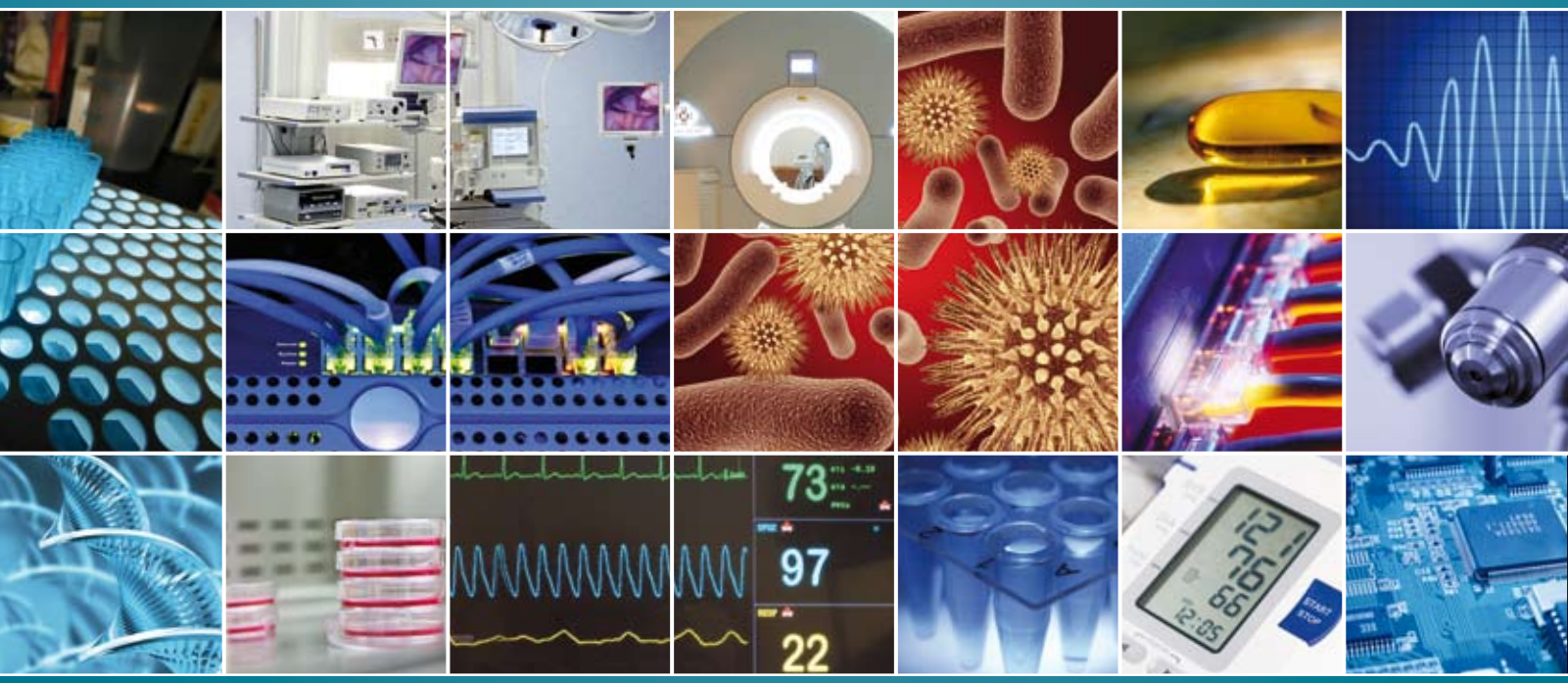
International Journal of Telemedicine and Applications

Electronic Health

Guest Editors: Hui Chen, Arnauld Nicogossian, Silas Olsson, Azhar Rafiq, Max E. Stachura, Mamoru Watanabe, Pamela Whitten, and Yang Xiao



Electronic Health

International Journal of Telemedicine and Applications

Electronic Health

Guest Editors: Hui Chen, Arnauld Nicogossian, Silas Olsson, Azhar Rafiq, Max E. Stachura, Mamoru Watanabe, Pamela Whitten, and Yang Xiao

Contents

Electronic Health, Hui Chen, Arnauld Nicogossian, Silas Olsson, Azhar Rafiq, Max E. Stachura, Mamoru Watanabe, Pamela Whitten, and Yang Xiao
Volume 2009, Article ID 308710, 1 page

Primary Care Physicians' Experience with Electronic Medical Records: Barriers to Implementation in a Fee-for-Service Environment, D. A. Ludwick and John Doucette
Volume 2009, Article ID 853524, 9 pages

Managing Requirement Volatility in an Ontology-Driven Clinical LIMS Using Category Theory, Arash Shaban-Nejad, Olga Ormandjieva, Mohamad Kassab, and Volker Haarslev
Volume 2009, Article ID 917826, 14 pages

Security Framework for Pervasive Healthcare Architectures Utilizing MPEG-21 IPMP Components, Anastasios Fragopoulos, John Gialelis, and Dimitrios Serpanos
Volume 2009, Article ID 461560, 9 pages

Agent-Oriented Privacy-Based Information Brokering Architecture for Healthcare Environments, Abdulmutalib Masaud-Wahaishi and Hamada Ghenniwa
Volume 2009, Article ID 101382, 17 pages

Enhancing E-Health Information Systems with Agent Technology, Minh Tuan Nguyen, Patrik Fuhrer, and Jacques Pasquier-Rocha
Volume 2009, Article ID 279091, 13 pages

An Adaptive Source-Channel Coding with Feedback for Progressive Transmission of Medical Images, Jen-Lung Lo, Saeid Sanei, and Kianoush Nazarpour
Volume 2009, Article ID 519417, 12 pages

Delay Analysis of GTS Bridging between IEEE 802.15.4 and IEEE 802.11 Networks for Healthcare Applications, Jelena Mišić and Xuemin (Sherman) Shen
Volume 2009, Article ID 987046, 13 pages

Using Discharge Abstracts to Evaluate a Regional Perinatal Network: Assessment of the Linkage Procedure of Anonymous Data, Catherine Quantin, Cyril Ferdynus, Paul Avillach, Béatrice Gouyon, Paul Sagot, and Jean-Bernard Gouyon
Volume 2009, Article ID 181842, 7 pages

Evaluation of a Telemedicine System for the Transmission of Morpho/Immunological Data Aiming at the Inclusion of Patients in a Therapeutic Trial, Jean-François Lesesve and Richard Garand
Volume 2009, Article ID 767145, 8 pages

Hindawi Publishing Corporation
International Journal of Telemedicine and Applications
Volume 2009, Article ID 308710, 2 pages
doi:10.1155/2009/308710

Editorial
Electronic Health

Hui Chen,[1] Arnauld Nicogossian,[2] Silas Olsson,[3] Azhar Rafiq,[4] Max E. Stachura,[5] Mamoru Watanabe,[6] Pamela Whitten,[7] and Yang Xiao[8]

[1] *Department of Mathematics and Computer Science, Virginia State University, P.O. Box 9068, Petersburg, VA 23806, USA*
[2] *Center for the Study of International Medical Policies and Practices, School of Public Policy, George Mason University, Fairfax, VA 22030, USA*
[3] *AAL International Association, Bischoffsheimlaan 25, 1000 Brussels, Belgium*
[4] *Center for Human Simulation and Patient Safety, School of Medicine, Virginia Commonwealth University, P.O. Box 980480, Richmond, VA 23298, USA*
[5] *Center for Telehealth, Medical College of Georgia, Augusta, GA 30912, USA*
[6] *Health Innovation and Information Technology Centre, Faculty of Medicine, University of Calgary, Calgary, AB, Canada T2N 1N4*
[7] *College of Communication Arts and Sciences, Michigan State University, MI 48824, USA*
[8] *Department of Computer Science, University of Alabama, P.O. Box 870290, Tuscaloosa, AL 35487-0290, USA*

Correspondence should be addressed to Hui Chen, huichen@ieee.org

Received 12 February 2009; Accepted 12 February 2009

Computing and networking techniques and technologies are gradually penetrating into every aspect of health care and medical practice. Many believe that important benefits can be achieved with the marriage of these two areas, including the delivery of high-quality health care at lower cost. Electronic health (e-Health) has become a very important area of focus and activity in multiple domains, such as health promotion, health care and maintenance, public health, medical science, health service, data management, image processing, telecommunication, wireless network, and operational research. The purpose of the special issue is to provide a high quality exposure to advances and experiences in e-Health related topics and to stir interest in advancing e-Health to the next level.

The call for papers for this special issue attracted numerous exciting responses from the research community. Submissions covered diverse areas of interests. Unfortunately, space limitations allowed only nine papers to be selected for inclusion in this special issue, after extensive peer review. In the following paragraphs, we briefly introduce the nine accepted papers.

Health information technology plays an important role in supporting decision making, health care delivery, and management of health services. Many sociotechnical factors affect physicians' adoption and implementation of health information systems. Having conducted semistructured interviews with 26 physicians from nine medical clinics in Alberta, Canada, D. A. Ludwick and John Doucette present a discussion of the barriers to implementing health information systems in a fee-for-service environment in their paper entitled "Primary care physicians' experience with Electronic Medical Records."

A potential obstacle impairing the adoption of health information systems might be simply that developed systems do not meet user needs. A. Shaban-Nejad et al. examined system requirement management in their paper "Managing Requirement Volatility in an Ontology-Driven Clinical Laboratory Information Management System (LIMS) Using Category Theory."

Security and privacy are important requirements for health information systems. For example, patients' privacy is protected by law in many countries. In Anastasios Fragopoulos et al. paper entitled "Context Aware Security for Pervasive Healthcare Architectures Utilizing MPEG-21 IPMP Components," MPEG-21 intellectual property management and protection components are used to achieve protection of transmitted medical information and enhance patient privacy.

Abdulmutalib Masaud-Wahaishi and Hamada Ghenniwa present an information brokering architecture that supports privacy-based information gathering in healthcare in their paper entitled "Agent-Oriented Privacy-Based Information Brokering Architecture for Healthcare Environments." In their proposed architecture, a brokering service is modeled as an agent with a specific architecture and interaction protocol appropriate to serve various requests.

Agent technology is gaining momentum in modeling and developing complex software systems. The next paper discusses major features and benefits of an agent-based approach to enhance a hospital laboratory legacy information system. This paper entitled "Enhancing E-Health Information Systems with Agent Technology" is contributed by Minh Tuan Nguyen et al.

The next two papers heavily focus on roles of communications and networking aspects in e-Health systems. In their paper entitled "An adaptive Source-Channel Coding with Feedback for Progressive Transmission of Medical Images," Jen-Lung Lo et al. studied source-channel coding problem with the consideration of characteristics of medical images.

IEEE 802.15.4 standard specifies the physical layer and the media access control layer for wireless personal area networks with low data rate and low power consumption. It forms the basis for many body sensor networks. IEEE 802.11 standards are the most widely adopted wireless local area network (LAN) standards. Jelena Misic and Xuemin (Sherman) Shen modeled an interconnected network consisting of IEEE 802.15.4 body sensor networks and IEEE 802.11 wireless LANs. Their paper is entitled "Delay Analysis of GTS Bridging between IEEE 802.15.4 and IEEE 802.11 Networks for Healthcare Applications."

The last two papers provide us with the examples of e-Health systems at work. C. Quantin et al. show that it is possible to set up a continuous and exhaustive recording system for linked perinatal data to assess the quality of care on a regional scale in their paper entitled "Using discharge abstracts to evaluate a regional perinatal network: assessment of the linkage procedure of anonymous data—using discharge abstracts to evaluate a regional perinatal network."

Jean-François Lesesve and Richard Garand demonstrate one of the first approaches to the use of telehematology for the quality control of diagnosis using the GOELAMS Chronic Lymphocytic Leukaemia 98 trial. Their paper "Evaluation of a Telemedicine System for the Transmission of Morpho/Immunological Data Aiming at the Inclusion of Patients in a Therapeutic Trial" details their experience and the benefits arising from the use of telehematology in clinical practice.

In summary, this special issue includes papers that span diverse areas of interests including system development, security and privacy, actual effect of e-Health in practice, as well as technological foundations. It was our honor to receive submissions from many authors. Without our unselfish reviewers, who provided us with extensive and constructive reviews, it would not be possible to run a special issue of such broad scope. Last, but not least, the publishing staff have worked diligently with us on this special issue. We are, therefore, indebted to all the authors, the reviewers, and publishing staff. We would like to express our sincere gratitude to them all.

Hui Chen
Arnauld Nicogossian
Silas Olsson
Azhar Rafiq
Max E. Stachura
Mamoru Watanabe
Pamela Whitten
Yang Xiao

Hindawi Publishing Corporation
International Journal of Telemedicine and Applications
Volume 2009, Article ID 853524, 9 pages
doi:10.1155/2009/853524

Research Article

Primary Care Physicians' Experience with Electronic Medical Records: Barriers to Implementation in a Fee-for-Service Environment

D. A. Ludwick[1,2] and John Doucette[1,3]

[1] *Department of Mechanical Engineering, Faculty of Engineering, University of Alberta, Edmonton, Alberta, Canada T6G 2G8*
[2] *Sherwood Park - Strathcona County Primary Care Network, Sherwood Park, Alberta, Canada T8A 4W6*
[3] *TRLabs, #401, 9426, 51 Avenue NW, Edmonton, Alberta, Canada T6E 5A6*

Correspondence should be addressed to D. A. Ludwick, dave_ludwick@shaw.ca

Received 29 April 2008; Revised 19 September 2008; Accepted 8 October 2008

Recommended by Hui Chen

Our aging population has exacerbated strong and divergent trends between health human resource supply and demand. One way to mitigate future inequities is through the adoption of health information technology (HIT). Our previous research showed a number of risks and mitigating factors which affected HIT implementation success. We confirmed these findings through semistructured interviews with nine Alberta clinics. Sociotechnical factors significantly affected physicians' implementation success. Physicians reported that the time constraints limited their willingness to investigate, procure, and implement an EMR. The combination of antiquated exam room design, complex HIT user interfaces, insufficient physician computer skills, and the urgency in patient encounters precipitated by a fee-for-service remuneration model and long waitlists compromised the quantity, if not the quality, of the information exchange. Alternative remuneration and access to services plans might be considered to drive prudent behavior during physician office system implementation.

1. Introduction

Aging populations with complex health conditions such as obesity and chronic disease place an increasing burden on primary care systems in many countries [1–5]. While demand escalates, health human resource supply is declining as Canada's health workforce retires earlier and the average age of the remaining working population increases [6–8]. Higher training requirements, tuition fees [9], certification requirements [8, 10] and a higher female to male enrolment ratio are leading to a decline in the primary care physician workforce. The adoption of health information technology (HIT) is seen as one way to address the widening health care demand and supply gap [11–13]. It seems intuitive that HIT would improve patient safety, improve physician office efficiency and mitigate shortages in health human resources, but studies have shown that such systems can compromise short-term physician office performance [14–16], intimidate physicians and their office staff [17], and have shown, on occasion, to increase medical errors [18–20].

Health information system adopters face several risks when implementing health systems [21]. The purpose of this project was to assess the relevance and impact of these risks in the context of primary care in Sherwood Park, Alberta. Due to an economic boom, the population of Sherwood Park grew 14% from 2001 to 2006 compared to a population growth rate of 10% in Edmonton [22] (the nearest city) and 5.4% in Canada [23] during the same period. The Sherwood Park circumstance offers a microcosm in which to study the effects of HIT adoption in primary care.

Primary care usually refers to family or general practice and is the first point of contact a person has with the health system [24]. An *electronic medical record* is a computerized health information system where providers record detailed encounter information such as patient demographics, encounter summaries, medical history, allergies,

intolerances, and lab test histories. Some EMRs support scheduling, billing, reporting, order entry, results management, and decision support [25–27]. Such systems are often referred to as *physician office systems* or *practice management systems* [25, 28].

2. Previous Research

In a previous comprehensive literature review [21], the implications of HIT were examined across a number of care domains. Health information technology implementation success depends on a number of factors. Implementers need to be aware of sociotechnical system fit to achieve success [29–32]. However, implementers perceive privacy [33, 34], patient safety, provider/patient relations, staff anxiety [31], time needed to implement [35–39], quality of care, financial [40–42], efficiency, and liability [43] factors as risks that can pressure or derail a project. Users' previous experiences with HIT affected their experience with a new system, both positively and negatively [44–46]. Users applied their previous experience to new systems and evaluated the usability and effectiveness of their new system against that of the previous system. Exam room layouts and computer monitor placement have been shown to affect, positively and negatively, the interaction between provider and patient [47–49]. Implementers can insulate the project from such risks by establishing strong leadership [16, 37, 45, 50–54], using project management techniques [50, 51, 55–61], establishing standards, and training their staff [13, 16, 35–39, 46, 52, 54, 58, 62–65] to ensure such risks do not compromise implementation success.

3. Methodology

This research project used one hour semistructured interviews to acquire information from primary care physicians' experience of selecting, implementing, and operating an EMR system. Physician candidates were selected from our local primary care network, in which 47 physicians are members. Inclusion criteria required physicians to be practicing full time in the community, have significant EMR experience and be a lead physician or influencer in clinic decision making. Physicians were paid a honorarium to acknowledge their lost revenue generating opportunity. An interview guide consisting of closed-ended statistical questions and several open-ended questions stimulated a qualitative conversation regarding the experience. The researcher recorded detailed notes that were later used for synthesis and analysis.

After the interview, the researcher documented the layout of exam rooms. Exam rooms were depicted in a floor plan conceptually locating the computer keyboard and monitor with respect to the patient exam table or chair. The researcher also asked physicians to recount the positioning of the patient with respect to themselves and the computer. Exam room layouts were subsequently analyzed and categorized into three general types for critical review. Figures 1, 2, and 3 depict the three exam room layouts which best illustrate the wide range of layouts. The researcher recorded the quantity of rooms in the clinics for statistical purposes.

4. Results

Of the 47 physicians in the primary care network, 19 physicians are clinic leads. Of the 19 clinic owners or influencers, there are 11 clinics with practical EMR experience. Nine interviewees were selected who represent a total of 26 physicians and were interviewed during the months of February and March 2008 using the interview guide shown in Table 1. Two interviewees were sole practitioners, 2 interviewees represented clinics with 2 physicians each in them, 3 interviewees represented practices with of 3 physicians, one interviewee represented a clinic of 5 physicians, and one interviewee represented a clinic of 6 physicians. Table 2 summarizes the key findings from the interviews' closed-ended questions.

