

Data Fusion and Aggregation Methods for Pre-Processing Ambulatory Monitoring and Remote Sensor Data for Upload to Personal Electronic Health Records

Bruce Moulton^{*1}, Zenon Chaczko², Mark Karatovic

^{*1} *Corresponding author* School of Electrical, Mechanical and Mechatronic Systems,

² School of Computing and Communication,

Faculty of Engineering and IT, University of Technology, Sydney, Australia

brucem@eng.uts.edu.au

doi: 10.4156/jdcta.vol3.issue4.14

Abstract

Several countries including Australia are developing standards for personal electronic health records, and in many cases the standards provide ways for patients to upload data from their own sensor systems to their electronic health records. This article explores issues relating to the design of systems where electronic medical records are updated with data or meta-data from in-home or on-body health monitoring sensor systems. A prototype system was designed, implemented and subsequently evaluated in a way intended to further understanding regarding collection and transfer of sensor data to an electronic health record. A key issue that arose was the extent to which levels of confidence in the data may be affected by the type, quality, installation, maintenance and calibration of the sensors. This is in part because when health monitoring devices are used in hospitals, procedures exist for all of these activities, however, such procedures are usually not applied to in-home or ambulatory monitoring systems. Issues relating to the adoption of data aggregation or fusion techniques were also considered, a reason being that certain techniques involving aggregation of data from different sensors may be suitable for improving the degree of confidence in the extent to which the data from any particular sensor can be relied upon. One of the findings was that teams that are developing systems that might perform some data fusion at the node or prior to storage would benefit from the inclusion of clinical specialists at an early stage of the development process, because these specialists may in certain circumstances require disaggregated or un-fused data to inform clinical decisions.

Keywords

Electronic health records, Ambulatory monitoring, personal health sensors, data fusion

1. Introduction

Medical sensing devices are relatively common in hospital settings, where clinicians regularly make on-the-spot decisions about which elements of data are sufficiently noteworthy to be included as comments (or some other form of meta-data) on a patient's medical record. Recent advances in computing, miniaturization and manufacturing have seen a trend towards in-home and on-body sensing devices that are worn outside the hospital, in settings that include assisted-living and outpatients' and aged people's homes.

One of the issues raised by increasingly ubiquitous health monitoring sensors relates to the storage of information on a person's electronic health record (EHR). For example, if there is no clinician to interpret the data from the sensors, questions arise concerning how to translate raw data from sensors into meta-data that is useful for clinical decision-making. Additional issues arise relating to maintenance and calibration of the sensors. While programs exist for these activities in hospitals, they do not ordinarily exist for in-home or ambulatory monitoring systems.

In-home systems for detecting medical emergencies can include wirelessly networked sensor-transceiver devices. In some of the systems that are currently being developed, fixed sensing devices such as sonar and passive infrared (PIR) are positioned in high-risk locations such as bathrooms and stairs. Other systems use wearable devices, or a combination of fixed and wearable devices. Systems that use wireless sensors fixed to walls or floors have certain advantages over on-body systems, especially for people who experience dementia.

Work undertaken at the University of Technology Sydney has led to the development of health monitoring systems that use wirelessly networked sensor/transceiver devices [1, 2, 3, 4], and we are

currently exploring issues associated with sensor-transceivers that stream data such as heart rate, electrocardiogram (ECG), oxygen saturation, body temperature, and body position (including fall detection).

A detailed analysis of remote monitoring use cases has been published by the conducted by the U.S. Department of Health and Office of the National Coordinator for Health Information Technology Human Services [5]. The analysis focuses on remote monitoring use cases, and provides descriptions of ways in which information may flow from on-body or in-home monitoring devices to a person's electronic health record. For example, a patient, family caregiver, advocate, or other party may use a remote monitoring device to gather measurements at home, work, school, while traveling, or in assisted living facilities, and "these measurements are then communicated to the appropriate clinician's EHR via an information intermediary which could be provided by a device manufacturer or other entity based upon authorization of the patient." [5:16]. Hence it is apparent that some of the difficulties associated with the uploading of data to an EHR may be handled by standards relating to intermediary information exchanges.

