INFORMATION QUALITY AND DATA MANAGEMENT WITHIN A PERVASIVE MEDICAL ENVIRONMENT

Paper Category: "Completed Academic Paper"

Topic Category: IQ in Sensor Networks; IQ in Situational Awareness; IQ Concepts, Tools, Metrics, Measures and Models

John O'Donoghue[1], John Herbert[2] David Sammon[1] and John Barton[3]

[1]Business Information Systems, University College Cork, Ireland

[2]Department of Computer Science, University College Cork, Ireland

[3]Tyndall National Institute, Ireland

john.odonoghue@ucc.ie, j.herbert@cs.ucc.ie, dsammon@afis.ucc.ie and john.barton@tyndall.ie

Abstract

Next generation pervasive medical domains will be made up of numerous quantities of autonomous: processing, communicating and sensing devices. These may include personal digital assistants (PDA) wireless sensor networks (WSN) or of a patient centric importance, Body Area Networks (BAN). Before any data management task may be executed, the context or situation of the user and their environment needs to be taken into account. This large paradigm shift from centralised decision making networks to remote autonomy create new challenges within the information quality community, particularly how to collect, correlate and disseminate this new information pool in an intelligent manner. Presented in this paper are the findings of the Data Management System-Data Consistency Model (DMS-DCM) software architecture within a pervasive medical environment. Five data management experiments were conducted to evaluate the DMS-DCM's effect on information quality.

1 INTRODUCTION

Poor data quality infrastructures are estimated to cost U.S. businesses more than $600 billion per year. Within a medical environment poorly designed clinical information systems result in inadequate access to electronic patient records. This can contribute to medical errors which may result in patient fatalities. For example according to the Institute of Medicine, 98,000 patients die in U.S. hospitals each year due to poor data management issues.

Patient sensors are playing an ever increasing role within medical environments, and these devices can generate large quantities of data. It is necessary to coordinate such data sources to provide the end users (i.e. medical practitioners) with context relevant information. Integration of mobile and sensing devices within a medical environment may provide continuous patient analysis facilities, improved data access and archival reports. These developments may help to improve the quality of information within clinical information systems. A number of architectures exist within the pervasive computing environments which are designed to tackle a number of context aware issues [2], [1].

Presented in this paper are the results of the Data Management System-Data Consistency Model (DMS-DCM), an agent based software component applied within a pervasive medical environment [10]. This component is part of the parent Data Management System (DMS) architecture (cf. figure 1). The DMS is capable of interacting with multiple sensors of various types and large data repositories.

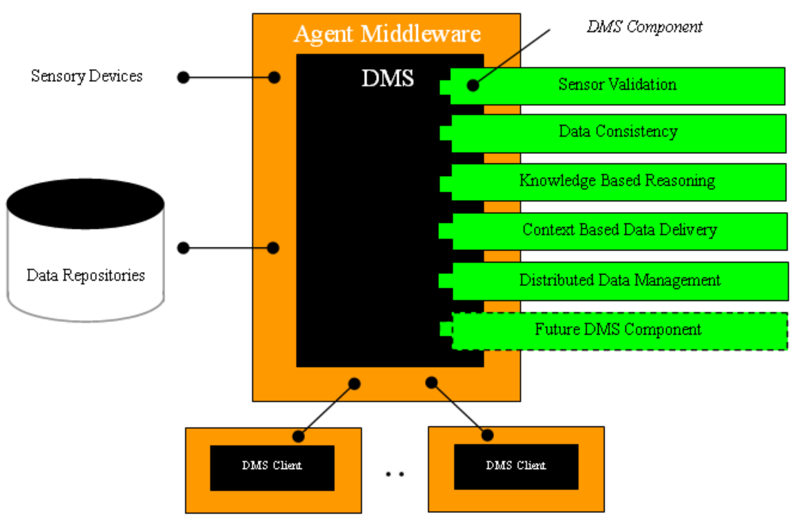

Figure 1: Data Management System Architecture. DMS-Client e.g. PDA or Mobile Phone with a built in agent layer.

The DMS-DCM is designed to ensure that vital datasets from the Tyndall Body Area Network (BAN) are synchronised between front end devices (PC and PDA) and backend servers. The DMS-DCM demonstrates that data inconsistency is acceptable if the datasets in question are of little or no importance. The DMS-DCM rules, underline this approach by ensuring that only relevant datasets are captured and synchronised in a timely manner. Therefore high levels of data inconsistency may exist within a number of devices. This helps to save on valuable bandwidth by only transmitting relevant datasets to the end user or central server. The results presented demonstrate this fact, as high levels of inconsistency do exist, however medical practitioners only receive the necessary datasets. This helps to reduce information overload and create a quality oriented data management systems.

2 Context Based Data Consistency

Within a pervasive medical environment multiple sources of static and dynamic datasets exist. A high degree of importance is associated with this data, as medical practitioners prescribe relevant patient care based on the information provided. Pervasive environments contain multiple points of access that allow medical practitioners to read and modify patient datasets through PCs, PDAs and other mobile devices. Enabling mobile medical practitioners to modify a patient dataset introduces a new data consistency problem. The DMS-DCM is designed to intelligently interact with servers, mobile computing devices and patient sensor nodes within a Wireless Sensor Network or Body Area Network (WSN or BAN). Effective data consistency is a fundamental requirement within health informatics. It provides the foundation to ensure that medical practitioners receive up-to-date high quality relevant information on time every time. In a distributed dynamic environment multiple views of the same dataset may exist. The DMS-DCM employs a Jade agent platform to ensure that all relevant medical practitioners share a consistent view of patient datasets in real-time.