All physicians have at least 10 years of practice experience. Two physicians were female, both of which are operating in multiphysician practices. All interviewees except one considered themselves owners or decision makers in the practice but all reported that they had a hand in selecting and implementing their EMR. Eight physicians are satisfied with their own computer and data entry skills, rating themselves a 3 out of 5 or higher. Physicians have 30 patient encounters per day but, often see as many as 40 patients at roughly 10 minutes per encounter.

Seven physicians routinely make encounter notes directly into their EMR during the interview, although occasionally they complete note taking outside the room after the encounter. The other 2 physicians make notes on paper. Six physicians have permanently stationed desktop computers located in exam rooms to make notes, while two use wireless laptop computers. One of the two physicians using a paper system has computers stationed in his exam rooms but has reverted back to record patient encounter data on paper. One clinic reported that physicians wrote encounter notes on paper and scanned them into their EMR as a way to kick start the implementation and develop their computer skills. Eight clinics use paper record systems prior to their EMR, while one clinic is now operating its second EMR.

Eight physicians did not follow a prescribed procurement plan while the other followed a procurement plan consisting of a market scan, price analysis, vendor demonstrations, and visiting colleagues' clinics. Four physicians invited vendors to demonstrate their products to them at their clinics. Two physicians completed a price comparison, while one called their professional association for procurement advice, another acquired his EMR through personal connections, and yet another could not remember how he had selected his EMR. Physicians did not have the time or experience to follow a detailed procurement plan. All physicians reported disorientation in the procurement process as they had not had any related experiences in the past.

Physicians did not report the breakpoint that McGrath had reported [47]. Even though physicians said patients rarely commented, some physicians felt a need to apologize for taking notes on computer, or at least to acknowledge it to patients. Those physicians who had owned their system for a while were more comfortable since most patients had rotated through and seen the system previously. Physicians

Table 1: Interview guide.

Interview questions for physician interviews		
Sherwood Park - Strathcona County PCN		
Interviewee(s):		
Interviewer:		
Date:		
Interview questions	Notes of candidate's answer	Interviewer's guide to answers
How long have you been in practice?		[Years or Months]
How many physicians are currently practicing in your office?		[Number]
How many non-physician clinicians do you employ?		[Number]
How many staff/admin do you employ?		[Number]
Are you the practice owner/key decision maker? If not, what is your role?		[Yes/No]. [If no, partner, contracted, part time]
How many patients do you typically see in a day?		[Number]
What is your target interview duration?		[Minutes]
What sort of health records system do you currently use?		[paper; electronic, but paper used to record notes first followed by transcription; electronic, desktop in exam room; electronic, laptop carried into exam room]
Can you describe the role your health information system plays when you are interviewing a patient		[take paper based notes as I go, take e-based notes as I go, don't take any notes in interview]
How long have you owned your EMR?		[Years or Months]
On a scale of 1 to 5, where 1 is poor and 5 is excellent, can you rate your computer skills (before and after the implementation)?		[1 to 5], [1 to 5]
When/where do you make your encounter notes?		[during interview in exam room, immediately after interview outside exam room door, at end of day either at the office or at home]
Prior to your current practice, what did you use for health information system to support your work?		[paper; electronic, but paper used to record notes first followed by transcription; electronic, desktop in exam room; electronic, laptop carried into exam room]
Can you describe the process you went through to buy your EMR? How did you gather market information?		[market scan, called vendors directly, talked to colleagues, talked to AMA/POSP/CPSA]
How did you select your EMR? What purchasing factors were most relevant to you?		[price, features, eligibility for financial support]
How did you install the EMR into your practice?		[big-bang, pilot, team-oriented integrative approach]
What do you use your EMR system for?		[Billing, scheduling, encounter note taking, lab results, order entry, contraindication management]
Where do you get your technical support?		[self, colleague, 3rd party]
What do you like/dislike about your current system?		—
Did you notice a change in your patient volumes after your implementation? If so, can you say what % age it dropped to and for how long? Why?		[%, months]
On a scale of 1 to 5, where 1 is completely dissatisfied and 5 is extremely satisfied, what would you say your overall satisfaction is with your system?		[1 to 5]
Knowing what you know now, would you still have bought the EMR? Why do you say that?		[yes/no]

Table 2: Closed-ended interview results: statistics describing the number and experience of physicians, patient throughput, years of experience using an EMR, computer skills, and clinic size.

Interview factor	Average ($N = 9$)	Range ($N = 9$)
Years in practice	20 years	10 to 33 years
Number of physicians practicing in clinic	3	1 to 6
Number of nonphysician clinicians in practice	1.75	0 to 8
Number of support staff in clinic	2.5	1 to 4
Target number of patients to be seen in a day	32.5	20 to 40
Target patient interview duration	8.4 minutes	7.5 to 15 minutes
Number of years owned an EMR	4 years	0 to 10 years
Personal rating of computer skills (range: 1 to 5)	3.25	2 to 4
Overall EMR satisfaction rating (range: 1 to 5)	2.9	2 to 4.5

felt compelled to stop typing if patients became emotional during the interview, although they did not always do so.

Our physicians complained about their training and postsale experience with their vendor. Instead of a training regimen similar to that described in the literature [21], physicians reported that their vendor simply offered one training session of one half to a full day in duration. Training was often too soon after implementation. Physicians had not developed sufficient experience with their new EMR to ask relevant questions or appreciate the answers.

Physicians reported that they could not always access vendor technical support. Even when they could get a person, they were not confident that the technical support person "knew how a clinical practice functioned." Physicians were concerned that the company did not appreciate the implications of a dysfunctional EMR. Physicians often procured supplementary local technical support at higher cost.

Physicians pointed to opportunities for more efficient data entry. Two physicians have made great use of the template features in their EMRs. They have spent significant time building templates which allow them to enter data or orders into their system for common ailments with a few key strokes. Two other physicians reported that they have made use of voice recognition software which emulates dictation, a familiar mode of data entry for physicians. Voice recognition software requires training and is not functional for clinicians with strong accents, but physicians who invested significant time training their software had achieved a satisfactory level of efficiency.

A total number of 19 examination rooms were viewed during the study representing 51% of the total 37 rooms in these physicians' offices. Figures 1 to 3 depict three exam room layouts which were categorized based on the following observations:

(i) the presence of an office desk, or not,

(ii) the presence of a patient interview chair, or not,

(iii) the general size of the room,

(iv) the orientation of the computer monitor with respect to the physician, the exam table, and the patient,

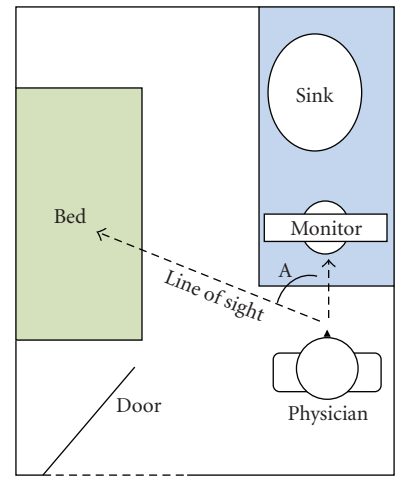


Figure 1: Exam room layout 1.

Table 3 below summarizes the key characteristics of the observed exam rooms. The layout-type column indicates the figure which best depicts the exam room. Eight out of 9 physicians interview patients while they are seated in chairs (one owns rooms numbered 13 and 14 which were too small for a chair and therefore interviewed and examined the patient on bed). Note that two columns in the table refer to an angle. Exam room observations note that an angle was created between the lines of sight from physician to monitor and physician to bed (Angle A) as well as between the lines of sight from physician to monitor and physician to chair (Angle B), if the chair existed.

Many brands of EMR are used in these clinics (Telin, Global Biometrics, Med Access, Practice Solutions, EMIS, and Wolf). Clinics use most system features including billing, scheduling, importing lab results, drug order entry, and encounter note taking. Drug-to-drug and drug-to-allergy contraindication management was used by many physicians when the data had been entered to support it. Three physicians do not use contraindication management because they leave this responsibility to the pharmacy. One practice reported that this feature had to be purchased separately so was not currently part of their system. Many physicians

Table 3: Exam room layout results.

Room no.	Layout type	System configuration	Has chair	Has office desk	Room size	Angle A (monitor to bed)	Angle B (monitor to chair)
1	2	Desktop	Yes	Yes	Medium	120	90
2	2	Desktop	Yes	Yes	Medium	120	90
3	2	Desktop	Yes	Yes	Medium	120	90
4	3	Desktop	No	Yes	Large	120	n/a
5	2	Desktop	Yes	Yes	Large	120	120
6	1	Laptop	Yes	No	Medium	0	0
7	1	Laptop	Yes	No	Medium	0	0
8	1	Laptop	Yes	No	Medium	0	0
9	2	Desktop	Yes	Yes	Small	180	90
10	2	Desktop	Yes	Yes	Small	180	90
11	3	Desktop	Yes	Yes	Large	180	180
12	2	Desktop	Yes	Yes	Large	120	90
13	1	Desktop	No	Yes	Small	90	n/a
14	1	Desktop	No	Yes	Small	90	n/a
15	2	Desktop	Yes	Yes	Medium	180	90
16	2	Desktop	Yes	Yes	Medium	180	90
17	1	Laptop	Yes	No	Medium	0	0
18	1	Laptop	Yes	No	Medium	0	0
19	1	Laptop	Yes	No	Medium	0	0

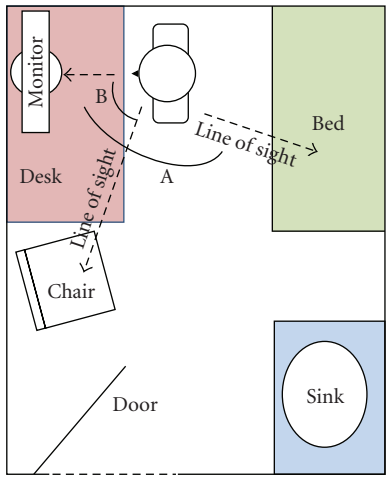


Figure 2: Exam room layout 2.

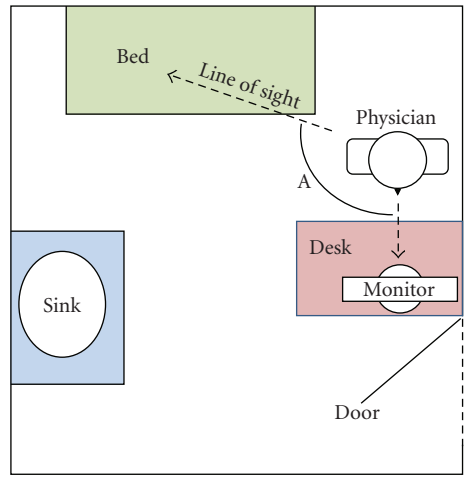


Figure 3: Exam room layout 3.

automatically receive lab test results electronically through an electronic mailbox system (ftp-based system) arranged by the RHA.

5. Discussion

The purpose of this project was to assess the relevance and impact of risk and insulating factors for HIT adoption in the context of primary care in Sherwood Park, Alberta. Our interviews showed that few physicians follow a complete procurement approach. Exam room layouts require computer systems to be situated such that physicians face away from their patients. Physicians struggle to get appropriate training and technical support for their systems. However, when physicians invest the time, they realize benefits to using their EMR.

Time constrains many physician offices when procuring and implementing HIT. In Canada, primary care physicians get paid on a fee-for-service basis. The more patients they see, the more revenue they generate. Further, Canada reports

large wait times for access to health services [66]. Physicians choose not to invest the time in systems procurement because they are uncomfortable with the process. Investigating systems during office hours reduces revenue generating opportunity and increases patient wait times. Interestingly, other reviews have shown that pay-for-performance models, as one strategy for payment, have worked well in driving to long-term national HIT adoption success [67]. More research on the effects of remuneration models on adoption is warranted and will be the subject of future research.

Interviews revealed that exam room layouts could compromise the quantity, if not quality, of information transfer from patient to physician. Our figures above attempt to simplify and categorize these into three types based on the type and placement of furniture, the type and placement of the computer monitor, as well as the positioning of the physician with respect to the patient. If the amount of interpersonal communication is a function of visual cues, as would be suggested by Mehrabian [68], then the Angle A, created between the two lines of sight from the physician to their computer monitor and the physician to the patient, would be critical to the success of communications. Layout 1 has a relatively small angle (estimated at 60 degrees). Layouts 2 and 3 show Angle A to be greater than 90 degrees. This situates the patients somewhat behind the physician as they face the monitor. Physicians operating in exam rooms similar to that of layout 1 expressed the least concern over eroded interpersonal communications. Furthermore, the two physicians using laptops could position themselves to look over their laptop monitor directly at the patient, effectively reducing Angle A to zero degrees. We did not interview enough physicians to be conclusive, but we assert that there could be a relationship between the quantity, and possibly quality, of information transfer from patient to physician and the size of Angle A, as would be supported by Robinson et al. [69]. The smaller Angle A is, the more direct patient eye contact is and, therefore, the more complete the interpersonal communication, possibly leading to higher quality of care. A few physicians appreciated this concept as one had previously taken advantage of pending renovations to accommodate her systems implementations and another was planning changes to his office furniture to close Angle A to zero degrees.

The above problem gets more aggravated when we consider our physicians' computer skills in the context of the complex EMR user interfaces and the time pressure of a patient encounter within the context of a fee-for-service remuneration model. Our physicians self-reported their computer skills rated at 3, on a scale of 1 to 5. Similarly, a US survey [70] reported that their physician survey respondents felt quite confident about their computer skills. We did not observe physicians using their EMR for note taking during patient encounters (exam room observation would have required significant ethics approval); however, extrapolating complaints they had about the usability of basic computer functions make us hypothesize that physicians, vendors, and HIT advocates have underestimated the level of computer skills required for this work (physicians reported that they hunt for menus and buttons to the extent they sometimes stop using the EMR in interviews because of the disruption). EMR user interfaces are complex and busy (reminiscent of an airplane cockpit). The skills needed to listen to patients' complaints, assess medical relevance, contemplate interventions as well as type notes—all at the same time–would require a significant level of concentration, typing skills, and familiarity with the application's user interface, not normally found in the most adept computer users. Therefore, we were not surprised to learn that physicians often had to complete note taking after the encounter or at the end of the business day. We hypothesize that HIT can disrupt the flow of information from patient to provider when computer monitors require the physician to face away from the patient. Physicians' eyes are focused on the computer system and not the patient which compromises information transfer especially in clinics with high-patient volumes and inexperienced physician computer users. We are concerned that this may compromise the physicians' implied and historic role as confidante. We are planning future research to investigate this concept further.

The study's most obvious weakness is its narrow field of interviewees. Sherwood Park PCN has over 40 physicians; however, only nine met our inclusion criteria. The small sample means that the discussion and conclusions outlined above can only be considered directional. They are not conclusive or statistically significant. Bias may result from interviewee selection. Ideally, interviewees would have represented more clinics from a greater geographical area. We interviewed physician leaders who influenced implementation decisions; yet, physician leaders' perceptions may not reflect those of their associates. Our physicians are members of a PCN; consequently, findings may not be applicable to primary care physicians who practice outside of an interdisciplinary team. We infer that there is a relationship between information transfer and the angles described above. Future research involving patients is required to confirm this. This Alberta study is influenced by provincial matters, such as health policy, remuneration approaches, and physician office system funding models, which may prevent results from applying in other jurisdictions.

6. Conclusions

Our interviews and previous research have shown time to be a precious resource for physicians in several facets of their day-to-day operations. Physicians do not take the time to properly become familiar with the available products, select an EMR, implement it, and then train to use it even though colleagues have invested time and realized great benefit. We wonder whether the current fee-for-service payment model in Alberta creates an urgency to maintain patient throughput. The opportunity to maximize clinic revenue and waiting rooms full of patients may discourage physicians from investing the time in EMR implementation activities. The Sherwood Park experience might point to a need for a change in remuneration approach and guidance for reducing wait times, at least for the purposes of selecting, acquiring, and implementing the system prior to returning to steady-state clinic operations.

Computer skills, complexity in EMR interfaces, and exam room layouts combine to affect physicians' encounter experience. Despite their strong self assessments, we are concerned that physicians do not have sufficient computer skills to take notes and navigate an EMR while listening to a patient in an encounter. Physicians might consider changing to laptop systems (even with wired networks), using voice recognition software and/or developing templates to permit more direct patient interaction and improve efficiency.

7. Relevance

Alberta, like other jurisdictions, is aggressively driving the adoption of HIT. Despite well structured and financed programs, factors such as computer aptitude in physicians and complexity in graphical user interfaces are not being considered as hindrances to adoption. Medical associations provide valuable coaching to physicians on system procurement and physician office design, but time constrains physicians from taking advantage. Vendor certification programs test and conform EMR applications for interoperability but need to increase scrutiny on vendor business and technical support qualifications. Although jurisdictions continue to finance adoption, organized effort needs to be applied to other points of friction. Training for physicians on computers, establishing user interface design standards and guidance on exam room design is also required.

Canada's fee-for-service payment model provides physicians with an opportunity to maximize patient throughput. Yet, HIT projects take physicians offline from their core activities as physicians. When physicians are remunerated based on patient volume, they are discouraged from spending the time needed to make their implementations a success. This paper does not advocate one payment model over another, but simply points to a pattern of behavior which seems to be caused by the current approach. Jurisdictions might consider the implications of the current payment model with regard to adoption and provide alternative vehicles which encourage physicians to invest the time to maximize outcomes from their investments.

References

[1] The Change Foundation, "Rising Tide—Understanding Demand in Health Care," Toronto, Ontario, Canada, May 2005.

[2] "U.S. Interim Projections by Age, Sex, Race, and Hispanic Origin: 2000–2050," Washington, DC: U.S. Census Bureau; 2004, http://www.census.gov/ipc/www/usinterimproj/.

[3] S. Starky, "The Obesity Epidemic in Canada," Library of Parliament. Parliamentary Information and Research Service, July 2005.

[4] Alberta Agriculture and Food, "Canadian Consumer Trends in Obesity and Food Consumption," February 2008, http://www1.agric.gov.ab.ca/$department/deptdocs.nsf/all/sis8438.

