International Journal on Advances in Life Sciences

2011 vol. 3 nr. 1&2

The *International Journal on Advances in Life Sciences* is published by IARIA.
ISSN: 1942-2660
journals site: http://www.iariajournals.org
contact: petre@iaria.org

Sponsored by IARIA
www.iaria.org

International Journal on Advances in Life Sciences
Volume 3, Number 1 & 2, 2011

- Steve Wheeler, University of Plymouth, UK

Advanced Knowledge Representation and Processing

- Freimut Bodendorf, University of Erlangen-Nuernberg Germany
- Borka Jerman-Blazic, Jozef Stefan Institute, Slovenia
- Andrew Kusiak, The University of Iowa, USA
- Selmin Nurcan, University Paris 1 Pantheon Sorbonne, France
- Jeff Riley, Hewlett-Packard Australia, Australia
- Lars Taxen, Linkoping University - Norrkoping, Sweden

International Journal on Advances in Life Sciences
Volume 3, Numbers 1 & 2, 2011

CONTENTS

Practical Sensor Network Management Technology for Healthcare Applications

Noriharu Miyaho, Tatsuro Nakamura, Takamasa Shimada, and Noriko Konno
Graduate School of Information Environment Technology
Tokyo Denki University (TDU)
2-1200, Muzai Gakuendai, Inzai, Chiba, Japan
e-mail: miyaho@sie.dendai.ac.jp, m7dec5c896sj8@t.vodafone.ne.jp, shimada@sie.dendai.ac.jp, konno@sie.dendai.ac.jp

***Abstract*—The implementation of telemedicine can make use of a sensor network that carries users' biometric information, as collected by tiny intelligent sensors attached to the human body. We propose a communication system that can be used for telemedicine to improve healthcare and quality of life by making use of bio-sensors, a sensor network and a database with mutual authentication to ensure security. This paper also describes the priority control mechanism in the sensor network, the experimental results, and clarifies the real-time performance of the proposed communication sensor network system.**

Keywords - healthcare, telemedicine, real-time processing, sensor, security, healing, ID-Based key sharing, EEG.

I. INTRODUCTION

Recently, the relaxing effects of visible light and aroma have been attracting attention. A network service in which aroma generators installed at users' locations are remotely controlled through the network has been proposed [1]-[3]. We expect that pain clinic doctors will be able to provide a health care service over the network using a combination of light and aroma. We have investigated a "healing" communication system in order to improve patients' quality of life (QoL). The system uses an aroma generator and a light wavelength controller, both of which are remotely controllable. We implemented a prototype system, and made both psychological and physiological measurements using an electroencephalogram (EEG).

First, the psychological effect of colour was examined in order to determine the actually comfortable healing environments. We examined the effect of different colors in lighting using a 7-point Likert scale from the viewpoint of mental health and the psychological effects that different visible light colors provide. The result is shown in Figure 1. The general evaluation for a healing effect was highest for the green light and the factor ratings were: Happiness factor (5.2625), Popularity factor (4.901), Healing factor (4.886), and Rest factor (4.781). The evaluation for restfulness was highest for blue light, with these factor ratings: Rest factor (4.946), Popularity factor (4.585), Healing factor (4.272), and Happiness factor (3.537), as shown in Figure 1 [4].

EEG is often used in a variety of fields to evaluate multiple cerebral states. It is well known that a specific EEG waveform is generated when the subject is in a specific state [5]. For example, in psychiatry, sleep stages and mental illnesses are diagnosed by their characteristic EEG waves [6].

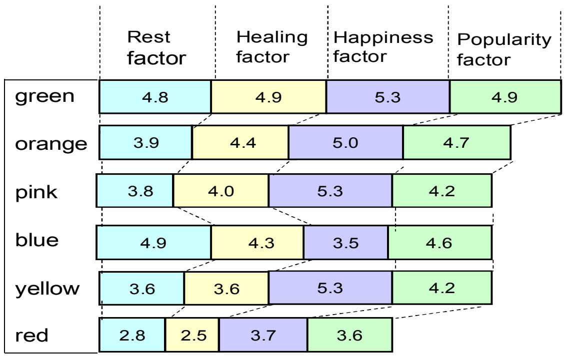

Figure 1. Mental Health factors of visible light color

For state conditions such as healing, it is well known that alpha and beta waves provide important evidence for medical diagnosis. The existence of alpha waves in an EEG has a strong correlation with the relaxation of a subject with eyes closed. Conversely, the existence of beta waves in an EEG has a strong correlation with a state of active concentration or excitement.

In our previous study, to estimate the healing effect of certain kinds of stimulation, we considered the presence of alpha waves in EEG records.

In our experiment, the international 10-20 system was used for EEG measurements and referential derivation was applied. The EEG data at electrode O1, where alpha waves are best observed, were used. The EEG data were sampled (at a rate of 200 Hz) with a 0.64 sec (128-point) Hamming window. Its logarithmic power spectrum coefficients were calculated by FFT and the spectral range of alpha waves (from 8 to 13 Hz) was used for analysis.

We found that the healing effect on subjects is larger when a specific combination of stimuli (lavender aroma, blue light, and music) is given to patients than when only a single stimulus is given. However, we found that green LED lighting also gives a significant human body healing effect when the brightness of the LED light was controlled to make the brightness vary inversely with frequency [7][8].

In this experiment, we compare the power of alpha waves recorded from subjects after irradiation for five minutes by green LED light with a continuous spectrum and also with light having a 1/f fluctuation. The number of subjects was 4. The values of the power of alpha waves of all subjects were averaged. Figure 2 shows the results of this experiment. The power in the alpha wave band after irradiation by green LED light with 1/f fluctuation was higher than that after exposure to a continuous spectrum. In particular, the power for two subjects showed a significant difference ($p=0.05$). This result indicates that green light with a 1/f fluctuation has a healing effect.

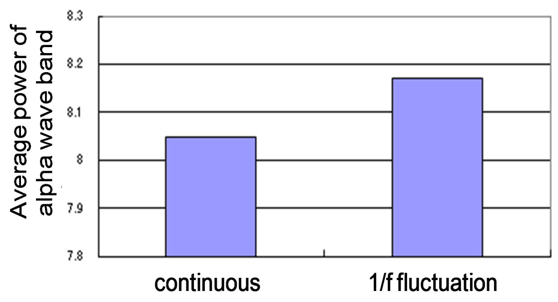

Figure 2. The average power in the alpha wave band versus the light exposure spectrum

We verified the corresponding effects using physiological and psychological techniques [2].

It will be possible to realize the above-mentioned healing environments inside a patient's premises via a network.

Currently, there is no objective method of evaluating the above-mentioned environments, although they will in the future be used to deliver "telemedicine". Therefore, healing communication systems that are reliable enough for use in commercial applications are yet to be developed. Our final goal is to realize healing environments and safe telemedicine services by making use of sensor networking technology and biomedical sensor node communication technology with a high level of security.

We proposed a mutual authentication method which relies on GPS (Global Positioning System) information [9][10]. Figure 3 shows the mechanism of this mutual authentication method. In conventional telemedicine, the hospital authenticates patients, to prevent leakage of personal information, but the patient does not authenticate the hospital.

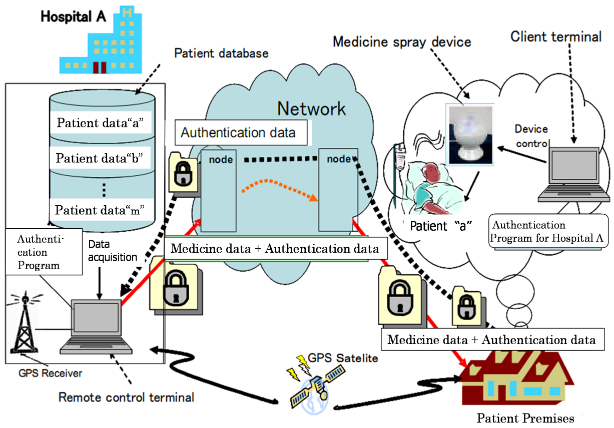

Figure 3. Proposed mutual authentication mechanism

Today, the patient should be able to authenticate the hospital because there is the potential of spoofing of the hospital and because authentication is necessary to ensure the security of network services. A reliable mutual authentication method can be realized by effectively combining GPS information, passwords, IDs, etc.

A secure remote diagnosis service can also be achieved by not only using a combination of GPS information and the user's and the hospital's passwords but also encrypting ID codes that are transmitted over the network. Today, the level of location precision determined by GPS is in the range of several centimeters, so GPS can be reliably used for authentication.

Examples of authentication data being exchanged by the patient and the hospital before any treatment commences are shown in Figure 3. Information sent by the hospital contains authentication information as well as the prescription. Information sent from the client contains authentication information and data about the patient's current physical condition, such as blood pressure and temperature.

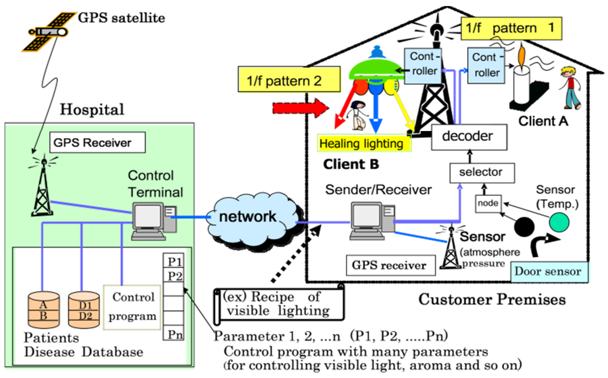

Figure 4. Telemedicine system for healing environment

Figure 4 shows the concept of the proposed telemedicine system for realizing a more comfortable and effective healing environment. The remote hospital has the individual patient's database and a disease-related database, which provide the appropriate recipe to control the visible light brightness, aroma type, 1/f fluctuation type, and so on, in the remotely located patient's premises. These various kinds of stimulation need to be properly selected, to take account of the patient's current physical and mental condition. The control program has many different parameters for controlling the visible light. Environment sensor information such as the temperature, atmospheric pressure, and humidity will also be utilized to assess the patient's environmental conditions accurately. The patient's physical parameters such as brain waveforms, blood pressure, heart rate, and so on can be fed back by making use of a high speed network and the sensor network. By utilizing this network mechanism, the optimum combination of stimuli can be appropriately and dynamically provided in real time, improving the patient's physical and mental condition. The hospital, or other organizations related to various medical treatments, and the patient's premises can be mutually authenticated by the use of detailed information provided by a GPS information system to ensure security.

Even if the patient is being transported in an ambulance, the patient and the hospital can authenticate each other using a GPS database. Furthermore, mutual authentication can be applied between hospitals to establish a communication link and to prevent a situation in which two hospitals might both provide the same prescription to the same patient.

The remainder of this paper is structured as follows. In Section II, we describe the practical healing communication system which makes use of an ID-based key sharing scheme. Then, in Section III, we clarify the performance evaluation results obtained when using the proposed priority control mechanism and the sensor database efficiently using a stream cipher. Finally, we provide our conclusions from these studies in Section IV.

Encrypted data

AιAuthentication Header			Prescription
Hospital Identification ID	Hospital password	Hospital GPS Information	Recipe data for prescription
ID 1	Pass1s1	(X_1,Y_1,Z_1)	R_1
⋮	⋮	⋮	⋮
ID N	PassN	(X_N,Y_N,Z_N)	R_N

Encrypted data

Authentication Header			Patient biomedical data
Patient Identification ID	Patient password	Patient premises (GPS information)	Blood pressure, pulse wave, temperature
id 1	p 1	(x_1,y_1,z_1)	Data 1
⋮	⋮	⋮	⋮
id m	p m	(x_m, y_m, z_m)	Data m

Figure 5. Examples of authentication data formats

II. HEALING HEALTHCARE COMMUNICATION SYSTEM

If the proposed system is to be used for medical treatment, the following requirements must be met. First, the health care organization, for example a hospital, must prevent the leakage of personal information. Second, from the user's standpoint, reliable mutual authentication is necessary. The proposed system achieves reliable authentication using the position information from a GPS system in addition to a user ID and a password.

1) Configuration of the proposed system

The configuration of the proposed system is shown in Figure 6. The system is composed of a Monitoring Centre and a Remote Node. The former is equipped with a GPS device, a sensor database and a recipe database. The Remote Node is equipped with a sensor node, which includes a GPS device and a pulse wave sensor which measures the user's biometric data, and a Gateway Node, which controls actuators such as an aroma generator and a light wavelength controller according to a recipe received from the Monitoring Centre, after mutual authentication has been established with the Centre.

The operational procedure applied to the proposed system is as follows.

(1) A medical specialist in both aroma and lighting accesses the Recipe Database, and selects the appropriate healing recipe data for individual users. There may be individual patient and disease-related databases for providing the appropriate recipe to select the color of visible light, aroma type and 1/f fluctuation pattern type, and to control the visible light brightness in the remotely located patient's premises. These different kinds of stimulation need to be selected appropriately based on the patient's current physical and mental condition as indicated by the biometric sensors. In addition to environmental sensor information such as ambient temperature, atmospheric pressure, light level and humidity, data obtained from body sensors and relating to the patient's body, such as pulse rate, blood pressure, respiratory rate and so on, will be utilized to accurately determine the patient's current physical and mental state. Data relating to the patient's current physical condition, such as brain waveform, blood pressure, heart rate, etc., can be fed back via the high-speed network, and the optimum combination of stimulations can be appropriately and dynamically provided in real time, so that the patient's physical and mental condition will improve.

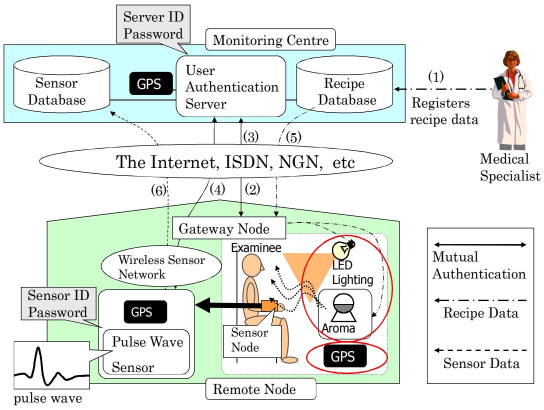

Figure 6. Configuration of the Proposed Healing Communication System

An example of the recipe data format is shown in TABLE I.

TABLE I. HEALING RECIPE DATA FORMAT

Aroma Recipe (1ch~6ch)	Brightness (Red color)	Brightness (Green color)	Brightness (Blue color)	1/f characteristics (pattern)

(2) A client (the Remote Node) accesses the Monitoring Centre via the Gateway Node, and sends his or her user ID, password, and geographical position as determined by GPS for mutual authentication.
(3) The Monitoring Centre sends the location information registered beforehand to the Remote Node for mutual authentication.
(4) The Monitoring Centre and the Sensor node execute mutual authentication as follows. The Monitoring Centre requests the Sensor node to send information about its geographical position, and determines whether the Sensor node's position coincides approximately with the position of the Remote Node.
(5) After mutual authentication has been established, the Monitoring Centre sends the appropriate healing recipe data via the Gateway Node.
(6) The Sensor node sends measured biometric sensor data to the Sensor Database via the Monitoring Centre. The sensor data acquired in real time are used to evaluate the patient's psychological and physiological status after the healing recipe has been applied. The results are fed back to the medical specialist to further improve the patient's condition.

2) Mutual Authentication and Key Sharing Scheme.

We propose the use of an ID-based key sharing mechanism [11] for mutual authentication. Several key sharing schemes have been developed for authentication over the Internet, such as Diffie-Hellman, and Kerberos [12].