The DMS-DCM is designed to optimise data management facilities within pervasive medical environments. It is essential within a medical domain that all generated datasets within the distributed environment (e.g. PCs, PDAs and patient sensors) are up-to-date. Classical data management employ two key operations read and write. In relation to data consistency a write operation can not be executed in isolation. It needs to verify that no other user is interacting with the current dataset and that correct datasets are replicated amongst all users. This is referred to as strong data consistency [11]. The DMS-DCM applies a similar approach to the view-based consistency of [6]. Following this method, data objects of a 'view' type are only required to be updated before they are accessed. This ensures that medical practitioners receive the latest datasets upon request. A multicast-based middleware is presented in [4]. The DMS-DCM provides similar data retrieval and dissemination techniques within a pervasive environment. The DMS-DCM is a novel approach in managing data consistency by employing context aware reasoning Jade agents within a data rich pervasive medical environment. The DMS-DCM shares similar qualities to [3] where cooperating mobile agents work in unison to ensure data consistency is maintained at all times under a diverse set of conditions.

The DMS-DCM architecture is built on two main datasets: core patient data (i.e. patient vital signs, patient history etc.) and context parameters (time, location, profile). By combining these two sets of data in the DMS-DCM model, data consistency techniques may be enhanced. For example consider a patient wearing the Tyndall-BAN. Pulse rate sensor readings are sampled and transmitted to the DMS-Server. One medical practitioner may update his/her PDA with extra information alongside the current real-time sensor values. Based on the context of the patient and medical staff, Jade agents with built-in reasoning may dynamically decide which datasets need to be transmitted to medical devices within the pervasive network. This approach takes advantage of all known real world information in relation to the environment and enhances the data management capabilities to achieve two key goals: to provide medical staff with relevant real-time information and to reduce information overload.

The data consistency techniques employed between medical practitioner mobile devices and central medical servers are built using an agent based architectural framework Jade. These techniques enable mobile medical devices to retrieve context relevant information and disseminate real-time datasets to the appropriate medical practitioners and medical servers. The communication links between medical practitioners and the DMS-Server are illustrated in figure 2. This diagram highlights a sequence of communication between agents and external patient sensors within a wireless patient sensor network.

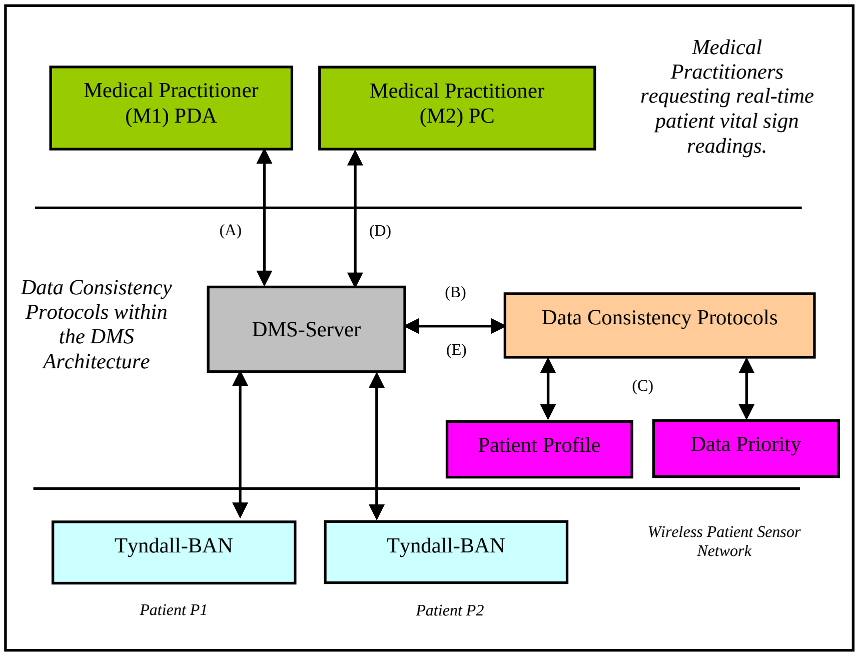

Figure 2: DMS-DCM Architecture.

The DMS-Server runs in parallel with DMS-DCM protocols (Jade agents). Here all datasets are prioritised, filtered, and transmitted based on the context of the patient, medical staff and their associated environments. A walkthrough of a number of communication links is outlined in figure 2:

A) A medical practitioner M1 requests patient data from patient P1.

B) Consistency protocols are executed before passing the real-time sensor data onto medical practitioner M1.

C) To save on bandwidth overheads only relevant data is replicated based on the patient profile and sensor data priority.

D) Medical practitioner M2 requests real-time patient data from Patient P1 in parallel to M1.

E) Jade software agents within the DMS-Server replicate all known datasets on M1s PDA and correlate this data with M2 and P1 real-time sensor readings, thus providing a higher QoS. M1 and M2 now share a consistent view of the patient dataset including real-time vital sign readings.

