International Journal on Advances in Internet Technology

2015 vol. 8 nr. 3&4

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 8, Number 3 & 4, 2015

Thierry Brouard, Université François Rabelais de Tours, France
Dumitru Dan Burdescu, University of Craiova, Romania
Carlos T. Calafate, Universitat Politècnica de València, Spain
Christian Callegari, University of Pisa, Italy
Juan-Vicente Capella-Hernández, Universitat Politècnica de València, Spain
Miriam A. M. Capretz, The University of Western Ontario, Canada
Ajay Chakravarthy, University of Southampton IT Innovation Centre, UK
Chin-Chen Chang, Feng Chia University, Taiwan
Ruay-Shiung Chang, National Dong Hwa University, Taiwan
Tzung-Shi Chen, National University of Tainan, Taiwan
Xi Chen, University of Washington, USA
Dickson Chiu, Dickson Computer Systems, Hong Kong
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
Alfredo Cuzzocrea, University of Calabria, Italy
Jan de Meer, smartspace®lab.eu GmbH, Germany
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
Marc Fabri, Leeds Metropolitan University, UK
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
Shang Gao, Zhongnan University of Economics and Law, China
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
Feliz Gouveia, Universidade Fernando Pessoa - Porto, Portugal
Kannan Govindan, Crash Avoidance Metrics Partnership (CAMP), USA
Bill Grosky, University of Michigan-Dearborn, USA
Vic Grout, Glyndŵr University, UK

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
Eva Hladká, Masaryk University, Czech Republic
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
Evangelos Kranakis, Carleton University - Ottawa, Canada
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
Nicolas Le Sommer, Université Européenne de Bretagne, France
Juong-Sik Lee, Nokia Research Center, USA
Wolfgang Leister, Norsk Regnesentral (Norwegian Computing Center), Norway
Clement Leung, Hong Kong Baptist University, Hong Kong
Man-Sze Li , IC Focus, UK
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
Sergio Martín Gutiérrez, UNED-Spanish University for Distance Education, Spain
Constandinos Mavromoustakis, University of Nicosia, Cyprus
Hamid Mcheick, Université du Québec à Chicoutimi, Canada
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
Vlad Nicolicin Georgescu, SP2 Solutions, France
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
Mark Perry, University of Western Ontario/Faculty of Law/ Faculty of Science - London, Canada
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
Joel Rodrigues, Instituto de Telecomunicações / University of Beira Interior, Portugal
Simon Pietro Romano, University of Napoli Federico II, Italy
Michele Ruta, Politecnico di 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
Didier Sebastien, University of Reunion Island, France
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
Åsa Smedberg, Stockholm University, Sweden
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
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
Parimala Thulasiraman, University of Manitoba, Canada
Pierre Tiako, Langston University, USA
Ioan Toma, STI Innsbruck/University Innsbruck, Austria
Orazio Tomarchio, Universita' di Catania, Italy

International Journal on Advances in Internet Technology
Volume 8, Numbers 3 & 4, 2015

CONTENTS

Investigation of Users Suspected of Manipulating Evaluations of Answers in a Q&A Site

Yasuhiko Watanabe, Hiroki Matsumoto, Kenji Umemoto, Ryo Nishimura, Yoshihiro Okada, and Shin Yamanaka
Ryukoku University
Seta, Otsu, Shiga, Japan
Email: watanabe@rins.ryukoku.ac.jp, t14m086@mail.ryukoku.ac.jp,
t11m074@mail.ryukoku.ac.jp, r_nishimura@afc.ryukoku.ac.jp,
okada@rins.ryukoku.ac.jp, t100450@mail.ryukoku.ac.jp

***Abstract*—Some users in a question and answer (Q&A) site use multiple user accounts and attempt to manipulate communications in the site. In this paper, we first show a proper reason why many users in a Q&A site use multiple accounts from the viewpoint of personal data protection. On the other hand, we show some users can use multiple accounts inadequately and manipulate communications in the site. In order to detect these inadequate multiple account users precisely, we investigate them from two new points of view. First point of view is the number of accounts for submitting questions and manipulating evaluations of their answers. Second point of view is the deviations of answer submission order. We show these points of view are useful for detecting inadequate multiple account users precisely. The results of this study will give us a chance to investigate communication strategies of users in a Q&A site, especially, purposes and methods of inadequate multiple account users.**

Keywords–multiple account; Q&A site; evaluation manipulation; anonymity.

I. INTRODUCTION

These days, many people use question and answer (Q&A) sites, where users share their information and knowledge. Q&A sites offer greater opportunities to users than search engines in the following points:

1) Users can submit questions in natural and expressive sentences, not keywords.
2) Users can submit ambiguous questions because other users give some supports to them.
3) Communications in Q&A sites are interactive. Users have chances to not only submit questions but give answers and, especially, join discussions.

As a result, Q&A sites are promising media. One of the essential factors in Q&A sites is anonymous submission. In most Q&A sites, user registrations are required for those who want to join the Q&A sites. However, registered users generally do not need to reveal their real names to submit messages (questions, problems, answers, comments, etc.). It is important to submit messages anonymously to a Q&A site. This is because anonymity gives users chances to submit messages without regard to shame and reputation. However, some users abuse the anonymity and attempt to manipulate communications in a Q&A site. For example, we showed some users use multiple user accounts and submit messages to a Q&A site inadequately [1]. Manipulated communications discourage other submitters, keep users from retrieving good communication records, and decrease the credibility of the Q&A site. As a result, it is important to detect users suspected of using multiple user accounts and manipulating communications in a Q&A site. In this case, identity tracing based on user accounts is not effective because inadequate users are likely to hide their true identity to avoid detection. A possible solution is authorship identification based on analyzing stylistic features of messages. In recent years, a large number of studies have been made on authorship identification [2] [3] [4] [5] [6], however, few researchers addressed the identification issues of authors suspected of using multiple user accounts and manipulating communications in a Q&A site. To solve this problem, we proposed methods of detecting

- multiple account users suspected of submitting questions and their answers repeatedly [7], and
- multiple account users suspected of submitting many answers to the same question repeatedly [8].

However, little is known about the purposes and methods of inadequate multiple account users. As a result, it is important to investigate these inadequate multiple account users from various points of view. To solve this problem, we proposed two new points of view for investigating inadequate multiple account users [1] [9]. In this study, we discuss these two points of view further and introduce them together for detecting users suspected of manipulating evaluations of their answers precisely.

First point of view is whether these inadequate users use two or more accounts for submitting questions and manipulating evaluations of their answers [1]. It is natural for them to use multiple accounts for submitting questions. This is because too many good evaluations from one account may give strange impressions to other users and operators in the site. Inadequate users do not want to draw attention to themselves. As a result, in this paper, we investigate users suspected of using multiple user accounts for submitting many questions and manipulating evaluations of their answers.

Second point of view is whether these inadequate users use multiple user accounts in different ways [9]. Suppose that one user intends to advocate or justify his/her submitted answer and uses multiple user accounts as follows:

- a main account, and
- secondary accounts for advocating or justifying his/her answer submitted by the main account.

In this case, the deviation of answer submission order is likely to occur. As a result, we investigate user pairs who had large deviations of answer submission order and discuss the reasons why and how the deviations occurred.

Finally, we should notice that it is difficult to verify the credibility of our investigation. This is because there is no reliable information about users who used multiple user accounts and manipulated communications in Q&A sites. In order to discuss the credibility of our investigation, we show the results of our investigation in detail. The results of this study will give us a chance to investigate communication strategies of users in a Q&A site, especially, purposes and methods of inadequate multiple account users.

The rest of this paper is organized as follows: In Section II, we survey the related works. In Section III, we describe Yahoo! chiebukuro for an example of Q&A sites. In Section IV, we describe how inadequate users use multiple user accounts in Q&A sites. In Section V, we investigate user pairs suspected of using two or more accounts for submitting questions and manipulating evaluations of their answers. In Section VI, we investigate user pairs who had large deviations of answer submission order and discuss the reasons why and how the deviations occurred. In Section VII, we introduce the proposed points of view together and show that they are useful for investigating inadequate multiple account users in a Q&A site. Finally, in Section VIII, we present our conclusions.

II. Related works

One of the essential factors of the Internet is anonymity. The author of [10] discussed the anonymity on the Internet from various points of view. The author of [11] pointed that anonymity is necessary in almost any protocol, application or service used in wired or wireless networks, and showed a survey on anonymity preserving solutions. These days, many users abuse the anonymity. Take a Sybil attack for example. In a Sybil attack, the attacker intends to gain large influence on a peer-to-peer (P2P) network by creating and using a large number of pseudonymous identities [12] [13]. Sybil attacks are cheap and efficient way to gain large influence on P2P networks [14]. Similarly, in human online communities, such as, web-based bulletin boards, chat rooms, and blog comment forms, many users are thought to use multiple user accounts inadequately and submit inadequate messages, such as, deceptive opinion spams. In recent years, a large number of studies have been made on authorship identification [2] [3] [4] [5] [6], however, few researchers addressed the identification issues of authors suspected of using multiple user accounts and manipulating communications in the Internet. One of the difficulties of this problem is that we did not have sufficient number of examples of inadequate multiple account users. To solve this problem, some researchers tried to extract inadequate submissions by using heuristic methods based on text similarities and ranking results [15] [16]. On the other hand, the authors of [17] pointed that these heuristic methods were insufficient to detect inadequate submissions precisely, and showed they could detect inadequate submissions precisely when they used large number of examples of inadequate submissions. However, they obtained examples of inadequate submissions by using Amazon Mechanical Turk [18]. The examples of inadequate submissions created by workers in Amazon Mechanical Turk have the following problems.

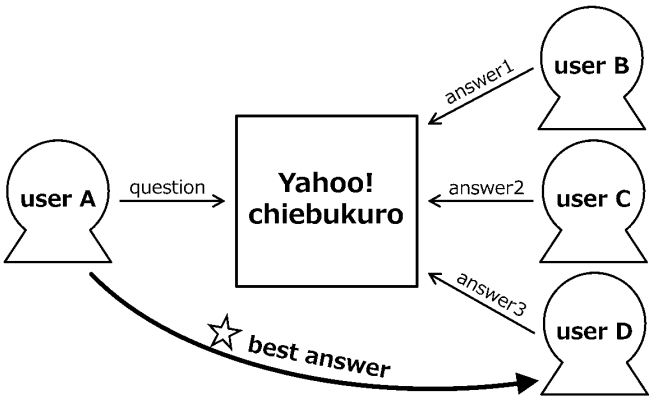


Figure 1. An example of how to use Yahoo! chiebukuro.

- Little is known about the purposes and methods of inadequate submissions. As a result, it is possible that their instructions to workers in Amazon Mechanical Turk were insufficient.
- There are unreliable workers in Amazon Mechanical Turk [19].

As a result, it is important to obtain inadequate submissions from the Internet. To solve this problem, we proposed methods of detecting inadequate multiple account users and their submissions [7] [8]. However, as mentioned, little is known about the purposes and methods of inadequate multiple account users. As a result, it is important to investigate these inadequate multiple account users and their inadequate submissions from various points of view.

III. Yahoo! chiebukuro

Yahoo! chiebukuro is one of the most popular community sites in Japan. Users of Yahoo! chiebukuro submit their questions and answers in the next way.

- User registrations are required for those who want to join Yahoo! chiebukuro.
- Users do not need to reveal their real names to submit their questions and answers.
- User accounts of submitters are recorded and shown in their questions and answers.
- Each user can submit his/her answer only one time to one question.
- The period limit for accepting answers is one week. However, questioners can stop accepting answers before the time limits.
- After the time limits, questions with no answers are removed and cannot be referable. On the other hand, questions with answers can be referable.

TABLE I. THE NUMBERS OF USERS AND THEIR SUBMISSIONS TO PC CATEGORY, SOCIAL ISSUES CATEGORY, AND ALL 286 CATEGORIES IN YAHOO! CHIEBUKURO (FROM APRIL/2004 TO OCTOBER/2005).

category	number of questioners	number of questions	number of answerers	number of answers
PC	43,493	171,848	27,420	474,687
social issues	13,259	78,777	25,766	403,306
all 286 categories	165,064	3,116,009	183,242	13,477,785

- Each questioner is requested to determine which answer to his/her question is best and give a *best answer* label to it.

Figure 1 shows that user A submitted one question to Yahoo! chiebukuro and three users, user B, user C, and user D answered the question, and then, user A selected user D's answer as a best answer. In this study, we used the data of Yahoo! chiebukuro for observation and examination. Chiebukuro means pearls of wisdom. The data of Yahoo! chiebukuro was published by Yahoo! JAPAN via National Institute of Informatics in 2007 [20]. This data consists of about 3.11 million questions and 13.47 million answers, which were posted on Yahoo! chiebukuro from April/2004 to October/2005. In the data, each question has at least one answer because questions with no answers were removed. In order to avoid identifying individuals, user accounts were replaced with unique ID numbers. By using these ID numbers, we can trace any user's questions and answers in the data. Table I shows

- the numbers of questioners and their questions in the data, and
- the numbers of answerers and their answers in the data.

In Table I, the number of questioners is the number of users who submitted one or more questions to Yahoo! chiebukuro from April/2004 to October/2005. Also, the number of answerers is the number of users who submitted one or more answers to Yahoo! chiebukuro from April/2004 to October/2005.

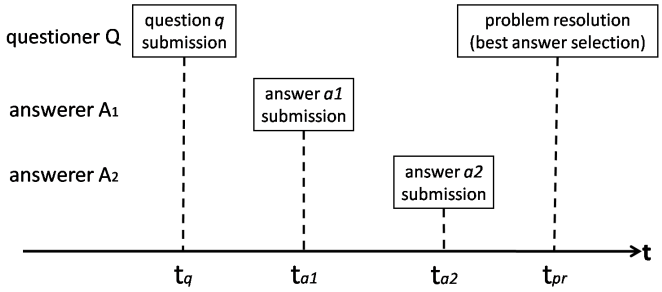
Furthermore, the following kinds of information are described in the data.

- submission time of question
- submission time of answer
- problem resolution time

Figure 2 shows an example of a series of events that occur after a questioner submits his/her question to Yahoo! chiebukuro. In Figure 2, the submission time of question q is t_q. Also, the submission time of answer a_1 and a_2 are t_{a1} and t_{a2}, respectively. Finally, the problem resolution time of question q is t_{pr}. At the problem resolution time, questioner Q stopped accepting answers and determined which answer was the best answer. By using these kinds of time information, we measured

- submission time lags between questions and their answers (e.g., $t_{a1} - t_q$ and $t_{a2} - t_q$ in Figure 2),
- submission time lags between answers submitted to the same question (e.g., $t_{a2} - t_{a1}$ in Figure 2), and
- answer submission order.

Figure 3 shows the cumulative relative frequency of the submission time lags between questions and their answers in the data of Yahoo! chiebukuro. Also, Figure 4 shows the cumulative relative frequency of the submission time lags between answers submitted to the same question. As shown in Figure 4, the median of the submission time lags between answers submitted to the same question in social issues category was greater than those of PC category and all 286 categories. In social issues category, there were many answers criticizing or against previous answers. As a result, many answerers in this category made and submitted answers after they read other answers to the same question. We think this is one of the reasons why the median of the submission time lags between answers submitted to the same question in social issues category was greater than those of PC category and all 286 categories.



Questioner Q submitted question q at t_q. Also, answerer A_1 and A_2 submitted their answers at t_{a1} and t_{a2}, respectively. Finally, questioner Q stopped accepting answers and determined which answer was the best answer at t_{pr}.

Figure 2. An example of a series of events that occur after a questioner submits his/her question to Yahoo! chiebukuro.

IV. SUBMISSIONS BY USING MULTIPLE USER ACCOUNTS

There are many reasons why users in a Q&A site use multiple user accounts. First, we discuss a proper reason. In Yahoo! chiebukuro, users do not need to reveal their real names to submit their questions and answers. However, their submissions are traceable because their user accounts are attached to them. Because of this traceability, we can collect any user's submissions and some of them include clues of identifying individuals. Each clue (age, gender, location, occupation, hobby, health, and so on) is not sufficient to identify individuals, however, the combination of them is sometimes sufficient. As a result, to avoid identifying individuals and protect personal information, it is reasonable and proper that users change their user accounts or use multiple user accounts.

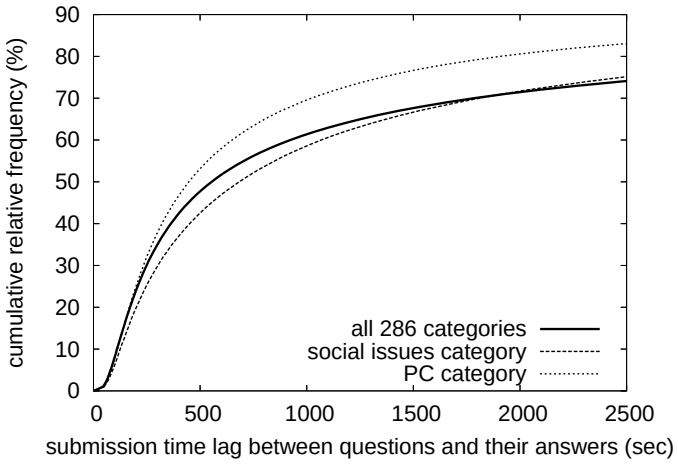


Figure 3. The cumulative relative frequency of the submission time lags between questions and their answers in social issues category, PC category, and all 286 categories of the data of Yahoo! chiebukuro.

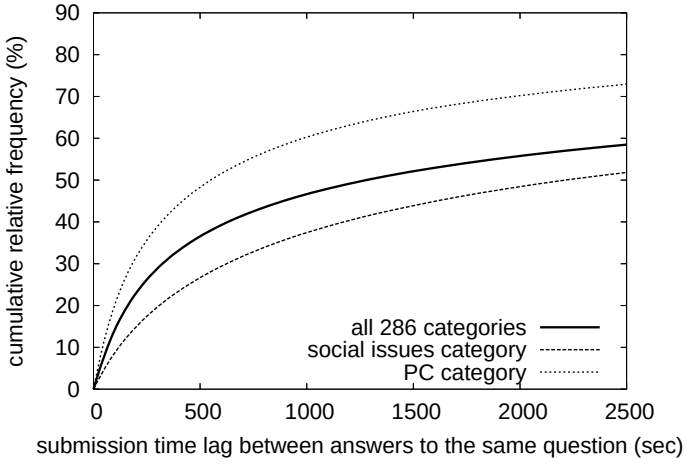
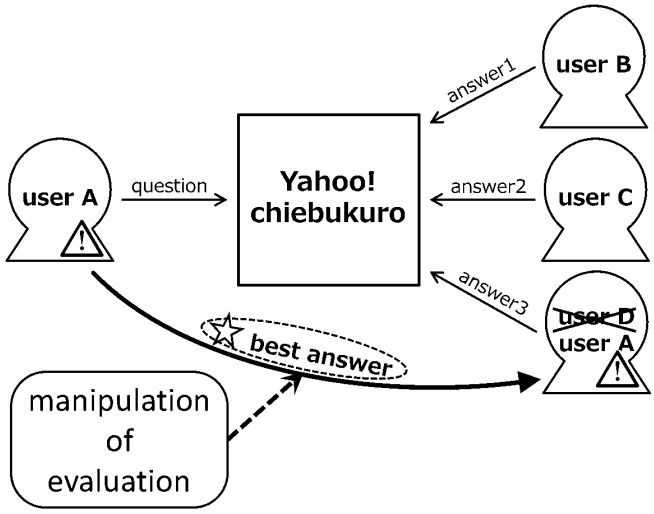


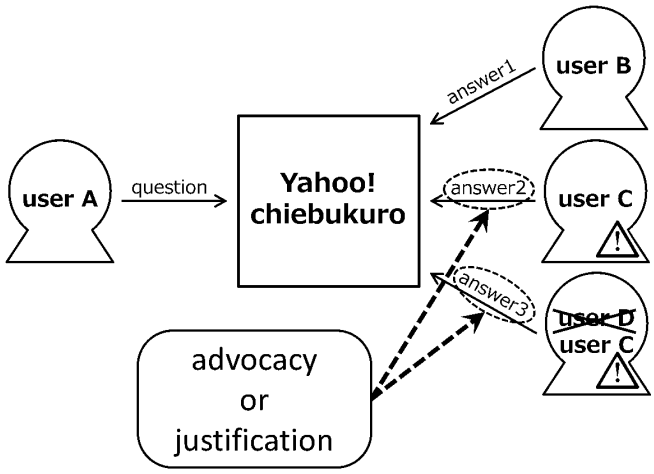
Figure 4. The cumulative relative frequency of the submission time lags between answers submitted to the same question in social issues category, PC category, and all 286 categories of the data of Yahoo! chiebukuro.

However, the following types of submissions by using multiple user accounts are neither reasonable nor proper.

TYPE QA One user submits a question and its answer by using multiple user accounts (Figure 5 (a)).
We think that the user intended to manipulate the submission evaluation. For example, in Yahoo! chiebukuro, each questioner is requested to determine which answer is best and give a *best answer* label to it. These evaluations encourage answerers to submit new answers and increase the credibility of the Q&A site. We think that the user repeated this type of submissions because he/she wanted to get many best answer labels and be seen as a good answerer.



(a) TYPE QA: one user submits a question and its answer by using multiple user accounts. (In this case, user A submits a question and its answer by using two user accounts.)



(b) TYPE AA: one user submits two or more answers to the same question by using multiple user accounts. (In this case, user C submits two answers by using two user accounts.)

Figure 5. Two types of inadequate submissions: TYPE QA and TYPE AA.

TYPE AA One user submits two or more answers to the same question by using multiple user accounts (Figure 5 (b)).
We think that the user intended to dominate or disrupt communications in the Q&A site. To be more precise, the user intended to

- manipulate communications by advocating or justifying his/her opinions, or
- disrupt communications by submitting two or more inappropriate messages.

TYPE AA submissions are more similar to Sybil attacks in P2P networks than TYPE QA submissions. The more answers inadequate users submit by using multiple user accounts, the easier they manipulate or disrupt communications in a Q&A site.

These two types are not all types of inadequate submissions. However, these kinds of submissions seriously disrupt communications in a Q&A site. For example, TYPE QA submissions are serious because users can manipulate evaluations of messages by repeating TYPE QA submissions. Manipulated evaluations discourage other submitters, keep users from retrieving good communication records, and decrease the credibility of the Q&A site. Furthermore, we think we cannot use knowledge and countermeasures obtained in studies of Sybil attacks in P2P networks because TYPE QA submissions are different from Sybil attacks. In a Sybil attack, the more pseudonymous identities the attacker uses, the easier he/she gain large influence on a P2P network. On the other hand, in a TYPE QA submission, the inadequate user can get a best answer label by using only two user accounts. To solve this problem, we proposed methods of detecting multiple account users suspected of repeating TYPE QA submissions [7]. However, little is known about the purposes and methods of inadequate multiple account users. As a result, it is important to investigate these inadequate multiple account users from various points of view. In this study, we introduce the following two points of view for investigating inadequate multiple account users.

- whether these inadequate users use multiple accounts for repeating TYPE QA submissions, and
- whether these inadequate users use multiple accounts in different ways for repeating TYPE AA submissions.

Furthermore, we investigate users detected based on these point of view and discuss whether these points of view are useful.

V. Detection of users suspected of using two or more accounts for submitting questions and manipulating evaluations

From the standpoint of the number of user accounts for submitting questions and manipulating evaluations of their answers, inadequate multiple account users who repeat TYPE QA submissions can be classified into two types:

- inadequate users each of whom uses only one user account for submitting questions and manipulating evaluations of his/her answers, and
- inadequate users each of whom uses two or more user accounts for submitting questions and manipulating evaluations of his/her answers.

In this section, we investigate the latter type of users, in other words, users suspected of using two or more user accounts for submitting questions and manipulating evaluations of their answers.

A. Basic idea

Suppose that one user intends to manipulate evaluations of his/her answers, submitted by using user account a, and repeats TYPE QA submissions by using two user accounts, q_1 and q_2. In this case, it is expected that we observe the following unusual submissions:

- user a submits too many answers to questions submitted by user q_1 and q_2,
- user q_1 and q_2 receive too many answers from user a, and
- user q_1 and q_2 give too many best answer labels to user a's answers.

Taking account of these points, we detect users suspected of using two or more user accounts for submitting questions and manipulating evaluations of their answers in the next way. We first detect user pairs suspected of repeating TYPE QA submissions. Then, we detect answerers who are found in two or more of the detected user pairs. The detected answerers are suspected of using two or more user accounts for submitting questions and manipulating evaluations of their answers.

B. Detection of users suspected of using two or more user accounts for submitting questions and manipulating evaluations of their answers

In order to detect users who intend to manipulate evaluations of their answers and submit many questions by using two or more user accounts, we propose a method that consist of the following two steps:

1) We first detect user pairs of a questioner and an answerer, who are suspected of repeating TYPE QA submissions, as shown in Figure 5, by using three hypotheses: Hypothesis QA1, QA2, and QA3.
2) We detect users who are answerers in two or more user pairs detected by using Hypothesis QA1, QA2, and QA3.

Hypothesis QA1, QA2, and QA3 are as follows:

[Hypothesis QA1] If user a did not submit unusually too many answers to user q's questions, we would expect that user a submitted at most $N_{QA1}(q, a)$ answers to user q's questions.

$$N_{QA1}(q, a) = P_{QA1}(q) \times ans(a) \tag{1}$$

where $ans(a)$ is the total number of answers submitted by user a and $P_{QA1}(q)$ is the probability that an user selects one question randomly and the question is one of user q's questions. Because each user of Yahoo! chiebukuro can submit his/her answer only one time to one question, $P_{QA1}(q)$ is

$$P_{QA1}(q) = \frac{qst(q)}{N_{qst}} \tag{2}$$

where $qst(q)$ is the number of questions submitted by user q and, as shown in Table I, N_{qst} is the total number of questions in the data of Yahoo! chiebukuro. If this hypothesis is rejected by an one-sided binomial test, we determine that user a submitted unusually too many answers to user q's questions.

The binomial test is an exact test of the statistical significance of deviations from a theoretically expected binomial distribution of observations into two categories [21]. There are two types of binomial tests: one sided binomial tests or two sided binomial tests. When the critical area of a distribution is one-sided, in other words, it is either greater than or less than

TABLE II. THE DETECTION RESULT OF USERS SUSPECTED OF USING TWO OR MORE USER ACCOUNTS FOR SUBMITTING QUESTIONS AND MANIPULATING EVALUATIONS OF THEIR ANSWERS

significance levels for QA1, QA2, QA3, and QA3aux	UP_{QA}	UP_{QA2+}	A_{QA}	A_{QA2+}
5×10^{-5}	814	329	581	96
1×10^{-5}	603	222	450	69
5×10^{-6}	537	188	408	59
1×10^{-6}	424	135	333	44
5×10^{-7}	407	129	319	41
1×10^{-7}	337	104	266	33
5×10^{-8}	325	101	257	33
1×10^{-8}	278	86	220	28

UP_{QA} is the number of user pairs who are detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux. UP_{QA2+} is the number of user pairs the answerers of whom were found in two or more user pairs detected by binomial tests. A_{QA} is the number of answerers who are found in user pairs detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux. A_{QA2+} is the number of answerers who are found in two or more user pairs detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux.

a certain value, but not both, only an one-sided binomial test is generally applicable. In this study, the distribution area is one-sided, we use the one-sided binomial test.

[Hypothesis QA2] If user q did not receive unusually too many answers from user a, we would expect that user q received at most $N_{QA2}(q, a)$ answers from user a.

$$N_{QA2}(q, a) = P_{QA2}(a) \times qst(q) \tag{3}$$

where $qst(q)$ is the total number of questions submitted by user q and $P_{QA2}(a)$ is the probability that an user received one answer from user a when user a selected one question randomly and answered it. Because each user of Yahoo! chiebukuro can submit his/her answer only one time to one question, $P_{QA2}(a)$ is

$$P_{QA2}(a) = \frac{ans(a)}{N_{qst}} \tag{4}$$

where $ans(a)$ is the number of answers submitted by user a and, as shown in Table I, N_{qst} is the total number of questions in the data of Yahoo! chiebukuro. If this hypothesis is rejected by an one-sided binomial test, we determine that user q received unusually too many answers from user a.

[Hypothesis QA3] If user q did not give unusually too many best answer labels to user a's answers, we would expect that user q gave at most $N_{QA3}(q, a)$ best answer labels to user a's answers.

$$N_{QA3}(q, a) = P_{QA3}(q) \times f_{QA}(q, a) \tag{5}$$

where $f_{QA}(q, a)$ is the number of answers submitted by user a to user q's questions, and $P_{QA3}(a)$ is the best answer ratio of user a.

$$P_{QA3}(a) = \frac{bestans(a)}{ans(a)} \tag{6}$$

where $ans(a)$ is the number of answers submitted by user a and $bestans(a)$ is the number of best answers in user a's answers. However, if user j satisfies one of the following conditions:

- all user a's answers were selected as best answers, in other words,

$$ans(a) = bestans(a) \tag{7}$$

- Hypothesis QA3aux, the auxiliary hypothesis for Hypothesis QA3, is rejected, in other words, it is considered that user a received too many best answer labels,

we set $P_{QA3}(a)$ as follows:

$$P_{QA3}(a) = \frac{N_{bestans}}{N_{ans}} = \frac{N_{qst}}{N_{ans}} \tag{8}$$

where $N_{bestans}$ is the total number of best answers. $N_{bestans}$ is equal to N_{qst} because each question has one best answer. If this hypothesis is rejected by an one-sided binomial test, we determined that user q gave unusually too many best answer labels to user a's answers.

