future.

TABLE III. INTEROPERABILITY LEVELS COVERAGE BY THE EXAMINED IOT FRAMEWORKS.

	SOURCE	Technical level	Syntactic level	Semantic level	Organizational level
AGILE	[24][45]	Yes (Makers Gateway)	Yes (Makers Gateway)	No	No
Open-IoT	[46][47]	No	No	(Extend SSN ontology, Linked Data)	(Extend SSN ontology, Linked Data)
VICINITY	[43]	Yes (Generic Gateway supports common networks: Wifi, ZigBee)	Yes (OWL Language)	Yes (VICINI-TY Ontologies)	Yes (Interoperability as a service)
BiG-IoT	[38][39][40]	No	No	Yes (Expand the standards of WoT, vocabulary management for handling semantics)	Yes (BiG-IoT API)
INTER-IoT	[24][41][42]	Yes (DS2DS)	Yes (DS2DS)	Yes (DS2DS)	Yes (INTER-API)
Machine to Machine (M3) Framework	[48][49]	No	Yes (Data acquisition layer)	Yes (Knowledge management layer, Reasoning layer)	Yes (Application layer)
FIESTA IoT Project	[51]	Yes (Increase interoperability among platforms)	No	Yes (Reasoning and Linking technics)	Yes (FIESTA API, Middleware-Application layer)
SymbIoTe	[51][52]	Yes (Interworking protocol between the IoT platforms, gateways and smart devices)	Yes (Interoperable language)	Yes (A semantic IoT search engine)	Yes (IoT platform federation)

To resolve research question RQ2, we summarize in Table IV the shortcomings of the examined IoT Frameworks, by classifying them based on the interoperability level. At the technical interoperability level, a typical drawback of many frameworks, is the lack of focus on common communication standards between devices and systems. Furthermore, in several architectures it is imperative to implement interoperable IoT gateways, where raw data will be collected from different heterogeneous sensors supporting open source, and messaging systems. Moreover, at the level of syntactic interoperability, a common gap identified between these frameworks, is the lack of syntactic translation tools that convert the heterogeneous data in a unified way, such as RDF, XML and JSON.

At the semantic level, the ontologies that are created in most of the IoT frameworks are complicated and are not interoperable with each other and focus mainly on the interoperability regarding specific fields rather than on a general solution. Besides that, tools for ontology alignment and ontology merging have not been particularly emphasized on solutions that can radically improve interoperability levels. Certain future research should focus on this direction so that future ontology engineers are given powerful and "lightweight" tools, such as ontology alignment tools for low-power devices, tools to implement "lightweight" ontologies for cross-domains, and semantic reasoning tools.

At the organizational interoperability level, there is a lack of IoT platform federations, i.e., associations between more than two platforms facilitating their secure interaction, collaboration and bartering of resources. Moreover, collaboration and social interaction mechanisms that provide open and cooperative IoT systems have not been particularly emphasized. It is considered necessary to create tools that will manage the collaborations between IoT devices and systems, as well as manage the social relationships between IoT devices, with the aid of semantic techniques. Consequently, supporting collaboration and social interaction mechanisms between IoT systems will improve the organizational interoperability (research question RQ4).

TABLE IV. SHORTCOMINGS OF THE EXAMINED IOT FRAMEWORKS.

Technical level	Syntactic level	Semantic level	Organizational level
• Incompatibility of different versions. • Different communication protocols or formats (IEEE 802.11, IEEE 802.15, LoRaWan, SigFox). • Lack of a common standard of communication between devices and systems. • Lack of interoperable IoT Gateways.	• Not well-defined syntactic metadata schema and their mapping mechanisms. • Lack of syntactic translation tools that convert the heterogeneous data in a unified way, such as RDF, XML and JSON. • Solutions include the messaging protocols CoAP, XMPP, AMQP, MQTT offer cross-domain compatibility. • Lack of a common syntactic format identification, registration and management mechanisms.	• Semantically incompatible information models (incompatible general ontologies) • Lack of common standards • Lack of 'lightweight' semantic tools. (Ontology alignment, ontology matching, reasoning), and lightweight interoperable ontologies. • Incompatible reasoning approaches to deduce new knowledge from data produced by objects/things.	• Lack of Collaboration Mechanisms. • Lack of collaboration management methods. • IoT platform federations, i.e., associations between more than two platforms facilitating their secure interaction, collaboration and bartering of resources.

IV. TOWARDS AN INTEROPERABLE AND COLLABORATIVE IOT FRAMEWORK

In this section, we report design requirements and open research challenges that our review of the existing frameworks has highlighted. Moreover, we present a high-level design description of the proposed IoT framework (the SSNT framework). Although the evaluation of this framework is not covered in this paper, a proof-of-concept implementation scenario is provided.

A. Requirements and Open Research challenges

Existing IoT Frameworks have the potential to provide numerous solutions for improving multi-level interoperability, but many challenges have not yet been fully addressed and require collaboration from standardization committees, hardware manufacturers, software developers and IoT stakeholders. This section discusses several challenges related to multilevel interoperability in the context of IoT. Moreover, our review on the existing frameworks answers research questions RQ2, and RQ3, and suggests that a novel IoT framework needs to support specific functional features, as the ones outlined in the following paragraphs.

1) IoT Resource Management

IoT Systems collect data from different distributed sensors. These data are multimodal, including heterogeneous data, such as video streams, images, audio, and simple text [2]. How to integrate these distributed data from multiple sources is a key challenge for IoT development and for the implementation of new innovative smart applications.

Moreover, communication between heterogeneous devices generates a large volume of real-time, high-speed, and uninterrupted data streams. These data streams include structured, semi-structured and unstructured data. When heterogeneous and various sensor data are acquired, multisource data should be merged to create a comprehensive and meaningful view for further utility [53].

2) Lightweight Semantic Tools

As mentioned in the previous section, ensuring semantic interoperability is very important to address the inability to exchange and reuse data. Unfortunately, even today, IoT systems consist of semantically incompatible information models, such as incompatible general ontologies that offer different descriptions or even understandings of resources and processes, and thus are a barrier to the development and adoption of the IoT.

Most of the existing semantic tools and techniques, such as Linked Data, ontology alignment and ontology matching [54][55] have been created primarily for Internet resources. Existing models provide the basic description frameworks, but alignment between different models and frameworks are required. In addition, the capacity of the natural environment and the resource constraints on IoT systems have not been taken into account [56]. Future work in this area should provide lightweight semantic tools that are easily adapted to environments with limited and distributed resources.

3) Standardization

In the new world of IoT, standards will be more important due to the greater interoperability demands. As more systems, devices, systems and platforms are connected we will see that this is only possible if all agree on common standards [29][40][57].

Firstly, one standard has no direct control over other standards, which means that changes for one standard will not automatically be propagated to other standards. Secondly, in order to support interoperability among several standards, a large number of adapters have to be developed, which is clearly inefficient. There are distinct missing standardization activities related to data models, ontologies, and data formats

to be used in IoT applications for service-level interfaces and protocols. Machine-to-Machine Measurement (M3) framework [48][49] is offered to supplement existing semantic standards by adding common format, nomenclature and methods for data interpretation. A semantic approach is aimed at resolving the issue of lack of standardization by introducing common ontologies, data models, and vocabularies; however, currently the application methods are non-unified, complicated, and require further improvement. So, a novel IoT framework should be based on common standards only and refrain from developing its own proprietary solutions.

4) Scalability

The exponential growth of connected objects to the Internet produces a massive quantity of data called" Big data". According to [58][59][60], the big data generated by IoT has different characteristics like large-scale data, heterogeneity, strong time and space correlation. Therefore, the main challenges encountered during the development of IoT applications/systems are the semantic IoT event processing, real-time processing of data streams and reasoning in a complex and dynamic context (spatiotemporal reasoning) in a scalable and secure way, etc. Consequently, these new requirements drive the need for the deployment of a scalable IoT system. Thereby, applying Semantic Web technologies (SPIN rules, SWRL, SPARQL, DL safe rules, RIF, etc.) to the IoT domain faces a new challenge on how to manage and interpret such heterogeneous data during a limited period in a scalable way.

5) Collaboration Mechanisms

Providing collaborative smart objects with interpretation and analytics methods to process and evaluate events in their surroundings is important for building new IoT-based applications [61]. Semantic descriptions serve the purpose of transforming large amounts of observed and perceived data created by users and things/objects into high-level concepts that are meaningful for establishing automated decision-making processes. However, the non-human perception contributes to existing pool of challenges in IoT. Similar to problems faced by the artificial intelligence research community, in IoT the challenges are data integration and amalgamation from different sources, rules of data aggregation, defining borders and thresholds, as well as describing events, actors and objects. Solutions are needed to integrate data from various environments, and patterns for further fusion of new knowledge based on learnt rules. So, a novel IoT Framework must have innovative mechanisms of cooperation between IoT devices and systems, not only to connect and interact, but also to socialize and collaborate with each other to achieve some specific task(s). In this way the organizational interoperability will be increased, an element that is missing from the IoT framework so far. This kind of social interaction requires cooperation among IoT devices.

B. SSNT Framework

To address the multifaceted problem of interoperability, and partially answer the research question RQ4, equal emphasis should be placed on all levels of interoperability as they have been presented in this work. It is necessary to create tools and software modules that will seamlessly confront the interoperability problem targeting all levels, and also provide solutions that are available for devices with constrained resources. In this vision, an indispensable, interoperable, global IoT ecosystem can be created in the form of an SSNT. Taking under consideration the open issues and shortcomings of the state-of-art frameworks, as discussed previously, an SSNT framework is proposed that consists of modules and tools to overcome interoperability issues.

Firstly, at the level of technical interoperability, new data collection and raw data filtering tools should be added to the system, so that data transferred to the cloud can be edited with edge computing techniques. Additionally, these new technologies should be also compatible with the new wireless technologies of the LPWAN family (LoRaWan, SigFox, NB-IoT). Following, at the level of syntactic and semantic interoperability, the SSNT architecture should include new tools creating interoperable ontologies that will extend the existing solutions. Initially, it is necessary to create an interoperable middleware framework with new semantic modules, through which heterogeneous devices will be interconnected. Moreover, with the successful implementation and development of the SSNT framework through which heterogeneous devices and systems can communicate seamlessly, many innovative applications could be spawned in various fields leveraging on the raw data collected. Consequently, the level of organizational interoperability will increase rapidly. For example, platforms can be enabled to perform collaborative sensing/actuation tasks to complement each other's infrastructure, and to interact directly in a decentralized way without exposing their business relationship to a centralized authority. Reasons for such a collaboration can vary e.g., similar IoT platforms that operate in different locations can federate to offer seamlessly to their clients IoT services in other locations, or collocated platforms can benefit from each other by forming partnerships to offer cross-domain solutions.

The SSNT architecture, as shown in Figure 3, is structured on four layers: Perception, Transmission, Middleware and Application.

The *Perception Layer* contains all the IoT heterogeneous physical devices, such as Beacon sensors, ZigBee sensors, LoraWan sensors, actuators, etc. from which all heterogeneous data are derived.

The *Transmission Layer* includes the following modules:

1. ***SSNT Data Acquisition***, which gets data from different types of sensor devices.

This module is responsible for the collection and filtering of raw IoT data from various heterogeneous IoT devices with IoT Gateways. It consists of two components:

- *Data Collection*: Obtains raw data from various heterogeneous sensors using interoperable architectures that support distributed, open-source and messaging systems (Apache Kafka, ThingsSpeak etc.). This section supports different data sources and executes multiple processes at the same time.
- *Data Filtering*: Verifies the field of data collected from the previous section. Filtering requires a database search and applies filtering rules. With this function, the "bad" values are discarded minimizing storage costs and ensuring fast data transmission.

2. ***SSNT Data Integration,*** which converts the heterogeneous data in a unified way, such as RDF, XML and JSON. It consists of four components:

- *Metadata Creation*: Some important metadata objects are obtained, like data type, measuring units, time stamp, and geolocation. This module also describes the specific industrial environment, data, and applications.
- *Communication Interface*: Communication between each module of the data collection component is organized. Various types of data are translated into a single format so that the system can understand. For example, the data coming from various devices with different formats are translated into JSON message structure first and then sent to the next phase for data aggregation.
- *Data Aggregation*: The pre-processed data is transmitted to the aggregation component for further summarization. The aggregated data is more significant than the raw data collected by factory devices. The data stream coming from the physical layer is separated into data summarization modules as described below.
- *Data Summarization*: The datasets of various devices are represented into groups according to time-period. It reduces computational and storage cost and improves consultation performance by minimizing the volume of data. So, the event table generated by the data collection.

The SSNT *Middleware Layer* contains components and functionalities that can be divided into several functional modules as follows:

1. ***Data Storage***, which contains a) tools for storing semantic IoT data to a cloud database and to NoSQL databases such as GraphDB, Cassandra; b) functionalities for querying and searching in a different kind of databases.

2. ***Lightweight Ontology Creator/Annotator***, which contains:

- Tools for designing interoperable "lightweight" ontologies and semantic structures, according to standard ontologies that can be interpreted, shared and reused by other ontologies
- Methods to change an isolated ontology to a reusable and interoperable ontology (such as IoT-Lite, SSN ontology)
- Methods to enrich metadata and create reusable data, to enable semantic interaction and interoperability between the various heterogeneous "things", offering a significant advantage compared to existing syntactic interactions.

3. ***Connector***, which provides Open Linked Data interfaces e.g., SPARQL (SPARQL Protocol and RDF Query Language) over ontologies for internet-connected objects within the physical world abstracted by the middleware to interact with an SSNT.

4. ***Reasoner***, which includes tools and components for the automated data configuration filtering, fusion and reasoning mechanisms, which obtain higher-level actionable knowledge from low-level sensor data

5. ***Ontology Alignment for Resource-Constrained Devices***, which includes tools for ontology merging, matching, and alignment related to the dynamics and complexity of the IoT systems.

6. Social Collaboration Generator / Manager

This component is responsible for building and managing social relationships between various heterogeneous IoT devices. The social relations that will be created at each level will improve the various issues of interoperability as mandated by the research question RQ4. It consists of tools for automatically building relationships between things, and methods to manage SSNT relationships. These tools integrate information into IoT devices so that they can make "friends", start a relationship, update a situation, and terminate a relationship. It is our proposed approach in the context of the answer to the 4 questions. The social relations that will be created at each level will improve the various issues of interoperability.

These relationships between IoT devices are classified according to the level of interoperability that they are addressed as follows:

i. ***Relationships between things in the level of device interoperability.***
These relationships take place for example between IoT devices that are on a different IoT network but are close together and can work together to achieve a common goal.

ii. ***Relationships between things in semantic and syntactic interoperability levels.***
These relationships are made at the level of semantic or syntactic interoperability and relate to IoT devices that represent data with common

vocabularies, ontologies (shared ontologies) or in a different way (different ontologies).

iii. ***Relationships between things in the organizational interoperability level.***
These relationships are made at the level of organizational interoperability between IoT devices belonging to different IoT platforms of organizations. Relationships between platforms can be enabled to perform collaborative sensing/actuation tasks.

Finally, the *Application Layer* leverages on the solutions provided by the underlined layers to accomplish disparate applications of IoT devices. The Application Layer is a user-centric layer which executes various tasks for the users. It represents innovative smart applications in various fields, such as smart homes, smart cities, smart healthcare, smart agriculture, smart buildings, etc. The provision of end user tools that enable people to engage in the formation of such applications by affording high level metaphors are also important [62].

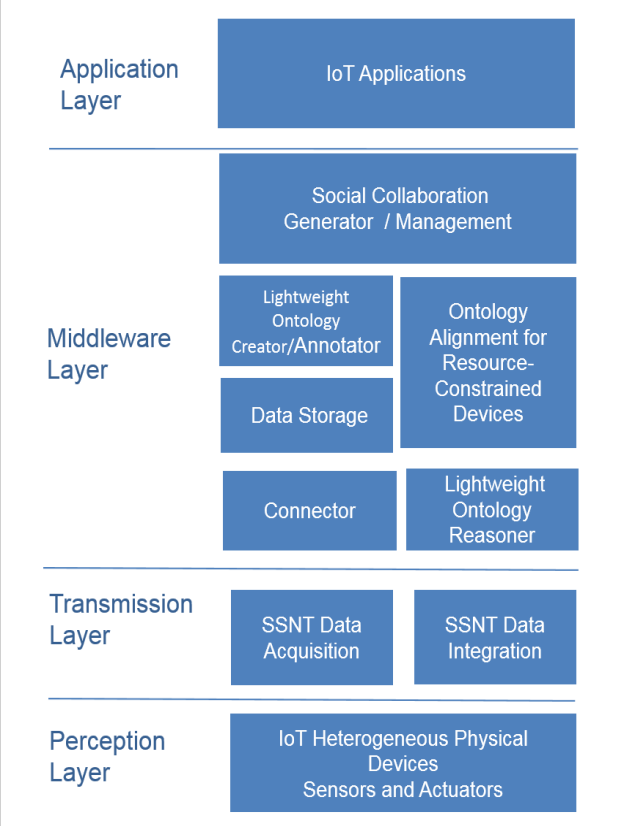


Figure 3. SSNT architecture overview.

A simple application scenario is given to illustrate part of the SSNT framework functionality. A smart lightweight application is designed as a result of the collaboration of different smart objects. In this case the SSNT consists of a smart desk, a smart chair, a smart book and a smart lamp. The application logic is that when the chair is occupied and is nearby the desk and the book is open above the desk the application infers that a study activity takes place and as service the application regulates the light depending on the brightness sensed on the book. Each smart object is described by properties in the form of an ontology (Figure 4, Figure 5, Figure 6, and Figure 7). Such ontologies may be independently developed and thus can be heterogeneous. The semantic interoperability support of the SNNT framework through ontology alignment may be required in this case to deduce the use of similar terms or structures between the ontologies.

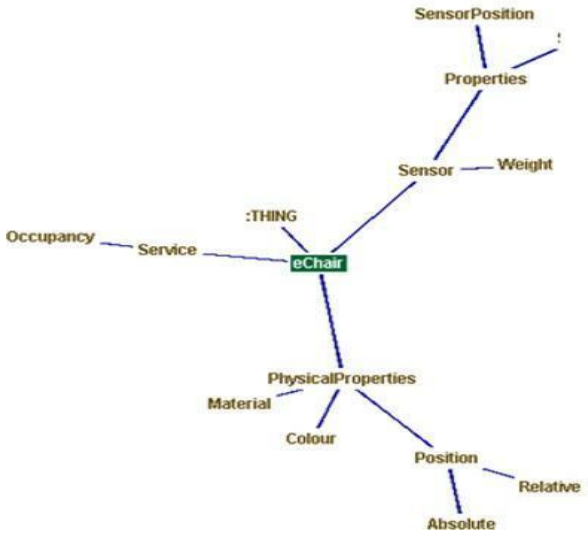


Figure 4. eChair Ontology.

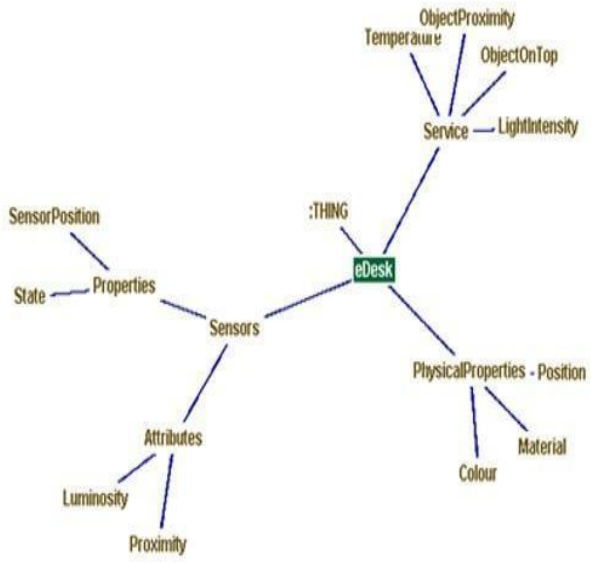


Figure 5. eDesk Ontology.

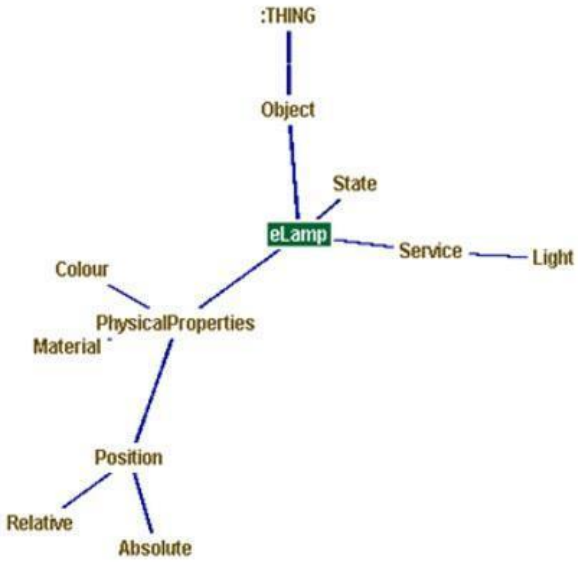


Figure 6. eLamp Ontology.

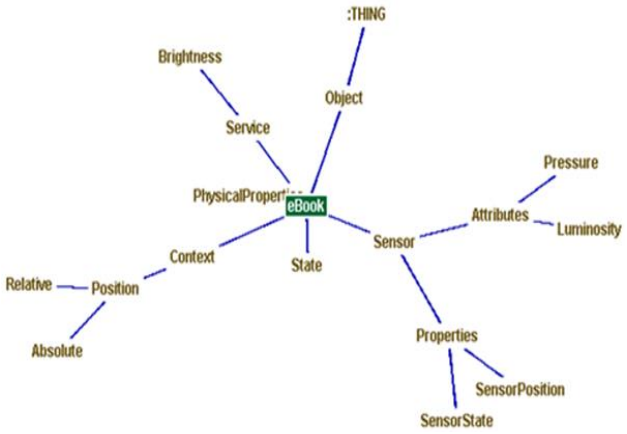


Figure 7. eBook Ontology.

In this context, an automatic ontology alignment module may apply linguistic and graph matching techniques [29]. Figure 8 shows an example of a linguistic similarity between two given ontologies based on descriptive information, like the property names. These similarities form the linking basis between smart object ontologies and provide the support mechanism to answer service discovery requests for a specific functionality that is required to instantiate an application. Similar questions may involve, for example, looking up a device that provides a light service or whether an IoT entity is of type desk. Such questions can be answered via the ontologies detailing semantically the smart objects and their alignments.

	Ontology 1	Ontology2	Similarity	Relation
0	SensorPosition	Position	0.715341	"="
1	hasProperties	hasPhysicalProperties	0.737579	"="

Figure 8. Example of similarities between two smart objects ontologies.

More rich knowledge can be acquired when individual ontologies are merged. Figure 9 illustrates the result of the merged ontology acquired using the ontologies of the smart objects involved in the smart light application. This merged ontology reflects the interconnected entities and can be used to infer knowledge regarding the collective behavior which can appear from the collaboration of the smart object services. Consequently, composite questions can be answered like whether a specific IoT environment is suitable for fulfilling the requirements of the smart light application.

