Health Technol. (2012) 2:43–55
DOI 10.1007/s12553-012-0015-6

ORIGINAL PAPER

Use of collaboration techniques and classification algorithms in personal healthcare

Vladimir Trajkovik · Elena Vlahu-Gjorgievska · Igor Kulev

Received: 3 November 2011 / Accepted: 17 January 2012 / Published online: 2 February 2012

Abstract Adoption of mobile devices and technology in the field of medical monitoring and personal health care systems is very important nowadays, especially when it comes to certain categories of people with chronicle diseases who need 24 hour access to medical care. The collaborative Information system model we present in this paper, gives a new dimension in the usage of novel technologies in healthcare. Using mobile, web and broadband technologies enable the citizens to have ubiquity of support services where ever they may be. The model incorporates collaboration techniques and classification algorithms in order to generate recommendations and suggestions for preventive intervention. In addition, the system enables the patient (system user) to contact other people with similar condition and exchange their experience. This system improves the terms of home care treatment of the patient and allows the user to adapt his/her physical activities to improve own health condition.

Keywords Personal health care system · Social network · Collaborative techniques · Classification algorithms

V. Trajkovik (✉) · I. Kulev
Faculty of Computer Science and Engineering,
University "Ss.Cyril and Metodius",
Rugjer Boshkovikj 16, P.O. Box 393,
1000 Skopje, Macedonia
e-mail: trvlado@finki.ukim.mk

I. Kulev
e-mail: Igor.Kulev@finki.ukim.mk

E. Vlahu-Gjorgievska
Faculty of administration and information systems management,
University "St.Kliment Ohridski",
Partizanska bb,
7000 Bitola, Macedonia
e-mail: elena.vlahu@uklo.edu.mk

1 Introduction

Advances in communication and computer technologies have revolutionized the way health information is gathered, disseminated, and used by healthcare providers, patients, citizens, and mass media. This led to the emergence of a new field and new language captured in the term "e-health" [1].

The importance of healthcare to individuals and governments and its growing costs to the economy have contributed to the emergence of healthcare as an important area of research for scholars in business and other disciplines [2]. The solution to decrease the cost of healthcare services and the load of medical practitioners requires necessary changes: moving from reactive to preventive medicine, concentrating on the long term care rather than only acute care, patient-centered care rather than hospital centered care, including remote care delivery mechanisms where the patient is taking a bigger role in his/her treatment and lifestyle management [3]. So, in addition to the embedded role of information technology in clinical and diagnostics equipment, information systems are uniquely positioned to capture, store, process, and communicate information to decision makers for better coordination of healthcare at both the individual and population levels [2]. The recent trend in healthcare support systems is the development of patient-centric pervasive environments in addition to the hospital-centric ones [4]. Such systems enable healthcare personnels to be able to timely access, review, update and send patient information from wherever they are, whenever they want [5].

In that way, pervasive health care takes steps to design, develop, and evaluate computer technologies that help citizens participate more closely in their own healthcare [6], on one hand, and on the other to provide flexibility in the life of patient who lead an active everyday life with work, family and friends [7].

The collaborative information system model (COHESY) we present in this paper, gives a new dimension in the usage of novel technologies in the healthcare. This system use mobile, web and broadband technologies, so the citizens have ubiquity of support services where ever they may be, rather than becoming bound to their homes or health centers as pointed out by different authors [8]. Broadband mobile technology provides movements of electronic care environment easily between locations and internet-based storage of data allows moving location of support. Cohesy has simple graphical interfaces that provide easy use and access not only for the young, but also for elderly users. System model has more purpose and includes use by multiple categories of users (patients with different diagnoses). Some of its advantages are scalability and ability of data information storing when communication link fail. Cohesy is interoperable system model that allow data share between different systems and databases.

Cohesy use collaborative algorithms that are implemented in the social network [9]. These algorithms allow connecting users with same or similar diagnoses, sharing their results and exchanging their opinions about performed activities and received therapy. At the same time, they also generate average values based on filtering large amounts of data about concrete conditions as are geographical region, age, sex, diagnosis, etc. The different levels of the validity of the data used by the collaborative algorithms and thus the validity of obtained results are elaborated in more details in [10].

The most important benefits of our proposed system model are: increased medical prevention, more immediate time response at emergency calls for doctors, 24 hour monitoring of the patients' condition, possibility for patient notification in different scenarios, transmissions of the collected biosignals (blood pressure, heart rate) automatically to medical personnel similar to the work of Komnakos, Vouyiokas, Maglogannis and Constantinou [11], increased flexibility in collecting medical data. Our system model creates the opportunity for increasing patient health care within their homes by 24 hour monitoring on the one hand, and increasing medical capacity of health care institutions on the other hand. This results in reducing the overall costs for patients and hospitals and improves the patient's quality of life [12].

We believe that the usage of the social network including the algorithms for classification and the collaborative techniques are the main components and advantages of our system which differentiates it from other health care systems. These components provide a new perspective in the use of information technologies in pervasive health care and make Cohesy more accessible to users. Meanwhile, the algorithm for recommendations bridges the gap between users, clinical staff and medical facilities, strengthening the trust between them and providing relevant data from a larger group of users, grouped on the basis of various indicators.

Related work is the next section. A brief overview of the system model is given in the third section. Collaboration techniques are discussed in the fourth section and in the fifth section some classification algorithms are given. The sixth section explains the design of the recommendation algorithm. Implementation remarks are given in the seventh section. Security issues are covered in the eighth section, while the ninth section concludes the paper.

2 Related work

There are many companies, researches, laboratory working on solutions for better health care of patients and systems that will help to have continuous monitor of the health of the patients. Pervasive and context-aware applications have been generally recognized as promising solutions for improving quality of life for the patients, which means integration of wireless sensor networks with pervasive computing devices in the creation of intelligent environment for providing unobtrusive monitoring, prompting or reminding of desirable activities and correcting or assisting the patients in their daily life [13].

MobiCare is a remote wireless patient monitoring system that provides better healthcare services [5]. It has three important building blocks: a body sensor network consisting of wearable sensors and actuators, a Body sensor network Manager (MobiCare client) that connects the network to a wide-area communication interface (using cellular wireless link) and back-end infrastructure support (MobiCare servers) at healthcare providers to implement necessary healthcare functionalities. This system enables healthcare personnels to be able to timely access, review, update and send patient information from wherever they are, whenever they want. Therefore, use of such mobile healthcare system can lead to significant economic benefits and cost savings for a patients as well as for the clinics.