[Hypothesis QA3aux] If user a did not receive unusually too many best answer labels, we would expect that user a received at most $N_{QA3aux}(a)$ best answer labels.

$$N_{QA3aux}(a) = P_{QA3aux} \times ans(a) \tag{9}$$

where P_{QA3aux} is the average best answer ratio.

$$P_{QA3aux} = \frac{N_{bestans}}{N_{ans}} = \frac{N_{qst}}{N_{ans}} \tag{10}$$

If this hypothesis is rejected by an one-sided binomial test, we consider that user a received unusually too many best answer labels.

C. Result of the Investigation

To evaluate our method, we conducted the detection of users suspected of using two or more user accounts for submitting many questions and repeating TYPE QA submissions, and manipulating evaluations of their answers in a Q&A site. In this experiment, the target users were all submitters in the data of Yahoo! chiebukuro. As shown in Table I, the numbers of the target questioners and answerers in the data of Yahoo! chiebukuro are 165,064 and 183,242, respectively.

In our method, we varied the significance levels for Hypotheses QA1, QA2, QA3, and QA3aux from 5×10^{-5} to 1×10^{-8} . They were extremely low because we intend to detect extreme unusual submissions. Table II shows the results of this experiment.

As shown in Table II, 59 users were detected when the significance level was 5×10^{-6}. We should notice that 28 users of them were detected when the significance level was 1×10^{-8}. It shows that many users were detected although the significance level was extremely low. As we expected, there are many users suspected of repeating TYPE QA submissions and manipulating evaluations of their answers by using two or more user accounts for submitting questions.

We checked questions and answers submitted by the detected user pairs and found that some questioners were criticized for their unfair best answer selections. For example, user 233650 was criticized that he/she selected user 678451's answers as best answers repeatedly and unfairly. After criticized for his/her unfair best answer selection, user 233650 stopped submitting any questions to Yahoo! chiebukuro. Our method is useful for detecting these suspicious users. Furthermore, if we detect and take care of these suspicious users, we can avoid unnecessary frictions between users.

VI. Detection of too large deviations of answer submission order

Inadequate users repeating TYPE QA submissions are likely to use multiple user accounts as follows:

- main accounts, and
- secondary accounts for submitting questions and manipulating evaluations of main accounts.

On the other hand, little is known how inadequate users repeating TYPE AA submissions use multiple user accounts. To solve this problem, in this study, we investigate whether inadequate users repeating TYPE AA submissions used multiple user accounts in different ways as inadequate users repeating TYPE QA submissions did, and discuss the purposes and methods of them.

If one user uses multiple user accounts in different ways, some deviations are likely to occur. Suppose that one user intends to advocate or justify his/her submitted answer and uses multiple user accounts as follows:

- a main account, and
- secondary accounts for advocating or justifying his/her answer submitted by the main account.

In this case, the user is likely to submit first answers from his/her main account and other answers from their secondary accounts. In order to detect this kind of inadequate users, we introduce *deviation of answer submission order*.

[deviation of answer submission order] Suppose user i and user j submitted their answers to the same N questions, and, user i submitted N_i answers earlier than user j and user j submitted N_j answers earlier than user i. The deviations of answer submission order of this user pair is $N_i - N_j$.

In this study, we investigate user pairs who had large deviations of answer submission order. This is because, we think, deviations of answer submission order give us a chance to learn inadequate multiple account users from the new point of view.

In Yahoo! chiebukuro, there were many questions the purpose of which was to collect opinions. For example,

(Q) What do you think about Prime Minister Koizumi? He has maintained high approval ratings and does well in his work.

This kind of question often had many answers. Some of them were criticizing or against previous answers. Because of such critical submissions, some users were likely to use multiple user accounts and submit new answers for advocating or justifying their previous answers. We think some users used multiple user accounts as follows:

- main accounts, and
- secondary accounts for advocating or justifying their answer submitted by the main accounts.

This is because it is easy to manage multiple user accounts. When multiple user accounts were used as above, it is easy to avoid submitting new answers that were inconsistent with the previous answers. Inconsistent answers often gave suspicious impressions to others. However, if multiple user accounts were used in this way, the deviation of answer submission order is likely to occur. As a result, in this study, we investigate user pairs who had large deviations of answer submission order and discuss the reasons why and how the deviations occurred.

A. Basic idea

In order to detect users who were suspected of repeating TYPE AA submissions by using multiple user accounts in different ways, we introduce two ideas. If one user repeated TYPE AA submissions too many times by using two user accounts, user i and user j, it is expected that

(idea 1) user i and user j submit too many answers to the same questions together.

Furthermore, if the user used these two user accounts in different ways, it is expected that

(idea 2) there are too large deviations of answer submission order between user i and user j.

Based on these two ideas, we determine whether users repeated TYPE AA submissions by using multiple user accounts in different ways.

B. Most frequently encountered user

In order to explain our method of detecting too large deviations of answer submission order, we introduce the term *most frequently encountered user*.

Many users have other users who submitted answers to the same questions with them. We will use the term *most frequently encountered user* of a certain user to refer to a user who submitted answers to the same questions most frequently with the user.

[most frequently encountered user] Suppose user i submitted answers to the same questions with N users ($k = 1, \cdots, N$), and among them, most frequently with user j.

$$mfe(i) = \max_{k \in N} N_{ans\_together}(i,k) = j$$

where $mfe(i)$ is the most frequently encountered user of user i and $N_{ans\_together}(i,k)$ is the number of questions to which user i and k submitted answers together. N_{mfe} is the total number of each user's answers that were submitted with his/her most frequently encountered user. For example, N_{mfe} is 74,781 in social issues category. As a result, it is expected that, when a user submitted 100 answers to social issues category, the user and his/her most frequently encountered user submitted

$$\frac{N_{mfe}}{N_{ans}} \times 100 = \frac{74781}{403306} \times 100 = 18.5$$

answers together to the same questions.

C. Detection of user pairs who submitted too many answers to the same questions

As mentioned, if one person used two user accounts, user i and user j, and repeated TYPE AA submissions in a Q&A site too many times, it is expected that we observe unusual submissions:

- user i submitted unusually too many answers to the same questions with user j.

To detect these unusual submissions, we test one hypothesis: Hypothesis AA.

[Hypothesis AA] If user i did not submit unusually too many answers to the same questions with user j, we would expect that user i submitted at most $N_{AA}(i)$ answers to the same questions with user j.

$$N_{AA}(i) = \frac{N_{mfe}}{N_{ans}} \times ans(i)$$

where $ans(i)$ is the total number of answers submitted by user i. N_{ans} is the total number of answers submitted to the category, and N_{mfe} is the total number of each user's answers that were submitted with his/her most frequently encountered user. If this hypothesis is rejected by an one-sided binomial test, we determine that user i submitted unusually too many answers to the same questions with user j.

D. Detection of user pairs who had too large deviations of answer submission order

If one user repeated TYPE AA submissions by using two user accounts, user i and user j, in different ways, it is expected that we observe

- too large deviations of answer submission order between user i and user j.

To detect too large deviations of answer submission order between user i and user j, we test one hypothesis: Hypothesis AASO.

[Hypothesis AASO] Suppose that there are $N_{AA}(i,j)$ cases where user i and user j submitted their answers to the same question. If one of these users did not submit answers too many times before the other did, we would expect that there are at most $N_{AASO}(i,j)$ cases where one user submitted his/her answer before the other did.

$$N_{AASO}(i,j) = P_{AASO}(i,j) \times N_{AA}(i,j)$$

where $P_{AASO}(i,j)$ is the probability that one user submitted an answer before the other did. In this study, $P_{AASO}(i,j)$ was set to 0.5. In other words, user i and user j have equal probability that one user submitted an answer before the other did. If this hypothesis is rejected by a two-sided binomial test, we determine that one of these users, user i or user j, submitted answers unusually too many times before the other did.

E. Result of the investigation

In order to detect too large deviations of answer submission order, we test Hypothesis AA and AASO. In this study, the target user pairs are 23,053,308 user pairs each of whom submitted answers to at least one same question in Yahoo! chiebukuro. The significance level for Hypothesis AA was extremely low: 5×10^{-6}. This is because we intend to detect extreme unusual submissions. On the other hand, the significance level for Hypothesis AASO was 1×10^{-2}.

In this experiment, we first applied Hypothesis AA on 23,053,308 user pairs in Yahoo! chiebukuro, and detected 790 user pairs who repeated submitting answers to the same question too many times. Then, we applied Hypothesis AASO on these 790 user pairs and detected 382 user pairs who had too large deviations of answer submission order. In order to discuss the detection results in detail, we take the result in social issues category for example. This is because there were many discussions between answerers in this category. As a result, it seems more likely that some multiple account users intended to advocate or justify their answers and repeated TYPE AA submissions in this category. The target user pairs in social issues category are 828,812 user pairs. We applied Hypothesis AA on these 828,812 user pairs and detected 20 user pairs who repeated submitting answers to the same question too many times. Then, we applied Hypothesis AASO on these 20 user pairs and detected 7 user pairs who had too large deviations of answer submission order. Table III shows the result of the investigation on these 7 user pairs. In Table III, user A_1 mainly submitted answers before user A_2 did. $N_{AA}(A_1,A_2)$ is the number of questions to which both user A_1 and user A_2 submitted answers. $NE_{AA}(A_1,A_2)$ is the number of questions where user A_1 submitted answers before user A_2 did. $T_{QA}(A_1,A_2)$ is the median of submission time lags between questions and the earlier of their answers of A_1 or A_2. $T_{AA}(A_1,A_2)$ is the median of submission time lags between answers of A_1 and A_2 submitted to the same question. Figure 3 shows the cumulative relative frequency of submission time lags between questions and their answers. Also, Figure 4 shows the cumulative relative frequency of submission time lags between answers submitted to the same question. By considering the similarity of writing styles and opinions, we determined whether each user pair is one and the same user or not. Decision shows our judgements. We discuss

- whether each of these 7 user pairs is one and the same user or not,
- the purposes of inadequate multiple account users, and
- the reasons why and how the deviations of answer submission order occurred

in detail below.

TABLE III. THE RESULT OF THE INVESTIGATION ON THE 7 USER PAIRS WHO HAD TOO LARGE DEVIATIONS OF ANSWER SUBMISSION ORDER.

A_1	A_2	$N_{AA}(A_1, A_2)$	$NE_{AA}(A_1, A_2)$	$T_{QA}(A_1, A_2)$	$T_{AA}(A_1, A_2)$	decision
691911	802184	47	43	5.1 min.	83 sec.	same user
267614	76731	62	44	22 min.	22 min.	same user
458523	518681	86	61	9.0 min.	26 min.	different users
414445	733881	20	18	4.0 min.	2.3 hrs.	different users
649164	622996	40	30	6.6 hrs.	30 hrs.	same user
471690	471692	12	11	16 hrs.	50 hrs.	same user
622996	471692	12	11	18 hrs.	74 hrs.	different users

$N_{AA}(A_1, A_2)$ is the number of questions to which both user A_1 and A_2 submitted answers. $NE_{AA}(A_1, A_2)$ is the number of questions where user A_1 submitted answers before user A_2 did. $T_{QA}(A_1, A_2)$ is the median of submission time lags between questions and the earlier of their answers of A_1 or A_2. $T_{AA}(A_1, A_2)$ is the median of submission time lags between answers of A_1 and A_2 submitted to the same question. Decision shows our judgements. By considering the similarity of writing styles and opinions, we determined whether each user pair is one and the same user or not.

User pair (267614, 76731) submitted many answers to the questions about foreign residents in Japan. We determined that user 267614 and 76731 were one and the same user. This is because their writing styles and opinions were quite similar and their answers often included special words, for example, personal HP and comic artists, which other users did not cover in this category. These accounts were likely to be used for repeating the same words. For example,

[Q: 654871] I found this exhibitor in the auction [URL]. I think it is against the rule.
[A: 76731] It is scratchbuild. Let it go. You are a snitch.
[A: 267614] You are like a snitch in North Korea. Or a hound.

We thought there was plenty of this kind of inadequate users in Yahoo! chiebukuro. Our method detected user pair (267614, 76731), however, it was not enough. This is because we did not think of any reasons why this kind of users used their multiple user accounts in this way. We are searching more examples of this kind of inadequate users and intend to find the reasons.

Also, in case of user pair (691911, 802184), we determined these users were one and the same user. This is because the median of submission time lags between their answers was only 83 seconds although user 691911 submitted answers at different times of a day. In addition, user pair (691911, 802184) had too large deviations of answer submission order in 9 categories, including social issues category. Furthermore, when user 691911 submitted questions, user 691911 selected user 802184's answers as best answers in too many times in various categories. In addition, like the case of user pair (267614, 76731), these accounts were likely to be used for repeating the same words.

In contrast, in cases of user pair (458523, 518681) and (414445, 733881), we determined that the users of each pair were different users. This is because we found many opinion conflict between the users of each pair. Each pair used Yahoo! chiebukuro almost at the same time of each day. For example, user 458523 and 518681 mainly used Yahoo! chiebukuro from 8:00 am to 5:00 pm. Also, user 414445 and 733881 mainly used Yahoo! chiebukuro from 8:00 pm to 1:00 am. As a result, the users of each pair read questions almost at the same time. On the other hand, the median of submission time lags from questions to user 458523's answers and user 518681's answers were 9.9 minutes and 28 minutes, respectively. Also, the median of submission time lags from questions to user 414445's answers and user 733881's answers were 7.4 minutes and 66 minutes, respectively. We think that these time lags gave the deviations of answer submission order between the users of each user pair.

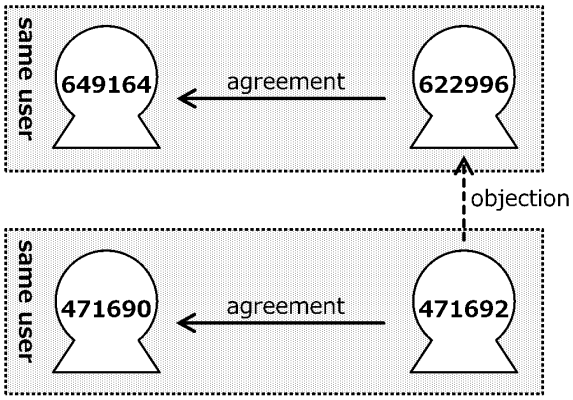


Figure 6. The relation between user pair (649164, 622996) and (471690, 471692).

Both user pair (649164, 622996) and (471690, 471692) submitted answers repeatedly to questions about a certain religious group. We determined that the users of each pair were one and the same users. This is because they had similar writing styles and opinions respectively. Especially, there was only one opinion conflict between user 649164 and 622996 just after they were pointed out that they were one and the same user. As shown in Table III, user 622996 and 471692 mainly submitted their answers after user 649164 and 471690 did, respectively. In both cases, two user accounts were used in different ways as follows:

- main accounts (user 649164 and 471690), and
- secondary accounts (user 622996 and 471692) for criticizing other users' answers, or advocating or justifying answers submitted by the main accounts

and repeated TYPE AA submissions. Especially, user 471692 often criticized user 622996's answers. As a result, user pair (622996, 471692) was detected although the users of this pair were different users and had different opinions. Figure 6 shows

TABLE IV. THE DETECTION RESULT OF USER PAIRS DETECTED BY USING THE METHODS DESCRIBED IN SECTION V AND SECTION VI.

significance levels for QA1, QA2, QA3, and QA3aux	UP_{QA}	UP_{QA2+}	UP_{QA2+AA}
5×10^{-5}	814	329	6
1×10^{-5}	603	222	6
5×10^{-6}	537	188	6
1×10^{-6}	424	135	5
5×10^{-7}	407	129	5
1×10^{-7}	337	104	5
5×10^{-8}	325	101	5
1×10^{-8}	278	86	5

UP_{QA} is the number of user pairs who are detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux. UP_{QA2+} is the number of user pairs the answerers of whom were found in two or more user pairs detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux. UP_{QA2+AA} is the number of user pairs (1) the answerers of whom were found in two or more user pairs detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux and (2) who were detected by binomial tests based on Hypothesis AA and AASO.

TABLE V. THE DETECTION RESULT OF ANSWERERS BELONG TO USER PAIRS DETECTED BY USING THE METHODS DESCRIBED IN SECTION V AND SECTION VI.

significance levels for QA1, QA2, QA3, and QA3aux	A_{QA}	A_{QA2+}	A_{QA2+AA}
5×10^{-5}	581	96	4
1×10^{-5}	450	69	4
5×10^{-6}	408	59	4
1×10^{-6}	333	44	4
5×10^{-7}	319	41	4
1×10^{-7}	266	33	4
5×10^{-8}	257	33	4
1×10^{-8}	220	28	4

A_{QA} is the number of answerers belong to user pairs detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux. A_{QA2+} is the number of answerers who were found in two or more user pairs detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux. A_{QA2+AA} is the number of answerers (1) who were found in two or more user pairs detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux and (2) who belong to user pairs detected by binomial tests based on Hypothesis AA and AASO.

the relation between user pair (649164, 622996) and (471690, 471692).

VII. DETECTION OF INADEQUATE MULTIPLE ACCOUNT USERS FROM NEW POINTS OF VIEW

In this section, we introduce the two points of view, described in Section V and Section VI, together for investigating inadequate multiple account users in a Q&A site and discuss the effectiveness of them.

Table IV shows the number of user pairs

- the answerers of whom were found in two or more user pairs detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux and
- who were detected by binomial tests based on Hypothesis AA and AASO.

Table V shows the number of answerers

- who were found in two or more user pairs detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux and
- who belong to user pairs detected by binomial tests based on Hypothesis AA and AASO.

We take the result under the significance level of 5×10^{-6} for example. As shown in Table IV, when the significance level was 5×10^{-6}, 537 user pairs were detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux. Among them, we found 188 user pairs the answerers of whom were found in two or more user pairs detected by the binomial tests. Among them, the following 6 user pairs were detected by binomial tests based on Hypothesis AA and AASO.

- user pair (236956, 32780),
- user pair (494870, 434516),
- user pair (302153, 434516),
- user pair (652657, 528067),
- user pair (479505, 528067), and
- user pair (691911, 802184).

The first users of them were questioners and the second users were answerers when they were detected by binomial tests based on Hypothesis QA1, QA2, QA3, and QA3aux. The accounts of these answerers are thought to be main accounts of inadequate multiple account users. In this experiment, we varied the significance levels for Hypotheses QA1, QA2, QA3, and QA3aux from 5×10^{-5} to 1×10^{-8}. However, as shown in Table V, four answerers (user 32780, 434516, 528067, and

TABLE VI. THE NUMBER OF ANSWERS AND BEST ANSWERS OF THE 6 USER PAIRS DETECTED BY USING THE METHODS DESCRIBED IN SECTION V AND SECTION VI.

Q	A	$ans(A)$	$N_{BA}(Q, A)$	$N_A(Q, A)$
236956	32780	457	275	282
494870	434516	88	67	69
302153	434516	20	9	9
652657	528067	743	95	199
479505	528067	448	55	110
691911	802184	146	143	146

$ans(A)$ is the total number of answerer A's answers in Yahoo! chiebukuro. $N_A(Q, A)$ is the number of answerer A's answers that were submitted to questioner Q's questions. $N_{BA}(Q, A)$ is the number of answerer A's best answers that were selected by questioner Q.

TABLE VII. THE DEVIATIONS OF ANSWER SUBMISSION ORDER OF THE 6 USER PAIRS DETECTED BY USING THE METHODS DESCRIBED IN SECTION V AND SECTION VI.

A_1	A_2	$N_{AA}(A_1, A_2)$	$NE_{AA}(A_1, A_2)$	$T_{QA}(A_1, A_2)$	$T_{AA}(A_1, A_2)$	category
236956	32780	22	2	5.0 min.	6.6 min.	healthcare
494870	434516	54	0	20 min.	2.3 hrs.	language
302153	434516	19	0	31 min.	5.0 hrs.	language
652657	528067	89	73	6.5 min.	38 min.	Yahoo! auction
479505	528067	104	87	11 min.	33 min.	Yahoo! auction
691911	802184	281	242	6.5 min.	27 min.	general issues
691911	802184	180	131	12 min.	9.8 min.	TV and radio
691911	802184	129	94	11 min.	4.7 min.	entertainer
691911	802184	81	54	11 min.	7.7 min.	people in the news
691911	802184	66	49	6.3 min.	2.7 min.	baseball
691911	802184	47	43	5.1 min.	83 sec.	social issues
691911	802184	36	27	10 min.	15 min.	language
691911	802184	10	10	19 min.	10 min.	mental health
691911	802184	38	30	8.7 min.	4.2 min.	domestic issues

$N_{AA}(A_1, A_2)$ is the number of questions to which both user A_1 and A_2 submitted answers. $NE_{AA}(A_1, A_2)$ is the number of questions where user A_1 submitted answers before user A_2 did. $T_{QA}(A_1, A_2)$ is the median of submission time lags between questions and the earlier of their answers of A_1 or A_2. $T_{AA}(A_1, A_2)$ is the median of submission time lags between answers of A_1 and A_2 submitted to the same question.

802184) were detected in each case. The result shows that there are few users who used secondary accounts for repeating both TYPE QA and TYPE AA submissions. This is because inadequate users do not want to draw attention to themselves.

Next, we investigate the detected 6 user pairs. Table VI shows the number of answers and best answers of them. On the other hand, Table VII shows the deviations of answer submission order of them. We determined that each user pair was one and the same user. We show the reasons in detail below.

First, we discuss user pair (691911, 802184). In Section VI, we determined user 691911 and user 802184 were one and the same user. This is because the median of submission time lags between their answers in social issues category was only 83 seconds although user 691911 submitted answers at different times of a day. Furthermore, as shown in Table VI, user 691911 received 146 answers from user 802184, and selected 143 answers of them as best answers. Also, as shown in Table VII, this user pair had large deviations of answer submission order in 9 categories. In addition, user 691911 submitted many questions about the meaning of technical terms. For example,

[Q: 691911] I have a question about earthquakes. What is the difference between earthquake intensity and magnitude?

[A: 802184] I show you a nice site for information about earthquakes. Visit the site at http://www.kishou.go.jp/know/faq/faq7.html

http://www.kishou.go.jp/know/faq/faq7.html is the uniform resource locator (URL) of the frequently asked questions (FAQ) page on the site of the Japan Meteorological Agency. User 691911 submitted 10 questions (including this question) in a row and received 10 answers from user 802184, each of which showed the URL of the Japan Meteorological Agency FAQ page. User 691911 selected all these answers as best answers. The reason why this user submitted trivial questions like them is thought that the user wanted to increase the number of his/her best answers without drawing attention to himself/herself.

Next, we discuss user pair (236956, 32780). As shown in Table VI, user 236956 received 282 answers from user 32780 and selected 275 answers of them as best answers. Also, user pair (236956, 32780) had large deviations of answer submission order. Furthermore, user 236956 received the following two answers from user 443403.

[A: 443403] I am quite sure that the first answerer will commit a criminal act!

[A: 443403] Are you guys all fools??????????????????

In both answers, user 443403 criticized other answerers. It is noteworthy that user pair (443403, 32780) was detected by binomial tests based on Hypothesis QA1, QA2, and QA3. We think that user 32780, user 236956, and user 443403 are one and the same user. It may be not enough for the user to select his/her answers as best answers and show that he/she is a good answerer. The user could not help but criticize other users. In other words, the user did TYPE QA submission and TYPE AA submission at the same time.

Next, we discuss user pair (494870, 434516) and user pair (302153, 434516). User pair (302153, 434516) was not detected when the significance level was less than 5×10^{-6}. However, we think user 302153, user 434516, and user 494870 are one and the same user. This is because we found examples where the user did TYPE QA submission and TYPE AA submission at the same time. For example, user 494870 received an answer in which user 434516's answer was criticized. Then, user 302153 submitted the following answer as if he/she was criticized.

[A: 302153] What are you talking about? Can you show BA right now? This is a grammar problem. "need to study more", who do you think you are?

User 302153 submitted 14 answers to user 494870's questions and in many of them he/she criticized other users as he/she did in this answer.

Finally, we discuss user pair (652657, 528067) and user pair (479505, 528067). As shown in Table VI, user 652657 received 199 answers from user 528067 and selected 95 answers of them as best answers. Also, user 479505 received 110 answers from user 528067 and selected 55 answers of them as best answers. In both cases, the best answer ratios were unusually high, however, they were lower than those of the other 4 detected user pairs. As a result, in order to determine whether user 479505, user 528067, and user 652657 are one and the same user or not, it is important to discuss the reason why the best answer ratios were low. Until August 2005, user 652657 submitted his/her questions to various categories, such as adult issues category and baseball category, and received 22 answers from user 528067 and selected 20 answers of them as best answers. On the other hand, from September 2005, user 652657 and user 479505 submitted their questions mainly to Yahoo! auction category and often did not select user 528067's answers as best answers. In Yahoo! auction category, user 652657 and user 479505 often submitted questions about violative items in Yahoo! auction, especially, used underwears. For example,

[Q: 652657] Oh, No! Violative item! Do something! [URL]
[A: 528067] So, report it by yourself. [URL] [URL]

[URL] is the URL to the page of the violative item. User 528067 showed the URL twice in this example. The question seemed to be the report of the violative item in Yahoo! auction. However, many other users thought that the user aimed to advertise the violative item. As a result, many answers of protest were submitted to user 652657's subsequent questions.

[Q: 479505] I found a traditional store in this field. How amazing! [URL]

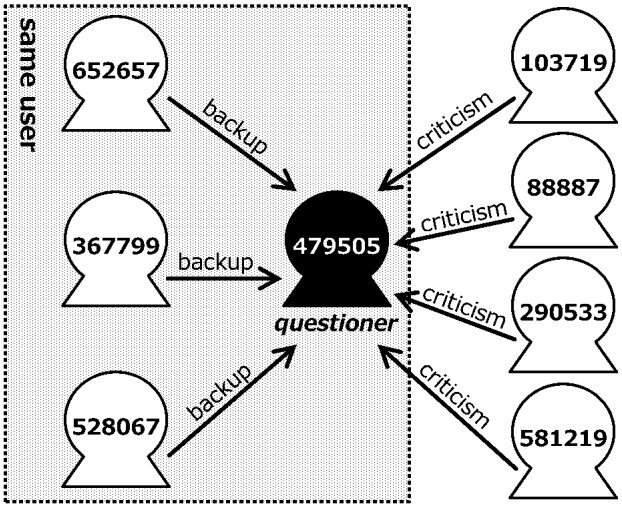


Figure 7. The relations of users who submitted answers to user 479505's question.

[A: 103719] You dummy. So what?
[A: 652657] Prestige store !!!!!!!!!!!!!!!!!!!!!!!! [URL]
[A: 88887] No advertisement. No advertisement. No advertisement.
[A: 290533] What do you mean? I didn't get your point. Is it because the seller received high evaluations?
[A: 581219] Advertisement?
[A: 367799] As long as there are violations, please report them. I'll be rooting for you.
[A: 528067] Get it eliminated. [URL] [URL]

In this example, user 479505 received 7 answers from 7 users. 4 answerers of them criticized user 479505. User 88887 and user 581219 accused user 479505 of advertising violative items. Also, user 103719 and user 290533 criticized user 479505 for submitting the ulterior question. On the other hand, two of the other 3 answerers, user 652657 and user 528067, were suspected of being the same user. Also, user 528067 was suspected of being the same user as the questioner, user 479505. Furthermore, user 367799 was suspected of being the same user as user 528067. This is because user pair (367799, 528067) was detected by binomial tests based on Hypothesis QA1, QA2, and QA3. As a result, user 367799, user 479505, user 528067, and user 652657 were suspected of being the same user. Figure 7 shows the relations of users in this example. Until September 2005, the user seemed to select his/her answers as best answers eagerly. However, in the example above, the user seemed to submit one question and three answers although he/she could select only one answer as a best answer. Actually, in this example, user 367799's answer was selected as a best answer. It is thought that his/her submission reason was changed from the best answer collection to the advertisement of auction items. As a result, the best answer ratio was lower than those of other user pairs suspected of being inadequate multiple account users. In our previous studies [7] [8], we put off the decision about whether

user 479505, user 528067, and user 652657 were one and the same user or not because the best answer ratios of them were lower than those of other detected user pairs. However, by introducing new points of view, we determined that these users were the same user. Furthermore, we collected the examples that show that the advertisement is one of the reasons for using multiple user accounts inadequately. As a result, it is important to investigate inadequate multiple account users from various points of view.

VIII. Conclusion

In this study, we investigated users suspected of using multiple user accounts and manipulating evaluations of their answers from various points of view. We first discussed reasons why users in a Q&A site use multiple user accounts. We think many users use multiple user accounts reasonably and properly, however, some users use them improperly. For example, there are many users suspected of using two or more user accounts for submitting questions and manipulating evaluations of their answers. In order to detect inadequate multiple account users, we proposed two detection methods [7] [8]. However, little is known about the purposes and methods of inadequate multiple account users. Actually, we found some suspicious users in our previous studies, however, it was difficult to determine whether they were inadequate multiple account users because of insufficient information. To solve this problem, it is important to investigate them from various points of view. As a result, in this study, we introduced two new points of view for investigating inadequate multiple account users:

- the number of user accounts for submitting questions and manipulating evaluations of their answers.
- the deviations of answer submission order.

Then, we investigated users in the data of Yahoo! chiebukuro from these points of view and found that

- the introduced points of view were useful for detecting inadequate multiple account users precisely,
- there were few users who did TYPE QA submission and TYPE AA submission at the same time, and
- some users might use multiple user accounts for purposes other than evaluation manipulation.

We intend to use the results of this study for further investigation of purposes and behaviors of inadequate multiple account users in Q&A sites. Especially, we intend to avoid unnecessary frictions between users in Q&A sites by detecting and taking care of these inadequate users.