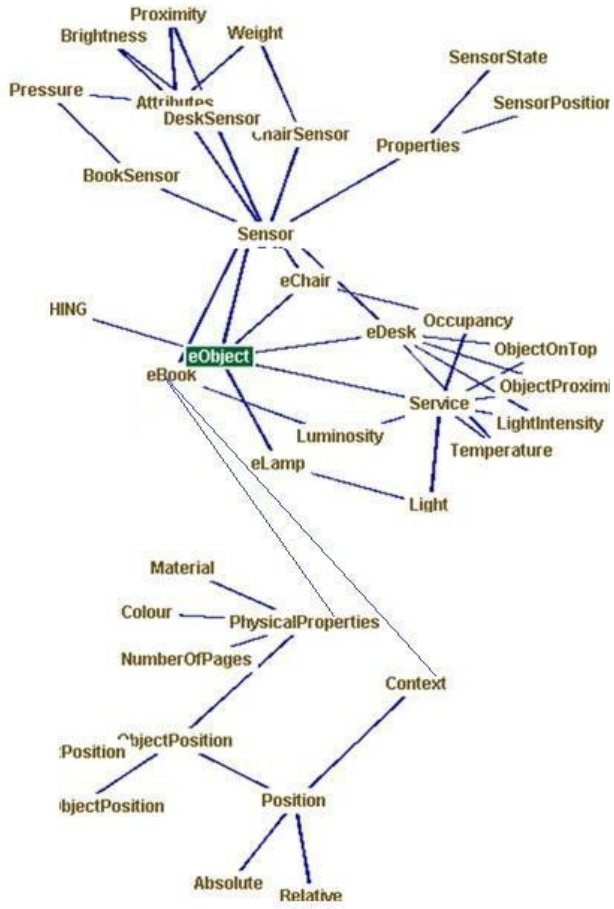


Figure 9. Merged ontology for the smart light application.

Figure 10 shows our future proof-of-concept implementation of SSNT in the domain of smart agriculture. The raw data streams will be collected by IoT sensors as they will be enriched with semantic annotation and will be modeled in ontologies with SSNT framework tools. Then, with semantic reasoning rules, social semantic groups will be created between the semantic data that aim to achieve a common goal, such as Greenhouse automation, crop management, and Monitoring of climate conditions. Finally, with SSNT semantic tools such as SPARQL queries, ontology alignment module etc., and new knowledge will be produced, and new services and applications will be created.

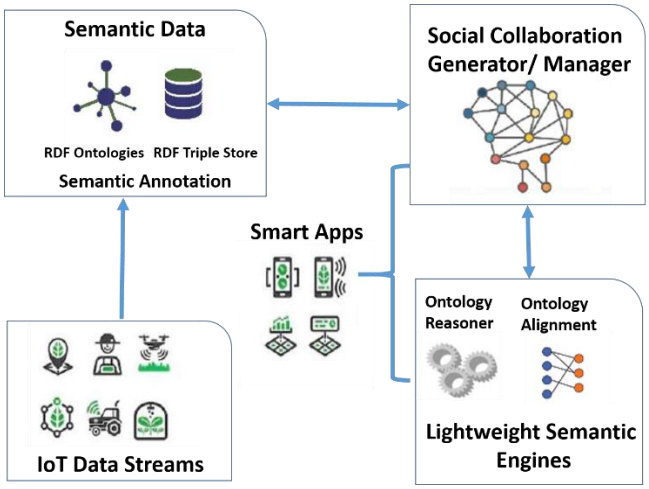


Figure 10. SSNT Framework in Smart Agriculture.

The SSNT framework has the capabilities of combining and analyzing data streams helping the farmers or agronomists in more informed decision-making in near real-time and fast reaction to changes and unpredictable events. For example, by automatically integrating sensory data about soil humidity with web services for weather forecasting, better decisions could be made about more precise irrigation and fertilization of the crops.

A basic application scenario of a smart crop management system is given to illustrate part of the SSNT framework functionality. We assume that an SSNT is deployed in a rural area. The system consists of various sensors (temperature, humidity, and thermal cameras), web services for weather forecasting, and actuators that help in the smart management of crops. Initially, heterogeneous IoT data streams are collected via the SSNT Data Acquisition module. Then, with the SSNT Data Integration module, the heterogeneous data are converted in RDF format. In this way, the raw IoT data modelled in interoperable ontologies that will be created with the Lightweight Ontology Creator/Annotator module. Furthermore, Data Storage module, storing semantic IoT data to a cloud database such as GraphDB, which is an enterprise-ready Semantic Graph Database, compliant with W3C standards.

Furthermore, with the application of semantic techniques using the appropriate tools of the SSNT framework, it is possible to create social groups of common interest which will be responsible for achieving a specific user goal. For the needs of the scenario, let us assume that two of the goals of the crop management system are: to increase fertility and to predict crop disease. After the goals are set by the user, through the Social Collaboration Generator / Manager Module, two social smart objects groups of interest will be created. In these groups, social relationships between things are created at the level of device interoperability, as well as relationships between things at semantic and syntactic interoperability levels. The first group will consist of soil humidity sensors, temperature sensors, data from web services for weather forecasting, and actuators such as solenoid valves. This group of smart objects will aim to collaboratively increase soil fertility. Through semantic functionality (SPARQL queries, reasoning rules), the semantically annotated data will feed special agricultural applications that will achieve the goal of increasing crops fertility. The second group of smart objects will consist of a thermal camera, and leaf wetness sensor. In the same way, the goal of disease control of cultivated plants will be pursued.

V. CONCLUSION

In contrast to other surveys of IoT research, this review study focuses on interoperability achieved by approaches in a multilevel perspective. Contemporary IoT frameworks have been systematically researched and their capability to achieve multi-layer interoperability between applications, services, and software platforms has been reported. Different solutions addressing interoperability issues at discrete levels have been studied, analyzed and compared to identify their limitations, such as lack of semantic lightweight tools, poor scalability and lack of collaboration mechanisms, while open issues and challenges were also identified. These limitations provide research opportunities and have motivated the Semantic Social Network of Things (SSNT) framework design. In this context, the concept of SSNT has been introduced for specifying device-to-device collaborative services based on the social interaction between smart objects while supporting interoperability at different levels and taking into account the limitations of IoT systems. Furthermore, a proof-of-concept application in the smart agriculture domain has been discussed to demonstrate important features of the presented approach.

Future activities will focus on implementing, deploying and evaluating the modules of the SSNT framework in real IoT environments. For instance, Generator / Manager social collaboration software will be evaluated in the agricultural domain where many heterogeneous IoT devices can be found. Software libraries and APIs related to the semantic data management (e.g., Jena, http:// https://jena.apache.org/), and open source IoT frameworks (e.g., openIoT framework, http://www.openiot.eu/), will be used to implement the proof-of-concept system of SSNT. Our future work aims also to address limitations of existing solutions such as the lack of lightweight semantic tools and the lack of tools for evaluating collaboration and social interaction mechanisms in order to assess how effectively such mechanisms can address multi-level interoperability issues in open IoT environments.

REFERENCES

[1] A. Pliatsios, C. Goumopoulos, K. Kotis, “Interoperability in IoT: A Vital Key Factor to Create the Social Network of Things,” The Thirteenth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies UBICOMM 2019, pp. 63-69, Porto, Portugal, 2019.

[2] L. Atzori, A. Iera, and G. Morabito, “The Internet of Things: A Survey,” Comput. Networks, pp. 2787–2805, 2010.

[3] F. Shi, Q. Li, T. Zhu, and H. Ning, "A survey of data semantization in internet of Things," Sensors, vol. 18, no 1, pp. 313, 2018.

[4] I. Lee, and K. Lee, "The Internet of Things (IoT): Applications, investments, and challenges for enterprises," Business Horizons, vol. 58, no 4, pp. 431-440, 2015.

[5] K. N. Kumar, V. R. Kumar and K. Raghuveer, "A Survey on Semantic Web Technologies for the Internet of Things," 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC), Mysore, pp. 316-322, 2017.

[6] B. Kitchenham, "Procedures for performing systematic reviews," Keele, UK, Keele University, vol. 33, pp. 1-26, 2004.

[7] A. Gyrard and M. Serrano, "A Unified Semantic Engine for Internet of Things and Smart Cities: From Sensor Data to End-Users Applications," 2015 IEEE International Conference on Data Science and Data Intensive Systems, Sydney, NSW, pp. 718-725, 2015.

[8] D. Guinard, V. Trifa, F. Mattern, and E. Wilde, "From the internet of things to the web of things: Resource-oriented architecture and best practices", in Architecting the Internet of things, Springer, Berlin, Heidelberg,pp. 97-129, 2011.

[9] A. Gyrard, P. Patel, S.K. Datta, and M. I. Ali, "Semantic web meets the Internet of Things and Web of Things," in Proceedings of the 26th International Conference on World Wide Web Companion, pp. 917-920, 2011.

[10] Cisco [online], "The Internet of Everything" Available from: www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoE_Economy_FAQ.pdf. 19.2.2020.

[11] B. Afzal, M. Umair, G. A. Shah, and E. Ahmed, "Enabling IoT platforms for social IoT applications: vision, feature mapping, and challenges," Future Generation Computer Systems, vol. 92, pp. 718-731, 2019.

[12] A.M. Ortiz, D. Hussein, S. Park, S.N. Han, and N. Crespi, "The cluster between internet of Things and social networks: Review and research challenges," in IEEE Internet of Things Journal, vol. no. 3, pp. 206-215, 2014.

[13] Y. Saleem, N. Crespi, M.H. Rehmani, R. Copeland, D. Hussein, and E. Bertin, "Exploitation of social IoT for recommendation services," In 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT), pp. 359-364, 2016.

[14] L. Atzori, A. Luigi, et al., "The social internet of Things (siot)–when social networks meet the internet of Things: Concept, architecture and network characterization," in Computer networks vol. 56, no 16, pp. 3594-3608, 2016.

[15] L. Atzori, A. Iera, and G. Morabito, "Siot: Giving a social structure to the internet of Things," IEEE communications letters, vol. 15, no 11, pp. 1193-1195, 2011.

[16] J. Radatz, A. Geraci, and F. Katki, "IEEE standard glossary of software engineering terminology," IEEE Std, vol. 610121990, no. 121990, pp. 3, 1990.

[17] A. Tolk, and J. A. Muguira, "The levels of conceptual interoperability model," In Proceedings of the 2003 fall simulation interoperability workshop, Citeseer, vol. 7, pp. 1-11, 2003.

[18] M. Noura, M. Atiquzzaman, and M. Gaedke, "Interoperability in internet of things: Taxonomies and open challenges," Mobile Networks and Applications, vol. 24, no. 3, pp. 796-809, 2019.

[19] A. Glória, F. Cercas and N. Souto, "Comparison of communication protocols for low cost Internet of Things devices," 2017 South Eastern European Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM), Kastoria, pp. 1-6, 2017.

[20] K. Rose, S. Eldridge, and L. Chapin, "The internet of Things: An overview," The Internet Society (ISOC), pp. 80, 2015.

[21] M.A Razzaque, M. Milojevic-Jevric, A. Palade, and S. Clarke, "Middleware for internet of Things: a survey," in IEEE Internet of Things journal, vol. 3, no. 1, pp. 70-95, 2015.

[22] P.P. Ray, "A survey on Internet of Things architectures," in Journal of King Saud University-Computer and Information Sciences, vol. 30, no. 3, pp. 291-319, 2018.

[23] R. S. Sinha, Y. Wei, and S. H Hwang, "A survey on LPWA technology: LoRa and NB-IoT" Ict Express, vol. 3, no. 1, pp. 14-21, 2017.

[24] A. Bröring et al., "Enabling IoT Ecosystems through Platform Interoperability," in IEEE Software, vol. 34, no. 1, pp. 54-61, 2017.

[25] C. Bormann, A. P. Castellani, and Z. Shelby, "CoAP: An Application Protocol for Billions of 8 https://www.w3.org/WoT Tiny Internet Nodes," IEEE Internet Comput., vol. 16, no. 2, pp. 62–67, Mar. 2012.

[26] IBM and Eurotech, "MQTT V3.1 Protocol Specification." [Online]. Available: http://public.dhe.ibm.com/software/dw/webservices/ws-mqtt/mqtt-v3r1.html. [Accessed: 24- Apr-2014].

[27] Open Mobile Alliance, "Lightweight Machine to Machine Technical Specification, Candidate," OMA, 2015.

[28] A. Bröring, J. Echterhoff, S. Jirka, I. Simonis, T. Everding, C. Stasch, S. Liang, and Rob Lemmens, "New Generation Sensor Web Enablement," Sensors, vol. 11, no. 3, pp. 2652– 2699, 2011.

[29] J. Swetina, G. Lu, P. Jacobs, F. Ennesser, and J. Song, "Toward a standardized common M2M service layer platform: Introduction to oneM2M," Wirel. Commun. IEEE, vol. 21, no. 3, pp. 20–26, 2014.

[30] L. Seremeti, C. Goumopoulos, and A. Kameas, "Ontology-based modeling of dynamic ubiquitous computing applications as evolving activity spheres," Pervasive and Mobile Computing, vol. 5, no. 5, pp. 574-591, 2009.

[31] M. Noura, A. Gyrard, S. Heil and M. Gaedke, "Automatic Knowledge Extraction to build Semantic Web of Things Applications," in IEEE Internet of Things Journal, 2019.

[32] P. Murdock et al., "Semantic interoperability for the Web of Things," 2016.

[33] H. Veer, and A. Wiles, "Achieving Technical Interoperability-the ETSI approach, European Telecommunications Standards Institute," Accessed: Sep 2008, 20, 2017.

[34] P. Barnaghi, W. Wang, C. Henson and K. Taylor, "Semantics for the Internet of Things," Int. J. Semant. Web Inf. Syst. vol. 8, no. 1, pp. 1–21, 2012.

[35] K. Kotis, A. Katasonov, and J. Leino, "Aligning smart and control entities in the IoT," In Internet of Things, Smart Spaces, and Next Generation Networking, Springer, Berlin, Heidelberg, pp. 39-50, 2012.

[36] M. Ma, P. Wang and C. Chu, "Ontology-Based Semantic Modeling and Evaluation for Internet of Things Applications," 2014 IEEE International Conference on Internet of Things (iThings), and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom), Taipei, pp. 24-30, 2014.

[37] I. P. Zarko et al., "Towards an IoT framework for semantic and organizational interoperability," 2017 Global Internet of Things Summit (GIoTS), Geneva, pp. 1-6, 2017.

[38] T. Jell, A. Bröring and J. Mitic, "BIG IoT – interconnecting IoT platforms from different domains," 2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC), Funchal, pp. 86-88, 2017.

[39] G. Hatzivasilis et al., "The Interoperability of Things: Interoperable solutions as an enabler for IoT and Web 3.0," 2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), pp. 1-7, Barcelona, 2018.

[40] S. Žitnik, M. Janković, K. Petrovčič and M. Bajec, "Architecture of standard-based, interoperable and extensible IoT platform," 2016 24th Telecommunications Forum (TELFOR), Belgrade, pp. 1-4, 2016.

[41] G. Fortino et al., "Towards multi-layer interoperability of heterogeneous IoT platforms: The INTER-IoT approach," In: Integration, interconnection, and interoperability of IoT systems. Springer, p. 199-232, Cham, 2018.

[42] M. Elkhodr, S. A. Shahrestani, and H. Cheung, "The Internet of Things: New Interoperability, Management and Security Challenges," 2016.

[43] Y. Guan et al., "An open virtual neighbourhood network to connect IoT infrastructures and smart objects — Vicinity: IoT enables interoperability as a service," 2017 Global Internet of Things Summit (GIoTS), Geneva, pp. 1-6, 2017.

[44] L. Daniele, F. den Hartog, and J. Roes, "Created in close interaction with the industry: the smart appliances reference (SAREF) ontology," In International Workshop Formal Ontologies Meet Industries, Springer, Cham, 2015.

[45] G. Hatzivasilis et al., "The Interoperability of Things: Interoperable solutions as an enabler for IoT and Web 3.0," 2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), pp. 1-7, Barcelona, 2018.

[46] J. Soldatos et al., "Openiot: Open source internet-of-Things in the cloud. In Interoperability and open-source solutions for the internet of Things," Springer, Cham, pp. 13-25, 2015.

[47] OpenIoT Consortium. [Online]. Available from: http://www.openiot.eu/ 2019.07.25.

[48] A. Gyrard, S. K. Datta, C. Bonnet and K. Boudaoud, "Standardizing generic cross-domain applications in Internet of Things," 2014 IEEE Globecom Workshops (GC Wkshps), Austin, TX, pp. 589-594, 2014.

[49] A. Gyrard, M. Serrano, J. B. Jares, S. K. Datta, and M. I. Ali, "Sensor-based linked open rules (S-LOR): An automated rule discovery approach for IoT applications and its use in smart cities," In Proceedings of the 26th International Conference on World Wide Web Companion, pp. 1153-1159, 2017.

[50] R. Agarwal, D. G. Fernandez, T. Elsaleh, A. Gyrard, L. Sanchez, and V. Issarny, "Unified IoT ontology to enable interoperability and federation of testbeds," In 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT), pp. 70-75, 2016.

[51] F. Carrez, T. Elsaleh, D. Gómez, L. Sánchez, J. Lanza, and P. Grace, "A Reference Architecture for federating IoT infrastructures supporting semantic interoperability," in 2017 European Conference on Networks and Communications (EuCNC), pp. 1-6, IEEE, 2017.

[52] I. Gojmerac, P. Reichl, I.P. Žarko, and S. Soursos, "Bridging IoT islands: the symbIoTe project," in e & i Elektrotechnik und Informationstechnik, vol. 133, no. 7, pp. 315-318, 2016.

[53] C. Agostinho, Y. Ducq, G. Zacharewicz, J. Sarraipa, F. Lampathaki, R. Poler, and R. Jardim-Goncalves, "Towards a sustainable interoperability in networked enterprise information systems: trends of knowledge and model-driven technology." Computers in Industry, vol. 79, pp. 64-76, 2016.

[54] G. Hatzivasilis, K. Fysarakis, O. Soultatos, I. Askoxylakis, I. Papaefstathiou, and G. Demetriou, "The industrial internet of Things as an enabler for a circular economy Hy-LP: a Novel IIoT protocol, evaluated on a wind park's SDN/NFV-enabled 5G industrial network," Computer Communications, vol. 119, pp. 127-137, 2018.

[55] P. Barnaghi, W. Wang, C. Henson, and K. Taylor, "Semantics for the Internet of Things: early progress and back to the future," International Journal on Semantic Web and Information Systems (IJSWIS), vol. 8, no. 1 pp. 1-21, 2012.

[56] M. B. Doumbouya, B. Kamsu-Foguem, H. Kenfack, and C. Foguem," Telemedicine using mobile telecommunication: towards syntactic interoperability in teleexpertise," Telematics and informatics, vol. 31, no 4, pp. 648-659, 2014.

[57] M. Marjani, F. Nasaruddin, A. Gani, A. Karim, I.A. Hashem, A. Siddiqa, and I. Yaqoob, "Big IoT data analytics: architecture, opportunities, and open research challenges," IEEE Access, vol. 5, pp. 5247-5261, 2017.

[58] Y. Ai, M. Peng, and K. Zhang, "Edge computing technologies for Internet of Things: a primer. Digital Communications and Networks," vol. 4, no 2, pp. 77-8, .2018.

[59] O.B. Sezer, E. Dogdu, M. Ozbayoglu, and A. Onal, "An extended iot framework with semantics, big data, and analytics," In 2016 IEEE International Conference on Big Data (Big Data), pp. 1849-1856, IEEE, 2016.

[60] H. Cai, B. Xu, L. Jiang and A. V. Vasilakos, "IoT-Based Big Data Storage Systems in Cloud Computing: Perspectives and Challenges," in IEEE Internet of Things Journal, vol. 4, no. 1, pp. 75-87, 2017.

[61] I.P. Žarko, J. Iranzo, C. Ruggenthaler, J. a. S. Murillo, J. Garcia, P. Skočir, & S. Soursos, "Collaboration mechanisms for IoT platform federations fostering organizational interoperability," In 2018 Global Internet of Things Summit (GIoTS), pp. 1-6, 2018.

[62] C. Goumopoulos, and I. Mavrommati, "A framework for pervasive computing applications based on smart objects and end user development," in Journal of Systems and Software, vol. 162, p.110496, 2020.

Approximate Dynamic Programming for Optimal Direct Marketing

Jesper Slik and Sandjai Bhulai

Vrije Universiteit Amsterdam
Faculty of Science, Department of Mathematics
Amsterdam, The Netherlands
Email: jesper.slik@pon.com and s.bhulai@vu.nl

***Abstract*—Email marketing is a widely used business tool that is in danger of being overrun by unwanted commercial email. Therefore, direct marketing via email is usually seen as notoriously difficult. One needs to decide which email to send at what time to which customer in order to maximize the email interaction rate. Two main perspectives can be distinguished: scoring the relevancy of each email and sending the most relevant, or seeing the problem as a sequential decision problem and sending emails according to a multi-stage strategy. In this paper, we adopt the second approach and model the problem as a Markov decision problem (MDP). The advantage of this approach is that it can balance short- and long-term rewards and allows for complex strategies. We illustrate how the problem can be modeled such that the MDP remains tractable for large datasets. Furthermore, we numerically demonstrate by using real data that the optimal strategy has a high interaction probability, which is much higher than a greedy strategy or a random strategy. Therefore, the model leads to better relevancy to the customer and thereby generates more revenue for the company.**

Keywords–email marketing; Markov decision processes; approximate dynamic programming; evolutionary computing; recommender systems.

I. INTRODUCTION

Customer communication is crucial to the long-term success of any business [1]. Research has shown communication effectiveness to be the single most powerful determinant of relationship commitment [2]. Companies can choose from multiple channels in reaching their customers. The recent rise of social media has expanded the possibilities immensely. Most research focuses on email communication, though, because it is relatively easy to collect data of every email sent and every interaction resulting from the email on a customer level. Therefore, a thorough analysis of email communication effectiveness is possible.

Currently, in most companies, domain experts determine the email strategy. Customers are selected for emails based on business rules. These rules can be deterministic, such as matching the email's language or gender with those of the customer. However, they can also be stochastic, such as matching the (browsing) activity categories of a customer to the email category. Measurements suggest that a large fraction of the emails are unopened, a larger portion of the emails do not even direct customers to the company's website, and almost all emails are not related to direct sales. An increase in the interaction probability, therefore, directly leads to additional revenue. This probability can be increased by a better recommendation process of deciding which email to send at what time to which customer.

The challenge faced in this research can be classified within the research field of recommender systems. A recommender system has as purpose to generate meaningful recommendations of items (articles, advertisements, books, etc.) to users. It does so based on the interests and needs of the users. Such systems solve the problem of information overload. Users might have access to millions of choices but are only interested in accessing a fraction of them. For example, Amazon, YouTube, Netflix, Tripadvisor, and IMDb use recommender systems to display content on their web pages [3]. Similarly, one can use recommender systems to recommend certain emails to users, thus, to determine when to send which email to which user.

Recommender systems have traditionally been classified into three categories: content-based filtering, collaborative filtering, and hybrid approaches [4]. Content-based filtering is a recommendation system that learns from the attributes (or the so-called contents) of items for which the user has provided feedback [5]. By doing so, it can make a prediction on the relevancy of items for which the user has not provided feedback. Collaborative filtering looks beyond the activity of the user for which a recommendation needs to be made. It recommends an item based on the ratings of similar users [4]. Hybrid recommender systems make use of a combination of the above-mentioned techniques in order to generate recommendations.

Although recommender systems might seem a good way to address the direct marketing problem, they have some shortcomings. One of the major problems for recommender systems is the so-called cold-start problem. This concerns users or items which are new to the system; thus, little information is known about them. A second issue is that traditional recommender systems take into account a set of users and items and do not take into account contextual information. Contextual information might be crucial for the performance of a recommender system [6]. A third issue is an overspecialization: "When the system can only recommend items that score highly against a user's profile, the user is limited to being recommended items that are similar to those already rated" [4]. Lastly, recommender systems must scale to real data sets, possibly containing millions of items and users. As a consequence, algorithms often sacrifice accuracy for having a low response time [3]. When a data set increases in size, algorithms either slow down or require more computational resources.

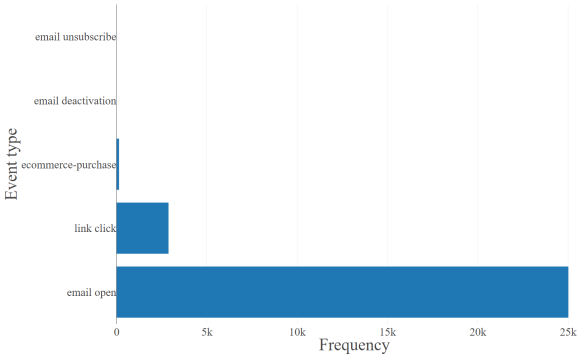


Figure 1. Frequency of event types.