Similar system, that facilitates the remote monitoring of the patients, is Personal Care Connect [14]. Personal Care Connect covers the connection and communication between the patient (with biomedical devices), data collection hub, medical server and medical personnel. Main advantage of Personal Care Connect is that it is a standard-based open (client and server) platform for interfacing to a wide variety of biomedical devices and sensors, collecting data from the devices and sensors, storing the data in a server repository, and making the data available to applications through a documented API, that allows Clinical-information system developers to have a uniform view of a wide variety of medical sensor data while have no need to know how to connect to each new device that comes on the market.

Unlike MobiCare and Personal Care Connect, our model is extended and includes communication between patient (with bionetwork—body network sensors), mobile applications, social networks, clinical centers and policy makers, which allows greater complexity of the data, extensive analysis and satisfying reliability.

Shopov, Spasov and Petrova [15] provide an overview of the building blocks and architecture models for Web-based Personal Health Systems. They are distinguishing several major blocks in Personal Health Systems: Network of biosensors, Personal gateway (Personal server), Clearinghouses, Medical Servers, and Medical Web Portal. Some of these blocks have similar features with elements used in our model, but some blocks, for example Clearinghouses, we did not use at all. Personal Server is similar to Mobile application explained in our model. It can do a local reasoning to determine user's health status based on data from multiple sensors and provide feedback through a user-friendly and interactive graphical or audio interface. But, the Mobile application in our model is wider because it also provide an emergency call (based on user condition) and enables the user to communicate with social network's users and use others patients experience.

The need for quality of service support in wireless e-health and e-emergency services is discussed by Gama, Carvalho, Alfonso and Mendes in [16]. They had analyzed some projects and the quality of service requirements that the respective authors of that projects have considered important to incorporate in their implementations. Firstly Gama, Carvalho, Alfonso and Mendes are considering the vital signals monitoring and suggest enhancing sensor node intelligence, available memory, processing power, and only enable on-line solicited requests for results in order to reduce the traffic load and the power consumption of a body sensor network. According to them, quality of service can be considered as guarantee for the right number of sensors required for monitoring the vital signals according to the patients' emergency state. They emphasize that the network must prioritize the transmission of critical data when sudden change occurs in the patient medical condition, and because of this it is important to distinguish all collected information. Available energy in the body sensor network is another very important parameter, because if energy is carelessly consumed, the body sensor network may rapidly become completely useless by lack of power. To save battery power, they recommend reduction of the sampling rate of the sensors while patient is in normal state or if the battery charge becomes low then its energy to be reserved to the more vital tasks of the patient. In order to provide a pervasive and trustable assistance to patients under health risk, they underline that, Personal Health Systems have to provide quality of service support such guarantee for bandwidth, reliability, end-to-end delay, jitter and loss. In our work we do not consider the quality of service. But, because of its importance in future work we will consider it and we will implement some quality of service in presented Collaborative Health Care System Model.

Most similar with our system model is Jog Falls system [17]. Jog Falls system is an end to end system to manage diabetes that blends activity and energy expenditure monitoring, diet-logging, and analysis of health data for patients and physicians. This is an integrated system for diabetes management providing the patients with continuous awareness of their diet and exercise, automatic capture of physical activity and energy expenditure, simple interface for food logging, ability to set and monitor goals and reflects on longer term trends. Its interface gives physicians comprehensive and unbiased visibility into the patients' life styles with respect to activity and food intake, as well as enabling them to track their progress towards agreed goals. The main emphasis authors place on its novel method for fusing heart rate and accelerometer data that improves the accuracy of energy expenditure estimation (a key feature in enabling weight loss). The first tier in Jog Falls is consisting of the sensor devices responsible for collecting the physiological and activity data. The second tier consists of a smart phone, which is responsible for communicating with the sensors via bluetooth, aggregating and storing the sensor data, calculating the energy expenditure and intake, providing the user interface for logging, alarming and data review, and communicating with the third tier through GPRS. And the third tier in Jog Falls is consisting of a backend server that is responsible for aggregating and storing the data from all users, and providing the user interface for the physician.

Although the three tiers of the Jog Falls system are like the first and the third layers in Cohesy, this system has no social network and collaborative algorithms, which are the main advantages of our system.

We must emphasize that the presented system model in our work consists of three levels and several different parts that mutually exchange information, which is not the case in the aforementioned examples. These examples include detailed research of systems that cover only parts of the presented model. Indeed, such comprehensiveness of our model is its advantage. Namely, linking different parts and their communication, broads their exchange of information which leads to more extensive data about various treatments, therapies and activities of patients with the same or similar diagnoses. This leads to further analysis and research that would result in improved treatment of patients with these diagnoses.

3 System layers

The mobile technologies (devices and applications) in this system are used to support and enable collaboration. The installed mobile application, using various sensors (bionetwork), performs readings regarding users health parameters (e.g. blood pressure, blood-sugar level, heart rate, weight) during his physical activities (e.g. walking, running, cycling) and based on them, gives appropriate instructions, proposals and constraints of their execution, in order to improve his own health.

The installed mobile application has access to the social network where it can store users' data and read average data readings on bio and physical activities of all users. Social network allows direct communication between users (if approved by the user and stored in the user profile) and sharing their results. This network can provide interface and use data from a variety of environmental databases (e.g. temperature, wind speed, humidity).

The medical personnel can remotely monitors the patient's medical condition, reviewing the medical data (health parameters) arriving from the mobile application of the patient. In this way, medical personnel can quickly respond to the patient by suggesting most suitable therapy as well as when to receive it, focusing on activities that are necessary for his rehabilitation and maintenance of his health, sending him various tips and suggestions for improving his health. The social network can learn from this recommendation and generate notifications and recommendation based on the most successful scenarios.

Simultaneously, clinical centers, medical databases and policy makers can exchange data and information with a social network and thus have access to a larger group of patients that can share research, recommendation and suggestion of the medical personnel. The social network has incorporated collaborative filtering that allows filtering large amounts of data on concrete condition. This complex structure of data from a social network along with the data arriving from different clinical centers can be used by medical databases for further analysis and research.