References

[1] H. Matsumoto, Y. Watanabe, R. Nishimura, Y. Okada, and S. Yamanaka, "Investigation of inadequate users in a Q&A site who use two or more accounts for submitting questions and manipulating evaluations," in Proceedings of the Sixth International Conference on Evolving Internet (INTERNET 2014), Jun. 2014, pp. 40–45. [Online]. Available: http://www.thinkmind.org/index.php?view=article&articleid=internet_2014_2_30_40036 [accessed: 2015-11-25]

[2] O. de Vel, A. Anderson, M. Corney, and G. Mohay, "Mining e-mail content for author identification forensics," SIGMOD Rec., vol. 30, no. 4, Dec. 2001, pp. 55–64. [Online]. Available: http://doi.acm.org/10.1145/604264.604272 [accessed: 2015-11-25]

[3] M. Koppel, S. Argamon, and A. R. Shimoni, "Automatically categorizing written texts by author gender," Literary and Linguistic Computing, vol. 17, no. 4, Nov. 2002, pp. 401–412. [Online]. Available: http://dx.doi.org/10.1093/llc/17.4.401 [accessed: 2015-11-25]

[4] M. Corney, O. de Vel, A. Anderson, and G. Mohay, "Gender-preferential text mining of e-mail discourse," in Proceedings of the 18th Annual Computer Security Applications Conference (ACSAC '02), Dec. 2002, p. 282.

[5] S. Argamon, M. Šarić, and S. S. Stein, "Style mining of electronic messages for multiple authorship discrimination: First results," in Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '03), Aug. 2003, pp. 475–480. [Online]. Available: http://doi.acm.org/10.1145/956750.956805 [accessed: 2015-11-25]

[6] R. Zheng, J. Li, H. Chen, and Z. Huang, "A framework for authorship identification of online messages: Writing-style features and classification techniques," Journal of the American Society for Information Science and Technology, vol. 57, no. 3, Feb. 2006, pp. 378–393. [Online]. Available: http://dx.doi.org/10.1002/asi.v57:3 [accessed: 2015-11-25]

[7] N. Ishikawa, Y. Watanabe, R. Nishimura, K. Umemoto, Y. Okada, and M. Murata, "Detection of users suspected of using multiple user accounts and manipulating evaluations in a community site," in Proceedings of the 6th IEEE International Conference on Natural Language Processing and Knowledge Engineering (IEEE NLP-KE'10), Aug. 2010, pp. 600–607.

[8] N. Ishikawa, K. Umemoto, R. Nishimura, Y. Watanabe, and Y. Okada, "Detection of users in a Q&A site who suspected of submitting multiple answers to a question by using multiple user accounts," in Proceedings of the fourth International Conferences on Internet Technologies and Applications (ITA 11), Sep. 2011, pp. 236–244.

[9] K. Umemoto, N. Ishikawa, Y. Watanabe, R. Nishimura, and Y. Okada, "Investigation of inadequate multiple account users in a Q&A site by considering deviations of answer submission order," in Proceedings of the Fourth International Conference on Evolving Internet (INTERNET 2012), Jun. 2012, pp. 51–57. [Online]. Available: https://www.thinkmind.org/index.php?view=article&articleid=internet_2012_2_40_30063 [accessed: 2015-11-25]

[10] A. N. Joinson, Understanding the Psychology of Internet Behaviour: Virtual Worlds, Real Lives. Palgrave Macmillan, Feb. 2003.

[11] G. Kambourakis, "Anonymity and closely related terms in the cyberspace: An analysis by example," Journal of Information Security and Applications, vol. 19, no. 1, Feb. 2014, pp. 2–17. [Online]. Available: http://dx.doi.org/10.1016/j.jisa.2014.04.001 [accessed: 2015-11-25]

[12] J. R. Douceur, "The Sybil attack," in Proceedings of the First International Workshop on Peer-to-Peer Systems (IPTPS '02), Mar. 2002, pp. 251–260. [Online]. Available: http://research.microsoft.com/pubs/74220/IPTPS2002.pdf [accessed: 2015-11-25]

[13] L. A. Cutillo, M. Manulis, and T. Strufe, "Security and privacy in online social networks," in Handbook of Social Network Technologies and Applications, B. Furht, Ed. Springer, Nov. 2010, pp. 497–522.

[14] L. Wang and J. Kangasharju, "Real-world sybil attacks in BitTorrent mainline DHT," in Proceedings of the 2012 IEEE Global Communications Conference (GLOBECOM 2012), Dec. 2012, pp. 826–832. [Online]. Available: http://dx.doi.org/10.1109/GLOCOM.2012.6503215 [accessed: 2015-11-25]

[15] N. Jindal and B. Liu, "Opinion spam and analysis," in Proceedings of the 2008 International Conference on Web Search and Data Mining (WSDM '08), Feb. 2008, pp. 219–230. [Online]. Available: http://doi.acm.org/10.1145/1341531.1341560 [accessed: 2015-11-25]

[16] G. Wu, D. Greene, B. Smyth, and P. Cunningham, "Distortion as a validation criterion in the identification of suspicious reviews," in Proceedings of the First Workshop on Social Media Analytics (SOMA '10), Jul. 2010, pp. 10–13. [Online]. Available: http://doi.acm.org/10.1145/1964858.1964860 [accessed: 2015-11-25]

[17] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (HLT '11) - Volume 1, Jun. 2011, pp. 309–319. [Online]. Available: http://dl.acm.org/citation.cfm?id=2002472.2002512 [accessed: 2015-11-25]

[18] "Amazon Mechanical Turk," URL: http://www.mturk.com/ [accessed: 2015-11-25].

[19] C. Akkaya, A. Conrad, J. Wiebe, and R. Mihalcea, "Amazon Mechanical Turk for subjectivity word sense disambiguation," in Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk (CSLDAMT '10), Jun. 2010, pp. 195–203. [Online]. Available: http://dl.acm.org/citation.cfm?id=1866696.1866727 [accessed: 2015-11-25]

[20] "Distribution of "Yahoo! Chiebukuro" data," URL: http://www.nii.ac.jp/cscenter/idr/yahoo/tdc/chiebukuro_e.html [accessed: 2015-11-25].

[21] M. Hollander and D. A. Wolfe, Nonparametric Statistical Methods, 2nd Edition. Wiley-Interscience, Jan. 1999.

A Review of Internet-based Technologies and Applications in the Food Industry

Saeed Samadi
Food Machinery dept.
Research Institute of Food Science & Technology (RIFST)
Mashhad, Iran
s.samadi@rifst.ac.ir

Abbas Pourzaki
Electrical Eng. dept.
Asrar Institute of Higher Education
Mashhad, Iran
a_pourzaki@yahoo.com

***Abstract*— In the modern world, information technology (IT) has been incorporated in most development activities. The food production industry is one of the recent industries to embrace IT in their major daily operations. The aim of this paper is to review the technical and scientific state of the art of internet-based technologies and future trends in the Food industry. These technologies are mainly classified into Radio Frequency IDentification (RFID) for supply chain management, quality and safety monitoring, e-commerce, robotics, Wireless Sensor Networks (WSN), and Geographic Information Systems (GIS). Since all emerging technologies are coupled with challenges, the study addresses both challenges and benefits of incorporating IT in the food industry. This paper discusses how IT can be integrated to enhance the safety and quality of food products. Internet standards and the role of social media and smartphone applications in agriculture and food industry are the other important aspects that are discussed in separate sections. Also, this paper highlights the potential of Internet-based applications, which may be efficient in future as well as the role of IT in support of sustainable food choices. The paper concludes by arguing that awareness be raised within the agro-food industry on the importance of the adoption of Internet-based technologies as a critical success factor in the twenty-first century.**

Keywords - food industry; information technology; Internet; RFID; standards; social media; e-commerce; SME.

I. INTRODUCTION

This paper is an improved and expanded version of the AFIN 2014 conference paper "Applications and Opportunities for Internet-based Technologies in the Food Industry" [1].

Information and communication technology has been defined as capturing, exchange, processing, transforming, and emission of the data. Its advances are based on penetrating of internet and mobile set in mass people all around the world.

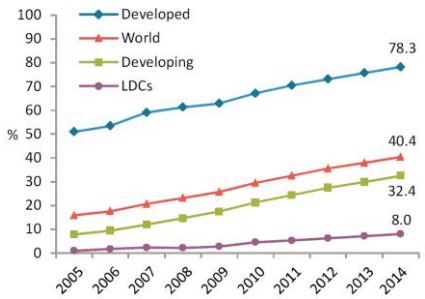


Figure 1. Individuals using the Internet, 2005-2014 [2]

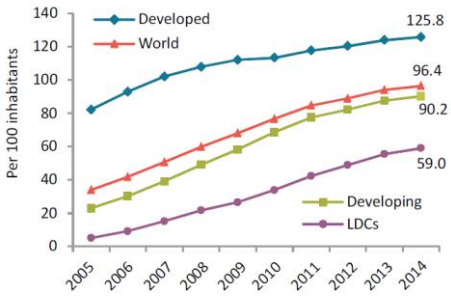


Figure 2. Mobile-cellular subscriptions, 2005-2014 [2]

Information technology (IT) is one of the individual forces that has contributed to globalization and advancement of life standards. These advancements have been occurring rapidly due to the rate of innovation from the IT industry. Significant incorporation of IT in most of the developmental activities is proof of the spread and importance of this technology. IT has been globally incorporated in construction industries, production, manufacturing, healthcare, education, information management, security, and food and agricultural production. However, IT has been embraced at different levels by the fields mentioned above. Information management ranks as the most advanced field concerning use of IT [3]. On the other hand, agriculture ranks as the least innovative field as far as incorporation of IT is concerned. Other fields of production besides agriculture and food obtain maximum potential production from their fields due to highly incorporated IT systems [4]. Unfortunately, agriculture and food production do not extract their maximum potential because of their low level of IT incorporation. Most of the yield available in the agricultural sector is retrieved from small and medium sized enterprises (SME). Therefore, high scale production firms from the food and agriculture production industry are not constructive parties in the business [5]. SME are characterized by either low or medium financial capacity. This financial background is not able to fully fund state-of-the-art technologies such as radio frequency identification (RFID), wireless sensor networks (WSN), and integration to e-commerce. These technologies are available for application by the food and agriculture industry, and once incorporated, agricultural and food production would be able to maximize its potential [6]-[8].

The rest of this paper is organized as follows. Section II summarizes a review of literature. Section III discusses the use of the most important and available Internet-based technologies in the food industry, the main aim of the paper.

It is includes: RFID, Tracking and Tracing, monitoring, E-commerce, Robotics technology, Wireless sensor network, and Farm management. Section IV discusses internet standards in the agro-food industry from EDI protocol to AgroXML and trace2p2 in finer detail. Section V includes applications of social media to smartphones. Section VI describes future internet based solutions and Section VII demonstrates information technology in support of sustainable food choices. Finally, Section VIII includes overall discussions and concluding remarks.

II. LITERATURE REVIEW

Currently, food and agricultural production has incorporated IT to a significant degree. Unfortunately, there still exist technical challenges that have resulted in the industry incurring losses and gaining a bad reputation. These technical challenges can be corrected through application of the mentioned technologies [9]. IT is signified by techniques that result in faster, efficient production with minimal human effort. Agricultural production is an economic activity that is more dependent on human input relative to machine input than other activities. This does not mean that technologies to minimize human effort and input in the industry are absent. Technologies that can result in reduction of human input exist in the industry, but the prevailing challenge is the cost of operation. As initially stated, SMEs comprise robust producers in the industry and lack sufficient capital to sustain these technologies [10]. IT applications relevant to the field of agriculture require high initial capital, but are cost effective. Areas within the field of agricultural production that can incorporate IT include: supply chains, harvest, standardization, marketing, soil fertility, and yield prediction [11]. These areas can be improved by the following technologies: RFID, WSN, GIS, robotics control, and e-commerce. These technologies are applied in the agricultural and food production industry to fulfill different objectives. These technologies utilize networks for communication. However, some technologies such as RFID have more than one application in the industry. It can be used in supply chain management and also in traceability for standardization [12].

III. AVAILABLE TECHNOLOGIES

In this section some main technologies on it-based food topics are discussed.

A. RFID Technology

This technology uses radio frequency to identify or retrieve information from production. It operates using the same mechanism as barcodes with magnetic strips [13]. Instead, of a barcode, RFID uses microchips that are embedded on the product of interest. RFID has two main advantages over barcodes. In the case of a barcode, it has to be on the line of sight of the barcode reader for information to be obtained from it. RFID is advantageous because the chip and the reader do not have to be on a line of sight to retrieve information from the chip, because the chip produces specific radio frequencies. The other advantage of RFID is that the chip is more reliable than the barcode [10]. This is because validity of barcodes is ruined once the code is scratched or removed. RFID microchips are not easily removed because they are not attached to the surface of the product.

RFID technology can be used in supply chain management and standardization. Food quality has been the cause of controversy in the food industry. Food that has not been properly stored has higher chances of going bad and once food has attained this status, it can become toxic. Food toxicity is dangerous as it can result in complex health disorders or even death. Therefore, a compromise on the quality of food is likely to ruin the reputations of the supplier and manufacturer, and this translates into losses [5]. RFID enables the user to establish the amount and type of ingredients contained in the food product. In addition, it also provides the time elapsed from the time of manufacture to the time of first use. This information is imperative to both the retailer and the consumer. Cases of food poisoning as a result of consuming expired food or allergic substances would be substantially reduced.

The other core challenge in the food production and agriculture industry is supply chain management. Some food products are essential for humans, but their production is unique in specific regions. Therefore, a comprehensive supply chain should be established so as to benefit both the manufacturer and consumer. The supply chain involves the food transit process from harvest, to processing, to distribution to the retailer [14]. Food undergoes this process before reaching the end user. Despite the extensive route, which is undergone before a product's use, monetary value has to be established. This means that the end user is not overcharged and the manufacturer is not underpaid [15]. RFID technology establishes an infrastructure that tracks a food product's location and ingredients, thus enhancing reliability of the end product. Farmers, specifically involved in food production, have been discouraged from expanding their investment due to limited profit from their enterprise. Previously, middle-men have benefited more than either the farmer or the consumer, minimizing profits to these constituents. Currently, with the employment of RFID, profits and satisfaction have improved because the supply chain of the goods has been bolstered by the technology. RFID technology has to be applied from the point of production (farmer) to the consumer. This reduces the bulk cost that could have been incurred by the distributor or supplier [16]. Wal-Mart is among the supply companies that have encouraged manufacturers to incorporate RFID to increase their profits. They encourage manufacturers through financing part of the RFID implementation. This practice is prevailing in most developing countries, as SMEs are financed to increase agricultural food production performance in the international market.

B. Tracking and Tracing

Traceability is the track or trace product ability in a production line, distribution processes, etc. It indicates that product flows are recognizable and can be logged uniquely and systematically (see Figure 3). The main reasons for growing attention to traceability in the food technology are new legislative requirements and the increasing demand for administrative systems and quality [16].

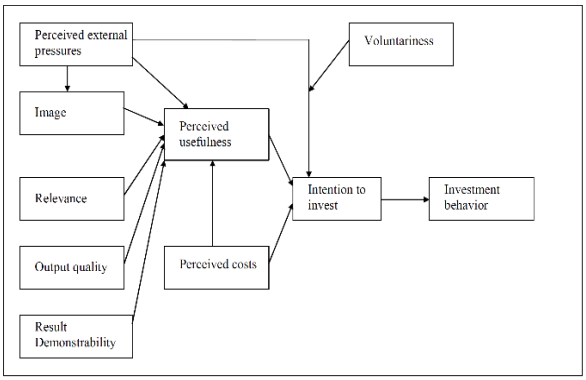


Figure 3. Tracking and Tracing Systems Investment Model [16]

Tracking and tracing are similar in some features but different in main characteristics. Tracing can be divided into upstream and downstream. Tracking follows the flow of downstream wares in in real-time (e.g., for creating situation data) while downstream tracing is not due to time. Tracing can be used in marketing or it can active goods recalls. Tracing has the ability to follow a product reverse direction, from the end product towards its start point. This can detect potential fail origin in food source networks [17]- [18].

In Japan, for safety assurance, a food traceability design has been constructed, which got knowledge about food production and distribution chains. Toward this object, the Japan Agricultural Cooperatives (JA), local governments, and the Japanese government have been incorporating in the proceeding of national projects, i.e., food traceability systems (FTS).

Filling the data forms of farm management is a labor work for almost farmers. Nowadays, Internet services and mobile phones are spread among peoples, especially farmers. Thus, for managing the agricultural production process an efficient "Farming Diary System" is developed. It use web compatible mobile phones [19]. As a commercialized version of this system, an internet software, "Application for Agricultural Methodological Analysis" (AFAMA) was developed in 2002.

Farmers, the major users of AFAMA, can insert their agro-food process information such as farm works and material applications in their cultivated fields while he/she do his activities in farm. The individuals can select input options such as materials, machines, fields, and crops. Using the AFAMA can be stepped as following:

a) Connecting a computer or mobile phone to the AFAMA Website log in using user ID and password.
b) Insert own data to customize one's materials, machines, fields, and crops.
c) Input the designs of material and cultivation and the operation data to choice the scheduled items.

C. Monitoring

Intervention in hot situations can be succeed if real-time data available. This avoids the cargo have been missed because of quality deficits. For example, fruits and vegetables transport from the countries of Southern European to Germany suffer transport time of about 50 hours (Spain to Germany). In this interval, the quality of products can decreases considerably under undesired conditions. In failing cooling system the monitoring system generates instantaneous information to the owner of the cargo. The cargo can be delivered to a near cold storage until the cooling system of the vehicle have been repaired.

Advances in the sensor network adding to communication allow the ability to real time observe of trucks positions and status. These technologies are developed more by new logistic service suppliers. These service suppliers get real time monitoring data and are able to convert them as a service for fresh produce customers. Especially, distribution chain data such as approximated arrival times and the transport status activate development in the commercial system of distribution shops. These developments decrease time of arranging reception of products and the shipping of them from distribution shops.

The position due to quality monitoring is such that cause the communication and processing will be unsuitable. High quality monitoring of food needs complicated sensor and information communication technology. Fresh goods sense out changes.

Temperature and humidity are the most important parameters for maintain desired food quality. They must be controlled and communicated within the information service, throughout the distribution chain from start (harvest) to end (retail outlets). Other interests in monitoring are far from forward product quality and include more environment data (e.g., CO_2 or NH_3 emissions).

Food quality and safety monitoring

In China, profits in the food industry have declined by 50 percent as a result of contaminated food. This shows the sensitive nature of the food industry as a single flaw has the potential to bring down the whole industry. In addition, China's food production is also mainly extracted from small scale investors. This means the sector is not fully exploited. One of the technologies that the country has embraced to enhance the status of the food supply is RFID food packaging. This technology uses RFID to identify the ingredients and the inventory of food products [21]. It utilizes disposable biosensors that produce an antigen-antibody reaction to identify any bacterial cells in the food product. When bacteria thrive in an enclosed food product, the result is a bio-chemical reaction that would either make the food product stale or poisonous. Therefore, this technology helps the food industry upgrade their monitoring systems, and the quality and safety of food products is enhanced. As a result, IT has aided in the restoration of trust between consumers and manufacturers. Furthermore, since the introduction of RFID food packaging, the number of health issues associated with food poisoning or food quality has declined by more than 50 percent. RFID technology also has an additional use as biosensors used in the tags containing inventory information that can be used in supply chain management. Traceability of food products from the farmer to the consumer is the other main concern in the food production industry. Effective supply management is a barrier that prevents SMEs from maximizing their potential. RFID detection technology poses a remedy to this barrier; RFID stores ingredients, destination, and the appropriate

geographic location of products [21]. This helps the food industry realize their market extent and as a result increase or reduce their production where necessary, thus minimizing losses. This technology enables rapid detection of poisons or derailed quality of finished food products. It also enables automatic identification of food products along a supply chain.

D. E-commerce

Internet technologies within the context of e-commerce have provided a more interactive market that enhances communication between manufacturers and consumers. This can be accomplished through existing social networking sites such as Facebook and Twitter. Manufacturers append their social networking websites on containers of food products so that in case of a complaint or compliment, the user can directly communicate with the food company [22].

When there is a reliable communication pathway between service or good providers and the end user, performance of the product is likely to be high. This is relative to a scenario where there is no elaborate communication between the user and manufacturer. IT provides better database management systems that portray the accurate needs of consumers. E-commerce expands the food market as the Internet is able to establish new consumers from regions where a specific food product has not yet been sold. E-commerce serves to benefit SMEs more because of their otherwise insufficient capital to market their food products. E-commerce is cheaper than hiring a marketing firm. This system requires less than five users to conduct online marketing and thus is affordable for SMEs [23].

As a result, SMEs can access a larger market without seeking additional financial assistance to facilitate marketing. Therefore, Internet technologies ensure development of a more reliable supply chain, higher quality food products, and a larger market for food products [24].

SMEs and consumer groups in local agribusiness

The main activities in retail is focused on decreasing the number of firms controlling portions of the retail bazar [25]. It is motivation of agrifood SMEs to give competitive advantages in their businesses. Some of their ways were joint initiatives and new methods for co-working [26]. A new method to exploit the benefits of repositioning is the new food networks for exchanging coalitions of consumers and clusters of agrifood producers in order to overcome the large-scale chain of retails. These arrangements have become increasingly important as a tools of decreasing cost, or mitigating economic risk [27].

If the compatibility of a commerce ICT proposal allows to purchase/sell agrifood goods more economic and efficiently, the main activator is "trust". Trust is a key in e-commerce and that its lack is one of the main factors for buyers to not use the Internet market. From the view of customers, for an online business trust in sellers is required but not sufficient [28]. It is investigated that customers trust almost are based on the reputation of the intermediary, web usability, and transaction security. So, intermediary should guarantee completion of the transaction chain and the sellers trust.

The trusted third party (TTP) is a systematic unit that uses equal trust from both the agrifood suppliers and buyers. Its goals are influence on innovation decisions, facilitate transactions, systemize the agrifood trade network, and, more important, provide and manage the 'trust structure', where e-commerce motivated by agrifood supplier clusters and buyer groups can be sufficiently guaranteed.

TTP has three roles in intermediation between suppliers and buyers groups:

- Technology intermediary, whose duty is to supply the ICT components such as security, hardware, and communication;
- Transaction intermediary services such as consultancy, hosting, and applications software. Among its roles, there are the co-working and direction of data flows (from agribusiness to consumers and vice-versa) and the logistic process (from agribusinesses to consumers application).
- Ensure authority that defines an "ethical code" and behavioral laws in transaction chains.

In practice, it reserves cumulative buying orders from consumers and reform individual orders for each agrifood producers cluster by analysis them. Once products get from clusters, they have been packed due to each consumers order and send to the primary consumers group location;

TTP is a main tool in obtaining the commitment of future participants, as groups and/or individuals, to insert the aggregation. It supplies a wide governance duty, actives the promotion of co-working among groups and controls interactions to guarantee transactions behavioral correctness in the network nodes.

The consumers group (CG): it is a special no-profit buying group that is self-systemized. The members are typically households, who need to buy cheap agrifood products. However, due to advances in discovery of regional traditions as tools to trust the quality and origin of goods, consumers request to gain secured local foods, where trace and producers reputation can be obtain assure and trust. For accessing purchasing power to gain the favorite traditional off, consumers must share "shopping lists" in one order. In creating the cumulative order, the CG follows previous ethical code and behavioral laws.

E. Robotics technology

Opposed to human input, machine input (as a result of IT) is both uniform and reliable. In addition, it is faster and produces more profit than human input. Robotics is applied in land preparation, planting, and weeding [24]. A series of corporative IT devices can sufficiently handle agricultural production leaving human application to solely play an oversight role. A combine harvester is one of the machines that has replaced human involvement in harvesting activity (see Figure 4). In cases where the machine has been used, there has been a greater than 100 percent advantage in yield compared to regions where human effort was used in harvesting. This prevailed in areas with the same size and climatic conditions.

Figure 4. Diagram showing robotic harvesters

F. Wireless sensor network

Wireless sensor network (WSN) differs from RFID in that it is able to integrate with other network devices in the field while an RFID tag can only be read with the RFID tag reader. WSNs comprise of Wi-Fi, Bluetooth, and ZigBee. The latter two operate within the Industrial Scientific and Medical (ISM) band of 2.4 GHz, which provides license-free operations, enormous spectrum allocation, and global compatibility. Other devices deployed on a farm to aid agricultural activities [30].

Diagram showing robotic harvesters is used in this industry for monitoring and surveillance of crops within a farm. However, weather variation is the sole challenge that affects performance of WSN in the agriculture industry. The technology utilizes radio frequencies that can be interfered with by weather conditions [31]. The technology is used in maintenance and monitoring of farmlands. This is achieved through installation of sensors and cameras on the field. These devices are linked to the control station on the farm via the mentioned wireless technology. Monitoring fields enables identification of severe conditions on the soil and weather. When soil moisture content is below the minimum, the information is transferred to the control that commands the irrigator to sprinkle the soil. Phytophotra is a disease that affects potatoes and is influenced by temperature and humidity conditions. Between 868MHz and 916MHz, motes can be used in determining moisture content on air and temperature [32]. Extreme temperatures can be reflected and relayed to the control station, which initiates spray of pesticides.

G. Farm management

In the new societies, from any place the farmer can connect to the network. It is done by powered wireless links. Nowadays, all farm equipment, machinery, and animals are supplied with various sensors and computers, which connected to the network; so he/she can observe any data in all point of the farm.

In East Asia, i.e., Japan and Korea, greenhouse farming has been extended. Shin et al. access the data by remote control by connecting the computer of the controlled greenhouse weather to the Internet [33]. Nonomiya and Kouno developed similar conjuration with a remote camera to visualize the crop conditions impression as a valuable marketing aspect [34].

The microchips can compute in much farm machinery and equipment, e.g., automatic feeding, in-house weather control. These computers will have networking so easy in the future. The farm boss will have remote control ability, and he may monitor the system functions anywhere in the farm.

A particularly non-neglectable extension may take place in animal husbandry [35]. A microelectronic herdsman can get information about the position by GPS (Global Positioning Systems) and about the health and welfare situation of cows, sows or sheep. In hot conditions, the microelectronic herdsman retains the animal at a feed suspensor and alarm the farm staff by a mobile set call.

Train animals respond to impulses of the electronic herdsman, which forces a suitable behavior of the animal. The perspectives of such a development are increasing animal welfare and higher outdoor care, because the need of keeping a control on animals systems, in which animals are integrated in small regions. This technology allows a cheap goods with no animals' natural need for free space and movement or the observation of the animals' welfare and health.

The farmers can communicate with out sources via networked farms, at any position and time. For example, the farmer can obtain weather updated observations data and its decision support throughout the day. For precise plant protection, the chemicals apply abilities can be enhanced. Plant weakness and illness alarming systems are new applications for local offices by the Internet.

Some advice can reach to former about where and when by natural speak powered by photos or video recordings; time and position are saved automatically. Then the problem can be send by email to the extension operator and receive a response later, or they may dialogue with each together.

GIS applications

A GIS uses unique colors and shades of colors to represent different atmospheric and soil conditions. It also uses the same set of unique colors to depict different terrains and ground cover. They utilize satellites to obtain aerial images of the Earth's surface. These satellites exist exclusively for GISs as the colors of objects and surfaces are different from ordinary depiction and representation. For instance, a water body would appear blue from ordinary satellites, whereas a GIS satellite depicts water bodies in dark. Food production and agriculture is governed by atmospheric conditions and soil fertility. Globally, farmers' yields are affected by changes in weather and climate. This is because of poor decisions that are dependent on farm activities [33].

For instance, harvesting time is signified by dry weather and medium to high temperatures. Therefore, when a farmer harvests during other atmospheric conditions, the resultant yield will be low. Through GIS technology farmers have been able to obtain atmospheric conditions in real-time diagram of remote-sensed image of the soil (GIS image) that have enabled them to conduct farming activities appropriately.

GIS images are specific to natural, physical features. As a result, farmers or investors are able to locate ideal regions that will favor their agricultural investments.

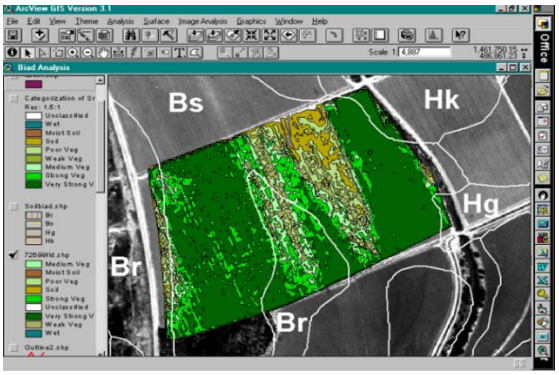


Figure 5. Diagram of remote-sensed image of the soil (GIS image)

Planting on the wrong soil will result in lower yield and losses. Analyzing a soil using the naked eye does not necessarily yield an accurate description of the soil's composition [37]. Therefore, advanced methods induced by IT would result in better land use, thus maximizing yield from the food production and agricultural industry (see Figure 5).

Mapping technologies

The accessibility of Internet-based atlas facilities (for example Google Maps) have motivated home location-based services as a whole. These are based on GIS tools to accumulate, form and register spatial records and maps. Such tools incorporate and spread over spatial and non-spatial databanks for combined examination, and are used to interconnect and bit geographic information.

In nutrition aid, as shown in Figure 6, these tools funding procedures related to, food security examination, eventuality and crisis planning, first cautionary organizations and logistics arrangement, by incorporating records taken in the ground with spatial records to funding result- making.