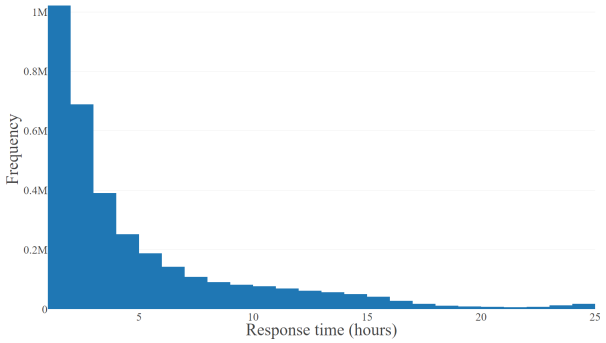


Figure 2. Distribution of time until the first interaction with an email.

The main contribution of this paper is that we address the mentioned shortcomings of the traditional recommender systems by formulating the direct marketing problem as a Markov Decision Process (MDP). This framework deals with context and uncertainty in a natural manner. The context (such as previous email attempts) can be specified in the state space of the MDP. The uncertainty is addressed by the optimal policy as an exploration-exploitation trade-off. The scalability of the algorithm is addressed by limiting the history of the process to sufficient information such that the state space does not grow intractably large. Furthermore, we test our model with real data on a greedy and random policy as a benchmark. The results show that our optimal strategy has a significantly higher interaction probability than the benchmark.

In this paper, we expand on our work in [1] by doing a more thorough data analysis, implementing an alternative method to solve the MDP, and by further elaborating on the discussion.

The organization of this paper is as follows. In Section II, we describe the data used for our data-driven marketing algorithm. Section III describes the model and introduces the relevant notation. In Section IV, we analyze the performance of the model and state the insights from the model. Finally, in Section V, we conclude and address a number of topics for further research.

II. Data

In this section, we describe the data used for this research. We explain the data, comment on the data quality, filtering, and processing. Finally, we explore the data by showing relevant statistics and visualizations.

The data is gathered from five tables of an international retailer from one complete year and concerns: *sales* data, *email sent* data, *email interaction* data, *customer activity* data, and *customer* data.

The *sales* table contains all orders that have been placed by each customer. This includes information on the product, price, and date. The *email sent* table contains all emails sent to each customer. An email is characterized by attributes such as title, category, type, gender, and date. The *email interaction* table is structured similarly to the *email sent* table, however, it contains an interaction type. An interaction type can be email open, link click, online purchase, email unsubscribe, or email deactivation. The *customer activity* table contains for each customer its activity on the retailer's platform, such as browsing or clicking on the website. Finally, the *customer* table contains characteristics of a customer, such as date of birth, country, city, and gender.

Quality

The data used for this research is, for the large part, automatically generated. However, this does not guarantee its quality. Some issues appear when inspecting the data.

First, according to the data, 232 countries exist. Although there is discussion on the number of countries in the world, the United Nations (UN) recognizes a little under 200 countries. Business rules can explain the high number of countries in the database, such as classifying a part of the business (e.g., customer services) as a separate country. We tackle this issue by filtering on countries recognized by the UN.

Second, some physical stores are classified as individual customers. This results in these customers making hundredths of orders every year, creating much revenue. For these reasons, they can easily be identified.

Last, a large part of the customers does not place orders or show activity. This might be because one physical customer might have multiple accounts or devices through which interactions are made. Additionally, bots or spam accounts might be classified as customers. Business rules and logic is applied to identify and consolidate; however, this logic is not 100% accurate.

Processing

In this research, we analyze a vast amount of data. After filtering, we analyze approximately just over a million customers, but millions of emails, orders, and email interactions. Just the size of the raw email table is larger than 200GB. Such amounts of data cannot be processed on a standard, local machine. Thus, we used cloud technologies to process the data. The tables were queried using the Presto query engine. The query results were analyzed using various Python scripts making use of Spark (PySpark). In total, 14 queries and 22 Python scripts were written to explore data, process data, and build models.

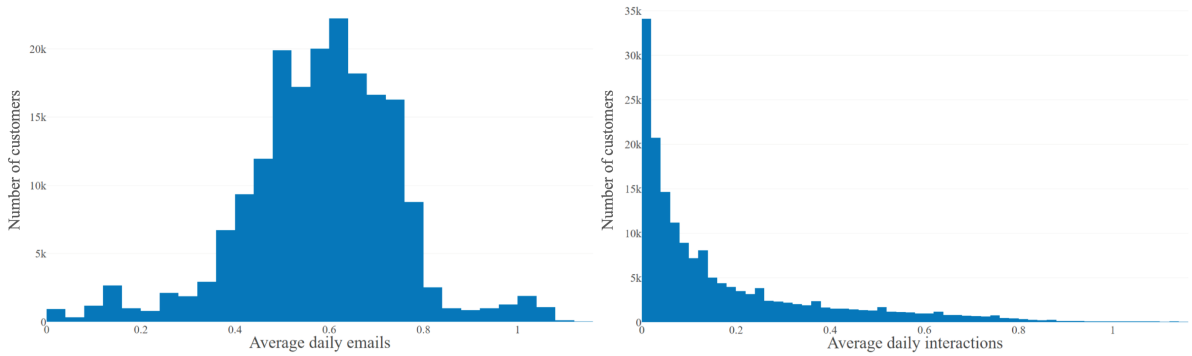


Figure 3. Distributions of emails received and emails interacted with by customers: average daily emails (left) and average daily interactions (right).

Figure 4. The various filters applied to the customer data.

Filtering

Given the data size, data quality challenges, and to focus on relevant customers, we filter the data. In the raw data, we have approximately 240 million unique customers. However, this does not correspond to reality because business definitions, falsely identified customers, or inactive customers inflate this number. We reduce this to approximately 1 million customers by applying various filters. This procedure is visualized in Figure 4. First, we exclude stores by limiting the number of purchases and the total order value of each customer. Next, we focus only on customers that have ever placed an order, are registered (this excludes guest accounts), are flagged as customer (excludes duplicate accounts), have indicated that they can be contacted online, are situated in Europe, and have shown activity on the online platform in the previous year.

Data exploration

The retailer has over 1 million unique active customers in its database. In total, a little more than 132 million emails were sent, leading to around 34.5 million interactions. The main interaction category is 'email open', which occurs over five times more frequently than the second interaction category, 'link click'. This is intuitive, as an email needs to be opened in order to click a link. Even fewer emails are related to direct online sales, and rarely an email leads to an unsubscribe or deactivation (see Figure 1). The customers that interact with an email, usually do so within a few hours. The majority even within one hour, with the number of interactions declining by the hour afterward. Only after 24 hours, there is a slight increase in the number of interacting customers (see Figure 2).

With the current email strategy, the retailer does not send the same emails to the same customers. The average customer receives an email every other day and interacts with an email every 10 days. Interestingly, some customers interact with more than 1 email per day on average. The email interaction rate varies between the email category and email type. The interaction rate of individual emails shows even larger differences. This rate ranges from 3.4% to 67%. Figure 3 shows the average daily emails received and the average daily interactions per customer. The distributions of both statistics differ much. The average daily interactions look exponentially distributed by visual inspection, whereas the average daily emails received looks more normally distributed.

In this research, we are mainly interested in delivering relevant communication to the customers. Whether an email is relevant to a customer can be expressed by whether the customer interacted with the email. We investigate two correlations related to the email interaction rate. We do this by visualizing the relation with a scatter plot (plotting a random sample of the data) and including a 95% confidence interval for the mean. The confidence interval is created through a bootstrap procedure.

Figure 5 (left) visualizes the correlation between the average number of emails received and the number of interactions. The average daily interactions is positively correlated with the average daily emails. This is intuitive, as it would benefit no strategy to send more emails to a customer that does not interact with emails. Also, it is impossible for a customer to interact with two emails if the customer only received one. However, sending more emails does not necessarily mean more interactions. Figure 5 (right) visualizes the correlation between the interaction probability and total order value of a specific customer. The interaction probability is defined as the number of interactions divided by the number of received emails for a specific customer. The graph indicates that a higher interaction probability is correlated with a higher-order value. When looking at the interaction probabilities of 0.3 and 0.4, the confidence intervals for the mean total order value (averaged over all customers) are non-overlapping. For a probability of 0.3, the confidence interval is $[174.68, 180.71]$ and for a probability of 0.4 this yields $[189.02, 195.11]$. Thus, customers that have a higher interaction probability have a

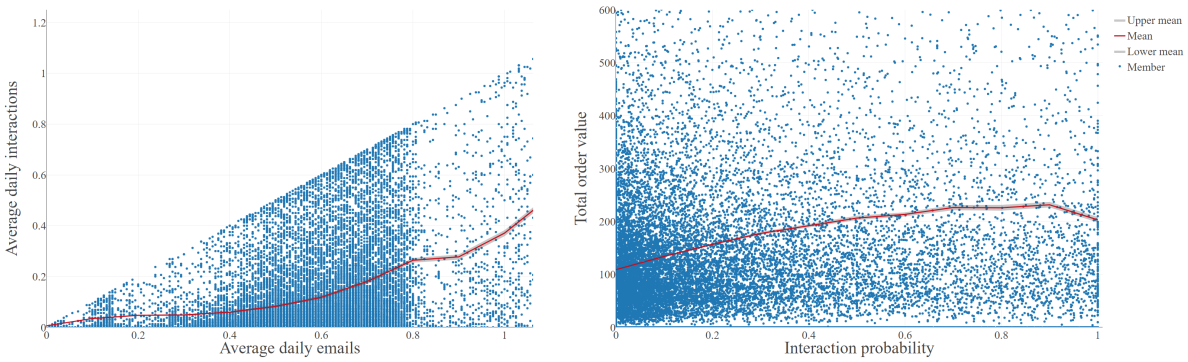


Figure 5. Scatter plots diagrams: # emails vs # interactions (left) and interaction probability vs customer order value (right).

higher customer value (for interaction probabilities smaller than 0.8).

III. Model description

We implement a discrete-time Markov decision process (MDP) for our email marketing process. The MDP is defined by four entities: the state space $\mathcal{S}$, the action space $\mathcal{A}$, the reward function r, and the transition function p.

We define a state $s \in \mathcal{S}$ as a vector of the form $s = (x_0, x_1, x_2, y_0, y_1)$. Here, x_i represents the $(3-i)^{\text{th}}$ previous interaction of the customer for $i \in \{0, 1, 2\}$. Similarly, y_j is defined as:

$$y_j = \begin{cases} 1, & \text{if } (2-j)^{\text{th}} \text{ previous action led to an interaction,} \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

for $j \in \{0, 1\}$.

This choice for the state is partially inspired by [7], in which the state is defined as the sequence of the past k items bought. We make a clear distinction between actions and interactions, an action meaning sending an email to a customer and an interaction meaning the customer interacting with an email. The x_i's of the state space represent a customer's preference in content, and the y_i's represent the customer's sensitivity to emails. The parameters $i = 3$ and $j = 2$ have been empirically chosen, leading to an approximate model. There is a trade-off between tailoring the model for individuals and, more accurately, estimating the model parameters. The size of the state space grows exponentially as i and j are increased, since $|\mathcal{S}| = |\mathcal{A}|^i\, 2^j$.

We define an action $a_i \in \mathcal{A}$ as an integer. This integer represents a combination of email category and email type. An example of a category is 'household products' and an example of type is 'special event'. In our data, 20 categories and 21 types exist. However, not all combinations of category and type appear in the data. Therefore, we focus on the 20 actions that occur most frequently. In this way, we reduce the size of the action set by 95% at the cost of discarding 21% of the data.

The reward function represents the reward (business value) of a customer visiting a state. We aim to maximize the communication relevancy to the customers. This can be measured by customers interacting with emails. Thus, the reward function should measure email interactions. We define the reward function as $r(s) = y_1$ for $s = (x_0, x_1, x_2, y_0, y_1)$. This function expresses whether the previous action leads to an interaction. Conveniently, the last element in the state vector already does so.

The transition probabilities are estimated by simply counting the occurrences of a transition in the data. Specifically,

$$p(s, a, s') = \frac{C(s, a, s')}{\sum_{s' \in \mathcal{S}} C(s, a, s')},$$

in which $C(s, a, s')$ is a function that counts the number of occurrences of transitioning from state s to state s' when applying action a. To create the data to estimate these probabilities, three steps are required. First, we collect on a daily level which action and interaction was registered with which customer. Next, we compute the state of each customer based on this information. Lastly, we aggregate all state changes of all customers into one final table. These steps are visualized in Figure 6.

To summarize the implementation of the MDP, we present an example. This example is visualized in Figure 7. The example highlights that when a customer is in state $s_t = (14, 6, 10, 0, 0)$ and action $a_t = 17$ is applied, we have a 19% probability of transitioning to state $s_{t+1} = (6, 10, 17, 0, 1)$ (since $p(s_t, a_t, s_{t+1}) = p\big((14, 6, 10, 0, 0), 17, (6, 10, 17, 0, 1)\big) = 0.19$) and an 81% probability of transitioning to state $s_{t+1} = (14, 6, 10, 0, 0)$. Note that for any s_t, only two possibilities exist for s_{t+1}.

Modeling considerations

Multiple challenges arise when modeling the problem as an MDP. Most of these have been tackled by defining an appropriate MDP as done in the previous paragraphs. However, some modeling choices remain, which are described next.

A. *The unichain condition*

In order for solution techniques to work for our model, the MDP needs to be unichain. The unichain property states that there is at least one state $s \in \mathcal{S}$, such that there is a path from any state to s [8]. A path from z_0 to z_k of length k is defined as a sequence of states $z_0, z_1, \ldots, z_k$ with $z_i \in \mathcal{S}$ with the property that $p(z_0, z_1) \cdots p(z_{k-1}, z_k) > 0$.

customer id	date	action	interaction
a	1	18	0
a	3	15	15
a	5	3	3
a	6	14	0
a	7	6	6
a	10	20	0
…	…	…	…

customer id	date	state	action	state next
a	1	(1, 1, 1, 1, 0)	18	(1, 1, 1, 0, 0)
a	3	(1, 1, 1, 0, 0)	15	(1, 1, 15, 0, 1)
a	5	(1, 1, 15, 0, 1)	3	(1, 15, 3, 1, 1)
a	6	(1, 15, 3, 1, 1)	14	(1, 15, 3, 1, 0)
a	7	(1, 15, 3, 1, 0)	6	(15, 3, 6, 0, 1)
a	10	(15, 3, 6, 0, 1)	20	(15, 3, 6, 1, 0)
…	…	…	…	…

state	action	state next	frequency
(11, 9, 17, 1, 1)	9	(9, 17, 9, 1, 1)	5197
(11, 9, 17, 1, 1)	9	(11, 9, 17, 1, 0)	828
(11, 9, 17, 1, 1)	11	(9, 17, 11, 1, 1)	6561
(11, 9, 17, 1, 1)	11	(11, 9, 17, 1, 0)	1042
(11, 9, 17, 1, 1)	12	(9, 17, 12, 1, 1)	10
(11, 9, 17, 1, 1)	12	(11, 9, 17, 1, 0)	2
…	…	…	…

Figure 6. The three data processing steps required for estimating the transition probabilities.

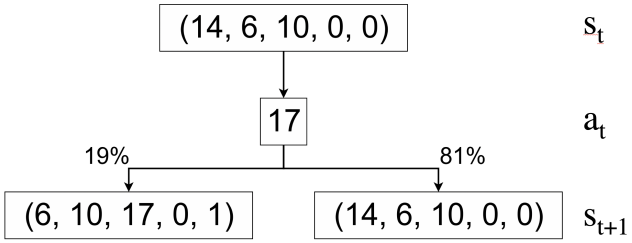


Figure 7. Example transition.

The unichain property does not automatically hold when we take all states and state transitions directly from the data. This is because the chain is partially observed, so for some states, it is not observed that a specific action causes an interaction. For some states, it might only be observed that the next possible state is the current state. We solve this problem by removing all states for which fewer than 2 next states are observed.

B. Estimation of transition probabilities

In our implementation, making the MDP unichain reduces the number of observed states. A problem with the estimation of the transition probabilities is that some probabilities are based upon thousands of observations, whereas others only on a few observations. This introduces noise in the transition probabilities. To tackle this challenge, we recursively remove state transitions that occur fewer than 50 times and, if this leads to states being impossible to transition to, we also remove those and transitions to those states.

The MDP is partially observed; we initially observe 86% of the theoretically possible states. After filtering, we are left with 39% of possible states. This is a large reduction in the number of observed states. However, it does ensure we focus on the most relevant and frequently observed states. Figure 8 shows the distribution of the number of observed transitions per state before filtering.

C. Exponential growth

Lastly, defining and solving an MDP can be difficult because of the exponential growth of the state space due to the multiple components of the state, as discussed before, when setting the values of i and j. If the state space becomes too large, solving the MDP might not be realistic. To ensure the MDP can be solved within a feasible time period, we implement a custom version of the value iteration algorithm, taking into account the following issues.

In our case, the set of possible next states, defined as $E(s, a)$, only consists of 2 states. This significantly reduces the run time of the algorithm. If we did not do this, the algorithm would have to check the transition probabilities to and values of all 32,000 possible states.

We implemented the action set, $\mathcal{A}$, as being dependent on the state, thus redefining it as $\mathcal{A}(s)$. For some states, not all 20 actions are observed. So it is unknown to the model what the transitions would be. Not taking into account these unknown actions improves the performance of the algorithm.

Finally, we initialize $E(s)$, $\mathcal{A}(s)$, and $p(s, a, s')$ for all s, a, and s' in memory using Python dictionaries. This allows for $\mathcal{O}(1)$ lookup steps of any probability, action set, or the set of next states within the algorithm.

Finding the optimal policy

We find the optimal policy to the MDP by using two methods: value iteration and Evolutionary Computing (EC). The field of Evolutionary Computing (EC) can be seen as a family of algorithms that acts as a meta-heuristic. They can be applied to finding the optimal value to an MDP and potentially generating near-optimal solutions efficiently. We have not observed this way of using EC in the literature.

Inspired by nature, EC works with notations of an individual, population, generation, selection, recombination, and mutation [9]. A generation is a population in a certain time period, a population consists of multiple individuals, and an individual represents some solution to a problem. Individuals can recombine with other individuals to create offspring, and individuals can be mutated. Selection operators determine which individuals pass on to the next generation.

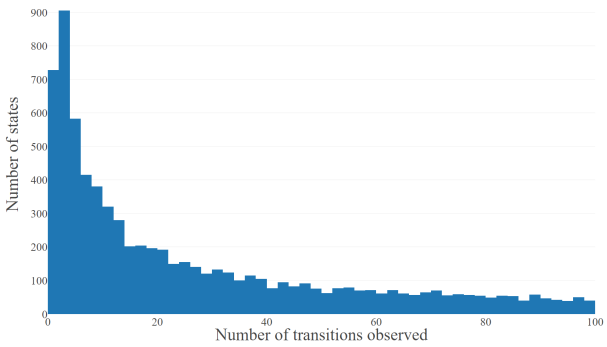


Figure 8. Distribution of the number of observed transitions per state.

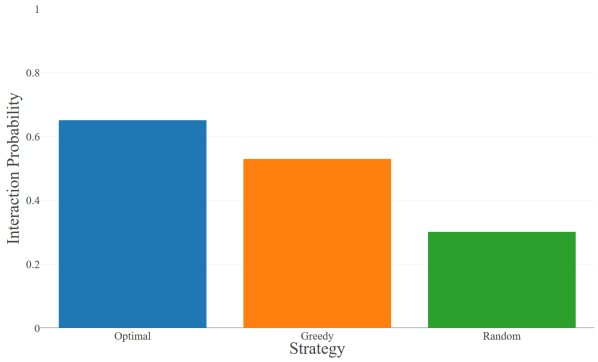


Figure 9. Comparing strategy performance: optimal vs. greedy vs. random.

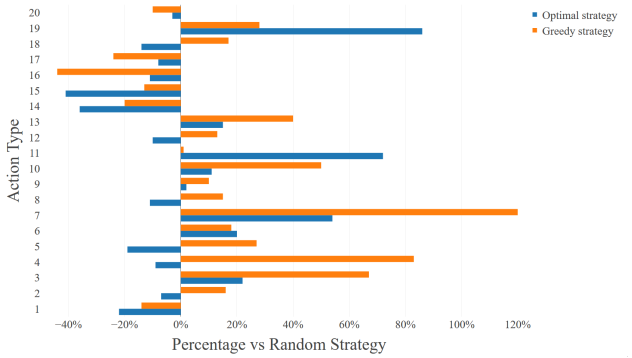


Figure 10. Action performance: frequency of an action within the optimal or greedy strategy divided by the expected frequency of that action.

We view an individual as a strategy, thus, as a mapping from state to action. Therefore, we represent an individual as a vector of integers $\langle s_1, \ldots, s_n \rangle, n = |S|, s_i \in A$, in which s_i is a gene specifying which action to take in the state at index i in the list of possible states. We can assign a fitness f_i to individual i by applying a variation of the value iteration algorithm, in which we do not follow the optimal but the current strategy.

We initialize the population by randomly generating individuals. We choose a population size of $\mu = 100$, fitness proportionate parent selection, the uniform recombination operator, and the random reset mutation operator.

Regarding survivor selection, we use a λ-μ ratio of 2, which means we generate twice as many offspring as we have parents. We use a $(\mu + \lambda)$ survivor selection technique, we select survivors from the union of the current population and the children. The survivor selection operator we use is a tournament selection procedure of size $k = 6$. We sample k individuals from the set of parents and children, and choose the individual with the highest fitness as a survivor. We repeat this process until our new population size equals μ. Additionally, we use elitism, i.e., the best individual from the old generation is always selected for the new generation.

Our termination condition is based upon time; we stop generating offspring after running the algorithm for 24 hours.

Modeling Considerations

Next to the considerations of creating the MDP, two more challenges arise when modeling the EC approach. They are described as follows.

D. Choice of components

The main challenge is choosing the components and the parameters they imply. We make these choices based upon a grid search procedure. This procedure is time-consuming, as most parameters influence the balance between exploitation and exploration, which concerns the algorithm's performance in the short- and long-term. It might seem that a parameter positively affects the fitness in the early generations. However, when looking at a longer horizon, this might not be true.

E. Evaluation of strategies

The evaluation of strategies is time-consuming as well. To improve the performance, we use a modified version of the value iteration algorithm, in which we only follow the given strategy. This has the advantage that not every action in every state should be taken into account, and thus, the algorithm converges faster. To further reduce the evaluation time, we decrease the convergence threshold ε over the generations. In this way, the evaluation of the population is faster in the first generations and gradually slows.

IV. Results

In this section, we present an analysis of the performance of the models. We analyze the strategy performance by comparing three different strategies, all based on the MDP framework: the optimal strategy, a greedy strategy, and a random strategy (benchmark). The optimal strategy is calculated through value iteration, the greedy strategy by choosing in each state the action with the highest interaction probability, and the random strategy by randomly choosing an action in each state.

Figure 9 shows the resulting performance of the three strategies. The optimal strategy has the highest long-run interaction probability, corresponding to a value of 65%. The greedy strategy is second with a rate of 53%, and the random strategy with 30%. Interestingly, the interaction rate of the optimal strategy is 23% higher than the rate of the greedy strategy, showing that taking into account delayed rewards can highly increase the strategy value. Both the optimal and greedy strategy perform better than the random strategy, showing that using advanced strategies has a large impact on the interaction rate.

Figure 10 highlights the effectiveness of each action type. This effectiveness is measured by dividing the frequency of an action within the optimal or greedy strategy over the expected frequency of that action. It is measured in this way, since an absolute measure would not be accurately representing the action performance, as in some states only one action might be possible. So the absolute measure would not represent how much the action is preferred over other actions. A comparison between the greedy and optimal strategy is made to highlight the difference between short- and long-term rewards of the corresponding action.