These are not suitable for sensor network applications because the processing power of a sensor node is too limited. Another conceivable scheme for mutual authentication is a key pre-distribution scheme, in which each sensor node receives common keys in advance from other sensor nodes. Usually, we have been considering a sort of client-server type system, with each sensor connected to the hospital. However, when we make use of multiple sensors in a ubiquitous environment sensor network, the different sensors need to exchange information with each other. However, the number of keys that each node must manage increases in proportion to the number of sensor nodes, while the memory capacity of each node is limited. In addition, this scheme lacks extendibility. To avoid the problems of these schemes, we propose an ID-based key sharing scheme. This allows prompt mutual authentication because this key sharing scheme does not require interactive key information exchange operations. The configuration of the ID-based key sharing scheme is shown in Figure 7.

The key sharing procedure uses secret private keys. The key management server generates secret private keys based on users' ID information and transmits the keys to each node in advance over a predetermined secure communication path. The procedure for sharing keys between communication nodes is handled only by the nodes which have been authenticated by the key management server in advance.

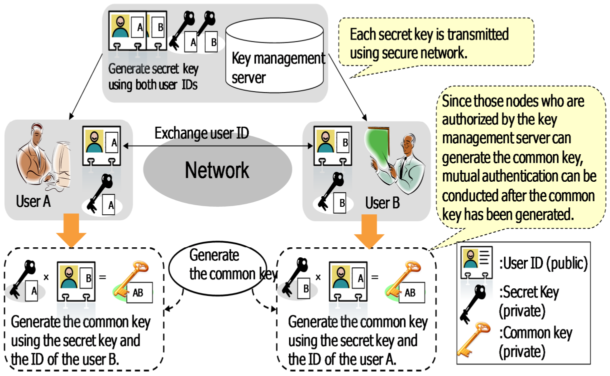

Figure 7. ID-based key sharing scheme configuration

This procedure is the basis of the proposed mutual authentication mechanism in which communication nodes share common keys. The common keys generated by one pair of communicating nodes differ from those generated by another pair of communicating nodes. This provides for network scalability. The mathematical background of the proposed common key generation method is as follows [13].

Let D be a private symmetrical matrix generated by the key management server.

The matrix I is a row vector, containing values which are equivalent to the ID information of each node. G is calculated based on matrix D and matrix I. The column vector G corresponds to the secret keys of the ID information. Matrix I is a row vector containing a set of values which are equivalent to the ID information of each node.

$$
\begin{aligned}
G &= (D \cdot I)^T \\
G \cdot I &= (D \cdot I)^T \cdot I = I^T \cdot D^T \cdot I = I^T \cdot D \cdot I \\
&= (G \cdot I)^T
\end{aligned}
$$

Therefore, a symmetrical matrix can be obtained. An example of the common key generation is provided below, based on this calculation.

Two nodes, say Node A and Node B, exchange their IDs, Ia and Ib, with each other and generate keys, Kab and Kba, using the exchanged IDs and their private keys.

$$
\begin{aligned}
K_{ab} &= G_a \cdot I_b = (D \cdot I_a)^T \cdot I_b = I_a{}^T \cdot D^T \cdot I_b = I_a{}^T \cdot D \cdot I_b \\
&= I_a{}^T \cdot (G_b)^T = (G_b \cdot I_a)^T = G_b \cdot I_a = K_{ba}
\end{aligned}
$$

These expressions show that Kab is equal to Kba, and that the procedure for key sharing between two nodes can be executed. We can use this mutual authentication procedure to implement a sensor network that deploys the priority control mechanism on the proposed sensor network itself.

The security of the proposed scheme can be enhanced by increasing the length of the ID code in addition to adopting the elliptic curve method [14]. An appropriate level of security can be established to suit the required security policy and the sensor network environment.

3) Configuration of the Intelligent Sensor

We used a Sun SPOT from Sun Microsystems for the sensor node. Since its control program can be written in the Java language, it is easy to develop a prototype system. The secret key used by the user authentication server for mutual authentication is generated at each node. Using a pulse wave sensor and a GPS device, the Sensor Node can measure biometric data and acquire its geographical position information. The pulse wave sensor measures changes in the amount of hemoglobin in the blood. The pulse wave sensor circuit used is shown in Figure 8. The pulse wave sensor irradiates the tip of a finger with light, such as an infrared beam. The light falling on hemoglobin is absorbed, and the light falling on the sensor is light that has not fallen on hemoglobin, but is reflected from the blood-vessel/bone boundary. The light caught by the optical sensor is converted into an electrical current. After processing the corresponding electrical current through a differential amplifier it will be transmitted via wireless communication. The electrical current can be converted into blood pressure data using a current-voltage converter, a low-pass filter, and a differential amplifier. Changes in the amount of hemoglobin can thus be acquired.

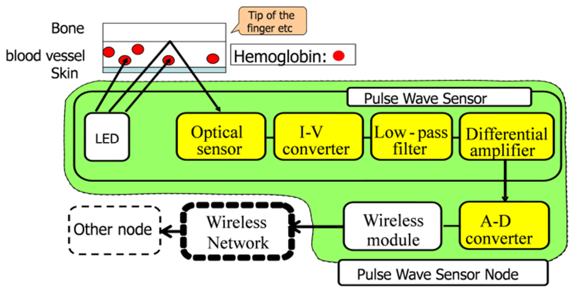

Figure 8. Pulse wave sensor circuit

4) Healing Recipes and the Evaluation of Physiological and Psychological Conditions

The proposed system uses pulse wave data to determine the specific healing recipe to be prescribed. Pulse wave data are useful for determining the condition of the autonomic nervous system. The autonomic nervous system is divided into the sympathetic nervous system, which represents the active state of a living body, and the parasympathetic nervous system, which represents the resting or relaxed state of the body. Since the autonomic nervous system directly affects the physiological and psychological conditions of a person, the psychological condition of a person can be determined objectively by measuring and analyzing his or her biometric data [15].

Some examples of the influence of the autonomic nervous system on the human body are shown in Table II.

TABLE II. PHYSIOLOGICAL TENDENCIES RELEVANT TO THE AUTONOMIC NERVOUS SYSTEM

		Sympathetic Nervous Activity	Parasympathetic Nervous Activity
Physiological Tendency	Heart	Increase of blood pressure, heart rate	Decrease of blood pressure, heart rate
	Blood Vessel	Contraction	Enlargement

The autonomic nervous system evaluation method is as follows. In the autonomic nervous system function, the sympathetic nerve is in a dominant state if the LF/HF value is high, and the parasympathetic nerve is in a dominant state if LF/HF value is low. Generally, for the analysis of frequency characteristics, the maximum entropy method (MEM) is used. MEM is a parametric technique and is used for estimating the power spectrum by making use of a linear prediction model. Even if only a few data points are measured, an adequate spectrum analysis with high resolution can be obtained by making use of MEM. That is the reason why MEM is suitable for analyzing the frequency spectrum just by using the minimum amount of data. In our experimental system, we used MEM to analyze the autonomic system function [16]. The autonomic nervous system function is measured by the acceleration pulse wave [17].

The acceleration pulse wave can be obtained using the second differential of the finger plethysmograms. The waveform pattern of the acceleration pulse is shown in Figure 10.

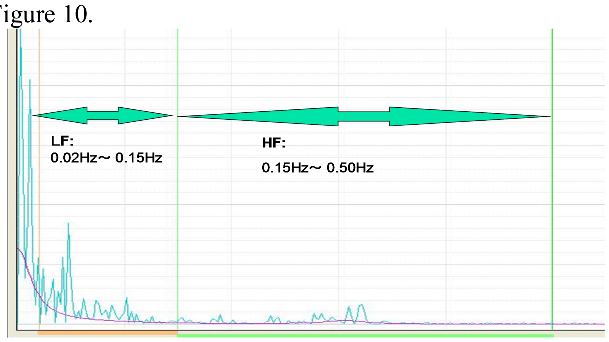

Figure 9. LF and HF bands of the autonomic nervous system

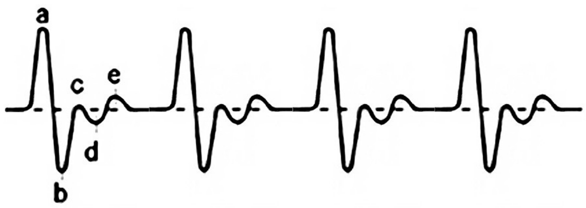

Figure 10. Waveform of acceleration pulse wave

The acceleration pulse wave form has the wave features a, b, c, d and e in Figure 10. Here, the value of b/a generally increases with the age of the subject, while d/a decreases with age. The "waveform index" is the degree of aging of the blood vessel as measured by the change of the shape of the waves [18]. Generally, the b wave feature becomes shallow, and the d wave feature becomes deep with aging. Therefore, we can calculate the change of the waveform index due to aging by making use of the values d/a and b/a. Then, an arteriosclerosis degree (score) that describes the blood vessel can be obtained.

The calculated score expresses the degree of hardening of the artery wall and the degree of functional strain. A score of under 20 indicates that the degree of blood vessel aging is low, and the flexibility of the blood vessels is much higher than average. A score from 20 to 34 also indicates a low blood vessel aging rate, and the flexibility of the blood vessels is notably higher than average. A score of 35-39 indicates a low blood vessel aging rate, and the flexibility of the blood vessels is slightly higher than average. A score of 40-59 indicates that the flexibility of the blood vessels is average. A score of 60-64 indicates that the blood vessel aging rate is a little faster than average, and the flexibility of the blood vessels is slightly lower than average. A score of 65-69 indicates a rapid blood vessel aging rate, and the flexibility of the blood vessels is much lower than average. A score of over 70 indicates a very rapid blood vessel aging rate, and the flexibility of the blood vessels is much lower than the average [19]. The arteriosclerosis degree (score) evaluation formula is as follows.

Score = 50 + 10 (X1 mean - X1 peak value) / X1 standard deviation

*X1 (Waveform index) = d/a - b/a

For evaluating the stress value in a human being it is often effective to measure the alpha amylase activity in saliva.

With the alpha amylase activity in saliva, the degree of stress can be measured as a biomarker. The density of the alpha amylase, which is one of the digestive enzymes, changes with any active change in the sympathetic system. There are several biomarkers that measure the degree of stress, and cortisol is one of the typical ones. In the case of alpha amylase, the time interval from the stimulation to the secretion is shorter than in case of the cortisol [20].

Therefore, in a situation with significant mental stress (acute stress) a quantitative analysis of the degree of stress is possible by measuring the alpha amylase in the saliva.

There are several advantages to assessing the stress value by measuring the alpha amylase activity in saliva. The procedure is:

(1) Non-invasive
(2) Rapid measurement
(3) Rapid analysis
(4) Simple and portable measurement
(5) Low cost

It should be noted that the alpha amylase activity in saliva differs for each individual and also changes during the day, so measuring the degree of difference before and after the stimulation is essential [21].

A spectrum analysis of historical data of pulse cycles is known to be an effective method for determining the state of the autonomic nervous system [22]. Past studies indicate that the fluctuation of pulse cycles of low-frequency components (LF: 0.02 to 0.15 Hz) is related to the strength of the activities of both the sympathetic nervous system and the parasympathetic nervous system. On the other hand, high-frequency components (HF: 0.15 to 0.50 Hz) are related to the strength of the activity of only the parasympathetic nervous system.

5) Priority Control Mechanism used in the Sensor Network

There have been no in-depth studies on how to achieve real-time processing for the collection of sensor data, such as pulse data, as is required of a sensor network. To achieve real-time performance in addition to ensuring network security, we propose to use the above-described ID-based key sharing scheme for implementing priority control and for ensuring security in the sensor network. We have evaluated the network performance of the proposed priority control mechanism while taking various elements of the network environment into consideration.

6) Functional Requirements for the proposed system

The functions required for priority control in the sensor network are described below.

To protect the sensor data from eavesdropping and tampering, the data must be encrypted in an effective manner. For encrypting sensor data, it is necessary to ensure security of key sharing between the communication nodes using mutual authentication between the communication sensor nodes. The proposed encryption mechanism implemented on an experimental system is shown in Figure 11.

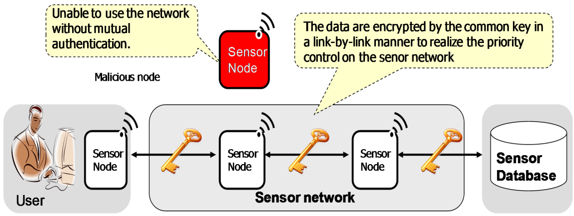

Figure 11. Encryption mechanism of the experimental system

Figure 12 shows an example of the format of the encrypted transmission data, which provides adequate confidentiality and security. To achieve mutual authentication between the sensor node and the sensor database, the sensor data must first be promptly encrypted over the end-to-end path using common keys. In addition, the corresponding data can also be encrypted in a link by a link procedure using different common keys. Figure 13 shows an example of the format of encrypted transmission data between sensor node 3 and the sensor database, which is the final destination. The encrypted data can finally be decrypted in the sensor database using both the key which is the common key with sensor node 3 for link-by-link communication and the different common key used between sensor node 1 and the sensor database.

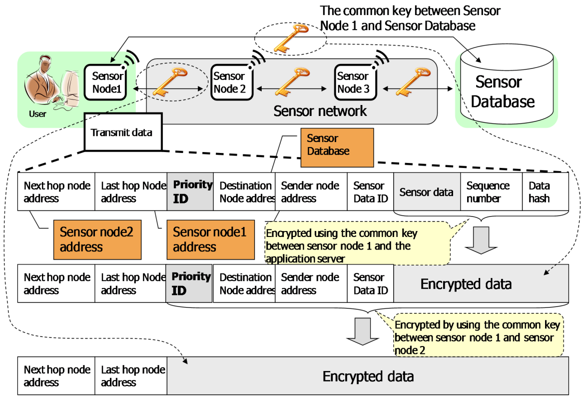

Figure 12. Composition of the transmitted sensor data

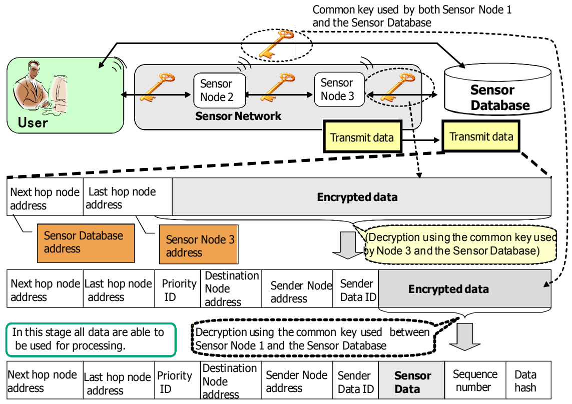

Figure 13. Decryption of the encrypted transmitted sensor data

III. EVALUATION OF SYSTEM PERFORMANCE

A. Evaluation of the proposed priority control mechanism

1) Conditions of the experiment

Traditionally, sensor network applications for environmental monitoring, precision agriculture, and so on, do not require a prompt response, and do not have to handle operations with different QoS. However, real-time communication services that make use of wireless sensor networks are becoming indispensable, in particular for healthcare monitoring [23]. Against this background, the priority queue monitoring mechanism and the corresponding communication protocols have been investigated [24].

However, most of the results obtained are based on simulations, and the substantial tradeoffs between the different QoS requirements have not yet been established. Consequently, we first examined the ability of the commercially available and economical Sun SPOT sensor nodes to send data while making use of the priority control mechanism, and ascertained the feasibility of realizing a queuing control system easily. Communications in sensor networks can be broadly divided into two categories: local coordination and sensor-based communication. Before sending information to the transit sensor node, sensors within a local area should cooperate in order to aggregate data and ensure reliable data transmission. However, in this paper, we disregarded these extra conditions in order to analyze the transmission effectiveness of a sensor node that employs the priority control mechanism. Based on these considerations, we evaluated the delay time performance under the following conditions.