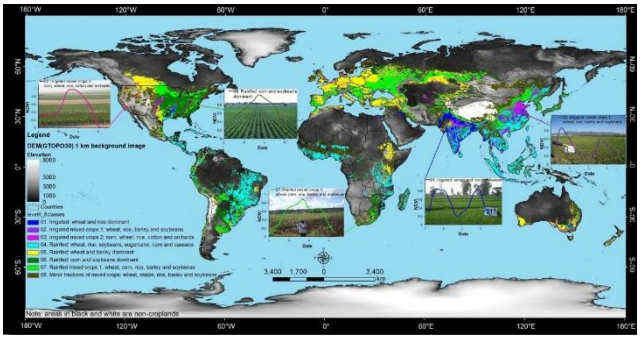


Figure 6. Map of global croplands and their water use for food security

IV. INTERNET STANDARDS CONSIDERATIONS

In this section some important notes on it-based standards in food industries are described.

A. *Electronic Data Interchange*

Electronic Data Interchange (EDI) is the technique of using computers to exchange commercial docs among firms. Before, facsimiles or customary mail was used to exchange docs. Mailing and faxing are still used in commercial, but EDI is a much quicker method to get the same object. EDI is used by an enormous number of companies. Over 100,000 businesses have substituted the more old-style means by EDI. This innovative scheme has a number of aids; price is one of them. Computer to computer exchange is much cheaper than old-style means of doc exchange.

EDI structures are expected to be Internet-based in the future. Although now only very insufficient initiatives use Web EDI, over 50 percent of the initiatives strategy to appliance this technology in the future. A main stage in decrease of costs of EDI answers is the use of the Internet with its current communication substructure as a means of carrying for EDI messages. A number of dissimilar communication rules can be used for the transmission of EDI messages via the Internet. Depending upon the task, the exchange can be made via FTP (File Transfer Protocol), HTTP (Hypertext Transfer Protocol) or SMTP (Simple Mail Transfer Protocol), whereas the files are encoded either with PGP (Pretty Good Privacy), S/MIME (Secure Multipurpose Internet Mail Extension) or SSL (Secure Socket Layer). Cheap combination with the huge number of emerging technologies for the Internet possibly will create a significant involvement to the rise of EDI users and particularly permit SMEs to join in EDI networks. The clear benefit of using the Web as a tool for EDI communication is that the single precondition is an Internet linking and a web browser. All communication uses the omnipresent HTTP-protocol. Safety subjects can be addressed by using SSL, for instance. Thus, all mandatory substructure is most possibly nearly anyplace accessible minus obliging the partners to participate bulky amounts of cash. In this situation, form-based EDI evidences to be a respectable hint for large businesses looking for means of having their small users send their documents in a regular layout.

B. *Extensible Markup Language*

The extensible markup language (XML) has the ability to be the information format of choice used together with the programming language of choice for the Web, Java, to allow the next stage in the development of EDI. Using open standards can significantly decrease the time and money consumed on realizing a solution. By escaping exclusive formats, the risk of asset ruins is decreased and future-oriented resolutions can be advanced. Though traditional EDI connections are regularly long-term and highly integrated, which are meaningful solitary with a large number of connections and for a long time, the readiness to devote into open, compatible IT-substructures is robust at any point of the Value Chain. Conventionally, the founding of compatibility between dissimilar EDI solution-systems was recognized over deep mixing of the EDI standard into the applications of the communication partners. XML is a guileless, very elastic text format resulting from SGML (Standard Generalized Markup Language). Formerly aimed to encounter the challenges of large-scale electronic publishing, XML has also a significant role in the exchange of an inclusive diversity of information on the Web and in another place. XML shares common origins with HTML and SGML (ISO 8879). The latter was envisioned for semantic rise that would assist computer classification and indexing. SGML delivered flexibility that had not been obtainable earlier and turn into very general beyond the

intentions of the original inventors. But it was very difficult and costly.

C. EDI - XML technologies in the agri-food industry

The continually cumulative requests for the level of tracing and tracking sequentially require even more innovative IT methods. Growing the speed of a tracing process is frequently the motive to accumulate data about loads and tracing in an ERP system, but these systems can also be applied to RF scanning and for EDI to communicate with customers and providers. In addition, speed also raises the efficiency of registration that guarantees recording the information totally suitable. Moreover, a data system that has been arranged well will decrease need to persons, since some data is kept in the information system itself. There are three main features of the donation of data:

- Identification of the lot, which is to be observed. A lot has an exclusive identity made by the mixture of item quantity and lot quantity.
- Recording and management of the lot past when provider send the raw material and when it was used in production, and so on.
- Link about the lot with other nodes in the chain network. EAN standards for barcodes and EDI have a significant role.

D. AGRO EDI Europe (AEE)

Later than 1992, Agro EDI Europe is working on association and regularization of information interchanges on the agricultural and its industrial parts. Nowadays, the association collects about 250 members from several sectors (farming input, agro equipment, management and accounting Sections, money support and reservation, insurance, packing, storage, quality control laboratories, etc.).

Since 2001, the commercial partners the Agro EDI Europe association introduced a regular data-processing format of exchange the data harvest page: the DAPLOS message. After rising this "plot message", AEE tries to solve the problems of supplement of farming works and computation of uncultured boundary, and benchmarking of harvest growing methods. Then, AEE faces to the domain of traceability, wholly tools allowing farm production administration being marketed nowadays as traceability resolutions. AEE formed the "plot message" (DAPLOS: Data Plot Sheet), which is a standard for relating data related to a particular social plot, to enable information exchange between different information systems. French software editors tried to twig to this standard though extending databases and programs, generally by executing some transfer data functions according to AEE message, which has focused consideration on the interactions of the farmer and service providers. Compared with AgroXML, the data description is not involved because it is an EDIFACT message.

E. AgroXML

Farmers must document agricultural activities. AgroXML is the consequence of a fitted assistance with agricultural software makers and online service suppliers, which mix AgroXML into their software. AgroXML presents a standard, which enables data exchange and storage. AgroXML is based on the worldwide standard XML and contains several content lists. Several claims in marketable software are surviving ever since two weeks. AgroXML is a language that assists the explanation of agricultural data that will assist a whole documentation of agricultural creation chains. The AgroXML object is to permit data exchanges minus dismissal among land proprietors, farmers, food industry, etc. AgroXML will be obtainable open on the web, and will be free of platform.

An XML Schema introduces electronic documents to exchange information. The AgroXML schema is founded on a model of the actual procedures in agricultural production. They are denoted in a tree-like hierarchy. Schema expansion will be taken place in the English linguistic. Presently delimited terms have previously been interpreted. Other subjects to be functioned upgrading of geo-data in addition to elements for farm animals and plants. At present, various corporations are applying AgroXML as a files exchange standard. Creators of agricultural tools like Claas and John Deere funding expansion of AgroXML. AgroXML has a high primacy for the agricultural software.

F. The benefits of AgroXML

Prospective users of AgroXML comprise someone along the making and providing process in the agricultural.

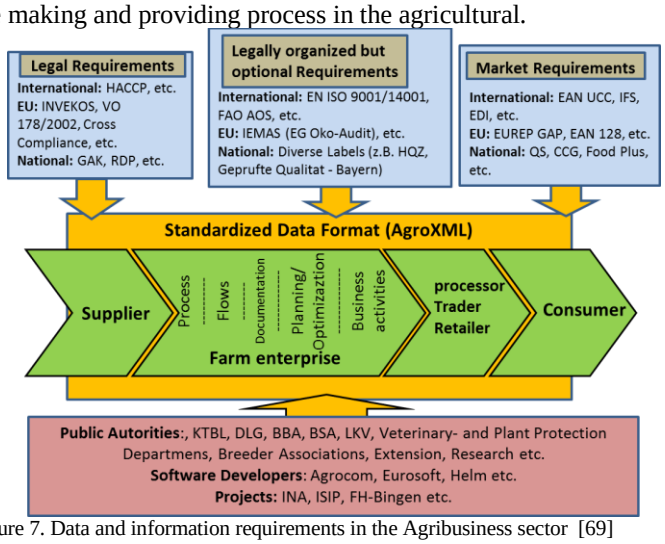


Figure 7. Data and information requirements in the Agribusiness sector [69]

1. Farmers

Hereafter farmers will not be knotted to a particular gathering or analyzing of data through the wide translation of documentation obligations. AgroXML does not register information and enables a struggle free communication with the farming management, accessing services and software corporations, without demanding an additional information input. Production procedures data, which must be stored, are accessible whenever must be used.

2. Consulting services

Combined vegetable production requires measuring agronomic must to be compatible to the local circumstances of the only one type. For adaptation there are many tools, available at the business position, or must be buy (monitoring section, weather station, warning tools, etc.). Farmers used the power of the combined vegetable production, if advisors and

development experts will access to type and specific oriented select aiding from internet. This method of obtaining advisor services makes advanced agriculture principally interesting [69].

3. Software companies

The requirement of a uniform information exchange protocol such as AgroXML will rise and quicken the development growth for farming software corporations. Agricultural software is progressively reliant on the information input from outer the commercial place. Although this worries the updating and care of forms and loads for request of subsidies, and the use of data relate the tools of the business. The matching updates need more work, so tend to more charge, which can be decreased greatly if this data will be accessible in a standardized format through the internet. Gradually, the clients will presume that the business software and type oriented online advisory services can networked suitably, so easy using AgroXML application.

G. Trace2p2

The P2P Project, co-supported by the European Community, goals to training, progress and test dissimilar procedures and apparatuses backup the speedy gathering of data from unlike traceability systems, in order to describe and extend a joined method to gather data from the different traceability systems working at the various corporates. In this method every corporation can pick its personal traceability system with fit to buyers, people and official personnel that will take a united line to ask automatically for goods information through Internet.

The key aim of the P2P plan are defining a Procedure (Trace Methodology) and a Software Style as the base of solution the crucial topics of incorporation. It means to regard and grow an answer intended at backup rapid gathering of data from dissimilar traceability systems of corporations of assumed farming cost procedure, presenting solution to the entire cost procedure traceability. An experimental system is developed for the swine value procedure. The most significant section of the software is the link standard to share data, i.e., the TRACE-XML protocol. The TRACE-XML defines the least set of data to trace all goods collection.

The TRACE-XML protocol will characterize the first stage to describe standard of automatic catching traceability data and documents of all corporation of the farming sector. For example, it could comprise of a subgroup of data and comments of the ebXML protocol. The TRACE-XML could denote an open protocol that could report the swine and especially all farming section. In brief, the key results of the project are:

- Trace Methodology: a procedure backup traceability data group in the corporations (especially what and how data);
- Trace-XML protocol: a least set of data to trace each collection and the associated communication standard;
- Trace-SW: a software assistant protocol and organization of data and documents associated to the traceability;
- Trace-Browser: a tool for browsing the traceability data obedient with the TRACE-XML protocol.

V. FROM SOCIAL MEDIA TO SMARTPHONE APPLICATIONS (APPS)

Each societal medium has an important role in passing data to publics via diverse internet tools, it has added improved usage in new times [38]. Social media service mobile and web-based tools to generate communicating platforms through which persons and groups share, debate, and adapt operator- produced content [39]. It includes a various kind of word-of-mouth, online forums including blogs, corporation backed argument boards and chat rooms, customer-to-customer e-mail, customer goods or service grades homepages and forums, Internet argument boards and forums, sites encompassing digital images, movies, or photos), and communal networking webs, to term a few [40] Administrations big and minor have hopped on the communal media fashion, feeling their method everywhere to make sense of its utility. They have peeped on Twitter, shaped pages on Facebook, and sent videos on YouTube [41] Internet has many sites for interacting users altogether and deliberate numerous viewpoints on nutrition. These variety from the farm and its situation of food making to the end artefact, investigation where, when and how to eat food, through furthermost sites based about an unambiguous objective or sphere. For example, Allotment Veg Growing [42] gives people data on vegetables, a forum for this and a webpage section for showing comment of users. Allrecipes.com [43] is similar, whereby people can add a storeroom to select recipes with ingredients and they can provide assured standards when browsing such as superior nutritional wants, season, or meal time kind. Using websites people can also instruction, catch and assessment food formations, JustEat [44] is a site that provides data on the adjacent carryout, whereby user can observe the menu, other user's comment and instruction the food to be conveyed. Urbanspoon [45] is a site to review eating locates, the people can division on liking the place, leave written comments and add data about the menu. EatPost [46] is like previous but it appraisals the meal, users send photographs of the plate with info about its origin place, other plates from the similar place, evaluations of the plate and proposals of extra alike founding.

Farm Blogs from Around the World has the purpose of gathering in one locate the best farm blogs, suggested farm blogs are requested to guide a short email on their farm and blog, and to contain private suggested farm blogs [47]. It provides attracted parties a gamble to study more about farming, with bloggers inclining to post at most one-time per day, and their posts tending to be one page length to less work for keeping up-to-date. Blogs are similarly obtainable that appearance at food in a more overall method, the Guardian's Word of Mouth Blog provides users data from the profits of red shoots [48], and the damage of Harry Ramsden's famous Fish and Chip Shop [49], these are all posted by dissimilar users sending the blog a broader comment.

Twitter is a microblogging site concentrating on donation real time updates with an middling of 90 million daily 'tweets', every one containing of 140 characters or fewer, these are typically small status updates of what people are undertaking, where they are, how they are feeling, they have confidence in

attention to the follower [38]. Many big food administrations have twitter pages, e.g., McDonalds [50], where the user can study about McDonald's history, food, persons and news everywhere in the world. It passes on data via twitters to 17,157 followers generous the positive social sensitivity, cooperating with a more 'human' singing than a 'corporate' singing [41] that McDonalds demands people to hear [50]. Burger King [51], a competing to McDonalds also has a twitter account and provides users data from the corporation. These corporations can show their wider worries through twitter, as shown from Burger King 'Through our BK CROWN Program, kids can learn about the National Parks Conservation Association. What's your favorite park?' (ibid.). There are also tweets that customers use for chatting eating food. Fastfoodie [52] in title of 'This is what we eat' where users tweet eating date and location online. Food.com [53] has 450,000 recipes, millions of reviews and photos, it is uses the twitter feed to direct people to the website for different recipes. There are also bloggers using these microblogs, where Irish Foodies is one of the largest who created a site for Irish and Irish-based food bloggers to link to others [54].

Food on Facebook is seen typically in two methods either via some initiative desiring to inspire the potential consumer to purchase their product or observe their locations [41]. This can been seen from the minor farm bazars, for example, Becketts Farm [55] , through to the larger supermarkets and professional shops, such as Camp Cupcake [56]. These corporations use facebook to shape a community and get members, they present new products, proposal discounts vouchers and competitive through the site. Facebook has different pages on awaring of food production, for example Farmers weekly, where one can find all the strangest news from the landscape [57]. Dairy Farming UK is a people cluster of all ages who support dairy agriculture in the UK [58]. There are more overall clusters who love Arable, Dairy, Poultry, Livestock, Horticulture, Mud, Tractors, Wellies, and The Crunchyside [59].

VI. FUTURE INTERNET BASED SOLUTIONS

In this section, a set of potential FI or electronic based results and expected applications were identified, which may be efficient in the future. In the focus clusters, these thoughts were discussed in order to try to acquire if a reason is applicable or not.

In Smart Farming area, the most applicable ideas uniformly were the "System for inessential and foreign bodies' detection ", the "Recommending system", and the "Development of the day-to-day work of the farmer "; beside these the idea of "Barcode/RFID system -Traceability system facilities" also was initiate significantly by the contributors of emphasis groups.

Recommending system for choosing the refined vegetation based on a databank

A large databank on diverse farming techniques must be accessible to notify every farmer about his/her refined crops. The system will compare the information provided by the farmers with the effects of dust trainings and principles and must recommend for vegetation, which could be grownup well on the explicit zone.

Watching surroundings for farms and vegetation - recommending system

The young farmers want access to a dependable and regularly updated watching / recommending system for their plants' or animals' health. The farmer should give information that may contain sensors' information, on and off time video, pictures, activities etc. The watching system should trigger an alarm when some thresholds have been passed; suitable advisory could be showed to the corresponding farmer in order to do further activities. For illustration if the temperature is too high and the humidity is too low, an approval can find a contractor for spraying, or a fertilizer contractor since those situations are hazardous for the vegetation to be infected by a illness.

Barcode/RFID system -Traceability system facilities

The farmer requires printing an elementary barcode tag for his/her ending product in the past it's storing or delivery. This barcode tag must encompass data such as the title of the corporation- farm, its origin area, the title of the product, the epoch of the production, etc. An internet based system may be beneficial for farmers with small production. They cannot spend a lot of money for resident software to make and printing the needed barcode tags. This automation must propose online barcode producer and RFID facilities, error tolerant without the load of managing hardware, organizing patches and upgrades, or watching performance.

Fundamentally, the accused named the most key functions of the FI as receiving more data (climate and ambient circumstances, dust conditions, etc.), and accumulating this united into an associated databank. Though, it should be considered that as a giant amount of data can be composed, it must can be shared. Accomplishment the right data or sharing the data and knowledge with the adjacent farmers - via a shared substructure - was similarly essential.

Many of these above stated systems are previously applied, but not common for their costs. Habits of such great numbers of sensors or applying such automatic interference systems have quite a high price.

A QR or RFID founded system, with developed traceability, is similarly excessively classy currently, though it is vital for the farmer. After the prices, two restrictions were recognized in the present exercise for a well- functioning, QR/RFID founded traceability system. One is the absence of a complex and public database, where the data can be composed and integrated. The other one is that in farming exercise it is rather hard to recognize a group or a smaller unit of crops in the harvest. In additional, Grease contributors agreed that farmers could use this type of systems if the program translated into the nationwide linguistic and had an actually modest user interface.

In Smart Agro-logistics area the maximum appropriate thoughts totally were the "Road watching application", the "Dock booking system" and the thought of "Combined cargo and navy organization for selling machines and small retail outlets" in a more wide-ranging way as "Combined cargo and navy organization in overall". The thought of "secure bank system" also was mentioned as being quite important.

Road observing application

The scope of this example is to share online monitoring information from trucks during the transport of cargo. Current practice allows monitoring trucks during transport with individual software applications and collects the monitoring data with available telematics systems. However, the access to such monitoring data is not organized on standards, which makes the exchange of data a complex task. Due to a divers spectrum of possible events disturbing the transport process (e.g., traffic jams or technical malfunctions) information needs to arise from uncertainties about arrival times and complications for further distribution planning as well as warehouse dock organization. The example shows an idealistic aggregation of information from different systems (order management system, online monitoring and event management system). This application can be opened for customers contracting a specific logistic service provider and enable a real-time event management in order to support decisions and planning.

Dock booking system

Current society systems of cross-docks are concentrating on first-come-first-serve philosophies. Online applications for dock booking are just applied for a short time. These applications permit reservation of dock places for a definite time ahead, but regularly need the recording up to 24 hours before passage entrance. The logistic service providers require an elastic solution, which has benefits for all joining enterprises. The procedure presented is based on the credentials of trucks and their mission in a particular topographical area (geo-fence) based on GPS directs around the storeroom/cross-dock. The telecommunication between storeroom and truck association needs the exchange of data on the truck (identification data based on license dish) and its duty (loading or unloading cargo) in addition to the recording and communication of a dock interplanetary and time epochs for the truck user approaching the storeroom.

Combined cargo and navy organization for selling machines and small retail outlets

Farmers expect a combined management system, which can improve logistics assets and to develop the typical control and production organization.

A software and/or internet reinforced stockholding and storing system - which assistances the corporation to enhance its stock, and the stock recording and the stock picking are automatic - is a common request by the people, but it is used previously at a number of corporations. Furthermore, this organization should grip the essential connections (worrying, re-ordering etc.) automatically too.

In small sales markets, which cannot reserves large stock such as selling machines and containers, it is expected that at the reduction of the stock to a regular level, an alarm indicator should be on or an automatic re-order should be produced for the provider. The problem of vending machines is that the provider has to convey lesser quantities of dissimilar products to numerous sites, the automatic orders encompass as much data as thinkable (what kind of product is wanted, in what quantity and how many portions can be served from the remained stock). Thus, the distribution way can be programmed after accumulating and processing the data from the diverse selling machines.

Refining the stock control is hope of the manufacturers, retailers and logistic service supplier too, then they could advantage by the better estimate and prediction (production plan, distribution ways), by the decrease of distribution and production budgets.

So, outside a GPS system, it needs a straight, real-time and long-range communication and information transmission among the solo units, the provider and the sole trucks of the provider. A sole selling machine should capable to show its data. This automatic alarming and re-ordering system may be used in smart households for refining the stock control in the storeroom or in the fridge, and for supplying input for the real shop list.

As a whole, all the particular applications have the same real profits as charge lessening, better organization and better data for choosing, and the practical control of procedures tending to cumulative efficacy.

Thoughts of a road watching application and the dock booking system essentially belongs to the opinion of navy organization - all these collected looked to be relatively applicable for two main reasons. First, numerous application previously exist and are applied, as GPS based navigation system, telematics systems or dock booking systems. Instead, most of the corporations considers needed in having these applications or systems for the easier association and more effective accomplishment of the conveyance procedures.

VII. INFORMATION TECHNOLOGY IN SUPPORT OF SUSTAINABLE FOOD CHOICES

Nowadays, attentions to application for data on agriculture and food choices are rising. Many local and international associations, events, and papers are developed to support its advance [60].

This kind of ingesting is a vital form of political commotion. A concentration sustainable ingesting may lead customers towards specific farming. Cohen [61] lists five contradictions intrinsic in the perception of sustainable ingesting: it encounters the universal delusion of customer dominion that occurs in most industrialized countries; the efforts to generate a distinct space for examination of sustainable ingesting are challenged by labors to place it in sustainable production area; superior ingesting efficiency and the development of technically "green" customer predilections together loan themselves to unwanted ricochet things; customer policymaking in greatest industrial countries is actual muddled and controlling answerability is fragmented; and public explanations of "customer interest" do not usually approve policy programs that intend to reduce the capacity of ingesting or the variety of elections.

Should one attentive in sustainable ingesting always choose organic, regardless of how it was created and qualified? The honestly new growth of large, conservatively accomplished, organic farms has produced subjects around strengthening. Guthman [62] pronounces how customary organic farms are challenging force to exaggerate their operations, which undercuts their skill to practice their accurate form of organic farming. Duchin [63] designates the

use of input-output and life-cycle examines in assessing food choices. For instance, Faist et al. [64] recognize that efficient cooling applications had more ability to drive requirements than a wholesale change to organically produced foods. Pretty et al. [65] do a cost analysis based on food miles and the various charges of conveying food over long road. The March 2, 2007 of TIME magazine issued that read "Forget Organic. Eat Local." Such a note is too unsophisticated, but the cover story [66] presented more steadiness. Merger of these numerous compound models in decision provision tools for customers interested in sustainable ingesting.

Globalization is an affluence of food choices, but it have a bad influence on food safety, security, and dominion. How does a customer choose between these options? Epistemic distance [67] is significant in setting primacies for customers, but how can one successfully progression all the variables? Waldfogel [68] discovers that customer dominion is necessary in convinced definite circumstances including choices between accustomed objects, but when user want to make inter-chronological ingesting choices, for example, "authoritarian involvements could advance their conclusions."

The producer must be sustainably operate before a customer can choose an ecological food item. Carolan [67] dialogs about the requirements to make the assistances of sustainable agriculture more noticeable to farmers. The assistances of sustainable agriculture are much less noticeable than those of conservative agriculture. Likewise, the charges of conservative agriculture are much less noticeable than those of ecological agriculture. What sustainable farmers might call "wildflowers" or "biodiversity", conservative farmers might demand "weeds". Cumulative the scale of thoughtful, from a solo farm to the entire watershed or food shed helps to visualize the benefits of ecological agriculture and the charges of conservative agriculture. Classification and "food miles" calculations are other techniques the corresponding benefits and charges can be made more noticeable. Caution is vital because of the incomplete viewpoint any one inventiveness can deliver.

Excessive transaction of potential in considering tend to use the internet for existing data in a method that develops easily criminal. The propensity is to fragment and the challenge is how to keep the whole thing together. This larger community must be addressed in some way. To get a model of an umbrella organization that can exist on the web and handle communication is also worthwhile. To move from a concept of "no wrong answers" to an available procedure for counting the sustainability of diverse food picks is significant. This data can be developed over the progress and use of complete computer models.

The internet can be used to advance the expediency of ecological food choices, for makers and customers by generating online clearinghouses for food objects that enable both commercial to commercial and commercial to customer interconnection. Cheap tries to assist traceability of products will expand the customer's belief in makers and affluence the acceptance of these another food chains. Although fruitful local food creativities may become, but everybody convoluted will progress individual relations with each other. So, as source chains grow, it is significant to stability local with expediency and sanctuary.

VIII. DISCUSSION

The primary goal of IT incorporation in the food industry is to foster food security and extract maximum sustainable yield. Once the primary role has been fulfilled, there are numerous secondary goals that IT ensures are effectively addressed. They include: processing, distribution, marketing, and storage [69]. IT, through the technologies previously discussed, fulfills each of these goals successfully. RFID ensures comprehensive results from supply management, which constitutes a secondary goal of the industry [71]. Regions that have incorporated RFID in their supply chains receive more revenue from the agriculture and food production industry than regions that have not applied RFID technology [72]. Similarly, regions that have incorporated WSN practice sustainable farming on a larger scale than in regions where the technology has not been applied. After production, the other barrier to extracting maximum potential from agriculture is the marketing of harvested goods. Large scale producers in farming have extensive marketing strategies that cover almost ninety percent of their produce. On the other hand, SMEs in agricultural and food production lack elaborate marketing avenues that can ensure intake of their products in the market. The first obstacle is the cost, which is a requirement for establishing an elaborate marketing network. The other obstacle is technology. Technology now offers a solution to its initial problem in that the Internet has contributed positively towards establishing global villages. Farmers are able to establish first person contact between the manufacturer and the user or processing firm. For instance, the Kenyan association of coffee growers has established a direct link to coffee processing firms in England and the United States. This ensures that farmers obtain maximum compensation for their products, and therefore represents an appropriate motivation for farmers to expand their farms. As a result, the potential of food and agricultural production is optimized.

Food investors who have embraced robotics and e-commerce receive more income from the food industry than food investors not aware of the technology or those who have shunned it. Consequently, in countries where these practices have been encouraged and are prevailing at significant levels have a better economy than in countries where IT application is limited.

Another added value of internet-based technologies in the agro-food industry is the improvement of efficiency and reactivity from real-time management of supply chains from farm to fork [73]. From a “food miles” point of view, this could result in a reduction in greenhouse gas emissions and in the carbon footprint, e.g., decrease of transport kilometers or empty vehicles, less waste, and better decay management.

The digital divide is a challenge that might hinder the applicability of the technologies discussed in this paper. Digital divide is mainly the gap between those with and those without access to ICT technologies and/or skills necessary to take advantage of ICT services. In addition, there is a widening gap between the urban and rural sectors on utilizing

advanced and emerging technologies [74]. To overcome this, measures should be taken to strengthen informatics in the agro-food industry by fostering the development of national information capacity and new databases, linking national and international databases, and adding value to information to facilitate utilizing them at various levels. Also, innovative ways of combining ICT-based information sources (such as agro-food information systems) with traditional ones should be considered.

IX. CONCLUSION

This paper gives an overview of major IT-based technologies and their impact on the food industry. It presents how selected fields of application can make a considerable contribution to food industry both in increasing efficiency and making data more available and easily managed. It discusses how these technologies can be integrated to enhance safety and quality of food products and provide advantages such as mobility, transparency and autonomy. The example technologies are mainly built on networked devices or utilize networks for communication. However, much additional work still should be done for a large scale integrated communication and scalable coordination throughout the agro-food networks.

Traditional industrial food manufacture has the objective of maximizing productivity and income. The motivation for acceptance of sustainable approaches comes first from customers who influence significantly on the obtainability of sustainable food choices by advantage of their purchases. By emphasis on food, ingesting efficiency and green partialities tend to very remarkable choices with doubtful sustainability.

Nowadays, farm smarting problems have been become more interesting and important for people. We can divide Smart Farming duties in main sections:

1. Recommending system for choosing the refined vegetation based on a databank
2. Watching surroundings for farms and vegetation - recommending system
3. Barcode/RFID system -Traceability system facilities
4. Shared substructure
5. Watching surroundings for animal welfare and sensors
6. e-commerce
7. System for inessential and foreign bodies' detection

Monitoring of product and their change is especially relevant for fresh products such as fresh fruits, warm meat, and daily vegetables, need monitoring of quality control in the distribution chain. Data of quality control is required to reserve the product quality during carrying.

In the last 20 years the structure of agrifood business industry has been improved considerably. For a lot of reasons such as advances in ICT, enhancing the customer quality prediction, nowadays hard functions, low agriculture commodity charges, the force of international retails and increasing of huge stores, all details of this industry has been changed and deformed. The disadvantage of Internet technology among SMEs is that the business owners and staff have to undergo training so that they can understand computer. This is an additional cost that a small scale investor aims to reduce by all means necessary. IT exposes SMEs to Internet hazards such as hacking and fraud, which can cause huge losses to investors.

Another important aspect of food networks, which is discussed in this paper is trade characteristics of a close relationship between customers and producers of goods/services using purchasing groups. Farmers can fill their agro-food process information forms on farm works, material applications, and so on, in their cultivated fields while he/she do his/her activities in farm. Using web connected mobile phones, the farmer store data in the databank. With this information, farmers make comprehensive decisions concerning planting activities. Wireless technology also enables pest control and irrigation activities that are essential when pursuing maximum yield. Sensors deployed on the soil are able to determine moisture content of the soil.