[5] Statistics Canada. The Daily, "Canadian Community Health Survey: Obesity among children and adults," February 2008, http://www.statcan.ca/Daily/English/050706/d050706a.htm.

[6] Canadian Institute for Health Information, "Analytical Bulletin: 2004 National Physician Survey Response Rates and Comparability of Physician Demographic," 2005.

[7] The College of Family Physicians of Canada, Canadian Medical Association, The Royal College of Physicians and Surgeons of Canada, "2007 National Physician Survey: National Demographics," Ottawa, Ontario, Canada, 2007.

[8] Canadian Institute for Health Information, "Canada's Health Care Providers," Ottawa, Ontario, Canada, 2001.

[9] Canadian Institute for Health Information, "Health Care in Canada," Ottawa, Ontario, Canada, 2001.

[10] The Change Management Foundation, "The Demand Management Think Tank. Backgrounder on Demand Management," Toronto, Ontario, Canada, January 2002.

[11] S. D. Backgrounder, "Infoway's Vision," Canada Health Infoway, 2006.

[12] D. K. McInnes, D. C. Saltman, and M. R. Kidd, "General practitioners' use of computers for prescribing and electronic health records: results from a national survey," *Medical Journal of Australia*, vol. 185, no. 2, pp. 88–91, 2006.

[13] Health Canada, "EMR Toolkit: implementing electronic medical records in primary health care settings," January 2008, http://www.emrtoolkit.ca.

[14] C. Bossen, "Test the artefact—develop the organization: the implementation of an electronic medication plan," *International Journal of Medical Informatics*, vol. 76, no. 1, pp. 13–21, 2007.

[15] B. Gamble, "Using IT to make primarycare reform work for you," *Future Practice*, pp. 10–11, April 2006.

[16] D. M. Schuster, S. E. Hall, C. B. Couse, D. S. Swayngim, and K. Y. Kohatsu, "Involving users in the implementation of an imaging order entry system," *Journal of the American Medical Informatics Association*, vol. 10, no. 4, pp. 315–321, 2003.

[17] "Personal observations and conversations with Sherwood Park—Strathcona County PCN primary care physicians," November 2007 to March 2008.

[18] Y. Y. Han, J. A. Carcillo, S. T. Venkataraman, et al., "Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system," *Pediatrics*, vol. 116, no. 6, pp. 1506–1512, 2005.

[19] R. Koppel, J. P. Metlay, A. Cohen, et al., "Role of computerized physician order entry systems in facilitating medication errors," *The Journal of the American Medical Association*, vol. 293, no. 10, pp. 1197–1203, 2005.

[20] D. A. Ludwick, "Computerized Physician Order Entry in Pediatric Intensive Care Units: A Critical Appraisal of Three Articles Illustrating the Importance of Organizational Considerations in a CPOE Implementation, Version 1.7," October 2007.

[21] D. A. Ludwick and J. Doucette, "Adopting electronic medical records in primary care: lessons learned from health information systems implementation experience in seven countries," *International Journal of Medical Informatics*. In press.

[22] Statistics Canada, "Census of Population and Housing. 2006 Census Profile for Edmonton—Sherwood Park, FED," Produced by BC Stats, October 2007, http://www.bcstats.gov.bc.ca/data/cen06/profiles/detailed/FED48016.pdf.

[23] Statistics Canada, "2006 Census release dates," March 2008, http://en.wikipedia.org/wiki/Canada_2006_Census.

[24] Primary Care Initiative, Alberta, "What is Primary Care?" September 2007, http://www.primarycareinitiative.ab.ca/content.asp?id=142.

[25] "Electronic medical record," Wikipedia, January 2008, http://en.wikipedia.org/wiki/Electronic_medical_record.

[26] "Practice management," Wikipedia, January 2008, http://en.wikipedia.org/wiki/Practice_management .

[27] The College of Family Physicians of Canada, "Electronic Medical Record," December 2007, http://toolkit.cfpc.ca/en/information-technology/electronic-medical-record.php.

[28] Physician Office System Program, "Testimonials," December 2007, http://www.posp.ab.ca/testimonials/.

[29] J. S. Ash, D. F. Sittig, R. H. Dykstra, K. Guappone, J. D. Carpenter, and V. Seshadri, "Categorizing the unintended sociotechnical consequences of computerized provider order entry," *International Journal of Medical Informatics*, vol. 76, supplement 1, pp. S21–S27, 2007.

[30] M. Reddy, W. Pratt, P. Dourish, and M. M. Shabot, "Sociotechnical requirements analysis for clinical systems," *Methods of Information in Medicine*, vol. 42, no. 4, pp. 437–444, 2003.

[31] A. X. Garg, N. K. J. Adhikari, H. McDonald, et al., "Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review," *The Journal of the American Medical Association*, vol. 293, no. 10, pp. 1223–1238, 2005.

[32] S. T. Rosenbloom, F. E. Harrell Jr., C. U. Lehmann, J. H. Schneider, S. A. Spooner, and K. B. Johnson, "Perceived increase in mortality after process and policy changes implemented with computerized physician order entry," *Pediatrics*, vol. 117, no. 4, pp. 1452–1455, 2006.

[33] Royal College of General Practitioners, "Putting IT into the NHS: all about NPfIT and the national "data spine"," *The New Generalist*, vol. 1, no. 3, pp. 59–60, 2003.

[34] E. Van Ekert and A. Took, "Cautionary Tales," July 2006.

[35] D. E. Trachtenbarg, "EHRs fix everything—and nine other myths," *Family Practice Management*, vol. 14, no. 3, pp. 26–30, 2007.

[36] R. Boaden and P. Joyce, "Developing the electronic health record: what about patient safety?" *Health Services Management Research*, vol. 19, no. 2, pp. 94–104, 2006.

[37] R. L. Wears and M. Berg, "Computer technology and clinical work: still waiting for godot," *The Journal of the American Medical Association*, vol. 293, no. 10, pp. 1261–1263, 2005.

[38] M. J. Ball, D. E. Garets, and T. J. Handler, "Leveraging information technology towards enhancing patient care and a culture of safety in the U.S," *Methods of Information in Medicine*, vol. 42, no. 5, pp. 503–508, 2003.

[39] N. Ferris, "IT no cure-all for prescription errors," Government Health IT. December 2007, http://www.govhealthit.com/online/news/89944-1.html.

[40] The College of Family Physicians of Canada, "Enabling Factors," December 2007, http://toolkit.cfpc.ca/en/information-technology/financial-support.php.

[41] J. Henderson, H. Britt, and G. Miller, "Extent and utilisation of computerisation in Australian general practice," *Medical Journal of Australia*, vol. 185, no. 2, pp. 84–87, 2006.

[42] Canadian Medical Association, "Have paper records passed their time?" *Canadian Medical Association Journal*, vol. 173, no. 7, p. 725, 2005.

[43] Certification Commission for Healthcare Information Technology, "CCHIT certified electronic health records reduce malpractice risk the case for offering liability insurance premium credits to physicians who use them," 2007.

[44] R. T. O'Connell, C. Cho, N. Shah, K. Brown, and R. N. Shiffman, "Take note(s): differential EHR satisfaction with two implementations under one roof," *Journal of the American Medical Informatics Association*, vol. 11, no. 1, pp. 43–49, 2004.

[45] A. Ahmad, P. Teater, T. D. Bentley, et al., "Key attributes of a successful physician order entry system implementation in a multi-hospital environment," *Journal of the American Medical Informatics Association*, vol. 9, no. 1, pp. 16–24, 2002.

[46] J. Øvretveit, T. Scott, T. G. Rundall, S. M. Shortell, and M. Brommels, "Improving quality through effective implementation of information technology in healthcare," *International Journal for Quality in Health Care*, vol. 19, no. 5, pp. 259–266, 2007.

[47] J. M. McGrath, N. H. Arar, and J. A. Pugh, "The influence of electronic medical record usage on nonverbal communication in the medical interview," *Health Informatics Journal*, vol. 13, no. 2, pp. 105–117, 2007.

[48] B. N. Doebbeling, A. F. Chou, and W. M. Tierney, "Priorities and strategies for the implementation of integrated informatics and communications technology to improve evidence-based practice," *Journal of General Internal Medicine*, vol. 21, supplement 2, pp. S50–S57, 2006.

[49] J. Holtschlag, "Successful implementation strategies," in *Ubiquity: Technologies for Better Health in Aging Societies*, pp. 705–710, IOS Press, Amsterdam, The Netherlands, 2006.

[50] R. Delaney, "Smiling all the way," *Health Care & Informatics Review*, vol. 8, no. 2, 2004.

[51] J. S. Luo, "Computer physician order entry: to implement or not?" *Primary Psychiatry*, vol. 13, no. 3, pp. 19–21, 2006.

[52] A. S. Gosling, J. I. Westbrook, and R. Spencer, "Nurses' use of online clinical evidence," *Journal of Advanced Nursing*, vol. 47, no. 2, pp. 201–211, 2004.

[53] E. G. Poon, D. Blumenthal, T. Jaggi, M. M. Honour, D. W. Bates, and R. Kaushal, "Overcoming barriers to adopting and implementing computerized physician order entry systems in U.S. hospitals," *Health Affairs*, vol. 23, no. 4, pp. 184–190, 2004.

[54] D. Protti, "International experiences in implementing 'single care' record approaches," NHS Connecting for Health, Management Brief no. 6, October 2007.

[55] M. A. Del Beccaro, H. E. Jeffries, M. A. Eisenberg, and E. D. Harry, "Computerized provider order entry implementation: no association with increased mortality rates in an intensive care unit," *Pediatrics*, vol. 118, no. 1, pp. 290–295, 2006.

[56] D. Martin, J. Mariani, and M. Rouncefield, "Implementing an HIS project: everyday features and practicalities of NHS project work," *Health Informatics Journal*, vol. 10, no. 4, pp. 303–313, 2004.

[57] D. Miranda, W. Fields, and K. Lund, "Lessons learned during 15 years of clinical information system experience," *Computers in Nursing*, vol. 19, no. 4, pp. 147–151, 2001.

[58] Doctor's Office Quality—Information Technology, "Common Implementation Problems," 2007.

[59] Doctor's Office Quality—Information Technology, "EHR Implementation Options Incremental vs. Big Bang: Comparing Approaches," 2007.

[60] Health Informatics Society of Australia, "A Vision for an Australian Healthcare System Transformed by Health Informatics," November 2007.

[61] C. Duggan, "Implementation evaluation: HIM professionals share their experiences bringing health IT online," *Journal of the American Health Information Management Association*, vol. 77, no. 6, pp. 52–55, 2006.

[62] American Academy of Family Physicians, Keys to Success, 2007.

[63] R. Didham and I. Martin, "A review of computerised information technology systems in general practice medicine," *Health Care & Informatics Review*, vol. 8, no. 1, 2004.

[64] Government News, "UK: Staff disengagement could stall UK's health IT programme, says study," August 2005.

[65] F. R. Elevitch, "SNOMED CT: electronic health record enhances anesthesia patient safety," *AANA Journal*, vol. 73, no. 5, pp. 361–366, 2005.

[66] Health Council of Canada, "Wading through Wait Times. What Do Meaningful Reductions and Guarantees Mean? An Update on Wait Times for Health Care," June 2007, http://healthcouncilcanada.ca/docs/rpts/2007/wait_times/hcc_wait-times-update_200706_FINAL ENGLISH.pdf.

[67] Canadian Medical Association, "IT in General Practice: A 10-country comparison," March 2008, http://www.cma.ca/index.cfm/ci_id/49047/la_id/1.htm.

[68] A. Mehrabian, *Silent Messages*, Wadsworth, Belmont, Calif, USA, 2008.

[69] T. E. Robinson II, G. L. White Jr., and J. C. Houchins, "Improving communication with older patients: tips from the literature," *Family Practice Management*, vol. 13, no. 8, pp. 73–78, 2006.

[70] R. L. Edsall and K. G. Adler, "User satisfaction with EHRs: report of a survey of 422 family physicians," *Family Practice Management*, vol. 15, no. 2, 2008.

Hindawi Publishing Corporation
International Journal of Telemedicine and Applications
Volume 2009, Article ID 917826, 14 pages
doi:10.1155/2009/917826

Research Article

Managing Requirement Volatility in an Ontology-Driven Clinical LIMS Using Category Theory

Arash Shaban-Nejad, Olga Ormandjieva, Mohamad Kassab, and Volker Haarslev

Department of Computer Science and Software Engineering, Concordia University, 1455 de Maisonneuve Boulevard West, Montreal, QC, Canada H3G 1M8

Correspondence should be addressed to Arash Shaban-Nejad, arash_sh@cs.concordia.ca

Received 17 June 2008; Revised 30 October 2008; Accepted 29 December 2008

Recommended by Hui Chen

Requirement volatility is an issue in software engineering in general, and in Web-based clinical applications in particular, which often originates from an incomplete knowledge of the domain of interest. With advances in the health science, many features and functionalities need to be added to, or removed from, existing software applications in the biomedical domain. At the same time, the increasing complexity of biomedical systems makes them more difficult to understand, and consequently it is more difficult to define their requirements, which contributes considerably to their volatility. In this paper, we present a novel agent-based approach for analyzing and managing volatile and dynamic requirements in an ontology-driven laboratory information management system (LIMS) designed for Web-based case reporting in medical mycology. The proposed framework is empowered with ontologies and formalized using category theory to provide a deep and common understanding of the functional and nonfunctional requirement hierarchies and their interrelations, and to trace the effects of a change on the conceptual framework.

1. Introduction

The life sciences constitute a challenging domain in knowledge representation. Biological data are highly dynamic, and bioinformatics applications are large and there are complex interrelationships between their elements with various levels of interpretation for each concept. In an ideal situation, the requirements for a software system should be completely and unambiguously determined before design, coding, and testing take place. The complexity of bioinformatics applications and their constant evolution lead to frequent changes in their requirements: often new requirements are added and existing requirements are modified or deleted, causing parts of the software system to be redesigned, deleted, or added. Such changes lead to volatility in the requirements of bioinformatics applications.

In this paper, we deal with an important problem of requirements volatility in the context of an ontology-driven clinical laboratory information management system (LIMS) [1, 2]. A LIMS is a software application for managing information about laboratory samples, users, instruments, standards, and other laboratory functions and products. It forms an essential part of electronic laboratory reporting (ELR) and electronic communicable disease reporting (CDR). ELR is a key factor in public health surveillance, improving real-time decision making based on messages reporting cases of notifiable conditions from multiple laboratories [3]. Combining these reports with clinical experiments and case studies makes up a CDR system [4]. This framework, along with the active participation of physicians specializing in fungal infectious diseases, infection control professionals, and lab technicians, aimed at generating automated online reporting from clinical laboratories to improve the quality of lab administration, health surveillance, and disease notification. It provides security, portability, and accessibility over the Web, as well as efficiency and data integrity in clinical, pharmaceutical, industrial, and environmental laboratoryprocesses.

Research Problem. Requirements volatility is "a measure of how much program requirements change once coding begins" [5]. Bioinformatics applications with frequently changing requirements have a high degree of volatility,

while projects with relatively stable requirements have a low one [6]. Higher requirement volatility will result in higher development and maintenance costs, the risk of schedule slippage, and an overall decrease in the quality of the services provided. Therefore, requirement volatility is considered one of the major obstacles to using a LIMS. In this paper, we propose an innovative approach for the automatic tracing of volatile requirement changes based on their formal representation in an ontological framework using a solid mathematical foundation, namely, category theory [7].

Approach. Investigating the factors that drive requirement change is an important prerequisite for understanding the nature of requirement volatility. This increased understanding will minimize that volatility and improve the process of requirement change management. One of the most important volatility factors is the diversity of requirement definitions in the application domain, which may lead to confusing and frustrating communication problems between application users and software engineers [8]. Ontologies [9] are widely used as a vehicle for knowledge management sharing common vocabularies, describing the semantics of programming interfaces, providing a structure to organize knowledge, reducing the development effort for generic tools and systems, improving data and tool integration, reusing organizational knowledge, and capturing behavioral knowledge. Ontologies can describe software architectures and requirements, which are difficult to model with object-oriented languages [10]. Conceptualization of the requirements using an ontology formalized with category theory minimizes requirement volatility by providing a deep and common understanding of the requirements [11], which is essential in order for bioinformatics application developers to manage the changes successfully. This paper proposes a generic categorical model of LIMS requirements with an emphasis on nonfunctional requirements, their dependencies and interdependencies using category theory as an advanced mathematical formalism. The resulting categorical model represents the functional requirements (FRs) and nonfunctional requirements (NFRs) based on an investigation of their dependencies and interdependencies, which is considered critical to success in tracing requirement changes. Requirement traceability, defined as "the ability to describe and follow the life of a requirement in both [forward and backward directions] " [12], is an essential part in performing requirement maintenance and change management processes. Moreover, the extent to which change traceability is exploited is viewed as an indicator of system quality and process maturity, and is mandated by existing standards [13]. These changes have to be monitored for consistency with the existing categorical framework in the LIMS context. After capturing the LIMS requirements in an ontological framework—to provide a common shared understanding of the requirements—empowered with category theory, a novel agent-based framework for the representation, legitimation, and reproduction (RLR) of changes [14] is proposed for implementing volatile requirement identification, and integrated change management and consistency monitoring in a LIMS (Figure 1).

RLR framework assists and guides the software developer through the change management process in general, and in representing and tracing the changes, particularly through the use of category theory.

The rest of the paper is organized as follows. Our discussion will be illustrated through examples from the LIMS system case study introduced in Section 2. Our approach for recruiting category theory for formalizing the conceptual framework of the requirements is presented in Section 3. The RLR framework for managing changes is described in Section 4. In Sections 5 and 6, we demonstrate the applicability of our categorical method for representing and tracking requirement changes and formalizing the interaction of agents in the RLR framework through an application scenario. We describe the evaluation phase in the proposed multiagent framework and review related work in Sections 7 and 8, respectively. The paper concludes with the list of contributions and an outline of research directions in Section 9.