To help achieve an effective high quality information data management system, three key data elements were utilised:

1. End User Profile

A pervasive medical environment may generate large amounts of data. This may result in data overload, poor QoS and potentially, poor patient care. The introduction of user profiles adds a new dimension to patient data management within a pervasive environment. Presented in [9] is the DMS-User Profile (DMS-UP). User Profiles are utilised to effectively manage the delivery of relevant information to a practitioner's mobile device, thus improving the required QoS. Similar approaches have been applied to overcome network disconnection and/or limited bandwidth [7].

2. Priority

Data consistency may employ various forms of data replication, where specific or entire datasets are copied to a variety of clients and/or servers. In relation to the DMS-DCM, patient, staff and environment variables are given priorities or data quality weightings. This enables the DMS-DCM to update scheduled medical practitioners with relevant real-time information.

3. The Role of Context

The advantage of integrating a context sensitive data caching technique is outlined in [8]. Here relevant data is cached to neighbouring nodes based on the locality of the mobile user within a wireless network. The DMS-DCM protocols are based on the same philosophy. Merging the profile and priority of patient/staff with the context of the environment (e.g. identity of staff, location and time) decreases the levels of inconsistency as only important datasets are synchronised.

2.1 Agent Based Data Management

Agent technology is the enabling middleware utilised by the distributed components within the DMS-DCM. In the context of software engineering, an agent can be defined as [12]: "An entity within a computer system environment that is capable of flexible, autonomous actions with the aim of complying with its design objectives" A mobile agent adheres to this definition as well as having the added capability of traversing networks. The field of agent technology is seen as a highly suitable paradigm and inter-communication infrastructure for the analysis and design of mobile telemedicine systems [5]. This work views agent technology as a vast improvement to the traditional client-server approach for developing complex telemedicine applications. These systems can be defined as communities of interacting entities that aim to support collaboration and resource sharing in a medical environment. This observation is especially prevalent for mobile telemedicine systems which have continuously appearing and disappearing components within their distributed network. An agent embodies characteristics of autonomy and sociality which make the multi-agent paradigm highly appropriate to develop mobile telemedicine systems. An outline of some of the key software agents and their interaction between medical staff and specific datasets is presented in figure 3. An overview of three of the main software agents is given:

- **Mobile Device Manager Agent (MDMA)**

This is a resident software agent (i.e. JADE) designed to arbitrate between the central DMS-Server and medical practitioner. The agent deals with medical staff requests and incoming server updates. The MDMA interacts with the Mobile Device Data Consistency Agent for pre and post data consistency checks.

- **Mobile Device Data Consistency Agent (MDDCA)**

Incoming and outgoing datasets to and from the mobile device pass through the MDDCA. This agent is responsible for ensuring that not only is the current dataset updated but that all related datasets are resident on the device. It achieves this by interacting with the Central Data Consistency Agent.

- **Central Data Consistency Agent (CDCA)**

The CDCA manages all real-time sensor data streams and medical practitioner read/write update/requests. CDCA executions are based on the central expert system, which contains a formal set of data consistency rules. All DMS-Server datasets are stored in the Medical Data Store; this includes sensor readings and database records.

