Management of Multiple Data Streams in Sensor Networks for Medicine

Jan Sliwa and Emmanuel Benoist
Bern University of Applied Sciences, Quellgasse 21, CH-2501 Biel, Switzerland

Keywords: Sensor Networks, Data Management, Quality Assurance, Medical Research.

Abstract: In this paper we discuss the privacy aware management of data streams generated by wireless sensor networks (WSN) in medical applications. Such data are primarily used to ensure the direct health support for a specific patient. They can also be used for other purposes, like quality assurance of applied devices and comparative evaluation research of medical treatments. Considering the novelty of the WSN-based medical application, every information helping to improve the service is extremely useful. As privacy in medicine is an important issue, care should be taken to harmonize the legal requirements with the medical necessities. Data used for different purposes have different characteristics, different users and different access rules. Therefore the raw data stream has to be split and processed (condensed, anonymized, etc.) in various ways dependent on the use. The access to data has to defined differently for different output data streams.

1 INTRODUCTION

This paper shows the problems regarding data management in wireless sensor networks (WSN). We will concentrate our attention on the applications in the medicine. For brevity, we use the term 'sensor networks' but we should not forget that such networks often also contain actors (or actuators) - devices acting directly on the patient, like defibrillators or insulin pumps. In this paper however we focus on the incoming data streams. In our previous papers we have discussed the ethical aspects of enhancing human organisms with technical instruments (Sliwa and Benoist, 2011b) and presented this overwhelming development in a historical perspective (Sliwa and Benoist, 2011a). We present here a basic introduction to this subject, a broader overview (with further references) is given in the aforementioned papers.

The development in the area of medical WSN based systems is driven by some enabling technologies. New devices permit measuring physical and chemical properties that were until now difficult to treat. They can also act on the human body in new ways, like stimulating the heart or dosing a medicine. Local networks permit the devices to cooperate and wide area networking allows them to exchange data with a hospital or a data center. Reducing power consumption, energy harvesting and remote powering permit a long term functioning.

The basic configurations are: Body Area Networks (BAN) and Ambient Assisted Living (AAL).

A BAN consists of a set of wearable and implantable devices carried by the patient. They can measure the health parameters (sensors) or act on the body (actors or actuators). They exchange data with a data aggregator (typically a smartphone) that sends periodical reports to the hospital or alerts it in the case of an emergency.

In the AAL scenario the patient lives in a home equipped with stationary medical devices that provide a supporting environment. In this way an elderly person may live longer in his/her familiar environment and still have an adequate level of security.

In this paper, first we set the problem of the data management in a productive deployment of medical WSN systems. Then we explore the various uses of the data generated by them: direct health support, quality assurance and medical research. For each of these uses we discuss the characteristics of the related data stream and analyze to whom, when and how should they be available. Finally, we outline some research challenges related to the data management facing the specialists who develop, deploy and operate such systems.

2 SETTING THE PROBLEM

In this paper we want to discuss the aspects of the managements of the data generated by the sensor networks in medicine. Presently, medical applications of

Sliwa J. and Benoist E..
Management of Multiple Data Streams in Sensor Networks for Medicine.
DOI: 10.5220/0004325403990402
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2013), pages 399-402
ISBN: 978-989-8565-37-2

the WSN are in the experimental phase. The technical challenges are immense, therefore most efforts concentrate on solving specific problems, ranging from the design of the sensors resistant to a hostile chemical environment via studying the propagation of the electromagnetic waves in the human tissue, energy harvesting for a long-term unattended operation to the development of efficient signal processing algorithms - and many more. In the experimental phase, the systems are installed and monitored by skilled and motivated technical and medical specialists, including the designers themselves. The data sets are small and are accessed by a small number of trusted scientists.

After the large scale deployment the situation changes radically. The systems are installed and operated by less specialized personnel. Their operation is not supervised. Data sets are large and the number of possible users increases. It may be tempting to integrate the systems using the bottom-up approach, just combining the available elements in a working system. It seems however to be now the right time for some more general reflections, setting the goals and analyzing the requirements.

In the area of data management, we argue that following uses of the collected data should be considered:

- direct health support
- quality assurance
- medical research

A reference scenario, used throughout this paper, is a Body Area Network for treating heart problems. A hospital diagnoses the patient, defines a treatment based on a WSN system and delivers a set of devices to him/her. It may be just a wearable device handed out with a short instruction, or an implantable device that requires an operation. The data aggregator (a smartphone) is configured by the hospital staff. The used devices are well defined, the patients and their devices are registered. A patient is related to a main hospital handling his/her case, called here a home hospital, where the patient's records are stored and which is called in case of an emergency. Thanks to a cooperation with other hospitals, the patient may be also treated in a remote location, if an emergency occurs there. The hospital, or a group of hospitals, operate a registry that helps to evaluate statistically the quality of the implemented devices and the correctness of the applied treatments.