The basic configuration of the experimental system in which the proposed priority control mechanism was implemented is shown in Figure 14. A Sun SPOT node from Sun Microsystems was used in the experiment.

In the experimental system, sensor data were transmitted from two sensor nodes to the sensor database, which collected sensor data via two intermediate hop nodes in order to ascertain the performance of the priority control mechanism. We set two levels of priority in all relay sensor nodes: high and low priority levels. For example, a high priority level was assigned to pulse wave data to reduce both delay and delay fluctuation. A low priority level was assigned to temperature data because neither its transmission delay nor its delay fluctuation needs to be small. A different data buffer was used for each priority level.

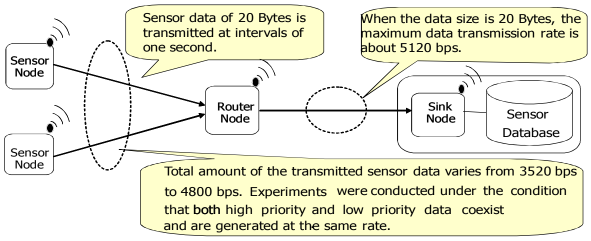

Figure 14. Configuration of the Experimental System

A photo of the actual sensor network is shown in Figure 15. Here several Sun SPOT nodes are used for a Sensor Node, a Router Node (Transit node), and a Sink Node. According to the specifications of the Sun SPOT it adopts the IEEE802.15.4 wireless standard (2.4-2.4835 GHz), and can transmit packet data using QPSK modulation technology.

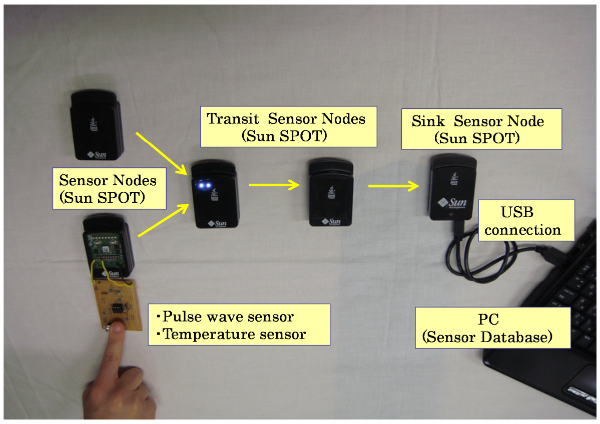

Figure 15. Photo of the experimental sensor network system

The size of each packet of sensor data was 20 bytes (8 bit quantization and sampled at 20 Hz). The sensor data were transmitted once every second. The maximum transmission speed of data transmitted over two hops was about 5120 bps as dictated by the specification of Sun SPOT. Two nodes can communicate by transmitting sensor data at speeds ranging from 3520 bps to 4800 bps.

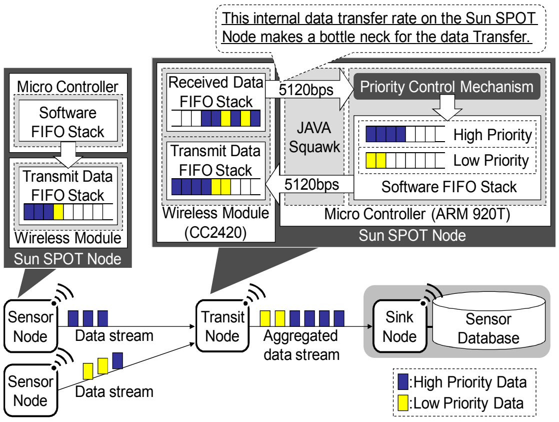

Figure 16. The mechanism providing priority control

In the sink nodes and transit nodes, we adopted a simple priority control mechanism in which the operation of switching to the lower priority output buffer can be done only after the higher priority output buffer has become empty, although the input data are stored continuously inside the buffer in a FIFO manner. Operation will be switched back again to the higher priority output buffer if higher priority data are stored in the FIFO buffer, even when the low priority data are being output, in order to ensure real-time performance of the higher priority data, as shown in Figure 16.

In this case, we found that the communication bottleneck among the Sun SPOT nodes was due to a restriction in the internal data transfer rate between the wireless module (C2420) and microcontroller (ARM920T), which are controlled by Java squawk in the Sun SPOT node. When the data size is 20 bytes, the maximum data transmission speed is measured as 5120 bps. However, even if the data transmission rate exceeds the specified limit for a short time, the data can be transmitted provided that the excess amount of transmitted data does not exceed the Sun SPOT FIFO buffer size.

The entire queuing control was implemented in the software. The Sun SPOT adopts the 180 MHz 32 bit micro-processor (ARM 920T) and can make use of 512 Kbyte RAM for programs and the FIFO control buffer.

The processing mechanism of the proposed priority control system is mainly composed of two independent processes, as shown in Figure 16.

Figure 17 (a) shows the process that registers the received sensor data in the appropriate priority control stack by identifying the priority of each data item by checking the Priority ID of the data frame, which was previously described in Figure 12.

Fig.17 (b) shows the process flow that transmits the data stored in each control stack according to its inherent priority level. The high priority control stack is identified first, and all high priority data are transmitted until no high priority data are left in the stack. In the next step, the process checks the low priority control stack and sends the low priority data if any exist. After finishing this procedure, the data transmission process reverts to the first operation and these operations are repeated continuously.

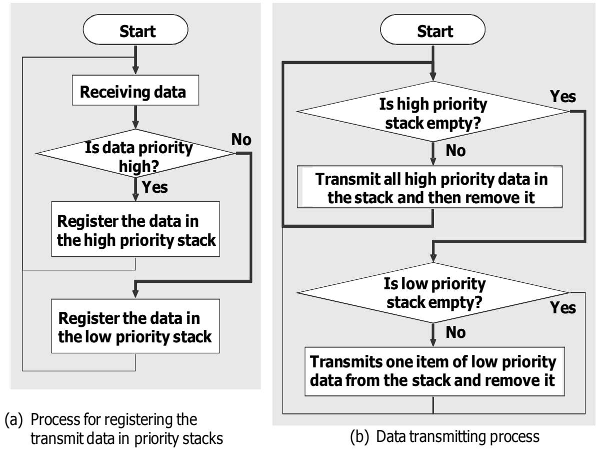

Figure 17. Processing flow of the proposed priority control mechanism

We carried out two experiments, one using priority control and the other without it. The transmission rates of both the high priority and the low priority data were the same. We measured the average data-receiving interval at the sink node, the average delay time and the average delay fluctuation of the transmitted sensor data. Sensor data were transmitted continuously for 12 seconds. Considering that about 1 second is required for call set up, we used data measured between 2 seconds and 12 seconds.

2) Experiment Results

We classified the experimental results into two parts: first where the total transmitted data rate was less than 5120 bps, and second where it was more than 5120 bps. As shown in Figure 18, the transmitted data rate of 5120 bps corresponds to the Sun SPOT internal transfer rate, so the maximum transmission throughput cannot exceed this limit. However, if the period of continuous data transmission is sufficiently small, then the use of the internal buffer may allow the transmission to be successful in some circumstances. Figure 19 shows the experimental results of the interval between receiving one packet and the next, the average delay time and the average delay fluctuation.

The average receiving interval was approximately one second, irrespective of whether priority control was used or not. We confirmed that no errors were observed at the destination sink node. The average delay and the average delay fluctuation increased in proportion to the total amount of sensor data transmitted from all nodes. In addition, the difference in the average delay and the average delay fluctuation between high priority data and low priority data became greater as the total amount of sensor data increased. The reason is due to the simple priority queuing control mechanism, implemented by the software as mentioned above. From these experiments, we obtained the following conclusions when we made use of Sun SPOTs as ad hoc sensor nodes.

(1) Restriction of the total volume of sensor data and aggregation of this data are very important in ensuring adequate real-time performance.

(2) The software control priority queuing scheme is very efficient in minimizing both the delay time and the delay variation simultaneously.

(3) When the total rate of sensor data transmitted is less than the pre-assigned maximum transmission speed of 5120 bps, the priority control mechanism is effective.

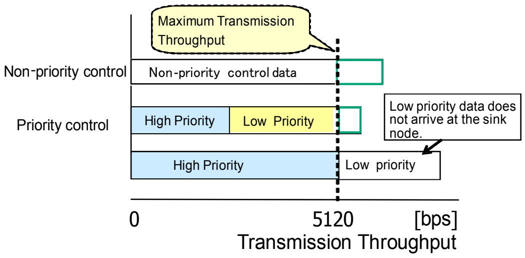

Figure 18. Priority and non-priority control concepts

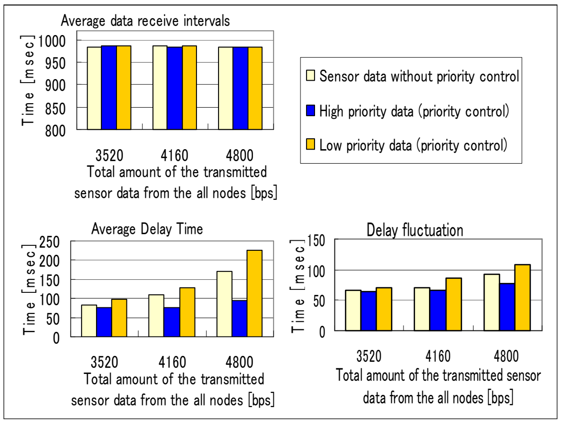

Figure 19. The average delay and the average delay fluctuation

B. *Evaluation of sensor database performance*

To ensure the confidentiality of personal information, such as measured sensor data, the sensor data were encrypted, and a secure method was used for accessing the sensor database. We evaluated quantitatively the real-time performance of the Sensor Database [25][26].

1) Configuration of the Database

The proposed database consisted of a Sensor Database, which stored encrypted data, and a Monitoring Centre, which stored encryption keys. A stream encryption method that requires low processing power was employed for data encryption, as shown in Figure.20.

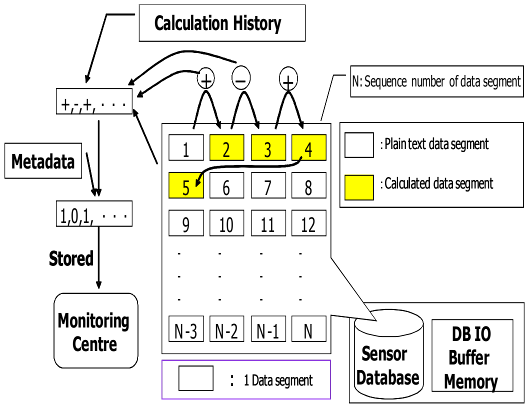

Figure 20. Proposed Encryption Method

In the proposed encryption method, sensor data, which consist of a series of plain text characters, is divided into segments of a certain number of bits. The simplest case is using a string of multiple bits as one word. One of the possible reversible logical operations, such as addition, subtraction or exclusive-OR operation, is applied to combine the binary values (the string of bits of value 0 or 1) of one word with those of the following word. The result of the calculation is then stored in the location of the second of these words and the sequence then repeated moving on to the next word. A different reversible operation may be selected for each word in the sequence, depending on the encryption key information stored in the authentication server.

The result of successive execution of a reversible operation on the original binary data can be replaced by the original data in each word. Although the above example uses a "word" as the minimum unit of encryption, any size can be selected as the minimum size of the bit length. Efficient encryption and decryption can be achieved by selecting an appropriate bit length execution. The information history associated with the sequence of selected reversible operations mentioned above is provided as metadata managed by the monitoring center, and the actual operations were executed on the test data for the evaluation.

For example, in encrypting an item of data consisting of three consecutive words, suppose that "1" corresponds to a binary addition, and "0" to a binary subtraction. If the metadata is "101" for example, a binary addition is executed for the first and second words, and the result is stored in the second word. After that, a binary subtraction is executed on the second and third words, and the result is stored in the third word. After that, a binary addition is executed on the third and fourth words, and the result is stored in the fourth word. In this way, metadata is used until an operation is executed on the last word in the sequence. To extract plain text from data encrypted as above, in the absence of the metadata, the number of random attempts required could be as high as 2 to the power of the number of words in the worst case. It is therefore difficult for a malicious third party to decrypt the data. In addition, since the size of data segment can be changed regularly, it can be made even more difficult to decrypt the data.

2) Experimental Methodology

In accordance with the proposed method, the processing power (including encryption) of the sensor database in the monitoring center was evaluated in line with the following principle. The parameters used in the experiments were the number of simultaneous connections to the sensor database (users), and the capacity of the I/O buffer used for direct access to the sensor database via a network interface. The rate at which data was stored in the sensor database was used as the measurement criterion. We assumed quasi-client terminals that contained sensor nodes.

We further assumed that 100 to 300 such terminals accessed the database simultaneously. When a sensor node accessed the monitoring center, it conducted mutual authentication using the user authentication server.

After the sensor node had received the healing recipe data, it transmitted a series of items of sensor data composed of pulse wave data (1kB, 8-bit data with a sampling rate of 1 kHz, for example) to the Sensor Database at intervals of one second for several seconds. In order to examine the real-time performance of the Sensor Database, we conducted evaluation experiments to determine the conditions under which all sensor data from 300 terminals were successfully stored, the operation being completed as quickly as possible. The criterion used was the rate at which the data transmitted continuously from each sensor node was stored. Normally, a continuous measurement of 5 minutes is sufficient to infer whether a person is in a relaxed state or not, so experiments were conducted for 5 minutes.

3) Results of the Experiment

Figure 21 shows the evaluation results. Parameter A used in this experiment is the total number of sensor nodes that accessed the database. This number is equal to the number of users. Parameter B is the capacity of the I/O buffer in the database. Experiments were conducted using these as the parameters. The results are shown below.

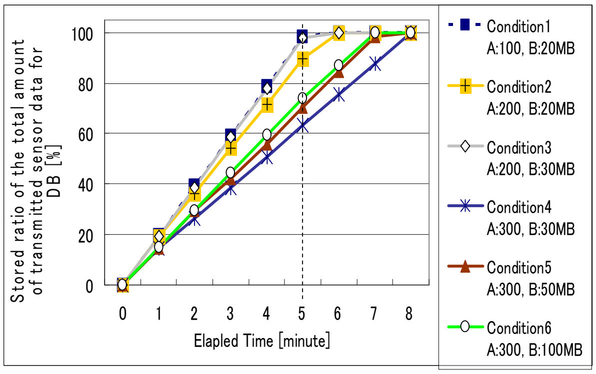

Figure 21. Results of the System Performance Evaluation

(1) Only for Conditions 1 and 3 did the amount of sensor data stored reach 100% of the total transmitted within 5 minutes of the start of the experiment. However, if the number of users is greater than that applied in the above conditions, it will be necessary to adjust the capacity of the I/O buffer in the database.

(2) The experiments using Conditions 4 to 6 revealed that for a fixed number of users, the effect of increasing the I/O buffer capacity for collecting data beyond 30 MB is relatively small.

(3) The results of these experiments identified the size of the DB I/O memory required to implement the system. It was found that if the number of users is in the range of 100 to 200, a 20 MB I/O buffer in the database was sufficient, even when sensor data collecting services were provided simultaneously to all sensors.

This result can be used to build a sensor network that has real-time performance sufficient for applications for future telemedicine.

IV. CONCLUSION

We have proposed a health care communication system that uses GPS technology and ensures confidentiality of personal information. The system can be used for telemedicine to improve a user's QoL.