The utility of Internet-based atlas facilities have motivated home location-based services. Social media service mobile and web-based tools to generate communicating platforms through which persons and groups share, debate, and adapt operator- produced content. Weblogs naturally define a private diary, reserved on the web, which can be corrected by people with few web edition abilities, as well as blogs to increase wakefulness of food goods. An online public forum can combine a somewhat large amount of rich, including data on a subject. These societies can cover a range of dissimilar zones.

The paper also highlights that there is great opportunity for internet-based applications in developing countries. However, in most developing countries, strategies should be employed to overcome technical and societal barriers that can hinder further development of these technologies in agro-food sector. Therefore, it is a mandate of the ministry of agriculture and/or other governmental authorities to ensure IT techniques are being used in the food and agriculture sector to boost production and create an extensive market for the produced goods.

REFERENCES

[1] S. Samadi, "Applications and Opportunities for Internet-based Technologies in the Food Industry," In The Sixth International Conference on Advances in Future Internet, AFIN 2014, Lisbon, Portugal, Nov. 2014, pp. 67-71.

[2] “Measuring the Information Society Report,” International Telecommunications Union, 2015.

[3] N. P. Mahalik and A. N. Nambiar, “Trends in food packaging and manufacturing systems and technology,” Trends in food science & technology, vol. 21, 2010, pp. 117-128.

[4] Z. Abbasi, N. Islam, and Z. A. Shaikh, “A review of wireless sensors and networks' applications in agriculture,” Computer Standards & Interfaces, vol. 36, 2014, pp. 263-270.

[5] D. Prajogo and J. Olhager, “Supply chain integration and performance: The effects of long-term relationships, information technology and sharing, and logistics integration,” International Journal of Production Economics, vol. 135, 2012, pp. 514-522.

[6] A. Suprem, N. Mahalik, and K. Kim, “A review on application of technology systems, standards and interfaces for agriculture and food sector,” Computer Standards & Interfaces, vol. 35, 2013, pp. 355-364.

[7] M. Cinque, D. Cotroneo, C. Di Martino, S. Russo, and A. Testa, "Avr-inject: A tool for injecting faults in wireless sensor nodes," In IEEE International Symposium on Parallel & Distributed Processing, IEEE IPDPS 2009, May 2009, pp. 1-8.

[8] D. Martino, G. D'Avino, and A. Testa, "icaas: An interoperable and configurable architecture for accessing sensor networks," International Journal of Adaptive, Resilient and Autonomic Systems, vol. 1 , no .2 , 2010, pp. 30-45.

[9] J. Von Braun, The world food situation: new driving forces and required actions. Intl Food Policy Res Inst, 2007.

[10] M.E. Yüksel and A. S. Yüksel, "RFID technology in business systems and supply chain management," Journal of Economic and Social Studies, vol. 1, no. 1, 2011, pp. 53-71.

[11] J. Wolfert, C. N. Verdouw, C. M. Verloop, and A. J. M. Beulens, "Organizing information integration in agri-food—A method based on a service-oriented architecture and living lab approach," Computers and electronics in agriculture, vol. 70, no. 2, 2010, pp. 389-405.

[12] T. Kelepouris, K. Pramatari, and G. Doukidis, "RFID-enabled traceability in the food supply chain," Industrial Management & Data Systems, vol. 107, no. 2, 2007, pp. 183-200.

[13] E. A. Soujeri, R. Rajan, and A. Harikrishnan, "Design of a zigbee-based RFID network for industry applications," Proceedings of the 2nd international conference on Security of information and networks. ACM, 2009, pp. 111-116.

[14] A. Sarac, N. Absi, and S. Dauzère-Pérès, "A literature review on the impact of RFID technologies on supply chain management," International Journal of Production Economics, vol. 128, no. 1, 2010, pp. 77-95.

[15] J. Fontanella, "Finding the ROI in RFID," Supply Chain Management Review, vol. 8, no. 1, 2004, pp. 13-16.

[16] L. Theuvsen and T.Hollmann-Hespos, 2005, The Economics of raceability: A Model of Investments in Tracking and Tracing Systems in Agriculture and the Food Industry, Vila Real (Portugal), pp. 914-921.

[17] J. H. Trienekens and J. G. A. J. van der Vorst, 2006. Traceability in food supply chains. Wageningen Academic Publishers, Wageningen, pp. 439–470.

[18] M. Fritz and G. Schiefer, "Tracking, tracing and business process interests in food commodities: a multi-level decision complexity," International Journal of Production Economics vol 117, no. 2, pp. 317–329, Feb.. 2009.

[19] K. Sugahara, "Farming Diary System using Internet-enabled cellular phones". In Internet Workshop 2001, Proceeding II Applications. pp. 247-252.

[20] K. Butner, "The smarter supply chain of the future," Strategy & Leadership, vol. 38, no. 1, 2010, pp. 22-31.

[21] X.Zhu, S. K. Mukhopadhyay, and H. Kurata, "A review of RFID technology and its managerial applications in different industries," Journal of Engineering and Technology Management, vol. 29, no. 1, 2012, pp. 152-167.

[22] N. P. Mahalik, "Processing and packaging automation systems: a review," Sensing and Instrumentation for Food Quality and Safety, vol. 3, no. 1, 2009, pp. 12-25.

[23] M. J. Meixell, "Quantifying the value of web services in supplier networks," Industrial Management & Data Systems, vol. 06, no. 3, 2006, pp. 407-422.

[24] R. A. Tabile et al., "Design and development of the architecture of an agricultural mobile robot," Engenharia Agricola, vol. 31, no. 1, 2011, pp. 130-142.

[25] A. C. Vias, "Bigger stores, more stores, or no stores: paths of retail restructuring in rural America," Journal of Rural Studies, vol. 20, no. 3, pp. 303-318, July 2004.

[26] G. Schiefer, "New technologies and their impact on the agri-food sector: an economists view. Computers and Electronics in Agriculture," Industrial Management & Data Systems 106.1, 2006, pp. 5-20.

[27] P. Wang and A. Watts, "Formation of buyer-seller trade etworks in a quality-differentiated product market," Canadian Journal of Economics, 2006, pp. 971-1004.

[28] K. Myoung-Soo and A. Jae-Hyeon, "A Model for Buyer's Trust in the E-marketplace," Proceedings of ICEC'05, August 15-17, 2005, Xi'an, China.

[29] L. Ruiz-Garcia, P. Barreiro, and J. I. Robla, "Performance of ZigBee-based wireless sensor nodes for real-time monitoring of fruit logistics," Journal of Food Engineering, vol. 87, no. 3, 2008, pp. 405-415.

[30] A. Testa, M. Coronato, and J. C. Augusto, "Static verification of wireless sensor networks with formal methods," In Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference on, IEEE, November 2012, pp. 587-594.

[31] M. C. Cinque, D. Di Martino, and A. Catello Testa, "An effective approach for injecting faults in wireless sensor network operating systems," Computers and Communications (ISCC), 2010 IEEE Symposium on. June 2010.

[32] Baggio, "Wireless sensor networks in precision agriculture," ACM Workshop on Real-World Wireless Sensor Networks (REALWSN 2005), Stockholm, Sweden. 2005.

[33] J. H. Shin, W. S. Hahn, and B. W. Lee, "Development of greenhouse control environment analysis system on World Wide Web," Second European Conference of the European Federation for Information Technology in Agriculture, Food and the Environment, 27-30 September, 1999, Bonn, Germany, pp. 83-87.

[34] S. Nonomiya and T. Kouno, "Internet remote camera system and its application in agriculture," Second European Conference of the European Federation for Information Technology in Agriculture, Food and the Environment, 27-30 September, 1999, Bonn, Germany, pp. 75-81.

[35] W. Rossing,"Electronic animal identification. Computers and Electronics in Agriculture," (Special Issue), 1999.

[36] I. Thysen and P. K. Haneveld, "Evading murphy: A sensor network deployment in precision agriculture," O teu-Delft, Xuño 2007.

[37] T. Kalaivani, A. Allirani, and P. Priya, "A survey on Zigbee based wireless sensor networks in agriculture," Trendz in Information Sciences and Computing (TISC), 2011 3rd International Conference on. IEEE, 2011.

[38] D.Westerman, P. Spence, and B. Van Der Heude, "A social network as information: The effect of system generated reports of connectedness on credibility on Twitter," Computers in Human Behavior, 28(1), pp. 199-206.

[39] Kietzmann J.H., Hermkens K., McCarthy I.P. and Silvestre B.S., "Social media? Get serious! Understanding the functional building blocks of social media. Business Horizons," 54(3), pp. 241-251.

[40] W. G. Mangold and D. J. Faulds, "Social media: The new hybrid element of the promotion mix. Business Horizons" vol. 52, pp. 357-365.

[41] B. Weinberg and E. Pehlivan, "Soical spending: Managning the social mix. Business Horizons," 54, pp. 275-282.

[42] Allotment Veg Growing, Retrieved September 2, 2015 from http://www.allotment.org.uk.

[43] Allrecipes, 2011. Retrieved September 2, 2015 from http://Allrecipes.com.

[44] JustEat, 2011. Retrieved September 2, 2015 from http:// www.just-eat.co.uk .

[45] Urbanspoon, 2011. Retrieved September 2, 2015 from http://www.urbanspoon.com.

[46] EatPost, 2011. Retrieved September 2, 2015 from http://www.eatpost.com.

[47] R. Kwai Fun and C. Wagner , 2008. Weblogging: A study of social computing and its impact on organizations. Decision Support Systems 45, 242-250.

[48] O. Thring, 2015. Benefit of the (red) sprout. Guardian.co.uk Word of Mouth Blog, [Blog] 28 September. Retrieved September 2, 2015 from http://www.guardian.co.uk/ lifeandstyle/wordofmouth.

[49] M. Wainwright, 2015. Harry Ramsden's famous original fish and chip shop faces closure after losses. Guardian.co.uk Word of Mouth Blog, [Blog] 30 September. Retrieved September 2, 2015 from http:// www.guardian.co.uk/lifeandstyle/ wordofmouth.

[50] McDonaldsCorps, 2015, 28 September. Happy Cyber Monday! Scoring any good deals today? It's chilly in McDonald land so I'm gonna shop w/a McCafe Hot Chocolate in hand. :) [Twitter post]. Retrieved September 2, 2015 from http://twitter.com/ #!/McDonaldsCorp>

[51] BurgerKing, 2011, 28 September. Through our BK CROWN Program, kids can learn about the National Parks Conservation Association. What's your favourite park? [Twitter post]. Re trieved September 2, 2015 from http:// http://twitter.com/ #!/BurgerKing>

[52] Fastfoodie, 2015, 30 September. [Twitter Post]. Retrieved September 2, 2015 from http://twitter.com/#!/fastfoodie>

[53] Food.com, 2015, 30 September. [Twitter Post]. Retrieved September 2, 2015 from http://twitter.com/#!/Fooddotcom>

[54] IrishFoodies, 2015, 30 September. [Twitter Post]. Retrieved September 2, 2015 from (http://twitter.com/#!/IrishFoodies>

[55] Becketts farm, 2015, 30 September. [Twitter Post]. Retrieved September 2, 2015 from http://twitter.com/#!/BeckettsFarm>

[56] Camp Cupcake, 2015, 30 September. [Facebook Update] Retrieved September 2, 2015 from http://www.facebook.com/campcupcakecakes>

[57] Farmers weekly, 2015c, 30 September. [Facebook Update] Retrieved September 2, 2015 from https://www.facebook.com/farmersweeklyuk>

[58] Dairy Farming UK , 2015, 30 September. [Facebook Update] Retrieved September 2, 2015 from https://www.facebook.com/group.php?gid=6220759330#!/gro up.php?gid=41237147430>

[59] Farming, 2015, 30 September. [Facebook Update] Retrieved September 2, 2015 from https://www.facebook.com/pages/Farming/1021816?sk=info>

[60] D. Stolle et al., Politics in the Supermarket: Policitical Consumerism as a Form of Political Participation, International Political Science Review, Volume 26, Number 3, 2005, pp. 25-269.

[61] M. J. Cohen, Sustainable Consumption Research as Democratic Expertise, Journal of Consumer Policy, Springer Verlag, Volume 29, 2006, pp. 67-77.

[62] J. Guthman, The trouble with 'organic lite' in California: a rejoinder to the 'conventionalisation' debate, Sociologia Ruralis, Volume 44, Number 3, 2004, pp. 3-16.

[63] F. Duchin, Sustainable Consumption of Food: A Framework for Analyzing Scenarios about Changes in Diets, Journal of Industrial Ecology, MIT Press, Volume 9, Number 1-2, 2005, pp. 99-114.

[64] M. Faist et al., The impact of household food consumption on resource and energy management, International Journal of Environment and Pollution, Volume 15, Number 2, 2001, pp. 181-199.

[65] J. N. Pretty et al., Farm costs and food miles: An assessment of the full cost of the UK weekly food basket, Food Policy, Volume 30, 2005, pp. 1-19.

[66] J. Cloud, Eating Better Than Organic, TIME Magazine, Time Inc., March 2, 2007. http://www.time.com/time/magazine

[67] M. S. Carolan, Do You See What I See? Examining the Epistemic Barriers to Sustainable Agriculture, Rural Sociology, Volume 71, Number 2, 2006, pp. 232-260.

[68] J. Waldfogel, Does Consumer Irrationality Trump Consumer Sovereignty?, Review of Economics and Statistics, Volume 87, Number 4, September 2005, pp. 691-696.

[69] R. Doluschitz, 2004. Der Beitrag der Informationstechnologie zu Produktionsmanagement, Qualitatssicherung und Ruckverfolgbarkeit in der Agro-Food-Chain - Notwendigkeit, Strategien und Perspektiven: In: FAT-Schriftenreihe, Agroscope FAT Tanikon/CH, 2004, H. 59, 5-25.

[70] D. Restuccia, et al, "New EU regulation aspects and global market of active and intelligent packaging for food industry applications," Food Control, vol. 21, no. 11, 2010, pp. 1425-1435.

[71] M. Bolic, D. Simplot-Ryl, and I. Stojmenovic (Eds), RFID systems: research trends and challenges. John Wiley & Sons, 2010.

[72] M. Canavari, R. Centonze, M. Hingley, and R. Spadoni, "Traceability as part of competitive strategy in the fruit supply chain," British Food Journal, vol. 112, no. 2, 2010, pp. 171-186.

[73] G. Schiefer, R. Reiche, and J. Deiters, "Transparency in Food Networks – Where to Go," International Journal on Food System Dynamics, vol. 4, no. 4, 2014, pp. 283-293.

[74] R. Bertolini, "Making information and communication technologies work for food security in Africa," International Food Policy Research Institute (IFPRI), No. 11, 2004.

Digital Inclusion - The Vision, the Challenges and the Way Forward

Leela Damodaran, Teresa Gilbertson, Wendy Olphert
Loughborough University
Loughborough, UK
L.Damodaran@lboro.ac.uk, t.gilbertson@dtsi.org.uk,
C.W.Olphert@lboro.ac.uk

Jatinder Sandhu
Nottingham Trent University
Nottingham, UK
jatinder.sandhu@ntu.ac.uk

Mary Craig
University of Edinburgh
Edinburgh, UK
m.craig@ed.ac.uk

***Abstract*—This paper considers the vision and aspiration of digital inclusion, and then examines the current reality. It looks beyond the rhetoric to provide an analysis of the status quo, a consideration of some facilitators and challenges to progress and some suggestions for moving forward with renewed energy and commitment. The far-reaching benefits of digital inclusion and the crucial role it plays in enabling full participation in our digital society are considered. At the heart of the vision of universal digital inclusion is the deceptively simple goal to ensure that everyone is able to access and experience the wide-ranging benefits and transformational opportunities and impacts it offers. The reality is a long way from the vision: inequality of access still exists despite many national campaigns and initiatives to reduce it. The benefits and beneficiaries of a digital society are not just the individual but all stakeholders in the wider society** [1]**. Research evidence has shown that the critical success factors for successful digital participation are (i) appropriate design and (ii) readily available and on-going ICT (Information and Communication Technology) support in the community. Challenges and proven solutions are presented. The proposition of community hubs in local venues to provide user-centred ICT support and learning for older and disabled people is presented. While the challenges to achieve digital inclusion are very considerable, the knowledge of how to achieve it and the technologies which enable it already exist. Harnessing of political will is necessary to make digital inclusion a reality rather than a vision. With the cooperation and commitment of all stakeholders actualisation of the vision of a digitally inclusive society, while challenging, can be achieved and will yield opportunities and rewards that eclipse the cost of implementation.**

Keywords-Digital society; digital inclusion; accessibility; participation.

I. DIGITAL INCLUSION: FRAMING THE VISION

The raison d'être of this paper is to provide an evidence-based roadmap to achieve the vision of digital inclusion embedded in the 'Declaration of Principles' presented at the World Summit on the Information Society, Geneva, 10-12 December 2003. This vision of digital inclusion is that everyone in the world will have equal access to the knowledge and information they require to enable them to live their lives to their full potential. The inspirational and aspirational [2] declaration states the following:

"We, **the representatives of the peoples of the world** …declare our common desire and commitment to build a people-centred, inclusive and development-oriented Information Society, where everyone can create, access, utilise and share information and knowledge, enabling individuals, community and peoples to achieve their full potential in promoting their sustainable development and improving their quality of life…."

While the Declaration sets crucially important and ambitious goals for our Information Society, relating to equality of opportunity in the digital world, clearly the goals will not be realised without equally ambitious development and implementation strategies to achieve an inclusive digital society which has been defined as one in which all members of a community are **able** to **access**, **use**, and **understand** digital technologies [3]. Building on the Declaration of Principles which underpins the vision of universal digital inclusion, the aspirations and principles of such a society would be characterised by global access to ICT (Information and Communications Technology) to make a significant contribution to improving health, wellbeing and quality of life for all [4][5].

This paper provides a clear basis for plans and strategies to progress towards the vision of an inclusive digital society.

A. Approach and methodology

This paper first considers the vision and aspiration of digital inclusion and its major and fast-growing importance, informed by published reports and case study examples of the transformative power of digital technologies. Section II gives an analysis of the status quo regarding the current reality of digital inclusion based upon an extensive array of published studies and supplemented by empirical data collected on the RCUK (Research Council UK) funded NDA (New Dynamics of Ageing) Programme, through the Sus-IT (Sustaining IT use by older people to promote

autonomy and independence) project. From this detailed examination, some of the key challenges to digital inclusion are identified and examined in Section III. Section IV then presents evidence-based recommendations on how to meet these challenges in order to close the gap between the vision and the reality. Finally, in Section V, new solutions are brought together with established good practice in an integrated strategy for achieving digital inclusion. This offers a holistic sociotechnical approach, and includes recommendations for design, for meeting ICT support needs in the community and for addressing ethical concerns.

The methodology applied included literature reviews, empirical research studies using a wide array of mixed research methods which includes surveys, interviews, focus groups and case studies. In addition to reporting well-established barriers to inclusion, the paper provides the basis for significant innovation by providing new insight into a key and fundamentally important critical success factor for effective and sustained digital participation, namely the provision of community-based ICT support.

In summary, the paper looks beyond the rhetoric on digital inclusion to provide an analysis of the status quo, a consideration of some important facilitators and inhibitors to progress and some carefully considered evidence-based suggestions for moving forward with renewed energy and commitment.

B. Why is Digital Inclusion Important?

Digital literacy is fast becoming a fundamental requirement for full participation in society. In our emerging e-society, ICTs are an important aspect of daily life. Primarily through the medium of the Internet, an ever-expanding range of information, goods, services, entertainment/leisure, educational and social networking opportunities are available. Digital technologies make possible transformations which enhance quality of life for individuals, increase life chances, prolong independence and autonomy and improve social connectedness [6][7]. Such transformations offer great potential to improve society and boost the economy [5][8].

Digital inclusion is essential for the realisation of many health benefits of technology. These are already being leveraged in the management of chronic conditions in ways that require a level of digital literacy. For example, for management of diabetes requiring insulin injections, people with this condition can already use insulin pumps, which slowly inject insulin into the body. While the pumps are mechanical, they utilise device management software [9]. Although there are other ways of managing insulin-dependent diabetes, for an individual to have the fullest range of choices, digital literacy is becoming a requirement. Personal assistants and reminder systems may also play a role in helping people maintain their independence. Indeed, many of the technologies under development that will enable people to live independently for longer have digital literacy implications for the person living independently, or for their family or their care assistants. In combination these technological capabilities offer the potential to lessen demands on other (both formal and informal) support systems for older people; improving the quality of life of older and disabled people while also reducing costs of care. By enabling rehabilitation, remote assessment, diagnosis and treatment delivery, the need for residential care could be delayed, or for some, avoided by supporting policies of early intervention.

Digital inclusion is also a prerequisite for capitalising on the social opportunities afforded by connectivity which can help to reduce social isolation and loneliness. Increasing digital participation allows people to stay connected with friends and family, their local community and the wider world in a variety of ways. For people who are geographically removed from their family or friends, services such as Skype are not merely cost-effective and convenient; they can become a component in maintaining social connectivity and cohesion in a way that telephones cannot. An example showing the power of the Internet to reduce social isolation is the experience reported by 80 year old Lucy: "With my computer I was there in the house with them when he opened his Christmas presents, my grandson came and kissed the computer screen. It stops you feeling alone" [4].

The role of digital inclusion in enabling social contact is highly significant and important in helping to maintain good mental health by reducing the depression, stress and anxiety associated with social isolation and loneliness. Numerous published articles by governments, academics, practitioners and others seek to show that being part of the digital world can improve life in a number of ways, means and forms [2][3][10]. The emphasis of such publications is often on the financial savings for the state and for individuals. While monetary gains are important, it is the transformations for individuals that are vastly more far-reaching and fundamentally important for our digital society and economy. This transformational capability of digital technologies is perhaps best illustrated by case studies which serve to convey the scale and scope of the power and significance of digital connection. For example, one case study reported on the website of Digital Unite describes the experiences of a woman who had lost her sight six years previously and now had to rely on her husband to read the post, write letters and other tasks that required reading such as checking recipes. When she discovered the magnification tool on her computer as well as text to speech software, her life was transformed. Through having access to the software and hardware of sophisticated assistive technologies (ATs), which could magnify the screen, read aloud both website content and scanned letters, she was able to regain her independence [11]. Many examples (e.g., from Leicestershire CareOnLine) show transformations of comparable magnitude and significance [12]. Their importance for the individuals concerned and for their carers and wider family cannot be over-stated.

It will be evident that the importance of digital inclusion extends way beyond an individual's social group. Digital technologies enable greater civic engagement, participation and influence in wider society. For example, through online voting, access to information about what is going on in your community – even for those who are housebound and opportunities for people (who may otherwise be heard) to have a voice on the issues that affect them. This not only connects them to their local community but also increases their visibility to those who are already digitally connected. Enabling social networks thus create digital opportunities for having fun and increasing and maintaining social contacts. Similarly, social media is fast becoming a means of disseminating local news. In the UK, for example, The West Bridgford Wire Facebook page [13], Loughborough Echo Facebook page [14], police alerts and appeals (e.g., Nottinghamshire Police Facebook page [15]) give information and warnings that affect people on a local level. It is also the case that social media can be an alternative and very visible means of influencing corporate behavior by publicising poor service, mistakes made and the consequences experienced (e.g., the viral Facebook post of a Virgin media bill posted by a man whose dead father-in-law had been charged a £10 late charge to his account, despite the company having been notified of his death) [16]. The account was subsequently closed. The post has been shared on Facebook a total of 97,677 times. This suggests that public Facebook/Twitter exposure is now mightier than the pen.

It will be evident from the extensive ramifications of digital connection considered here that the advantages of being part of the digital world are vast and increasing all the time – as are the 'dis-benefits' of being digitally excluded. The authors have sought to show that beyond the increased ease of use, enhanced communication, reduced costs and so on enabled by the internet and use of ICTs in everyday life, it is their transformational capability that makes the quest for universal digital inclusion truly important and significant.

II. DIGITAL INCLUSION: THE REALITY

A. Global Internet Usage

Statistics show that access to the Internet is currently unevenly distributed amongst the global population, and older people in most countries are less likely to be Internet users than younger people [17][18]. ITU World Telecommunications predicted that in 2010 approximately 70% of the world's population (6.9 billion) were not using the Internet [19]. In 2014, the figure for non-users had dropped to 60% [19]. World-wide, large numbers of older adults are reported to be non-users of technology [20]. In 2011, approximately 45% of the world's Internet users were below the age of 25 [19]. Statistics from 2013 show that while 99% of people in the UK aged between 16 and 34 were Internet users, this drops to 88% for people aged between 55-64, 71% of people aged between 65-74 and a very low 37% for people aged over 74 years [21].

B. The Digital Divide

The gap between those who do and who do not enjoy the benefits of access to the Internet has been termed the 'digital divide'. The term "digital divide" was adopted by the Clinton/Gore administration in the US in the late 1990s. Many studies have sought to explore the factors that underlie this phenomenon. These have shown that the digital divide is not a simple binary division between the 'haves' and 'have-nots' [10] but that digital divides arise from three main sources of inequality which occur both between nations and within them. These are differences in (i) *connectivity* (i.e., access to appropriate infrastructure, hardware, software, and services including fast and reliable Internet access), (ii) *capability* (i.e., appropriate education, tailoring for ability/disability, digital literacy and skills) and (iii) *content* (i.e., the availability of accessible, meaningful, relevant material which, combined with the 'pull' of compelling functionality provides powerful motivation to become digitally 'engaged' [5].

The digital divide persists despite the attention it has attracted in society from academics, politicians, social activists and many others for almost two decades. The causes and factors associated with the phenomenon have been considered extensively and it appears that multiple factors are involved and these are considered in the section below. The large digital divide between older and younger people is explained partly because older people may not have the financial means to pay for equipment and services, partly because they may not have acquired the necessary skills either through education or in the workplace and partly because they may not have any motivation or interest to use these new technologies. Chronological age however, is clearly not a factor in itself, since many older people do use, and enjoy using, computers and the Internet [22]. It has been suggested that this so-called 'grey digital divide' [23] will eventually close, as people who are established computer users move into retirement, or as a consequence of the many initiatives and programmes designed specifically to encourage and teach older people to use computers and the Internet, or perhaps as a consequence of improvements in the ease of use and accessibility of the technology. However, there are counter arguments: in 1998 President Clinton's commencement speech at the Massachusetts Institute of Technology asserted that: "...the digital divide has begun to narrow, but it will not disappear of its own accord. History teaches us that even as new technologies create growth and new opportunity, they can also heighten economic inequalities and sharpen social divisions" [24]. Further, recent research suggests that a fourth digital divide exists, namely that caused by digital disengagement [25].

The salient point is that, irrespective of age or any other defining characteristic, the reality for those who are, or who become, digitally disengaged, is a range of adverse social

and economic consequences associated with reduced access to on-line government and commercial services, to health information and social support as well as to researching and purchasing goods using such services as price comparison and review websites.

C. Factors Associated with the Digital Divide

In addition to the digital divide which exists between the old and the young, there are other digital divides based upon education [26], gender [27], social class [28], ethnicity [6] and disability [29].

Access to technology and the acquisition of the basic skills to use it were initially seen as the key to bridging the digital divide. However, as access to computers has increased across all industrialised societies, it has become clear that the digital divide is not just about access and the acquisition of basic skills and knowledge, but also about the ability and the confidence of individuals to utilise digital technologies effectively. Numerous studies have identified barriers to uptake of the Internet, and this has led to the proposition that non-users of the Internet can be transformed into users (the 'digitally engaged') by an additive model which addresses each of the three barriers – connectivity, content and capability). Consequently, governments and other bodies in many countries are investing significant resources into providing technical infrastructure, awareness and training initiatives, and the development of digital content and digitally-delivered services, with the aim of increasing access to the Internet and promoting digital engagement. Despite many widely publicised initiatives and programmes, the number of people not using the Internet has reduced only slightly, and the overall figure remains fairly stable (73% in 2009 and 72% in 2015 of people sampled by the Office for National Statistics (ONS) reported using the Internet daily (or almost every day for 2015) [30][31].

D. Beneficiaries

The beneficiaries of a digitally inclusive society are numerous and include key stakeholders in the following categories: government – national and local, service providers, retailers of on-line services and products, designers and developers of ICT products and services, AT providers, voluntary organisations and social movements, individuals and society.

1) National and Local Government

National governments benefit from widespread digital inclusion as a result of an increasing proportion of the populations being competent and confident to access centralised government services online. By reducing the burden on traditional services in this way, national and local governments are direct beneficiaries of a digitally-engaged populace with access to the health and wellbeing benefits enabled by technology. The ability for all members of society, especially older people, who are the heaviest users of many services, to utilise these online, would allow for cost-savings arising from greater individual self-care, care delivery and societal participation. Local government could expect see a lessening on the burden on some local services as a result of the benefits of digital inclusion that enable independent living, greater well-being and the reduction of social isolation.

2) Service Providers

In the case of the public sector, the increased levels of confidence and capability of older people in digital participation should increase the uptake of on-line services (e.g., Universal Credit in the UK) and this will be of great value to local government bodies attempting to 'achieve more with less'. Whether delivered by a nationalised or privatised service or a mix of the two, service providers stand to benefit from the increased inclusion enabled through use of accessible technologies. As an ageing society, chronic diseases and their management is increasingly important and needs to be cost-effective to maintain pace with the growing proportion of people living into old age. Health care is likely to see mobile technology becoming increasingly used as a management tool for such conditions [32]. In terms of proportion of spending, globally, the treatment of chronic diseases currently counts for 60% of healthcare spending [32]. Healthcare providers have an opportunity to reduce their costs and ameliorate the effects of the growing ratio of older people to members of the workforce in society by ensuring that they make use of the wealth of information available about how to make their current and emerging products more inclusive. It will become increasingly important that new technologies are accessible to the non-technologist, whether that is the community nurse, or the carer or the person with the health condition.