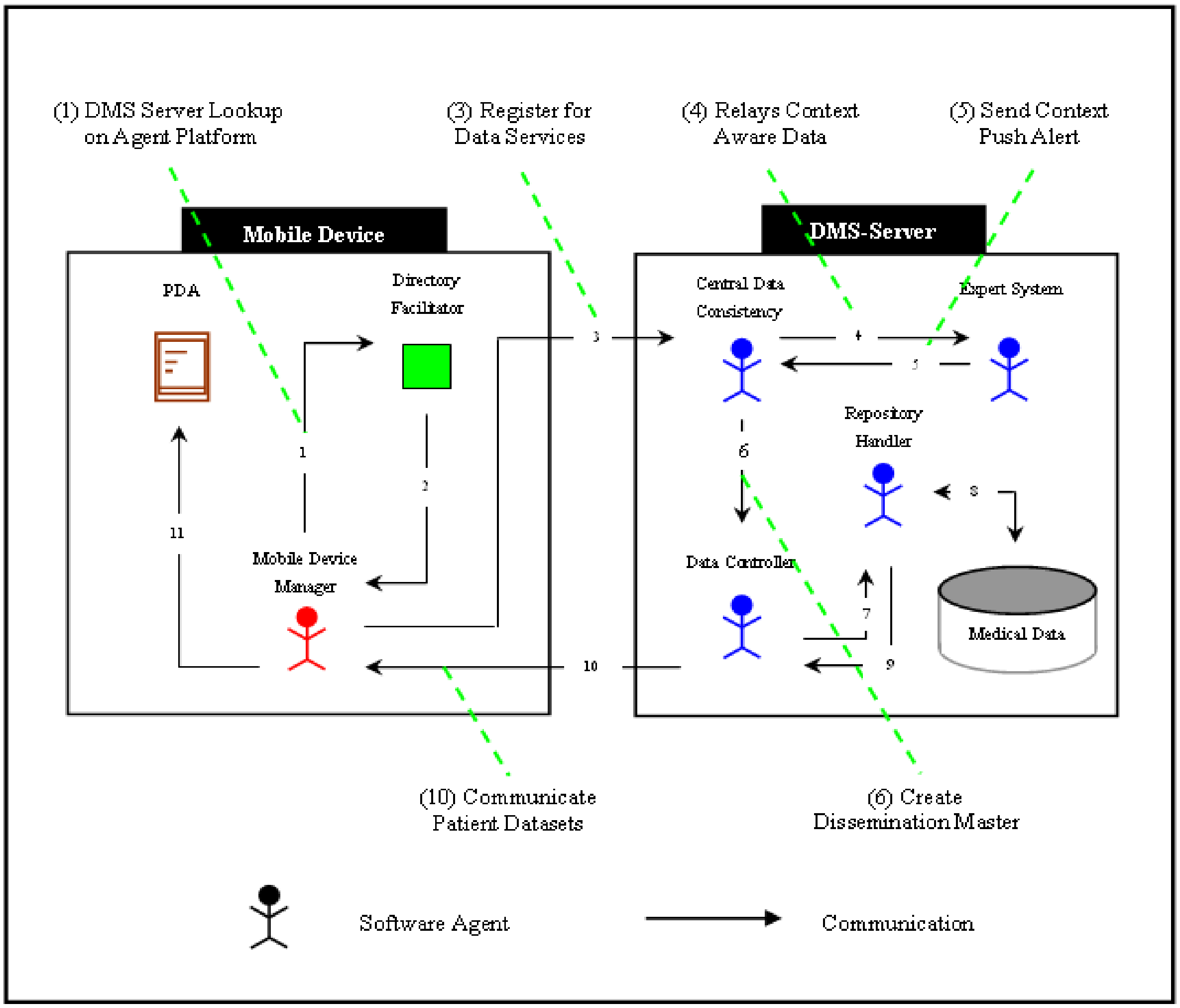

Figure 3: The DMS-DCM Agent Based Architecture.

2.2 *Evaluation*

The current DMS-DCM prototype has been evaluated for the performance of the agent based data consistency protocols in helping to deliver high quality datasets within a pervasive medical environment. Experiments were conducted to evaluate the effectiveness of DMS-DCM's broadcasting (server to client) and sampling (based on client requests) capabilities. A patient's pulse rate reading (over a four hour period) which was taken using the Tyndall-BAN is used as the primary patient sensor data source.

Five data management scenarios are evaluated:

1) Periodic Server to Client Updates. The DMS-Server periodically updates a mobile user (DMS-Client) with patient information datasets.

2) Medical Practitioner's Server Interaction. A mobile medical practitioner requests patient datasets (simulated environment / Poisson distribution).

3) Multiple User Interaction. A single patient may be monitored by more than one medical practitioner. It is important that data concerning that patient is maintained not only on the DMS-Server but also on the relevant medical practitioner's mobile devices.

4) A user profile may contain the wellbeing associated with each patient (i.e. low or high risk). This can greatly affect the quantity of datasets which need to be transmitted.

5) Data Priority. The rate at which a patient's sensor reading may change over a period of time can indicate potential patient risks. It is important to recognise such risks in context of supporting sensor or historical values this in turn may help to reduce false alarm generation.

2.3 Test Case Environment

All experiments are conducted on patient pulse rate readings in an offline mode. The pulse readings are sampled over a 4 hour period (240 minutes). Sensor readings are stored within the medical database. 240 readings are stored, one reading for each minute.

The definition of a patient's pulse region in this paper is defined as generic or expected (cf. table 4). For improved accuracy a set of pulse regions should be designed for each individual patient. This will take into account their medical history, age and level of activity. This will help to decrease false alarm generation and provide the medical practitioners with an improved view of the patient's state of health.

2.4 Evaluation Results

The evaluation of the five data management experiments are as follows:

2.4.1 Periodic Server to Client Updates

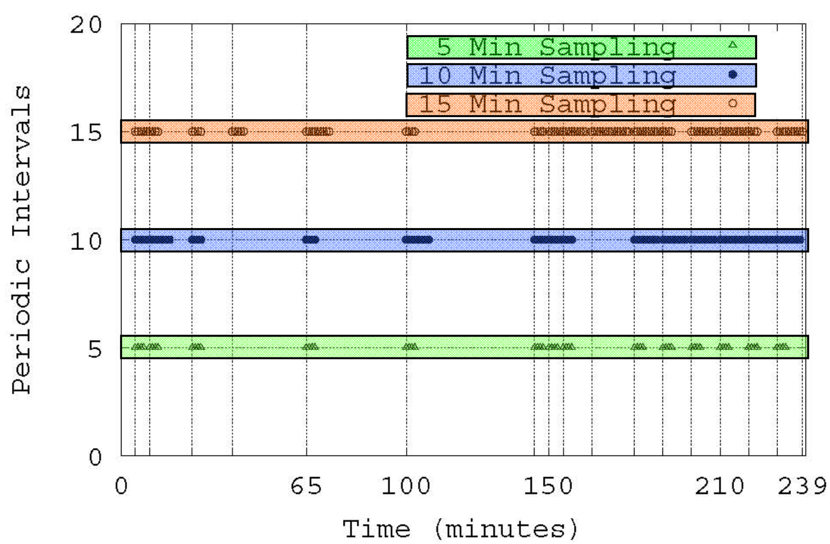

Figure 4: Periodic Sampling within a Simulated Environment (Poisson distribution).