In addition, we have proposed a priority control mechanism that uses an ID-based key sharing scheme in the proposed sensor network both to achieve adequate real-time performance and to ensure security of sensor data. Evaluation using an experimental system confirmed that the average delay time and the delay fluctuation of delay-sensitive data were reduced effectively when such data were transmitted with high priority compared with the method of non-priority control. We have also confirmed that a sensor database that uses an efficient encryption mechanism for ensuring security and maintaining the confidentiality of users' sensor data shows excellent real-time performance.

In the future, we will collaborate with medical hospitals and social welfare centers to realize practical healthcare applications.

ACKNOWLEDGMENT

This study was supported in part by Tokyo Denki University Research Center for Advanced Technologies (Q07J-13) and a grant of Strategic Research Foundation Grant-aided Project for Private Universities from Ministry of Education, Culture, Support, Science, and Technology, Japan (MEXT) (2007-2011)

REFERENCES

[1] N. Miyaho, T. Nakamura, N. Konno, and T. Shimada, "Sensor Network Management for Healthcare Applications", ICSNC2010, ICSNC1 SENET I, pp.14-20 Aug., 2010.

[2] N. Miyaho, N. Konno, and T. Shimada, "Study on Healing Environment using Green, Blue and Red LED and Aroma", J. Light & Vis. Env. Vol. 32, No. 2, pp. 47-52, May, 2008.

[3] N. Miyaho, N. Konno, and T. Shimada, "Study on Healing Environment using green, blue, red LED and aroma" (Invited), First International Conference on White LEDs and Solid State Lighting (White LEDs-07), Nov., pp. 307-310, 2007

[4] N. Konno, "The psychological effect of color lighting with regard to mental health", Bulletin of Tokyo Denki University, Arts and Sciences, No.6, pp. 55-62, 2008. (In Japanese).

[5] Mary A. Carskadon and Allan Rechtschaffen, "Monitoring and Staging Human Sleep", Principles and Practice of Sleep Medicine (Fourth Edition), pp. 1359-1377, 2005.

[6] Alex L. van Bemmel, "The link between sleep and depression: The effects of antidepressants on EEG sleep", Journal of Psychosomatic Research, vol.42, Issue 6, pp.555-564, June, 1997.

[7] T. Nakamura, T. Shimada, N. Konno, and N. Miyaho, "Study on Healing Environment Conditions by making use of 1/f Fluctuation", IEICE National Convention, BS-12-6, 2007.

[8] N.Shinzawa, M. Komazaki, and N. Miyaho, "Study on Applications of Using Visible Light Communications", IEICE National Convention 2008 A05_004, 2008. (In Japanese)

[9] S. Kurokawa, Y. Iwaki, and N. Miyaho, "Study on the distributed data sharing mechanism by making use of the mutual authentication and meta database", IEEE APCC2007, pp. 215-218, 2007.

[10] S. Kurokawa and N. Miyaho, "Study on a distributed data sharing mechanism with mutual authentication", IEICE IN Technical Paper IN 2006, No.214, pp 203-206, 2006. (In Japanese)

[11] S. Tsujii, K. Araki, and T. Sekine, "A New Scheme of Non-Interactive ID-based Key Sharing with Explosively High Degree of Separability (Second Version)", Tokyo Institute of Technology Technical Report, Dept. of CS, TR-0020, 1993. *(In* Japanese)

[12] http://www.rfc-editor.org/rfc/rfc1510.txt (Sep.1993), <2011.7.10>.

[13] R. Blom, "Non-public key distribution," Proceeding of Crypto'82, pp. 231-236, 1982.

[14] O. Kiyoshi, R. Ryuichi, and M. Kasahara, "Notes on ID-based Key Sharing System over Elliptic Curve", ISEC99-57. pp.37-42, 1999. *(In* Japanese)

[15] J. W. Severinghaus, and J. F. Kelleher, "Recent developments in pulse oximetry", Anesthesiology 76 1018-1038, 1992.

[16] N. Konno, "Influence that Irradiation of color Lighting exerts on Autonomic Nervous system Function", Bulletin of Tokyo Denki University, Arts and Sciences, No.5, pp. 67-74, 2007.(In Japanese)

[17] H. Takada, K. Okino, and Y. Kiwa, "An Evaluation Method for Heart Rate Variability, by using Acceleration Plethysmography", Health & Prom., 31(4), pp. 547-551, 2004.

[18] H. Takada, K. Washino, and H. Iwata, "Accelaration plethysmography to evaluate aging effect in cardiovascular system.Using new criteria of four wave patterns", Medical progress through technology 21(4), pp. 205-210, 1996-1997.

[19] H. Takada, "Proposal of Aging Score Method by Acceleration Plethysmography", Health Evaluation and Promotion, 29(5), pp. 855-861, 2002.

[20] Weibel L, "Methodological guidelines for the use of salivary cortisol as biological marker of stress", Presse Med 32., pp. 845-851, 2003.

[21] M. Yamaguchi, M. Deguchi, J. Wakasugi, S. Ono, N. Takai, T. Hihashi, and Y. Mizuno, "Hand–held monitor of sympathetic nervous system using salivary amylase activity and its validation by driver fatique assesment", Biosens Bioelectron, 21, pp. 1007-1014, 2006.

[22] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology: "Heart rate variability: standards of measurement, physiological interpretation and clinical use", Circulation, Vol. 93, pp. 1043-1065, 1996.

[23] T.Bratan and M.Clarke, "Optimum design of Remote Patient Monitoring System", 28th International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'06, pp.6465-6468, 2006.

[24] M. Basta, R.Severino and M. Alves, "Supporting different QoS levels in Multiple-Cluster WSNs", Proceedings of 10th Portugeuse Thematic Network on Mobile Communications Workshhop (RTCM), Porto, Portugal, June, 2009.

[25] T. Kobayashi, S. Kurokawa, N. Konno, and N. Miyaho, "Study on Database Server Performance Evaluation for Realizing Aroma Information Delivery", IEICE Student National Convention, p.86, 2007. (In Japanese)

[26] T. Nakamura, T. Shimada, N. Konno, and N. Miyaho, "Study on the sensor database for realizing Healing Communication System", IEICE Society Convention, BS10-25, pp.S113-114, 2009.

MoVEing Forward: Towards an Architecture and Processes for a Living Models Infrastructure

Michael Breu, Ruth Breu, Sarah Löw
Quality Engineering Research Group
Institute of Computer Science, University of Innsbruck, Austria
{*Michael.Breu,Ruth.Breu,Sarah.Loew*}*@uibk.ac.at*

***Abstract*—Development, maintenance and operation of complex IT–systems, involving various stakeholders are challenging tasks. Managing such IT–systems with model–based approaches can help to cope with this complexity. As IT–systems are changing during their lifetime, as do the models describing certain aspects of such systems. Document and model versioning repositories are the preferred infrastructure to maintain the documents and models, reflecting the evolution of the IT–system. However, there are more complex requirements to model versioning compared to classical source code or document versioning: Depending on the types of models different modelling tools may be employed and must interface to the repository. The consistency between models must be ensured, and finally, since various stakeholders are involved, changes must be propagated between models. In this paper, we analyse these requirements and present the basic architectural concepts for a Living Models infrastructure that supports the evolution of models.**

***Keywords*-model versioning; model evolution; software engineering; change based process; tool integration; automated change management.**

I. Introduction

This paper is based on an initial presentation in [1]. Model engineering is a widely accepted engineering discipline [2], and a lot of models are developed in manifold contexts in practice.

Within software projects models form a basis for (manual or automatised) development of software. In addition models may also be used in broader context, e.g., security models help to analyse and document security concepts in critical IT–systems, business process models document interleaved business processes and IT–landscape models help to manage IT–landscapes in the context of business processes and organisational structures.

Our own engineering experience in commercial projects shows that it is hard to keep models up–to–date, if no active continuous maintenance process is in force. After the initial development phase (or after a radical renovation), a system evolves step-by-step through change requests. Quite often models die slowly in such a process, i.e., they are getting gradually out–dated, since the physical systems are evolving, without further maintenance of the model.

The reasons for dying models are manifold. One major reason is that model management is in many cases centralised to a designated stakeholder. Thus only a limited set of persons can maintain it. In real life there are various stakeholders, that have the knowledge to maintain a model, but do not have the authorisation to update the model or no tool access. Also if the modelling tool supports multiple users, concurrent maintenance of a model is not always well supported. Neither do classical version management systems like subversion support adequate conflict resolution for structured data like models. Finally, models can become quite complex, dependencies between modelled elements may be intricate, which makes it difficult to maintain the consistency of a model.

Facing this problem, [3] proposes ten principles to ensure an agile and flexible way to maintain models embedded in a change–driven engineering process. In this publication we discuss the requirements for and impacts of these principles on the implementation of an Living Models infrastructure.

Since the ten principles of Living Models [3] are the major motivation for this work, and to make the presentation self–contained, we recapitulate these ten principles shortly in the following:

Persistence Models should be stored persistently and their further evolution shall be supported.

Common System View All models of a current revision should be related by a common system view. I.e., each model maintained in the repository is considered part of a common system model. Such a system model may not exist as an explicit artifact, rather it may exist as an abstraction of the maintained models, ensured by consistency rules between these models.

Information Consistency and Retrieval Based on the common system meta–model consistency rules can be defined. Also new views on the information stored in the system model should be retrievable.

Bidirectional Information Flow between Models and Code and/or the Runtime System Code should be aware of the models and there should be an information flow from the code and the runtime system back to the models in order to enable monitoring and analysis. For source code this is quite often also known as “round trip engineering”.

Close Coupling of Models and Code Changes to code artifacts should be reflected back to the model, in order to keep them consistent.

Model Element States Each model element can have a state that reflects certain aspects in its life cycle (e.g., a risk assessment may have the states *draft, under review* and *final*, or a hardware component may have the states *under acquisition, operative, faulty* and *decommissioned*).

Change and Change Propagation The state of a model element can change between one system model version and the following. A state change may trigger other changes to the model, e.g., if an risk assessment element in a *final* state depends on another element that has changed, its state may be also changed and be reset to *draft*.

Change–Driven Process The software and system development process in a Living Models environment is driven by change events, the states of the model elements and their interrelationships. These change events may either trigger further internal changes or may be forwarded to the resonsible stakeholders to react on such changes.

Stakeholder–Centric Modelling Environments The environment should involve all relevant stakeholders with different goals (and different expectations on the type and abstraction–level of the models, as e.g., an aggregated IT–landscape model, a risk model, or a database model).

Domains and Responsibilities A variety of stakeholders operates on the system model. In order to coordinate the work on the models in an organised way, there should be assigned responsibilities for each model element.

For details about the Ten Principles we would like to refer the reader to the original publication. In this paper, we outline an infrastructure that supports those principles.

The rest of this publication is structured as follows: In the next section, we will discuss related research approaches that cover aspects of the ten principles. In the main section we introduce the major concepts and requirements of an infrastructure for Living Models, its basic architecture, and discuss the most important use cases for (meta–) model versioning and state management. We report on first experiences on a case study carried out in the context of the SecureChange project. A discussion on some collaboration aspects and conflict reduction strategies follows, before we conclude with an overview of the implementation of the initial infrastructure prototype.

II. Related Approaches

The ten principles embrace research topics that are already part of intensive research such as model versioning and merging, and meta model management.

Model repositories with versioning support are a major topic of academic and industrial research projects as, e.g., ModelBus [4], or AMOR [5] show. Both projects try to establish a central repository where models and meta models can be stored and retrieved via adaptors from various tools. Bëzivin et al. [6] coined the term MegaModel for the registry of models and meta models.

Versioning models in a distributed environment leads to the problem of conflicts due to concurrent commits. Thus adequate *model merging* algorithms are an important topic. Kolovos et al. [7] has compared the most important algorithms, as e.g., EMF Compare [8] or UMLDiff [9]. However there is still ongoing research in this area (see e.g., [10], [11]).

Related to the context of model repositories is the concept of *model integration* as a basis for (modelling) tool integration, as shown e.g., in Unicase [12], iRM [13], MOFLON [14], or (again) ModelBus [4], AMMA [6] and AMOR [5]. Mainly the integration is achieved by the implementation of suitable adaptors that connect the tools with a model repository. The category of model integration also comprises the concept of model transformation as ATL [6] or as discussed by Strommer at al. [15] and *round trip engineering*, which is already well established in many modelling tools, as e.g., Eclipse MDT [16] MagicDraw [17], or Rational Software Architect [18].

Atkinson and Stoll [19] choose a different solution by managing a common system view together with derived views.

Managing a common system model also includes the aspect of meta model management as e.g., proposed by [20]. Since also meta models may evolve there is a need for tools that allow for the co–evolution of meta models together with the associated models as e.g., COPE [21], or the work concerned by EMF Refactor [22].

A topic that is not yet sufficiently covered by research is the topic of change–driven modelling. There are certain (industrial–scale) tools that include aspects of change–driven requirements engineering, as e.g., DOORS [23] or in–Step change management [24].

These tools allow the requirements engineer to define state–based transition systems to model changes and their consequences to requirements.

The project MoVE (Model Versioning and Evolution) aims to establish an infrastructure to maintain Living Models. It does not concentrate on a single research topic, rather it strives to combine existing techniques (partially still under research, partially well established mechanisms). MoVE is based on a novel model based approach to implement a change–driven process.

III. Requirements for a Living Models infrastucture

In this section, we will map the generic principles for Living Models of Section I to requirements of a working Living Models infrastructure.

A. Concepts

Before going into details we define some concepts, that we need for the precise definition of the requirements.

A *m*odel captures a view of a physical system. It is an abstraction of the physical system, with a certain purpose. This purpose determines what is to be included in the model and what is irrelevant. Thus the model completely describes those aspects of the physical system that are relevant to the purpose of the model, at the appropriate level of detail. (Taken from [25] Section 17.3.1). Models are expressed in specific concrete syntax, which can be graphical, textual or other suitable notation. Examples are state charts, business process diagrams, lists, trees or graph presentations, or any other domain specific language.

So MoVE does not require models to be expressed in specific representation (as e.g., UML notation). However, as we see later in Section V, UML/EMOF or ecore are the main candidates for the MoVE internal model representation in our prototype. All external representations are mapped to these internal representations.

A *model element* is an atomic constituent of a model [26].

A *system model* is an abstraction of all relevant concepts and their relationships in a system. Thus a system model can be seen as a set of consistent models, or vice versa, a model represents a specific perspective on the system model. We speak of a *partial* model, if we want to emphasise that a model is part of the overall system model.

A *meta model* is a model that defines model element (types) and their relationships for expressing a model [27]. In extension to this definition a MoVE meta model additionally associates state machines to model elements, in order to model an event driven process.

Expressing it in terms of the MOF layer model [28], the meta model refers to layer M2, the system model and the (partial) models refer to layer M1.

Being a little bit more formal we consider a MoVE meta model as a tuple $(MM, \mathcal{C}, \mathcal{SM}, m{:}\mathcal{SM} \rightarrow Attributes(MM))$, where

- MM is a meta model expressed as EMOF–model [26], together with
- a set $\mathcal{C}$ of OCL statements,
- a set $\mathcal{SM}$ of state machines (expressed as UML–based behavioural state machines [25]), and
- a mapping m that maps each state machine to a distinguished state attribute in the meta model.

Set $\mathcal{C}$ defines additional OCL constraints that cannot be expressed in a standard EMOF-model (as e.g., logical dependencies of model elements). $\mathcal{SM}$ and m form the definition of a change–driven model maintenance process, which will be explained in Section III-D.

B. Basic Conceptual Architecture

One of the major principles of a Living Models infrastructure is that various stakeholders can cooperate through a set of tools via the common system model. The common system model reflects the actual state of the physical system. Each stakeholder has its own tool (set) to express and to maintain his/her own view of the complete picture. Figure 1 depicts this interaction.