3) Retailers of On-line Products and Services

A major market exists in the 50+ age group which offers significant, largely untapped, commercial opportunities to many retailers and other commercial companies. Understanding and being able to 'segment' the older market appropriately to achieve better tailoring of products and services offers competitive advantage to business. For businesses such as banks and retailers that provide online services, participation in community schemes to provide ICT help and support would provide a venue for retailers and service providers to demonstrate their online services (without applying sales pressure), provide information and support, build capacity in the older population, encourage customer loyalty and develop customer engagement in an environment that is comfortable to their potential customers [33]. The benefit of customer engagement could also extend to online stores which allow people with mobility or transport issues, to comparison shop and buy online with greater ease. By participating in community venues, retailers have the opportunity to work in partnership with customers, allowing them to identify products that best meet the needs of their customers, highlight where changes would enhance the user experience as well as provide them with testing

opportunities for new products and services with a demographic of people with whom they might not otherwise engage. This reciprocal relationship would not only confer benefits to both parties, but would create the environment for good products to become great products that are inclusive as well as appealing to the mainstream.

4) Designers and Developers of ICT Products and Services

Meeting the design needs identified represents a commercial opportunity for ICT designers as well as offering an exciting intellectual challenge. Instead of designing for a typical user, (i.e., someone like the designer; young and, typically, male) designing for inclusion will take into account the variability of user characteristics, as well as the changes to individual capabilities over both the short and long term [34]. Such inclusive techniques not only benefit novice users, as well as people with disabilities, but also experienced users [35]. Indeed, the argument has been made that barriers that affect novice users such as indecipherable error messages, unexpected crashes, confusing menus, and site disorganisation, also present problems to experienced users [35]. By designing for inclusion, more people can use the service and the whole user-base benefits from a better product.

5) Assistive Technology Providers

Many ATs for people with recognised disabilities can help everyone in demanding or extreme usage situations. Innovative application of accessibility principles offers the promise of increased profitability. Successful examples of ATs that have wider appeal include such applications as text to speech; shorthand for text messaging; image stabilisation and closed captions in video games. Such accessibility technologies open up new market opportunities in every sphere of life, including healthcare, homecare, commerce, education and recreation.

6) Voluntary Organisations and Social Movements

The internet has also provided a much needed voice for people to take part in social movements and voluntary campaigns. Web-based activism organisations such as Avaaz [36], 38 Degrees [37], SumOfUs [38] have not only given people a voice with online petitions, but have also provided a platform for individuals to stage their own petitions. 38 Degrees also allows for grass-roots interactions through regular polling of members about the direction of campaigns as well as polling members on preferred campaigning strategies. In order for these organisations to allow all people to exercise their democratic rights, digitally inclusive strategies to give each citizen a voice in online activism are imperative.

7) Individuals

Individuals will benefit from inclusion in numerous ways elaborated in earlier parts of the paper. In addition to the generic advantages of digital inclusion described above, there are specialised developments which will be of particular benefit to older and disabled users. For example, creating adaptable interfaces could be particularly important for older users who want to use what everyone else is using rather an AT, either because of the steep learning curve involved in learning to use some ATs, or because of a personal preference to use what everyone else is using [39]. The experience of built-in personal customisation is of benefit to all. Similarly, adaptivity features that aid in automatic customisations can reduce or eliminate the learning involved in making such changes, thereby removing a number of access barriers. By ensuring stable, intuitive, usable and adaptive design, individuals will benefit in terms of not only accessing goods and services but also engaging in personal pursuits and living independently.

8) Society

The breadth of individual economic, health and wellbeing and social benefits combined with the reach of benefits across the private and public sectors also benefits society as a whole. Moreover, the interaction between all of these advantages creates a synergy such that the total benefit to society is potentially far greater than the sum of the individual benefits. A genuinely digitally inclusive society offers transformations which range, for example, from the empowerment that results from an individual who learns a specific skill on YouTube being able to solve a particular problem/meet a need, to becoming economically active through online courses and social support, to being enabled to participate fully in civil society.

III. Some Key Challenges to Digital Inclusion

The challenges to achieving digital inclusion are extensive and include: getting online; staying online; design barriers; inadequacy of support; a culture of disinterest; and emerging ethical considerations. Each of these is described below.

A. Getting Online

There are significant challenges involved in getting online – especially those faced by many older and disabled people. In one study which examined the barriers to learning to use ICTs, the older people who participated reported that the process of learning to use a computer required significant effort, time and patience and demanded considerable help and support [40]. Novice users described feeling ‘not in control’ and sometimes feeling overwhelmed by the complexity of learning to use a computer [40]. For example the processes required to turn a computer off, or to connect to the Wi-Fi network, or to upload software to connect a printer to the computer were not felt to be intuitive and were found to be ‘too confusing’ and complicated. The struggle of learning to use a computer was exacerbated by sudden, unexpected and unwanted ‘pop-ups’ which were described as, ‘confusing’, ‘annoying’ and ‘distracting’ [40]. Fear of using a computer was also a barrier, not only for novice users but for established users who feared doing something wrong or ‘breaking’ the device. For instance, learners were left uncertain regarding what

remedial action to take in response to error messages and in some cases the distressing experience deterred them from continuing their efforts to gain digital literacy skills [40]. Additionally, getting online was found to be a challenging process since learning to use ICTs is not a one-off exercise but an on-going process necessary in order to cope with the complexity and rapidly changing nature of technology. For instance, updates to operating systems meant that learners had to re-learn how to do tasks they had previously comfortably done at ease [40]. The cost of training/learning support sessions/classes is a further challenge to getting online, especially for novice users who often require sustained support over a period of time to consolidate their learning, accommodate the implications of their disabilities and memory issues and to develop confidence in operating in a digital world which is often unfamiliar to them. Another challenge experienced by novice ICT users included making the transition from using devices in the training environment to those they had in their own homes [40]. Their comments included: "the buttons are in different places" and "I can't find things that were obvious before". The lack of appropriate and readily available ICT learning support (discussed further below in D) was also found to be a major challenge associated with getting online [40][41].

B. Staying Online

Once the initial learning challenges have been overcome sufficiently to permit the novice user to participate in the digital world, further challenges arise in the face of capability and other changes. It is well established that older adults are vulnerable to capability changes associated with cognitive and physical changes in later life. Changes in psychological and cognitive aspects will impact on the capability to remember sequential processes and on confidence levels in using ICTs. Changes in physical aspects, such as changes in vision make it difficult to see what is on the screen, and dexterity issues will create problems for controlling the mouse. Social changes, such as family members moving away, also impact on the support available to older ICT users. Further, older adults encounter some or a range of the following barriers in learning to use and sustain use of ICTs: confidence and fear of using ICTs; problems with understanding technical jargon and dealing with pop-ups and spam; problems with updates, drivers, and software; dealing with the rate of change of technology and coping with poorly designed software and hardware [41].

In recognition of the issues facing older ICT users, which may lead to people giving up use of computers and other digital technologies, the 2004 UK Digital Inclusion Panel Report [5] stated that "there is a real risk that in the medium to long term, significantly more citizens will migrate from being digitally engaged to being unengaged than the other way round, as their capabilities change." It is a common myth that 'once people are online, they stay online'. There is evidence to show that some people who have used the Internet at some point, and for some period of time, have subsequently stopped doing so [5][17][29]. This phenomenon is a potential, but largely unrecognised, 'fourth digital divide' [25], i.e., it cannot be explained by a simple interpretation of lack of access, lack of skills or lack of interest or motivation, because the people in this category have formerly been users. One in ten people are reported to have given up on using computers and it has been found that the older generation are more likely to be the ones that 'give up' [26].

C. Design Barriers

The design of ICTs continues to pose many problems for older and disabled people – particularly relating to the speed of change and unnecessary complexity of software and products. Despite the existence of inclusive design concepts and general usability guidelines and heuristics for ICT developed since the 1980s [42], design barriers still present barriers to digital inclusion. Although there has been extensive work on accessibility, it is still the case that users who most need operating system-based accessibility tools are often unaware of their existence [43], arguably this is a design failing rather than a failure of the AT. Typically accessibility features enable the user to adjust for decline in eyesight and dexterity enabling them to change font characteristics, contrast settings, text colour and size, mouse speed etc. Additionally, even when ATs such as screen readers are selected and purchased by a user, such technology must then be checked for compatibility with the computer as well as for any unintended side-effects [43]. For example, magnification software will create the need for increased use of the mouse or touchscreen, and the need for more scrolling is likely to pose problems for a user with motor difficulties [44]. Compatibility issues can have implications for usability, for example, not all AT software will work with all operating systems, and sometimes the user will be expected to install updates and patches in order to make such software run. For people with basic digital literacy, it's the ICT equivalent of asking them to service their car, i.e., beyond the capabilities of many people.

There is also a lack of ATs that help people with minor to moderate decline in working memory. This lack is a significant barrier to the ageing population [45]. Technical jargon and generic manuals are also barriers to the successful use of ICTs in the ageing population [45]. The paper also identified a number of additional barriers that are exacerbated by age, which include:

a. keyboard characteristics (e.g., "too cramped" to accommodate motor function decline caused by osteo-arthritis and other age-related conditions)
b. software updates (difficulty understanding license agreements and their implications)
c. "invisibility" of accessibility features [45].

It is clear that no one solution will solve all problems. For some older people, direct interaction devices such as tablets and smartphones may indeed provide an easier means of interaction than a laptop, however, others may

have problems using such direct interfaces due to parallax and the crudity of the finger as an input device [46]. Additionally, while direct interaction may work for some, the practicalities of access will need to be addressed to ensure that this technology is available to all who wish it. For example, while smartphone and tablet use is increasing among the over 65s in the UK, accessing the internet away from home via portable devices, i.e., use of laptops or tablets is still only 22% and an even lower 16% for smartphones [31]. When compared to 54% and 87% of 35-44 year olds who use laptops/tablets and smartphones respectively to access the internet away from home [31], it is clear that this disparity of use and the reasons for such disparity need to be addressed as a part of the solution. Smart TVs with voice control and other input modalities may also serve a useful role in keeping some older people online. However, this may not be appropriate for all users for any number of access reasons including the fact that speaking can be very tiring for some older people; and personal choice and budget are relevant deterrents for others. Smart TV ownership in the UK is 29% as of January 2015 [47], making it a potential tool in the digital inclusion toolbox, but not a panacea, even in consumer-driven developed nations.

In 2001, Shneiderman argued that designing for diversity not only benefits all people by promoting quality but also that usability and design are key to success [35]. Crucially, Shneiderman points out that designers, ***"rarely see the pain they inflict on novice and even expert users"*** [35]. Unfortunately, 14 years later, this is often still the case.

D. Inadequacy of Support

Findings from a survey of older ICT users conducted as part of the Sus-IT project, make clear the importance of ICT support with 56% of older people saying they regarded support as the most important factor in sustaining their digital participation. Respondents reported heavy reliance on support from family members or friends both to learn and to solve problems. Approximately one quarter of respondents said that human support and encouragement was the most important thing to help them use technology successfully [45]. Yet results from another study on the Sus-IT project showed that the opportunities for ICT learning and on-going support are extremely inadequate [41].

The poor ICT support for older users in the community contrasts sharply with that available in the work place. For ICT users in the workforce, the majority of workplaces will have a dedicated member of staff tasked with ICT support i.e., setting up and maintaining the infrastructure, selecting which technology platform is used, installing updates etc. The ICTs are installed and maintained for all those in the workplace although most are likely to be the able bodied, and reasonably technologically au-fait members of society. What exists beyond the workplace for the many without such institutional support is very different. In contrast to the situation in many workplaces, many ICT users who are at home, unemployed, retired, living with disabilities, living on a reduced income, living with a reduced social circle and possibly living with reduced health find themselves having to cope with all these demands themselves.

Organisations such as charities for older people and public libraries do offer limited help, but typically this is piecemeal and variable, and as such, is perceived by older users as inadequate and unsatisfactory [41]**.** While instruction in digital skills was offered typically through structured classes, many older participants expressed their discomfort with formal classes and some articulated their negative associations with their school days [40]. The majority of ICT courses and campaigns offered in the UK to promote IT use focus on the acquisition of basic skills to go online rather than on promoting confident participation in the digital world and on sustained usage. The consequence of this narrow provision is that once initial training is over, older people can feel alone, anxious and frustrated when experiencing problems with on-going ICT use [40]. Ongoing support for many outside of work currently means having to rely on whatever support they can find from friends and family or in the community and, typically, experiencing the frustration associated with using these forms of support is the norm for those outside of work.

The poor quality of available support is also a challenge to sustaining digital inclusion. It is a common complaint that when support is offered from children, they are too fast or resolve the issue for the user without explanation [40]. Further, studies have identified teaching styles that tend to undermine learners' attempts to learn in formal settings. Some tutors fail to project their voice adequately [48] while some proceed too quickly. The speed at which people learn and process information tends to decline with age [49]; consequently, mixed age classes may fail to meet the needs of all. A fast pace of learning may leave some learners behind while a slow pace may cause more able learners to lose interest [50]. Furthermore, the use of jargon often confounds ICT novices [48].

Another gap in adequate support is dealing with anxiety caused by the threat of cyber-crime and bullying. While is is possible to counter some of the threats of viruses and cyber-crime through technical solutions such as firewalls and anti-virus software, the need to address the real fear that can arise as a result of the possibility of cyber-bullying and cyber-crime (e.g., phishing, viruses etc.) requires actual human support to ensure confident use of ICTs.

E. Culture of Disinterest

The culture in many commercial companies that develop ICT software and products fosters the belief that ATs do not create a significant return on investment [43] that the related lack of interest in the hidden nature and, in some cases, paucity of accessibility features often remains unchanged and unchallenged. It has been observed that developers experience difficulty convincing both clients and their management of the importance of accessibility [51]. While

many web developers view accessibility as important and consider it at the early stages of projects, there are a considerable number of people in project management roles who do not see ageing as an accessibility issue [52]. This lack of interest and awareness of age-related capability change impacting on the use of technology does not create a climate conducive to the development and visibility of accessibility features to assist those with declining working memory, dexterity, contrast vision and other age-related changes. The prevailing culture in many countries is often one in which compliance with established standards on accessibility is resisted. For example, a study of the webmaster attitudes to the implementation of Section 508, (Amendment to the Rehabilitation Act 1973 in the United States that mandates that government electronic content be accessible) on American e-government sites found that most sites did not conform to Section 508 [53]. If websites that have a legislated requirement to conform to published standards fail to meet this requirement, it is clear that simple legislation is not a cure-all for accessibility and inclusion and that a culture change is needed. Since research has shown managerial attitude to Section 508 compliance to be very important in determining whether attempts to make sites accessible were actually made [53], this finding is concerning and attitude change at all levels is urgently needed. It is clearly very difficult to persuade organisations of the importance of accessibility tools if there is widespread disregard of existing standards. The challenge of getting ICT professionals, particularly at management levels, to embrace digital inclusion (and usable and discoverable accessibility features) is a continuing barrier, despite the passion and commitment of many designers and developers to the development of inclusive products and services.

F. Emerging Ethical Dilemmas

There are very many ethical issues associated with different aspects of digital inclusion, e.g., as posited by Eccles et al. [54], and it is clearly beyond the scope of this paper to deal with these comprehensively. The focus of this section is on the exponential growth of automated collection of data about users, and its potential use and abuse as another threat to digital inclusion. Of particular concern are the ethical issues associated with the modern web and emerging smart technologies such as digital assistants and smart televisions. These technologies rely on information about the user to make suggestions and process interactions which pose risks to the privacy of their users. Information about the user is an important part of the functionality of the web. Browser-based tracking enabled by the use of cookies can be benign, for example, allowing a user to logon to a website for e-commerce, social media or similar purposes. However, third-party tracking cookies, which can be used to track people and to monitor their web behaviour in to order to tailor in-page advertising to them and to draw a picture of their browsing habits, are also prevalent, but not crucial to the workings of a webpage. For an individual struggling with only rudimentary capabilities in digital literacy, it is very difficult to distinguish between the benign and the invasive. While it is now a requirement in the EU for websites to give people the option to reject the saving of cookies on their browser in a clear and concise manner, lack of user understanding of cookies may prevent informed consent. When the consultation in the UK regarding changes to the law was held, it was found that 37% of people consulted did not know how to manage cookies [55]. This number is likely to be substantially higher in the general population due to the disproportionate number of "Internet-savvy" users taking part in the consultation [55].

The data collection that takes place at the level of the operating system also raises related ethical questions regarding privacy. Do people understand the privacy implications of certain operating system features? For example, the express installation settings for Windows 10 (i.e., when a user chooses to use the default settings for installation that only require clicking the "next" button rather than choosing to go through each setting before they are installed), allows for the collection of personal data including location, calendar and contact list details in order to tailor advertising and other features by default. For those with just basic digital literacy (and 21% of people in the UK do not have even basic online skills [56]) it is clearly not appropriate to assume understanding about what information their IT devices are collecting, why such information is needed, who sees this information and how they can control it. Important ethical questions about automated data collection are clearly a major challenge to digital inclusion. Possible safeguards must be considered, e.g., should default installations expressly state which data is collected and under what circumstances this happens? Should there be a master opt-out for all data collection?

Data collection is required for personalisation features for adaptive technologies. For example, Cortana, Microsoft's personal assistant software, (which can set and manage your calendar and even remind you of errands based on your location), relies on data collection to work. Taking into account the assistive capabilities of personal assistant technologies such as Microsoft’s Cortana and Apple's Siri as well as the emerging capabilities of some smart televisions, which can allow you to control your television using your voice for people with disabilities, it is vital that both the assistive capabilities of such technology as well as the privacy implications of using such technology are known and transparent. As the situation currently stands, people who use these services for the accessibility capabilities they provide are faced with a very serious dilemma about how they manage their own data. Assuming that people using these products for their assistive features are aware of the privacy implications of collection of their data, such people may not have a choice in foregoing the technology for the sake of greater management of their privacy. This lack of choice between access and privacy for people who need the

assistive capabilities of a product is a considerable barrier to overcome.

Excessive complexity and confusing wording, as well as the volume of text in terms and conditions and end user license agreements (EULAs) pose major challenges for very many users – not only the old and disabled. In light of the complexity presented by such agreements, the prospective user is presented with another considerable dilemma, namely to use technology and risk their data being used in unknown ways, or to abandon use of the technology and maintain their privacy. Many older people opt for the latter [26]. One such example of the complexity of privacy policies is that of the Samsung smart television. Following the launch of Samsung's smart television with voice command features and the publication of Samsung's smart television's privacy policy for televisions with this feature, questions arose highlighting the concern about the implications of privacy for users of the voice recognition feature. Questions surrounded the privacy policy of the Samsung smart television was first highlighted by the Daily Beast [57], an online magazine, in 2015. The policy, as originally published, implies that users are being monitored by their televisions, "*Please be aware that if your spoken words include personal or other sensitive information, that information will be among the data captured and transmitted to a third party through your use of Voice Recognition.*" [58]. Wording like this fosters fear of 'Big Brother' in your television threatening your privacy and can leave people with a number of questions about just how safe it is to use their technology. For example, is it safe to have a conversation in front of your television? Who has access to this information? Can this information be hacked? Although in response to the press attention and concern of privacy experts, the Samsung privacy statement has since been reworded to better explain the process of converting the speech to text as well as how to disable this functionality [58], questions still remain. How comprehensible is the privacy policy to the average end user? Who actually reads the terms of service?

From the brief analysis of some of the issues presented in this section, it will be evident that the challenges to digital inclusion are immense. Ways of addressing these are considered in Section IV.

IV. HOW TO MEET THE CHALLENGES?

To meet the many significant challenges identified in the previous section associated with getting and staying online, design barriers, inadequate support, the prevailing culture of disinterest in inclusion and accessibility and emerging ethical dilemmas will demand a raft of integrated solutions. There is no 'silver bullet' to meet the complex array of challenges and many contributions will be needed, including, for example, from businesses; government (both national and local); the third sector; policy-makers; educationalists and research and design communities.

The solutions which evolve will need to be implemented by policy makers, manufacturers, retailers, service providers, designers, education providers and others. Relevant stakeholders will need to be identified, empowered and engaged to work together to influence policy makers, manufacturers and service providers to convince them of the vast significance and value of achieving the vision of digital inclusion. Crucial to success is the necessity for relevant stakeholders to work together to increase access and develop widespread capability in the population and also to make understood the benefits to individuals.

To narrow the many digital divides identified in Section II will require not only the necessary telecommunications infrastructure to enable connection, but also the equipment to utilise the technological capabilities and the human capability to capitalise on both of these components. Evidence from Sus-IT indicates that two key determinants of sustained digital engagement of older people – and therefore of their digital inclusion in the long term – are (i) the appropriateness of design and (ii) the adequacy of ICT support available to them.

A. How to meet the design challenges

To meet the design challenges identified in Section III requires processes which enable older people to participate in shaping the design of products; increased investment in 'elder friendly design'; and greater awareness and education in accessibility. Experience on the Sus-IT project offers lessons/good practice in engaging older people in shaping design decisions showing that co-design 'sandpits' can be used successfully to involve older people in the research and design of digital products rather than as passive research 'subjects'. The outcome of these sandpits included custom computers for older people, products to support memory and identity in later life, products to combat isolation and iPad apps designed specifically for older people [59]. Further good practice relevant to tailoring design to the needs of users experiencing capability change and often unaware of the accessibility features available, or with cognitive difficulties with adjusting font size, the development and testing was carried out of an adaptive software framework which explored how to how to connect existing operating system accessibility solutions to the people who needed them, at the time they needed them [60]. The framework developed monitored moment by moment interactions of the user with their ICT, balancing the pros and cons of any changes and suggesting improvements such as zooming text size and changing mouse speed in a manner that would aid their interactions as a way of meeting design challenges and potentially prevent disengagement with their ICT.

To promote greater awareness and education in accessibility a number of proven methods are available. For example, embedding accessibility education throughout the undergraduate level [61], encouraging students to evaluate their own projects using ATs [62] and exposure to issues pertaining to diversity and inclusion [61] all raise awareness

of the diverse needs of Internet users. Similarly, engagement with students and designers through interactive video such as the UTOPIA Project (Usable Technology for Older People: Inclusive and Appropriate) [63] proved successful in changing the attitudes towards older people's use of technologies of both undergraduates, and Human Computer Interaction (HCI) professionals [64]. Another established method for promoting awareness and changing attitudes is the use of interactive theatre: in an evaluation study of a pilot, over 85% of respondents agreed with the importance of highlighting accessibility issues and the use of the theatre format [65]. By ensuring that the diversity of needs is known to, and understood by, students and professionals, content designers and developers are encouraged to follow user-centred design principles in an inclusive and participative approach to the design of their products, systems and services which then becomes embedded in the culture. Interactive theatre also allows older people to contribute to the shaping of design.

Goals for designing to achieve universal inclusion were first presented by Shneiderman in 2000. These are:

- increased availability of low cost ICT with better quality of service
- the reduction of system complexity and user frustration
- ensuring older technology and slower networks can participate fully
- ensuring that services for low-income and poorly educated users are understandable and usable, while allowing expert users to use more novel strategies
- ensuring accessibility for people needing support for visual, auditory, physical and other disabilities
- enabling designs to work in multi-lingual and device diverse environments [35].

These points are as germane today as they were in 2000 in the reduction or elimination of design barriers.

With regards to ethical issues, there needs to be a clear and possibly legislated procedure on getting informed consent from people with regard to when their data could be collected, how it is to be used, who sees such data and how long it is stored across smart technologies. There is a very strong argument for the service provider to ensure that privacy data is understood and clearly stated at the time of first use of the software, with step by step configuration of each setting with explicit information on the consequences of agreeing or disagreeing to the collection of personal details. This also holds true for operating systems. Rather than configuring privacy at the time of installation, where the gathering of data is not linked to the operating system itself but to bundled software, such settings should be off by default and configured upon the first launch of the software in question.

There will clearly not be one generic solution to overcoming design barriers. Device availability, device capability, cost, interaction modality and personal choice will all play a role in solving interaction problems experienced by diverse individuals. Regardless of the device used, what is key is that the person using the device is able to carry out the tasks they wish to perform.

B. How to meet the challenge of inadequate support

The evidence described in Section III that document the challenges, makes clear that support has to address the needs of getting online, staying online and dealing with a vast array of problems that arise on an on-going basis. The research findings cited have shown that getting online is only one of the challenges. Once the initial excitement of connection has passed, older users' new found-confidence can be eroded very rapidly by the early challenges of coping with technology at home. We have seen that if older people are to become confident participants and contributors in the digital world, they need ongoing, friendly and reliable support that meets their needs. The challenge of inadequate support for ICT users has been identified and explained. The pressing requirement is to fill the ICT support void beyond the workplace. There are two components involved in this: (i) enabling non-users to become users and, (ii) supporting existing users in sustaining their learning and use of fast-changing technologies.

Informed by participative research with over a thousand older people across the UK, a co-design process developed a user-specification of the users' support needs. This had the following characteristics:

- readily available, trusted and sustained
- delivered in familiar, welcoming and local venues
- embedded in social activities / personal interests
- free of time pressure and assessments
- inclusive of problem solving / troubleshooting
- offering impartial advice and 'try before you buy'.

Informed by this user-led specification, a co-design process followed, with older people and other stakeholders to deliberate and develop solutions to the challenges of inadequate support [33]. Consensus emerged that a socially embedded model of provision was required. From an extensive consultation and validation process, a detailed proposition for provision of ICT support in the community evolved. In this model (aspects of which already exist in limited ways), socially-embedded community-hubs provide flexible and adaptable learning to match people's range of needs, learning speeds and styles. They use existing community venues and are locally run, providing ad-hoc trouble-shooting assistance as problems arise. The social nature of the hubs allows people to share tips and support each other, all within a relaxed, social atmosphere allowing for intergenerational exchange and access to professional expertise. The value of venues/hubs is in their capability to:

- Support and sustain on-going ICT use
- Empower and inspire people in use of ICTs
- Promote adventurous use of ICTs

- Make online participation enjoyable, rewarding and 'angst –free'.

There was also consensus that in order to be effective in the long-term, ICT hubs should:

- be widely available across the UK
- available on an on-going and consistent basis
- user-driven and locally run and make use of local venues
- have access to appropriate training resources and professional expertise
- be sustainable by securing multiple revenue streams – which will differ according to local circumstances and assets.

The proposition has a strong evidence base. It has been developed through a programme of roundtable discussions with multiple stakeholders, culminating in an intensive multi-stakeholder consultation.

The advantages that community-based ICT support hubs will provide users can also be expected to benefit the beneficiaries identified in Section II. For instance, the public sector could expect to achieve greater uptake of online government services, e.g., Universal Credit in the UK. For companies, hubs will give them access to older people who will be enthusiastic testers of new products and services as well as providing a venue to conduct research on user needs. Similarly online retailers can enjoy increased customer engagement and loyalty as well as the opportunity to demonstrate and allow people to try new products in an environment where people will not feel pressured.

Additionally, five single-sector roundtables were convened with the following stakeholder groups:

1. Businesses developing or delivering accessible solutions
2. Designers and developers of ICT products
3. Retailers of ICT products or assistive technologies for older people
4. Government departments, the Public sector and Local Government
5. Older Peoples' Groups, e.g. 50 + forums, Age UK and U3A.

The roundtables refined the proposals for community-based ICT support and these were subsequently the focus of a consultation entitled 'Falling off the bandwagon' at St. George's House (SGH). The proposition was considered and validated further over the 24 hour consultation with experts representing a variety of stakeholders at SGH, Windsor Castle, where the feasibility of implementing the proposition as proposed was extensively deliberated, including how it might be funded and managed [66].

The extensive process to find and develop a solution to address effectively the challenge of inadequate support culminated in the publication entitled 'Promoting Digital Participation. The proposition: Community hubs, Meeting older people's technology support needs; developing social communities and reducing isolation' under a creative commons license [67] requiring only acknowledgement of the source to use the material freely.

This proposition is now informing digital literacy plans of some local councils in the UK and the major joint initiative of Barclays and BT to introduce pop-up community hubs in public libraries and other local venues.

Therefore, community hubs in highly varied forms tailored to local circumstances and needs can be seen to offer a powerful generic way of addressing the challenge of inadequate support. The wide-ranging outcomes include:

- friendly informal help for beginners and advanced technology clinics for troubleshooting problems, e.g., computers, cameras, online shopping, smartphones, bill paying
- readily available, trusted advice in simple language
- opportunities to try before you buy, e.g., smartphones, tables, laptops etc.
- help with choosing software that meets individual needs
- hobby-based support, e.g., family tree, digital photography
- opportunities to swap knowledge and tips
- reducing social isolation and loneliness through social gatherings.

V. Discussion and Conclusion

In this paper we have presented the vision of universal connectivity and inclusion. The authors regard digital inclusion as fundamental to a flourishing democracy and to the full participation of people in society. At the individual level, digital inclusion is crucial to sustaining and enhancing independence and autonomy. However, the transformative potential of such inclusion transcends the individual and has wide-reaching benefits for the whole of society and the economy. To achieve the vision will require coordinated policies, strategies and practices led or endorsed by national governments and coordinated and implemented by local government, service providers, businesses and third sector organisations. Encapsulating the essence of the 2003 Declaration of Principles allows the vision of an inclusive society and economy to be articulated as "the enhancement of the quality of life for all, extending autonomy and independence through the use of digital technologies. This vision would be characterised by:

- Empowered people experiencing the benefits of digital inclusion
- Widespread participation in society and the economy
- Readily available support in the community for engaging with and managing all aspects of the "digital world".

To create the digitally inclusive society encapsulated in the vision described above requires structural, political, and

social change on a vast scale - which perhaps helps to explain the slow progress to date.