Prior research suggests that data from on-body and in-home sensors can be combined with data such as vitals statistics or patient-discomfort to allow predictive analyses of patient levels of risk. For example, one study aggregated data from passive infrared (PIR) sensors (placed in the home of an 86 year over a two month period) with observations by nurses concerning measures including "bad sleep" and "faintness". The study found that certain characteristics in the PIR sensor data correlated with observations included in the medical health record of the patient [6].

One of the difficulties of such work relates to methods for processing or visualizing sensor data so that it can be meaningfully used by clinicians for decision-making. To this end various efforts have been undertaken and documented. These include the *LifeLines* system which summarizes a medical record as a set of lines and events on zoom-able timeline. Data is grouped according to categories including problems, allergies, imaging and medications, and searches can be performed on the entire record which allow relationships between different types of data to be analyzed [7]. Such work is in part motivated by the authors' view that even though more attention is now paid to developing standards for gathering medical records, too little effort is devoted to designing appropriate visualization and navigation techniques for

aggregating, presenting and exploring data contained within personal history records. It has been proposed that a referent tracking system may address such goals, but a drawback is that its implementation can require new approaches that differ greatly from those that are commonly used for clinical record keeping [8].

Data aggregation and data fusion techniques can be used to overcome individual sensor calibration issues and improve confidence in observations that are derived from multiple sensors [9]. For example, O'Donoghue et al. provide a real time framework for in-home health monitoring systems that aggregates data from various sensors together with patient information. Information relating to the characteristics of the various different types of sensors is combined with the data from the sensors to provide data-validation reports which can be used by the client to estimate sensor reliability [10]. Other projects have made use of different forms of contextual information to attempt to improve the reliability of aggregated data from remotely located health sensors [11].

An EHR ordinarily stores information relating to the health status of an individual. The information within an EHR can be used for several purposes. For example, it can not only be used to support the provision of health care to the owner of the record, but it may also be used to benefit the wider community when analyzed for purposes of quality management, research or education.

Several countries are moving towards national and international standards for EHRs (e.g. ISO and HL7). In Australia, the June 2009 *Final Report of the National Health and Hospitals Reform Commission* recommends that by 2012 every Australian will have a personal EHR that will at all times be *owned and controlled by that person*. In addition each patient will be able approve designated health care providers and caregivers to have authorized access to parts or all of their personal EHR [12]. The report also recommends that the Australian commonwealth government should oversee a national policy which ensures open technical standards framework for e-health, and that the standards should encompass issues such as interoperability, security and data quality [12]. An objective of the standards is to facilitate the eventual convergence of the current local systems into an integrated national health information system.

It is thought that the standardization of EHRs will help to bring about improvements in patient safety, quality of health care, consistency in reporting of test results, and information transfer between health providers [13].

An Australian IT-014-09 Electronic Health Records Interoperability Subcommittee is developing standards to facilitate the sharing of EHR information, and aims to enable improved alignment with proposed global ISO, HL7 and other EHR interoperability standards. These standards are intended to accommodate new forms of self monitoring/sensor data and new forms of clinical knowledge/meta data.

The specifications are intended to permit on-body sensors or in-home health monitoring systems to transfer different types of data to an EHR, either directly or via an information exchange.

Given the recent developments, we report on part of a project which investigates the integration of data from sensors with an electronic health record.

2. Assumptions/Method

Data models and block diagrams were developed, and a prototype was constructed and evaluated. Initial evaluations of the prototype led to revisions and the development of a second prototype system. The prototype designs and associated architectural considerations are described in [14]