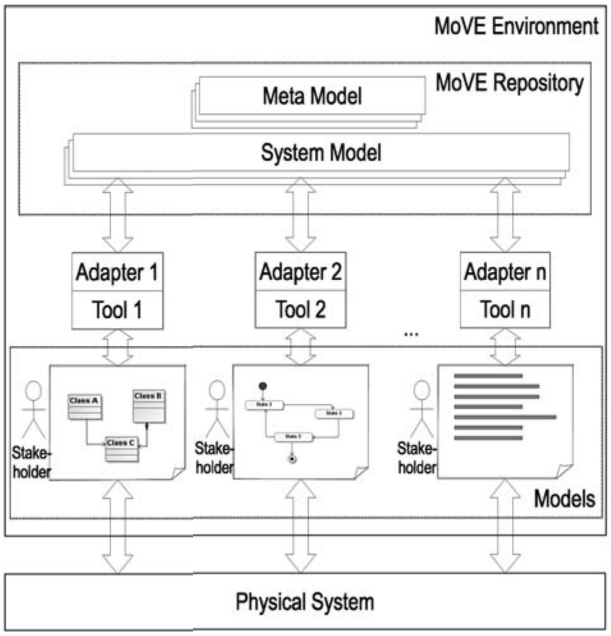

Figure 1. The Living Models infrastructure MoVE: Various tools interact with a common model repository

Partial models from different stakeholders may overlap, e.g., in the security model risks elements are related to model elements in the enterprise model, i.e., two views onto the system model may overlap. As long as this view is generated from the system model it can be easily handled as a projection on the relevant concepts of the system model, i.e., by just ignoring irrelevant concepts of the system model.

However, we accept that partial models are the basic artefacts to collect information about the real system, we have to merge modified partial models into the common system model. Depending on the type of underlying meta–meta–model as e.g., EMOF, UML, XML different options for merging algorithms exist as explained in [29].

Merging of models leads to the problem of merge conflicts that need to be resolved manually. In Section VI we propose some heuristics how to reduce such conflicts.

C. Model Versioning

The use case diagram in Figure 2 shows a high level view on the major use cases of a MoVE infrastructure. We can roughly group the use cases into model versioning use cases and change management use cases.

The actors are typically (human) stakeholders, that are in charge either to maintain the meta model as a meta model

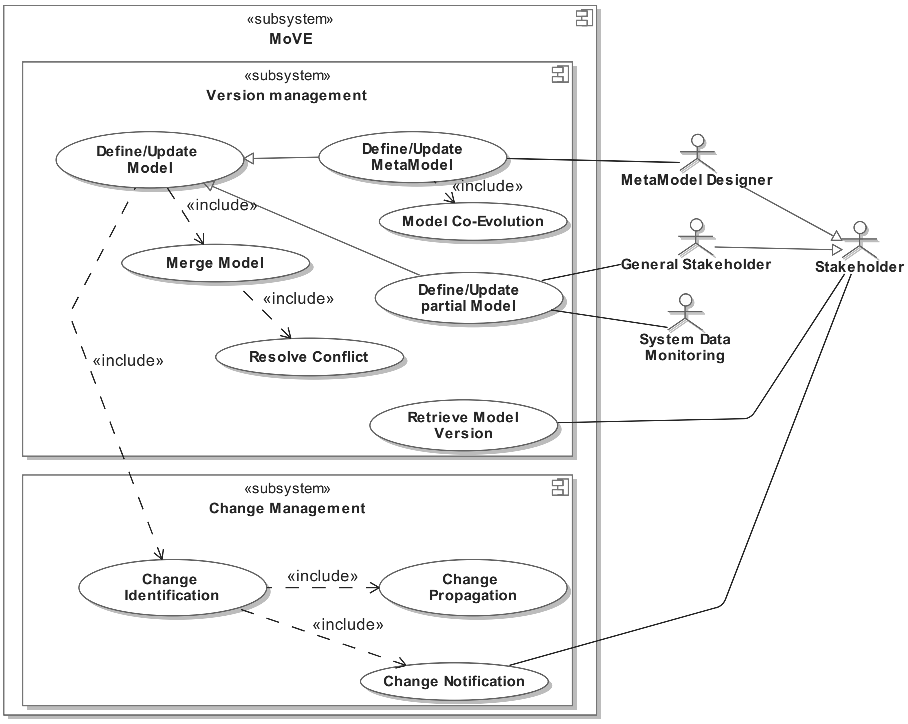

Figure 2. The actors and use cases of MoVE

designer, or to maintain his/her partial model.

Besides human stakeholders there may also be special (automatised) processes that monitor the physical system and forward system data automatically into models. An example could be an inventory process that monitors the state of all IT–applications running on a component and reports these states back into the enterprise model.

One major use case is *Define/Update Model*, where the model can be either a partial model or the meta model.

Updating a partial model may cause conflicts with the system model. Conflicts may result from classical versioning conflicts that are related to concurrent changes by different stakeholders. In this case, classical conflict resolution techniques, e.g., EMF Compare [8], may be applied. Other conflicts may result from the fact that constraints $\mathcal{C}$ from the MoVE meta–model are violated. The most consequent option for handling constraint violations would be, not to allow the commit of the partial model into the system model unless the conflicts are resolved. Since a violation is not necessarily restricted to a single partial model, or requires the interaction of several stakeholders (e.g., a cascading delete crossing model boundaries), we will just notify the concerned stakeholders that commits the change about this violation.

Partial models typically exist both in a proprietary representation used by the specific tool and as a part of the system model. So the tool adaptor must manage the mapping of the proprietary representation to the system model, support conflict resolution, and must keep the proprietary representation in synchrony with changes in the system model.

Besides the system model, also the meta model may evolve over time. In this case, mechanisms have to be provided to propagate changes in the meta model to the model layer. This is related to the large field of model migration, co–adaptation and co–evolution.

D. State Management

The second major functionality of a Living Models infrastructure is the change identification, propagation and notification.

Each class may have one or more distinguished attributes that represent the state of each instance. State transitions of such an instance are governed by an associated state machine. The set of state machines $\mathcal{SM}$ and the mapping m of the state machine to the distinguished attributes are defined in the MoVE meta model $(MM, \mathcal{C}, \mathcal{SM}, m{:}\mathcal{SM} \rightarrow Attributes(MM))$.

The MoVE repository monitors state changes of that dedicated model element attributes in a currently committed partial model and identifies an corresponding transaction in the associated state machine. The transaction may result in further transactions that changes the state of other model element (change propagation). Not all consequences of a change may be carried out automatically. Thus transactions may also trigger a notification of the responsible stakeholder. This stakeholder is then in charge of reacting on the change and to take further actions.

In order to illustrate these concepts, we present a case study in the following section.

IV. ATM Case Study

In this section, we present one of our actual case studies to give a clearer picture of how MoVE works and how it supports its users. The presented case study is taken from the area of Air Traffic Management (ATM) initiated by the EU project SecureChange [30]. It is concerned with the operational processes of managing air traffic in terminal areas, focusing on risk assessment and security analysis.

A. General Parts of the ATM Case Study

Within the SecureChange project three relevent types of models were designed, developed and maintained. In the following we will introduce the different types:

- *Enterprise Model.* This model defines the business processes, the information, the organisational units, etc. of the SecureChange project. The meta model is depicted as the bottom part of Figure 3.
- *Security Model.* This model defines all security (analysis) related objects. It contains risks, threats, security controls, security requirements and business security objectives. The (simplified) meta model is depicted as the upper part of Figure 3.
- *Common Meta Model.* This meta model describes as well the concepts of the Enterprise and the Security Model, as the interrelations between them (see Figure 3). Each element of the Enterprise Model is a generalisation of the "ModelElement"[1] of the Security Meta Model. This means that each object of an instance specification has a "state" attribute, which will be important in case security analysis started.

Both the common meta model and the enterprise model were documented with MagicDraw. The security model was documented with a project specific tool and persisted in a MySQL database. Based on the principle "Information Consistency and Retrieval" (see Section I), the Living Models infrastructure helps to keep the consistency across different models and modelling tools.

The usage of different tools in this project not only originated from different requirements of each model–purpose, but also from the different stakeholders involved, such as the maintainer of the system meta model, the responsible for IT–systems (typically the organisation's CIO or his/her deputy) and the responsible for security (e.g., the organisation's CSO). This is a consequence of the principle "Stakeholder–Centric Modelling Environments".

[1]Unfortunately the SecureChange meta model also used the term *ModelElement* as a type in its meta model, which is a slight inconsistency with the concept of *model element* in [26].

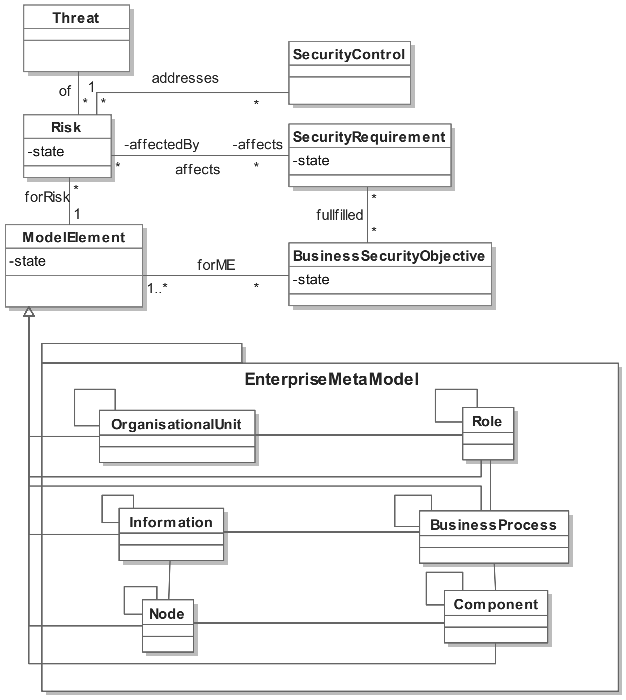

Figure 3. Enterprise Meta Model related to Security Meta Model

The security analysis conducted in the SecureChange project is based on regular system checks as soon as an element of one of the above mentioned models changes or is initially added. These checks may change the (security) state, attached to each model element. To support this security analysis the principles "Model Element States", "Change and Change Propagation" and "Change–Driven Process" have to be encompassed in the underlying infrastructure.

B. Change Management in Action

One of the main requirements of MoVE is to support Change Management and Propagation accross different models and their elements (see Section I). To fullfill this requirement we introduced the following three concepts:

- *"state" attribute:* Each model element has an attribute called "state". This attribute will be set and changed by the MoVE framework automatically (more details will follow in Section IV-C).
- *state machine definition (graphical and textual):* For each class of the Security Meta Model, a UML state machine is defined. These state machines define the possible state transitions from each model element from one state to another. To make the state machines more easy to process for MoVE, we translate them into SCXML [31], which is an XML representation, using OCL [32] for conditional expressions.

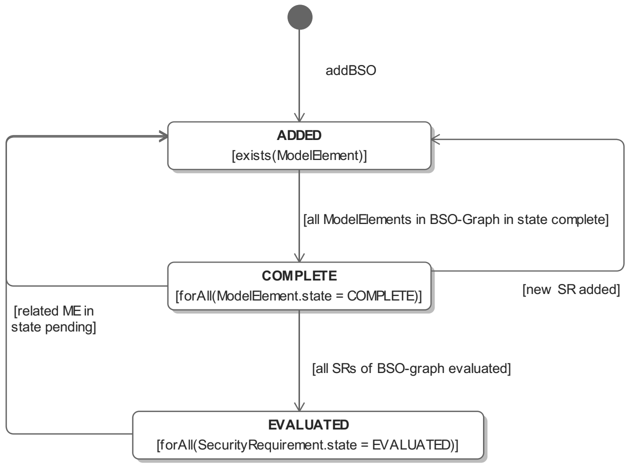

Figure 4. Example state machine for a SecureChange Business Security Objective (BSO) with informal description

- *set of states:* For each class of the Security Meta Model, a set of possible states is defined. For example, a BusinessSecurityObjective can have the states “ADDED”, “COMPLETE” and “EVALUATED”.

Figure 4 shows the UML state machine for the model element BusinessSecurityObjective.

Each BusinessSecurityObjective is in one of the following states:

- **ADDED**: This business security objective is identified by somebody, but not yet evaluated.
- **COMPLETE**: All model element instances in the enterprise model associated (indirectly by SecurityRequirements) with this BusinessSecurityObjective are identified.
- **EVALUATED**: the implementation of the security objective is evaluated.

As you can see, the conditions for entering a certain state, do not only depend on BusinessSecurityObjectives, but also on various other model elements, such as SecurityRequirements. This means that changes of one model element, can enhance changes of many other model elements, triggered by the state machines.

The same state machine as in Figure 4, is given in the following Listing 1. This textual representation is written with the language SCXML (State Chart XML), an XML dialect to define state machines.

```
1<scxml version="1.0" initialstate="initstate">
  <state id="initstate">
3  <initial/>
   <transition event="addBSO" target="ADDED"
5    cond="let result: Boolean = self.forME->asSet()->
       notEmpty()"/>
   </state>
7
  <state id="ADDED">
9  <transition event="all ModelElements in BSO-Graph in
       state complete" target="COMPLETE"
       cond="
11
     public definitions context EObject
13   def treeIterator: ...
     def getAttributeValueByName: ...
15   enddefinitions

17   public queries context EObject
     result:Boolean = treeIterator(self)->forAll(o1 |
19   o1.eClass().name = 'SR' implies
     getAttributeByName(o1,'State') = 'COMPLETE')
21   endqueries
   </state>
23
  <state id="COMPLETE">
25  <transition event="all SRs of BSO-graph evaluated"
       target="EVALUATED"
       cond=" ... "/>
27  <transition event="related ME in state pending" target
       ="ADDED" cond="
     let result: Boolean = self.forME->asSet()->exists(
       ModelElement me| me.state='pending')"/>
29  <transition event="new SR added" target="ADDED" cond="
       let result: Boolean = self.fulfilled ->
     asSet()->exists(SecurityRequirement r| r.state='
       ADDED')"/>
31 </state>

33 <state id="EVALUATED">
   <transition event="related ME in state pending" target
       ="ADDED" cond="
35   let result: Boolean = self.forME->asSet()->exists(
       ModelElement me| me.state='pending')"/>
   </state>
37 </scxml>
```

Listing 1. SCXML version of Business Security Objective state machine

The structure of Listing 1 looks as follows. Except of the SCML root element, the outermost elements are identified with the tags “state” (see Lines 2, 8, 24 and 33). These states have an ID, which represents the name of the states (compare Figure 4). Within each state element, we can define several “transition” elements. Each transition has two to three attributes:

- *event:* This attribute denotes the name of the transition, e.g., compare Line 25 of the above SCXML and the name of the transition between the states “COMPLETE” and “EVALUATED” in Figure 4, named *[all SRs of BSO-graph evaluated]*.
- *target:* The target attribute denotes the destination state after transition, or also called transition target.
- *cond:* This attribute is optional. In case of existence, it contains valid OCL code, incorporating the guard condition to follow a transition. In case of absence, the transition is an “immediate” transition. It is also possible to build more complex OCL statements including “definitions” and “queries”. Whith the help of these constructs (see Lines 12-21). The definitions “treeIterator” and “getAttributeValueByName” describe helping functions to parse the commited model. The OCL queries are invoked to gather the target ModelElements and the transitions to be triggered. For example, the query starting in Line 17 uses the functions defined above to check if all ModelElements attached to the actual Business Security Objective are in the state

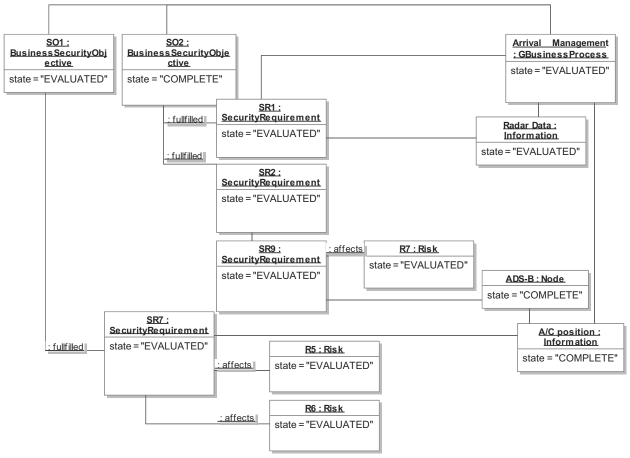

Figure 5. Commited instance model of a possible ATM case

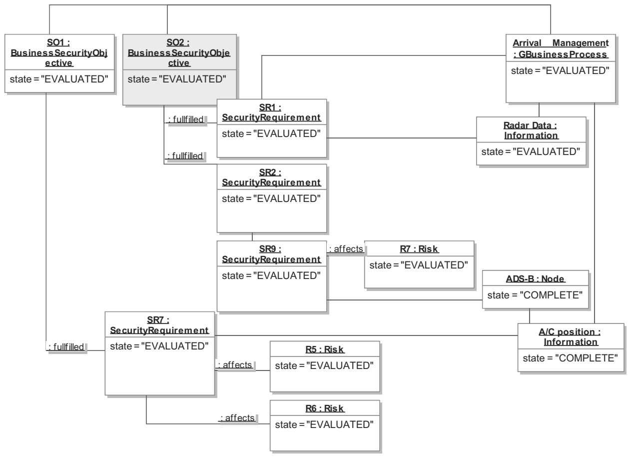

Figure 6. Instance model of Figure 5 after automatic state changes

"EVALUATED". In case the query returns the boolean value "TRUE", the transition is fired and the Business Security Objective is now in the state "COMPLETE".