2. The MYCO-LIMS Requirements Overview

The mycology laboratory information management system (MYCO-LIMS) is software for managing information about laboratory samples, users, instruments, standards, and other laboratory functions and products, and provides security, portability, and accessibility over the Web, efficiency, and data integrity in clinical, pharmaceutical, and industrial laboratoryprocesses. The MYCO-LIMS is an ontology-driven object-oriented application for a typical fungal genomics lab performing sequencing and gene expression experiments in the domain of medical mycology. Based on Gruber's definition [9], an ontology is a "specification of conceptualization", and provides an underlying discipline for knowledge sharing by defining concepts, properties, and axioms. The term "conceptualization" includes conceptual frameworks for analyzing shared domain knowledge which are necessary for knowledge representation in the domain of interest. In our context, the conceptual framework for requirement management outlines possible courses of action and patterns for describing a system's specifications and requirements. In complex biomedical systems development, a bioinformatics requirement change typically causes a ripple effect and forces the categorical requirements model to be altered as well.

MYCO-LIMS is used in the FungalWeb [14] integrated system to respond to queries regarding the clinical, pharmaceutical, industrial, and environmental processes related to pathogenic fungal enzymes and their related products. It is estimated that laboratory data account for 60–80% of the data generated during the entire clinical trial process [15].

The FungalWeb semantic Web infrastructure [16] (Figure 2) consists of the FungalWeb ontology, skin disease ontology (SKDON), a text mining framework, and intelligent agents. In addition, several external applications such as the MYCO-LIMS, the MYCO-LIS, and mutation miner [17] have been designed for knowledge exchange.

Microarrays are produced in different proportions, depending on the specific requirements of the gene expression

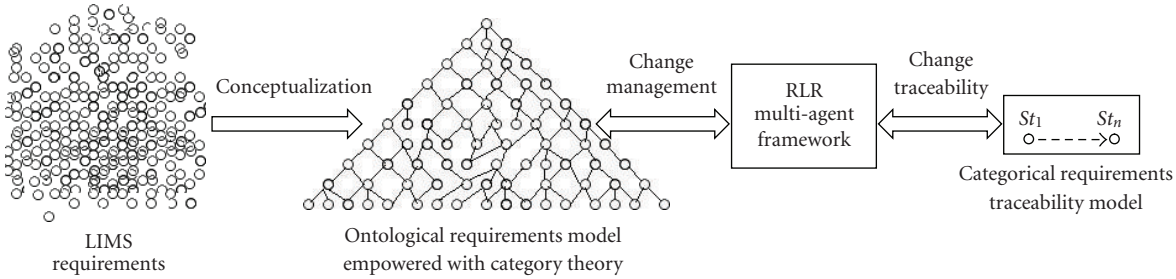


Figure 1: General view on the proposed approach for managing requirement volatility.

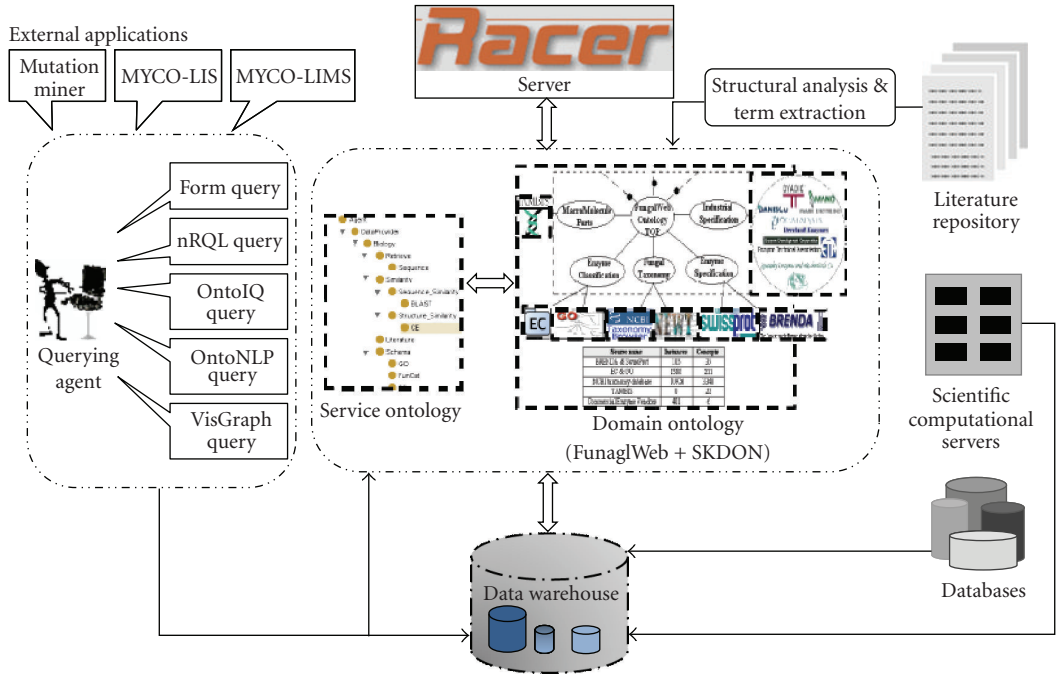


Figure 2: The FungalWeb infrastructure.

study being initiated. A typical microarray may include thousands of distinct cDNA probes [18]. Preparation of an array begins with the clone set deliverance in the form of plates or tissue samples (with associated data) from a vendor or other source [18]. The MYCO-LIMS will be able to maintain the taxonomy for each plate or sample in the system; such that a user can easily see the life cycle of the entity. The LIMS is based on MGED-specified [19] microarray data exchange standards, such as MIAME [20] or MAGE-ML [21].

Software systems in general and MYCO-LIMS in particular are characterized both by their functional behavior (what the system does) and by their nonfunctional behavior (how the system behaves with respect to some observable attributes like reliability, reusability, maintainability, etc.). Both aspects are relevant to software development and are captured correspondingly as functional requirements (FRs) and nonfunctional requirements (NFRs).

2.1. LIMS Functional Requirements (FRs). MYCO-LIMS is a Web-based system capable of providing services such as managing microarray gene expression data and laboratory supplies, managing patients, physicians, laboratories supplies or vendors' information, managing and tracking samples information, and managing orders. Figure 3 summarizes some of the main actors and services of the MYCO-LIMS application in a standard use case diagram.

MYCO-LIMS is capable of receiving multiple orders or cancelation requests at the same time. It requires its users to have a certain level of privileges to access any of the functionalities, except when searching for a product. The privileges are granted automatically upon successful authentication. In this paper, we limit the scope of the discussion to one functional requirement, "manage order", and further decompose it into two more specific sub-NFRs, "view orders," and "place order". In each decomposition, the offspring FRs contribute toward satisfying the goal of

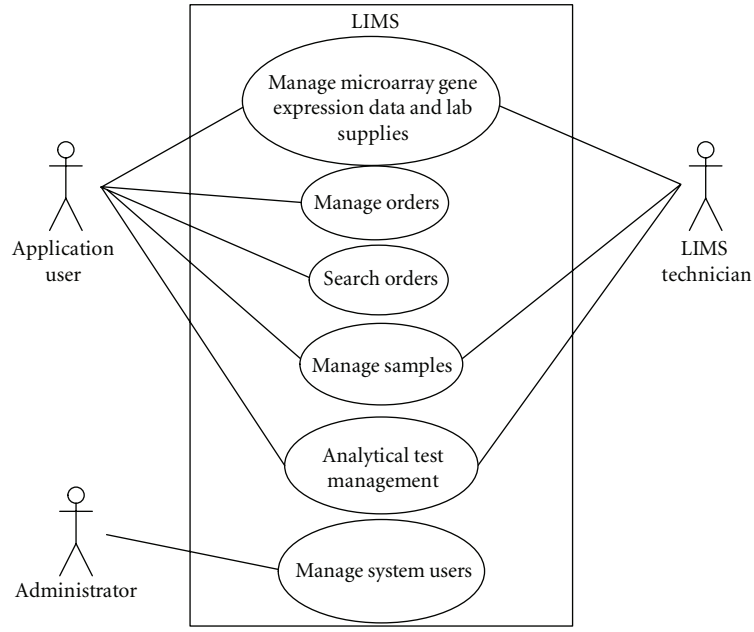


Figure 3: LIMS use case diagram.

the parent. Figure 4 presents the functional model and shows that an FR is realized through the various phases of development by many functional models (e.g., in the object-oriented field, a use case model is used in the requirements engineering phase, a design model is used in the software design phase, etc.). Each model is an aggregation of one or more artifacts (e.g., a use case and sequences of events representing scenarios for the use case model, classes and methods for the design model). For instance, view order use case is refined to a sequence of events < enter order number, visualize order > illustrating an instance of view order service; each event is refined as a method (viewOrderSession.view and viewCatalogue.view correspondingly) in the design phase. Modeling FRs and their refinements in a hierarchical way gives us the option of decoupling the task of tracing FRs change from a specific development practice or paradigm. Figure 4 visualizes the FR hierarchical model for the chosen case study through the hierarchy graph that forms a primary taxonomy for analyzing ontological relationships between requirements.

2.2. LIMS Nonfunctional Requirements (NFRs). The use case diagram shown in Figure 3 specifies the FRs of MYCO-LIMS services. At the same time, compliance with the NFRs, such as performance, scalability, accuracy, robustness, accessibility, resilience, and usability, is one of the most important issues in the software engineering field today. NFRs impose restrictions by specifying external constraints on the software design and implementation process [22] and therefore need to be considered as an integral part of the process of conceptual modeling of the requirements. The goal of this section is to build a systematic, quantitative, and formal approach to NFR modeling, impact detection, and volatility evaluation/decision-making from the early stages of the software development process.

We decompose a high-level NFR into more specific sub-NFRs. In each decomposition, the offspring NFRs can contribute partially or fully toward satisficing the parent. Let us consider the requirements of "managing orders with good security" and "maintain the users' transactions with good performance". The security requirement constitutes quite a broad topic. To effectively deal with it, the NFR may need to be broken down into smaller components, so that an effective solution can be found. We can decompose the security NFR into the sub-NFRs integrity, confidentiality, and availability. In the security example, each sub-NFR has to be satisfied for the security NFR to be satisfied. The sub-NFRs are refined (operationalized) into solutions that will satisfice the NFR. These solutions provide operations, processes, data representations, structuring, constraints, and agents in the target system to meet the goals stated in the NFRs. In the confidentiality example, a solution can consist of either implementing authorization or the use of additional ID. Figure 5 visualizes the NFR partial hierarchy resulting from the decomposition and operationalization relations for the NFRs chosen in the LIMS.

NFRs pose further challenges when it comes to determining their relationships with FRs. The tendency for NFRs to have a wide-ranging impact on a software system services and the strong interdependencies and tradeoffs that exist between them and the FRs leave typical existing software modeling methods incapable of integrating them into software engineering. In Section 2.3, we propose a new generic ontological framework for conceptualizing the NFR and FR requirements, their decompositions, and the corresponding associations.

2.3. Integrating FRs and NFRs into a Generic Ontological Framework. Hardly any requirement is manifested in

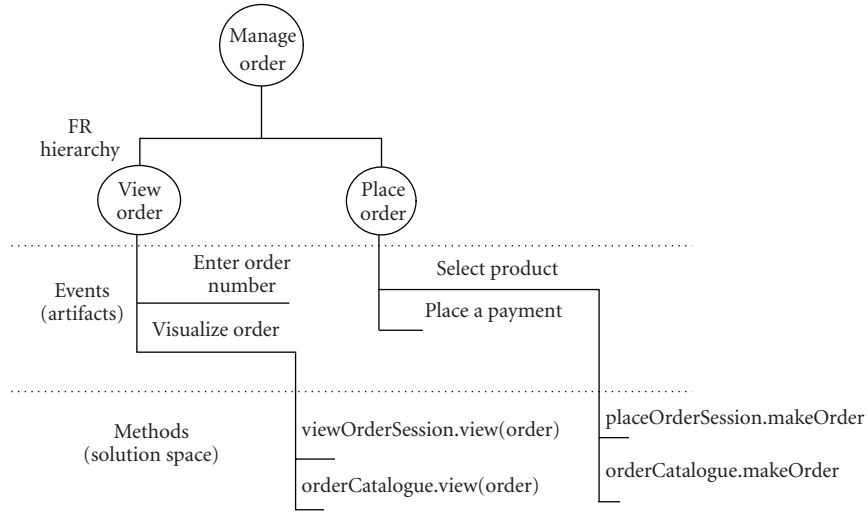


Figure 4: Illustration of the MYCO-LIMS FR traceability model.

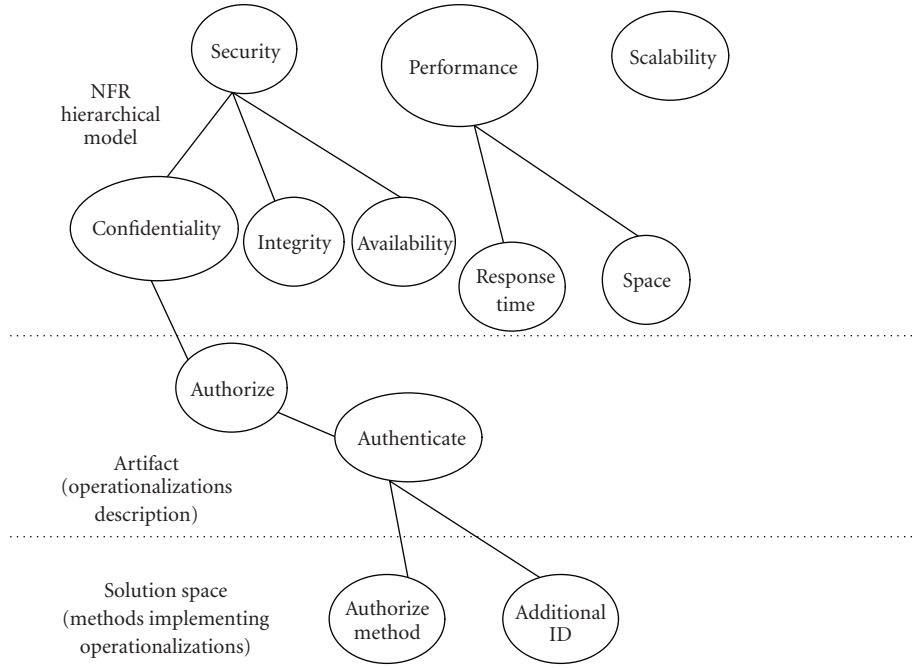


Figure 5: Illustration of the MYCO-LIMS NFR traceability model.

isolation, and normally the provision of one requirement may affect the level of provision of another. Understanding FR/NFR relations is essential to influencing the consistency and change management of the requirements. Once a software system has been deployed, it is typically straightforward to observe whether a certain FR has been met or not, as the ranges of success or failure in its context can be rigidly defined. However, the same is not true for NFRs as these can refer to quantitative statements that can be linked to other elements of the system. In fact, NFRs are not standalone goals as NFRs and their derived design solutions (operationalizations) can be associated to FRs throughout the software development process.

While tracing requirements is a major activity for change management of the system requirements, it has, by and large, been neglected for NFRs in practice. This area needs a special attention because NFRs are subjective in nature and have a broad impact on the system as a whole. In this section, we illustrate our approach toward finding an effective method for conceptualizing NFRs based on their hierarchy and their interrelations with FRs in the MYCO-LIMS invoicing system case study.

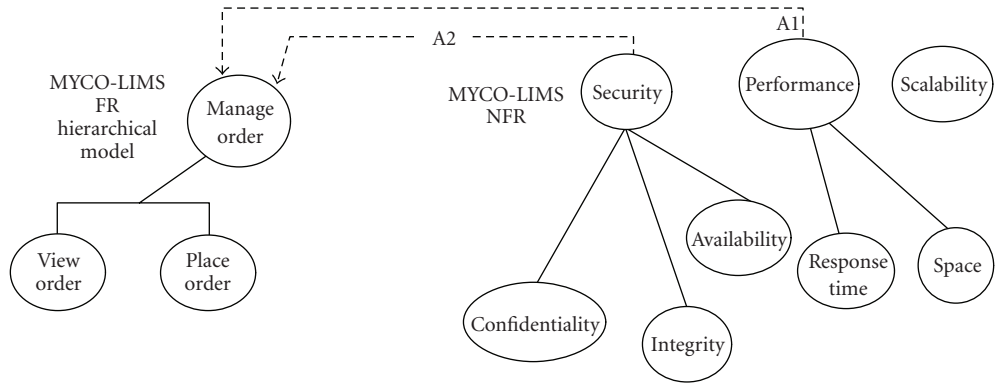


Figure 6: Illustration of MYCO-LIMS NFRs/FRs dependencies hierarchical model.

For example, associating response time NFR to view order use case would indicate that the software must execute the functionality within an acceptable duration (see association A1, Figure 6). Another example is associating security NFR to the "manage order" FR, which would indicate that the interaction between user and the software system in the "manage order" service must be secured (see association A2, Figure 6), which also precisely implies that user interface for other interactions is not required to be secured.

If an association exists between a parent NFR and a functionality (e.g., association *A2* between *security* and *manage_order*, or association *A1* between *performance* and *manage_order*) (see Figure 6), there will be an association between operationalizations derived from NFRs and methods derived from the functionality (e.g., *authorize* derived from *security*, and *placeOrderSession.makeOrder* derived from *manage_order*) (see Figure 7).

Figure 7 illustrates the refinement of the interactions. The complete change management model would require the refinement of performance and scalability into operationalizations and methods, and the identification of the associated interaction points to which they are mapped.

A change in FRs or NFRs can be authorized if and only if that change is consistent with the existing requirements model. Our future work includes the development of more consistency rules based on a formal presentation of the FR and NFR hierarchies and their relations, and these rules will be checked automatically before a change is authorized.

The conceptualization of FR and NFR hierarchies and their interconnections form the bases for analyzing ontological relationships between requirements in the service ontology (see Figure 2). The NFR/FR ontological framework introduced in this section can be visualized through a categorical hierarchical graph, which makes it possible to keep track of the required behavior of the system using dynamic views of software behaviors from requirements elicitation to implementation.

The following subsection proposes a generic categorical model of requirements with an emphasis on NFRs and their interdependencies and refinements through using category theory as an advanced mathematical formalism, and this model will be independent of any programming paradigm.