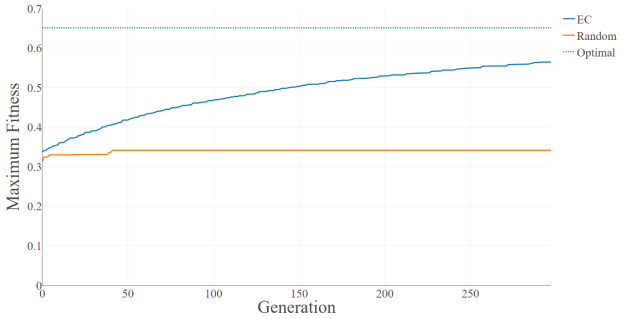


Figure 11. Performance of EC: maximum fitness per generation

Large differences are visible in action performance. Actions that perform well on both the short- and long-term are action 7: the type retail clearance, 19: weekly product releases in a specific category, and 6: new releases. Interestingly, some actions are highly beneficial for the long-term, but not beneficial for the short-term, see., e.g., action 4. Actions that perform poorly are action 1, 14, 15, 16, or 17, which are all weekly product releases. It seems that only the weekly product release in a specific category (action 19) performs well.

Figure 11 shows the fitness of the best individual throughout the generations during the run of the EC algorithm. This fitness curve increases rapidly in the first generations. However, the increase slows down as the generations pass. The algorithm did not find the optimal strategy within the time limit of 24 hours, which corresponds to running 297 generations. Given more time, the algorithm will find better individuals and converge towards the optimal solution. The value of the optimal strategy is highlighted by the dotted line. The orange line (the line with the lowest maximum fitness) shows the maximum fitness found by a random search.

V. Conclusions and Discussion

This research shows that the retailer can increase its relevance to its customers by applying a different email strategy. Hereby, it possibly increases the revenue it generates. However, the strategy we developed is based on the data generated from the retailer's current email strategy. If the retailer starts experimenting with different strategies, this might uncover patterns unknown to the current model and potentially improve the optimal strategy we presented.

An interesting result of this research is the difference between the optimal and the greedy strategy. The interaction rate of the optimal strategy is 23% higher, relatively. Thus, the balance between short- and long-term rewards should be taken into account when dealing with similar problems. If we had chosen to use traditional methods, such as content-based or hybrid filtering, this result would not have been directly visible. These methods do not explicitly include this balance, so during the modeling process, it will be beneficial to try to include this balance.

Moreover, the results indicate a 'reality gap' between theory and practice. The interaction rate of the random strategy (30%) is higher than the interaction rate of the retailer's current strategy (27%). This is probably because our model has fewer restrictions compared to real life. However, with the interaction rate of the optimal strategy being 65%, the model shows to have potential.

Throughout this research, all data concerns the past. However, to measure the impact of strategies more accurately, it would be better to measure the performance in real-time. For example, through an A/B testing procedure. Additionally, an algorithm like reinforcement learning could be used to learn the value of strategies in real-time. This algorithm is known to balance short- and long-term rewards and balance the trade-off between exploration and exploitation. It hereby tries to both learn a better strategy and apply the best-known current strategy whilst executing certain strategies.

Furthermore, we can extend the model by redefining actions. In this research, we focused on emails. However, this channel is not tied to the model. In the future, the same model can optimize push notifications of mobile applications, in exactly the same manner as the current model does.

Our research results show that evolutionary computing is less efficient in finding the optimal solution than the value iteration algorithm. The value iteration algorithm converges below ϵ within 20 minutes on the same machine. Potentially, the EC approach can be improved by choosing different operators. However, the algorithm needs to be improved largely in order to match the speed of the value iteration algorithm. On our MDP, the EC approach seems inadequate; however, in other cases, it might still be a good idea to implement. For example, an MDP where the action space is larger and, therefore, the value iteration algorithm might have difficulties to converge. In this case, the EC approach can deliver better strategies than random, and if given enough time, approach the optimal solution.

Research opportunities

As with any model, the model we presented in this research is a simplification of reality. The main impact is that, compared to real life, the model can choose between more actions. In reality, not every action can be undertaken in every time period. This can be improved by further restricting the action set, based upon the state. For example, incorporating the previous action in the state and restricting the action set based on this previous action.

Furthermore, the estimate of transition probabilities can be improved. At the moment, this estimation is based upon counting frequencies. However, when transitions are not observed or observed infrequently, this estimation is unreliable and these transitions are filtered. This leads to a further restricted state space. Instead of removing these transitions, we could initialize a default probability from transitioning from a state to any other state. Or we could use machine learning techniques to estimate these probabilities, as a transition probability might say something about the transition probability of a similar action.

References

[1] J. Slik and S. Bhulai, "Data-driven direct marketing via approximate dynamic programming," in Proceedings of the 8th International Conference on Data Analytics. IARIA, 2019, pp. 63–68.

[2] N. Sharma and P. Patterson, “The impact of communication effectiveness and service quality on relationship commitment in consumer, professional services,” Journal of Services Marketing, vol. 13, 1999.

[3] F. Ricci, L. Rokach, B. Shapira, and P. Kantor, Recommender Systems Handbook. Springer, 2011.

[4] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions,” IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 6, 2005.

[5] M. Pazzani and D. Billsus, “Content-based recommendation systems,” in The adaptive web. Springer-Verlag, 2007, pp. 325–341.

[6] G. Adomavicius, B. Mobasher, F. Ricci, and A. Tuzhilin, “Context-aware recommender systems,” AI Magazine, 2011.

[7] G. Shani, R. Brafman, and D. Heckerman, “An MDP-based recommender system,” Journal of Machine Learning Research, no. 6, 2005.

[8] S. Bhulai and G. Koole, “Stochastic optimization,” September 2014.

[9] Eiben, A.E., Smith, J.E.: Introduction to Evolutionary Computing. Natural Computing Series. Springer, Heidelberg, 2015

Surveying the Incorporation of IoT, SCADA, and Mobile Devices into Cybersecurity Risk Management Frameworks

Aaron Pendleton
Graduate Cyberspace Operations
Air Force Institute of Technology
Wright-Patterson AFB, Ohio 45433
Email: Aaron.Pendleton@afit.edu

Richard Dill
Dept. of Electrical and
Computer Engineering
Air Force Institute of Technology
Wright-Patterson AFB, Ohio 45433
Email: Richard.Dill@afit.edu

James Okolica
Dept. of Electrical and
Computer Engineering
Air Force Institute of Technology
Wright-Patterson AFB, Ohio 45433
Email: James.Okolica@afit.edu

Dillon Pettit
Graduate Cyberspace Operations
Air Force Institute of Technology
Wright-Patterson AFB, Ohio 45433
Email: Dillon.Pettit@afit.edu

Marvin Newlin
Graduate Cyberspace Operations
Air Force Institute of Technology
Wright-Patterson AFB, Ohio 45433
Email: Marvin.Newlin@afit.edu

***Abstract*—This paper reviews the state of the art in cyber risk management with a focus on the adaptations in methodology to account for Mobile Devices, Industrial Control Systems, and Internet of Things systems into present risk analysis framework models. Internet of Things devices present unique risks to a network due to their highly connective and physically interactive nature. This physical influence can be leveraged to access peripherals beyond the immediate scope of the network, or to gain unauthorized access to systems which would not otherwise be accessible. A 2017 Government Accountability Office report on the current state of Internet of Things device security noted a lack of dedicated policy and guidance within the United States government cybersecurity risk assessment construct and similar private sector equivalents. The purpose of this paper is to expand that work and assess additional risk models. Surveyed in this paper are 30 original frameworks designed to be implemented in enterprise networks. In this research, the comparison of frameworks is analyzed to assess each system's ability to provide risk analysis for Internet of Things devices. The research categories are level of implementation, quantitative or qualitative scoring matrix, and support for future development. This survey demonstrates that there are few risk management frameworks currently available which attempt to incorporate both cyber-physical systems and enterprise architecture in a large scale network.**

***Keywords*—IoT; Mobile; Cybersecurity; Risk; ICS.**

I. Introduction

This paper is a continuation of the work "Surveying the Incorporation of IoT Devices into Cybersecurity Risk Management Frameworks" presented in the 2019 SECURWARE proceedings [1]. The paper assesses the extent that risk management frameworks have adapted to Industrial Control Systems (ICS) and Internet of Things (IoT) devices which have infiltrated most networks that would traditionally be classified as enterprise networks. The transient or multi-connected nature of IoT devices poses a challenge to security methods based on creating a secure baseline. The unprecedented rise in popularity of mobile and interconnected IoT devices has made it challenging for companies to assess and mitigate the additional risk presented by incorporating them into networks implementing risk management frameworks. Frameworks from specific industries such as online services, critical infrastructure, research and design, and enterprise risk management have been evaluated an effort to fully assess the state of the art across the security and risk industry.

IoT devices present unique risks to a network due to their highly connective and often cyber-physical nature. Enterprise networks that are not equipped with methods of assessing vulnerabilities across less traditional interfaces or protocols such as Bluetooth or remote location devices with unsecured external connections are exposed to unaccounted risks. This physical influence can be leveraged to gain unauthorized access to systems which would not otherwise be accessible [2]. Similarly, they have been shown to exhibit several widespread security challenges that require special consideration. Many IoT devices are difficult to patch, do not have consistent software updates, or lack strong encryption. This creates vulnerabilities in networks that require authentication, access control, or data privacy [3]. It is also difficult to identify IoT devices that already exist on a network due to many autonomous and passive applications [3].

The United States (U.S.) Government Accountability Office (GAO), an independent and nonpartisan U.S. Congressional watchdog organization, provides objective and reliable information to the government regarding work and spending prac-

tices. GAO focuses on identifying problems and proposes solutions [2]. In July 2017, GAO released a report titled *Internet of Things: Enhanced Assessments and Guidance Are Needed to Address Security Risks in DOD* in order to highlight shortcomings in most current operational risk assessment frameworks to include those implemented by the U.S. Department of Defense (DOD). The report includes security concerns with Mobile Devices, Supervisory Control and Data Acquisition (SCADA), Programmable Logic Controllers (PLC), and Remote Terminal Units (RTU) in the U.S. DOD [2].

GAO noted a lack of dedicated policy and guidance within the U.S. government cybersecurity risk assessment construct and similar private sector equivalents. In the report, GAO defines IoT devices as any personal wearable fitness device, portable electronic device, smartphone, or infrastructure device related to industrial control systems [2].

Present DOD Instructional Guidance does not address IoT devices sufficiently [2]. Furthermore, no single DOD entity is responsible for the security of IoT systems, and the primary guidance on IoT security is a strategic directive to establish an operations security program. This paper furthers the research done by GAO in order to expand the scope of analysis beyond the U.S. DOD and into the greater field of published cyber risk solutions.

A risk analysis methodology must account for more than just traditional enterprise network components in order to mitigate the risks presented by an unregulated or loosely defined set of devices on an otherwise secure network [2]. The purpose of this survey is to analyze the pace of development and compare the strengths and weaknesses of each analyzed framework with regard to IoT and ICS devices. The extent of advancements in risk management is assessed in order to highlight current knowledge and research gaps. 30 original risk assessment and management models are compared based on their method of risk scoring, level of implementation, and future development plans. These metrics are used to gauge the effectiveness of a framework when accounting for devices which may not be consistently part of the secure baseline, or may not be easily patched and secured. The ability of a risk analysis model to incorporate these common, but otherwise difficult to attribute systems is compared in order to establish the state of the art in currently employed systems. These methodologies are compared to recently proposed frameworks to assess the current gap in risk management. Frameworks published from as early as 2002 were identified and assessed for their ability to adapt to IoT devices. This paper analyzes the extent that network risk analysis and management frameworks have adapted to this evolving threat terrain. Section II outlines the risk framework models and their attributes, Section III presents the methods used to analyze and evaluate the frameworks in order to make accurate comparisons, and Section IV provides an assessment of the current state of the art in order to then make recommendations for future research. We conclude this work in Section V with recommendations for future work.

II. Related Work

This section reviews elements of 30 risk frameworks and provides background information used in the analysis and assessment. Specific methodology is discussed in order to establish the basic elements of each model and to ascertain the level of effectiveness observed.

A. National Institute of Standards and Technology (NIST)

The United States uses a centralized risk framework system based on application. NIST is tasked with creating and maintaining effective cyber risk modeling and management frameworks implemented on millions of government and civilian devices [4].

1) Risk Management Framework (RMF): The primary risk assessment and management framework used by the U.S. government, military and DOD to conduct mission assurance is the cybersecurity Risk Management Framework (RMF) developed by NIST. The NIST RMF process shown in Figure 1 is a six step qualitative analysis method for assessing risk. RMF uses a strict adherence to process management to establish a secure baseline through identifying controls that are to be updated as changes are detected [5]. The strength of the RMF process is that it allows for a network to grow and evolve without a complete re-evaluation of its security posture. Best practices are evaluated and selected as security controls and solutions when new devices are added to the existing baseline. The weakness in this method is it sacrifices micro-level visibility of device interactions in favor of broad security measures. NIST RMF implementation policy requires end users to disable the impertinent network components of IoT devices, but not physical removal. This leaves the opportunity for subversion of the RMF process in personal and government devices by dis-associating some capabilities from the network and the secure baseline without fully mitigating the threat. IoT and mobile devices present heightened risk levels that are left unaccounted for in the overall assessment [2]. Qualitative frameworks such as RMF rely on scanning tools and strict Information Assurance (IA) policy to prevent unauthorized activity. These security measures can be subverted by IoT devices because they often have limited up-time, minimal support, a notable lack of associated scanning tools, and a smaller footprint for vulnerability testing [2].

2) Cybersecurity Framework (CSF): The CSF is designed to provide a higher level of protection specific to the unusual or irregular systems common in Critical Infrastructure (CI). CSF is considered one of the premiere risk management models for CI, and provides a five step, tiered, qualitative approach to modeling risk to networks both small and large. The CSF framework is a guide for security measures to be implemented and allows classification of the current security posture in order to highlight pressing weaknesses. Many academic institutions, government and DOD entities, and private companies have implemented CSF. CSF continues to struggle with the same weaknesses identified in RMF despite offering significant improvements over previous generations of risk framework [7].

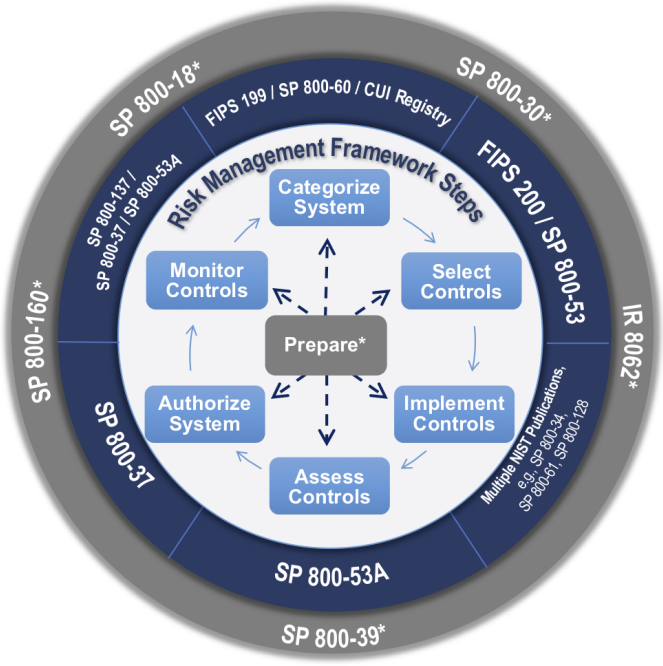


Fig. 1. Risk Management Framework Process and Governance [6]

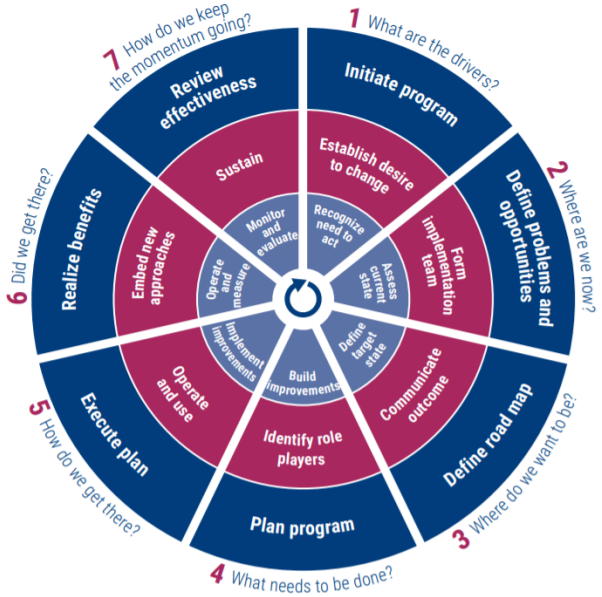


Fig. 2. COBIT 2019 Process Overview [11]

B. Control Objectives for Information and related Technology (COBIT) 5

COBIT 5 is the latest COBIT version analyzed. It was developed by the Information Systems Audit and Control Association (ISACA) and is a qualitative framework designed to provide top-down security of a business sized network. It relies on control objectives to build out the security requirements, and the level of security is assessed by maturity models. COBIT follows a purpose built model which is intended to allow for only necessary systems to be on the network in order to minimize risk [8][9]. COBIT 5 incorporated elements of the NIST CSF structure, but did not greatly mitigate the weaknesses of CSF when IoT devices are introduced to the network environment without adequate vulnerability scanning and assessment methods [7]. Initial information and methodology publications introducing COBIT 2019 have been released, but the framework implementation is not mature enough to analyze at this time [10]. Figure 2 introduces the process of implementing COBIT 2019. COBIT follows a process-based approach similar to other qualitative methods such as NIST RMF. A comparison of Figure 1 and Figure 2 shows the extent of the similarities between qualitative process-based risk management methodologies. The primary focus of the process is to identify the problem by outlining each device and defining its potential interactions with the previously established secure baseline. Security risks are then mitigated and monitored. This general approach is observed in each leading enterprise solution assessed in this survey.

C. ISO Risk Management Frameworks

The International Organization for Standardization (ISO) is an independent and international organization dedicated to developing international standards. The standards created by ISO are not inherently designed for cybersecurity applications, but they are tools for assessing risk across multiple domains.

1) ISO31K Series: The ISO 31000 standard is a general risk standard mandated in some information technology applications built off of the Australian/New Zealand risk management standard AS/NZS 4360. It identifies specific language to be used when classifying risk, but is not a strong methodology for addressing it. It is not based on quantifiable probabilities or decision points, but a qualitative assessment conducted at key points in the risk management cycle. It is important to note that the standard is specifically not intended for purposes of certification. ISO 31000 alone cannot be considered sufficient for a risk assessment framework within an enterprise network, but frameworks have been designed to provide compliance with this standard [12] [13].

2) ISO27K Series: The ISO/IEC 27000 series is a large framework of best practices published by the ISO and the International Electrotechnical Commission (IEC). It provides a security control based qualitative framework with significant modularity for varying levels of implementation similar to the NIST RMF and COBIT. The strength of this model is its inherent ability to scale to the needs of the network, but allows for weaknesses where the framework is not fully implemented. Implementation is conducted through a six step qualitative process that assesses the current state of the network. Governance of the network is through the assignment of controls using a methodology similar to the NIST RMF. ISO 27K is a contemporary of the NIST RMF, COBIT 5, and other

qualitative networks which are the operational state of the art. It is currently in extensive use across the European Union [14] [15] [5].

D. Information Security Maturity Model (ISMM) (2011)

The ISMM model was created by analyzing eight existing models: NIST, Information Security Management Maturity Model (ISM3), Generic Security Maturity Model (GSMM), Gartner's Information Security Awarness Maturity Model (GISMM), SUNY's Information Security Initiatives (ISI), IBM Security Framework, Citigroup's Information Security Evaluation Maturity Model (ISEM), and Information Security Management System (ISMS) Maturity Capability Model. ISMM assesses the security requirements of an organization and then assigns a maturity level that will provide the correct balance of security and accessibility. They propose a method of quantifying risk at a very abstracted level, but the model itself is primarily a qualitative system to initiate compulsory levels of security [16].

E. Information Security Maturity Model (ISMM) (2017)

This ISMM model was also created following a comparison of several current implementations of risk modeling frameworks to include NIST RMF, COBIT, and ISO 27001. ISMM attempts to directly map each capability provided by current models to determine the most mature framework. The findings discovered weaknesses in all frameworks, and a single composite framework was introduced as a solution which provides all capabilities of currentimplementations in one system. The framework is still at a theoretical stage of implementation, but has the potential to create a more complete qualitative solution [5].

F. Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE)

1) OCTAVE (original): OCTAVE is a self directed risk management solution for large enterprises. It makes assumptions regarding the network staff's knowledge of critical systems and components to create a secure baseline. The weakness of this system is it is outdated (2003) and relies on having an expert team with significant resources. There have not been significant updates to OCTAVE following the release of OCTAVE-Allegro, and it could now be considered a legacy framework [17].

2) OCTAVE-S: OCTAVE-S is designed as a smaller scale implementation of OCTAVE, but suffers from several similar pitfalls. A manually created baseline that is updated as changes are observed cannot be easily adapted. OCTAVE-S provides additional structure for a less experienced team, but at the expense of significant system constraints as the implementation matures [17].

3) OCTAVE-Allegro: Allegro attempts to make risk management system more approachable than the original models. The complexity level of OCTAVE Allegro is lowered and the system is shifted to a more information-centric container based approach. Allegro is one of the first qualitative systems to incorporate an abstracted level of quantitative analysis using the containers as network elements. Due to the still largely qualitative nature of Allegro, it can have issues with implementation consistency. This can be especially challenging when accounting for IoT devices [18].

G. Holistic Cyber Security Implementation Framework (HCS-IF) (2014)

Atoum [19] introduces HCS-IF in an attempt to create a more complete approach to risk management that avoids the fragmented stovepipe nature that developed over several iterations of abstracted quantification in many risk management frameworks of the current state of the art. Frameworks that have used metrics of qualitative adherence to create a security score give some users the perceived confidence of a quantitative system without the overhead of a fully mapped risk framework. HCS-IF identifies core issues with bringing quantitative risk modeling back to cyber risk frameworks, but rather than create a fully quantitative methodology it attempts to advance the state of the art in qualitative models. The potential value added by their research must assessed in future studies before making any significant assertions of overall effectiveness [19].

1) HCS-IF Implementation Case Study (2017): The National Information Assurance and Cyber Security Strategy (NIACSS) of Jordan analyzed the HCS-IF in 2017 to determine if it could be applied at a national level. An implementation of HCS-IF is anticipated within the next three years following an evaluation by the Jordanian National Information Technology Center (NITC). Primary areas of improvement identified prior to adoption included change management and human resource issues [20].

H. IoT/M2M

Cisco introduces the IoT/M2M framework in order to address the rising challenge of securing networks saturated with relatively insecure IoT devices. The downside to this model is the cost and difficulty in building a network from essentially the ground up as opposed to introducing new security measures to an existing network. IoT/M2M employs a qualitative zero trust approach to security that attempts to limit the access of IoT devices in order to prevent them from being leveraged to influence otherwise secure devices. Live network evaluation has not been published. The proprietary nature of this framework significantly hinders any further testing in a research environment [21]. IoT/M2M builds a compelling argument for the success of the theoretical model and it may serve as the basis for future IoT security research.

I. Mobius

Mobius is a legacy framework included as an example of quantitative systems. It creates a quantifiable model which allows for risk calculations to be made using custom designed profiles for each device. The weakness of this methodology the poor scaling and implementation relative to more modern tools. As enterprise networks have grown in both size and diversity it is not financially advantageous to manual create a threat model for each device. This requires extensive expertise to properly employ, and additional development to account for IoT devices [22].