Based on the state machines defined in SCXML, MoVE can check after each commit whether new state transitions are possible or not.

It may be the case that there are more than one possible transition (e.g., several model elements can change their state). In this case the out–come of the change propagation is non–deterministic. It depends on the selected sequencing of changes by the change propagation algorithm. It is in the responsibility of the meta model designer to define the state machines, such that this non–determinism does not lead to undesired results or even into circular state change loops.

C. Example Models

In this section, we introduce an example instance model of the ATM Case Study. Although the security model is stored in a relation MySQL database, the respective management tool synchronizes the content with the MoVE repository. MoVE represents the contents of the database internally by instance specifications with relationships modelled as slots [25]. Figure 5 shows a possible instance model (already converted into an internal representation) that will be commited by the user to the MoVE repository. Having a look at the states of all elements of this model, we can see that within the state machine "Business Security Objective" the condition for one transition is fulfilled. The according event is called "all SRs of BSO-graph evaluated", which means that in case of the Business Security Objective "S02", all attached Security Requirements (namely "SR1", "SR2" and "SR9") are in the state "EVALUATED". As soon as MoVE receives the instance model via a commit by the user, it triggers the according state transition. The resulting instance model after the transition is depicted in Figure 6. As soon as the transition finished its execution, MoVE parses the model element states again for possible transitions. Only if no further changes are possible and the set of model element states is stable, change propagation is finished and this version of the model will be stored in the repository.

V. Current Prototype

We have implemented an initial prototype in order to study the usage scenarios of Living Models.

The analysis of requirements presented in Section III implies a flexible architecture with exchangeable components. Therefore MoVE consists only of a small number of core components and a large number of exchangeable and extendible components. The architecture of MoVE is twofold: it consists of client–side and server–side components (see Figure 7).

To provide typical features of a standard VCS (version control system) and a stable and well tested communication protocol, Subversion (SVN) [34] is used as a core component on the client– and server–side.

The MoVE repository is built on top of an SVN server to update and commit models and meta–models. The main advantage of this decision is that both MoVE models and other artifacts can be versioned in the same repository. We use SVN properties to tag artifacts that are models, which need special handling by the MoVE infrastructure. MoVE hooks onto subversion by a subversion pre–commit hook, that invokes the dedicated MoVE functionality on the repository contents.

Conflict resolution is based on EMF Compare. State machine support is included based on SCXML.

In context of our case study (see Section IV) we have implemented three adaptors on the client–side, one for MagicDraw, one for Eclipse and another one for a proprietary

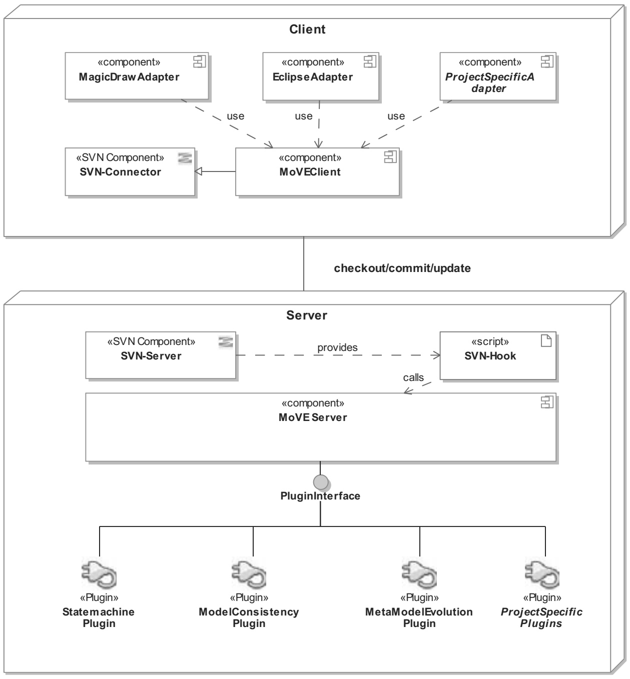

Figure 7. The MoVE Architecture [33]

application developed on top of a MySQL data base. A somewhat arbitrary decision was, to use the UML meta model as a representational base for the partial models. This was the best choice for a research prototype, because a lot of UML–support is available in the context of eclipse EMF. Model instances of the SecureChange Security Model were mapped from the database to instance specifications in UML. However the performance issues of handling a UML–model, compared to navigating through an optimized relational database, would not be acceptable in a professional application context.

VI. Conflict Reduction Strategies

Naturally when working jointly in a team on a set of documents (or partial models), concurrent conflicting modification of versioned items may occur.

Finding differences (and conflicts) between models is not as simple as in linearly structured documents as e.g., text files [35]. There exist also algorithms [7] to find differences and conflicts in models (as e.g., EMF Compare). However it turns out that in practical applications conflict resolution in models can be fairly cumbersome, if changes are quite complex. This get's even more complex if a system model is represented in various partial models.

In order to reduce the complexity of changes to the system model, we propose some rules to constrain most changes locally to partial changes and to manage changes that cross the boundaries between partial models.

MoVE currently uses the UML/EMOF meta model as its internal meta meta model. A MoVE meta model (e.g., as in Figure 3 is an instance of the EMOF meta model. The elements in the partial models are typically on the instance level, e.g., instances of SecurityRequirements or BusinessSecurityObjectives. The main elements for change on the instance level are instances of classes, attributes and associations. In Table I we identify atomic operations on this instances that can be sequentially combined to a more complex change.

object type	change type	
class	create	a new instance is created
	change	basic type attributes (e.g., the name) are changed
	delete	the instance is removed (together with its attributes and associations to other class instances)
attribute	create	an attribute is added to a class instance
	change	an attribute value is changed
	delete	an attribute is removed from a class instance
association	create	an association instance is added between two class instances
	change	not relevant
	delete	an association instance is deleted between two class instances

Table I
Atomic changes to a (partial) model

The main rule is, that each class, attribute, and association has its owner tool. Changes to instances of classes, attributes, and/or associations are only allowed by the tool that "owns" this class or attribute.

Such a rule is not unusual in practical applications, because every model element typically has its stakeholder that manages all instances of this type. A *security objective* instance is maintained by the CSO in his/her tool to maintain the security model, a *service* instance is maintained by the CIO in his/her tool.

However there may be situations where changes to some object type may inflict immediate inconsistencies, as e.g., an class instance cannot be deleted, because it would result in a constraint violation of the multiplicity of an association. A simple example could be that the class instance is partner in an association that must hold exactly one object of this type. Such a constraint can be compared to a foreign key constraint violation in a relational database.

A simple example could be that every *BusinessSecurityObjective* instance must be assigned to at least one *ModelElement* instance as in association *forME* shown in Figure 3. Since *BusinessSecurityObjectives* are managed by a stakeholder A, who is different from stakeholer B in charge for *ModelElements*, a *ModelElement* could never be deleted by B, before the stakeholder B deletes the corresponding

BusinessSecurityObjective. This would be very unfunctional and counter–intuitive.

This type of problem can occur with any consistency constraint that crosses model boundaries (i.e., involves instances from several partial models). The solution is to "extend" the final state of the class' state machine by adding a extra state named "toBeDeleted" (see Figure 8) into the state machine of a ModelElement. A plug–in that is in charge for maintaining the consistency between partial models, converts the deleted instance into an object of state "toBeDeleted".

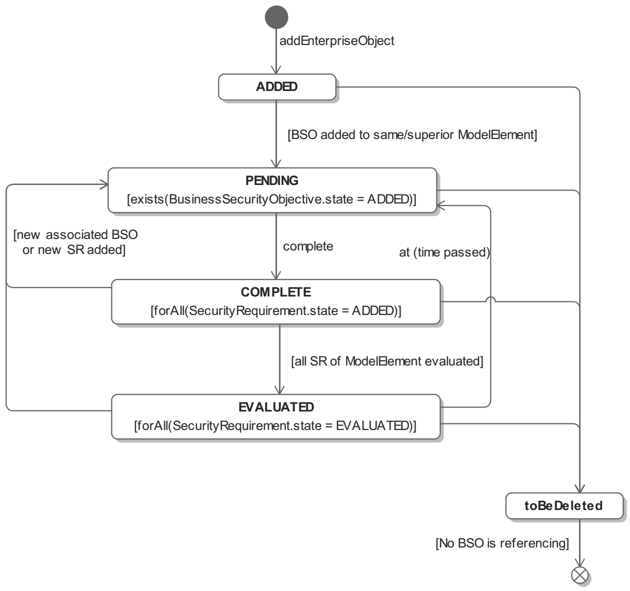

Figure 8. Extended state machine for ModelElement

Due to the notification mechanisms in the MoVE infrastructure the relevant stakeholder can be notified that he has to correct his partial model, in order to allow for the final deletion of this instance.

VII. Conclusion and Future Work

Managing IT systems with *Living Models* and to keep them alive, requires a rigorous development and maintainance process. This process must be change driven, in order to effectively coordinate the interaction of various stakeholders, taking jointly the responsibility for a complex system.

We have shown the major requirements and architectural implications for a *Living Models* infrastructure. Modelling relevant aspects of an IT–system and keep them up–to–date, needs an appropriate tool and cooperation support.

It strongly depends on the project or system context how comprehensive such a common system model can be. Establishing a "universal" system model where all tools describing any aspect of a system can contribute, will be doomed to failure. Our experience is that one should find a level of detail in the common system model, which is enough to control the relationships between domains and responsibilities of the stakeholders.

The central issues are model versioning and change–driven model evolution. Especially the second issue needs excellent tool support in order to ensure a lively modelling culture. Manual checks would fail, because models and their interdependencies tend to become more and more complex.

The challenge is to both combine results from different model research areas, and from well established software engineering disciplines, as e.g., classical versioning, flexible plug–in architectures together. The MoVE infrastructure represents an initial prototype that implements a working environment.

The prototype allowed us to study first impacts of a change–driven system management process based on models. We are sure that this is a natural evolution of the ideas of "model driven software engineering". Besides technical challenges like scalability, and extendibility, we still see a lot of open research questions. Challenges are:

- The implementation of adapters for various modelling tools is still a tedious work. The reason for this is, that not all tools technically provide usable interfaces to implement adapters, but also that the semantic gap between the tool's modelling representation and the chosen internal MoVE representation in EMOF can be quite huge.
- Change propagation can only be done partially automatic. I.e., efficient change notification mechanisms are needed to forward required actions to the relevant stakeholders. So MoVE should interface to work flow systems, which have a lot of mechnisms to handle event notification, and process control.
- A change–driven process also has social implications. It is a cultural shift from a capability based process to a change–driven process. What is the best way to bring such a process into an existing organisation?

This will be the starting point for further research on "Living Models".

This work is partially sponsored by the projects SecureChange (EU Prj. No. ICT-FET-231101), QELaB, and COSEMA (funded by the Tiroler Zukunftsstiftung).

References

[1] M. Breu, R. Breu, and S. Löw, "Living on the MoVE: Towards an Architecture for a Living Models Infrastructure." in *The Fifth International Conference on Software Engineering Advances (ICSEA 2010)*, IEEE Computer Society. Nice, France: IEEE Computer Society, 08/2010 2010, pp. 290–295.

[2] R. Huber, "Agile Practices and Model Driven Development in Avionics Software Development," in *Software Engineering Live*. Valtech GmbH, 2009.

[3] R. Breu, "Ten Principles for Living Models — A Manifesto of Change–Driven Software Engineering," in *International Conference on Complex, Intelligent and Software Intensive Systems (CISIS))*, 2010, pp. 1–8.

[4] A. Aldazabal, T. Baily, F. Nanclares, A. Sadovykh, C. Hein, and T. Ritter, "Automated Model Driven Development Processes," in *ECMDA Workshop on Model Driven Tool and Process Integration*. Stuttgart: Fraunhofer IRB Verlag, 2008.

[5] K. Altmanninger, G. Kappel, A. Kusel, W. Retschitzegger, W. Schwinger, M. Seidl, and M. Wimmer, "AMOR — Towards Adaptable Model Versioning," in *1st Int. Workshop on Model Co-Evolution and Consistency Management, in conjunction with Models' 08*, 2008.

[6] J. Bezivin, F. Jouault, P. Rosenthal, and P. Valduriez, "Modeling in the Large and Modeling in the Small," in *Model Driven Architecture*. Springer, 2005, pp. 33 – 46.

[7] D. S. Kolovos, D. D. Ruscio, A. Pierantonio, and R. F. Paige, "Different models for model matching: An analysis of approaches to support model differencing," in *Proc. of the Int. Workshop on Comparison and Versioning of Software Models (CVSM) in conjunction with ICSE, Vancouver*. IEEE Computer Society, 2009, pp. 1–6.

[8] Eclipse Foundation, "EMF Compare," July 2010, http://wiki.eclipse.org/index.php/EMF_Compare. Accessed on Jan 05, 2011.

[9] Z. Xing and E. Stroulia, "UMLDiff: An Algorithm for Object–Oriented Design Differencing," in *ASE '05: Proceedings of the 20th IEEE/ACM international Conference on Automated Software Engineering*. New York, NY, USA: ACM, 2005, pp. 54–65.

[10] M. Schmidt, S. Wenzel, T. Kehrer, and U. Kelter, "History–based Merging of Models," *ICSE Workshop on Comparison and Versioning of Software Models*, vol. 0, pp. 13–18, 2009.

[11] M. Kögel, J. Helming, and S. Seyboth, "Operation–based Conflict Detection and Resolution," in *CVSM '09: Proceedings of the 2009 ICSE Workshop on Comparison and Versioning of Software Models*. Washington, DC, USA: IEEE Computer Society, 2009, pp. 43–48.

[12] B. Brügge, O. Creighton, J. Helming, and M. Kogel, "Unicase – An Ecosystem for Unified Software Engineering Research Tools," in *Third IEEE International Conference on Global Software Engineering, ICGSE*, 2008.