A goal of the analysis was to explore design issues relating to a system that would be suitable for both data aggregation and data fusion. *Data aggregation* was taken to mean a process by which several values are combined (for example, weight, physical activity and time). The derivation of a measure of physical activity from the data from a PIR sensor is another example of data aggregation. *Data fusion*, on the other hand, was treated as a process by which data from several sensors, perhaps of different types, is integrated in a way that permits more efficient or more accurate inferences. Data fusion may entail some reduction, goals typically being that the fused data set is smaller, has less noise, or has more predictive capacity than that of any single sensor. The evaluation explored issues relating to the design of systems where electronic medical records are updated with data or meta-data from remotes sensing and on-body health monitoring sensor systems. Requirements analyses included the consideration of EHR remote sensing use-case scenarios. For example, one use-case scenario was that information from different types of in-home sensors, such as weight and daily movements around the home, would be daily updated to the EHR. The data could be used to track changes in the person's weight, which might in itself have important clinical decision-making implications, but if it were available, a clinician might be especially interested in informing decisions with a visualization

which aggregates variation in weight together with movement around the home over a period of months or years. The analysis also considered general requirements for medical records including requirements relating to objectivity, accountability, and freedom from ambiguity [15]. A sensor simulator was implemented for testing purposes. The simulator implements a sensor protocol whereby the simulated sensors download configurations for simulated patients and send sensor events containing different types of data, including data from in-home remote sensing and ambulatory monitoring systems [16].

3. Results/Discussion

The prototype was used to test key ideas relating to the design considerations of such systems. For the prototype, the JINI framework was considered potentially suitable because of the way it handles broadcasts to different devices and allows for device discovery. In addition, the distributed event mechanism allows devices to subscribe to and observe other devices on the network.

To develop a new architecture built around a TCP/IP socket connection framework was considered difficult. However, an advantage over JINI would be that if a connection were to be dropped, the framework could notify the server immediately rather than waiting for the next lease discovery phase.

The Java Message Service (JMS) could allow asynchronous broadcasts across multiple subscribers similar to JINI, but would not allow for automatic discovery of services and would not give any warning if an intended destination was unavailable.

It was determined that the prototype would implement a new framework via a custom protocol which allows customized polling, centralized information distribution and a customized security protocol.

An initial prototype includes an "event server" running in a Java VM and Java EE server side components running on the Glassfish server. The event server communicates with the Glassfish server to collect patient data and to send events to clients such as clinicians. After making a JNDI request to the lookup server using the JNP protocol, clients can connect to the event server via RMI.

Sensors issue events via their TCP/IP connections to the server, and the server can alert each client by way of a callback RMI call.

A client can also create a new event for any specific patient. In this case, the EJB Glassfish server sends a

message to the event server via JMS, and the event server fetches the new event cache using an EJB call via RMI.

The EJB Glassfish server exchanges messages with its MySQL server database via an SQL TCP/IP

connection. EJB calls that alter a patient's record are sent to the medical history server.

The data model is given in Figure 1.