The DMS-Server (PC) is configured to periodically transmit a patient's pulse rate every 5, 10 and 15 minutes to a DMS-Client (PDA). The vertical lines (cf. figure 4) represent an increase or decrease in the patient's pulse level (a significant pulse change). The solid horizontal segments during the periodic intervals 5, 10 and 15 minutes represent the inconsistency between the DMS-Server and the mobile device.

	Update Interval Period (Server to Client)		
	5 Minutes	10 Minutes	15 Minutes
Sensor Activity (Frequency of Change)	Percentage of Inconsistency		
Low (20 Minutes)	10%	35%	40%
Medium (10 Minutes)	20%	70%	64%
High (2 Minutes)	80%	90%	93%
Simulated Environment (Poisson Distribution)	24%	40%	46%

Table 1: Percentage of inconsistency during periodic updates.

During a simulation evaluation for 5, 10 and 15 minute periodic updates the mobile device is inconsistent with the DMS-Server for 24%, 40% and 46% respectively (cf. table 1). For a 5 minute periodic update with sensor activities of 20, 10 and 2 minutes the practitioner's mobile device is inconsistent for 10%, 20% and 80% during a 4 hour period. The fundamental issue with periodic updates stems from its lack of context awareness. It does not take into account the meaning of the sensor values and what they represent. If the medical practitioner is to receive the correct information on time every time then a greater understanding as to what each value represents and its context is required.

2.4.2 Medical Practitioner Server Interaction

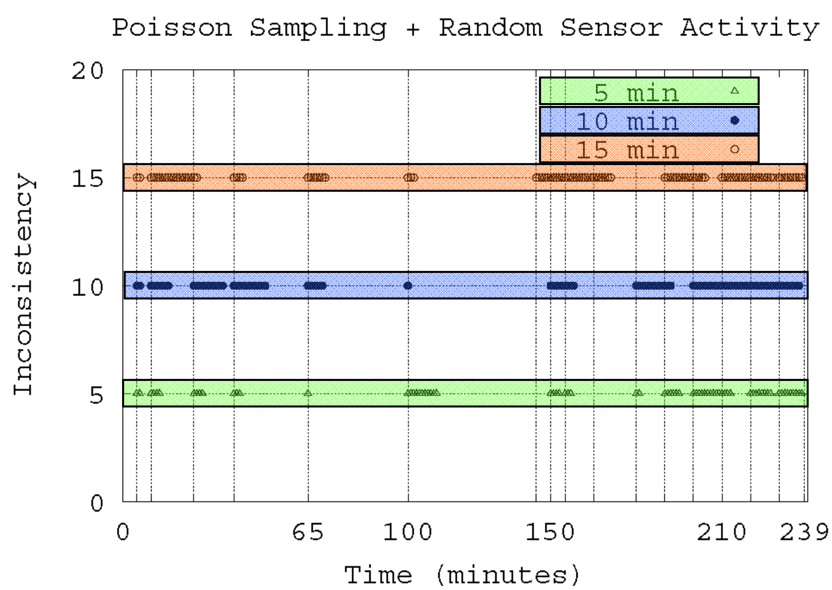

Figure 5: Poisson Server Interaction within a Simulated Environment.

A medical practitioner may request patient vital signs randomly throughout the day. Presented in figure 5 is the level of inconsistency generated with a Poisson distribution mean of 5, 10 and 15 minutes within a simulated environment (cf. table 2).

	Poisson Server Request		
	5 Minutes	10 Minutes	15 Minutes
Sensor Activity (Frequency of Change)	Percentage of Inconsistency		
Low (20 Minutes)	32%	50%	38%
Medium (10 Minutes)	49%	63%	59%
High (2 Minutes)	74%	85%	88%
Simulated Environment (Poisson Request)	30%	41%	43%

Table 2: Percentage of inconsistency during Server Interaction.

The percentage of inconsistency of the periodic update and a client request is very high. Both approaches do not evaluate the patient sensor readings against known context elements. If a patient's vital signs were to elevate to a high risk level it is possible that the required data may not reside on the medical practitioner's device. This has the potential to interfere with the level of patient care.

2.4.3 Multiple User Interaction

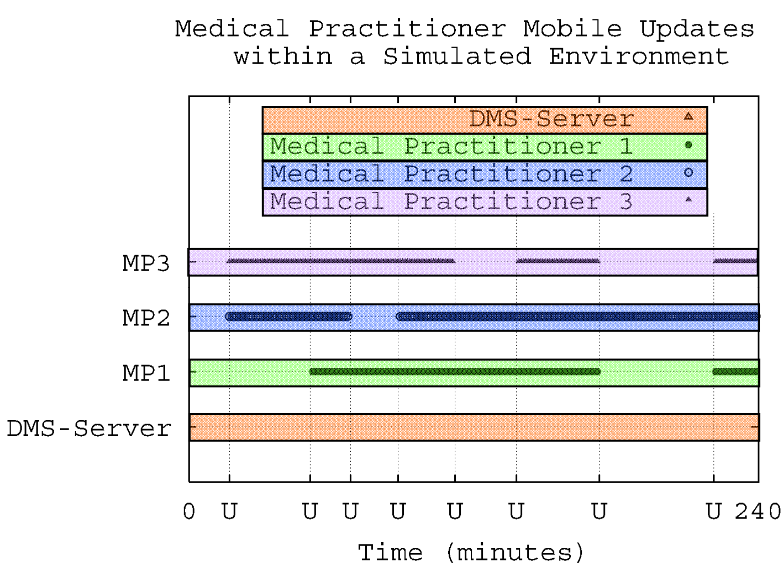

Figure 6: Multiple Medical Practitioner Data Updates within a Simulated Environment. U implies a DMS-Server to DMS-Client update. MP1 is medical practitioner 1.