[13] I. Petrov and S. Jablonski, "An OMG MOF Based Repository System with Querying Capability — The iRM Project," *Proceedings of iiWAS*, vol. 4, 2004.

[14] C. Amelunxen, F. Klar, A. Königs, T. Rötschke, and A. Schürr, "Metamodel–based tool integration with MOFLON," in *Proc. of the 30th Int. Conf. on Software Engineering. New York, NY : ACM*. ACM, Januar 2008, pp. 807–810.

[15] M. Strommer and M. Wimmer, "A framework for model transformation by–example: Concepts and tool support," *Objects, Components, Models and Patterns (TOOLS). LNBIP. Springer, Heidelberg*, pp. 372–391, 2008.

[16] Eclipse Foundation, "Model Development Tools (MDT)," http://www.eclipse.org/modeling/mdt/?project=uml2. Accessed on May 17, 2010.

[17] No Magic, Inc., "MagicDraw," http://www.magicdraw.com/. Accessed on May 17, 2010.

[18] IBM, "Rational Software Architect," http://www-01.ibm.com/software/awdtools/swarchitect/websphere/. Accessed on May 17, 2010.

[19] C. Atkinson and D. Stoll, "Orthographic Modeling Environment," in *FASE*, 2008, pp. 93–96.

[20] B. Baudry, C. Nebut, and Y. Le Traon, "Model–driven engineering for requirements analysis," in *11th IEEE International Enterprise Distributed Object Computing Conference, 2007. EDOC 2007*, 2007, pp. 459–459.

[21] M. Herrmannsdoerfer, S. Benz, and E. Juergens, "COPE — Automating Coupled Evolution of Metamodels and Models," in *23rd European Conference on Object-Oriented Programming (ECOOP)*, ser. Lecture Notes in Computer Science, S. Drossopoulou, Ed., vol. 5653. Springer, 2009, pp. 52–76.

[22] T. Arendt, F. Mantz, L. Schneider, and G. Täntzer, "Model Refactoring in Eclipse by LTK, EWL, and EMF Refactor: A Case Study," in *Models and Evolution, Joint MoDSE–MCCM 2009 Workshop*, 2009.

[23] IBM, "Rational DOORS," http://www-01.ibm.com/software/awdtools/doors/. Accessed on May 17, 2010.

[24] micro TOOL, "in–Step," http://www.microtool.de/instep/. Accessed on May 17, 2010.

[25] OMG, "Unified Modeling Language 2.2," 2009. [Online]. Available: http://www.omg.org/spec/UML/2.2/Superstructure/PDF/

[26] OMG, "Meta Object Facility (MOF) Core Specification, Version 2.0," January 2006.

[27] OMG, "Unified Modeling Language 2.0 Infrastructure Specification," September 2003.

[28] OMG, "Meta Object Facility (MOF) Core Specification, Version 1.4," April 2002.

[29] L. M. Rose, R. F. Paige, D. S. Kolovos, and F. A. Polack, "An Analysis of Approaches to Model Migration," in *Models and Evolution, Joint MoDSE–MCCM 2009 Workshop*, 2009.

[30] A. Armenteros, B. Chetali, M. Felici, V. Meduri, Q.-H. Nguyen, A. Tedeschi, F. Paci, and E. Chiarani, "D1.1 Description of the Scenarios and their Requirements," SecureChange, Tech. Rep., 2010, http://securechange.eu/sites/default/files/deliverables/D1.1_Description_of_Scenarios_and_their_requirements.pdf. Accessed on January 14, 2011.

[31] J. Barnett, R. Akolkar, R. J. Auburn, M. Bodell, D. C. Burnett, J. Carter, S. McGlashan, and T. Lager, "State Chart XML (SCXML): State Machine Notation for Control Abstraction," World Wide Web Consortium, Working Draft WD-scxml-20080516, May 2008.

[32] OMG, "UML 2.0 OCL Specification," *(ptc/03-10-14)*, 2003.

[33] P. Kalb, "UML Model Merging in the Context of a Model Evolution Engine," Master's thesis, Leopold-Franzens-University Innsbruck, 2010.

[34] C. M. Pilato, B. Collins-Sussmann, and B. W. Fitzpatrick, *Version Control with Subversion*. O'Reilly Media, 2008.

[35] The GNU Team, "Comparing and Merging Files," April 2002, http://www.gnu.org/software/hello/manual/diff.html. Accessed on Jan 05, 2011.

Sustained Weight Loss during 20 Months using a Personalized Interactive Internet Based Dietician Advice Program in a General Practice Setting

Vibeke Brandt
University of Southern Denmark
Odense, Denmark
e-mail: vibra@student.sdu.dk

Carl J. Brandt
Stenstrup Lægehus
Svendborg, Denmark
e-mail: carl_brandt@get2net.dk

Dorte Glintborg
Department of Endocrinology M
Odense University Hospital
Odense, Denmark
e-mail: dorte.glintborg@dadlnet.dk

Cecilia Arendal
Clinical Dietician
Nyborg, Denmark
e-mail: cia-erik@adslhome.dk

Søren Toubro
Reduce
Roskilde, Denmark
e-mail: st@reduce.dk

Kirsten Brandt
Human Nutrition Research Centre, School of Agriculture, Food and Rural Development
Newcastle University
Newcastle upon Tyne, United Kingdom
e-mail: kirsten.brandt@newcastle.ac.uk

Abstract—**Obesity is an increasing drain on the resources of general practitioners, who have few effective options for treatment other than surgery and (often prohibitively expensive) personal dietician advice. This pilot project investigated the weight loss efficacy and the cost of an interactive internet-based weight loss program in a Danish medical center setting. The study comprised an initial weight loss period of approximately 4 months, consisting of frequent online consultations with a dietician and an exercise coach supported by electronic diaries and establishment of an online community, where the patients exchanged experiences with other users of the program. This was followed by a 16-month maintenance treatment providing less intensive counseling. Of 46 obese patients offered participation, 32 patients were enrolled in the study and 21 completed the full course. The mean weight at inclusion was 104 kg with a BMI of 36.4 kg/m^2. After 4 months of treatment and an average of 17 consultations the participants lost on average 7.0 kg, p<0.001. During the 16-month maintenance period, the average weight did not change and 81% of the participants retained or increased their initial weight loss. The cost of the initial treatment was calculated as 165 DKK (approx. €22) per kg weight lost. These results indicate that e-mail consultations can produce comparable weight loss as conventional weight loss treatments in general practice at a lower cost, particularly for sustaining the weight loss over a longer period of time. The results of this preliminary uncontrolled study with few participants indicate that future randomized clinical trials with more participants comparing the e-consultations with relevant conventional practices are justified, in order to quantify effect and long-term cost-efficiency of e-consultations as an intervention against obesity.**

Keywords-obesity; internet community; treatment; preventive medicine

I. INTRODUCTION

Obesity is a growing problem, resulting in an increasing demand for efficient weight loss treatments suitable for use in general practice settings [1]. According to the Framingham study, obesity shortens life by 3 to 8 years for a 40-year-old person [2] demonstrating the urgent need for effective ways to reduce obesity.

Form and content of communication are important in order to modify life style factors [3], which is the general recommendation to obtain a sustainable weight reduction. Several Cochrane reviews infer that advice from general practitioners (GPs) by itself does not have a long-term effect on weight loss compared to placebo [4]. Dietician guidance and the establishment of group meetings have a significant effect in the short term, but only few long-term studies are available [4][5]. Surgical intervention is an effective long-term weight loss option, but in Denmark it is reserved for very obese patients [5]. Consequently no documented effective non-surgical weight loss offers are available to the GPs of the Danish National Health Service for patients with simple overweight [6].

There is now a vast literature on how the internet can be integrated as a consultation tool [5][7][8][9]. For example, several studies suggest that interaction over the internet with experts in an internet based community is the most effective way to lose weight [10][11]. Recent studies suggest that online contacts are an economically attractive contact form for optimizing guidance on diet and exercise and in keeping the patients motivated [12][13][14]. The internet furthermore substantially facilitates the use of motivational tools such as self-monitoring, which has been used successfully in other approaches of internet based weight loss interventions [16].

In Denmark, 86 % of the population has internet access at home [15], which makes it possible to reach out to most of the patients by online intervention.

Conventional dietician advice is costly, therefore it is important to ensure that resources are being used in the best way possible. In Denmark it is now possible to employ dieticians in general practice and health care centers. However many practices experience difficulties to organize the activities in a way that makes it economically feasible to offer diet treatment to patients within the rates provided by the Danish National Health Service.

The present paper reports on the methods used and results achieved in a preliminary uncontrolled prospective survey of weight loss and weight maintenance among obese patients, who received advice and support about diet and exercise using a personalized interactive internet-based dietician advice program in a clinical practice setting.

II. METHOD AND SUBJECTS

A. Patients

One medical center with primary care participated in the study. In May 2008, new patients and patients who were already enrolled in weight stabilization courses were offered the opportunity to participate. Initially 46 patients attended a consultation with a dietician, of which 32 patients agreed to participate in the full study and signed the informed consent form. Patients then received information on how they could log on to the program. Before attending the dietician, the patients filled out name, address and e-mail address. The study was approved by the South Danish regional committee on biomedical ethics.

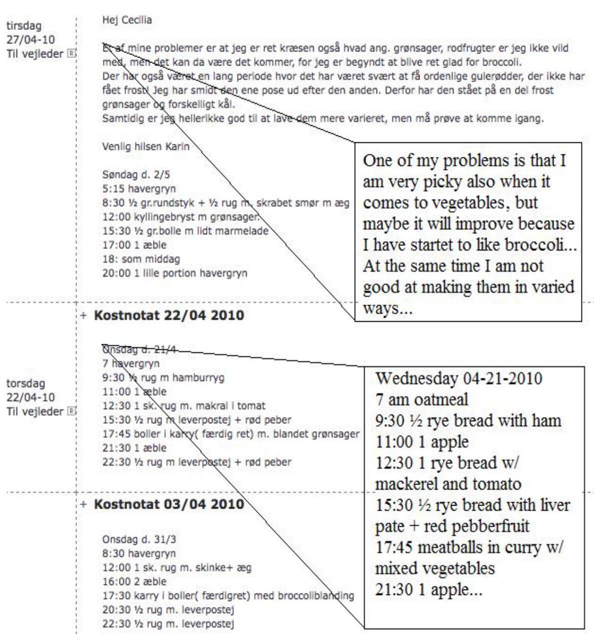

Figure 1. Screenshot of dietary notations on a daily basis from one of the participants using the website.

B. Study design

The pilot study was designed as an uncontrolled prospective survey of the efficacy of using an existing commercial weight loss program [17] for obese patients in a general practice setting. At the first login, patients filled out a comprehensive 16-page medical history with information regarding their health, education and medicine intake. Completing the forms gave e-access to consultations with a dietician and an exercise coach.

E-access also allowed e-mail chats with the other patients participating in the study.

During the first week the patients recorded a diet and exercise history on a day-to-day basis (see Figure 1). Based on these records, the patients received a diet plan, weekly advice from a dietitian (see figure 2) as well as an exercise plan and advice from an exercise coach once a month. Treatment principles in both online and physical consultations were based on the Danish Board of Health's recommendations from "The 8 dietary guidelines" [6].

The aim was to enhance the daily intake of vegetables and fruits, choose whole-grain options for bread and other cereal products, replace products rich in fat with lean alternatives and distribute food intake into several smaller meals throughout the day.

The treatment consisted of providing simple and manageable guidelines and tools that gave the patients substantial freedom in planning their meals while the dietitian could supervise and advise each patient individually on where improvements could be achieved. Patients who according to the dietician's professional assessment needed face-to-face consultation with the dietician during the study period, were seen by a dietician in the medical center 2 to 3 months into the study period. Only one dietitian was connected to the project, which means that the patients were met by the same dietician at the medical center as online.

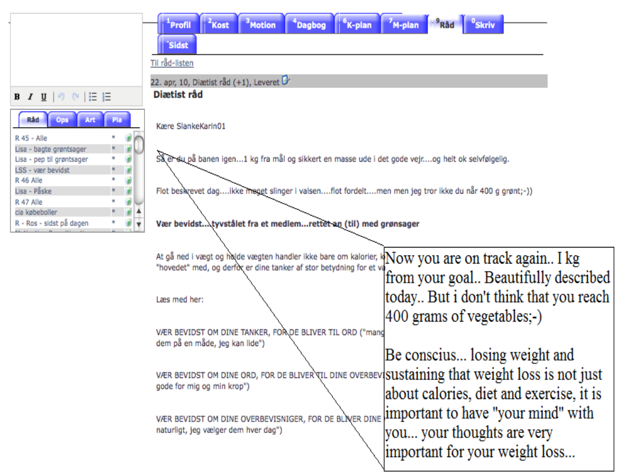

Figure 2. Screenshot of personal advice from the dietician as a response to the information in Figure 1.

The internet tools in the program encouraged the patients to record their exact dietary intake on a day-to-day basis, enabling the dieticians at the face-to-face consultation to focus more on serving the patient's needs and spend less time to simply clarify recent food intake. The patients could also write about any complication or worry that they might have during a day, as illustrated in the example in Figure 1. Dietary notes and commentaries from the patients were used by the dietician, the exercise coach as well as peers (other users of the program) to intervene and relate to problems when they appeared.

The patients were supported by their peers on the website by using an internet community (presented as 'forums for debate), consisting of all users of the internet based program, both the patients enrolled in the study and other users who pay for the service privately to lose weight in a non-clinical setting. The members of this internet community were encouraged to contact each other for support, as seen in the example in Figure 3.

The internet community was very intimate as only patients with a weight problem had access. The patients could communicate via discussion forums and internet chat forums. Communication was also available in specially designed inboxes on the website, as comments to food and exercise records or via personal pictures.

To illustrate this we have chosen some typical comments from the patients: "I know I don't use as many fiber rich vegetables in my salad as I would like, but there is no room when I eat my regular salad that I love…" "…the weather makes it difficult for me to exercise because of my gout. What can I do?" or "my dog died. I'm so sad…". The dietician would get to know every participant individually, and be able to guide them in a way that suited their lifestyle, as seen in Figure 2.

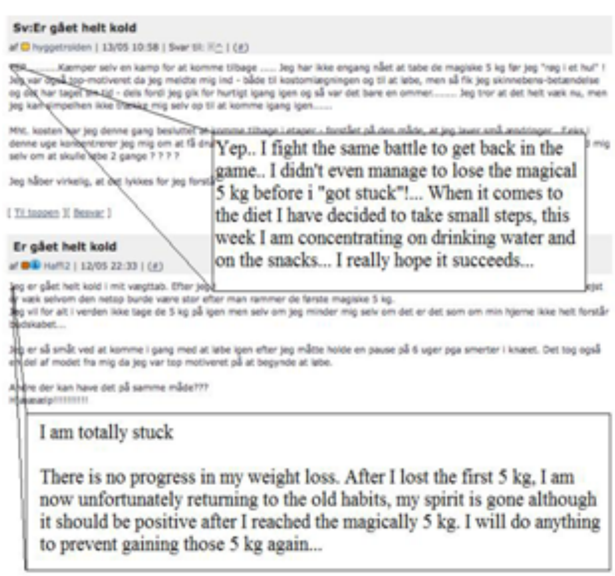

Figure 3. Screenshot of participants using the debate forums from the internet community.

Waist and hip measurement, weight and clinical analysis values were obtained at baseline and after 4 month initial treatment. These assessments were performed by the study team at physical consultations, see table 1. The following approx. 16 months (maintenance period) most of the patients continued to use the program, but only received internet consultations when requested by the patient or by the dietician, and weight measurements were recorded whenever the patients attended the health centers for other reasons.