Our experience in managing medical data used for multiple purposes by various user's groups is based on the platform for medical registries that we have developed and operated for about 10 years (Röder et al., 2006). Its goal is to store the patients' records and to make the anonymized extracts available for research. The architecture based on the physical separation of personal and clinical data enhances the privacy protection (Sliwa and Benoist, 2012). Currently we input data using fairly static online forms. Nevertheless, this experience is useful for solving similar problems arising in the WSN field, regarding storing and protecting data, controlling the access and presenting in various ways for various user groups.

3 DATA FOR HEALTH SUPPORT

3.1 Data Characteristics

The basic goal of medical WSN systems is providing health support for specific patients. The delivered information can vary with respect to its quantity and temporal characteristics. It may consist of single values, signal waveforms or images. A local unit can process the continuous waveforms and extract features from them - either calculate general properties or detect events. Data may come from one or from many sensors, it may be also a combination of them or a sequence of events. For example, when a drug is automatically delivered into blood, the timing of its propagation and of other chemical reactions to it may be important.

Sensor data may just periodically report the health parameters with no special action required. They may also detect emergencies - in this case an action has to be taken. This action may be executed by actuators on the patient remotely controlled from his/her home hospital, or it may involve real people and equipment. In the latter case it is necessary to know the actual location of the patient, and the task has to be assigned to the nearest cooperating hospital. This example shows us that real-life deployment of a health support system requires important organizational measures and merely technical solutions are not sufficient.

3.2 Data Access

For regular treatments the patient's case can be handled by his/her home hospital. The difference in comparison to the traditional way is data transmission via a wireless channel and data storage that both occur automatically. Long distance data transmission from the data aggregator has to be protected like any secure wireless transmission. The problem is the transmission between the sensors/actors and the data aggregator. These miniaturized, low power devices cannot support the same level of security measures (like strong encryption) due to their limited capabilities.

When data reach the hospital, they are treated in a similar way as other data in a clinical information system and have to be protected accordingly. They are stored on disks accessible to the system operators and are visible to the personnel that normally is allowed to see them.

A different class of problems appear when an emergency service is needed in a remote location as in the example mentioned above. First, a network of cooperating institutions that agree for a mutual assistance has to be defined. Then, procedures of the transfer of the patient's history are necessary. The intervening team has to know all relevant information, and have it available fast. When the case is closed, the home hospital needs the update of the patient's record, data at the remote hospital have to be removed as no more necessary. Similar problems are being solved in the European e-Health Project epSOS[1].

4 DATA FOR QUALITY ASSURANCE

4.1 Data Characteristics

Data collected from the medical WSN can be used to evaluate the quality of the deployed equipment. Although they are thoroughly tested and formally approved, only the actual operation can give us information about long term results, unforeseen adverse reactions and rare incidents. If an unexpected event happens it may have various causes, just to name a few: mechanical problem - fixture loosened, part broken; fluid sensor dirty, nozzle clogged; poor usability of the user interface - display unreadable, small keys, dialog unclear; battery depleted too fast; no phone signal available; external attack.

Some faults can be diagnosed on the basis of the analysis of the sensor messages. In an optimal case, an intelligent device performs regular self tests and informs about the possible and actual problems, but often a problem analysis by a human is necessary. Industrial networks, like train control systems, periodically test their integrity. They have also well defined real-time properties. In a similar way, a health supporting system has certain temporal requirements, depending on the severity of the treated disease.

The named problems can be fixed in very different ways, like upload of a corrected software version, device replacement, device redesign or organizational changes. Therefore in order to keep track of the quality problems and solutions a registry is necessary.

Medical software based devices pose challenging problems to the statistical analysis. One of the problems is assuring a reasonable size of a statistical sample of comparable, uniform enough data. The products are often upgraded, therefore there are not so many identical devices deployed. Even when only a software bug on a device is corrected, the device behaves differently, so formally speaking, it is not the same device as before.

Moreover, a patient has not just one device, but a set of cooperating medical devices together with a specific model of a smartphone with specific phone applications loaded. Because of the interdependencies of the elements, any change of any element makes a different system, a different case to be evaluated. Because of this confused, dynamically evolving situation, the definition of the analysis processes requires an involvement of humans that understand the underlying problems.