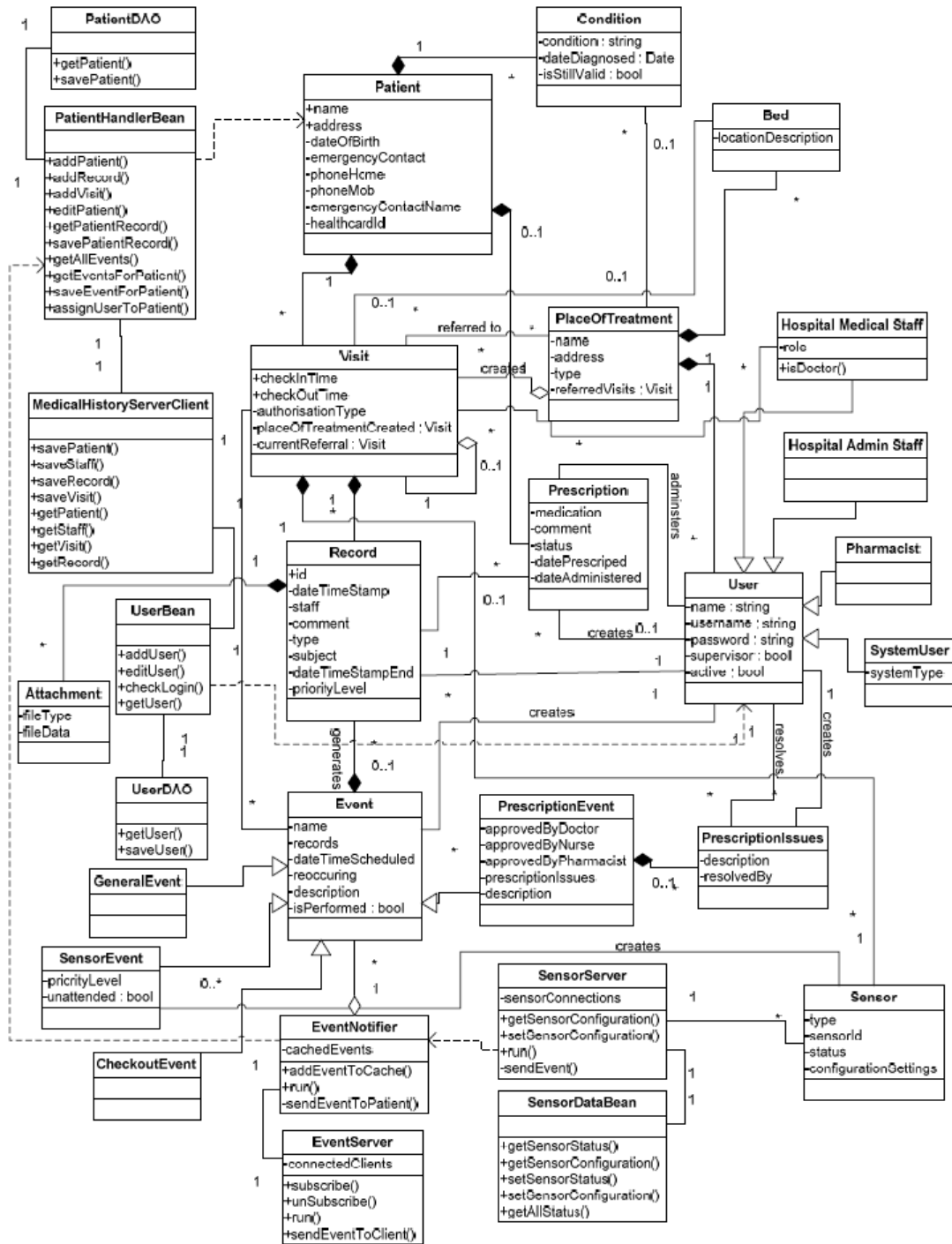


Figure 1. Data model

After a preliminary evaluation some of the requirements were relaxed and alterations were made to the prototype. The Event Server was moved to the EJB Glassfish server and implemented by a combination of EJB calls and JMS messages. JMS messages can broadcast requests for information to particular sensors, either to send current readings or to change their configuration.

JMS messages may be addressed to any particular sensor. A sensor that receives a JMS request will call on the Glassfish server to either send a sensor event or create a record of the current sensor reading.

To comply with the data model, each record created by a sensor is associated with a user.

The network security key is stored in the sensor system user object instance as its password. Sensors require the network security key to join the network, hence unregistered sensors will be unable to join unless they are given the key.

Records created by a sensor are marked with the user and current medical institution. SSL based connections were not implemented in the prototype, but it is recognized that in a production environment some form of encryption would be necessary to ensure that patient data transported across the network and across public networks would be less accessible to an unintended third party.

The resulting event subsystem is set out in Figure 2.