3. Generic Categorical Representation of Requirements and Their Traceability

An ontology is a categorization of things in the real world. It can be viewed in terms of an interconnected hierarchy of theories as a subcategory of a category of theories expressed in a formal logic [23]. Categorical notations consist of diagrams with arrows. A category consists of a collection of objects and a collection of arrows (called morphisms). Each arrow $f: X \rightarrow Y$ represents a function. Representation of a category can be formalized using the notion of the diagram. We have chosen category theory as the main formalism in our framework because it has proved itself to be an efficient vehicle to examine the process of structural change in living and evolving systems [24].

In fact, we can use category theory to represent ontologies as a modular hierarchy of domain knowledge. Categories capture and compose the interactions between objects, identify the patterns of interacting objects in ontologies, and either extract invariants in their action or decompose a complex object in basic components. Categories are also able to identify patterns that recur again and again in a changing system. Other reasons for using category theory in our framework, as stated by Adamek et al. [25], aretheabundance, precise language, and convenience of symbolism for visualization. Although category theory is a relatively new domain of mathematics, introduced and formulated in 1945 [7], categories are frequently found in this field (sets, vector spaces, groups, and topological spaces all naturally give rise to categories). The use of categories can enable the recognition of certain regularities in distinguishing a variety of objects, their interactions can be captured and composed, equivalent interactions can be differentiated, patterns of interacting objects can be identified and some invariants in their action are extracted, and a complex object can be decomposed into its basic components [26].

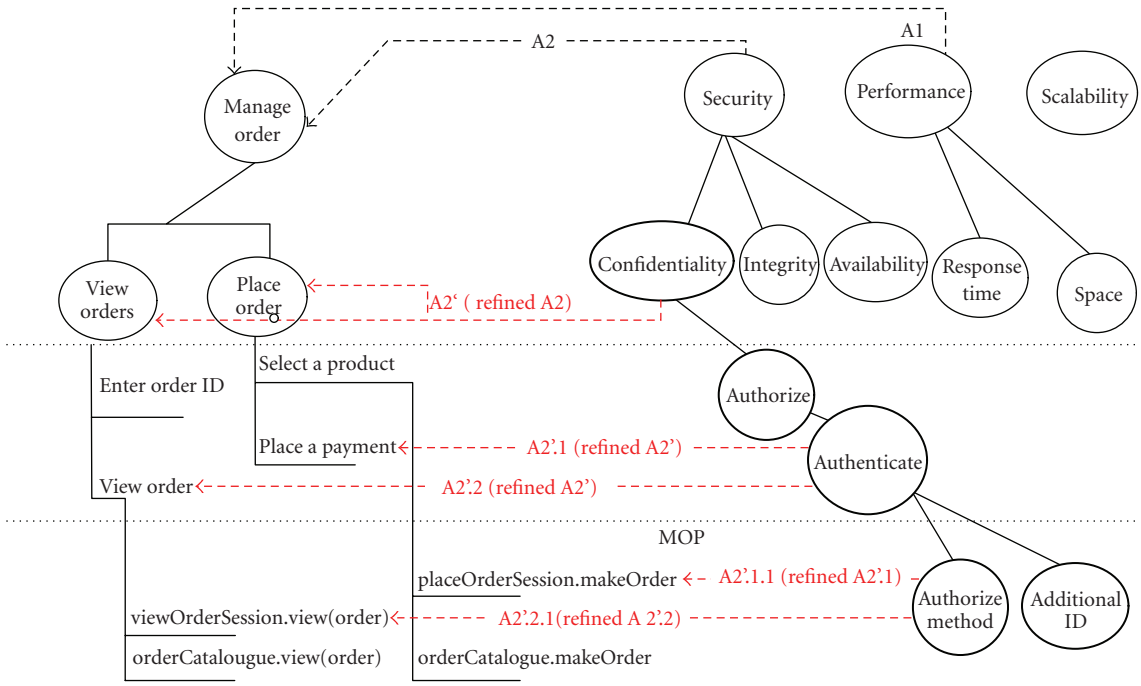


Figure 7: MYCO-LIMS requirements associations' refinement.

In order to explicitly reason about the impact of NFRs and their refinements on the project throughout the software development process, we explicitly represent NFRs, FRs, and their dependencies and refinements using the language of category theory. Figure 8 captures the generic view on the requirements modeling process where requirements group, hierarchical model, artifacts, and solution space are categories representing the project requirements, the analysis models, the refined representations of the project requirements, and the requirements implementation, respectively. The arrows are morphisms which capture the refinement processes, namely, decomposition, operationalization, and implementation defined as shown in Figure 8.

Figure 8 shows that a requirement is realized through the various phases of refinement by hierarchical models, where each model is an aggregation of one or more artifacts. The implementation arrow refines the artifacts into solutions in the target system that will satisfy the requirements. These solutions provide operations, processes, data representations, structuring, constraints, and agents in the target system to meet the requirements represented in the requirements group. High-level FRs are refined in the requirements analysis phase into more specific sub-FRs (use cases and their relations (FR hierarchy model), e.g.), which are then operationalized as use case scenarios describing instances of interactions between the actors and the software, and modeled as events (artifacts), which are implemented as methods (solution space). High-level NFRs are refined into an NFR hierarchy where the offspring NFRs can contribute fully or partially toward satisficing the parent. The sub-NFRs are operationalized into solutions (artifacts) in the target systems, which will satisfice the NFR. These procedures provide operations, processes, data representations, structuring, constraints, and agents in the target system to meet the needs stated in the NFRs, and are implemented as methods in the solution space.

The requirement refinements are then expressed formally in terms of the composition operator $\circ$, assigning to each pair of arrows f and g, with $\operatorname{cod} f = \operatorname{dom} g$, a composite arrow $g \circ f$: $\operatorname{dom} f \rightarrow \operatorname{cod} g$ ($\operatorname{cod} f$ is a notation for a codomain, and $\operatorname{dom} f$ is the notation used to indicate the domain of a function f). In this case, each requirement object belonging to the requirements group category will be refined to its implementation belonging to the solution space. The resulting solution forces preservation of the requirements and their relations, which are modeled with the *trace* arrows. The consistency between the solution and the original requirements can be guaranteed by the composition of categorical arrows representing morphisms. As a result, each change to a requirement or its refinement belonging to the domain of f will be traced to its refinement belonging to the codomain of g by means of the composition of the corresponding trace arrows.

3.1. Categorical Representation of FRs, NFRs Hierarchies, and Their Interdependencies. The category *FR, NFR hierarchies, and relations* (Figure 9) consist of objects representing FRs and NFRs, their decomposition into sub-FR and sub-NFR (which are also FR and NFR correspondingly), and their impact associations; above concepts are treated jointly and in an integrated fashion. We identify four critical areas for impact detection in which NFRs require change

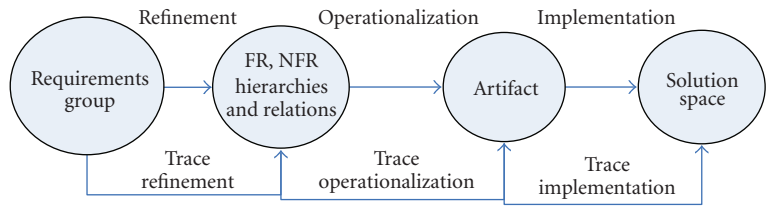


Figure 8: Generic categorical framework for requirement traceability.

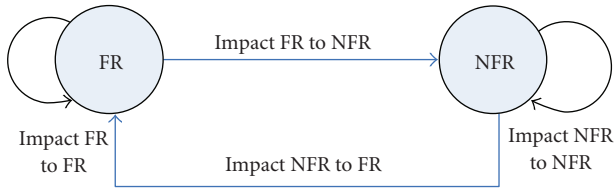


Figure 9: *FR, NFR hierarchies, and relations* category.

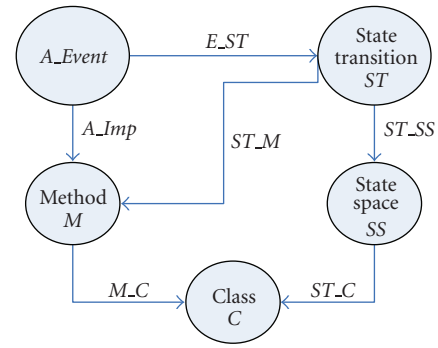


Figure 10: Tracing the changes to the state spaces, classes, and methods.

management support: (i) impact of changes to FRs on NFRs (intermodel integration), (ii) impact of changes to NFRs on FRs (intermodel integration), (iii) impact of changes to NFRs on sub-NFRs and parent NFRs (intramodel integration), and (iv) impact of changes to NFRs on other interacting NFRs (intramodel integration).

3.2. Categorical Representation of the Solution Space. The solution space category contains state space *SS* (all potential states including initial states), state transition *ST* (next state function), class *C* categorical objects, and methods arrows. The *trace implementation* morphism traces the effect of the changes to artifact objects on the solution space objects. In Figure 10, for instance, we illustrate the refinement of an event from the artifact category to a state transition object *ST*. Moreover, each state transition *ST* is defined on the state space *SS* (arrow *ST_SS*) linked by a function *ST_C: ST* → C to a class *C*. The state transitions are implemented by methods captured with the function *ST_M: ST* → *AP_M*, and belonging to a class *C* (see function *M_C*). The above functions support the tracing mechanism and are captured formally in Figure 10. The changes are then represented formally in terms of the composition operator ∘; for instance, *E_ST* ∘ *ST_SS* ∘ *ST_C* will trace a change in dom *E_ST* (which is *A_Event*) to the codomain of *ST_C* (which is class *C*).

As we presented in [27], category theory has great potential as a mathematical vehicle to represent, track, and analyze changes in ontologies. For example, it can be used in the taxonomical representation of requirements to help in the study of the ontological relationship between the various nodes within the hierarchy. After describing the ontological concepts within the categories representing a modular hierarchy of domain knowledge, we have employed category theory to analyze ontological changes and agent interaction in different stages of the RLR framework [14].

4. The RLR Framework

The RLR multiagent framework [14] (RLR stands for: representation, legitimation, reproduction) (Figure 11) aimed at capturing, tracking, representing, and managing the changes in a formal and consistent way, enabling the system to generate reproducible results using change capture agents, reasoning agents, learning agents, and negation agents. Change capture agents are responsible for discovering, capturing, and tracking changes in ontology, by processing the change logs. The change logs accumulate important data about various types of changes. In RLR, a learner agent uses these historical records of changes that occur over and over in a change process to derive a pattern to estimate the rate and direction of future changes for a system by generating rules or models. The reasoner (which verifies the results of a change) and negotiation agents can change the rules generated and send modifications to the learning agent. Negotiation takes place when agents with conflicting interests want to cooperate. In RLR, the negotiation agent acts as a mediator allowing the ontology engineer and other autonomous agents to negotiate the best possible realization of a specific change, while maximizing the benefits and minimizing the loss caused by such a change. A human expert may then browse the results, propose actions, and decide whether to confirm, delete, or modify the proposals, in accordance with the intention of the application. In RLR, negotiation is defined based on the conceptual model of argumentation [28], where an argument is described as a piece of information allowing an agent to support and justify its negotiation stance or affect another agent's position

through a communication language and a formal protocol [28]. The negotiation protocol can formally provide the necessary rules [29] (i.e., rules for admissions, withdrawals, terminations) for negotiation dialog among participants. In our approach, we have partially adapted the architecture of the argumentative negotiating agent described at [30].

Within the RLR argumentative architecture, the negotiation agent and the reasoning agent provide arguments for the acceptance or rejection of a change proposal. The "argument generator" (Figure 11) determines appropriate responses based on the negotiation rules. Different arguments attack one another to impose their rules and defeat their peers by sending counter arguments. The inferred arguments can increase the possibility of higher-quality agreement [30, 31]. The negotiation protocols in the RLR architecture contain the negotiation protocol's rules, which dictate a protocol. As an application is used and evolves over time, the change logs accumulate invaluable data and information about various types of changes. A learner agent can use these historical records of changes that occur over and over in a change process to derive a pattern out of the rules generated. The reasoner and the negotiation agents can change the rules—if necessary—and send modifications to the learning agent. The learning agent starts with limited, uncertain knowledge of the domain and tries to improve itself, relying on adaptive learning based on the semantics provided by the ontological backbone.

5. Employing Category Theory in the RLR Framework

We have used categories in various stages of the RLR multiagent framework for representing and tracking changes in NFRs and FRs.

5.1. Category Theory for Representing and Tracking Changes. The categorical representation enables the progressive analysis of ontologies and can be used to represent the evolutionary structure of an ontology, to provide facilities for tracking each change and to analyze the impact of these changes by the following.

(a) *Comparing different states of a class*. We have used "functor", which is a morphism in the category of all small categories (where classes are defined as categories) to describe the set of state space (set of all possible states for a given state variable set) for a class as a cross product of attribute domains and the operations of a class as transitions between states for ontological elements indexed by time. Using the functor, the transition from O_t to $O_{t'}$, where the time changes from t to t', can be represented and analyzed. For more information see [27].

(b) *Measuring coupling*. Coupling indicates the complexity of evolving structure [27]. When coupling is high, it indicates existence of a large number of dependencies in an ontological structure which must be checked to analyze and control the chain of changes. Following [32], to analyze the coupling we consider three types of arrows, namely, precondition, postcondition, and message-send arrows in category theory to analyze various conditional changes [27].

(c) *Using Pushout and Pullback*. When a change is either integration or mergence, one can use two categorical constructors: pushout and pullback [33]. The pushoutfor two morphisms $f : A \rightarrow B$ and $g : A \rightarrow C$ is an object D, and two morphisms $i_1 : B \rightarrow D$ and $i_2 : C \rightarrow D$, such that the square commutes (Figure 12(a)). D is initial object in the full subcategory of all such candidates D' (i.e., for all objects D' with morphisms j_1' and j_2', there is a unique morphism from D to D').

The *pullback* (also known as "cartesian square") for two morphisms $f : A \rightarrow C$ and $g : B \rightarrow C$ is an object D, and two morphisms $i_1 : D \rightarrow A$ and $i_2 : D \rightarrow B$, such that the square commutes. Here D is the terminal object in the full subcategory of all such candidates D' [34] (Figure 12(b)). Hitzler et al. [35] and Zimmermann et al. [36] also used pushout for ontology alignment.

5.2. Category Theory for Representing Agent Interactions and Conflict Resolution. Intelligent agents perform actions in a context by using rules. Changing the rules is a main adaptation principle [37] for learning in RLR framework. The adaptive agents in the RLR have been defined following Resconi's method [37]. The rules consist of a set of semantic unity symbolized by S_1, IN, P_1, and OUT, representing the input statement, the domain of the rule, the rule, and the range of the rule (denoting the value of an agent's action), respectively. When we are working in a dynamic environment, it is likely that these rules change into other rules. Therefore, a single change in the primary structure triggers other changes in rules and contexts. A communication channel [37] between those rules and between different adaptive agents is needed to manage all the necessary interactions.

In the RLR we have used category theory formalism, along with general systems logical theory (GSLT) [38], to formalize agents' communications. For instance, the communication between different semantic unities [14] can be represented as in Figure 13.

In addition, category theory can be used for modeling agent interactions [39], yielding a practical image of adaptive learning agents, their semantic unities, and adaptation channels [37].

We have also followed the approach presented in [40] for representing the product and coproduct of objects, to categorically represent the integration and merging of NFR objects, which are defined as ontological elements. The negotiation agent in RLR can negotiate to determine the best of several methods of integration. For example, an integration can be implemented as the product $A \times B$ (all possible pairs <elements from A, elements from B>), or the coproduct of the objects $A + B$ (all elements from A and all elements from B) for both categorical objects and arrows

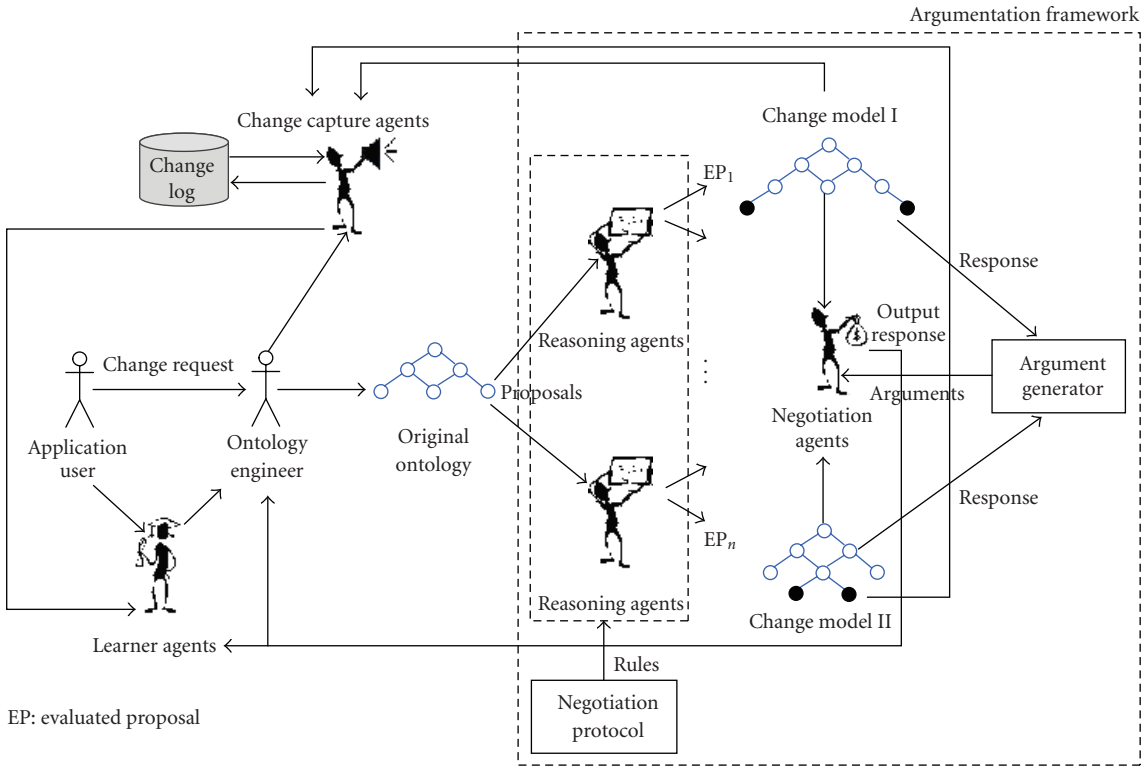


Figure 11: The RLR framework for change management and conflict resolution.