J. Online Services Security Framework (OSSF)

The OSSF framework is designed to manage risk in an enterprise network offering online services and remote access. It provides a structure or guide to create a secure baseline for both the provider and the consumer, but inherently must be configured by the end user. It accounts for highly mobile IoT devices, but it is currently limited in its application until it can be expanded to more diverse networks [23].

K. The CORAS Method

The CORAS approach is an 8 step model-based solution which allows a great deal of flexibility in implementation. It is built on the ISO 31000 standard for risk management as a self contained risk management solution for information technology systems. A risk evaluation matrix is populated using the CORAS tool that provides both high and low level analysis, but at the cost of significant labor as the baseline is constantly redefined when new assets are introduced. It uses a threat diagram to estimate risk based on past experience [24] [12].

L. Threat Agent Risk Assessment (TARA) (2009)

TARA was created by Intel and uses a calculation matrix to predict which agents pose the highest risk to the network. The output is then cross-referenced with known vulnerabilities and controls to mitigate risk. A meaningful published application of the TARA system has not been identified during this survey. TARA offers high levels of security, but the tradeoff is high operating costs. TARA attempts to bridge the gap between quantitative and qualitative systems, but the framework is not simplified enough to gain prominence [25].

M. Threat Assessment & Remediation Analysis (TARA) (2011)

The MITRE Corporation created the TARA system to secure specific networks known to be of interest to potential threat actors during the system design and acquisition phase. TARA uses a scoring model to identify probability of attack and potential attack vectors. It is difficult to scale, but can provide very sophisticated assessments if the cybersecurity budget is sufficiently large. This method attempts to create stovepipes that can be tracked and modeled quantitatively. TARA is designed to be used primarily as an assessment tool to establish a risk baseline before a more sustainable qualitative tool is employed for the operational phase of the network life-cycle [26].

N. CCTA Risk Analysis and Management Method (CRAMM)

CRAMM is a framework designed by the United Kingdom (UK) Central Computer and Telecommunications Agency (CCTA). It is a relatively outdated method of providing qualitative analysis across multiple asset groups and requires them to be built out on a per-network basis. This makes the modular construction useful, but at the cost of significant overhead to implement. It has been implemented in many countries, but has not been updated since CRAMM 5 in 2003 [27].

O. Cyber Assessment Framework (CAF) 2.0

Created by the UK National Cyber Security Centre (NCSC), the CAF is a model based risk assessment system similar to NIST RMF which provides extensibility across many devices and network types including SCADA [28]. The intent is to provide support from NCSC to adoption of the European Union (EU) Network and Information System (NIS) directive. The framework is new, without published academic assessment, but it has been adopted at an international level with a particular focus on SCADA and business IT systems. CAF is implemented through the 14 principles of cybersecurity and resiliency identified by the NCSC. CAF provides an approachable methodology, but does not yet have the validated technical controls of more mature qualitative frameworks [29].

1) CAF 3.0 Release: The release of CAF 3.0 makes no changes to the structure or technical content of the CAF, but replaces specific NIS Directive terminology with simpler language better suited to users outside the direct purview of NIS [30].

P. Cyber Risk Scoring and Mitigation (CRISM)

CRISM was developed in 2018 as an effort to reintroduce quantifiable metrics into cyber risk assessment in order to mitigate the information advantage of the network owner in cyber insurance applications. The model uses Bayesian graphs to build an end-to-end automated capability which can provide security scores and prioritized mitigation plans. The primary goal is to identify the exploitable attack surface of the network, and then to assess the risks of lateral propagation. With this information, a risk mitigation plan can be created and implemented. CRISM relies on network scanning tools to analyze the attack surface, which can struggle to detect IoT devices. The likelihood of device exploitation is based on CVSS to access the Common Vulnerability Exposures (CVE) library. The weakness in this method is that a CVE entry must exist for the vulnerability [31]. CRISM leverages a high level of automation to make implementation much simpler for small teams, but live network testing has not been published. Additional testing and development is necessary before CRISM is deployed to an enterprise network [32].

Q. Network Security Risk Model (NSRM)

NSRM relies on establishing a secure baseline and comparing risk levels after the introduction of each new device. This method is relatively outdated and labor intensive, but can provide good results if it is effectively implemented. It is targeted at Process Control Networks (PCN) which have less variance and is not suitable for a large enterprise network [33].

R. Cyber-Physical Systems Security (CPSS)

DiMase [34] identified the need for a Cyber-Physical System (CPS) centric risk framework to account for the rise in CPS devices across enterprise networks. It relies on a heuristics based approach rather than a secure baseline to provide an initial level of security, and over time creates an operational baseline. The model does not yet employ a holistic approach, but it is anticipated in future research and development. Additional standardization is also necessary in order to allow the framework to function across multiple domains. The concept has not yet been tested on a live netowrk. Despite the need for extensive future development, the framework attempts to solve many current issues with cyber-physical system security [34].

S. Harmonized Threat & Risk Assessment (HTRA)

Published by the Canadian Government, HTRA provides a risk management framework which expounds rapid adjustments to account for quickly evolving threat terrain, but still implements a traditional secure baseline structure. HTRA suffers from the same pitfalls of most large frameworks in that the size of the network often determines how effectively the model is implemented. HTRA follows the NIST model closely in an attempt to preserve scalability and consistency, but does not implement the rigorous controls used by RMF [35].

T. System-Fault Risk (SFR)

The qualitative framework created by Ye employs systems engineering, fault modeling, and risk assessment to classify cyber attacks. It accounts for several layers of interconnection by creating multiple attack origin classification models. The framework is modular and capable of extension into nearly any device that operates on a network, but at extreme cost. SFR takes the form of a checklist taxonomy which requires manual assessment and identification of devices in order to populate the risk matrices. It is not intended to be used as a full enterprise solution in its current form, but provides attack classification and characterization tools. Future research intends to provide further development toward a functional system [36].

U. Hierarchical Model Based Risk Assessment

Baiardi introduces a quantifiable framework based on security dependency hypergraphs which have the capability to identify attack paths which an analyst may miss in a qualitative assessment, but the model does not account for the inner state or operations of components. Risk is modeled and predicted within the graph. This allows for risk assessment and mitigation for each individual node or device. Tools for basic implementation were developed but not widely tested in a live network [37].

V. Patel & Ziveri Model

The model is a quantitative system which depends on predetermined types of attacks and devices to populate a risk matrix. This is accomplished by identifying the level of vulnerability each device has to each type of attack across several levels of effect. The model accounts for equipment loss, control loss, time loss, potential damage, and cost of prevention. A case study is performed in a small laboratory with several ICS devices. Additional research would be required in order to account for anything outside of the current scope of the model. It is presently designed for implementation in SCADA networks, and does not account well for IoT or any attack that is not within the matrix [38].

W. IBM Security Framework

The IBM security blueprint stovepipes security into domains which are broken down further into distinct objectives and services. The IBM model is specific to proprietary implementations of IBM hardware and products, but includes applications with devices from other vendors. Network sub-domains are defined by the framework in order to give the network managers sufficient segmentation for their environment. IBM relies heavily on operating according to industry best practices [39]. An update in 2014 showed successful results in several live networks [40].

1) Additional Publications (2016): IBM has published a series of books [41] to address practical application of the IBM security framework. They recognize the theoretical nature of the original publication [40] and introduce controls to assist in implementation of the framework. Each security domain is broken down into individual elements and appropriate security solutions are advocated. IoT devices are only accounted for through host and endpoint security measures and Access Control Lists (ACLs). The security model is simplistic, but operates at a level equivalent to current generation frameorks [41].

X. Information Security Risk Analysis Method (ISRAM)

ISRAM is an attempt to bridge the gap between the overwhelming challenge of implementing a quantitative model on a complex network and the inconsistencies of a qualitative model. While sound in theory, the product still suffers from the extensibility issues faces by quantitative models. It operates by using one of the fundamental risk calculations, a function of probability and consequence. ISRAM relies heavily on surveys to populate risk tables. The case study was limited to a 20 device Local Area Network (LAN). The primary weakness of ISRAM is that it is blind to risk that is not identified through the surveys [42].

Y. Cyber-Physical Security (CPS) Model

Amin [43] employs elements of game theory to estimate security risks using technology based security defenses grounded in information security tools and fault tolerant controls in an attempt to create a more quantitative framework to address the risks presented by cyber-phsyical systems on a network. The methodology struggles to account for all components simultaneously in a large composite model, and lacks extensibility. Amin argues that the inter-dependencies of cyber-physical systems is not well documented, and the risks they pose to an established network are not assessed accurately due to the lack of research in cyber-physical system vulnerabilities [43].

Z. Cybernomics

Cybernomics is an attempt to incorporate cyber risk management and economic modeling to build a more quantifiable framework which can be scaled to a larger enterprise network using a formally proposed unit of cyber risk. It provides a more network centric portfolio, and in turn may be capable of providing sound IoT accountability. This framework is reliant on large scale adoption as a means to populate common threat indexes and create informed risk models. Live network testing is anticipated in a future publication [44].

III. Methodology

Four primary elements common to each framework are evaluated. This establishes a basic standard used to make comparisons, and highlights several key differences between otherwise similar methods. These attributes are mapped and graded to determine the level of efficacy provided. It is challenging to conduct a full pairwise comparison between any two models due to their inability to target IoT devices specifically. Nearly all models surveyed neglected to take special measures towards securing IoT devices versus other enterprise components. Models which account for IoT/mobile/ICS often highlight that they are a security challenge, but do not have specific countermeasures in place to mitigate the threats they introduce. This led to a largely qualitative analysis of the merits of each model, with models that have a particularly outstanding system being highlighted in Section IV.

A. Quantitative vs. Qualitative

Each framework surveyed was classified as either primarily qualitative, or quantitative. The constraints of the quantitative model are similar to the strengths of a qualitative model, and vice versa. Quantitative models can provide unparalleled threat modeling at the expense of scalability. Popular methods of quantitative modeling require manual analysis of each device to identify network interfaces and operating systems. For the purpose of this assessment a framework must demonstrate device specific risk or attack probability considerations to be classified as quantitative. Frameworks employing specific architecture requirements, implementation controls, and vulnerability assessments were categorized as primarily qualitative. Any system that used a method of device abstraction for a quantitative analysis is classified as qualitative.

B. Level of Implementation

Models are assigned an enterprise network implementation score of high, low, or N/A in order to account for the broad range of real-world testing frameworks have received. It is considered irresponsible to recommend an untested framework for use in production networks prior to significant live testing. A framework with hundreds of implementations and years of feedback will similarly have more data points to evaluate than a network which is conceptual or in its first live network test. Many surveyed frameworks have not yet been employed in a significant capacity on a live network, but they are included in this survey. Untested frameworks are examined in order assess approaches that have been tried in previous research, or are on the cutting edge of risk management development.

C. Age and Support Level

Risk assessment frameworks which no longer have a robust implementation or supporting entity may no longer be viable. It is important to consider that legacy models may no longer provide adequate security, but they are important to consider when examining the current state of IoT adaptation. Several analyzed methodologies have been iterated over the course of years and decades. The version of a methodology selected for this paper is reflected by the date and any version release information discussed in Section II. When applicable, the individual publications are cited and referenced with the specific iteration selected for analysis.

D. Overall Rating

The current industry standard for a risk assessment framework is the a qualitative model. This method of assessment relies on robust security policy and patching processes alongside vulnerability scanning and security controls. Examples of these frameworks include the NIST RMF, NCSC CAF 3.0, and ISACA COBIT. These methods are suitable for securing a traditional enterprise network, but have weaknesses to IoT devices that are introduced without being fully incorporated to the baseline. Any framework that meets, but does not have the potential to exceed the current state of the art implementation is rated “Yellow”. Yellow rated models are a relatively good assessments of cyber risk, but they do not manage IoT devices well. Any framework which is unable to achieve the same level of network protection as the current generation of frameworks is rated “Red”. Models which have made a meaningful step towards properly accounting for IoT devices within enterprise networks will be rated “Green”. Several methodologies rated green have not been fully deployed in a live test, but have demonstrated that they manage IoT devices with a higher level of effectiveness.

IV. Analysis of Risk Assessment Frameworks

A live test and assessment of each risk model is beyond the scope of this survey. Each selected methodology is broken down according to the criteria outlined in Section III. The assessment of each framework allows for comparison across methodology, age, implementation level, and effectiveness rating. This breakdown is introduced in Table I.

TABLE I. RISK FRAMEWORK COMPARISON

Reviewed Framework	Framework Analysis		
	Rating	*Implementation*	*Year*
†CAF [28]	Yellow	High	2018
†COBIT 5 [14][9]	Yellow	High	2012
†CORAS [45]	Red	Low	2003
*CPS Model [43]	Red	N/A	2013
†CPSS [34]	Red	N/A	2015
*CRAMM [27]	Red	Low	2003
*CRISM [32]	Green	N/A	2018
*Cybernomics [44]	Green	N/A	2017
†HCS-IF [19]	Green	N/A	2014
†*Hierarchical Model[37]	Red	N/A	2009
†HTRA [35]	Yellow	High	2007
†IBM Framework [39]	Yellow	Low	2010
†IoT/M2M [21]	Green	N/A	2016
†ISO27K [14][15]	Yellow	High	2005
†ISO31K [13]	Yellow	High	2009
*ISRAM [42]	Red	N/A	2005
†ISSM [5]	Green	N/A	2017
†ISSM [16]	Yellow	Low	2011
*Mobius [22]	Red	N/A	2002
†NIST CSF [7]	Yellow	High	2014
†NIST RMF [46]	Yellow	High	2015
*NSRM [33]	Red	N/A	2009
†OCTAVE [17]	Red	Low	2003
†OCTAVE-S [17]	Red	Low	2003
†OCTAVE-Allegro [18]	Red	Low	2007
†OSSF [23]	Green	N/A	2017
*Patel & Ziveri Model [38]	Red	N/A	2010
†SFR [36]	Red	N/A	2005
†*TARA (Intel) [25]	Yellow	Low	2009
†*TARA (MITRE) [26]	Yellow	Low	2011

†Indicates Qualitative *Indicates Quantitative

A. Common Framework Pitfalls

Initial assessment standards required a significant implementation instance in order to merit a "green rating", but no surveyed models with production implementation were designed to account for IoT devices. This requirement was removed as a result each model that rated "green" for IoT advancement has not been implemented in a live network. Similarly, all models rated "high" for implementation scored "yellow" in IoT advancement. This overwhelmingly indicates that the state of the art has not yet accounted for IoT properly, and no single framework can be recommended as an immediate solution to the IoT problem. The current model of a qualitative risk assessment may no longer be viable as IoT devices continue to become more critically integrated into networks. Each qualitative model surveyed attempts to use only existing resources to secure the IoT threat vector. In order to continue using existing risk models, it is necessary to either invest in new risk assessment architecture to account for the largely unknown vulnerabilities presented by current off the shelf IoT systems, or incorporate only IoT systems which have been subjected to a much higher degree of security analysis. The current model of minimal support and small device market share footprint is unsustainable if security is to be prioritized.

B. IoT Advancements

It is imperative that security development be proactive due to the increasingly vital role that IoT devices have in enterprise networks. Among the most promising proposed models is the zero trust approach in the IoT/M2M framework. Rather than attempt to impose enterprise security methods on IoT devices, it attempts to section them off as much as possible into other network segments. This is not a full solution, but it may prove more effective than current implementations. The frameworks that have the ability to accurately model risks to ICS and IoT systems have primarily implemented a quantitative risk assessment approach, but no solution has been able to provide cost-effective coverage to a larger network. Most quantitative models draw from the CVE database, which is reliant on vulnerability publications. Due to the obscurity of IoT systems, many face less rigorous assessment and have fewer published CVE findings. The primary weakness to this solution is some devices will eventually have to have a trusted relationship, and this will lead to inevitable unmitigated vulnerabilities. This method is at best a technique to shrink the attack surface of a network, and does not fully mitigate the risk of IoT devices.

C. Proposed Solutions

Two courses of action for securing IoT devices based on the analysis of the 30 frameworks surveyed are proposed based on short term and long term research goals. The trend of predominately quantitative risk assessment frameworks in early models was primarily rendered obsolete due to implementation costs rather than level of effectiveness. A short term approach focused on bolstering the IoT specific security controls of qualitative methods is recommended based on current developments in IoT and ICS security best practices. The long term approach recommended by this paper is based on reintroducing elements of quantitative risk assessment and mitigation models through the use of Artificial Intelligence (AI) solutions designed to perform risk modeling and attack probability extrapolation.

1) Short Term: Use network segmentation and a zero trust model: IoT devices cannot be considered trusted or secure by a risk analysis model until a more robust vulnerability assessment process can be developed. IoT and ICS devices both utilize interfaces which are not assessed by most current enterprise network vulnerability assessment tools. Physical access on remote devices must also be considered by a risk methodology. Designing network architecture to create the smallest foothold possible for compromised IoT devices may be an effective short term solution, but would need to be accompanied by policy and control updates. Potential examples of this would include creating requirements

to implement an IoT device Virtual Local Area Network (VLAN), De-Militarized Zone (DMZ), or using bastions as IoT interface servers. Similarly, isolating IoT devices from domain credentials and trust settings is vital to ensuring that a vulnerable IoT device does minimized damage if exploited. Due to the inherent hidden vulnerabilities in many IoT devices, the threat of lateral attack propagation is extremely high. These strategies focus on limiting an attackers influence in the event that they do gain access to a device. This strategy has been well documented and proposed in several IoT risk management models, but have not been implemented at the scale of a large enterprise in any research studies. Models such as the Cisco IoT/M2M [21] provide an overview of this concept. The focus of the network security controls is placed on regulating and limiting the level of interaction a device can have with other elements of the network.

2) Long Term: Increase viability of quantifiable risk assessment frameworks with Machine Learning: Quantitative frameworks have demonstrated the highest level of accuracy when employed to assesses cyber risk, but are not capable of modeling large networks in their present state. The next iteration of quantitative framework research, currently underway, relies on existing CVE score data to calculate risk, and requires significant oversight to operate. This model still suffers from the scalability issues observed in past threat-quantification based methodologies. This problem must be solved in order for quantifiable frameworks to become viable.

Potential methods for achieving this could include the use of machine learning (ML) in order to implement risk classification and develop individual device profiles. This direction requires significant future research with live testing and development, but could yield lower operating costs when applied at an enterprise level. Building the threat profile and identifying logical/physical location of a device are currently the areas that reliant on the effectiveness of a human input to the system. Creating a method capable of employing passive device detection automatically adjusted to compensate for the additional network systems offers significantly higher reliability at the cost of adding nodes to each subnet. This increases reliance on initial configuration, rather than reliance on network data inputted through survey. Additional scanning tools would be necessary to provide oversight of external network interfaces created by IoT devices similar to proposed solution 1).

ML Tasks typically fall into two categories: regression and classification. Regression involves predicting a real-valued output while classification involves predicting a categorical value [47]. A regression task that could be applied for cyber risk frameworks is to predict values of risk using inputs like those that go into the CVE score along with other risk features. Using these features as input, an ML algorithm could be applied to predict risk values much in the same way as the CVE score. This system would also allow for very accurate projections of security level in proposed architecture developments, as well as software migrations and patching. A classification approach could be applied in coupling with items such as an Intrusion/Anomaly Detection System. The IDS can monitor traffic and create traffic profiles and then they can be fed in as inputs. Using these features, a classification of risk level could be made using a classification algorithm such as Support Vector Machines, Logistic Regression or Random Forest.

Using ML for risk classification and device profiles would require a multi-level approach. For developing device profiles, a classification task could be applied to classify the traffic for each device. With these classifications, then, using a separate ML algorithm, risk level could be classified using the device profiles and passive network traffic such as Snort logs [48]. Coupling this with an input such as CVE scores for known vulnerabilities visible in the traffic could allow for classification of successive levels of risk. Regardless of the ML algorithm used, an approach such as this would require a significant amount of time and data to be useful. The data would also have to be labelled so as to be useful for training and testing an ML algorithm. Thus, this would not be a quick solution, but could be quite powerful if implemented.

V. Conclusion

The assessment of 30 cyber risk assessment frameworks shows significant shortcomings in all state of the art risk methodologies. No developmental model was identified that could be considered deployment ready with capabilities clearly exceeding those of the current generation of qualitative system. Developmental models with the ability to incorporate both cyber-physical systems and enterprise architecture in a large scale network were reviewed, but none have been tested in a live environment. At this time, there is still a significant need for research on methods to incorporate IoT devices into enterprise networks while maintaining necessary levels of accessibility balanced with security. The scale and diversity of IoT has been insurmountable for qualitative models, but future research developing Proposed Solution 1). may yield significant advancements that do not require substantial changes in architecture. At this time there is not a methodology shown to be able to quantify the additional risk presented by IoT devices. A significant change in funding or advancement in implementation methods will be necessary in order to drastically alter the current risk assessment terrain away from qualitative models. Minimal published research on the application of machine learning to cyber risk assessment was identified, but this avenue of research outlined in Proposed Solution 2). offers a potential way forward to make the quantitative model viable again. The development of quantifiable risk methodologies is well regarded, but most current research avenues are still reliant on known vulnerabilities. Additional research in IoT vulnerability assessment is needed in order to accurately populate the risk matrices employed by most proposed quantified frameworks.

ACKNOWLEDGMENT

Disclaimer: The views expressed in this paper are those of the author and do not necessarily reflect the official policy or position of the Air Force, the Department of Defense, or the U.S. Government.

REFERENCES

[1] A. J. Pendleton, D. Pettit, and R. Dill, "Surveying the incorporation of IoT devices into cybersecurity risk management frameworks," *SECURWARE 2019, The Thirteenth International Conference on Emerging Security Information, Systems and Technologies*, pp. 128—-133, 2019.

[2] Government Accountability Office, "Internet of things: Enhanced assessments and guidance are needed to address security risks in dod," *Publication No. GAO-17-668*, 2017.

[3] M. Conti, A. Dehghantanha, K. Franke, and S. Watson, "Internet of things security and forensics: Challenges and opportunities," 2018.

[4] K. L. Dempsey, G. A. Witte, and D. Rike, "Summary of nist sp 800-53, revision 4: Security and privacy controls for federal information systems and organizations," Tech. Rep., 2014.

[5] S. Almuhammadi and M. Alsaleh, "Information Security Maturity Model for NIST Cyber Security Framework," *Computer Science & Information Technology*, vol. 51, 2017.

[6] National Institute of Standards and Technology. (2016) Risk management framework steps. Last Accessed 2019-12-2. [Online]. Available: https://csrc.nist.gov/CSRC/media/Projects/Risk-Management/images-media/NIST-RMF.png

[7] N. Keller, "NIST Cybersecurity Framework (CSF) Reference Tool," 2014.

[8] M. Ahlmeyer and A. M. Chircu, "Securing the internet of things: A review." *Issues in information Systems*, vol. 17, no. 4, 2016.

[9] K. Wal, J. Lainhart, and P. Tessin, "A cobit 5 overview," 2012.

[10] J. Lainhart, "Introducing COBIT 2019: The Motivation for the Update?" 2018.

[11] ISACA, "Cobit 2019 framework: Introduction and methodology," 2018.

[12] M. S. Lund, B. Solhaug, and K. Stølen, *Model-driven risk analysis: the CORAS approach*. Springer Science & Business Media, 2010.

[13] M. Leitch, "ISO 31000: 2009—The new international standard on risk management," *Risk Analysis: An International Journal*, vol. 30, no. 6, pp. 887–892, 2010.

[14] W. Al-Ahmad and B. Mohammad, "Can a single security framework address information security risks adequately," *International Journal of Digital Information and Wireless Communications*, vol. 2, no. 3, pp. 222–230, 2012.

[15] T. Humphreys, "State-of-the-art information security management systems with iso/iec 27001: 2005," *ISO Management Systems*, vol. 6, no. 1, 2006.

[16] G. Karokola, S. Kowalski, and L. Yngström, "Towards an information security maturity model for secure e-government services: A stakeholders view." in *HAISA*, 2011, pp. 58–73.