A pervasive medical environment may contain a large number of mobile devices. Each device may hold specific datasets related to a group of patients. To help save on bandwidth and reduce data redundancy only active medical practitioners receive patient vital sign readings when the pulse level reaches an area of risk (cf. figure 6, table 3). The level of inconsistency is still very high; however the medical practitioners receive all of the patient datasets they require to make a well informed decision. This data management process helps greatly to improve the overall information quality within the system.

Medical User	MP1	MP2	MP3
Percentage of Inconsistency	58%	85%	61%

Table 3: Multi User Interaction (MP is Medical Practitioner).

2.4.4 Patient Profile

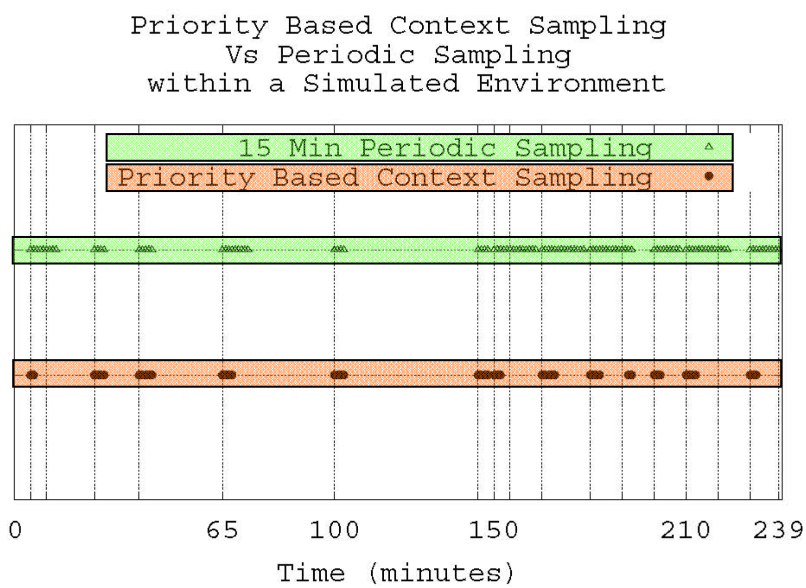

Figure 7: Patient Profile Priority Based Update Compared with a Periodic 15 Minute Update. Vertical Lines Represent Sensor Value Change.

A patient's resting pulse rate can differ significantly between an athletic and non-athletic. A patient's profile containing the type of patient (i.e. athletic or non-athletic) can assist is identifying which datasets need to be transmitted to the medical practitioner for analysis. Presented in figure 7 is a comparison between a patient profile priority update and a 15 minute periodic update. With a patient profile priority update, more datasets needed to be transmitted. As the patient's pulse rate was below the expected range, the patient priority based approach achieved 25% inconsistency compared with the periodic update of 46%. This approach ensured that the medical practitioner received the necessary information on time to provide better patient care.

The state of a patient allied with the rate of sensor change can significantly alter the required amount of data which needs to be transmitted to a mobile device. Presented in figure 8 is a simulated environment with three patient types; low, medium and high risk (in relation to their pulse regions (cf. table 4)) within a simulated environment.

Age/Risk	Low (BPM)	Medium (BPM)	High (BPM)
0-1	130+-20	100-109 / 151-160	<100 / >160
1-10	100=-30	60-69 / 131-140	<60 / >140
10+	80+-15	60-74 / 96-100	<65 / >100
Athlete	50+-5	40-45 / 55-60	<40 / >60

Table 4: Generic Pulse Regions (BPM is Beats Per Minute).

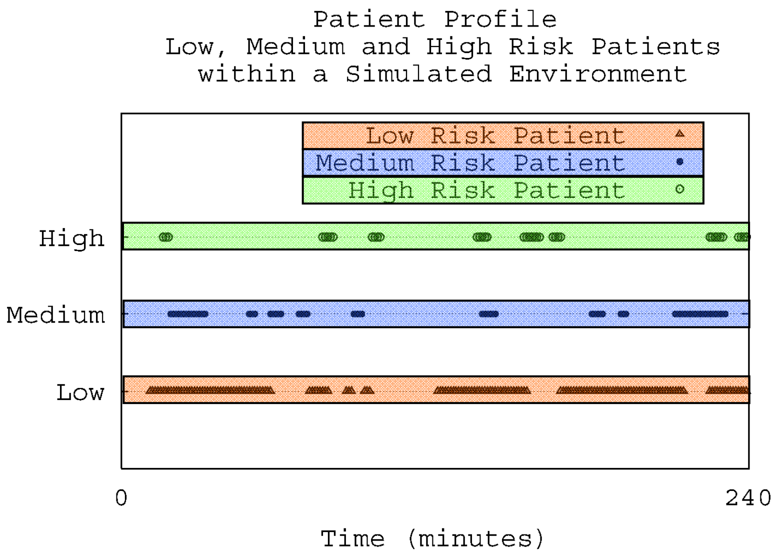

Figure 8: Data Transmission with various Patient States based on their level of Risk.