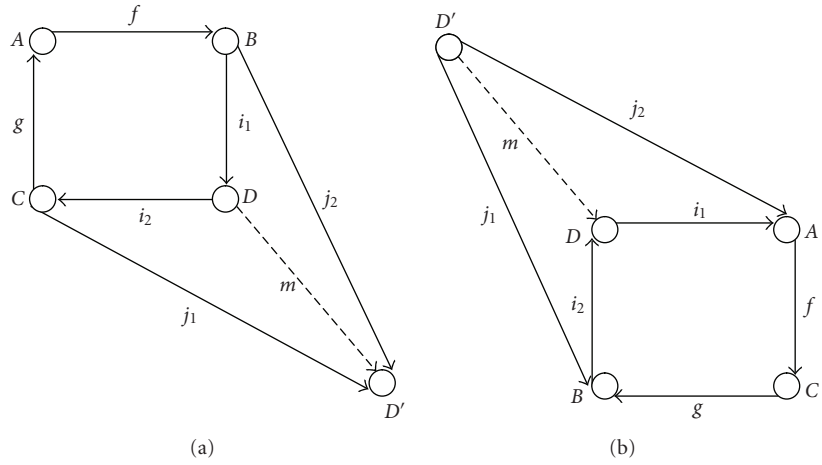


Figure 12: Two categorical constructors: (a) pushout, (b) pullback.

(denoting ontological elements). Assume that we define the following arguments for integrating ontological structures within a dialectical database [31] in the RLR framework:

$$a_1\colon A \times B, \qquad a_2\colon A + B, \qquad a_3\colon A, \qquad a_4\colon B. \quad (1)$$

Categorically speaking "a_1 defeats a_2" can be represented by an arrow from a_1 (domain) to a_2 (codomain) (Figure 14). By following categorical representation, an argumentation network will be generated, which can be used to formally describe negotiations and speed up inferences [31].

6. Application Scenarios

As shown in Section 5, category theory can be used in RLR to integrate time factor and represent and track changes in ontological structure in time through using the notion of state capturing an instance of system's FRs, NFRs, and associations at certain period of time. For example, a change in the authorize method would affect the method "placeOrderSession.makeOrder" in state St_1 of the system, which will be traced to changes in state St_2 (Figure 15). Explicitly capturing of the evolution of the requirements

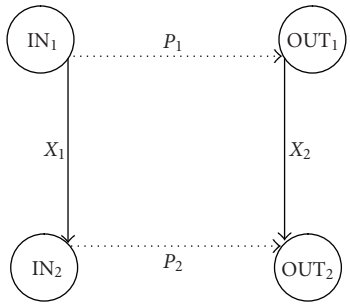


Figure 13: The categorical representation that shows how rules P_1 and P_2 enable the transformation of rule X_1 into rule X_2 (following [37]).

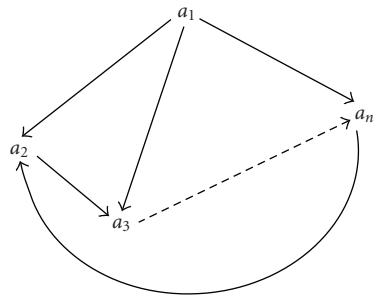


Figure 14: Categorical representation of the argumentation network.

in time can aid MYCO-LIMS developers and maintainers to deal with requirements change management in highly dynamic clinical applications.

Generally speaking, changes to each NFR would lead to changes in the conceptual framework. As mentioned in Section 3, we are monitoring the effect of FR or NFR changes through their refinement relations, that is (1) identifying the "slice" of the conceptual framework that will be affected by the change, (2) applying the consistency rules to make sure the change does not introduce any inconsistencies in the "slice," and (3) implementing the change, if authorized.

The RLR change management framework is modeled as an intelligent control loop, which has one state for each of the above stages (1), (2), and (3), the events modeling the change of state. Considering the requirements to be organized in a lattice-like ontological framework, in order to represent the various states of our conceptualization, we use a categorical discrete state model, which describes the states and events in the ontological structure using a diagrammatical notion. The discrete state model is specified by a state space (all potential states), a set of initial states, and a next-state function. Based on our application, we designed our class diagrams following the method described in [27, 32] (Figure 16), which can be used to create patterns for learning agents. The Op_i arrows in this figure represent the operations for the class, wherein the operation or event Op_1 causes an object in state St_1 to undergo a transition to state St_2. The operation Op_1 has no effect on the object if it is in any other state, since no arrow labeled Op_1 originates from any other state. The object $\varnothing$ in the diagram is the null state. The *create* arrow represents the creation of the object by assigning an identifier and setting its state to the initial defined state, and the *destroy* arrow represents its destruction [32].

Based on [32], a projection arrow for any attribute is drawn from the state space to the attribute domain and labeled with the name of the attribute (i.e., π_i represents the value of the ith attribute). A selection arrow for each state x (labeled as σ_x) is drawn from the state space to that state (i.e., σ_i gives the ith state).

Using category theory we represent the most common operations during requirement change management such as adding/deleting a class of requirements, combining two classes of requirements into one, adding a generalization/association relationship, adding/deleting a property or relationship. For more information see [27].

7. Evaluation of the Approach along with Change Verification

The legitimation phase in RLR verifies the legitimacy and consistency of a change in the domain of interest. This phase assesses the impact of a potential change before the change is actually made. Experts and logical reasoners study a change based on its consistency with the whole design, in varying degrees of granularity. Then, final approval is needed from the end users. Logical legitimation is obtained by a reasoning agent, which is a software agent that controls and verifies the logical validity of a system, revealing inconsistencies, misclassifications, hidden dependencies, and redundancies. It automatically notifies users or other agents when new information about the system becomes available. We use RACER [41] as a description logic reasoner agent, along with other semiformal reasoners in RLR. When the agent is faced with a change, it ought to revise its conceptualization based on the new input by reasoning about the consistency of the change using both prior and new knowledge. We also use a semiautomated reasoning system for basic category theory reasoning [42] based on a first-order sequent calculus [43], which captures the basic categorical constructors, functors, and natural transformations, and provides services to check consistency, semantic coherency, and inferencing [43]. Placing a new class of requirements in a system may sometimes lead to redundancy in the requirement taxonomy. One of the major issues in requirement analysis is finding and identifying logically equivalent classes and relationships which may differ in name but perform the same function. Employing category theory enables us to deal with this problem of logical equality in the evolving requirement hierarchy using isomorphic reasoning [44].

8. Related Work

Several efforts have been reported [45–48] during the last decade in the pursuit of inclusive frameworks for managing dynamic taxonomies, ontologies, and control vocabularies. Since existing knowledge representation languages, including well-established description logic, cannot guarantee the

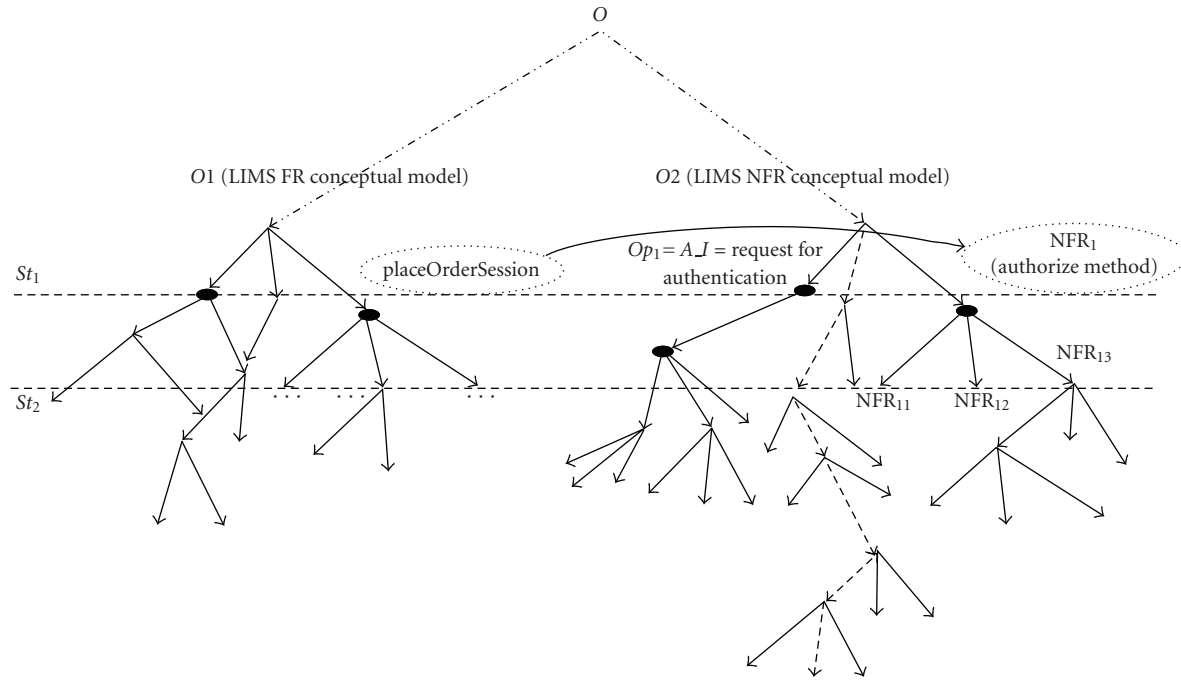


Figure 15: Categorical representation of evolving MYCO-LIMS functional requirements (FRs) and nonfunctional requirements (NFRs).

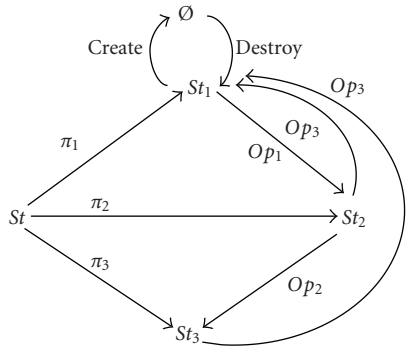


Figure 16: Class diagram for the part of the FR-NFR ontological structure that represents the transition between states.

computability of highly expressive time-dependent models, the current efforts have been entirely focused on time-independent ontological models. However, the real ontological structures exist in time and space. From another perspective, those who choose other knowledge representation formalisms, such as state machine [49], can cope with time-based models, but these formalisms fail to address ontological concepts and rules because they are much too abstract and have no internal structure or clear semantics. In our proposed framework, category theory, with its rich set of constructors, can be considered as a complementary knowledge representation language for capturing and representing the full semantics of evolving abstract requirements conceptualized within ontological structures. Rosen [50] was among the first to propose the use of category theory in biology, in the context of a "relational biology".

Category theory also has been used by MacFarlane [24] as an efficient vehicle to examine the process of structural change in living/evolving systems. Whitmire [32], Wiels and Easterbrook [51], and Mens [52] have examined category theory for change management in software engineering domain. Hitzler et al. [35] and Zimmermann et al. [36] also have proposed using this formalism in knowledge representation area.

9. Discussion, Challenges, and Future Work

Any attempt to successfully systematize and automate electronic communication in biomedicine—with its continuously changing nomenclature and requirements—needs to pay special attention to managing requirement volatility in various stages of the biomedical application life cycle. Due to the wide variety of requirements controlled by the LIMS across diverse industries, LIMS software needs to be inherently more flexible [15]. One of the issues in requirement evolution and change management is a lack of formal change models with clear and comprehensible semantics. In order to represent, track, and manage requirement changes throughout a LIMS software project, we have proposed an agent-based framework to handle evolving requirements, which are categorized in an ontological structure. An ontology provides a means for formally capturing the FR and NFR hierarchies and their interrelations, and for exhaustive tracing of the effects of a change on the conceptual framework. In addition, we have proposed using category theory—which is an intuitive and powerful formalism, independent of any choice of ontology language—to capture the full semantics of evolving hierarchies in various phases

of RLR. It also provides a language to precisely describe many similar phenomena that occur in different mathematical fields with an appropriate degree of generality. For example, category theory makes it possible to make a precise distinction between categories via the notion of natural isomorphism. It also provides a unified language to describe topological spaces via the notion of concrete isomorphism [25]. In addition, categorists have developed a symbolism for visualizing complicated facts by means of diagrams. Our proposed method for employing category theory to manage the evolving FR and NFR hierarchical structure can significantly help formalize agile requirement modeling in highly dynamic clinical applications. Moreover, this method can be easily adapted to different project situations and needs. The ontology-grounded categorical framework introduced here can be used to reduce requirement volatility by facilitating the definition of consistency rules for requirement change and supporting the automatic evaluation of consistency rule compliance with software requirements. The knowledge captured about requirement volatility and formalized using category theory is a suitable means to trace the effect of any requirement change on the specifications of the whole system.

In the process of employing category theory as the core formalism for our proposed framework, we had to deal with several challenges. Some of the major ones included the reasoning issues and managing conceptualization changes. Although we are able to provide some sort of basic reasoning and inferencing for categories, we still need to improve the reasoning capability to cover more advanced reasoning services. Also, the representation of changes in conceptualization due to the nature of NFRs, which needs to deal with abstract concepts and notions, is challenging. In order to overcome this issue, we are working on grammatical change algorithms in linguistics and language evolution. For future work, we plan to concentrate on the evolution of requirement calculation rules, which are based on the available requirement traceability information. Finally, a third field of study will address dashboard visualization and customization for various FungalWeb requirement management tools.

References

[1] R. D. Mcdowall, "A matrix for the development of a strategic laboratory information management system," *Analytical Chemistry*, vol. 65, no. 20, pp. 896A–901A, 1993.

[2] G. Avery, C. McGee, and S. Falk, "Implementing LIMS: a "how-to" guide," *Analytical Chemistry*, vol. 72, no. 1, pp. 57A–62A, 2000.

[3] J. M. Overhage, J. Suico, and C. J. McDonald, "Electronic laboratory reporting: barriers, solutions and findings," *Journal of Public Health Management and Practice*, vol. 7, no. 6, pp. 60–66, 2001.

[4] R. Wurtz and B. J. Cameron, "Electronic laboratory reporting for the infectious diseases physician and clinical microbiologist," *Clinical Infectious Diseases*, vol. 40, no. 11, pp. 1638–1643, 2005.

[5] M. Eisenbarth and J. Dörr, "Facilitating project management by capturing requirements quality and volatility information," in *Proceedings of Workshop on Measuring Requirements for Project and Product Success (MeReP '07)*, Palma de Mallorca, Spain, November 2007.

[6] Y. K. Malaiya and J. Denton, "Requirements volatility and defect density," in *Proceedings of the 10th International Symposium on Software Reliability Engineering (ISSRE '99)*, pp. 285–294, Boca Raton, Fla, USA, November 1999.

[7] S. Eilenberg and S. Mac Lane, "General theory of natural equivalences," *Transactions of the American Mathematical Society*, vol. 58, no. 2, pp. 231–294, 1945.

[8] K. E. Wiegers, *Software Requirements*, Microsoft Press, Wasthington, DC, USA, 2nd edition, 2003.

[9] T. R. Gruber, "A translation approach to portable ontology specifications," *Knowledge Acquisition*, vol. 5, no. 2, pp. 199–220, 1993.

[10] D. Deridder and T. A. D'Hondt, "Concept-centric approach to software evolution," in *Proceedings of the 19th Annual ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA '04)*, Vancouver, Canada, October 2004.

[11] G. Dobson and P. Sawyer, "Revisiting ontology-based requirements engineering in the age of the semantic web," in *Proceedings of the International Seminar on Dependable Requirements Engineering of Computerised Systems at NPPs*, Halden, Norway, November 2006.

[12] O. C. Z. Gotel and A. C. W. Finklestein, "An analysis of the requirements traceability problem," in *Proceedings of the 1st International Conference on Requirements Engineering (ICRE '94)*, pp. 94–101, Colorado Springs, Colo, USA, April 1994.

[13] N. Aizenbud-Reshef, B. T. Nolan, J. Rubin, and Y. Shaham-Gafni, "Model traceability," *IBM Systems Journal*, vol. 45, no. 3, pp. 515–526, 2006.

[14] A. Shaban-Nejad and V. Haarslev, "Incremental biomedical ontology change management through learning agents," in *Proceedings of the 2nd KES International Symposium on Agent and Multi-Agent Systems (KES-AMSTA '08)*, vol. 4953, pp. 526–535, Incheon, Korea, March 2008.

[15] E. D. Krasovec, Limber Labs, 2007, http://www.starlims.com/ICT_Autumn_2007.pdf.

[16] C. J. O. Baker, A. Shaban-Nejad, X. Su, V. Haarslev, and G. Butler, "Semantic web infrastructure for fungal enzyme biotechnologists," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 4, no. 3, pp. 168–180, 2006.

[17] C. J. O. Baker and R. Witte, "Mutation mining—a prospector's tale," *Information Systems Frontiers*, vol. 8, no. 1, pp. 47–57, 2006.

[18] D. Jackson and A. Fergus, "Gene expression informatics," *PharmaGenomics*, pp. 24–26, November-December 2002.

[19] Microarray and Gene Expression Data-MGED, http://www.mged.org/index.html.

[20] Minimum Information about a Microarray Experiment, http://www.mged.org/Workgroups/MIAME/miame.html.

[21] MAGE-ML: MicroArray Gene Expression Markup Language, http://xml.coverpages.org/MAGEdescription2.pdf.

[22] G. Kotonya and I. Sommerville, *Requirements Engineering: Processes and Techniques*, John Wiley & Sons, New York, NY, USA, 1998.

[23] M. J. Healy and T. P. Caudell, "Ontologies and worlds in category theory: implications for neural systems," *Axiomathes*, vol. 16, no. 1-2, pp. 165–214, 2006.

[24] A. I. MacFarlane, "Dynamic structure theory: a structural approach to social and biological systems," *Bulletin of Mathematical Biology*, vol. 43, no. 5, pp. 579–591, 1981.

[25] J. Adamek, H. Herrlich, and G. E. Strecker, *Abstract and Concrete Categories: The Joy of Cats*, John Wiley & Sons, New York, NY, USA, 1990.

[26] A. C. Ehresmann and J.-P. Vanbremeersch, "The memory evolutive systems as a model of Rosen's organisms-(metabolic, replication) systems," *Axiomathes*, vol. 16, no. 1-2, pp. 137–154, 2006.