Simple overview of our system model is presented in Fig. 1. Generally, this system is deployed over three basic usage layers. The first layer is consisting of the bionetwork (implemented from various body sensors) and mobile application (that collects users' bio data and health parameters). The second layer is presented by the social network implemented as a web portal which enables different collaboration within the end user community. The third layer enables interoperability with the primary/secondary health care information systems which can be implemented in the clinical centers, and different policy maker institutions.

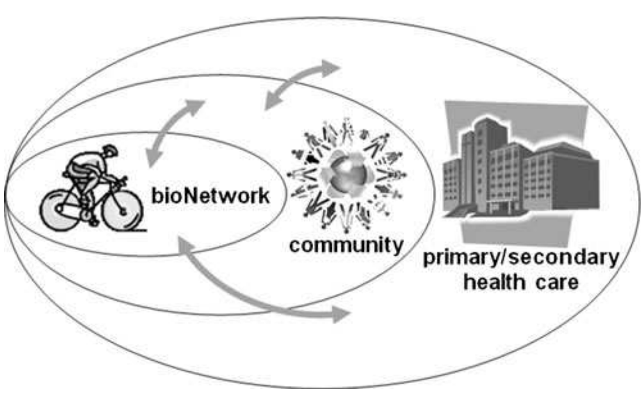

Fig. 1 Cohesy system layers

Communication between the first and the second layer is defined by users' access to the social network where user can store their own data (e.g. personal records, healthcare records, bionetwork records, readings on physical activities). Social network allows communication between users based on collaborative filtering techniques, thus connecting the users with the same or similar diagnoses, sharing their results and exchanging their opinions about performed activities and received therapy. Users can also receive average results from the other patients that share same conditions in a form of recommendation or notifications. These notifications can vary from the average levels of certain bio data calculated for certain geographical region, age, sex to the recommendation for certain activity based on the activities of other user. Collaborative filtering can be used to achieve different recommendations in these contexts.

Communication between the first and the third layer is determined with the communication between patient and health care centers. The patient has 24 hour access to medical personnel and possibility of sending an emergency call. The medical personnel remotely monitors the patient's medical condition, reviewing the medical data (e.g. blood pressure, blood-sugar level, heart rate) and respond to the patient by suggesting most suitable therapy (if different from the one that is encoded in the mobile application) as well as sending him/her various notifications (e.g. tips and suggestions) regarding his/her health condition.

The second and the third layer can exchange data and information regarding a larger group of patients group by any significant indicator (e.g. region, time period, sex, type of the activities) which can be later used for research, policy recommendations and medical campaign suggestions.

4 Collaborative filtering techniques

Collaborative filtering is a process of separation or evaluating of "objects" using the opinions of a group of people. The expression collaborative filtering is used more than a decade, although this method of filtering data people are using for centuries [18].

Namely, people often talk with each other and discuss about books they have read, movies they have watched, restaurants that have used. Based on the individuals who gave recommendations as well as on their interests, people choose whether the recommendation is good or a bad one. Over time, people "learn" which recommendations should be respected (and important) and how these recommendations help in determining the quality of the facility. Collaborative filtering is a tool that by using the speed of computers allows us to process these recommendations in real time and to decide not only what large group of people thinks of an object, but also to develop personal attitude of that object using opinions that are relevant from a particular user or user group.

Collaborative filtering, as one of the most successful approaches in building a recommendation system, use the known interests of the group of users to generate recommendations or predictions for unknown interests of other users [19].

For the first time the term collaborative filtering occurs in the Tapestry system, after which this concept is widely accepted, no matter that those who produce the recommendation need not explicitly to collaborate with those who receive it, so recommendations themselves can suggest interesting objects that may need to be discarded. Tapestry [20] is one of the first systems to generate recommendations based on collaborative filtering, developed by XeroxPARC, which incorporates the activities and opinions of the users in the message database and recommendation system. Later, was promoted a recommendation system by Maltz and Ehrilch [21] which allows users who read the document and found it good to send it to other colleagues to see. After that, on the market appear more collaborative filtering systems in different areas like: GroupLens [22, 23], Ring [24], Bellcore's Video Recommender [25]. These systems use the evaluations of users in order to compute the similarity or weight between users or objects and on the basis of these calculations generate predictions or recommendations. In these systems grouping of users is performed, according to their specific characteristics (tastes or interests). Meaning, the system finds "neighbors" (users with similar tastes) for each user and based on the given rates by the "neighbors" for certain object(s), the recommendation for the actual user is formed.

Collaborative filtering systems generate predictions or recommendations for a given user, recommending him one or more objects. The object is something for which the user can give his rate such as books, movies, CDs, magazine, photograph or holiday offer. That is, collaborative filtering techniques use databases of characteristics and interests for objects and users in order to predict a new range of facilities that would be of interest to users. Generally, there are m $\{u_1, u_2, .., u_m\}$ users and a list of n $\{i_1,i_2,..,i_n\}$ objects, so that each user u_i has a list of object I_{ui} that the user has evaluated or for which user has shown interest through his own behavior.

Rating techniques in collaborative filtering may be explicit or implicit. Explicit rates are those that the user provides when asked to give a direct opinion on the subject. These opinions may have different forms: Scalar ratings consist of numerically rates (e.g. from 1 to 5) or descriptive rates (e.g. very good, good, average, bad, very bad); Binary rating is the choice agree/disagree or good/bad; Unary rating indicates that the user has reviewed, purchased or otherwise positively evaluated an object. The lack of rating indicates that there is no information related to the specified user and object. Implicit evaluations are those that are generated on the basis of activities that the user makes. For example, if a user visits a web page for a specific products means he is interested in that product, while if a user order the product means he has a lot more interest for that product.

In recent years collaborative filtering algorithms have evolved from algorithms for tracking user interests into algorithms that have performance of large commercial applications. Most collaborative filtering algorithms are divided into: memory-based algorithms in which all rates, objects and users are stored in memory, model-based algorithms that periodically generate overall ratings offline and hybrid algorithms. In terms of organization collaborative filtering algorithms are divided into non-probabilistic and probabilistic algorithms.

Memory-based CF algorithms use the entire set of user-object ratings in generating a recommendation or prediction. Each user in these algorithms is part of a group of users with similar interests. Identifying the so-called neighbors of the actual user generates a recommendation or prediction for that user. Memory-based algorithms are easy to implement and have good performance when we have dense databases. Deficiency in these algorithms is their dependence on user ratings, have bad performance when it comes to sparse databases and have limited scalability with very large datasets. The most known memory-based algorithm is k-nearest neighbors.