A DMS agent operating within the DMS-Server evaluates the patient's pulse level and compares it against the generic pulse regions. After evaluation it transmits data to relevant medical practitioners based on the following data quality transmission rules.

Rule	Transmission Time
If pulse reading is within a low risk region.	Once every 60 minutes,
If pulse reading is within a medium risk region.	Once every 10 minutes.
If pulse reading is within a high risk region.	Once every minute.

Table 5: Data Quality Transmission Rules.

	Patient Risk Level		
	Low	Medium	High
Percentage of Inconsistency	67%	27%	15%
Number of Client Updates	6	36	114

Table 6: Patient Profile with Low, Medium and High Sensor Value Change.

High percentages of inconsistency can be tolerated based on the content of the data. For example in table 6 a low risk patient dataset is inconsistent for 67% of the time during a four hour period. The software agent operating within the DMS-Server uses data consistency rules to effectively evaluate the patient state and sensor readings and transmit accordingly. This helps to reduce information overload while, in parallel, frees up resources for higher risk patients whose data needs to be communicated, thus improving the overall QoS within the medical environment.

2.4.5 Data Priority

A patient's pulse level may rise or fall gradually over a period of time. It is therefore necessary to not only read the current value but to view it in the context of previous sensor readings. A new set of data consistency rules are presented (cf. table 7) to manage this data consistency requirement.

Pulse Region	Transmission Time
Pulse rate of 70-90 BPM	Once every 20 minutes
Pulse Rate 90-150 BPM	Update 1 every 10 minutes If longer than 10 minutes Update 1 every 5 minutes If longer than 15 minutes Update 1 every 2 minutes If longer than 20 minutes Update 1 every minute
Pulse Rate 150+ BPM	Update Once every Minute

Table 7: Gradual Data Consistency Rules.

This new set of data consistency rules were applied within a simulated environment with three patient levels of activity: resting, mildly active and active. The results are presented in figure 9 and table 8.

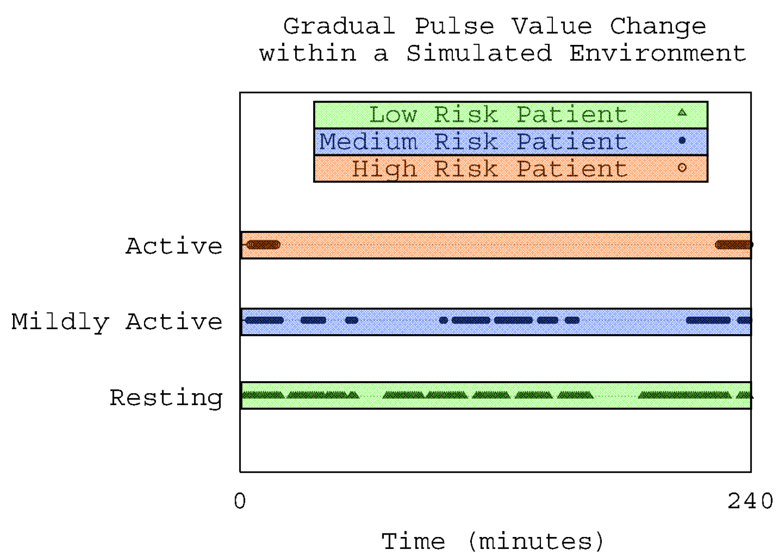

Figure 9: Gradual Sensor Reading Change under Various Patient States.

	Patient State		
	Resting	Mildly Active	Fully Active
Percentage of Inconsistency	72%	42%	12%
Number of Client Updates	24	76	208

Table 8: Gradual Patient Updates under Various Patient States.

The gradual data consistency rules have prioritised the critical data and transmitted accordingly. The overhead associated with high risk patients can be seen with a total of 208 DMS-Client updates over a four hour period. However this will ensure that the medical practitioners receive real-time datasets at the patient point of care.

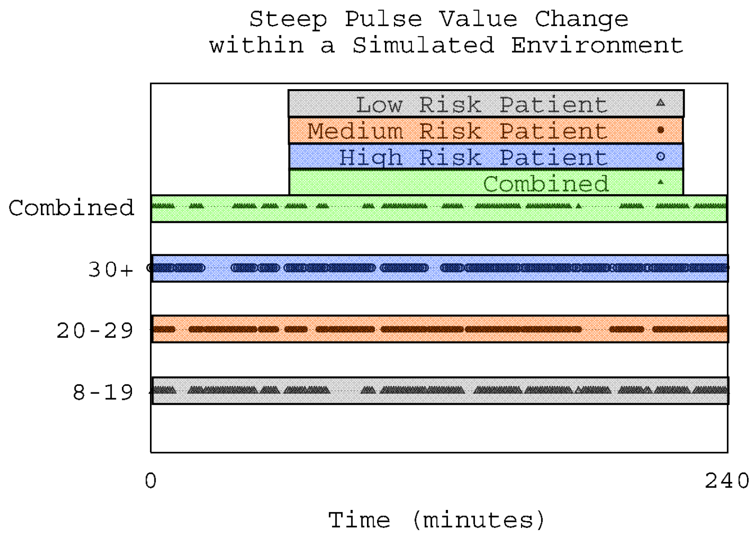

Figure 10: Steep Pulse Value Change.