The process of achieving such change needs itself to be inclusive of all stakeholders across society. This means that to succeed, the co-creation of an inclusive society is required. This will require collaboration on a grand scale to address the challenges through innovating, creating and evolving a digitally inclusive society that harnesses the power of ICT for the benefit of all. The knowledge to meet the challenges already exists.

The barriers to achievement of a digitally inclusive society are well understood and, while commitment and effort to overcome them will be required, solutions are available and the return on investment in implementing these will be extensive. With leadership and commitment, sustained digital connectivity for everyone is achievable. In particular, leadership is needed to promulgate the vision and to encourage the development of strategic alliances and partnerships within a framework of appropriate policies and strategies, involving all relevant stakeholders – especially older people and disadvantaged groups in society. Engagement at grass-roots is urgently needed to complement the 'top-down' digital inclusion campaigns currently in operation in some countries.

Ethical implications with regard to data collection must be clearly stated and visible to software users, in order for uptake of digital technologies to have the potential to reach all members of society.

Meeting the wide range of challenges of digital inclusion requires development and implementation of a strategy with the following components:

- A shared and compelling vision of what digital inclusion offers to the economy, to society and to individuals and communities
- Policies and strategies for implementing steps towards the vision
- Strategic alliances, partnerships and collaborations
- Leadership

Public events and workshops offer a compelling way of promoting awareness and excitement of digital inclusion and the resultant digitally inclusive society. Implementing the following steps may stimulate awareness and mobilise widespread support for digital inclusion:

- Showcase a ***realisable vision of an inclusive digital society*** and economy to inspire investment by ***all*** stakeholders in building ICT capability and confidence of all.
- Demonstrate 'proof of concept' by ***modelling best practice*** in government policies and strategies, i.e., 'Do as we do' to promote digital inclusion.
- ***Quality of Life strategies*** in the community: creative holistic strategies for achieving technology-enabled autonomy and independence of all.
- Establish ***intergenerational problem-solving forums*** to engage diverse sections of the community to scope problems, exchange ideas and co-create solutions.
- Enable the above by ***use of innovative techniques and methods*** (e.g., drama/interactive theatre, 'sandpits', story-telling etc).

There will inevitably be different perspectives and priorities among individuals, communities, businesses and government regarding what a digitally inclusive society looks like. However, the transformational outcomes envisioned by the Declaration of Principles help to inspire a shared vision and perspectives and encourage commitment to follow the roadmap to achieving this. The sharing of perspectives between all stakeholders, especially those in positions of authority such as national and local government and those in business developing and selling ICTs are crucial to success in creating and sustaining a digitally inclusive society. Inter- and intra- stakeholder co-operation, ongoing negotiation for the mutual benefit of varied stakeholders and respecting the voices of less influential/authoritative stakeholders is also key to this journey.

To provide a roadmap towards the vision of an inclusive digital society, the following steps are essential:

- Promote widespread awareness of the benefits – individual, societal and economic – of digital inclusion
- Engage and gain the 'buy-in' of key stakeholders to the vision of digital inclusion
- Encourage and reward adoption of inclusive design principles and promote them as the industry 'norm' for designers, developers and manufacturers of ICT systems, services and products
- Create expectations of and demand for inclusive design/digital inclusion amongst buyers and users of ICT
- Promote awareness that current ICT learning and support provision in the community is variable in quality and availability and not tailored to the requirements of users
- Recognise that sustaining people online is an even greater challenge than getting them online in the short term – and invest in community provision indicated above.
- Create a framework for ubiquitous provision of ICT support in the community, e.g., in public libraries
- Document and co-ordinate the various local initiatives that exist, e.g., some general medical practitioners (GPs) are now "social prescribing" (e.g., recommending patients make an appointment at a local library to obtain digital skills training).
- Utilise local resources to meet local needs
- Harness the political will to push this vision

- Recognise and celebrate what we can all achieve given the right access to ICTs.
- Promote the adoption of socio-technical systems as the industry norm

These steps offer a roadmap to a digitally inclusive society.

To summarise:

- We have the vision of universal connectivity and inclusion
- The potential economic and social rewards are vast
- The challenges are immense – *but we have the know-how to meet them*
- The opportunities for innovation and change in business and in society are even greater than the challenges
- With leadership and commitment we can tackle social inclusion through digital inclusion.

In conclusion, the path to achieving the Vision of Digital Inclusion is well-developed and waiting to be travelled. The rewards are significant and achieving the vision of digital inclusion presents opportunities for innovation and change in business and society that are even greater than the challenges.

ACKNOWLEDGMENT

Sus-IT ('Sustaining IT use by older people to promote autonomy and independence') is a New Dynamics of Ageing project, funded by the 5 UK Research Councils – AHRC, BBSRC, EPSRC, ESRC and MRC (grant No. RES-353-25-0008). Funding from the EPSRC for KT-EQUAL (Grant Number EP/G030898/2).

REFERENCES

[1] L. Damodaran, C. W. Olphert, T. Gilbertson, J. Sandhu, J. and M. Craig, "Digital Inclusion - The Vision and the Reality," in *ICDS 2015. The Ninth International Conference on Digital Society*, pp.22 -28. 2015.

[2] WSIS, "Declaration of Principles: Building the Information Society: a global challenge in the new Millennium," in *World Summit on the Information Society*, 2003.

[3] Information Policy and Access Centre, "Public Libraries and Digital Inclusion. Digital Inclusion Survey 2013," 2013.

[4] Race Online 2012, "Manifesto for a Networked Nation," 2010.

[5] UK Cabinet Office, "Enabling a digitally United Kingdom," 2004.

[6] K. Mossberger, C. Tolbert, and R. S. McNeal, "Digital Citizenship: The Internet, Society, and Participation," *J. Inf. Technol. Polit.*, vol. 5, no. 2, pp. 262–264, Aug. 2008.

[7] Department for Business Innovation and Skill, "Digital Britain," 2009.

[8] Business Innovation and Skills Committee, "Broadband Digital Exclusion," 2010.

[9] American Diabetes Association, "Insulin Pumps." [Online]. Available: http://main.diabetes.org/dforg/pdfs/2015/2015-cg-insulin-pumps.pdf. [Accessed: 17-Aug-2015].

[10] P. Norris, *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*, vol. 21. Cambridge University Press, 2001.

[11] Digital Unite, "Reclaim your independence by getting online!" 2013. [Online]. Available: http://digitalunite.com/blog/reclaim-your-independence-getting-online. [Accessed: 22-Dec-2014].

[12] Leicestershire CareOnLine, "Your Story." [Online]. Available: http://www.leicscareonline.org.uk/careonline_people. [Accessed: 26-Aug-2015].

[13] West Bridgford Wire Newspaper, "West Bridgford Wire Facebook page," *Facebook*. [Online]. Available: https://www.facebook.com/WestBridgfordWire?fref=ts. [Accessed: 17-Aug-2015].

[14] The Loughborough Echo Newspaper, "The Loughborough Echo Facebook page," *Facebook*. [Online]. Available: https://www.facebook.com/loughboroughecho. [Accessed: 17-Aug-2015].

[15] Nottinghamshire Police, "Notts Police Facebook page." [Online]. Available: https://www.facebook.com/nottspolice?fref=ts. [Accessed: 17-Aug-2015].

[16] J. Boyden, "Virgin Media Bill viral Facebook post," *Facebook*, 2013. [Online]. Available: https://www.facebook.com/photo.php?fbid=10152797001415192&set=p.10152797001415192&type=1&theater. [Accessed: 17-Aug-2015].

[17] W. H. Dutton, G. Blank, "Next Generation Users, The Internet in Britain: Oxford Internet Survey 2011," Oxford, 2011.

[18] W. H. Dutton, E. J.. Helsper, and M. M. Gerber, "Oxford Internet Survey 2009," Oxford, 2009.

[19] ITU World Telecommunications, "ICT Facts and Figures 2005-2014."

[20] S. Gorard and N. Selwyn, "Towards a le@rning society? the impact of technology on patterns of participation in lifelong learning," *Br. J. Sociol. Educ.*, vol. 26, no. 1, pp. 71–89, Jan. 2005.

[21] Office for National Statistics, "Statistical Bulletin Internet Access Quarterly Update, Q1 2013," *Stat. Bull.*, no. February, 2013.

[22] C. W. Olphert, L. Damodaran, and A. May, "Towards digital inclusion–engaging older people in the 'digital world,'" in *Design in the Digital World*, 2005, p. 7.

[23] P. Millward, "Perception, exclusion and barriers of access to the Internet for older people," *First Monday*, vol. 8, no. 7. Ghosh, Rishab Aiyer, 07-Jul-2003.

[24] B. Clinton, "Remarks by the President at Massachusetts Institute of Technology 1998." Boston, 1998.

[25] C. W. Olphert and L. Damodaran, "Older people and digital disengagement: A fourth digital divide?," *Gerontology*, vol. 59, no. 6. pp. 564–570, 2013.

[26] W. H. Dutton, G. Blank with D. Groselj, "Cultures of the Internet: Oxford Internet Survey 2013," Oxford, 2013.

[27] J. Cooper, "The digital divide: the special case of gender," *J. Comput. Assist. Learn.*, vol. 22, no. 5, pp. 320–334, Sep. 2006.

[28] P. F. Cleary, G. Pierce, and E. M. Trauth, "Closing the digital divide: understanding racial, ethnic, social class, gender and geographic disparities in Internet use among

school age children in the United States," *Universal Access in the Information Society*, vol. 4, no. 4. pp. 354–373, 2006.

[29] W. Young, J. Clarke, G. Klima, V. Gadag, L. Gien, and I. Hardill, "The Technology Collection Sustaining Information and Communication Technology use among Canadians with at Least One Activity Limitation," *Int. J. Technol. Knowl. Soc.*, vol. 7, no. 1, pp. 1–12, 2012.

[30] Office for National Statistics, "Internet Access Households and Individuals," Office of National Statistics, 2009.

[31] Office for National Statistics, "Internet Access - Households and Individuals, 2015," 2015.

[32] D. Lewin, S. Adshead, B. Glennon, B. Williamson, T. Moore, L. Damodaran and P. Hansell, "Assisted living technologies for older and disabled people in 2030."

[33] KT Equal, "Taming the Dragon; making technology work for us," 2010.

[34] P. Gregor, A. F. Newell, and M. Zajicek, "Designing for dynamic diversity: interfaces for older people," in *Diversity*, 2002, vol. Edinburgh, no. May, pp. 151–156.

[35] B. Shneiderman, "Design: CUU: bridging the digital divide with universal usability," *interactions*, vol. 8, no. 2. pp. 11–15, 2001.

[36] Avaaz.org, "Avaaaz.org The World in Action," 2014. [Online]. Available: http://avaaz.org/en/about.php. [Accessed: 11-Nov-2015].

[37] 38 Degrees, "38 Degrees: people. power. change," 2015. [Online]. Available: https://home.38degrees.org.uk/.

[38] SumOfUs, "SumOfUs: Fighting for people over profits," 2015. [Online]. Available: http://sumofus.org/.

[39] S. Sayago and J. Blat, "About the relevance of accessibility barriers in the everyday interactions of older people with the web," in *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibililty (W4A) - W4A '09*, 2009, p. 104.

[40] J. Sandhu, L. Damodaran, and L. Ramondt, "ICT Skills Acquisition by Older People: Motivations for learning and barriers to progression," *Int. J. Educ. Ageing*, vol. 3, no. 1, 2013.

[41] L. Ramondt, J. Sandhu and L. Damodaran, "Staying digitally connected – a study of learning and support provision for older people in seven cities in the UK and the implications for policy and practice," *Int J Educ Ageing*, vol. 3, pp. 95–114, 2013.

[42] A. F. Newell, "Interfaces for the ordinary and beyond," *IEEE Softw.*, vol. 10, no. 5, 1993.

[43] M. T. Atkinson, M. Bell, and C. Machin, "Towards ubiquitous accessibility: capability-based profiles and adaptations, delivered via the semantic web," … *Conf. Web Access.*, pp. 2–5, 2012.

[44] M. T. Atkinson, "Collaborative Adaptive Accessibility and Human Capabilities," Loughborough University, 2012.

[45] L. Damodaran, C. W. Olphert, and J. Sandhu, "Falling off the bandwagon? Exploring the challenges to sustained digital engagement by older people," *Gerontology*, vol. 60, no. 2, pp. 163–173, 2014.

[46] A. M. Piper, R. Campbell, and J. D. Hollan, "Exploring the accessibility and appeal of surface computing for older adult health care support," *Proc. CHI '10*, pp. 907–916, 2010.

[47] YouGov, "TV may be the key for device connectivity and content sharing," 2015. [Online]. Available: https://yougov.co.uk/news/2015/03/05/tv-may-be-key-device-connectivity-and-content-shar/.

[48] A. Dickinson, R. Eisma, P. Gregor, A. Syme, and S. Milne, "Strategies for teaching older people to use the world wide web," *Univers. Access Inf. Soc.*, vol. 4, no. 1, pp. 3–15, 2005.

[49] T. A. Salthouse, "The processing-speed theory of adult age differences in cognition.," *Psychol. Rev.*, vol. 103, no. 3, pp. 403–428, 1996.

[50] M. Csikszentmihalyi, *Flow: The psychology of optimal performance*. 1990.

[51] J. Lazar, A. Dudley-Sponaugle, and K. D. Greenidge, "Improving web accessibility: A study of webmaster perceptions," *Comput. Human Behav.*, vol. 20, no. 2, pp. 269–288, 2004.

[52] T. Gilbertson, "Attitudes and behaviours towards web accessibility and ageing: Results of an industry survey," *Gerontechnology*, vol. 13, no. 3, pp. 337–344, Dec. 2014.

[53] P. T. Jaeger and M. Matteson, "e-Government and Technology Acceptance: the Case of the Implementation of Section 508 Guidelines for Websites," *Electron. J. eGovernment*, vol. 7, no. 1, pp. 87–98, 2009.

[54] A. Eccles, L. Damodaran, W. Olphert, I. Hardill, M. Gilhooly, "Assistive Technologies: Ethical Practice, Ethical Research, and Quality of Life," in *Technologies for Active Aging*, A. Sixsmith and G. Gutman, Eds. Boston, MA: Springer US, 2013.

[55] Information Commissioners Office, "Guidance on the rules on use of cookies and similar technologies." 2012.

[56] Ipsos MediaCT, "Understanding Digital Capabilities follow-up September 2013 and March 2014," 2014.

[57] The Daily Beast, "Your Samsung SmartTV Is Spying on You, Basically," 2015. [Online]. Available: http://www.thedailybeast.com/articles/2015/02/05/your-samsung-smarttv-is-spying-on-you-basically.html. [Accessed: 11-Sep-2015].

[58] Samsung tomorrow, "Samsung Smart TVs Do Not Monitor Living Room Conversations." [Online]. Available: http://global.samsungtomorrow.com/samsung-smart-tvs-do-not-monitor-living-room-conversations/. [Accessed: 11-Sep-2015].

[59] D. Frolich, C. Lim, S. Woods, A. Ahmed, "What older people want: A catalogue of co-designed ICT concepts," 2012.

[60] New Dynamics of Ageing, "Sus-IT ('Sustaining IT use by older people to promote autonomy and independence')." .

[61] A. Waller, V. L. Hanson, and D. Sloan, "Including accessibility within and beyond undergraduate computing courses," in *11th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '09*, 2009, pp. 155–162.

[62] S. A. Youngblood, "Communicating Web Accessibility to the Novice Developer: From User Experience to Application," *J. Bus. Tech. Commun.*, vol. 27, no. 2, pp. 209–232, Sep. 2012.

[63] Usable Technology for Older People: Inclusive and Appropriate. UTOPIA, *Video Trilogy*. 2004.

[64] A. Carmichael, A. F. Newell, and M. Morgan, "The efficacy of narrative video for raising awareness in ICT designers about older users' requirements," *Interact.*

Comput., vol. 19, no. 5–6, pp. 587–596, 2007.

[65] M. Morgan, V. Hanson, C. Martin, J. Hughes, and A. Newell, "Accessibility Challenge - a Game Show Investigating the Accessibility of Computer Systems for Disabled People," in *Proceedings of ACM CHI 2008 Conference on Human Factors in Computing Systems*, 2008, vol. 2, pp. 2609–2610.

[66] St. George's House, "Consulation Report: Falling off the bandwagon? Sustaining digital engagement by older people," Windsor, 2012.

[67] KT-EQUAL, New Dynamics of Ageing, Sus-IT, "Promoting Digital Participation The Proposition: Community hubs. Meeting older people's technology support needs, developing social communities and reducing isolation.," Loughborough, 2012.

Analyzing Personalized Walking in Smart Cities Through a Multi-Modal Transportation Simulation Environment

Monsak Socharoentum
National Electronics and Computer Technology Center,
National Science and Technology Development Agency
Pathumthani, Thailand
e-mail: monsak.soc@nectec.or.th

Hassan A. Karimi
Geoinformatics Laboratory, School of Information Sciences, University of Pittsburgh
Pittsburgh, USA
e-mail: hkarimi@pitt.edu

***Abstract*— While current transportation simulations can be used to evaluate vehicle trips or neighborhood walkability, none is able to evaluate trips that require multi-modal transportation when walking is always one mode. In this paper, we address this gap by the Multi-Modal Transportation (MMT) with Multi-Criteria Walking (MMT-MCW) simulation. MMT-MCW simulation can be used to evaluate various aspects of smart cities, such as walkability. The premise of MMT-MCW is based on the observations that: (a) walking can be performed for other purposes besides merely reaching destinations, such as to maintain or improve health and (b) traveler's characteristics and preferences play an important role in determining optimal route choices. Selected MMT-MCW scenarios were simulated to evaluate walkability of several cities with respect to three criteria: inter-modal transfer locations (parking lots and bus stops) elevation of walking routes, and walking distance. The simulation results show that: (a) despite similar elevation range, cities may contain walking routes with significant different average calories burn and (b) walking paths connected to parking lots and bus stops are not necessarily correlated to number of routes, amount of calories burnt, and elevation.**

Keywords-smart cities; walkability; multi-modal transportation; routing; multi-criteria walking.

I. INTRODUCTION

The concept of Multi-Modal Transportation with Multi-Criteria Walking is discussed in [1]. MMT-MCW is based on information and communication technologies and is designed to offer personalized transportation services. In its initial version, MMT-MCW is built as a simulation environment to analyze scenarios where walking is always one mode and for various purposes (multi-criteria). Two sets of factors impacting MCW are environmental factors and traveler factors. Environmental factors, compared to driving cars or riding public transportation, may have a greater impact on walking. For example, people may prefer driving cars or riding buses over walking due to rain, snow, hilly terrain, or air pollution. Location is also an environmental factor that influences walking, for instance, fastest walking routes may be based on flat and short routes, which take priority over steep and long routes. However, when walking is for exercise, the steeper and/or longer route may be preferred. Traveler factors, such as individuals' characteristics, also have an impact on choosing walking routes [2]. Several studies, such as [3], reported a correlation between individual behaviors and walking. Studies by Leslie et al. [4] and Ewing et al. [5] are examples related to the urban area evaluation in terms of neighborhood walkability. Factors, such as land use and urban design [6], [7], aging [8], air pollution [9], body mass index [10], preference for passive transport [11], and pedestrian's preference [12], that influence walking have been studied. The findings of various research projects highlight the benefits, such as weight loss [13], of replacing short motorized trips with walking. In other studies, the relationship between the built environment and walking has been studied. A longitudinal analysis is discussed in [14], itineraries to destinations is considered in [15], level of service of urban walking environment and its influence are investigated in [16], GPS-measured walking, bicycling and vehicle time in adolescents are analyzed in [17].

In this work, MMT-MCW is built as a simulation environment to find connections between different modes of transportation by utilizing public transportation infrastructures (road networks and sidewalk networks) and provide the residents of smart cities with flexible and personalized walking places. The following two example simulation cases show how MMT-MCW can be used for smart cities.

By simulating various walking scenarios in MMT-MCW, walking clusters ("walking hotspots") in a city can be identified; these walking hotspots would help urban designers and planners determine locations where new information and sensors would benefit the residents. For example, by knowing locations of walking hotspots in a city, real-time information about activities and traffic on sidewalks of those hotspots can be made available on a website, or directly be sent to the residents. Similarly, appropriate sensors, such as those that can assist pedestrians with safety issues on sidewalks, can be installed. An example sensor could be one detecting within a specified range a sharp elevation change between two adjacent sidewalk segments.

In another simulation through MMT-MCW, various connection scenarios between road networks and sidewalk networks for driving first and then walking (driving-walking) can be analyzed in a city in order to identify most usable transfer nodes ("transfer hotspots"). By knowing locations of transfer hotspots, which could be parking garages or street

parking places, sensors can be installed to detect transfer node availability and inform the residents, through a website or notifications on mobile devices, in real time. An example sensor could be video images of specific parking places. Such video images when collected over a period of time can be mined so that transfer node (parking space) availability can be predicted and communicated to interested residents.

Despite the benefits of MMT-MCW for evaluating transportation options in smart cities, currently there is no research that is focused on evaluation of city' transportation infrastructures and utilities (e.g., parking locations and walking routes). To fill this gap, MMT-MCW simulation is proposed to evaluate three basic options: (a) inter-modal transfer locations (parking lots and bus stops); (b) elevation of walking routes; and (c) walking distance. The first option is related to MMT, and the last two are related to MCW.

MMT-MCW may be implemented in several ways for smart cities, for example, as a new service for individuals interested in finding routes that include walking components. [18] developed a prototype service (called Route2Health) that recommends walking sessions, if feasible, for any trip. By taking origin, destination, and traveler's conditions as input, Route2Health recommends a sequence of transportation modes along with specific details about each mode that is most optimal (personalized). MMT-MCW can also be implemented to simulate the design of smart cities.

The paper's contribution is a novel integration of new and existing Web techniques and technologies for evaluating and analyzing transportation options for smart cities. A Web-based tool (simulation) is developed to analyze mushing up data and find transportation solutions based on existing Web services (Google Map APIs). The rest of the paper is organized as follows. Section II describes MMT-MCW. Sections III and IV discuss MMT-MCW simulation and example scenarios and their results. The paper ends with a summary and suggestions for future research in Section V.

II. MMT-MCW

MMT-MCW is designed to find: (a) multi-modal transportation routes with walking as one mode and (b) optimal walking paths by considering multiple criteria. Walking transfer node and route score are the two factors that MMT-MCW considers in finding optimal solutions (routes).

Three modes of transportation are considered in MMT-MCW: walking, driving, and riding (bus). We define "walking transfer node" as a location where travelers switch from a pedestrian network to a vehicular network, or vice versa. In MMT-MCW, walking transfer nodes play an important role in finding suitable (personalized) routes. For example, change of one parking lot to another (as a walking transfer node) may result in a different (and desired) solution. With respect to public transportation, the choice of a bus stop (as a walking transfer node) determines a specific bus route. To identify a suitable walking transfer node, traveler's desired walking distance is separated into estimated upper and lower limits. The upper limit excludes walking transfer nodes that are located beyond a traveler's maximum preferred distance. The lower limit excludes walking transfer nodes that are located closer than the desired minimum walking distance. Accordingly, one or more suitable walking transfer nodes are identified.

Route score is used to quantify the suitability of a walking route in meeting traveler's criteria. To compute a route score, a relevant criterion must be identified and used to formulate a suitable metric function. Examples of route score criterion are: (1) traveler's desire to burn a specific amount of calories by walking and (2) traveler's preference for a certain level of elevation variation. The route score for the first criterion should reflect calories burnt on walking and for the second criterion should reflect elevation variation.

To calculate calories burnt on walking, the ACSM walking equation [19] can be used:

$$EE = (0.1 \cdot S + 1.8 \cdot S \cdot G + 3.5) \cdot BM \cdot t \cdot 0.005 \quad (1)$$

where EE is walking energy expenditure (kilocalories), S is walking speed (meters/minute), G is grade (slope) in decimal form (e.g., 0.02 for 2% grade), BM is traveller's body weight (kilograms), and t is walking time (minutes). The constant 0.005 is the amount of energy expenditure burnt per one kilogram per one minute.

To calculate elevation variation, walking surface roughness is used. The walking surface roughness refers to the standard deviation of the elevations along an entire walking route. The standard deviation of a flat walking route is zero, and the higher value of walking surface roughness refers to higher variation of elevations along the walking route.

III. Simulation and Data Collection

Several cities within the United States were used in the simulation. Twelve cities were selected based on three attributes: population, body mass index (BMI), and elevation range. The US Office of Management and Budget uses population to define a statistical area. A statistical area contains one or more cities (and/or counties) and can be classified as metropolitan (high-density population) or micropolitan (low-density population). Metropolitan has population greater than 50,000 and micropolitan has population between 10,000 and 50,000. The US Center for Disease Control and Prevention's definitions and categories were considered for normal weight ($18.5<BMI<24.9$) and obese ($BMI>30.0$) where individual's BMI is calculated by dividing the individual's weight (kilogram) by the square of the individual's height (meter); $BMI=weight/(height)^2$. BMI statistics are from the year 2012 provided by the CDC. Elevation range was classified into hilly (elevation range $\geq$100 meters) and flat (elevation range $\leq$ 50 meters), where elevation range is the difference between maximum and minimum elevations of the area. The elevations of 100 randomly selected positions within the city of interest were used to calculate the elevation range. The two threshold values (50 and 100 meters) were chosen for separating between hilly and flat terrains. The purpose of using the three attributes (population, BMI, and elevation range) is to

explore their influence on walking routes and walking transfer nodes.

For simulations, cities were selected based on statistical areas (ranked in descending order) and on BMI statistics (percentage of normal weight and obese people). To have realistic representatives for each BMI category, city selection was based on rank; a higher rank has a higher priority. To include as many states as possible in simulations, if the second city is located in the same state as the first one, then the next city that is located in a different state is selected. The selected cities are categorized and shown using a tree diagram in Figure 1. There are six possible combinations based on the three mentioned attributes (population, BMI, and elevation range), and up to two cities were selected for each combination.

Figure 2 shows the maximum, minimum, and average elevations of the selected cities, and the following abbreviations are used: Micropolitan (Mi), Metropolitan (Me), Obese (O), Normal weight (N), Hilly (H), and Flat (F).

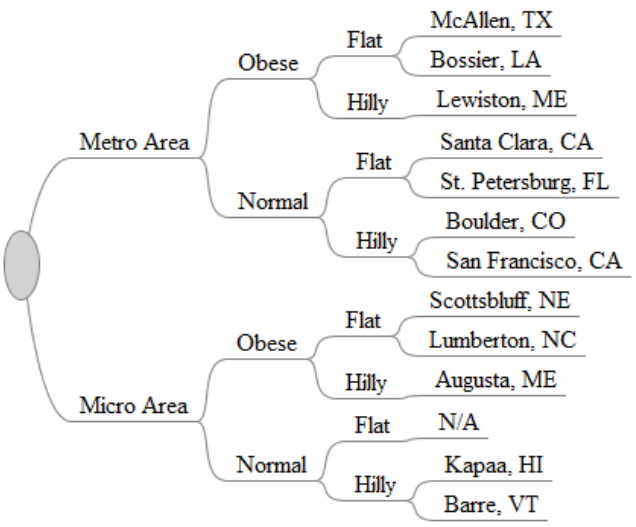


Figure 1. City and state selected for each category.

Two different MMTs were simulated: driving-walking and riding-walking. A driving-walking trip usually comprises (in sequence) driving, parking, and then walking, and return in the reverse sequence. Unlike driving-walking, travelers do not have to begin with riding (public transportation) in a riding-walking trip. The trip may start by walking from origin to a nearby bus stop then taking bus to destination. Walking can also be in the middle to connect two different bus routes, and the return trip can be in any sequence. For simplicity, the return trips were not considered and walking was assumed as the mode connecting the walking transfer nodes and destinations. To this end, walking transfer nodes and walking routes were simulated. Note that the vehicular route computation between origin and walking transfer node was not considered since it is not the MMT-MCW's main contribution. Parking lots, bus stops, walking routes, sidewalk slopes, and points of interest (POIs) were other data considered for simulations.

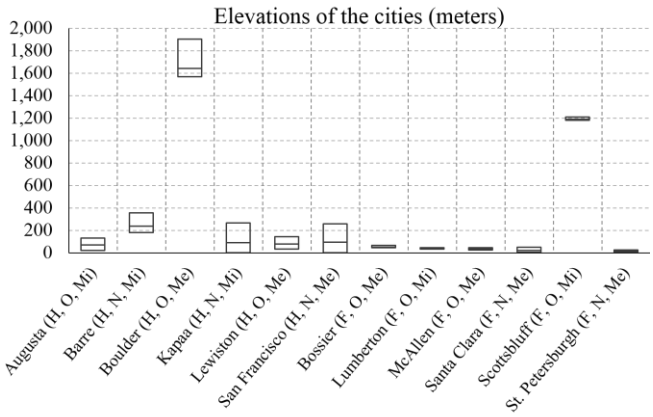


Figure 2. Maximum, minimum, and average elevations of the selected cities.