[27] A. Shaban-Nejad and V. Haarslev, "Categorical representation of evolving structure of an ontology for clinical fungus," in *Proceedings of the 11th Conference on Artificial Intelligence in Medicine (AIME '07)*, pp. 277–286, Amsterdam, The Netherlands, July 2007.

[28] I. Rahwan, S. D. Ramchurn, N. R. Jennings, P. McBurney, S. Parsons, and L. Sonenberg, "Argumentation-based negotiation," *Knowledge Engineering Review*, vol. 18, no. 4, pp. 343–375, 2003.

[29] N. R. Jennings, S. Parsons, P. Noriega, and C. Sierra, "On argumentation-based negotiation," in *Proceedings of the International Workshop on Multi-Agent Systems (IWMA'S '98)*, Boston, Mass, USA, July 1998.

[30] R. Ashri, I. Rahwan, and M. Luck, "Architectures for negotiating agents," in *Proceedings of the 3rd International Central and Eastern European Conference on Multi-Agent Systems and Applications(CEEMAS '03)*, pp. 136–146, Prague, Czech Republic, June 2003.

[31] M. Capobianco, C. I. Chesñevar, and G. R. Simari, "Argumentation and the dynamics of warranted beliefs in changing environments," *Autonomous Agents and Multi-Agent Systems*, vol. 11, no. 2, pp. 127–151, 2005.

[32] S. A. Whitmire, *Object Oriented Design Measurement*, John Wiley & Sons, New York, NY, USA, 1997.

[33] J. L. Fiadeiro, *Categories for Software Engineering*, Springer, Berlin, Germany, 1st edition, 2004.

[34] S. Easterbrook, "Category theory for beginners," in *Proceedings of the 3rd IEEE International High-Assurance Systems Engineering Symposium (HASE '98)*, Honolulu, Hawaii, USA, October 1998.

[35] P. Hitzler, M. Krötzsch, M. Ehrig, and Y. Sure, "What is ontology merging?—a category-theoretical perspective using pushouts," in *Proceedings of the 1st International Workshop on Contexts and Ontologies (C&O '05)*, pp. 104–107, Pittsburgh, Pa, USA, July 2005.

[36] A. Zimmermann, M. Krötzsch, J. Euzenat, and P. Hitzler, "Formalizing ontology alignment and its operations with category theory," in *Proceedings of International Conference on Formal Ontology in Information Systems (FOIS '06)*, pp. 277–288, Baltimore, Maryland, USA, November 2006.

[37] G. Resconi and L. C. Jain, *Intelligent Agents: Theory and Applications*, vol. 155 of *Studies in Fuzziness and Soft Computing*, Springer, Berlin, Germany, 2004.

[38] G. Resconi and G. Hill, "The language of general systems logical theory: a categorical view," in *Proceedings of the 3rd European Congress on Systems Science*, pp. 1091–1096, Rome, Italy, October 1996.

[39] M. W. Johnson, T. Miller, and P. McBurney, "Modelling multi-agent interaction protocols using categories enriched over pointed sets," Tech. Rep. ULCS-07-009, University of Liverpool, Liverpool, UK, 2007.

[40] E. Yanenko, "Evolving categories: consistent framework for representation of data and algorithms," http://arxiv.org/abs/cs/0412089.

[41] V. Haarslev and R. Möller, "Description logics for the semantic web: racer as a basis for building agent systems," *Künstliche Intelligenz*, vol. 17, no. 3, pp. 10–15, 2003.

[42] A. Poigné, "Elements of categorical reasoning: products and coproducts and some other (Co-) limits," in *Proceedings of the International Workshop on Category Theory and Computer Programming (CTCS '85)*, pp. 16–42, Guildford, UK, September 1986.

[43] D. Kozen, C. Kreitz, and E. Richter, "Automating proofs in category theory," in *Proceedings of the 3rd International Joint Conference Automated Reasoning (IJCAR '06)*, pp. 392–407, Seattle, Wash, USA, August 2006.

[44] B. Mazur, "When is one thing equal to some other thing?" June 2007, http://www.math.harvard.edu/~mazur/preprints/when_is_one.pdf.

[45] J. J. Cimino and P. D. Clayton, "Coping with changing controlled vocabularies," in *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, pp. 135–139, Washington, DC, USA, November 1994.

[46] D. E. Oliver, Y. Shahar, E. H. Shortliffe, and M. A. Musen, "Representation of change in controlled medical terminologies," *Artificial Intelligence in Medicine*, vol. 15, no. 1, pp. 53–76, 1999.

[47] N. F. Noy, A. Chugh, W. Liu, and M. A. Musen, "A framework for ontology evolution in collaborative environments," in *Proceedings of the 5th International Semantic Web Conference (ISWC '06)*, pp. 544–558, Athens, Ga, USA, November 2006.

[48] M. C. A. Klein and N. F. Noy, "A component-based framework for ontology evolution," in *Proceedings of the 18th International Joint Conferences on Artificial Intelligence (IJCAI '03)*, vol. 71, CEUR-WS, Acapulco, Mexico, August 2003.

[49] J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, Reading, Mass, USA, 2nd edition, 2001.

[50] R. Rosen, "The representation of biological systems from the standpoint of the theory of categories," *The Bulletin of Mathematical Biophysics*, vol. 20, no. 4, pp. 317–341, 1958.

[51] V. Wiels and S. Easterbrook, "Management of evolving specifications using category theory," in *Proceedings of the 13th IEEE International Conference on Automated Software Engineering*, pp. 12–21, Honolulu, Hawaii, USA, October 1998.

[52] T. Mens, "A formal foundation for object-oriented software evolution," in *Proceedings of IEEE International Conference on Software Maintenance (ICSM '01)*, pp. 549–552, Florence, Italy, November 2001.

Hindawi Publishing Corporation
International Journal of Telemedicine and Applications
Volume 2009, Article ID 461560, 9 pages
doi:10.1155/2009/461560

Research Article

Security Framework for Pervasive Healthcare Architectures Utilizing MPEG-21 IPMP Components

Anastasios Fragopoulos,[1,2] John Gialelis,[1,2] and Dimitrios Serpanos[1,2]

[1] *Industrial Systems Institute, R.C Athena, Stadiou Street, Platani, 26504 Patras, Greece*
[2] *Department of Electrical and Computer Engineering, University of Patras, 26500 Rio Patras, Greece*

Correspondence should be addressed to John Gialelis, gialelis@ece.upatras.gr

Received 29 May 2008; Accepted 6 November 2008

Recommended by Yang Xiao

Nowadays in modern and ubiquitous computing environments, it is imperative more than ever the necessity for deployment of pervasive healthcare architectures into which the patient is the central point surrounded by different types of embedded and small computing devices, which measure sensitive physical indications, interacting with hospitals databases, allowing thus urgent medical response in occurrences of critical situations. Such environments must be developed satisfying the basic security requirements for real-time secure data communication, and protection of sensitive medical data and measurements, data integrity and confidentiality, and protection of the monitored patient's privacy. In this work, we argue that the MPEG-21 Intellectual Property Management and Protection (IPMP) components can be used in order to achieve protection of transmitted medical information and enhance patient's privacy, since there is selective and controlled access to medical data that sent toward the hospital's servers.

1. Introduction

In modern computing era, health is an important aspect for society since it involves every citizen. The current health care system is designed to reward medical doctors and medical institutions to treat people when they are sick, and not acting proactively thus preventing them from being sick. Recently, in many countries the health care system is subject to reform since the rapidly growing populations of elderly combined with the increasing cost of health care services impose several challenges on health care providers, insurance companies, hospitals, and even on patients. All these suggest that first of all healthcare needs a major shift toward a more proactive direction focusing on prevention and/or early detection of acute events and afterwards more scalable and more affordable solutions should be offered. Therefore, all the above-mentioned requirements have accentuated the need for pervasive and ubiquitous embedded e-Health environments, given that limited financial and human resources will be committed.

The modern backbone communication infrastructures, networks, and technologies like GSM, UMTS, WCDMA, WiMAX, Internet, Wireless, and Bluetooth protocols provide extraordinary high data rates, thus allowing cost-efficient and time-efficient remote delivering of medical data that have been collected from the portable, embedded devices that reside onto the end-users (monitored patients) toward remote Medical Servers, for further processing. Furthermore, the pervasiveness and ubiquitous character of modern user-centric nomadic environments, in which the user is the central point surrounded by different types of embedded and small computing devices, like, for example, electrocardiograph (ECG) and pulse, oxygen saturation, and blood pressure sensors, add extra requirements concerning security aspects such as the protection of sensitive medical data and measurements, data integrity and confidentiality, and protection of patients privacy. In such environments, it is imperative to design and deploy efficient and effective network architectures as well as a generic, if possible, communication interface targeted to connect the external networks with the "smart" individual-person (equipped with different types of "smart" embedded medical devices (EMDs)) satisfying the basic security requirements for real-time, secure data communication, and protection of

sensitive medical data and measurements, data integrity and confidentiality, and protection of the monitored patient's privacy. The architectures and the interface must consider the limited resources of the interconnected embedded systems, especially in light of the significant resources required for implementing security, which, in general, are quite resource-hungry leading thus to significant technical problems.

MPEG-21 is a standard that defines mechanisms and tools as means of sharing digital rights, permissions, and restrictions for digital content from content creator to content consumer. It is an XML-based standard that is designed to communicate machine-readable license information in an ubiquitous, unambiguous, and secure manner between peer entities. Although MPEG-21 is a standard that is used for sharing multimedia digital content, under well-defined terms, we argue that it could be used for protection of medical data and user's privacy in a pervasive healthcare environment, assuming that the digital content is the medical data generated by the EMDs. The content creator is any "user" under monitoring and the "end-user" is anyone having access to a remote Medical Server located in a hospital either being supervisors, trained medical stuff, or doctors.

In our work, we propose a generic security framework based on the MPEG-21 standard adopted for wearable EMDs thus filling an existing gap which made these devices and the data acquired by them vulnerable to any kind of attacks. Although the MPEG-21 standard has been adopted for delivery of multimedia content, we argue that it can be used in our architecture in order to enhance users' privacy and to enhance security requirements that are applied in such environments. Moreover, since the medical measurements that taken from patients' EMDs can include different data types, besides raw data, like, for example, a snapshot of electrocardiograph data in a jpeg file, or a video file of electrocardiograph data, the MPEG-21 standard can help us in content delivery, with security enhancement toward different types of end-users. Our endeavor demonstrates that no general-purpose home server is required for processing medical data before sent toward end-users, that is, doctors, medical staff, thus making the proposed architecture a good candidate for portable devices in pervasive healthcare environments exposing limited resources.

We strongly believe that the outcome of this work contributes toward the establishment of a standard context-aware security environment in which parts of sensitive personal information can be viewed only by valid users even though many users can have access to them. The main characteristics of our approach, which directly affects medical safety and treatment efficacy, comprise standardization, minimalizm, and sophistication. Ultimately, affects reduction of total health care cost, thus leading to a better utilization of limited healthcare resources.

More specifically, Section 2 gives a presentation of the relevant work, while Section 3 analyzes the emerging of EMDs security requirements in combination with MPEG-21 standard. Section 4 describes our proposed architecture and Section 5 provides the conclusions and the future work projection.

2. Related Work

Gialelis et al. [1] propose a pervasive healthcare architecture into which a wearable health monitoring system is integrated into a broad telemedical infrastructure allowing high-risk cardiovascular patients to closely monitor changes in their critical vital signs and get experts feedback to help maintain optimal health status. Consistent with the major challenge to provide good quality and reliable health care services to an increasing number of people utilizing limited financial and human resources, they propose a person-based health care system which consists of wearable Commercial of-the-shelf nodes which are already used in the hospital environment, and they are capable of sensing and processing blood-oxygen, blood-pressure, ECGs, and other vital signs and can be seamlessly integrated into wireless personal area networks (WPANs) for ubiquitous real-time patient monitoring. Their architecture lacks safety, security, and privacy considerations, which may lead to serious breaches to architecture's and EMDs functionalities or to users' privacy.

Venkatasubramanian and Gupta [2] made a survey on security solutions for pervasive healthcare environments focusing on securing of data collected by EMDs, securing the communications between EMDs and investigation of mechanisms for controlling the access to medical data. They propose the use of cryptographic primitives, where measurements of physiological values are used for cryptographic keys, eliminating thus the necessity for key distribution, for securing data, and for the establishment of secure communications between two entities. Concerning the access control to medical data, they survey methods that are based on role-based access control (RBAC), extending it for usage in pervasive healthcare environments.

As far as we know, the only proposal for usage of MPEG-21 as a mechanism to access control to medical records has been proposed by Brox [3]. The author links patients records into MPEG-21 digital items and attempts to find access control mechanisms based on MPEG-21 standard. In this work, the author does not provide a clear architecture that implements the usage of MPEG-21 and also he does not use the MPEG-21 Intellectual Property Management and Protection (IPMP) components for the protection of medical records, but he mentions its use as a future and open research issue.

3. Security Requirements—Threat Model

The primary function of a pervasive healthcare architecture is to collect, to process, and to store medical data from different types of EMDs, which are located on the patient, locally into general-purpose computing devices—for further handling or to remote-located servers. In order to identify the security mechanisms that should be deployed into the proposed architecture for pervasive healthcare with the utilization of EMDs, we must identify and classify (i) possible attackers and malicious users of the aforementioned environment; (ii) safety, security, and privacy requirements.

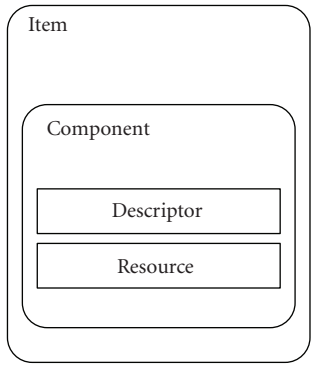


Figure 1: A schematic structure of an MPEG-21 digital item.

Halperin et al. [4] provide a short classification of possible malicious users and adversaries in a pervasive healthcare environment, identifying

(i) *active adversaries* which are malicious users that have explicit interference and physical access to an EMD (i.e., launching probing attacks, tampering attacks),

(ii) *passive adversaries* which are malicious users that have implicit interference with an EMD, through eavesdropping the communication link between an EMD and another system or through side-channel attacks,

(iii) *insiders* which comprise the most probable adversaries in such environments. Such attackers include healthcare personnel, nurses, doctors, IT professionals, and even the patients themselves.

Security is a general requirement in modern computing environments, but in pervasive healthcare systems security is an imperative requirement because those systems handle very sensitive data like medical data and electronic personal records (EPRs). In order to identify the security aspects and requirements in such environments, we have to identify the threat model; some of the most significant threats for pervasive healthcare environments are (a) *nonauthorized access to patient's medical data*, that is, a nurse and a doctor must have different authorities and access control to medical data; (b) *intentional alteration of medical data*, thus leading to incorrect diagnosis and patient's treatment; (c) *disclosure of medical data to third parties* (e.g., to insurance companies or any third parties that may use such records to gain profit), aiming to increase their revenue. In their work, Halperin et al. [4], also, describe a generic framework for analysis, design, and evaluation of security and privacy issues in implantable medical devices (IMDs). Although our architecture does not contain any type of IMDs, a subset of the security and safety goals that they describe in their work applies to us, also. They identify two directions, one concerning safety and utilization goals and the other concerning security and privacy goals, and furthermore they are trying to find tradeoffs and tensions between those directions. Safety and utility goals are traditional requirements in pervasive healthcare systems, involving some security aspects. Such goals are (a) *accessibility to medical data*, that is, ensuring that only appropriate entities must have access to EMDs and their medical data; (b) *accuracy of measurements and data*, that is, all data and measurements taken from the EMD must be accurate; (c) *traceability and identification of EMDs*, that is, an EMD must have mechanisms that allow it to make its presence clear to authorized entities, whenever it is necessary; (d) *maintenance and reconfigurability*, that is, authorized personnel may alter EMDs configuration, locally or remotely; finally, (e) *resources efficiency*, that is, extend EMDs battery life, through minimization of power consumption. Besides safety and utility goals, due to their nature, the modern pervasive healthcare environments have also to fulfill some security- and privacy-related goals and requirements. Those goals are not different from traditional security requirements of computing systems that basically rely on the attainment of authentication and authorization, data confidentiality, data integrity, availability and protection of users' privacy, and have to be extended and possibly reviewed in the context of pervasive healthcare environments. *Authentication.* Refers to methods and mechanisms which allow to an entity to prove to a remote end its identity, that is, in a transaction between two end-users over a possibly unsafe communication network, there must be mechanisms that assure that each part can be authenticated by a remote end. *Authorization*, Refers to access control mechanisms and to the ability of an entity to access shared resources. Two basic subcategories of authorization can be distinguished (a) *personal authorization*, that is, specific people or groups of people may access patients' data and perform specific action over it; (b) *role-based authorization*, that is, a person may have access to medical data based on a specific role that he has, for example, a doctor, a nurse, a caregiver. *Data integrity.* Mechanisms ensure that when there is an interchange of data between two peer entities, the received data and the original ones are the same, and that no intermediate alteration has occurred, for example, through interference of an eavesdropper. In a typical pervasive healthcare system, various messages and data are interchanged between different participating entities, so the integrity of the transmitted messages and data is a basic requirement. *Data confidentiality*, It assures that stored or transmitted data are well protected from possible disclosure. A mean to achieve data confidentiality is through cryptographic mechanisms. *Availability.* It is a security requirement which implies that a malicious user may not be able to passively or actively perform a denial-of-service attack to an EMD, for example, battery attack, memory overflow, jam the communication interface [5], thus preventing it from operating normal and smoothly. *Privacy*, It can be defined as an entity's ability to control how, when, and to what extent personal information about the entity will be communicated to third parties [6]. Anderson [7] defines privacy as the secrecy for the benefit of an individual entity, where secrecy refers to generic mechanisms that do not allow unauthorized usage and access of data and resources. Extending privacy in the context of pervasive healthcare environments, we refer to (i) *EMD-existence and EMD-type privacy*, that is, an EMD should not make its

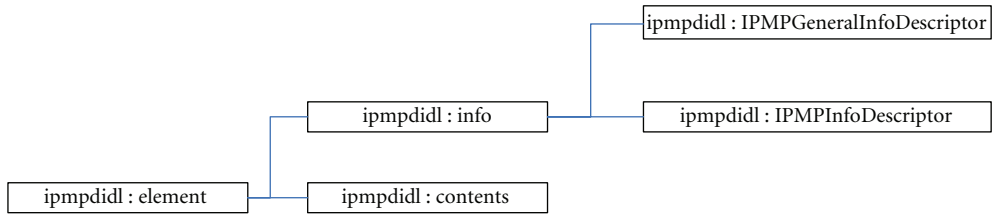


Figure 2: Structure of an IPMP DIDL element.

presence and its type clear to an unauthorized party, since the user (patient) might not want to know nonauthorized users what type of EMD they carry with them and what type of functions it serves.