4.2 Data Access

The analysis of the data for quality assurance needs no personal information about the patients, therefore should be performed on anonymized data. This is entirely true if we are interested just in the number of faults for a device model. If we want to find the cause of a problem, the procedure is more similar to an analysis of an airplane crash. Still, the name of the patient is not relevant, but we may need supplementary information, for example what is his/her diet, does he/she any physical exercises, what is his/her age, gender, education level. If in case of an emergency no help came, or it came too late, it may have been caused by a software bug on the device, a loss of phone signal, inefficient information flow at the hospital or no free ambulance. Therefore it may be necessary to perform the analysis of the case at a very detailed level, also with the access to the raw sensor data.

General statistical analysis is best performed by a team of independent researchers, trained in medicine and statistics, understanding information and communication technology. For a precise analysis of a technical fault, representatives of the producers may have to be involved. The scenario outlined above shows that an interdisciplinary problem needs an interdisciplinary solution.

We also see how important, and how difficult it is to define the rules what data should be available, to whom and under what conditions.

[1] http://www.epsos.eu/ *(visited on 2012-10-26)*

5 DATA FOR MEDICAL RESEARCH

5.1 Data Characteristics

Data collected by the WSN-based systems can be finally used for medical research. As the approval process for intelligent software-based devices is not as strict as for drugs (chemicals), it is important to evaluate the merits of various devices in their actual application. Apart of comparing specific models, it is interesting to determine if a treatment method provides expected results. In the case of drugs we expect healing or at least slowing down the progress of the disease. In the case of medical devices the effect may consist of enhancing the quality of life, reducing the costs and effort for routine measurements and consultations, improving the security or delaying the transfer to a nursing home.

The information about the effectiveness of such systems is important for the hospitals applying them and having to choose the best one. It is also important for approval institution in order to reevaluate the deployed systems (postmarket vigilance) and to assess the costs, especially if they are covered by public funds. Furthermore, this information can serve the device developers and scientists in choosing the viable directions for further research.

5.2 Data Access

In the medical research the identity of the patient has no relevance, therefore anonymized data should be used. It is much more important to accumulate data from many medical institutions, related to systems of many producers. Also combining data from various sources is useful. For example, correlating medical data with data about lifestyle or social environment may give clues about the causes of good or poor effectiveness of the systems. Of course, such correlated data permit to re-identify the patients if enough is known about their cases. It is however counterproductive to obfuscate data used for scientific research. Bad data may lead to bad, harmful conclusions. It is therefore much more important to control the group of people allowed to access data, recording suspicious queries and limiting the size of detailed, non-aggregated data sets accessible to any single person.

It may be also useful - under certain conditions - to retrieve specific patients. Normally statistics is interested in the typical cases, not in the outliers. It may however happen, that some patients or some groups of them respond differently to a treatment. In the case of drugs, it could be a patient having an unexpected resistance to the disease. In such a case, a more profound analysis of previously not considered factors could help to understand unknown aspects of the treatments.

6 RESEARCH CHALLENGES

Managing data - collecting, storing and distributing - will be a major issue in the real-life deployment of the medical applications based on the wireless sensor networks. We can name here some of the problems to be solved:

- compressing raw data streams for different recipients and for different purposes
- designing a hardware and software architecture for storing and querying large amounts of data
- protecting the data stream on its way end-to-end from the sensors to the final destinations
- removing the patient's identity from the data while still permitting to retrieve the detailed case if necessary
- defining a reasonable and enforceable privacy protection scheme permitting quality assurance and medical research
- developing secure and portable calculation and communication intensive applications for the smartphones
- organizing the cooperation between the data producers and data users

In this paper admittedly more questions are posed than resolved. We expect to take an active part in providing the answers.

REFERENCES

Röder, C., Müller, U., and Aebi, M. (2006). The rationale for a spine registry. *European Spine Journal*, 15:S52–S56.

Sliwa, J. and Benoist, E. (2011a). Pervasive computing - the next technical revolution. In *Developments in E-systems Engineering (DeSE), 2011*, pages 621 –626.

Sliwa, J. and Benoist, E. (2011b). Wireless sensor and actor networks: e-health, e-science, e-decisions. In *Mobile and Wireless Networking (iCOST), 2011 International Conference on Selected Topics in*, pages 1 –6.

Sliwa, J. and Benoist, E. (2012). A web architecture based on physical data separation supporting privacy protection in medical research. *International Journal of Reliable and Quality E-Healthcare*, 1(4).