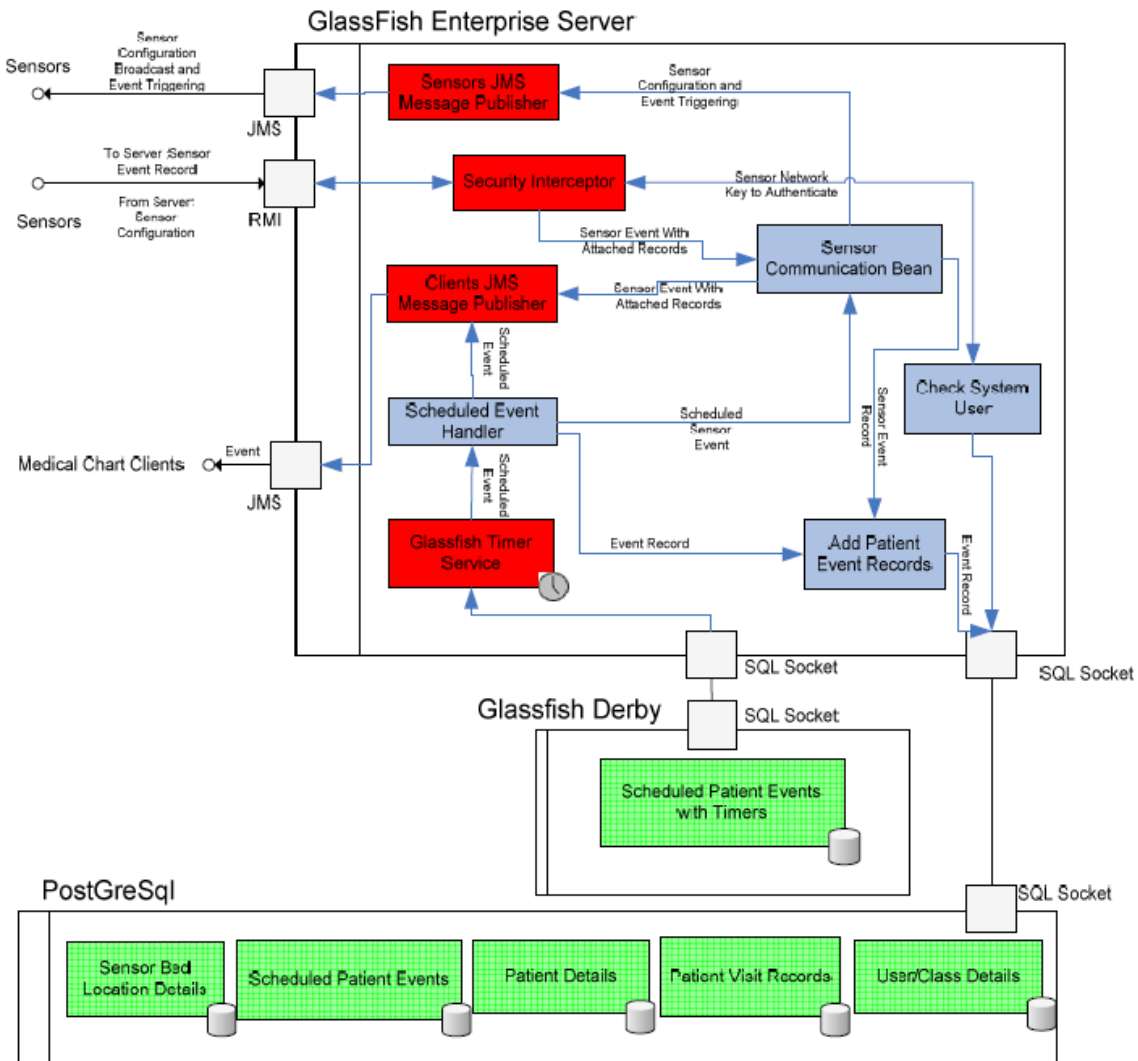


Figure 2. Prototype event subsystem architecture

All bean classes are treated as though they implement the EJB 3.0 standard, implement the correct interfaces, the correct annotations are defined, and are called on inside a Java EE container. The model assumes that the get() and set() methods for data classes are implemented for each attribute defined, and are not defined as operations in their respective classes. (This is common practice.) The bean classes are to conform to the Enterprise JavaBeans 3.0 specification and will interface with the Glassfish server.

Clients connect to the sensor server which is implemented by a class named SensorServer. The run() method is used because the class is threaded. When clients wish to customize sensor configurations they do so via the SensorDataBean class, which allows clients to get and set sensor configurations and connection status.

After some experimentation with the prototype it became clear that a system which aggregates data from heterogeneous sensors would benefit from a subsystem which classifies each sensor according to measures of the validity, reliability or quality of the data produced by each sensor. When large numbers of sensors are installed used by home users in an as required manner, additional issues arise relating to registration and calibration.