Pulse regions of Interest	8<= X <= 19	20<= X <= 29	X >= 30	Combined
Percentage of Inconsistency	74%	79%	82%	60%
Number of Client Updates	43	35	36	65

Table 9: Steep Pulse Value Change. A patient's pulse is monitored over a three minute period. If it increases or decreases by more then or equal to 8 BPM then data is transmitted.

Over a short period of time a patient's pulse rate may increase or decrease dramatically. Such a phenomenon may require immediate attention. Presented in figure 10 are the results based on the data consistency rules (cf. table 9) applied to a patient's pulse readings over a four hour period. The data consistency rules were applied to a patient in a resting state. These results help to identify peak periods during a patient's daily routine. Through the combination of all three areas of interest a total of 65 required updates were passed onto the medical practitioner.

3 Conclusion

Presented is the evaluation of the DMS-DCM (Data Management System-Data Consistency Model) under five data management experiments. These experiments were developed to examine the effects various data management rules may have on data quality within a pervasive medical environment. The DMS-DCM is designed to intelligently collect, process and disseminate real-time patient data within a pervasive medical environment. It achieves this by merging context related consistency rules with known context information (patient profile, rate of sensor change etc) to deliver relevant real-time information to medical practitioners.

The DMS-DCM ensures that a dataset which may reside on a medical practitioner's mobile device (which may be associated with other specific members of the medical personnel) is communicated appropriately. This ensures that any context update which occurs within the mobile device is transmitted to key members of staff in a timely manner. The current DMS-DCM prototype demonstrates that issues such as data overload and bandwidth usage may be improved. It also demonstrates that relevant real-time information may be contextually delivered to members of the medical staff, thus improving the overall quality of information within the medical environment.

The next increment of the DMS-DCM will include a greater array of patient sensors (e.g. beat-to-beat and accelerometers). This will enable the medical practitioners to gather an improved understanding of a number of real world situations specifically in the area of falls assessment among the elderly. In turn data mining tools and approaches will be employed to improve the off-line analysis capabilities.

References

[1] Biegel, G. and Cahill, V., “A Framework for Developing Mobile, Context-Aware Applications”. In Proceedings of the Second IEEE International Conference on Pervasive Computing and Communications (PerCom), pp. 361–365. (2004).

[2] Bu, Y., Gu, T., Tao, X., Li, J., Chen, S. and Lu, J., “Managing Quality of Context in Pervasive Computing”. In Proceedings of the 6th International Conference on Quality Software (QSIC '06), Beijing, China, October (2006).

[3] Cao, J., Chan, A.T.S. and Wu, J. “Achieving Replication Consistency Using Cooperating Mobile Agents”, In Proceedings of International Conference on Parallel Processing Workshops (ICPPW), pp. 453-458, (2001).

[4] Chrysanthis, P., Pruhs, K. and Liberatore. V., “Middleware Support for Multicast-based Data Dissemination: A Working Reality”, WORDS, (2003).

[5] Della Mea, V., “Agents Acting and Moving in Healthcare Scenario: A New Paradigm for Telemedical Collaboration”, In Proceedings of IEEE Transactions on Information Technology in Biomedicine, vol. 5, pp. 10-13, (2001).

[6] Huang, Z., Sun, C., Purvis, M. and Cranefield, S., “View-based Consistency and its Implementation”, In Proceedings of the First IEEE/ACM International Symposium on Cluster Computing and the GRID, (2001).

[7] Kambalakatta, R., Kumar, M. and Das, S.K., “Profile Based Caching to Enhance Data Availability in Push/Pull Mobile Environments”, In Proceedings of the First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous), pp. 74-83, (2004).

[8] Mishra, A., Shin, M-H. and Arbaugh, W.A., “Context Caching using Neighbour Graphs for Fast Handoffs in a Wireless Network”, in Proceedings of INFOCOM. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies, (2004).

[9] O'Donoghue, J. and Herbert, J. “Profile Based Sensor Data Acquisition in a Ubiquitous Medical Environment”, In Proceedings of UbiCare 2006 workshop in conjunction with IEEE PerCom, pp. 570-574, (2006).

[10] O'Donoghue, J., Herbert, J. and Kennedy R., “Data Consistency within a Pervasive Medical Environment”, In Proceedings of IEEE Sensors Conference, (2006).

[11] Pitoura, E., Bhargava, B. K., “Data Consistency in Intermittently Connected Distributed Systems” In Proceedings of IEEE Transactions on Knowledge and Data Engineering, pp. 896-915, (1999).

[12] Wooldridge, M., “Agent-Based Software Engineering”, In IEE Proceedings of Software Engineering, no. 144, vol. 1, pp. 26-37, (1997).