Model-based collaborative filtering techniques use the raw data for ratings to estimate or to learn a model that will help them to generate predictions. The design and development of model-based techniques are based on models such as machine learning or data mining, which allow the system to recognize complex samples based on previously established training set and then make intelligent prediction for test or real set of data, based on learning model. These models are explored in order to solve the shortcomings of memory-based models. Shortcomings of the model-based techniques are their impracticality when it comes to extremely sparse data, use of dimensionality-reduction techniques or transformation of several class data into binary (leading to reduced performance in generating recommendations), high cost of creating the model and the loss between predicted performance and scalability of some algorithms. Some model-based techniques are Baesyan belief Nets [26–28], clustering CF models [29], Latent semantic models CF [30].

Most hybrid collaborative filtering techniques combine collaborative filtering algorithms with context-based techniques or other recommendation systems to minimize the shortcomings and limitations of both systems and to improve the characteristics of the predictions and recommendations. Besides improved performance hybrid collaborative filtering techniques are based on external content information which are usually inaccessible and generally have increased complexity. As an example of hybrid collaborative filtering techniques we would mention content-boosted collaborative filtering algorithm [31] and Personality Diagnoses (PD) [32].

Most often, any referral system that provides rapid and precise recommendations would be very attractive and profitable for companies. For the collaborative filtering system to provide high quality recommendations or predictions, the system has to serve all the challenges and demands of the given tasks. Collaborative filtering algorithms should be able to operate with highly sparse data, to perform calculations as the number of users and objects grows, to generate good and accurate recommendation in a short period of time and deal with other problems such as synonyms (when same or similar objects have different names), shilling attacks, data noise, problems of private protection.

It is always desirable to design a collaborative filtering system that would be easy to implement, do not use many resources, provide accurate predictions and recommendations and solve any challenges, such as data sparsity, scalability, synonyms, in the real application. These are the goals we want to accomplish in our collaborative information system model Cohesy.

5 Classification algorithms

Classification in machine learning is defined as the identification of sub-population to which belong new observations. The classification is done based on training set of data containing observations which sub-population are known, while the identity of the classified sub-population is unknown. In this manner, according to the training set in which there are already made groupings, new objects are placed in groups according to quantitative information for one or more characteristics. To perform categorization, in our case, it is necessary to determine on which attributes (medical diagnosis, height, weight, region, blood pressure, blood-sugar level) the categorization of users will be performed.

The most used classifiers are naive Bayes algorithm, k-nearest neighbors, the decision trees, neural networks and support vector machines.

The naive Bayes classifier is simple probabilistic classifier based on the application of Bayes theorem with assumptions of independence between the characteristics by which classification is done [33]. In our case it can be assumed normal distribution for each of the attributes of the users (parameters can be approximated with relative frequencies from the training set) and after using data from the training population will determine which are the parameters of the normal distribution for each attribute, the posterior probability for each of the class labels will be calculated. The new user will be placed in the class where posterior probability is at maximum.

The algorithm k-nearest neighbors is used for classification of objects according to the closest training samples in the space defined by the characteristics and is one of the simplest algorithms in machine learning [34, 35]. Namely, the object (in our system user) is classified by majority “votes” of its neighbors and the object is assigned to the class that occurs most frequently among its k-nearest neighbors. Training set represents a set of labeled samples. In this algorithm there is no training phase, all labeled samples participate in the process of class decision for the new samples.

The decision trees are predictive model that maps observations about an instance in the conclusions of its class [36]. In these structures, leaves represent classifications and branches represent the terms of the characteristics that lead to these conclusions. The decision tree can learn by dividing the set of instances to subsets according to the test of the value of an attribute [37]. This process is repeated on each new subset, recursively. Recursion is completed when a subset of a node has the same class label.

One of the best classifiers, which can be used in our system, is the neural network. Each of the values of attributes of users should first be normalized within the limits from 0 to 1. Then the architecture of the neural network is defined. It should have an equal number of inputs as the number of attributes of the user (based on which classification is performed) while the number of outputs should match the number of class labels. The labeled population should be split into training, test and validation set. First the neural network should be trained using a training population in order to learn its weights [38]. It is necessary to perform cross-validation to determine the optimal number of iterations for the training of neural network. The output of the neural network should correspond to the class label for specific input. The label that corresponds to the output with maximum value is defined as the class label of the new sample.

Support vector machines are a concept for set of supervised learning methods that analyze data and detect patterns [39, 40]. They belong to the group of non-probabilistic binary linear classifiers. Support vector machines construct hyper plane or set of hyper planes in the multidimensional space which can be used later for classification. Good separation is achieved by hyper-plane that has largest distance to the nearest instance of any class, because with the increasing of the margin reduces the error of the classifier.

In Table 1 are given the advantages and disadvantages of five classsification techniques. We have explanations for every techiques, why it can or cannot be used as classifier in our system model.

We must emphasize that in the classification for our system model, belonging of a given object to a particular class will be presented with fuzzy logic.

Table 1 Advantages and disadvantages of several classifiers

Classifier	Advantages	Disadvantages	Use in our model
naive Bayes classifier	• probabilistic model • fast • space efficient (especially when the sample size is small) • robust to noise found in real data	• not capable of solving more complex classification problems • when the classification parameters has dependency between, NBC cannot be used	• Can be used in our model because of its simplicity, speed and space efficiency. In our model classification parameters are independent.
k-nearest neighbor algorithm	• simple to understand • easy to implement and debug • interpretabile	• can have poor run-time performance if the training set is large • very sensitive to irrelevant or redundant features	• Cannot be used in our model because of its slow computing time and the need to examine the entire training population in order to determine the class label of the new sample.
decision tree	• low computing time • interpretabile which leed to rapid classification • has the flexibility of choosing different subsets of features at different internal nodes of the tree • use a smaller number of features at each internal node	• overlap especially when the number of classes is large • two internal nodes contain at least one common class • error accumulation from level to level in a large tree • problem of minimizing the expected number of tests required to classify an unknown sample	• Can be used in our model because of its low computing time, but we must consider to have a smaller number of classes.
neural networks	• fast training process • can handle classification with very many parameters • training samples can be added or removed without extensive retraining • guaranteed to converge to an optimal classifier as the size of the representative training set increases • high performances on very difficult classification tasks	• notoriously slow (slow execution of the network) • large memory requirements • requires a representative training set	• Can be used in our model because of its high performance to handle with many classification parameters. We have to consider the slow execution of the network.
support vector machine	• robust • can work well with a small training dataset • can process high dimensional data • high performances on very difficult classification tasks • provide a good out-of-sample generalization	• limited speed • extensive memory requirements • high algorithmic complexity	• Cannot be used in our model because of its limited speed and large memory requirements.