[17] C. Alberts, A. Dorofee, J. Stevens, and C. Woody, "Introduction to the OCTAVE Approach," Carnegie-Melon Univ Pittsburgh PA Software Engineering Inst, Tech. Rep., 2003.

[18] R. Caralli, J. F. Stevens, L. R. Young, and W. R. Wilson, "Introducing OCTAVE allegro: Improving the information security risk assessment process," 2007.

[19] I. Atoum, A. Otoom, and A. A. Ali, "Holistic cyber security implementation frameworks: A case study of jordan," *International Journal of Information, Business and Management*, vol. 9, no. 1, p. 108, 2017.

[20] ——, "Holistic cyber security implementation frameworks: A case study of jordan," *International Journal of Information, Business and Management*, vol. 9, no. 1, p. 108, 2017.

[21] "Cisco: Securing the internet of things: A proposed framework." 2016.

[22] D. D. Deavours, G. Clark, T. Courtney, D. Daly, S. Derisavi, J. M. Doyle, W. H. Sanders, and P. G. Webster, "The mobius framework and its implementation," *IEEE Transactions on Software Engineering*, vol. 28, no. 10, pp. 956–969, 2002.

[23] J. Meszaros and A. Buchalcevova, "Introducing ossf: A framework for online service cybersecurity risk management," *computers & security*, vol. 65, pp. 300–313, 2017.

[24] M. S. Lund, B. Solhaug, and K. Stølen, "A guided tour of the CORAS method," in *Model-Driven Risk Analysis*. Springer, 2011, pp. 23–43.

[25] M. Rosenquist, "Prioritizing information security risks with threat agent risk assessment," *Intel Corporation White Paper*, 2009.

[26] J. Wynn, J. Whitmore, G. Upton, L. Spriggs, D. McKinnon, R. McInnes, R. Graubart, and L. Clausen, "Threat assessment & remediation analysis (tara): Methodology description version 1.0," MITRE CORP BEDFORD MA, Tech. Rep., 2011.

[27] Z. Yazar, "A qualitative risk analysis and management tool–cramm," *SANS InfoSec Reading Room White Paper*, vol. 11, pp. 12–32, 2002.

[28] United Kingdom National Cyber Security Centre, "Cyber assessment framework," 2020.

[29] T. Kevin, "Introducing the cyber assessment framework v2.0," 2018.

[30] ——, "Introducing the cyber assessment framework v3.0," 2019.

[31] Government Accountability Office, "Internet of things: Status and implications of an increasingly connected world," *Publication No. GAO-17-75*, 2017.

[32] S. Shetty, M. McShane, L. Zhang, J. P. Kesan, C. A. Kamhoua, K. Kwiat, and L. L. Njilla, "Reducing informational disadvantages to improve cyber risk management," *The Geneva Papers on Risk and Insurance-Issues and Practice*, vol. 43, no. 2, pp. 224–238, 2018.

[33] M. H. Henry and Y. Y. Haimes, "A comprehensive network security risk model for process control networks," *Risk Analysis: An International Journal*, vol. 29, no. 2, pp. 223–248, 2009.

[34] D. DiMase, Z. A. Collier, K. Heffner, and I. Linkov, "Systems engineering framework for cyber physical security and resilience," *Environment Systems and Decisions*, vol. 35, no. 2, pp. 291–300, 2015.

[35] Government of Canada, "Harmonized Threat and Risk Assessment (HTRA) Methodology," 2007.

[36] N. Ye, C. Newman, and T. Farley, "A system-fault-risk framework for cyber attack classification," *Information Knowledge Systems Management*, vol. 5, no. 2, pp. 135–151, 2005.

[37] F. Baiardi, C. Telmon, and D. Sgandurra, "Hierarchical, model-based risk management of critical infrastructures," *Reliability Engineering & System Safety*, vol. 94, no. 9, pp. 1403–1415, 2009.

[38] S. Patel and J. Zaveri, "A risk-assessment model for cyber attacks on information systems," *Journal of computers*, vol. 5, no. 3, pp. 352–359, 2010.

[39] A. Buecker, M. Borrett, C. Lorenz, and C. Powers, "Introducing the IBM security framework and IBM security blueprint to realize business-driven security," *IBM Redpaper*, vol. 4528, no. 1, pp. 1–96, 2010.

[40] A. Buecker, S. Arunkumar, B. Blackshaw, M. Borrett, P. Brittenham, J. Flegr, J. Jacobs, V. Jeremic, M. Johnston, C. Mark *et al.*, *Using the IBM Security Framework and IBM Security Blueprint to Realize Business-Driven Security*. IBM Redbooks, 2014.

[41] A. Buecker, B. Chakrabarty, L. Dymoke-Bradshaw, C. Goldkorn, B. Hugenbruch, M. R. Nali, V. Ramalingam, B. Thalouth, J. Thielmann *et al.*, *Reduce Risk and Improve Security on IBM Mainframes: Volume 1 Architecture and Platform Security*. IBM Redbooks, 2016.

[42] B. Karabacak and I. Sogukpinar, "Isram: information security risk analysis method," *Computers & Security*, vol. 24, no. 2, pp. 147–159, 2005.

[43] S. Amin, G. A. Schwartz, and A. Hussain, "In quest of benchmarking security risks to cyber-physical systems," *IEEE Network*, vol. 27, no. 1, pp. 19–24, 2013.

[44] K. Ruan, "Introducing cybernomics: A unifying economic framework for measuring cyber risk," *Computers & Security*, vol. 65, pp. 77–89, 2017.

[45] K. Stolen, F. den Braber, T. Dimitrakos, R. Fredriksen, B. A. Gran, S.-H. Houmb, M. S. Lund, Y. Stamatiou, and J. Aagedal, "Model-based risk assessment-the coras approach," in *iTrust Workshop*, 2002.

[46] R. S. Ross and L. A. Johnson, "Guide for applying the risk management framework to federal information systems: A security life cycle approach," Tech. Rep., 2010.

[47] G. James, D. Witten, T. Hastie, and R. Tibshirani, "Classification," in *An Introduction to Statistical Learning*. New York: Springer Science & Business Media, 2013, ch. 4, pp. 127–173.

[48] M. Roesch, "Snort-Lightweight Intrusion Detection for Networks," in *Proceedings of LISA '99*, 1999, pp. 229–238.

On Heterogeneity of Management and Orchestration Functional Architectures in 5G Slicing

Eugen Borcoci, Cosmin Contu, Andra Ciobanu
University POLITEHNICA of Bucharest - UPB
Bucharest, Romania
Emails: eugen.borcoci@elcom.pub.ro, cosmin.contu@elcom.pub.ro, andraciobanu90@yahoo.com

Abstract — **Management and orchestration functionalities are crucial activities in 5G slicing systems. Essentially, the basis is the integrated framework Management and Orchestration (MANO), of The European Telecommunications Standards Institute but enriched, in order to cope with slicing. In particular, supporting technologies like Network Function Virtualization and Software Defined Networks are considered, to deliver functional components, cooperating for 5G slicing management and orchestration. The multi-tenant, multi-domain, multi-operator, end-to-end features of the 5G slicing determine a high complexity for management and orchestration. Consequently, many different architectural variants have been already proposed, studied and developed in recent studies, standards and projects. This study is useful because, despite many efforts (spent in the last five years), much heterogeneity and different solutions still exist, even at the management and orchestration architectural level. However, the MANO is considered as a base starting point architecture. This paper is an extension of a previous study. It analyzes in more depth the existing common parts, differences and heterogeneity of several management and orchestration 5G slicing architectures, identifying the similar functionalities and also the factors leading to heterogeneity.**

Keywords — 5G slicing; Management and Orchestration; Software Defined Networking; Network Function Virtualization;. Service management; Resource management.

I. INTRODUCTION

This paper is an extension of a previous paper published at IARIA AFIN Conference, 2019 [1], dealing with management and orchestration of 5G sliced systems.

The emergent 5G mobile network technologies offer powerful features, in terms of capacity, speed, flexibility and services, to answer the increasing demand and challenges addressed to communication systems and Internet [2][3]. 5G can provide specific types of services to satisfy simultaneously various customer/tenant demands, in a multi-x fashion (the notation –x stands for: tenant, domain, operator and provider).

The 5G network *slicing concept* (based on virtualization and softwarization) enables programmability and modularity for network resources provisioning, adapted to different vertical service requirements (in terms of bandwidth, latency, mobility, etc.) [3]-[7]. In a general view, a *Network Slice* (NSL) is a managed logical group of subsets of resources, Physical/Virtual network functions (PNFs/VNFs), placed in the architectural Data Plane (DPl), Control Plane (CPl) and Management Plane (MPl). The slice is programmable and has the ability to expose its capabilities to the users.

Network Function Virtualization (NFV) [8]-[10] *and Software Defined Networks (SDN)* are two powerful technologies, which offer the basis for softwarization and virtualization. They are considered as cooperating tools [11] to manage and control the 5G sliced environment, in a flexible and programmable way.

Management and Orchestration (M&O) is a crucial subsystem in NFV framework and also in 5G. Such topics constitute the object of standardization organizations and forums among which the 3rd Generation Partnership Project (3GPP), the 5G Infrastructure Public Private Partnership (5G PPP), and European Telecommunications Standards Institute (ETSI) are representative [12]-[17]. Given the complexity of 5G systems, the above organizations cooperate in order to harmonize their specifications. For instance, the 3GPP-defined management system interacts with ETSI's NFV MANO system to enable the resource management for virtualized Core Network (CN), virtualized Radio Access Network (RAN) and network slicing. ETSI collaboration with 3GPP – especially the Service and System Aspects Fifth (SA5) Working Group – is a key throughout the specification work of both ETSI NFV Releases 2 and 3, to ensure interoperability between management systems.

ETSI NFV has recently designed new features to support 5G networks. 5G resource M&O aspects were added on top of the NFV Release 2 framework. New NFV Release 3 [10] topics related to 5G includes: "Support for network slicing in NFV", "Management over multi-administrative domains", and "Multi-site network connectivity". These features are essential to address the variety of applications and services expected to run on top of a 5G system, while using in a distributed way resources over single or multiple sites, or in centralized or a combination of both.

However, it is recently recognized that a complete understanding of the relationship of a M&O system and a slicing system is still missing [3]. Even more, there is not yet a general/common agreement on the slice definition itself; several definitions exist, having major impacts and relationships to the M&O.

In the simplest view, a slice is a service with resource guarantees. In such a case, the slicing system and the orchestration system are identical. At the other end of the spectrum, a slice is a complex entity, i.e., a collection of

resources (computing, networking, storage) – that constitute a virtual logical network (customizable), working on top of a physical networking infrastructure. Inside such a slice, the slice owner/tenant has partial or even full freedom to enforce its own management and control (M&C) policies and actions. Consequently, each slice will have its own M&O. Many studies and standards adopted the above complex definition of a slice; *this is also considered in this work*, given the high flexibility that it can offer to the tenants. On the other hand, the complex structure of such a slice induces M&O complexity and *leads to a large variety of possible architectural approaches.*

Given many architectural proposals, there is an interest to evaluate in what degree they have similar approaches of the main "core" architectural functional set of blocks and what are the factors that induce heterogeneity. The similarity degree of different architectures could be named "convergence", although this word has usually a richer semantic.

Among many architectural aspects, the focus of this paper is on M&O sub-systems. Due to space limitation, this text cannot afford to offer detailed explanations about the architectures presented; the objective here is to identify the major point of similarity and heterogeneity of different approaches.

Therefore, this paper is mainly an overview and analysis type. Its structure is described below. Section II outlines the stakeholder/actors roles, given that such definitions determine essentially the overall system architecture. Section III evaluates whether a core unified view exists at architectural level, expressed in so-called *meta-architecture*. Section IV performs an analysis of some factors that lead to heterogeneity of the refined M&O architectures. Section V summarizes a few relevant examples extracted from various studies and projects, to illustrate the heterogeneity of solutions. Section VI presents the conclusions and possible future work.

II. Business Model and Stakeholder Roles

The layered structure of the 5G slicing M&O strongly depends on the definition of the *business model (BM),* which defines the stakeholder/actors roles and their interactions. Different BMs aim to support multi-tenant, multi-domain end-to-end (E2E) and multi-operator capabilities. A basic model (see A. Galis, [18]) defines four roles:

Infrastructure Provider (InP): owns and manages the physical infrastructure (network/cloud/data center). It could lease to a slice provider its infrastructure (connectivity, computing and storage resources) as they are, or it can itself construct slices and then lease the infrastructure in a network slicing fashion.

Network Slice Provider (NSLP): can be, typically, a telecommunication service provider (owner or tenant of the infrastructures from which network slices are constructed). The NSLP can construct multi-tenant, multi-domain slices, on top of infrastructures offered by one or several InPs.

Slice Tenant (SLT): is the generic user of a specific slice, including network/cloud/data centers, which can host customized services. An SLT can request from a NSLP to create a new customized slice instance. The SLT can lease virtual resources from one or more NSLP in the form of a virtual network, where the tenant can assemble, manage and then provide *Network Services* (NS) to its individual end users. A NS is a composition of *Network Functions (NFs*), physical or virtual, defined in terms of the individual NFs and the mechanism used to connect them. A single tenant may have one or several slices in its domain.

End User (EU): consumes (part of) the services supplied by the slice tenant, without providing them to other business actors.

The InP, NSLP and SLT have, each one, a specific role in M&O activities. A powerful feature of the above business model is the recursivity (see Ordonez et al., [4]), i.e., a tenant can become itself a new slice provider; at its turn it can offer parts of its sliced resources to other tenants. Other variants of business models are presented in [18].

Several recent Public Private Partnership (PPP) Phase I/II collaborative research projects are running, having as objectives 5G technologies [18]. Some of them extended the list of role definitions to allow various possible customer-provider relationships between verticals, operators, and other stakeholders. The 5G PPP Architecture Working Group, "View on 5G Architecture", Version 3.0, June 2019, [3] has defined a more refined business model (Figure 1):

Service Customer (SC): uses services offered by a Service Provider (SP). The vertical industries are considered as typical examples of SCs.

Service Provider (SP): generic role**,** comprising three possible sub-roles, depending on the service offered to the SC: *Communication SP* offers traditional telecom services; *Digital SP* offers digital services (e.g., enhanced mobile broadband and IoT services) to various verticals; *Network Slice as a Service* (NSaaS) *Provider* offers an NSL and its services. The SPs have to design, build and operate high-level services, by using aggregated network services.

Network Operator (NOP): orchestrates resources, potentially coming from multiple *virtualized infrastructure providers* (VISP). The NOP uses aggregated virtualized infrastructure services to design, build, and operate network services that are offered to SPs.

Virtualization Infrastructure SP (VISP): offers virtualized infrastructure services and designs, builds and operates virtualization infrastructure(s) (networking and computing resources). Sometimes, a VISP offers access to a variety of resources by aggregating multiple technology domains and making them accessible through a single *Application Programming Interface (API).*

Data Center SP (DCSP): designs, builds, operates and offers data center services. A DCSP differs from a VISP by offering "raw" resources (i.e., host servers) in rather centralized locations and simple services for consumption of these raw resources.

The hierarchy of this model (in the top-down sense of a layered architecture) is: SC, SP, NOP, VISP, DCSP. Note that in practice, a single organization can play one or more roles of the above list.

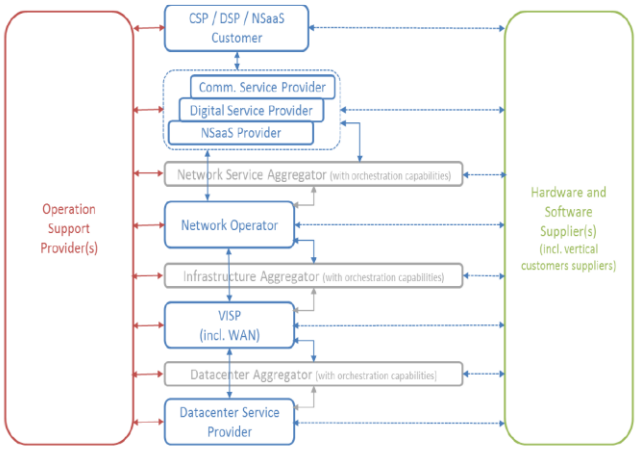


Figure 1. 5G PPP Business model [3]

III. A GENERIC 5G MANAGEMENT META-ARCHITECTURE

The analysis of many architectural proposals (in 5G and in particular, in 5G slicing) leads to the question: *is there any high-level consensus core architecture?* Recently, the document [3], authored by 5G PPP Architecture Working Group has identified a set of requirements for a consensus/meta 5G high-level architecture (collecting some M&O fundamental functionalities). The identified features are general for 5G and in particular applicable also to the slicing approach. This M&O architecture should be able to support:

a. individual control of NFs (number of instances to be created, their distribution/placement, deployment of an execution environment, start/stop the instances, management of the instances' states, etc.).

b. chaining of individual NFs into services (NF graphs, see [8]) facilitated by different control mechanisms at network level (e.g., the NFs chaining can be SDN – controlled - where the NFs are treated by the SDN controller like SDN forwarding nodes).

c. different underlying execution environments: various virtualization techniques (virtual machines (VM), containers, or plain processes) in clusters of different sizes (from a CPU board to an entire large-scale data center) over different, specialized "technological domains" - i.e., from some simple hardware, up to complex networking environments (wireless, optics, cable).

d. working across different "organizational", or administrative domains, i.e., owned by network operators or companies and using various business models (e.g., network operators can be separated from cloud infrastructure operators). Multiple operators and multi-domains operation are also a target, in order to provide services at vast geographic ranges.

e. a large range of applications with different specific requirements (in terms of resources, deployment, orchestration and optimization goals).

f. subdivision of the infrastructure in logical, separated and isolated slices – while offering different levels of guaranteed performance to their tenants.

Note that slicing capabilities – could be seen as part of a M&O system. However, there is no general consensus on this inclusion. There are also proposals to position a slicing system underneath or above a MANO system.

Several core roles of the involved entities have emerged from the above requirements: end user, function developer, application developer, validation and verification entity, tenant (owner of applications), operator (not necessarily encompassing slicing operator) infrastructure provider (network, cloud), etc., [3]. These can be mapped onto the roles described in Section II. Overlaps can exist between some of the above. Also, the mapping of the above roles on real organizations roles is flexible.

The requirements listed above actually drive the definition of the so-called M&O *meta-architecture*, in the sense that no matter how the particular architectural solution will be chosen, the six functionalities should be included. These define a general level of convergence from an architectural point of view. A particular architecture will be a refinement of the meta-one.

Another general aspect is related to the different time scales of different operations. One can distinguish between *"orchestration"* and *"control"* actions. The first are mid-long-time scales operations, relatively heavy-weight (e.g., optimization of the overall structure of a service, group of services, or slices). The second class comprises short time scales operations (e.g., light-weight operations, flow routing, etc.). We defend here the idea that such a logical separation should exist (it is natural) between functional elements performing the orchestration, w.r.t. those dedicated to control; however, in different refinements of the meta-architecture this separation is not quite obvious; this, again, leads to heterogeneity of approaches.

The basic framework for a high-level meta-architecture is offered by ETSI NFV (Figure 2) [8]. This has been defined as a general framework, before the 5G slicing concepts emerged. However, NFV Management and orchestration (NFV MANO) has been soon considered, by the standardization organizations, operators and research groups, as being appropriate to further develop M&O for 5G sliced systems.

The main M&O blocks are: the NFV Orchestration (NFVO), VNF Manager (VNFM) and Virtual Infrastructure Manager (VIM). If the principle of separation between the orchestration and control is applied, then the specific network configuration tasks (e.g., connectivity - related) can be outsourced to a separate SDN controller, working under command of the NFVO. An alternative could be to split the NFVO into two parts – orchestrator and controller.

We recall shortly the roles of the basic NFV functional blocks [9]:

NFV Orchestrator (NFVO) has two main responsibilities:

- the lifecycle management of Network Services (NS); thus, it fulfills the *Network Service Orchestration* functions;
- the orchestration of NFVI resources across multiple VIMs; thus, it fulfills the *Resource Orchestration (RO)* tasks.

VNF Manager (VNFM) is responsible for the lifecycle management (LCM) of VNF instances;

Virtualized Infrastructure Manager (VIM) is responsible for managing and controlling the NFVI resources, i.e., compute, storage and network resources.

To provide a more complete architectural assembly the following functional blocks may also be considered:

Operation/Business System Support (OSS/BSS) represents the combination of the operator's other operations and business support functions that are not otherwise explicitly captured in the architectural diagram. An Element Manager (EM) is responsible for traditional management functionality (fault detection, configuration, accounting, performance, security- FCAPS) for a VNF.

The NFVI represents all the hardware (e.g., compute, storage, and networking) and software (e.g., hypervisors) components that together provide the infrastructure resources where VNFs are deployed.

The NFV framework is added with new functions in order to support slicing (see Figure 2). The slicing support feature introduces significant differentiation between particular architectures. The slice management:

- could be included into the NFVO (because a network slice instance (NSLI) can be seen, in a simpler approach, as a guaranteed network service);
- or, a separate slice manager exists (controlled by NFVO). Note that the service management can be defined as separated from resource management (this option provides a cleaner architecture), or they can be treated together.

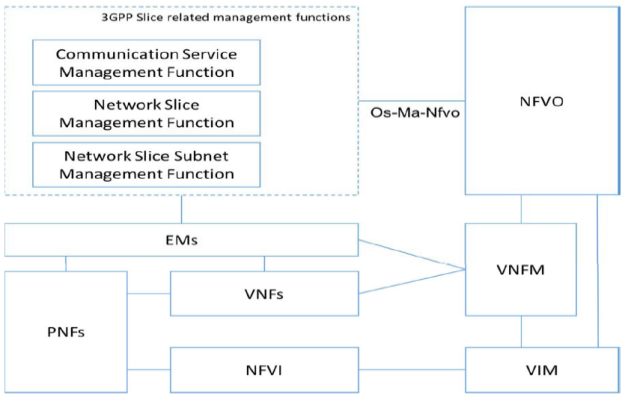


Figure 2. Network slice management in an NFV framework (ETSI GR NFV-EVE 012 V3.1.1, [15])

NFV -Network Function Virtualization; EM - Element Manager; MANO - Management and Orchestration (NFVO – NFV Orchestration; VNFM – VNF Manager; VIM Virtual Infrastructure Manager); VNF/PNF – Virtual/Physical Network Function; NFVI -NFV Infrastructure; NS-Network Service; OSS-Operations Support System.

In multiple domain cases, the NFVOs should federate in some form with peer NFVOs, placed in a single or in multiple organizations. In some approaches, a hierarchy of service management instances is developed, having on top a multi-domain manager (working at abstract level) and single-domain managers at lower level. The latter should perform also peer interactions.

A typical set of functional M&O blocks for a single-domain meta-architecture is [3] (the levels are top-down ordered): [Service management, Orchestrator, (MANO controller, SDN controller), VIM, Resources]. In a multi-domain environment, each domain should have the previous set and on top of them a multi-domain service manager should exist. Note that inter-domain (horizontal) peer interactions must exist between peers (e.g., Orchestrator_X <---> Orchestrator_Y).

The basic 5G slicing high level architecture proposed by ETSI [15] (Figure 2), can be considered as a meta-architecture comprising the six features exposed above. To the original ETSI NFV architecture [8][9], several new functional blocks have been added in order to support the network slicing (ETSI-NFV EVE 012 [15]).

The 3GPP TR 28.801 document [16] defines three new management functions:

- *Communication Service Management Function* (CSMF) – it translates the communication service requirements to NSL requirements;
- *Network Slice Management Function* (NSMF) - responsible for the management (including lifecycle of instances) of NSLIs (it derives network slice subnet requirements from the network slice related requirements);
- *Network Slice Subnet Management Function* (NSSMF) - responsible for the management (including lifecycle) of *Network Slice Subnet Instances* (NSSIs).