Results were analyzed as a one-sample t-test for the hypothesis that the weight loss or other change from one time point to the next was different from 0.

III. Results

The datasets from 22 of the 32 enrolled patients were sufficiently complete to be included in the outcome calculations. Of the remaining 10 patients, 2 only registered starting weight and the remaining 8 never logged on. All 22 completed the initial treatment period and 21 the maintenance treatment. One patient only completed a 12-month period due to pregnancy; we used the last observation carried forward. Baseline data and details of treatment for the 22 patients who participated in the are given in Table 1.

The average weight loss after the initial intensive treatment period of 115 days (95% CI: 101; 121),was 7.0 kg, with a standard error of the mean of 1.1 (95% CI: 4.6, 9.3), P<0.001. Nine participants achieved a weight loss of 5-10 kg and 4 participants lost more than 10 kg. There were no significant correlations between weight loss and duration of treatment period or between weight loss and number of consultations. Clinical and anthropomorphic data of patients enrolled are listed in Table 1. The mean age at inclusion was 43 years and the mean weight 104 kg with a BMI of 36.4 kg/m^2.

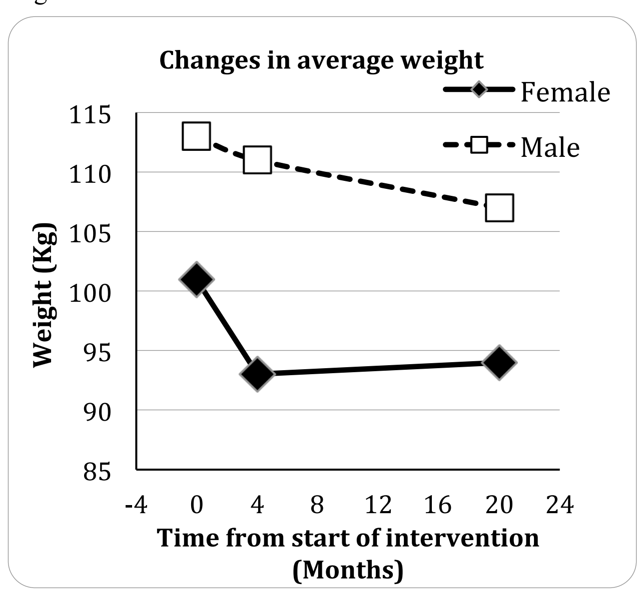

Figure 4. Average weight of the participants at baseline, after 4 months and after 20 months follow up.

The mean weight loss from baseline after the maintenance period, a total of approx. 20 months, 595 days (95% CI: 519; 671), was still 7.0 kg. 15 out of 21 achieved a weight loss between 5 and 29 kg. 4 lost between 0.3 and 2.2 kg, and the last 3 patients gained between 0.1 and 4.7 kg. One patient became pregnant and 1 patient was absent for the approx. 20 month assessment, they were excluded and we used the last observation carried forward. Seventeen out of the 21 i.e. 81% of the participants managed to sustain a weight loss of more than 1 kg after 20 months.

The dietician and some of the patients were interviewed about their experience with the program. Both parties agreed that one of the most important parts of the program was the continuity. The dietician was always available over the internet, which created an ongoing motivation for lifestyle changes instead of a short-term diet change. The patients also found that continuous emotional support and practical advice from peers had been very important during the study. Some patients found the internet community equally important as the dietician, as illustrated by a comment: "I spend most of my time on the internet community, I like to see how the others are doing and whether they have the same problems as me."

Feedback from the doctors and staff in the medical center indicated that they were satisfied with the cooperation with the dietician. It was seen as a benefit to offer dietician advice close to the patients without requiring frequent visits to the medical center. The main challenge mentioned by this group was the technical integration with the existing e-journal system of the Danish National Health Service.

The total cost of the initial weight loss treatment, including the personal (face-to-face) consultations and clinical assessments, was estimated to a total of approximately 35,000DKK. Without the cost of the study assessments the cost would have been approximately 25,700 DKK. This estimate corresponds to 165 DKK (€22) per kg weight lost for the treatment provided.

TABLE 1. Clinical and anthropomorphic data for the patients. Full set of data at baseline and after the initial treatment period, only weight measurements after the approx. 20 months maintenance period.

	Females (n=17)			Males (n=5)		
Age(years)	42 (39-46)			43 (37-50)		
	Start	4 months	20 months	Start	4 months	20 months
E-mail cons.	0	17 (14-20)	57 (45-70)	0	17 (12-21)	29(18-39)
Period (days)		111 (97-126)	568 (485-650)		130 (119-140)	683 (649-717)
Weight (kg)	101 (94-108)	93 (85-100)**	94*** (85-102)**	113 (108-117)	111 (105 -116)*	107 (101-113)*
BMI (kg/m^2)	35.6 (32.9-38.2)	32.6 (29.9-35.3)**	32.9 (29.9-36.0)**	39.1 (38.0-40.2)	38.4 (37.1-39.8)	37.3 (33.0-41.6)
Waist (cm)	103.2 (97.3-109.2)	95.4 (90.2-100.)*		122.7 (119.9-125.5)	120.3 (116.6-124.1)	
Hip (cm)	119 (112.1-125.8)	109.4 (104.2-114.6)*		113.5 (112.5-114.5)	113.5 (112.5-114.5)	
WHR	0.86 (0.82-0.93)	0.86 (0.82-0.93)		1.10 (1.07-1.13)	1.09 (1.04-1.13)	
Total cholesterol	5.2 (4.6-5.7)	4.9 (4.4-5.3)		5.8 (5.6-6.0)	5.0 (4.5-5.5)	
LDL (mmol/l)	3.1 (2.5-3.6)	3.0 (2.5-3.4)		3.3 (3.0-3.5)	3.5 (3.3-3.7)	
HDL (mmol/l)	1.7 (1.1-1.4)	1.7 (1.1-1.4)		1.0 (0.8-1.1)	1.1 (1.0-1.1)	
TG (mmol/l)	1.7 (1.3-2.1)	1.3 (1.0-1.6)		4.2 (3.3-5.1)	2.2 (1.6-2.7)	
HbA1c (%)	5.0 (5.3-6.8)	5.2 (5.2-6.3)		8.0 (6.6-8.7)	7.0 (6.0-7.8)	

Data presented as mean value (95% confidence interval)
* $p<0.05$ vs. before related to same sex
** $p<0.001$ vs. before related to same sex
*** Two patients were excluded due to pregnancy and absence; we used last observation carried forward (one after 4 months and one after 10 months)

The cost to the Danish National Health Service of an e-mail consultation was 49.68DKK, compared with 211.14DKK for a consultation in the medical center with the aim of assisting the patient to change lifestyle. The average price per patient for the 4-month weight loss process in this implementation study was 1165 DKK. Data on costs are not available for the subsequent maintenance period, but they were substantially lower than during the initial treatment period.

IV. Discussion

In the present study we used a combination of the expertise available to a Danish health center with an interactive e-consultation delivery tool and an internet community to accomplish a sustainable weight loss amongst obese Danes. Using this method we achieved an average weight loss of 7.0 kg during the first 4 months, which is comparable to other conventional treatments [10]. A maintained average weight loss of 7.0 kg after 20 months follow up is a strong indicator that this might be a way to efficiently and cost-effectively reduce weight for a large population. However a randomized controlled trial would be necessary to determine whether the results are reproducible.

Bennett *et al* recently reported on a randomized clinical trial of another web-based weight loss program in primary care in the USA, with a similar study population (baseline BMI 34.6 and age 54.4). This study showed a comparable efficacy with a 3.05 kg greater weight loss amongst cases compared with usual care. The trial period was 12 weeks and one of their conclusions was that trials of longer duration are necessary [16].

This study however displayed several differences from ours, i.e. they tried to create adherence to the program by offering the possibility of winning money and their internet program did not include facilitated peer support (online community). Since many of the patients used the community frequently during our study and found it very beneficial this may be an important difference. Also the patients in our study corresponded with the same dietician over the internet and during the counseling in general practice, in contrast to the study of Bennett *et al* [16], where the program was not designed to provide individual counseling. Using the patient's interaction with the community and e-mails from the patient, the dietician could follow the patients' progress and provide more accurate and effective advice, since she had the opportunity to build a greater understanding of every individual. Together with the longer weight loss period this might be a reason why our study appeared to show a larger weight loss, while their results are more significant due to a higher number of participants and more relevant due to the comparison with a control group. However in combination these two studies strongly indicate that as a concept internet based weight loss programs can be successful in the short term, may be useful in longer term maintenance of weigh loss and can be effectively introduced in health care.

Several studies suggest that keeping the patients in the program is as effective as frequent follow-up but cheaper [5]. In the present study, the low number of dropouts among those who progressed beyond the first week was remarkable; according to the patients this was mainly due to the community, where they established relationships with other patients. The low dropout rate could also be affected by the fact that the patients all received advice from the same dietician, and she could possibly be very good at keeping the patients motivated. It would be interesting to further investigate specifically the efficacy of the internet community, since the Bennet et al. study showed that the more patients were using the internet program, the greater their weight loss [16], even though no such correlation was found with the small number of participants in the present study. Future developments of the program could focus on the community and make it more attractive. We may achieve greater and in particular more sustained weight loss results if we could get the patients more involved in the program through the internet community.

The use of self-monitoring provided by the online program had both advantages and disadvantages. The patients were able to follow their own progress using the website, which can help keep motivation, and the data that the patients provided were essential as tools for the dietitian to achieve the very cost-effective provision of advice. In contrast, the self-reported data could not be used for, evaluation of the intervention outcome, due to potential bias such as under- or over-estimations or recall bias. Therefore only data from measurements that were carried out at the medical center by the dietician or a nurse are presented in the present paper.

Web-based interventions have the disadvantage that the participants must be fairly proficient at using the internet and have the required writing and reading skills for using the program, in addition to the obvious requirement for convenient internet access. In our study, 14 patients who were offered participation failed to go through the enrollment progress, and some of these could be due to lack of computer or writing skills.

Feedback from doctors and secretaries involved in the study was very positive. It was agreed that the program could potentially help to better utilize the scarce dietician resources by decreasing the need for consulting face to face. In relation to the implementation in the medical centers, it was important that there should be a technical integration that makes the internet portal an integrated part of the electronic journal system used by the Danish National Health Service. Communication between the medical center and dieticians could become an integrated part of the doctors daily work tool. It is especially necessary to establish a technical integration with billing and information exchange, to minimize the need for intervention by the other staff at the medical center.

With a total of 500 licensed dieticians in Denmark and approximately 50 newly educated every year, faced with the needs of 4000 medical centers to provide relevant treatment offers to ever increasing numbers of obese patients, the present pilot trial indicates the potential usefulness of this type of effective and economically attractive individual internet treatment for the large part of the population in need of dietary advice.

V. Conclusion and Future Work

The study showed that the internet based interactive weight management program may be a cost-effective way to produce a significant and sustained weight loss among patients with obesity in general practice. The internet can be used as a communication tool for lifestyle changes and provide a community for the patients to support them to maintain weight loss and healthier life style. We have developed a protocol for a randomized controlled trial to further investigate the efficacy of this weight-loss program in a more controlled setting, comparing the intervention with usual care [18]. Furthermore, we are working on a refinement of the internet platform to record how much each of the participants use the internet community, as a tool to assess the importance of this feature.

Acknowledgment

Thanks to medical student Mathilde Pedersen, University of Southern Denmark and to political science student Sara Katrine Brandt for editorial contributions to this paper. Thanks to the staff at Stenstrup Lægehus, Nyenstadlægerne and Marstal Lægehus for making this study possible. Thanks to Region Syddanmark and the KEU fund for input and financial support.

References

[1] C. J. Brandt, C. Arendal, D. Glintborg, S. Toubro, "e-Dietician in General Practice," etelemed, pp.29-32, 2010 Second International Conference on eHealth, Telemedicine, and Social Medicine, 2010

[2] A. Peeters, J. J. Barendregt, F. Willekens, J. P. Mackenbach, M. A. Al, and L. Bonneux. Obesity in adulthood and its consequences for life expectancy: a life-table analysis. Ann Intern Med 2003 Jan 7;138(1):24-32.

[3] C. J. Brandt, H. Ellegaard, M. Joensen, F. V. Kallan, A. D. Sorknæs, and L. Tougaard. Effect of diagnosis of "smoker's lung". RYLUNG Group. Lancet 1997 Jan 25;349(9047):253.

[4] R. L. Thompson, C. D. Summerbell, L. Hooper, J. P. Higgins, P. S. Little, D. Talbot et al. Dietary advice given by a dietitian versus other health professional or self-help resources to reduce blood cholesterol. Cochrane Database Syst Rev 2003;(3):CD001366.

[5] M. Neve, P. J. Morgan, P. R. Jones, and C. E. Collins. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. Obes Rev 2010 Apr;11(4):306-21.

[6] B. Richelsen, A. Astrup, G. L. Hansen, H. S. Hansen, B. Heitmann, L. Holm, et al. A report from the board of Nutrition. The Danish epidemic of obesity – Proposal for a prevention effort . 2003.

[7] T. J. Moore, N. Alsabeeh, C. M. Apovian, M. C. Murphy, G. A. Coffman, D. Cullum-Dugan, et al. Weight, blood pressure, and dietary benefits after 12 months of a Web-based Nutrition Education Program (DASH for health): longitudinal observational study. J Med Internet Res 2008;10(4):e52.

[8] D. J. Wantland, C. J. Portillo, W. L. Holzemer, R. Slaughter, and E. M. McGhee. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40.

[9] R. A. Winett, D. F. Tate, E. S. Anderson, J. R. Wojcik, and S. G. Winett. Long-term weight gain prevention: a theoretically based Internet approach. Prev Med 2005 Aug;41(2):629-41.

[10] B. C. Gold, S. Burke, S. Pintauro, P. Buzzell, and J. Harvey-Berino. Weight loss on the web: A pilot study comparing a structured behavioral intervention to a commercial program. Obesity (Silver Spring) 2007 Jan;15(1):155-64.

[11] D. F. Tate, R. R. Wing, and R. A. Winett. Using Internet technology to deliver a behavioral weight loss program. JAMA 2001 Mar 7;285(9):1172-7.

[12] J. Harvey-Berino, S. Pintauro, P. Buzzell, and E. C. Gold. Effect of internet support on the long-term maintenance of weight loss. Obes Res 2004 Feb;12(2):320-9.

[13] H. A. Haugen, Z. V. Tran, H. R. Wyatt, M. J. Barry, and J. O. Hill. Using telehealth to increase participation in weight maintenance programs. Obesity (Silver Spring) 2007 Dec;15(12):3067-77.

[14] L. P. Svetkey, V. J. Stevens, P. J. Brantley, L. J. Appel, J. F. Hollis, C. M. Loria, et al. Comparison of strategies for sustaining weight loss: the weight loss maintenance randomized controlled trial. JAMA 2008 Mar 12;299(10):1139-48.

[15] Statistics Denmark. Internet access in the home - devices with access to internet. http://www.statistikbanken.dk/BEBRIT01 (accessed January 2nd, 2010).

[16] G. G. Bennett, S. J. Herring, E. Puleo, E. K. Stein, K. M. Emmons, and M. W. Gillman. Web-based weight loss in primary care: a randomized controlled trial. Obesity (Silver Spring) 2010 Feb;18(2):308-13.

[17] SlankeDoktor. www.slankedoktor.dk. (accessed January 2nd, 2011)

[18] V. Brandt, M. Pedersen, C. J. Brandt, C. Arendal, D. Glintborg, S. Toubro, K. Brandt "Design and aims of study to measure long-term efficacy of interactive online diatician weight loss advice in general practice," etelemed, 2011 Third International Conference on eHealth, Telemedicine, and Social Medicine, 2011