6 Proposed design for the recommendation algorithm

Recommendations module, which is part of the social network (implemented as a web service), generate recommendations for users to carry out physical activities to improve their health.

This web service uses data read by bionetwork, data for the user's physical activity (get by the mobile application), medical records of the user (obtained from clinical centers) and data of the user profile on social network (so far based on knowledge of social network).

The main purpose of this algorithm is to find the dependency of the users' health condition and physical activity he/she perform. To achieve this we consider datasets from the health history of users and use classification algorithms on these datasets for grouping the users based on their similarity. This section will explain the algorithms used for functionality of this web service.

6.1 Basic concepts for the recommendation algorithm

The recommendation algorithm is based on the dependence between the values of user's health parameters (heart rate, blood pressure, blood-sugar level, weight) and his physical activities (walking, running, cycling). All activities between two readings of the same health parameter (e.g. blood pressure) affect the resultant change of that health parameter a little or a lot. The aim of the recommendation algorithm is to discover which activities affect change in the value of each health parameter individually. Once revealed, algorithm can use that information in situations it recognizes as same or similar to previous health conditions of a same or another user with similar medical problems. If there is information, in users' history, that after execution of a physical activity, the health situation had already changed for the better, it can be concluded that this activity can help him or other users with similar health condition and improve their health condition.

The periods between every two consecutive readings of the value of a particular health parameter are taken as intervals for testing physical activities that contribute changing the health status. All activities that occur within this period and determine their effect in the change of the health status are considered. Each parameter is viewed independently of others and all activities that affect him during the two consecutive readings are monitored.

In systems where cannot be made strict separation between the conditions of the variables, it is good to use fuzzy logic. Therefore, it is necessary to bring fuzzy logic in our system for recommendations, in order to better represent the reality.

6.2 Categorization of users

The main data used for recommending is derived from the data of users within the social network. Such database stores all the conclusions about changes in the parameters depending on the specific activities that are obtained from medical history and history of activities of social network users. This database gathers the knowledge of all users, but these data are now categorized and classified.

Use of classified data when generating the recommendation provides more relevant recommendations, because they are enacted on knowledge for users with similar medical conditions and reference parameters (Fig. 2). To perform the classification of the users, it is necessary to build profiles of users on which basis the classification will be done. There are several different ways of representation of users. A (simplest) representation of the user may be in one single form (such as: diagnosis, age, weight, height). Moreover, users who have similar values for these parameters will be grouped together.

Above we have explained several algorithms for classification. Based on their advantages and disadvantages we proposed which algorithms can be used in our system model.

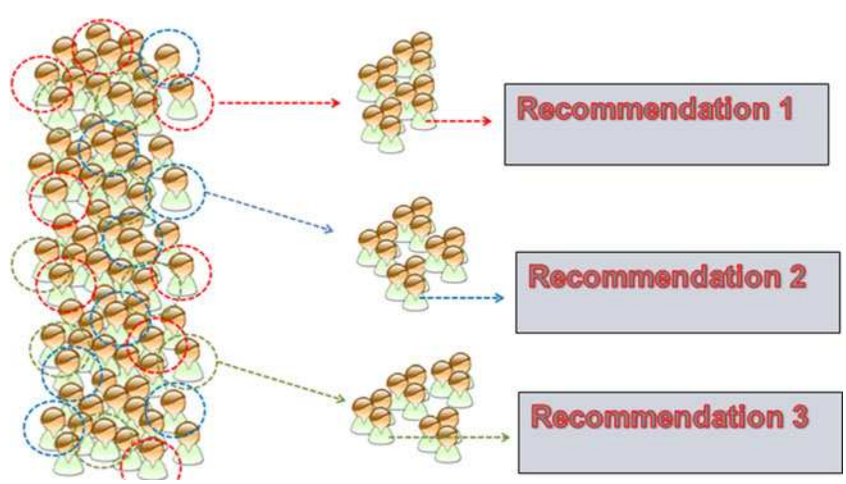

Fig. 2 Classification of the system users

6.3 Recommendation algorithm

Recommendation algorithms are very important element in the Cohesy model. The purpose of recommendation algorithm is to give a recommendation for performing a specific activity that will improve user's health, based on his given health condition and set of knowledge derived from the history of the user and users like him. The algorithm is shown in Fig. 3.

For each activity from the set of activities, efficiency is calculated before it is recommended. This efficiency is determined by points which are assigned to that activity. The argument of the recommending function is the current given health condition of the user in the form of values of the health parameters. All knowledge is gained from the history of the user that uses and the histories of other users with similar health parameters. This knowledge is filtered twice. First it is filtered according to whether the activity involved in the knowledge is equal to the activity which is currently under calculation of efficiency function. Then it is tested whether the health parameter, that is monitored, is disturbed. This function is shown in Fig. 4. If both conditions are met, points are added or subtracted, according to whether the activity would contribute to the improvement or deterioration of the current health condition of the user.

The test function whether the health parameter is disturbed checks whether the current value of that parameter is in the normal range or not. To determine the health parameter disturbance we use fuzzy logic, or algorithm that will deliver value in the interval [0, 1]. Test function for efficiency of the activity checks whether the direction of change in the value of the health parameter, which is in the knowledge base, is required to normalize the value of the disturbed parameter.

For example, if in the knowledge base is preserved that the old value was A, and the new value was B, and we know that A<B then this activity has contributed to increase the value of the parameter. This is useful when the value of the disturbed parameter is less than the lower normal limit, but has a bad effect if the value of the disturbed parameter is greater than the upper normal limit. In this case the activity would contribute to further disruption of the health parameter, and thus the overall health of the user. Therefore points are deducted for this activity.