An interface is defined, i.e., Os-Ma-NFVO Reference Point (RP) with ETSI NFV-MANO. To interact in an appropriate way with NFV-MANO, the NSMF and/or NSSMF need to determine the type of network service or set of network services, the VNFs and PNFs that can support the resource requirements for a NSLI or NSSI. Consequently, one should determine whether new instances of these NSs, VNFs and the connectivity to the PNFs need to be created, or existing instances can be re-used.

Starting from the above basic architecture and considering different visions (shortly presented in the Introduction section), several research groups and/or projects developed a large set of variants of refined architectures (see examples in A. Galis [18], [19]). Some of them are substantially different from each other. Currently there is a high heterogeneity seen in this area. The question analyzed in this paper is: *how much convergence/similarity* and how much *mutual compliancy* exists among them?

IV. WHERE DOES THE HETEROGENEITY COME FROM?

Despite the fact that many architectures essentially satisfy the requirements of the meta-architecture presented in Section III, significant heterogeneity can inherently appear in different proposals. This section will summarise the factors leading to heterogeneity in the area of particular architectures. Note that, given the topic's complexity and this limited paper space, the analysis cannot be exhaustive. Some aspects are not touched, or only briefly mentioned, such as: abstraction aspects, slice isolation and security, slice composition, monitoring issues and slice optimization, details on multi-domain interactions, technological and implementation details and so on.

A. Services deployment

This is inherently heterogeneous, depending on applications to be supported. An example is the traffic locality property (at the edge of the network/slice or crossing the core part). An orchestrator should be aware of such traffic properties and if necessary, deploy the corresponding network functions at the mobile edge. The orchestrator needs to have enough topology information of slices in order to be able to install appropriate functions at right places.

B. Assigning tasks to the edge or core network

Some tasks may be executed either in edge network or in the core. So, there are options how to share such burdens between edge and core network. Two options can be identified:

- to keep some administrative control functions (e.g., call management) in the core and only move data plane media-related functions to the edge;
- to move to the edge all relevant VNFs and services in both data, management and control plane.

Trade-offs are between: operational complexity, the need to run multiple instances of the same services, reduced tunneling overhead, and others. Content delivery services in mission-critical environments may require similar decisions.

C. Execution environments

At the infrastructure level, the e*xecution environments* could be heterogeneous. The infrastructure should provide an interface to the orchestrator, via which different functions execution can be started, stopped, paused, or migrated; the interface also provides means to influence the transport of data. Two variants are mentioned below:

- The infrastructure hides (to MANO) its information on the type of execution elements available. The infrastructure management chooses the right (i.e., "functionally possible") realization of a function (virtual machine (VM) or container, etc.). This abstraction simplifies the MANO tasks, but makes difficult for the infrastructure manager to decide what is "performance-optimal", given its lack of information about the performance requirements of an entire service and the relationships to other services.
- The infrastructure provides to the MANO information on available types of execution resources (quantity, locations, etc.). So, the MANO has enough information to optimize the execution environment. The price paid is a higher burden for MANO. Note that such an approach will have some additional issues: it should consider the degree of trust between the infrastructure provider and MANO entity, especially in multi-domain environment.

D. Hardware heterogeneity

At infrastructure level, the hardware heterogeneity can also determine many variants, e.g., virtualization methods and other factors (e.g., Field programmable gate arrays (FPGA), Graphics processing unit (GPU) implementations, hardware accelerators, etc.).

E. Vertical separation of services

The classical principle of separation in *network-related services* (i.e., connectivity–oriented) and *application-level services* (e.g., caching, video transcoding, content-oriented, web server, etc.) could be preserved or not. One can respectively speak about, segregated or integrated orchestration. . The separation will require one service orchestrator and separate network/service orchestrators.

Concerning slicing, one can define some slices offering essentially connectivity services and other dedicated to high-level services and applications. The clear separation of areas of responsibility over resources could be an advantage for operational stability (e.g., a segregated RAN orchestrator could still maintain basic RAN services even if an application-oriented orchestrator fails). On the other hand, the integrated orchestration could be attractive, in particular for operators, if both kinds of services (i.e., the high level and respectively the connectivity-oriented services) could be orchestrated in the same fashion (and possibly, even with the same orchestration infrastructure). These two options also determine heterogeneity at M&O architectural level.

Segregated orchestrators approach leads to a more complex overall architecture. One must assign areas of responsibilities from a resource perspective (i.e, which orchestrator controls - what resources); one should identify services pertaining to each orchestrator. The split of a service is also a problem, i.e., the service description should define the "network" and "application-facing" parts of the service. Aligning the control decisions taken by these two kinds of orchestrators in a consistent way is also not trivial. In an integrated orchestration approach, all these problems disappear. However, an integrated orchestrator might be very complex if required to treat substantially different services (an orchestration of type "one-size-fits-all" approach is rather not the best choice). An integrated orchestrator is a more challenging piece of software (from both dependability and performance perspectives) but would result in a more compact overall architecture.

Considering the above rationale, we defend the idea that from the slicing point of view, a segregate orchestrator approach is a better choice in the sense that it provides s more clear separation of orchestration tasks.

Note that in practice, both approaches have been pursued in different projects. *Currently, a final verdict on segregated*

versus integrated orchestration, commonly agreed by many communities is not yet available. Apparently, there is no evident need to standardize such an option, as long as both of them could be realized inside a meta-architecture. So, for the time being, we can state that M&O heterogeneity, from this point of view, will last.

F. "Flat" or "Hierarchical" orchestration

In the flat solution, a single instance of a particular orchestrator type is in charge to orchestrate all assigned resources. In the hierarchical solution, there are multiple orchestrators (a "hierarchical" model is needed, when orchestrators know to talk to each other). Note that a hierarchical orchestrator is *not necessarily* a segregated one, because all hierarchy members could deal with the same type of services.

In many projects and studies, the hierarchical M&O option is chosen [7][18]-[22]. However, several issues should be solved in each of the two solutions [3]:

- The *number of hierarchy levels* and each member responsibility area could be fixed or adaptive (upon load changes the responsibility areas can be split/merged; new hierarchy levels can be added/removed and new orchestrator instances can be started or some old ones can be stopped). However, the adaptive option is highly complex, given the inherent dynamicity capability required.
- *North/south vertical interfaces* between the orchestrators must be defined. In a flat model, the service requests are received by an orchestrator's northbound interface (NBI). At its south bound the orchestrator communicates with NBI of the abstracted infrastructure (VIM). These two NBIs are structurally different. In a hierarchical model, an orchestrator should be able to communicate with a lower level orchestrator through a different interface than for VIM. So, an orchestrator should be able to use different NBIs (NBI of a VIM, or NBI of a lower-level orchestrator). It is still in study how to create uniform interfaces; the advantage would be that from the perspective of a higher-level orchestrator, it always talks to a VIM-style interface. In such a case, recursive orchestration could be much easier implemented.
- *Horizontal interfaces* (east/west) should be defined between peer orchestrators (those who are on the same level), if they are allowed to negotiate directly with each other (for resources). Such interfaces are naturally to exist in cross-domain slicing scenarios.
- *Multi-domain scenarios* create new problems (e.g., in the case of a multi-domain "federated" slice) [6][18]. In a flat model, each orchestrator of a domain is actually multi-orchestration capable, i.e., it can discuss/negotiate with other domains' orchestrators. In the hierarchical model, a higher-level orchestrator could exist, in charge of harmonizing multiple organizations cooperation. However, several issues are not fully solved today: which entity would run that multi-domain orchestrator, trust issues, preservation of domains independency, assuring the fairness, etc.
- *Mapping* of the *orchestration entities* (and their areas of responsibility) onto *"domains"* (in a very general sense of the word "domain") is still an open research issue and it is also a factor of heterogeneity of the refined M&O architectures. For instance, one could have separate orchestrators for different technological domains (e.g., computational resources, optical networking infrastructure, wireless edge, etc.). However, the word "domain" can be associated as well, to organizations/companies boundaries. Such domains have overlap with the technological ones. A third possible semantic is that a "domain" could be a subdivision of a larger infrastructure into an edge domain, a core domain, etc. (each one spanning multiple technologies, possibly dealing with all kinds of services in a non-segregated way).

G. Relationship of the M&O system and the slicing system

This is another factor of architectural variability, depending on what the definition of a slice is. A largely agreed solution is to have a general orchestrator (configured offline), capable to trigger the construction of a new slice and then to install in this new slice a dedicated orchestrator (before the slice run-time). To still assure the basic services outside any slice (e.g., packet forwarding at network level) one can construct an additional special orchestrator installed outside of all slices. *Currently, many combinations have been proposed, and there is still no consensus on such matters.* The convergence of solutions will be determined probably by the adoption of a more unique definition of a slice – which could assure better inter-operability.

H. Different abstraction mappings applied between hiearchical levels

In a multi-level hierarchy levels of orchestrators, abstractions will be used between adjacent layers, to hide to the upper levels the details of the lower ones. However, it is not clear what the best mapping is, in order to produce a simplified view of a lower level to the upper one. Violations could appear when mapping high level services onto the resources of a lower level [3]. So, different mapping methods can lead to heterogeneity.

I. Conflict resolution

In 5G complex systems there will exist inherently conflicts between participating entities given the basic idea of resource sharing. Different specific choices to solve them will lead to heterogeneity of solutions. A few examples are given below.

Resource conflicts for shared resources: they can appear due to incorrect admission control or overly aggressive oversubscription. Architectural refinements are necessary to solve them.

Conflicting rules: e.g., when composing a service out of functions that specify mutually incompatible packet

forwarding behavior (this can happen both in NFV context or in SDN context).

Feature interaction conflicts: this is a classical issue in systems offering complex multi-feature services and being dynamically configurable as in the case of updating slices.

The conflicts need to be avoided or detected and resolved. Pre-fixed policies (limited approach), either for a platform, or for a service in particular, can help. More research effort points towards conflict resolution actions from inside an operational network is necessary.

J. Time scales (short vs. long-term actions)

It makes sense to separate short-term actions (e.g., actions on a flow level) from long-term planning actions (e.g., decision where to run which function). The refined functional architecture can reflect this separation, e.g., by splitting the MANO system into separate subsystems, each one responsible for different types of actions. A typical terminology would be: "control" for short-time scale operations vs. "orchestration" for operations on longer time scales. This separation is attractive from a software development and maintenance perspective (e.g., a SDN controller becomes a separated piece of software); however, this separation does introduce additional interfaces and operational dependency into an already complex architecture model. The decision on which actions are short-term and which are long-term can produce heterogeneity.

K. Traffic load variations

Some traffic spikes can happen which cannot be simply dealt by the short-term control system. Hence, the long-term orchestrator needs to be also able to deal with short-term changes (this is related with the control/orchestration separation). The MANO system's architecture should have the ability to bring up additional instances. The cloud computing can solve this (Function as a Service –concept (FaaS)) by bringing up functions on an as-needed, load-adaptive basis. However, this requires that the realized code is indeed a function, hence, stateless – there is no state maintained inside a function and it is not possible to move state between function instances.

V. Examples of sliced 5G Management and Orchestration Functional Architectures

This section will provide a few relevant examples to illustrate the major management and orchestration (M&O) options and also the heterogeneity of the refined architectures. Given the limited dimension of this paper, the depth of discussion on them is also limited to the essential aspects illustrating the main characteristics and heterogeneity factors.

A. Example 1

The 5G PPP Working Group [2] and NORMA European Project [20] have proposed a 5G multi-domain architecture by defining four planes: *Service, M&O, Control* and *Data* planes (Figure 3). Note that in [2] the above are called "layers"; however, we believe that the correct semantics is rather "planes". The architecture also includes a *Multi-Domain Network Operating System* containing different adaptors and network abstractions above the networks and clouds heterogeneous fabrics.

The *Service plane* comprises *Business Support Systems* (BSSs) and business-level Policy and Decision functions as well as applications and services operated by the tenant. This includes the end-to-end orchestration system (not detailed in this architecture).

The *M&O* plane comprises a general *Service Management,* the *Software-Defined Mobile Network Orchestrator* (SDMO) and the ETSI NFV lower level managers (i.e., VNFM and VIM). The SDMO is composed of a *Domain specific application management*, an *Inter-slice Resource Broker* and *NFVO*. The SDMO performs the E2E management of network services; it can set up slices by using the network slice templates and merge them properly at the described multiplexing point. The *Inter-slice Broker* handles cross-slice resource allocation and interacts with the *Service Management* (SM) function. The SM is an intermediary function between the service layer and the *Inter-slice Broker*. It transforms consumer-facing service descriptions into resource-facing service descriptions and vice versa. The SDMO has a complete knowledge of the network managing the resources needed by all the slices of all tenants. This enables the SDMO to perform the required optimal configuration in order to adjust the amount of used resources. The MANO accommodates domain-specific application management functions (e.g., in 3GPP, this comprises *Element Managers* (EM) and *Network Management* (NM) functions, including Network (Sub-) Slice Management Function (N(S)SMF). Those functions would also implement ETSI NFV MANO interfaces to the VNF Manager and the NFVO.

The *Control Plane (CPl)* is "horizontally" separated in two parts: intra and inter-slice control functions. "Vertically", it is organized in SDN style, i.e., with three planes: *Control applications* (inter and intra-slice); *SDN controller*s; *SDN nodes* (these are actually slicing control function blocks realized as physical or virtual network functions PNF/VNFs). Note also the flexibility of SDN-NFV cooperation: some slicing control functions are seen and realized as SDN nodes.

The SDN controllers are two types: *Software-Defined Mobile Network Coordinator* (SDM-X) and *Software-Defined Mobile Network Controller* (SDM-C). Following the SDN principles, SDM-X and SDM-C translate decisions of the control applications into commands to VNFs and PNFs. Each network slice has an SDM-C, to manage the network slice resources and building the paths to join the NFs taking into account the received requirements and constraints. The SDM-C and SDM-X take care of dedicated and shared Network Functions (NFs), respectively. SDM-X and SDM-C as well as other control applications can be executed as VNFs or PNFs themselves; this shows the flexibility of SDN/NFV cooperation.

The *Data plane* (DPl) comprises the VNFs and PNFs executing different tasks to carry and process the user data traffic. Following the NRFV principles VNF/PNF graphs are defined and configured in DPl.

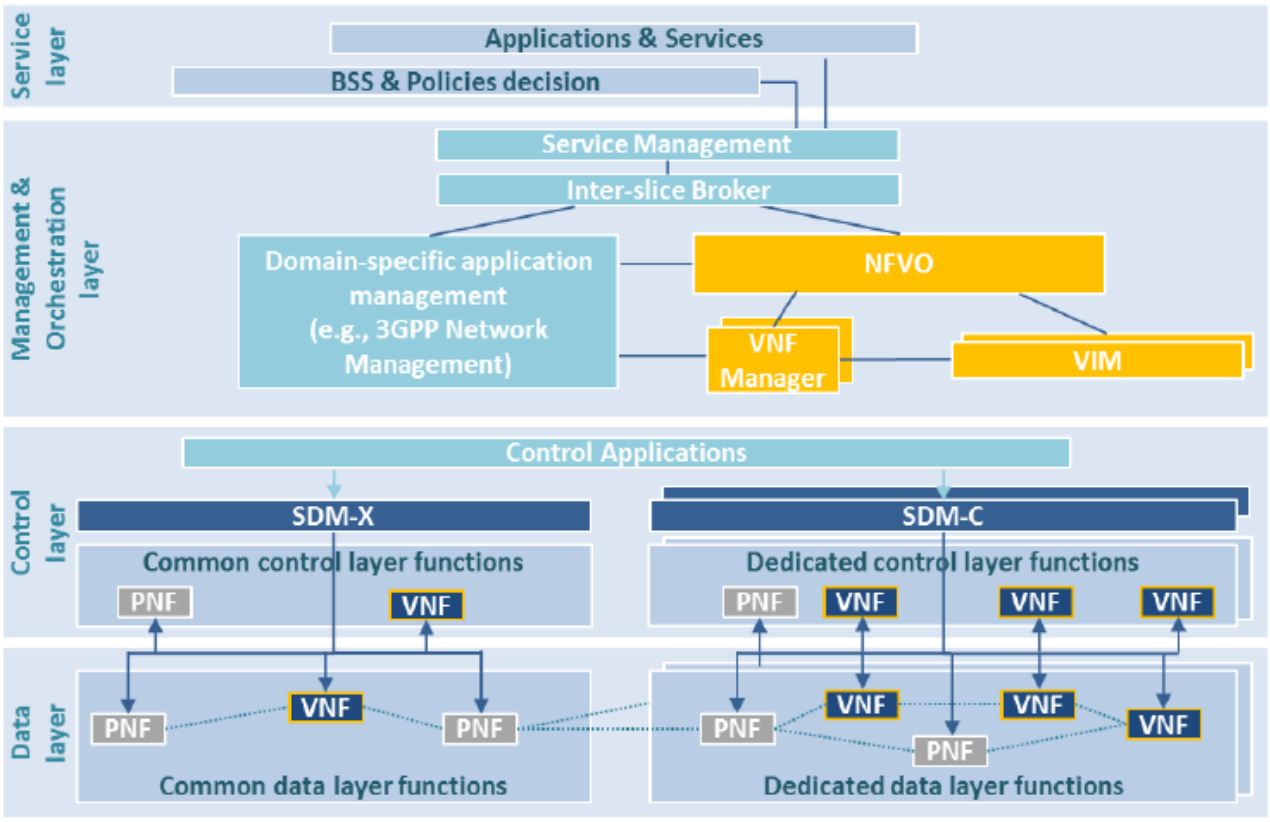


Figure 3. 5G PPP and NORMA project - proposed functional 5G slicing architecture (source: [2][20])

The *Multi-Domain Network Operating System* Facilities (not represented in Figure 3) is an additional subsystem which includes different adaptors and network abstractions above the networks and clouds heterogeneous fabrics. It allocates the (virtual) network resources and maintains network state; it also ensures network reliability in a multi domain environment.

Horizontally, the architecture should cover all segments: RAN (radio and edge), core and transport. 3GPP has defined [2] the 5G System (5GS) comprising a core network (CN) and one or more access networks, e.g., a RAN. The CN consists of NFs, NF services and the interaction between NFs to support data connectivity and other services. It is needed to provide infrastructure connectivity from the Access Points (APs) to the CN, also referred to as transport network connectivity. Transport networks are the foundation of 5GS as they provide the network fabric interconnecting NFs, CN and RAN and the units of RAN.

The architecture presented in Figure 3 is only high level defined. With respect to the meta-architecture capabilities exposed in Section III, it is evident that 5G PPP/NORMA architecture can generally satisfy the requirements a., b., e. and partially f. However, several options could be considered for c., d., f., if wanting to develop further refinements regarding:

c. Different execution environments

The architecture (Figure 3) does not functionally define the virtual infrastructure, neither in data plane nor in the management and control planes, except mentioning the usage of graphs of PNF/VNFs. Therefore, one of several refinement options can be selected.

d. *Working across different "organizational" or administrative domains*

Figure 3 does not define a mapping on a business models containing different actors. While a multi-domain feature is desired, the functional split between different actors is not yet defined. In [2] it is proposed the Mobile Network Service Provider as a main entity capable to serve several tenants with dedicated slices, based on the infrastructure offered by one ore more infrastructure providers, but without detailing the precise framework for resource management. Concerning the multi-domain capabilities one can assume that Inter-slice Broker can manage slices covering several domains [2] but it is not decided how such an Inter-slice Broker is mapped in flat style or hierarchical one onto business actors.

f. Subdivision of the infrastructure in logical separated and isolated slices.

The Figure 3 architecture shows the split of the control and data plane in two regions: common (shared) and respectively dedicated functions. However, the choice on how to separate the slices from point of view performance (observed in the data plane) and security for both data and control plane) can lead to different options for solutions.

Therefore, different or heterogeneous refinements (see Section IV and [2] for several possible solutions) can be selected for such matters.

B. *Example 2*

A multi-domain, multi-tenant hierarchized slicing architecture (viewed at run-time phase, i.e., after a slice instance has been created and activated) is presented in Figure 4, adapted from the proposal ETSI GR NFV-EVE 012 [15] and J. Ordonez-Lucena et al. [4][21]. We state that in comparison with *Example 1*, this architecture presents a more clear hierarchization of M&O functions and also a clear mapping onto a business model. It is adopted a solution with multiple levels of orchestrators and the principle of clear separation between service management versus resource management.

A multi-domain slice instance can span several InPs and/or administrative or technological domains belonging to different providers. Figure 4 shows several domains upon which multi-domain slices can be constructed. (the picture focuses on the transport and core network domains, omitting the RAN domain).

The main M&O entity is the *Network Slice Provider (NSLP)*. Inside NSLP, a highest layer multi-domain *NSL Orchestrator* (NSLO) (configured offline) has a main role, both in the *creation* phase of slices and also in the *run-time* phase. In the creation phase, NSLO receives from a tenant the order to deploy a NSLI (or the NSLP decides itself to construct a slice by provisioning actions). The NSLO should have enough information (including on multi-domain resource availability) in order to check the feasibility of the order. To accomplish this, it interacts with a lower level *Resource Orchestrator* (RO) (which aggregates resource information from several domains (InPs)), and also accesses the VNF and NS catalogues.

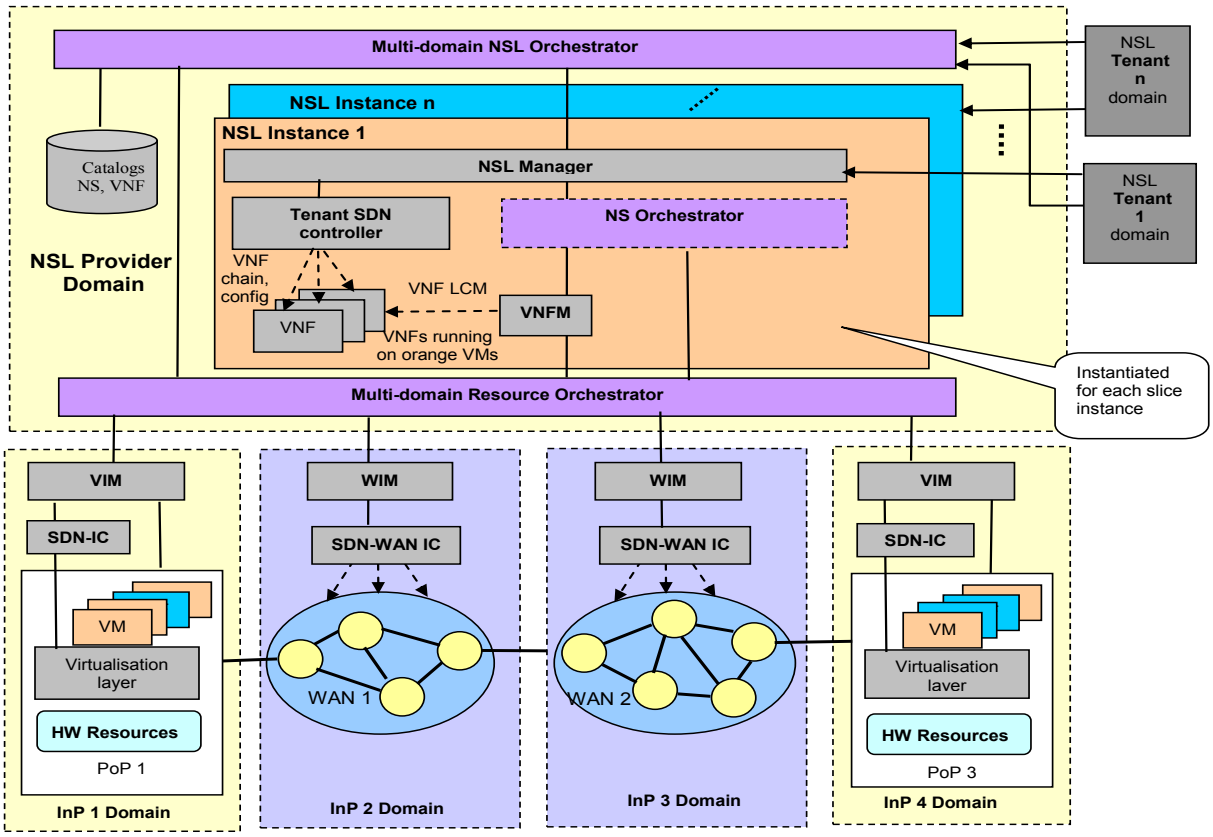


Figure 4. Run-time view of a multi-domain slicing hierarchical architecture example 2
(adapted from ETSI GR NFV-EVE 012 [15] and Ordonez-Lucena [4][21])

NS – Network Service; NSL - Network Slice; VNF – Virtualized Network Function; VNFM – VNF Manager; SDN - Software Defined Networking; LCM – Life Cycle Management; VIM – Virtual Infrastructure Manager; WIM- Widea Area Infrastructure Manager; SDN-IC- Infrastructure SDN controller; HW-Hardware; WAN – Wide Area Network; InP - Infrastructure Provider

The NSL provider plays a role of an infrastructure tenant; it rents the infrastructure resources owned by the underlying infrastructure providers and uses them to provision the NSL instances. The RO uses the set of resources supplied by the underlying VIMs/WIMs and optimally dispatches them to the NSL instances. All the NSL instances are simultaneously provided with the needed resources to satisfy their requirements and preserve their performance isolation. Note that in this high-level architecture proposal it is not detailed how the multi-domain capable RO is implemented in order to assure two important objectives: harmonizing the resource assignments per slice and per-domain and also to preserve the inter-domain independence in terms of management and control.