TABLE I. SIMULATION PARAMETERS

Parameter Names	Values	Units
1. Desired walking distance	1.0	kilometer
2. Inner radius	0.5	kilometer
3. Outer radius	1.5	kilometer
4. Maximum number of suitable parking lot locations	20	-
5. Maximum number of suitable bus stop locations	No limit	-
6. Body weight trial values	60, 80, 100, 120	kilogram
7. Walking speed trial values	60, 80, 100	meters/minute

The walking distance between a walking transfer node and a destination was assumed to be one kilometer. POI locations were selected from OpenStreetMap [20] and 100 destinations within each city were randomly selected (in case the number of POIs in a city was less than 100, all POIs were used). To identify suitable parking lots, a buffer (inner radius: 0.5 kilometer; outer radius: 1.5 kilometer) around each destination was created. For each destination, up to 20 parking lots within a buffer were selected as suitable walking transfer nodes (note that 20 is an arbitrary number and each destination may have a different number of parking lots). Bus stops and bus routes data were collected from Google Transit Feed Specification [21]. For each suitable parking lot and bus stop, up to three candidate walking routes were generated (ordered by travel time). Parking lot locations and walking routes were retrieved from Google Place API and Google Direction API, respectively. Once all candidate routes were computed, elevation of points along the walking route of interest was retrieved from Google Elevation API, and then Equation (1) was used to calculate calories burned for each candidate walking route. Walking surface roughness was also calculated using the elevations of route segments. To reflect multiple traveler's characteristics, four body weights (60, 80, 100, and 120 kilograms) and three walking speeds (60, 80, and 100 meters/minute) were simulated. Note

that these walking speeds are for experimentation; see [22] and [23] for research on walking speed. Table I summarizes the simulation parameters.

IV. Simulation Results

The simulation results are related to four entities: suitable parking lots, suitable bus stops, walking routes that connect parking lots and destination (PK routes), and walking routes that connect bus stops and destinations (BS routes). Suitable parking lots and suitable bus stops are the parking lots and bus stops that are located within a ring buffer (inner radius: 0.5 kilometer; outer radius: 1.5 kilometer) around each destination. All 12 cities contain suitable parking lots and acceptable PK routes, but only two cities (San Francisco and Santa Clara) contain suitable bus stops and acceptable BS routes (these are the only two cities, among the selected cities, that provide coordinates of their bus stops). PK and BS routes are considered acceptable if their distances are within 0.9 and 1.1 kilometer from destinations. Based on these four entities, results are separated into three sub-sections: (A) suitable parking lots and acceptable PK routes, (B) suitable bus stops and acceptable BS routes, and (C) acceptable PK routes vs acceptable BS routes.

A. Suitable Parking Lots and Acceptable PK Routes

Figure 3 shows the selected cities (on x-axis), the numbers of destinations (on y-axis), and the counts of destinations that have suitable parking lots (on y-axis). The following abbreviations are used in the figure: Micropolitan (Mi), Metropolitan (Me), Obese (O), Normal weight (N), Hilly (H), and Flat (F). Most cities in metropolitan areas have a large number of destinations with suitable parking lots except Bossier (1 out of 72) and McAllen (4 out of 99). Four cities (Barre, Kappa, Scottsbluff, and Bossier) have zero or only one destination with at least one suitable parking lot, which are considered outliers and excluded from the analysis. From the figure, both obese and normal weight groups fall within cities with both small and large number of destinations with suitable parking lots. This indicates that there is no obvious separation between the two groups with respect to number of destinations.

Figure 4 shows maximum, minimum, and average number of calories burned (top left) and walking surface roughness (lower left) for walking routes that connect to parking lots (PK routes). On x-axis, the first four cities are hilly and the latter four are flat. The graphs indicate that hilly cities have wider ranges of both calories burnt and walking surface roughness. This is because both calories burnt and walking surface roughness are directly related to the elevation range of hilly cities and the walking routes. An interesting observation is that most cities (except Boulder) in the left figure have a similar average calories burnt regardless of the elevation range. Although Boulder has a similar walking surface roughness compared to other hilly cities, its average calories burnt is significantly higher than the others. This indicates that walking routes in Boulder are better in terms of burning calories.

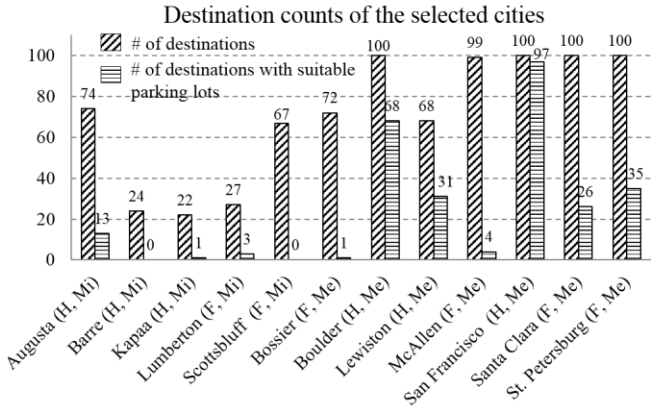


Figure 3. Number of destinations and counts of destinations that have ≥ 1 suitable parking lots.

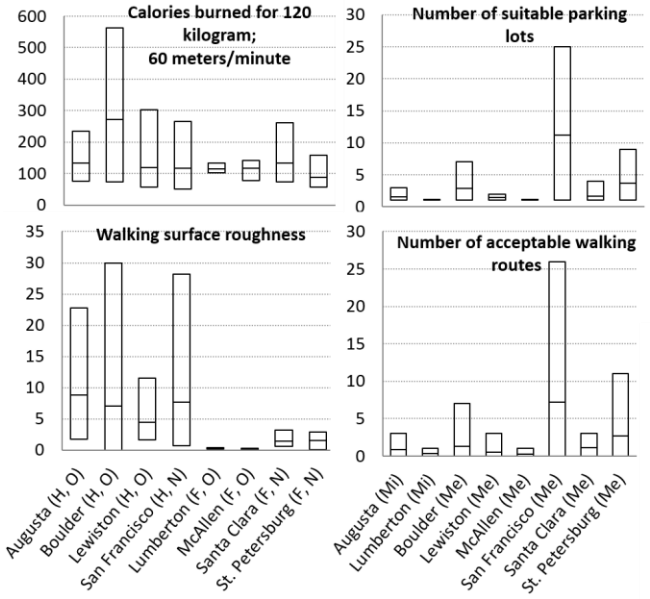


Figure 4. Comparison of attributes related to PK routes in different cities.

Figure 4 (top right) shows number of suitable parking lots and (lower right) shows number of acceptable walking routes. Acceptable walking routes refer to walking routes that have their distance within 0.9 and 1.1 kilometer (±10% of the designated one kilometer walking distance) from destinations. The graphs show that San Francisco has the highest average number of suitable parking lots, the highest number of acceptable walking routes, and the largest range on both attributes (largest variation of results); this is expected for a metropolitan city where transportation infrastructures are dense. Note that San Francisco is the 13th most populous city in the United States [24].

Figure 5 shows spatial distribution of destinations that have acceptable PK routes. Destinations in Augusta and Santa Clara have a very small number (between 1 and 3) of acceptable PK routes. St. Petersburg reveals a road (north-

south direction) that has destinations with a high number of acceptable PK routes. Destinations in Boulder spatially spread across the city and do not show an explicit pattern. San Francisco has most of its destinations with a large number of acceptable PK routes. Most destinations in San Francisco cluster together in the north-east region of the city because most of the POIs (which are used as destinations) in San Francisco are also located in the north-east region.

Figure 6 shows average calories burn for the acceptable PK routes grouped by destinations. Each map has legends showing minimum and maximum values with circle sizes. All the cities (except Boulder) have their average and maximum value lower than 200.

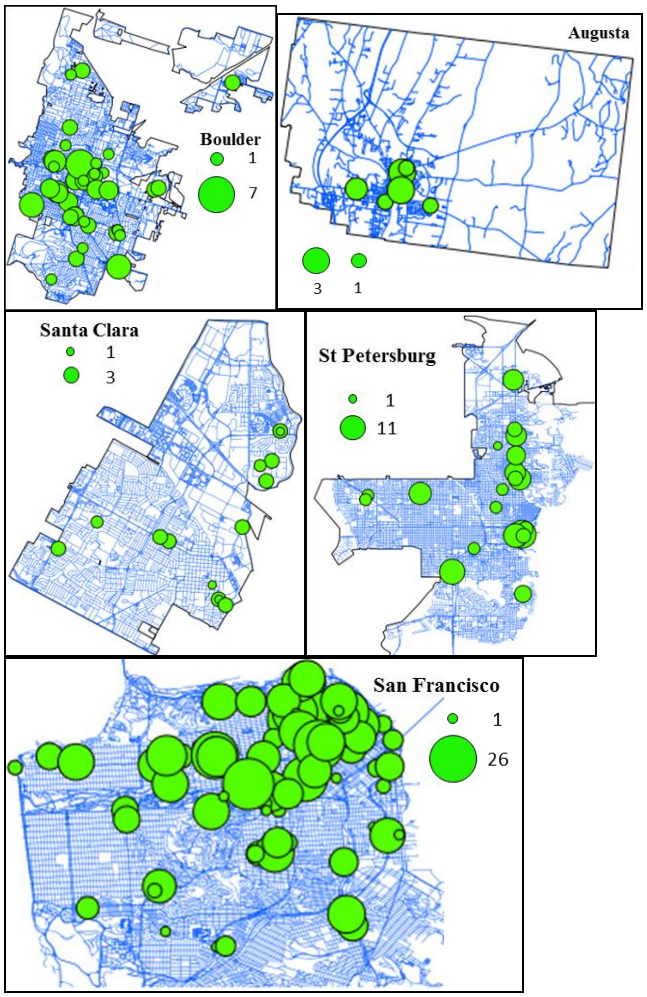


Figure 5. Destination distributions with number of acceptable PK routes.

This sub-section discusses and compares suitable parking lots and acceptable PK routes in five cities: Augusta, Boulder, San Francisco, Santa Clara, and St. Petersburg. San Francisco has the highest average numbers for both suitable parking lots (11.16) and acceptable PK routes (7.2). With similar walking distance and time, Boulder offers the best environment in terms of burning calories. Augusta has a high amount of calories burn only for destinations within the city's inner area, which has a dense road network. Even though Santa Clara is classified as a metropolitan city, it does not provide a large number of suitable parking lots. Santa Clara has 1.70 suitable parking lots on average, while Augusta (a micropolitan city) has 1.54.

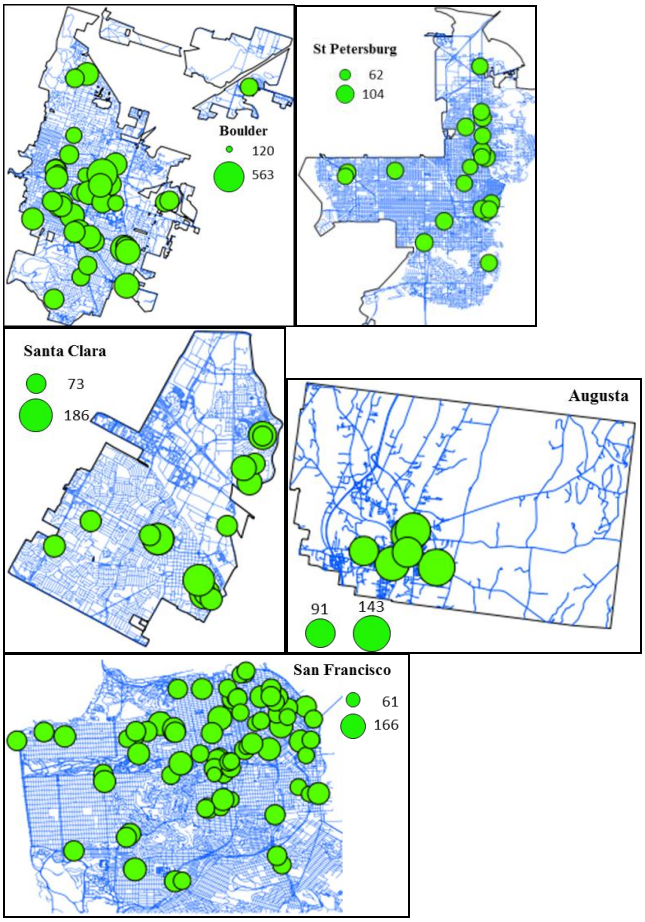


Figure 6. Destination distributions with calories burn for acceptable PK routes.

B. *Suitable Bus Stops and Acceptable BS Routes*

Figure 7 shows spatial distribution of destinations in San Francisco and Santa Clara. The destinations are classified into two groups: (1) destinations that have suitable bus stops (circle shape) and (2) destinations that have no suitable bus stops (triangle shape). According to Figure 7 (left), there is a clear distinction between the two groups in San Francisco. The first group densely clusters within the north-east region of the city, while the second group is surrounding the first group. All destinations in Santa Clara have at least one suitable bus stop.

Figure 8 shows spatial distribution of destinations with number of suitable bus stops. None of the destinations in San Francisco has more than four suitable bus stops, while the destinations in Santa Clara have much larger number of

suitable bus stops. Figure 9 shows spatial distribution of destinations with number of acceptable BS routes. None of the destinations in San Francisco has more than five acceptable BS routes, while the destinations in Santa Clara have wider range of number of acceptable BS routes. Most destinations with large number of acceptable BS routes in Santa Clara cluster within the inner region of the city. A counter intuitive observation is that despite the denser road network (which means a large number of road segment connections and route choices), the destinations in San Francisco still have fewer number of acceptable BS routes compared to Santa Clara. Average number of BS routes is 2.41 for San Francisco and is 9.29 for Santa Clara. Note that destinations that do not have acceptable BS routes were not included in calculating the averages. An interesting observation related to suitable bus stops and acceptable BS routes in Santa Clara is that destinations in the lower region of the city have a large number of bus stops but have a small number of BS routes, while destinations in the upper region have the opposite numbers.

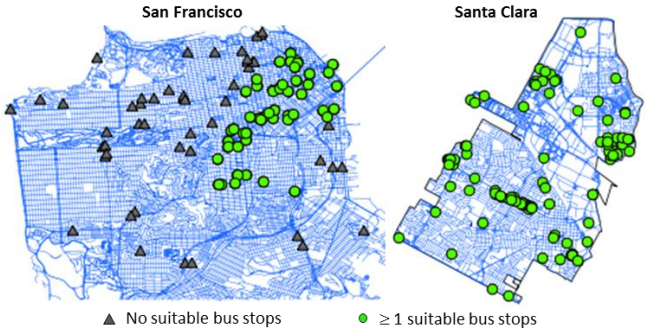


Figure 7. Suitable bus stops availability.

Figure 8. Number of suitable bus stops.

Figure 10 shows spatial distribution of destinations with number of bus routes. All destinations that have a suitable bus stop in San Francisco have four bus routes. Two and seventeen are the minimum and maximum numbers of bus routes for destinations in Santa Clara, and destinations in the northern region tend to have a larger number of bus routes than the others.

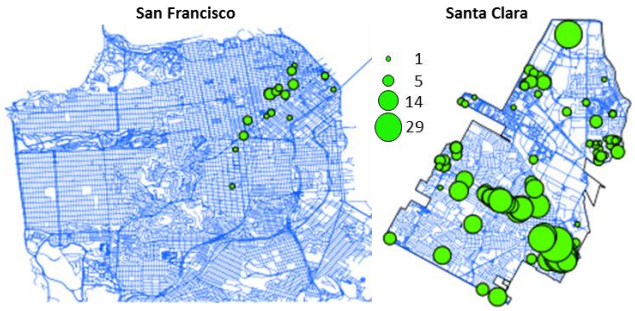


Figure 9. Number of acceptable BS routes.

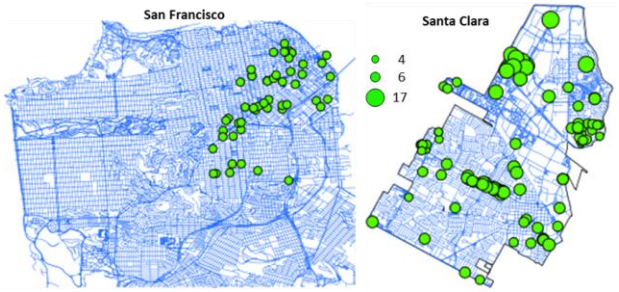


Figure 10. Number of bus routes.

This sub-section discusses and compares suitable bus stops and acceptable BS routes in San Francisco and Santa Clara. Only destinations in the north-east region of San Francisco have suitable bus stops, while all destinations in Santa Clara have suitable bus stops. Assuming a direct relationship between road and sidewalk densities, San Francisco, which has a denser road network, is expected to have a higher average number of BS routes, however, the results show otherwise. This observation indicates that road network density does not necessarily correlate with the number of acceptable BS routes.

C. Acceptable PK Routes VS Acceptable BS Routes

Figures 11 and 12 show the comparisons between acceptable PK routes and acceptable BS routes in San Francisco and Santa Clara, which are the only two cities (among the selected cities) that publish their bus stop coordinates. Each bar graph represents maximum, minimum, and average values. In San Francisco, PK routes have higher average values than BS routes, while, in Santa Clara, it is the opposite. This indicates that PK routes and BS routes are not necessarily correlated. Considering walking surface roughness, BS routes in San Francisco have narrower range than PK routes, meaning that BS routes are generally flatter than PK routes. This can be inferred that the acceptable BS routes are available mostly in the flat areas. However, an interesting observation is that BS routes in Santa Clara show opposite behavior such that PK routes are flatter than BS routes. It should also be noted that both PK and BS routes in San Francisco have a much larger walking surface roughness (by around 10 times) than their counterparts in Santa Clara, meaning that PK and BS routes

in Santa Clara are much flatter. A counter intuitive observation from calories bar graphs in both Figures 11 and 12 is that despite different elevation range (San Francisco: Hilly; Santa Clara: Flat) and a large walking surface roughness difference, both PK and BS routes in San Francisco still have average calories close to their counterparts in Santa Clara (~120 calories for PK routes; ~140 calories for BS routes). This is because the amount of calories burn was estimated using the same walking speed (60 meters/minute) in both cities. Therefore, despite the close amounts of estimated calories burn, walking routes in San Francisco help burn more calories within the same period of time compared to the walking routes in Santa Clara.

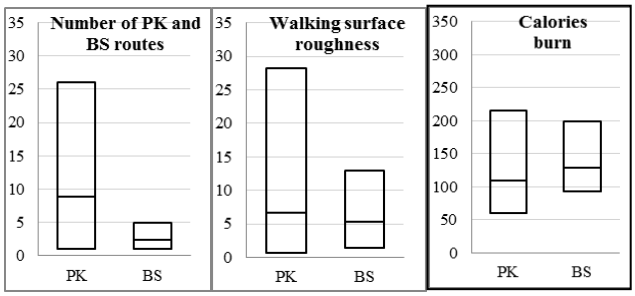


Figure 11. Comparisons between PK routes and BS routes for San Francisco.

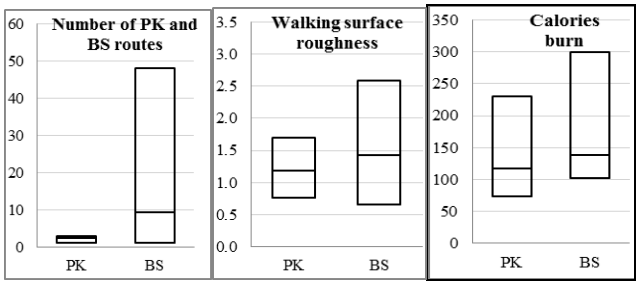


Figure 12. Comparisons between PK routes and BS routes for Santa Clara.

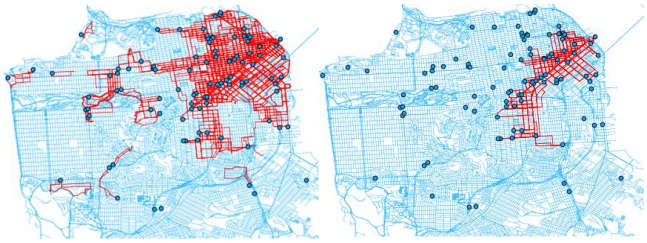


Figure 13. San Francisco: PK routes (left) and BS routes (right).

Figures 13 and 14 show the spatial distribution of destinations and the spatial coverage of PK and BS routes in San Francisco and Santa Clara overlaid on the cities' road networks. The maps indicate the inverse behavior between the two cities such that PK routes have more coverage than BS routes in San Francisco, and vice versa for Santa Clara. Figure 15 shows the comparisons of destination distributions with number of acceptable PK and BS routes in San Francisco and Santa Clara. The maps indicate the inverse behavior between the two cities such that the number of acceptable PK routes is much larger than that of BS routes in San Francisco, and vice versa for Santa Clara.

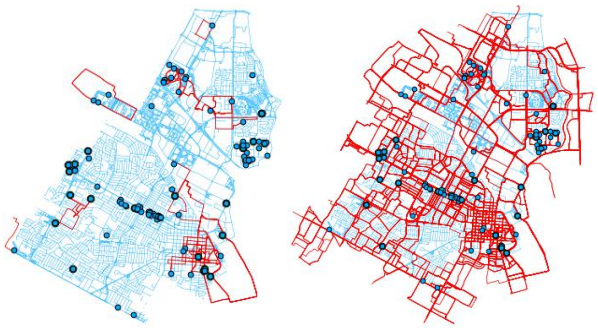


Figure 14. Santa Clara: PK routes (left) and BS routes (right).

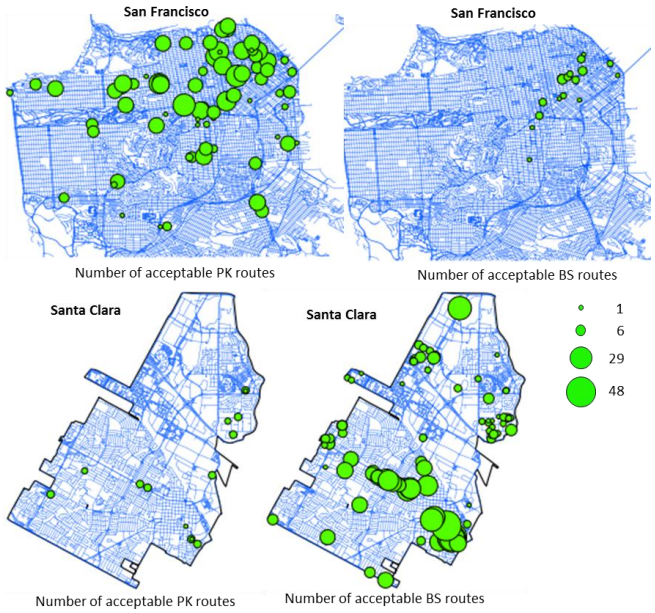


Figure 15. Comparisons between destination distributions with number of acceptable PK and BS routes.

In this sub-section, PK routes and BS routes of San Francisco and Santa Clara are compared. The results show that PK routes and BS routes are not necessarily correlated. For driving-walking mode, San Francisco has much more parking lots and acceptable PK routes than Santa Clara. For riding-walking mode, Santa Clara has much more bus stops and acceptable BS routes than San Francisco. Both PK and BS routes in San Francisco have much larger walking surface roughness (by around 10 times) than their counterparts in Santa Clara, meaning that PK and BS routes in San Francisco have much higher elevation variations.

V. SUMMARY AND FUTURE RESEARCH

This paper presented a new simulation approach for evaluating smart cities. Scenarios in selected cities were

simulated. The simulation results show that: (a) despite similar elevation range, cities may contain walking routes with significant different average calories burn and (b) PK routes and BS routes are not necessarily correlated (e.g., San Francisco and Santa Clara).

Considering that enhancing health and wellbeing of people, among other benefits, is one objective for building smart cities, our proposed approach can be used to evaluate smart cities for their environment infrastructures (roadways and sidewalks) and transportation infrastructures (different modes) and as a simulation tool to design new smart cities.

Some future research directions are:

- Investigating and developing MCW optimization algorithms for travelers, such as people with disabilities (e.g., wheelchair users and people who are blind or visually impaired), people with special physical conditions (e.g., people with joint problems), and people with health conditions (e.g., people who must be less exposed to air pollution or sun light).
- Investigating and developing a predictive MMT-MCW methodology that allows route request well in advance and can monitor the recommended route up to minutes before the route is taken and update the recommendation based on changes of environmental and individual factors.

REFERENCES

[1] M. Socharoentum and H. A. Karimi, "Simulation of Web-Based Multi-Modal Transportation with Multi-Criteria Walking for Smart Cities," In Proceedings of The Third International Conference on Building and Exploring Web Based Environments (WEB 2015), Rome, Italy, May 24-29 2015, pp. 14-18.

[2] A. W. Agrawal, M. Schlossberg, and K. Irvin, "How far, by which route and why? A spatial analysis of pedestrian preference," Journal of Urban Design, 13(1), 2008, pp. 81-98.

[3] J. Sun, M. Walters, N. Svensson, and D. Lloyd, "The influence of surface slope on human gait characteristics: a study of urban pedestrians walking on an inclined surface," Ergonomics, 39(4), 1996, pp. 677-692.

[4] E. Leslie, B. Saelens, L. Frank, N. Owen, A. Bauman,N. Coffee, and G. Hugo, "Residents' perceptions of walkability attributes in objectively different neighbourhoods: a pilot study," Health and Place, 11(3), 2005, pp. 227-236.

[5] R. Ewing, S. Handy, R. C. Brownson, O. Clemente, and E. Winston, "Identifying and Measuring Urban Design Qualities Related to Walkability," Journal of Physical Activity and Health, 3, 2006, pp. s223-s240.

[6] A. Forsyth and M. Southworth, "Cities Afoot—Pedestrians, Walkability and Urban Design," Journal of Urban Design, 13(1), 2008, pp. 1-3.

[7] R. Ewing and S. Handy, "Measuring the unmeasurable: Urban design qualities related to walkability," Journal of Urban Design, 14(1), 2009, pp. 65-84.

[8] A. C. King et al., "Aging in neighborhoods differing in walkability and income: associations with physical activity and obesity in older adults," Social Science and Medicine, 73(10), 2011, pp. 1525-1533.

[9] J. D. Marshall, M. Brauer, and L. D. Frank, "Healthy neighborhoods: walkability and air pollution," Environmental Health Perspectives, 117(11), 2009, pp. 1752-1759.

[10] K. R. Smith, B. B. Brown, I. Yamada, L. Kowaleski-Jones, C.D. Zick, and J. X. Fan, "Walkability and body mass index: density, design, and new diversity measures," American Journal of Preventive Medicine, 35(3), 2008, pp. 237-244.

[11] D. V. Dyck, B. Deforche, G. Cardon, and I. D. Bourdeaudhuij, "Neighbourhood walkability and its particular importance for adults with a preference for passive transport," Health and Place, 15(2), 2009, pp. 496-504.

[12] Barbara B. Brown, Ikuho Yamada, Ken R. Smith, Cathleen D. Zick, Lori Kowaleski-Jones, and Jessie X. Fan, "Mixed land use and walkability: Variations in land use measures and relationships with BMI, overweight, and obesity," Health and Place, 15(4), 2009, pp. 1130-1141.

[13] C. Morency, M. Demers, and E. Poliquin, "Shifting short motorized trips to walking: the potential of active transportation for physical activity in Montreal," Journal of Transport & Health, 1(2), 2014, pp. 100-107.

[14] M. W. Knuiman, H. E. Christian, M. L. Divitini, S. A. Foster, F. C. Bull, H. M. Badland, and B. Giles-Corti, "A Longitudinal Analysis of the Influence of the Neighborhood Built Environment on Walking for Transportation The RESIDE Study," American journal of epidemiology, kwu171, 2014.

[15] N. Karusisi, F. Thomas, J. Méline, R. Brondeel, and B. Chaix, "Environmental conditions around itineraries to destinations as correlates of walking for transportation among adults: the RECORD cohort study," 2014.

[16] T. Muraleetharan and T. Hagiwara, "Overall level of service of urban walking environment and its influence on pedestrian route choice behavior: analysis of pedestrian travel in Sapporo, Japan," Transportation Research Record: Journal of the Transportation Research Board, 2014, pp. 7-17.

[17] J. A. Carlson, B. E. Saelens, J. Kerr, J. Schipperijn, T. L. Conway, L. D. Frank, J. E. Chapman, K. Glanz, K. L. Cain, and J. F. Sallis, "Association between neighborhood walkability and GPS-measured walking, bicycling and vehicle time in adolescents," Health & place, 32, 2015, pp. 1-7.

[18] H. A. Karimi and M. Socharoentum, "Route2Health: A Novel Routing Service to Assist in Increasing Physical Activity," In Proceedings of the 7th International Conference on health Informatics, Angers, France, 2014, pp. 43-51.

[19] S. J. Tharrett and J. A. Peterson (Eds.), ACSM's health/fitness facility standards and guidelines. Human Kinetics, 2012, pp. 141.

[20] J. G. Benner and H. A. Karimi, "Geo-Crowdsourcing." In Advanced Location-Based Technologies and Services (Ed: H. A. Karimi)". Taylor and Francis, 2013.

[21] GTFS Data Exchange. Retrieved April 10, 2015 from http://www.gtfs-data-exchange.com/

[22] R. W. Bohannon, "Comfortable and maximum walking speed of adults aged 20—79 years: reference values and determinants." Age and ageing 26.1, 1997, pp. 15-19.

[23] J. E. Himann, D. A. Cunningham, P. A. Rechnitzer, D. H. Paterson, "Age-related changes in speed of walking." Medicine and science in sports and exercise 20.2, 1988: pp. 161-166.

[24] U.S. Census Bureau: State and County QuickFacts. (2010). Retrieved April 10, 2015 from http://www.census.gov/2010census/popmap/.

www.iariajournals.org

International Journal On Advances in Intelligent Systems
issn: 1942-2679

International Journal On Advances in Internet Technology
issn: 1942-2652

International Journal On Advances in Life Sciences
issn: 1942-2660

International Journal On Advances in Networks and Services
issn: 1942-2644

International Journal On Advances in Security
issn: 1942-2636

International Journal On Advances in Software
issn: 1942-2628

International Journal On Advances in Systems and Measurements
issn: 1942-261x

International Journal On Advances in Telecommunications
issn: 1942-2601