Function to calculate the points for an activity should give a value which depends on the change in the value of the health parameter and the period of change, range of values for that parameter and intensity of activity. The activity is more effective if the change is greater and if the period is smaller. In practical implementation period is measured in seconds and for less sensitivity of function vs. period, the change in value is divided by logarithm of the period, rather than directly with the period. The final formula is:

$$\frac{\left(\frac{\text{new\_value}-\text{old\_value}}{\text{upper\_limit}-\text{lower\_limit}}\right)}{ln(\text{period})} \cdot \text{intensity} \cdot \textit{coefficient}(\text{parameter})$$

The formula uses coefficient that indicates the importance of the health parameter. That is, if a parameter is more important and change of its value would be critical to the health of the user, then its coefficient should be larger. Once we get the number of points for each activity, we check if the maximum number of points is less than or equal to zero. If this condition is met, it means there is no activity that would be treated favorably in terms of improving the health condition of the user. In this case there is no returned recommendation. Otherwise, it

Fig. 3 Recommendation algorithm for a given current situation and set of knowledge

```
get_recommendation (condition current_condition, Activities, Knowledge_base)
begin
        for each activity in Activities
        begin
        points[activity] = 0
        for each knowledge in Knoledge_base where knowledge.activity == activity
        begin
                p = knowledge.parameter
                if (disturbed_parameter (current_condition, p) == yes)
                begin
                        if (efficient_activity (current_condition, knowledge) == yes)
                                points[activity] += calculate_points (knowledge, p)
                        else
                                points[activity] -= calculate_points (knowledge, p)
                end
        end
end

if max_sum_points (points) <= 0
        return nothing
else
        begin
                best_activity = activity_with_max_points (points)
                return choose_recommendation (current_condition, best_activity, Knowledge_base)
        end
end
```

```
disturbed_parameter (condition c, parameter p)
begin
        if c.value_of_p < p.lower_normal_limit
                return yes
        else
                if c.value_of_p > p.upper_normal_limit
                        return yes
                else
                        return no
end
```

Fig. 4 Test function for disturbance of a given parameter

selects the action with the highest number of points. According to the current health condition of the user, the recommendation is chosen from the knowledge base corresponding to the activity and the appropriate intensity of the activity.

7 Implementation remarks

The communication between the first and second level of Cohesy is implemented within the framework of SportyPal mobile collaborative system (http://www.sportypal.com) [41]. It has approximately 450000 active users that are connected to its dedicated social network.

SportyPal system is capable of reading parameters for a particular activity, such as path length, speed, time interval, consumed calories. With the help of the GPS service on the mobile phone, SportyPal application reads and writes a map of the path by which the activity is executed. By using an additional device that connects with the application, it can read health parameters of the user (e.g. blood pressure, blood-sugar level, weight and currently heart rate). Each execution of a specific operation is stored as a separate workout (Fig. 5). The user has access to all of his stored workouts and thus he is able to analyze and compare them later. SportyPal offers possibility to present each exercise in a map view, draw graphics charts and present its summary information.

The SportyPal system includes active social network at SportyPal.com where users can upload their results (Fig. 6). This social network additionally allows users to analyze their results, to compare them with the results of other users, to comment all results and to organize virtual competitions.

From this point of view, current functionality of SportyPal system corresponds with the services we offer in Cohesy, especially in terms of tracking the user's physical activity. Although we must stress that there are several differences which can be easily upgraded in SportyPal.

Like the application that we propose in Cohesy, the application that is actively used in SportyPal allows reading of several health parameters: heart rate, blood-sugar level,

Fig. 5 Application SportyPal displays details (current values of the parameters) about current users' exercise

blood pressure and weight. SportyPal application still has no publically available module that will signalize if an irregularity occurs while reading health parameters for the user. Also, this application has no communication with a medical center so it does not offers possibility of sending an emergency call, for sudden deterioration of patient's medical condition, to the medical staff. In terms of social network there is still not publically available connection of SportyPal.com with environmental database, so there is no data for weather conditions (temperature, wind speed, humidity) in which the activity was performed.

Although SportyPal system has basic profile information about users, such as height and weight, this system does not have data for a possible diagnosis and therapy of the user.

We would also like to emphasize that one of the advantages of Cohesy, its recommendations module, is still not implemented in the SportyPal system. However, we are convinced that in our further work, we will be able to implement the module for generating recommendations based on the recommendation algorithm that we have proposed, and will succeed to evaluate its functionality and benefits through the SportyPal system. Before that, we plan to implement specific collaborative algorithms within Cohesy in order to get some summary data, grouping the users by certain parameters.

Because currently we have no accesses to medical records and have no established cooperation (communication) with