Levels of confidence in the data are to some extent affected by the type, quality, installation, maintenance and calibration of the sensor. Data aggregation or fusion techniques may be suitable for improving the degree of confidence in the extent to which the data from any particular sensor can be relied upon. Some of the considerations that were recognized during the evaluation phase are given in Table 1.

Table 1. Summary of considerations relating to data aggregation and data fusion

Sensor type calibration and operation	Standards play a key role in ensuring that data from different types of sensors can be used within the system. Issues remain relating methods for determining the degree to which sensors from certain manufacturers fail to meet standards or stated specifications. Calibration is initially performed by the manufacturer, and in a hospital setting, medical device sensors are regularly recalibrated as required. This is not the case with many in-home systems, so alternative methods are required. There is a role
---------------------------------------	--

	for data aggregation and data fusion techniques in mitigating issues that arise due to in-home sensor drift and lack of recalibration.
Patient/user competence	Remotely located sensors are commonly installed or attached by non-technicians, whether they are fixed to the walls of a home, or attached to part of the body by the patient. Confidence in the data is affected by the degree to which the remotely located person is able to use the sensors competently. The degree of confidence in the extent to which a sensor is correctly attached or installed can be improved by combining data from other sensors. For example, data from a force sensor or strain gauge located next to an electrode to could be used provide an indication of whether the electrode is firmly attached, or data from an accelerometer co-located with a PIR sensor could be used to help determine whether the fixed sensor is in the correct orientation.
Convergence/divergence	It seems likely that convergence will see integrated systems that are able to store large amounts of different types of data. It may be the case that search tools will evolve to allow patients, caregivers, and perhaps people with limited specialist knowledge to use the data to inform health-related decisions. However, there may be occasions when a detailed knowledge of the operation of the sensors is required, and this may lead to specialists (much the same way as scans are interpreted by radiologists rather than general practitioners). It may take specialist skills to drill down into the data, and if raw data is not transmitted or stored, fused data sets would likely present problems for some specialists who need to perform certain analyses.

4. Implications

Current policy in Australia stipulates that by 2012 every Australian will have a personal electronic health record that will at all times be owned and controlled by that person. Similar trends are taking place in other

countries, and many countries are working towards the development of standardized and nationwide electronic health records to allow more-integrated storing and sharing of information. Newly developing international standards are being drafted in a way that will enable patients to upload data from their own sensor systems to their electronic health record. This raises many new and interesting questions concerning the design and operation of such systems.

One area for consideration relates to the trend towards pre-processing or integrating data at the nodes of wireless sensor networks. Many current system designs incorporate local processing because it can help nodes within the network to reduce transmissions and hence save battery power – this is because transmission ordinarily requires disproportionately more energy than local processing. However, one of the concerns with this is that disaggregated data can be more useful for certain analyses and decision-making tasks. If the resolution and separation of data from multiple sensors is sufficient, it seems like the future technologies will make use of it in ways that are currently poorly understood. Hence in some situations there is a strong case against the reduction or fusion of certain types of data prior to transmission/storage.

Another area for further consideration relates to the level of confidence in the data that is collected from sensors, in part due to the fact that consumer-level sensors vary greatly in type, quality, reliability, maintenance and (lack of) regular recalibration. This issue is particularly important because when health monitoring sensor devices are used in hospitals, procedures exist for maintenance and recalibration, but such procedures are usually not applied to fixed or on-body in-home health monitoring systems.

5. Conclusions

Much work is to be done to further understand in-practice considerations relating to the collection and transfer of sensor data to an electronic health record. A key issue is the extent to which levels of confidence in the data may be affected by the type, quality, installation, maintenance and calibration of the sensors – these variables are controlled in hospital settings, but typically not in ambulatory or in-home health monitoring systems. Data fusion and aggregation techniques may be required for improving the degree of confidence in the extent to which the data from any particular sensor can be relied upon. In addition, it seems likely that standards will play a key role in

ensuring the quality of data from different types of sensors.