For each new network slice instance (NSLI), an individual set of M&O entities is dynamically created and installed when the new slice instance is created. Each NSLI has its own M&O and control planes (this assures the slice isolation capability) composed of: NSL Manager, Network Service Orchestrator (NSO), Tenant SDN Controller and VNF Manager (VNFM).

The NSLP rents infrastructure resources owned by the underlying InPs to construct NSL instances. The *Resource Orchestration* (RO) manages the set of resources offered by different INPs (the resources are supplied under the control of the underlying VIMs/WIMs), and optimally dispatches them to the NSLIs aiming to satisfy their requirements but preserving their logical isolation. The RO *should have information on resource availability in each domain* whose resources will enter the multi-domain NSLI. To construct a multi-domain slice, inter-domain interactions are also necessary.

An SDN control is supposed to exist at domain level. The *SDN - Infrastructure Controller* (SDN-IC) manages and controls connectivity in its domain, under the directives of the corresponding VIM/WIM. The VIMs and WIMs can act as SDN applications, delegating the tasks related to the management of networking resources to their underlying ICs.

Does the above architecture satisfy the requirements of a meta-architecture (see Section III)? The answer is "yes", i.e.:

a. The individual control of NFs (their placement, LCM, etc.) can be realized due to existence of the pair- manager VNFM and tenant SDN Controller (at M&Olevel) and by the pair VIM and SDN-IC.

b. The chaining of individual NFs into services (NF graphs) can be assured by the same M&O blocks as above.

c. Different underlying execution environments: various virtualization techniques (virtual machines (VM), containers, or plain processes) can refine the architecture. Such details are not visible at this high level but are naturally possible to be embedded in each domain.

d. Working across different "organizational", or administrative domains, i.e., owned by network operators or companies and using various business models- is already emphasized in Figure 4.

e. A large range of applications with different specific requirements (in terms of resource, deployment, orchestration and optimization goals) can be supported given that a tenant has interfaces to NSLs, allowing it to express its requirements.

f. Subdivision of the infrastructure in logical separated and isolated slices with levels of guaranteed performance is possible to be achieved, given the mapping from services to the resources orchestrated by the RO.

Refinements of the above high-level architectures are possible [4][15][21], following different paths to go further towards the system design. Examples could be:

- how to split the RO functionalities between different operators' domains in the case of E2E multi-domain slices;
- the functional split among SDN-IC and WIM and consequently the interface/relationship between WIM and SDN-IC with respect to: (1) the style used by SDN-IC to upload information to VIM/WIM, about its available resources: *on demand* (OD) or in *proactive* (P) style (at SDN-IC initiative); (2) the amount and depth of information uploaded by SDN-IC on the network resources (graph, capacities, etc.).

The above example illustrates the inherent heterogeneity of particular refined architectures, while all starting from a "tree root" defined by the meta-architecture requirements.

C. Example 3

T. Taleb et al. [7] recently proposed a multi-domain slicing hierarchical, complex management and orchestration architecture (Figure 5). They use a powerful definition of a slice, i.e., "a set of network functions, and resources to run these network functions, forming a complete instantiated logical network to meet certain network characteristics required by the Service Instance(s)". An E2E NSL can be deployed across multiple networks, stretching across the RAN, transport and core network segments; belonging to the same or different administrative domains.

A NSLI typically consists of multiple Network Slice Subnet Instances (NSSIs) that represent a group of network function instances and/or logical connectivity. As an example, a *Fully-Fledged NSLI*, can consists of several NSSIs, each belonging to a different technology domain, e.g., Radio Access Network (RAN), transport and core. The RAN and core NSSIs are composed of VNF(s) interconnected over logical transport links.

The proposed architecture is structured into four major strata: *Multi-domain Service Conductor (MSC); Domain-specific Fully-Fledged Orchestration, Sub-Domain Management and Orchestration (MANO) and Connectivity, Logical Multi-domain Slice Instances.*

The architecture introduces (at top level) a novel architectural plane named *Service Broker* (SB), to handle incoming slice requests from verticals, for instance *Mobile Virtual Network Operators* (MVNO), and application providers. The main SB operations are: Network Service (NS) admission control and negotiation, considering service aspects; management of slice user/owner relationship enabling a direct tenant interface with the *Multi-domain Service Conductor* (MSC) plane; billing and charging; NSLI scheduling, i.e., start and termination instant of time, related with slice composition and decommission.

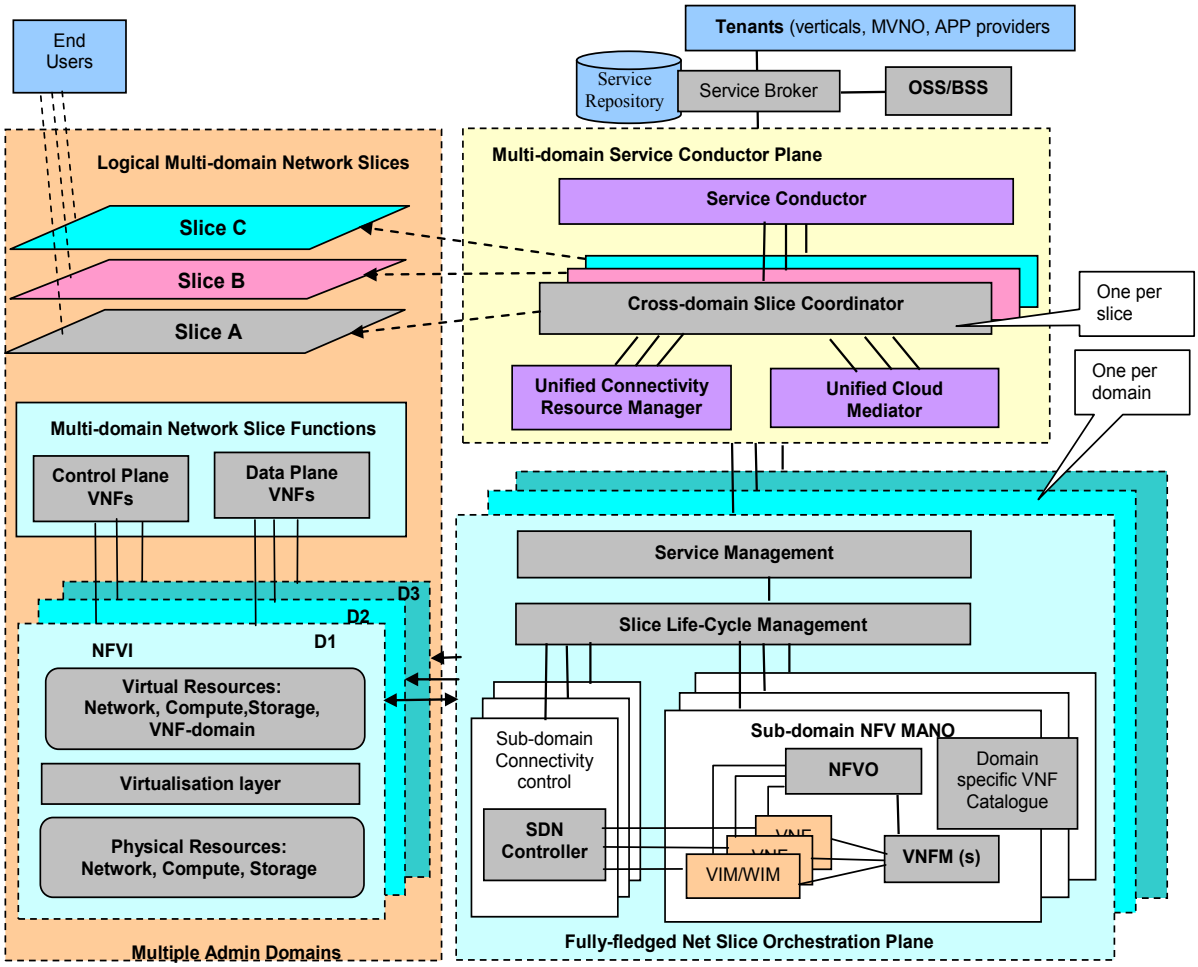


Figure 5 Multi-domain multi-tenant 5G slicing architecture example 2 (adapted from [7])

The *Multi-domain Service Conductor* is (functionally) placed under Service Broker and performs service management across *federated domains*. The MSC stratum analyzes and maps the service requirements of incoming multi-domain slice requests onto the one or several administrative domains. It also maintains the desired service performance throughout the entire service lifecycle. Inside MSC, a *Service Conductor (SC)* entity is placed on top; the SC analyses and maps the service requirements of incoming slice requests onto appropriate administrative domains and maintains the desired service performance during service lifecycle.

Below SC, a *Cross-domain Slice Coordinator* is defined for each slice, which aligns cloud and networking resources across federated domains and carries out the *Life Cycle Management (LCM)* operations of a multi-domain slice. It also establishes and controls inter-domain transport layer connectivity, assuring the desired performance. A multi-domain NSLI can combine several *Fully-Fledged NSLIs* that belong to distinct administrative domains, to get an E2E multi-domain (i.e., a federated NSLI).

A coordinated management system is required to facilitate an effective LCM of a Fully-Fledge NSLI. At minimum, the following management, orchestration and control M&O&C entities are necessary: _ *Network Slice Manager* for the configuration and operation of a mobile network service to a Fully-Fledge NSLI; _ NFV MANO to instantiate and orchestrate the requested VNFs considering the supported availability; SDN Controller that connects together VNFs forming service function chains and controls the transport layer connectivity.

For each domain a *Fully-fledged NetSlice Orchestration Plane* is constructed, dealing with specific operations associated to slices instance in that domain (such as service management and slice lifecycle management). The lower layers of this specific orchestration plane comprise NFV MANO functionalities (NFVO, VNFM and VIM). Low level connectivity tasks

between VNF/PNFs are performed by an SDN controller; this is a similar solution as in Example 2.

The above complex architecture *can satisfy all of a…f.* general requirements of the meta-architecture. Many specific refinements can be added to satisfy the A..K (Section IV) as presented in the work [7].

D. Example 4

The 5G-MoNArch H2020 project [22] develops a hierarchical architecture consisting of four layers: *Service, M&O, Controller and Network* layer (similar to that proposed in [2] by 5GPPP).

The main design goals of the 5G-MoNArch architecture design [22] has been among those defined by the meta-architecture described in Section III:

(1) Support for E2E network slicing: one can combine different options of slicing support across M&O and network layers for each slice instance. Several options are possible: a. slice-specific functions (i.e., dedicated/ customised functions that are not shared with others); b. functions (or function instances) that are shared by multiple slices and have the capability to address requirements from multiple slices in parallel).

(2) Split of control and user/data plane throughout all network domains, including RAN, Core Network and Transport Network.

(3) Flexible architecture customisation: this is performed by the management system which can modify the architecture and functionality used in existing slices. For example, this can include further deployment, management, orchestration, and control instructions for specialised NFs [22].

The overall functional architecture is presented in Figure 6. The Service layer comprises *Business Support Systems (BSS),* business-level *Policy and Decision* functions, and further applications and services operated by a tenant or other external entities.

The management and orchestration layer contains M&O functions from different network, technology, and administration domains (e.g., 3GPP public mobile network management, ETSI NFV MANO, ETSI Multi-access Edge Computing functions, management functions of transport network or enterprise networks). The M&O layer is divided into an End-to-End (E2E) service M&O sublayer and an additional sublayer containing domain-specific management functions. An E2E network slice is composed of *Network Slice Subnet Instances (NSSIs)*, typically each from a different network domain, including subnets from radio access network (RAN), transport, and core network domains, or private networks. The M&O layer performs cross-domain coordination actions.

Note again the architectural separation between the management and control. The Controller layer comprises two types of controllers- cross-slice and the intra-slice (XSC and ISC, respectively). On top of the controllers, there are *Control Applications*; together they realise the network programmability in SDN style. Each network domain has a dedicated controller that is aware of the domain technology and implementation characteristics.

Generally, the MoNArch architecture satisfies the requirements of the meta-architecture described in Section III. However, many (heterogeneous) refinements should be added in order to cover the A..K (Section IV) needs.

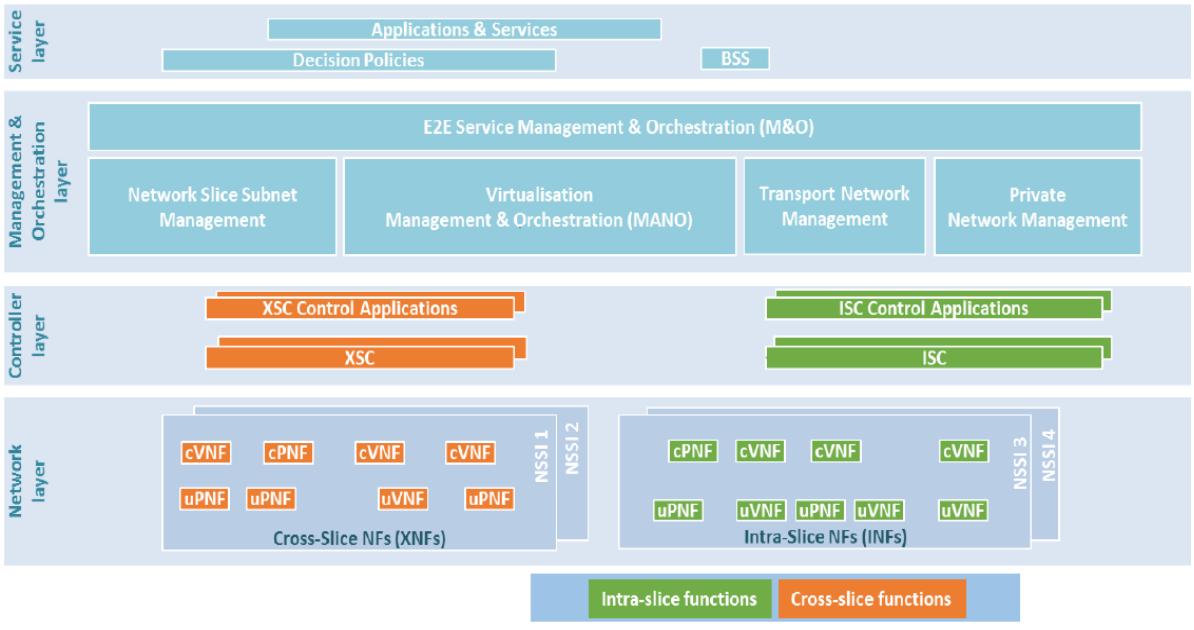


Figure 6. 5G-MoNArch high-level overall functional architecture (Source: [22])

Other variants f architectures are proposed and developed in different research projects [18]. Again, all of them satisfy the characteristics of the meta-architecture described in Section III. However, different specific developments are present in their refined version.

VI. CONCLUSIONS AND FUTURE WORK

This is an overview-type paper; it analyzed different M&O architectures for 5G slicing, in order to evaluate the degree of their similarity/convergence, given the large variety of proposals existing in various studies, standards and projects.

It has been shown that business model definitions (actors) and their roles (Section II) have an important impact on the high-level definition of the architectural assembly. Actually, the variety of business models is a primary factor of architectural heterogeneity, given the different definition of actors and roles, firstly adopted mainly from business reasons and only secondly from technical ones. *Also, the definition of a slice itself is still not yet globally agreed and this situation naturally leads to different architectures.*

However, a unifying meta-architecture has been defined (see Section III), answering to some basic requirements for 5G systems and, in particular, for 5G M&O slicing. It has been derived from ETSI MANO work complemented with additional functionalities slice-oriented. The most relevant architecture examples found in literature and developments are essentially compliant with the basic meta-architecture. It is important to note that all relevant architectures proposed in different studies, standards and projects, generally try to achieve the main meta-architecture capabilities.

On the other hand, many factors are inducing heterogeneity of the refined architecture variants, such as: multi-domain, multi-tenant, multi-operator, multi-technology.

Future work can go further to consider more deeply the multi-x aspects, implementation and performance. Future work can concentrate on M&O issues such as: an appropriate cooperation between slice-specific management functional blocks. Policies need to be captured in a way that they can be automatically validated. This automation enables slice-specific functional blocks to be authorized to perform the corresponding management and configuration actions in a timely manner.

Designing computationally efficient resource allocation algorithms and conflict resolution mechanisms at each abstraction layer is also a way to flexibly assign resource on-the-fly to slices.

Lastly, one should mention new approaches for 5G slicing M&O architectures: usage of artificial intelligence and in particular, machine learning techniques in order to provide more M&O automation, optimization and capabilities of dealing with big volumes of data [23]-[26]. This domain is only at its beginning, so it is an open field for further studies.

REFERENCES

[1] E. Borcoci, C. Contu, A. Ciobanu, "5G Slicing Management and Orchestration Architectures - Any Convergence?", The Eleventh International Conference on Advances in Future Internet, AFIN 2019, October 2019 - Nice, France https://www.iaria.org/conferences2019/ProgramAFIN19.html [retrieved February, 2020]

[2] 5GPPP Architecture Working Group, "View on 5G Architecture", Version 2.0, December 2017, https://5g-ppp.eu/wp-content/uploads/2017/07/5G-PPP-5G-Architecture-White-Paper-2-Summer-2017_For-Public-Consultation.pdf, [retrieved June, 2019].

[3] 5G PPP Architecture Working Group, "View on 5G Architecture", Version 3.0, June, 2019, https://5g-ppp.eu/wp-content/uploads/2019/07/5G-PPP-5G-Architecture-White-Paper_v3.0_PublicConsultation.pdf, [retrieved October, 2019].

[4] J. Ordonez-Lucena et al., "Network Slicing for 5G with SDN/NFV: Concepts, Architectures and Challenges", IEEE Communications Magazine, 2017, pp. 80-87, Citation information: DOI 10.1109/MCOM.2017.1600935.

[5] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network Slicing in 5G: Survey and Challenges", IEEE Communications Magazine, May 2017, pp. 94-100.

[6] I. Afolabi, T. Taleb, K. Samdanis, A. Ksentini, and H. Flinck, "Network Slicing & Softwarization: A Survey on Principles, Enabling Technologies & Solutions", IEEE Communications Surveys & Tutorials, March 2018, pp. 2429-2453.

[7] T. Taleb, I. Afolabi, K. Samdanis, and F. Z. Yousaf, "On Multi-domain Network Slicing Orchestration Architecture & Federated Resource Control", http://mosaic-lab.org/uploads/papers/3f772f2d-9e0f-4329-9298-aae4ef8ded65.pdf, [retrieved June, 2019].

[8] ETSI GS NFV 002, "NFV Architectural Framework", V1.2.1, December, 2014.

[9] ETSI GS NFV-IFA 009, "Network Functions Virtualisation (NFV); Management and Orchestration; Report on Architectural Options", Technical Report, V1.1.1, July, 2016.

[10] ETSI GR NFV-IFA 028, "Network Functions Virtualisation (NFV) Release 3; Management and Orchestration; Report on architecture options to support multiple administrative domains", Technical Report, V3.1.1, January, 2018.

[11] ONF TR-526, "Applying SDN Architecture to 5G Slicing", April 2016, https://www.opennetworking.org/wp-content/uploads/2014/10/Applying_SDN_Architecture_to_5G_Slicing_TR-526.pdf [retrieved January, 2020]

[12] "5G Network and Service Management Including Orchestration" v3.14.0, Project 5G NWMO, 2019, https://www.ngmn.org/fileadmin/ngmn/content/downloads/Technical/2019/190312_5G_Network_and_Service_Management__including_Orchestration_3.14.0.pdf, [retrieved June, 2019].

[13] "ETSI NFV Announces New Features to its Architecture to support 5G", https://www.etsi.org/newsroom/press-releases/1622-2019-07-etsi-nfv-announces-new-features-to-its-architecture-to-support-5g, [retrieved July, 2019].

[14] G. Daniels, "ETSI tightens 3GPP 5G compatibility with NFV" Release 3, July 2, 2019, https://www.telecomtv.com/content/nfv/etsi-tightens-3gpp-5g-compatibility-with-nfv-release-3-35653/, [retrieved July, 2019].

[15] ETSI GR NFV-EVE 012, Release 3 “NFV Evolution and Ecosystem; Report on Network Slicing Support with ETSI NFV Architecture Framework”, Technical Report, V3.1.1, December, 2017.

[16] 3GPP TR 28.801, "Telecommunication management; Study on management and orchestration of network slicing for next generation network", V15.0.0, September, 2017.

[17] 3GPP TS 28.530, “Management of 5G networks and network slicing; Concepts, use cases and requirements”, Rel. 15, April 2018.

[18] A. Galis, “Network Slicing- A holistic architectural approach, orchestration and management with applicability in mobile and fixed networks and clouds”, http://discovery.ucl.ac.uk/10051374/, [retrieved July, 2019].

[19] K. Katsalis, N. Nikaein and A. Edmonds, "Multi-Domain Orchestration for NFV: Challenges and Research Directions", 2016 15th Int'l Conf. on Ubiquitous Computing and Communications and International Symposium on Cyberspace and Security (IUCC-CSS), pp. 189–195, DOI: 10.1109/IUCC-CSS.2016.034, https://ieeexplore.ieee.org/document/7828601, [retrieved July, 2019].

[20] 5G NORMA D3.2 Network Architecture – Intermediate Report, https://5gnorma.5g-ppp.eu/dissemination/public-deliverables/ [retrieved September, 2019].

[21] J. Ordonez-Lucena et al., "The Creation Phase in Network Slicing: From a Service Order to an Operative Network Slice", European Conference on Networks and Communications (EuCNC), 2018, https://arxiv.org/abs/1804.09642 [retrieved July, 2019].

[22] H2020-ICT-2016-2, Monarch Project, 5G Mobile Network Architecture for diverse services, use cases and applications in 5G and beyond, Deliverable D2.3, “Final overall architecture”, 2019, https://5g-monarch.eu/wp-content/uploads/2019/05/5G-MoNArch_761445_D2.3_Final_overall_architecture_v1.0.pdf, [retrieved June, 2019].

[23] V. P. Kafle et al., “Consideration on Automation of 5G Network slicing with Machine Learning”, ITU Kaleidoscope Santa Fe, 2018.

[24] J. Moysen and L. Giupponi, “From 4G to 5G: Self-organized Network Management meets Machine Learning”, arXiv:1707.09300v1 [cs.NI] 28 July 2017.

[25] S. Ayoubi et al., “Machine Learning for Cognitive Network Management”, IEEE Communications Magazine, January 2018, pp.158-165.

[26] D. Lorenz et al., “SliceNet – Cognitive Slice Management Framework for Virtual Multi-Domain 5G Networks”, https://www.systor.org/2018/pdf/systor18-21.pdf, [retrieved September, 2019].