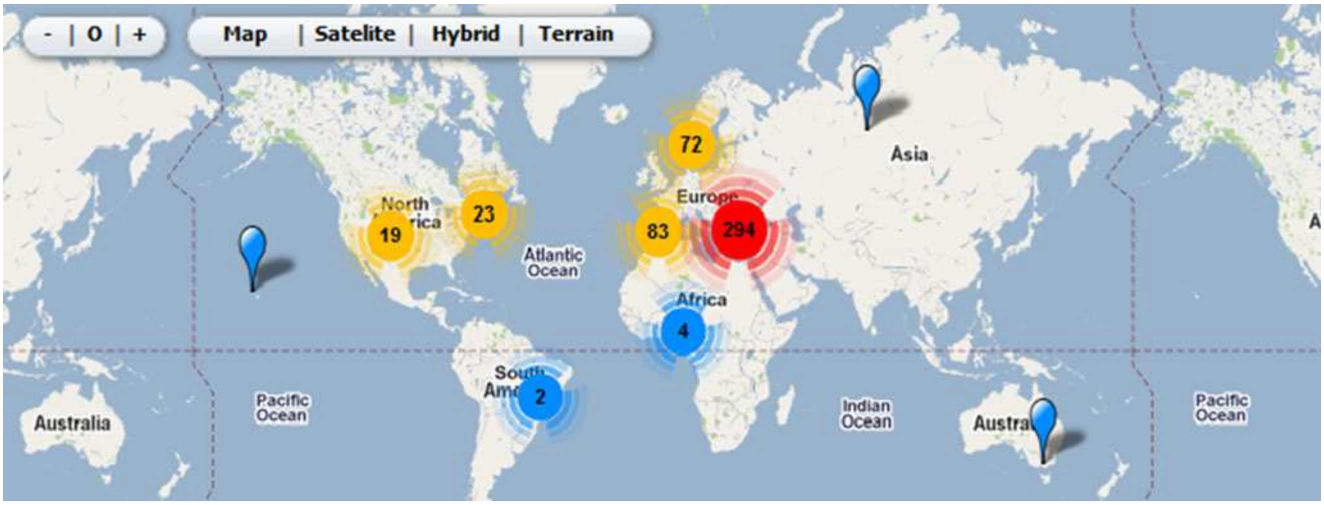

Fig. 6 SportyPal.com displays the latest 500 workouts from world

the information systems used in institutions of primary and secondary health care, we are not able to present practical results for the 3rd Cohesy level and its communication with the first and the second level. In the coming period we expect to realize the possibility of cooperation with certain institutions of primary and secondary health care which would open the opportunity for full exploration of the complete Cohesy architecture and analysis of the overall benefit from this collaborative information system.

8 Security issues

The fundamental goals of secure healthcare systems are safely exchanging the patient's information issued by mobile devices, and preventing improper use of illegal devices, such as intercepting transferred data, eavesdropping communicating data, replaying out-of-date information, or revealing the patient's medical conditions. Based on the potential threats of mobile healthcare [7], specific security requirements will have a significant influence on the performance of the system:

1) *Data Storage and transmission*: Local database (in mobile phone) stores data received by sensors, in case there is always back up of data (they will be saved only some period of time). When there are problems in sending data to clinical center, some of data is not going to be send, all transaction will be rolled back. In this way there will be always all sensors data and when service will be available data will be sending provide quality of service (QoS) facilities since these clearly demand for high reliability, guaranteed bandwidth and short delays [16].
2) *Data Confidentiality*: Most patients do not want anyone to know their medical information, except their family doctor or medical specialist. The solutions are to use a cryptographic algorithm to encrypt medical information and protect the necessary data.
3) *Authentication*: Only an authenticated entity can access the corresponding data that are available for that entity; unauthenticated entities are denied access when they visit data information that they do not have the rights to obtain. For example, asymmetric cryptography (i.e. PKI) is often used, because these private keys are credentials shared only by the communicating parties.
4) *Access Control*: In traditional network security models, access control determines whether a subject can access an object based on an access control list (ACL).
5) *Privacy Concerns*: Every user can choose what information can be private or public. User can choose his records to be public: (a) for medical purposes, (b) to all visitors of the Social network, (c) to users in his category, (d) to none. In order to have medical support the user has to agree to share personal information with clinical centers and medical databases, whose data are also protected. According to user agreement policy those data information would be exchanged through system.

Though many healthcare researchers are interested in collecting and recording medical sensor data, these data may contain many personal facts, meaning patients are not willing to reveal them [7]. Especially in an open wireless environment, an intruder may observe network traffic and thereby infer the relationships and identities of the communicating nodes.

One possible approach for solving this issue is to apply the theory of trust to identify malicious nodes and thereby exclude them from a presently healthcare network. As an emerging technique, trust can be defined as "the degree to which a node should be trustworthy, secure, or reliable during any interaction with the node" [42]. This means that if one node trusts another node to perform the intended operation, the trust relationship between these two nodes can be established reliably from the communicating initiator's point of view. So, the technique of trust evaluation without a centralized trust management authority can significantly improve the security and reliability of the network while also reducing the complexity of the traditional trust schemes and thus improving efficiency.

9 Conclusion

In this paper we are presenting Collaborative health care system model (Cohesy) implemented with help of the mobile and web technologies. The system provides tool for personal health care by generating different recommendation, notification and suggestion to the users. The recommendation algorithm, that generates these recommendations, uses some classification algorithms and different collaborative techniques. The generated information has different validity depending on the validity of the data that is used to generate information.

In addition to services for end users that our system model offers, the primary purpose of presented model is collecting different types of data and combining them into complex data structures based on collaboration. The survey, analysis and research of such structures allows to understand the impact and the influence of applied therapy, physical activity, time parameters and other factors on the development of the health condition of the patient.

Using knowledge the system model allows the user to adapt and align his physical activities while improving his health condition and overall way of rehabilitation, meaning to be fully able to take self care and professional concern about his health.