6. References

- [1] Fischer, M., Lim, Y.Y., Lawrence, E., Ganguli, L.K. "ReMoteCare: Health Monitoring with Streaming Video," 7th International Conference on, Mobile Business, 2008. ICMB '08. pp.280-286, 7-8 July, 2008.
- [2] Lim, Y., Messina, M., Kargl, F., Ganguli, L., Fischer, M., and Tsang, T. "SNMP-Proxy for Wireless Sensor Network," Fifth international Conference on Information Technology, Singapore, 29 August 2008.
- [3] Leijdekkers, P., Gay, V. and Lawrence, E. "Smart Homecare System for Health Tele-monitoring," First International Conference on the Digital Society, (ICDS'07), 2007.
- [4] Demirjian, S., Chaczko, Z., Moulton, B., Mahadevan, V. (2009) Evaluation of an approach for developing a touch screen UI for residential aged care services, International Conference on Computer and Network Technology (ICCNT 2009), July 24-26, 2009, Chennai, India.
- [5] United States Department of Health and Office of the National Coordinator for Health Information Technology Human Services. Remote Monitoring, detailed use case. Detailed use case document (2008) http://healthit.hhs.gov/portal/server.pt/gateway/PTARGS_0_10731_848114_0_0_18/RMonDetailed.pdf [accessed 19.8.2009]
- [6] Hadidi, T. and Noury, N, A Predictive Analysis of the Night-Day Activities Level of Older Patient in a Health Smart Home, Ambient Assistive Health and Wellness Management in the Heart of the City 7th International Conference on Smart Homes and Health Telematics, ICOST 2009, Tours, France, July 1-3, 2009. Proceedings 290-293 2009 Springer Berlin / Heidelberg
- [7] Plaisant, C, R. Mushlin, A. Snyder, J. Li, D. Heller, and B. Shneiderman. Lifelines: Using visualization to enhance navigation and analysis of patient records. In American Medical Informatics Association Annual Fall Symposium, pages 76-80, 1998.
- [8] Ceusters W, Smith B. (2006) Strategies for referent tracking in Electronic Health Records. Journal of Biomedical Informatics Volume 39, Issue 3: 362-378.
- [9] Tatbul, N, M. Buller, R. Hoyt, S. Mullen and S. Zdonik. Confidence-based data management for personal area sensor networks. In DMSN '04: Proceedings of the 1st international workshop on Data management for sensor networks, pages 24–31. ACM, 2004.
- [10] O'Donoghue, J, J. Herbert and D. Sammon. Patient sensors: A data quality perspective. Proceedings of the 6th International Conference on Smart Homes and Health Telematics, pages 54–61, 2008.
- [11] Mohamed, I, A. Misra, M. Ebling, and W. Jerome. HARMONI: Context-aware Filtering of Sensor Data for Continuous Remote Health Monitoring. PerCom 2008: 248-251

- [12] A Healthier Future For All Australians – Final Report of the National Health and Hospitals Reform Commission – June 2009
<http://www.health.gov.au/internet/main/publishing.nsf/Content/nhhrc> [accessed 31.6.09].
- [13] Standards Australia e-health IT-014 <http://www.e-healthstandards.org.au/cat.asp?catid=5> [accessed 30.6.09].
- [14] Moulton, B, Chaczko, Z., Karatovic, M. (2009) Updating electronic health records with information from sensor systems: considerations relating to standards and architecture arising from the development of a prototype system. In press, accepted 1 Oct 2009, .Journal of Convergence Information Technology.
- [15] Staunton & Whyburn (1997), Nursing and the law, 4th Edition, W B. Saunders, Sydney.
- [16] Moulton, B., Chen, J., Croucher, G., Lal, S., Lawrence, E., Mahendran, L., Varis, A. (2009) Ambulatory Health Monitoring and Remote Sensing Systems to be used by Outpatients and Elders at Home: User-Related Design Considerations. Proceedings of the 11th IEEE International Conference on e-Health Networking, Applications & Services, IEEE Healthcom 2009, 16-18 December 2009.