References

1. Viswanath K, Kreuter MW. Health disparities, communication inequalities and e-health: A commentary. Am J Prev Med. 2007;32 (5):S131–133.
2. Fichman RG, Kohli K, Krishnan R. Editorial overview-the role of information systems in healthcare: Current research and future trends. Inform Syst Res. 2011;22(3):419–28.
3. Laleci GB, Dogac A, Olduz M, Tasyurt I, Yuksel M, Okcan A. SAPHIRE: A multi-agent system for remote healthcare monitoring through computerized clinical guidelines. Agent technology and e-health. Babel, Switzerland: Birkhäuser Verlag; 2008. p. 25–44.
4. Koufi V, Malamateniou F, Vassilacopoulos G. A medical diagnostic and treatment advice system for the provision of home care. Proc. 1st international conference on PErvasive Technologies Related to Assistive Environments. USA: ACM; 2008. p. 1–7.
5. Chakravorty R. MobiCare: A programmable service architecture for mobile medical care. Proc. 4th IEEE Conference on Pervasive Computing and Communications Workshops (PerCom 2006 Workshops). IEEE Computer Society; 2006. pp. 532–536.
6. Ahamed SI, Haque MM, Khan AJ. Wellness assistant: a virtual wellness assistant using pervasive computing. Proc. Symposium on Applied Computing. USA: ACM; 2007. p. 782–7.
7. Ballegaard SA, Hansen TR, Kyng M. Healthcare in everyday life: designing healthcare services for daily life. Proc. Conference on Human Factors in Computing Systems. USA: ACM; 2008. p. 1807–16.
8. Taylor C, Dajani L. The future of homecare systems in the context of the ubiquitous web and its related mobile technologies. Proc. 1st international conference on PErvasive Technologies Related to Assistive Environments. USA: ACM; 2008. p. 1–4.
9. Xin X, King I, Deng H, Lyu MR. A social recommendation framework based on multi-scale continuous conditional random fields. Proc. 18th ACM conference on Information and knowledge management. USA: ACM; 2009. p. 1247–56.
10. Vlahu-Gjorgievska E, Trajkovik V. Towards collaborative health care system model—COHESY. Proc. 3th IEEE Workshop on Interdisciplinary Research on E-health Services and Systems (WoWMoM 2011). IEEE Computer Society; 2011. pp. 1–6.
11. Komnakos D, Vouyiokas D, Maglogannis I, Constantinou P. Feasibility study of a joint e-health mobile high-speed and wireless sensor system. Proc. 1st international conference on PErvasive Technologies Related to Assistive Environments. USA: ACM; 2008. p. 1–7.
12. Zimmerman T, Chang K. Simplifying home health monitoring by incorporating a cell phone in a weight scale. Proc. 1st international conference on PErvasive Technologies Related to Assistive Environments. USA: ACM; 2008. p. 1–4.
13. Shou-ming M, Ru-chuan W, Ning Y. Using context prediction for elderly health monitoring in pervasive computing environments. Int J Digit Content Technol App. 2011;5(1):16–25.
14. Blount M, Batra VM, Capella AN, Ebling MR, Jerome WF, Martin SM, Nidd M, Niemi MR, Wright SP. Remote health-care monitoring using personal care connect. IBM Syst J. 2007;46(1):95–113.
15. Shopov M, Spasov G, Petrova G. Architectural models for realization of Web-based Personal Health Systems. Proc. International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing. USA: ACM; 2009. p. 1–6.
16. Gama O, Carvalho P, Alfonso JA, Mendes PM. Quality of service support in wireless sensor networks for emergency healthcare services. Proc. 30th Annual Internacional Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Computer Society; 2008. pp. 1296–1299.
17. Nachman L, Baxi A, Bhattacharya S, Darera V. Jog falls: A pervasive healthcare platform for diabetes management. Pervasive Computing. 2010;6030:94–111.
18. Schafer JB, Frankowski D, Herlocker J, Sen S. Collaborative filtering recommender systems. In: Brusilovsky P, Kobsa A, Nejdl W, editors. LNCS: The adaptive web. Berlin: Springer-Verlag; 2007. p. 291–324.
19. Su X, Khoshgoftaar TM. A survey of collaborative filtering techniques. Advances in artificial intelligence. 2009; 2009 (Article ID 421425):1–19.
20. Goldberg D, Nichols D, Oki BM, Terry D. Using collaborative filtering to weave an information tapestry. Communications of ACM. 1992;35(12):61–70.
21. Maltz D, Ehrlich E. Pointing the way: Active collaborative filtering. Proceedings of ACM Conference on Human Factors in Computing Systems. USA: ACM; 1995. p. 202–9.
22. Konstan JA, Miller B, Maltz D, Herlocker J, Gordon L, Riedl J. GroupLens: Applying collaborative filtering to usenet news. Communications of the ACM. 1997;40(3):77–87.
23. Resnick P, Iacovou N, Suchak M, Bergstrom P, Riedl J. Grouplens: An open architecture for collaborative filtering of netnews. Proceedings of the 1994 ACM conference on computer supported cooperative work. USA: ACM Press; 1994. p. 175–86.
24. Shardanand U, Maes P. Social information filtering: Algorithms for automating "Word of Mouth". Proceedings of Conference of

Human Factors in Computing Systems. New York: ACM Press; 1995. p. 210–7.

25. Hill W, Stead L, Rosenstein M, Furnas G. Recommending and evaluating choices in a virtual community of use. Proceedings of ACM Conference on Human Factors in Computing Systems. USA: ACM Press; 1995. p. 194–201.
26. Breese J, Heckerman D, Kadie C. Empirical analysis of predictive algorithms for collaborative filtering. Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann Publishers; 1998.
27. Miyahara K, Pazzani MJ. Collaborative filtering with the simple Bayesian classifier. Proceedings of the 6th Pacific Rim International Conference on Artificial Intelligence. Berlin: Springer-Verlag; 2000. p. 679–89.
28. Su X, Khoshgoftaar TM. Collaborative filtering for multi-class data using belief nets algorithms. Proceedings of the International Conference on Tools with Artificial Intelligence. Washington: IEEE Computer Society; 2006. p. 497–504.
29. Ungar LH, Foster DP. Clustering methods for collaborative filtering. Proceedings of Workshop on Recommender Systems at the 15th National Conference on Artificial Intelligence. Menlo Park California: AAAI Press; 1998. p. 112–25.
30. Hofmann T. Latent semantic models for collaborative filtering. ACM Trans Inform Syst. 2004;22(1):89–115.
31. Melville P, Mooney RJ, Nagarajan R. Contentboosted collaborative filtering for improved recommendations. Proceedings of the 18th National Conference on Artificial Intelligence. AAAI; 2002. pp. 187–192.
32. Pavlov DY, Pennock DM. A maximum entropy approach to collaborative filtering in dynamic, sparse, highdimensional domains. Advances in Neural Information Processing Systems. USA: MIT Press; 2002. p. 1441–8.
33. Rish I. An empirical study of the naive bayes classifier. Proceedings of International Joint Conference on Artificial Intelligence workshop on Empirical Methods in AI. AAAAI; 2001. pp. 41–46.
34. Pan JS, Qiao YL, Sun SH. A fast K nearest neighbors classif ication algorithm. IEICE Trans Fundam Electron Commun Comput Sci. 2004;E87-A(4):961–3.
35. Zhang ML, Zhou ZH. A k-nearest neighbor based algorithm for multi-label classifcation. Proceedings of the 1st IEEE International Conference on Granular Computing. IEEE Computer Society; 2005, pp. 718–721.
36. Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. IEEE Trans Syst Man Cybern. 1991;21(3):660–74.
37. Mehta M, Agrawal R, Rissanen J. SLIQ: A fast scalable classifier for datamining. Lecture Notes in Computer Science: Advances in Database Technology. 1996; 1057/1996:18–32.
38. RipleyBD. Pattern recognition and neural networks. Cambridge University Press; 1996.
39. Burges CJC. A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov. 1998;2(2):121–67.
40. Cristianini N, Shawe-Taylor J. An introduction to support vector machines. Cambridge University Press; 2000.
41. http://www.sportypal.com/Home/Overview, 2011.
42. Boukerche A, Ren Y. A trust-based security system for ubiquitous and pervasive computing environments. Comput Comm. 2008;31:4343–51.