Our proposed architecture addresses authentication and authorization issues, since each user has a license that defines what he/she has access to. Moreover, data confidentiality is achieved through cryptographic mechanisms, that is, sensitive data are encrypted and data integrity is achieved transparently by the doctor's devices that have access to the data. Finally, the utilization of a DRM mechanism through the MPEG-21 framework helps in the protection of user's privacy, that is, different groups of users have access to different types of medical data.

3.1. MPEG-21 Standard. MPEG-21 is a new standard which has been proposed for primary use in the context of multimedia world, allowing the seamless, transparent, and universal delivery of multimedia content to the end-user, thus solving any interoperability issues. The core notion into the standard is the digital item (DI) that represents the digital asset (e.g., a movie, an audio track, or other data), and which contains the digital content and other related metadata (e.g., creator's details, usage rules and terms, licenses, security related information); the DI hierarchy is represented in Digital Item Declaration Language (DIDL) which is an XML-based document [8]. In the digital item hierarchy, see Figure 1, we identify (i) *resources* which are the individual multimedia data, that is, a picture, audio or video data, or any other data; (ii) *components* which are collection of resources with their descriptors. A component by itself is not an item, but the components are the basic building blocks of items, and (iii) *descriptors* which are metadata that accompany a resource, containing information that concerns all or part of the specific resource. Also, a *container* is a structure which comprises of items and containers, forming thus logical packages for interexchanges between entities.

3.2. MPEG-21 IPMP. The security problems may arise from the fact that the digital item's description, that is, its structure, contents, attributes, and metadata, is a clear XML document and it is easily visible to anyone and vulnerable to nonauthorized usage. Due to that fact, the MPEG-21 includes a part named Intellectual Property Management and Protection (IPMP) which provide mechanisms for protection of digital item. More specifically, MPEG-21 IPMP in conjunction with the MPEG-REL Rights Expression Language (REL) provide a framework that enables all users in the digital contents delivery chain to express their rights and interests in digital items and to have the assurance that those rights and interests will be persistently and reliably managed and protected across a wide range of networks and devices [9]. The core notion in MPEG-21 IPMP is related with the IPMP tools that are used to protect the digital item. Those tools are not predescribed by the standard, but each user, vendor, and so forth, may define and implement his own set of tools which perform basic security functions like encryption/decryption algorithms, authentication and data integrity mechanisms, watermarking, and fingerprinting. With the use of MPEG-21 IPMP components, we may protect the whole DI or a part of it through the encapsulation of the original DIDL elements that we want to protect, with additional information (IPMP Info) that refer to mechanisms and tools for the protection of the original elements. MPEG-21 IPMP defines a new set of IPMP DIDL elements which have the same role and semantics as an element defined in DIDL. The structure of an IPMP DIDL element can be seen in Figure 2.

The *ipmpdidl : info* element contains information about protection and usage rules of the digital content, which may be categorized to (i) information about protection of the whole digital item, which is included in the child element *ipmpdidl : IPMPGeneralInfoDescriptor* and (ii) information about protection of a certain part of the digital item's hierarchy, which is included in the child element *ipmpdidl : IPMPInfoDescriptor*, see Figure 2. Both prementioned child elements have two purposes of existence (a) to describe the tools that are used for digital items protection, and (b) to provide a set of licenses that accompany the content and define its usage rules. More specifically, the *ipmpdidl : IPMPInfoDescriptor* element contains the following child elements (see Figure 3): (a) the *ipmpinfo : Tool* child element which is used to specify a tool (or tools) that protect the specified part of the DI's hierarchy. Each tool has a unique ID and may be either referenced, if it is located in a remote IPMP tools server or it may be included in line to the digital item (in the latter case, the base-64 encoded version of the tool is included); (b) the *ipmpinfo : RightsDescriptor* child element which contains the licenses that govern the usage of contents. Part 5 of MPEG-21 standard [10] defines the MPEG Rights Expression Language (REL) which a language that is used to create licenses that express usage rules and rights set by the creators of the digital items, regarding actions over them. Moreover, the license can be used to convey some other sensitive data, like, for example, decryption keys for the encrypted (protected) contents of the digital item.

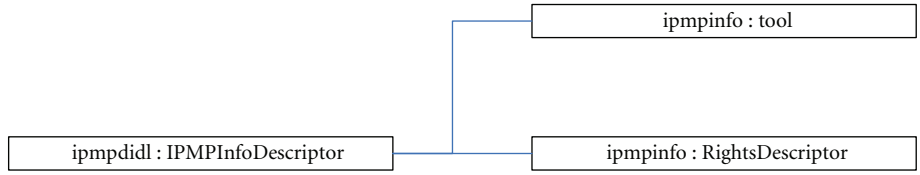


Figure 3: Schematic view of the structure of the *ipmpdidl : IPMPInfoDescriptor* element.

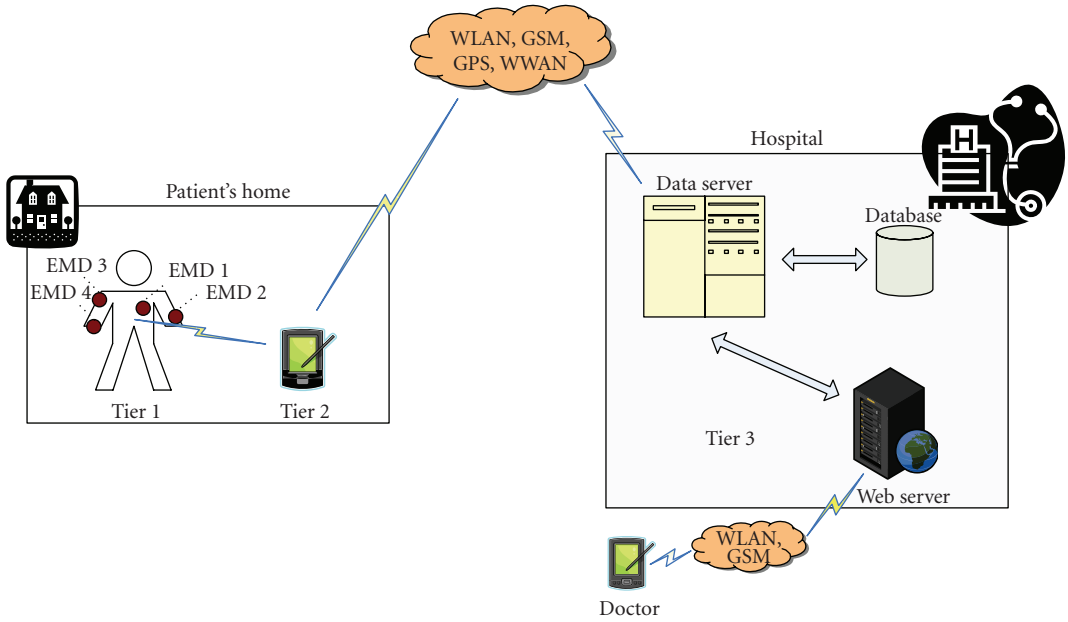


Figure 4: A high-level view of the proposed architecture.

The MPEG-REL defines the *encryptedLicense* element which contains the license's contents in an encrypted form and it is used when the issuer wants to keep the content of the license confidential. When decrypted, the original license's contents are unveiled.

The child element *ipmpdidl : contents* contains the protected digital asset itself which has been protected with a set of tools (e.g., cryptographic tools, watermarking, hashing, digital signatures) that have been described in the child element *ipmpdidl : info*.

4. Proposed Architecture

In the high-level view of the proposed architecture as depicted in Figure 4, we can identify the user at home, equipped with different portable EMDs that transmit their measurements into a PDA which aggregates them for temporal storage and further processing (creation of encapsulated MPEG-21 IPMP-protected metadata file). Then, the generated file is sent to a data server which is located at the patient's hospital and it is stored to a database. If it is necessary, the patient's doctor may extract data from that file remotely through his PDA device, acting and informing accordingly the patient.

More specifically, Tier 1 encompasses a number of portable EMDs equipped with the corresponding sensors, easy to handle and maintain, which are integrated into a wireless wearable personal area network (WWPAN). Each of them can acquire, sample, and process one or more physiological parameters. More specifically, a small (5 cm × 5 cm) portable electrocardiogram device is utilized for the heart activity monitoring by providing a full 12-lead ECG by collecting signals through ten sensors and electrodes with fixed position. The ten input channels (VLA, VRA, VLL, VRL, V1-V6) are sampled at 500 Hz with a 10-bit resolution analog to digital (A/D) converter. A small (watch size) portable wrist (and arm if chosen) blood pressure device is utilized for monitoring patient's blood pressure upon request. Oxygen saturation is monitored through a small (2 cm × 2 cm) portable device which its sensor fits to one of patient's figures. Pulse rate is also measured either from the portable wrist or the ECG device.

Tier 2 encompasses a personal server application running on a personal digital assistant, a cellular phone, a laptop, or a home PC. The personal server undertakes a number of tasks such as a transparent interface to the WWPAN (portable devices), an interface to the user and an interface to Tier 3 (Medical Server & Thin Client). The interface to the WWPAN comprises network configuration and management tasks. The network configuration task, among other functions, supports device and sensor registration and initialization, customization, and network communication

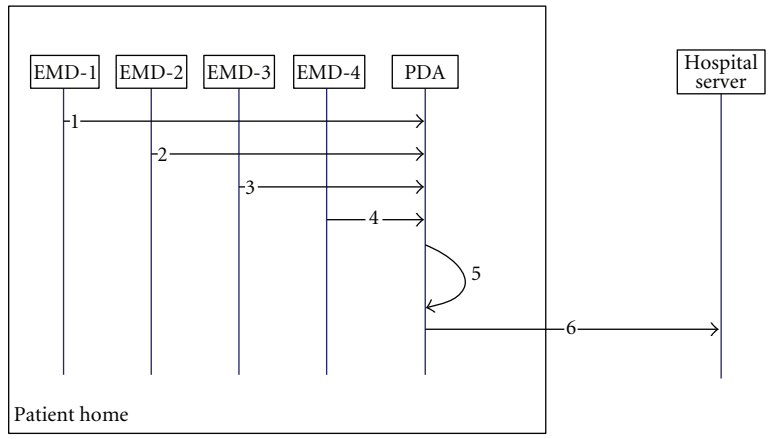


Figure 5: A sequence diagram of flow of medical data measurements from patient's home toward the data server which is located at the patient's hospital.

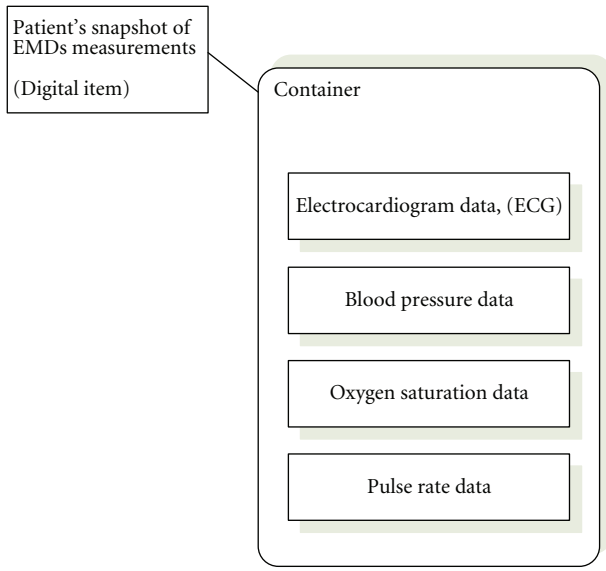


Figure 6: Abstract view of MPEG-21 digital item that presents a snapshot of EMDs measurements for a specific patient.

setup. After the network is configured, the personal server application is initiated in order to manage the network, that is, channel and time synchronization, data acquisition, data processing, data fusion, and so forth. Furthermore, upon the availability of a communication channel, the personal server establishes a secure link toward Tier 3 and transmits reports or files which are being further processed and integrated in the patient's health record. In case the availability of a communication channel is not possible, the personal server stores the data locally until the channel becomes available.

Tier 3 comprises the Medical Server which, actually, encompasses other servers such as emergency server, healthcare server, medical records server, mail server, and the "thin-client" used for remote accessing of Medical Server applications utilizing a "visualization" applet under secure link.

The Medical Server runs a large diversity of applications comprising setting up communication channel to personal servers, collecting users' reports, integrating data into patients' medical records data bases, processing medical records upon authorized doctors and/or specialized personnel demand, processing patients messages and other critical applications.

As soon as the medical data measurements are collected by the EMDs, they are sent toward the PDA (Tier 2) which is responsible for the creation of the MPEG-21-protected container that contains the medical measurements with appropriate usage licenses, as it can be seen in Figure 5.

In steps 1–4, medical data measurements from different user-located EMDs (ECG, blood pressure and oxygen saturation, and pulse rate) are taken and sent toward the PDA, which in turn, aggregates them and creates an initial packaged data file that contains the previous data measurements. Afterwards, the packaged file is amended to further processing in order to be "encoded" according to MPEG-21 terms, with appropriate metadata, protection information, and patient's instructions (relevant usage licenses) (step 5). Then, the new MPEG-21-based packaged file is sent to the hospital's servers for further handling, that is, storage in data servers, and informing physicians and doctors about patient's situation (step 6).

In the following figures, we depict an abstract view of an MPEG-21 digital item that contains EMDs measurements, Figure 6, and the schematic encapsulation of them into an IPMP container, see Figure 7. As we have stated before, the MPEG-21-based encapsulation of data protects it from unauthorized view and possible malicious usage.

In Section 4, we have emphasized that in such pervasive healthcare environments the protection of patient's privacy and medical data confidentiality are key issues. So when the MPEG-21-based packaged file comes to the hospital, its structure and its contents should not be fully accessible to nonappropriate users. We identify three groups of people that should have different access rights to the protected contents (i) the IT supervisors which are responsible for storing the received packaged files to databases, (b) the

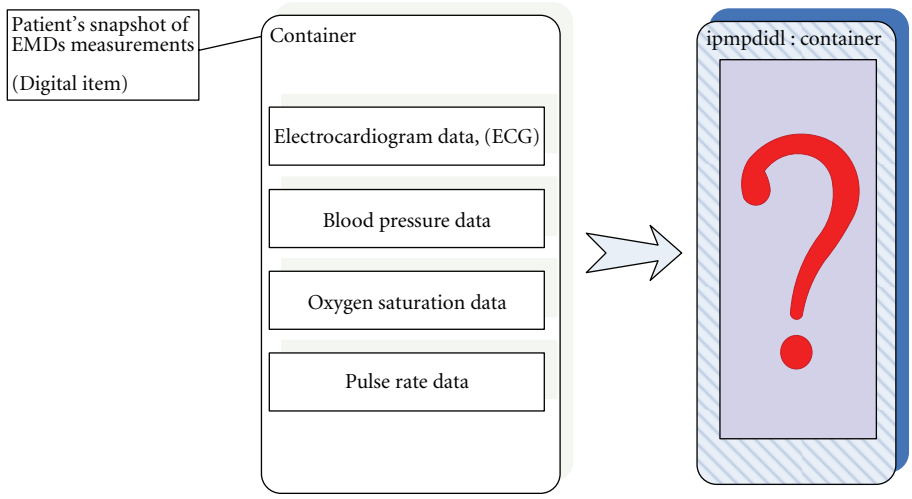


Figure 7: High-level view of encapsulation of a digital item using IPMP framework.

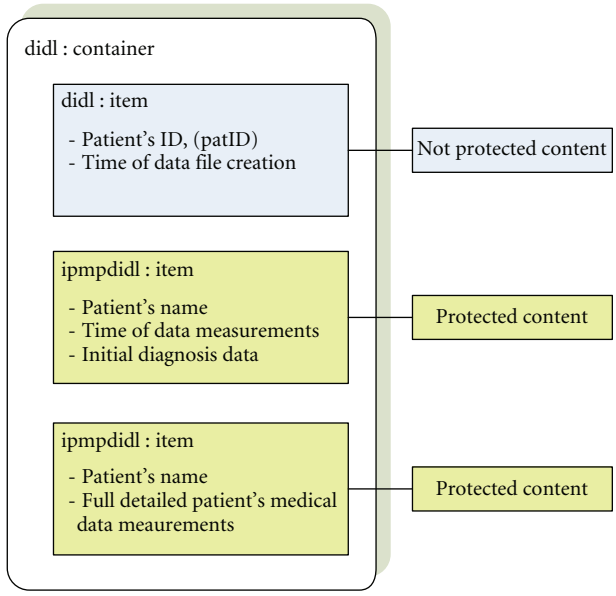


Figure 8: A schematic view of the encapsulated medical data into an MPEG-21 container. The first item is not protected, and it contains some data which can be used by the IT personnel in order to store the protected file into appropriate databases. The second item is protected, and it contains some part of personal data for the patient. Finally, the last item is also protected and includes the detailed medical data measurements. Each item is encrypted using different encryption keys, so each item can be seen only by users that possess the appropriate key for decryption.

specialized medical personnel like, for example, the nurses, and (c) the doctors, who are responsible for the attendance of the patients. Each of these groups must have different access rights to protected digital contents. For example, the IT stuff may just see the patient's identification number, (patID), which is unique for each patient, in order to store the packaged file in data server and update patient's electronic record. On the other hand, a nurse or trained medical stuff must be able to see some details about the patient, for example, a part of patients' personal data like surname, name, and maybe an initial diagnosis, which may have automatically taken part at patient's home after the occurrences of data measurements, in order to inform the patient's doctor for further handling of the patient's situation. Finally, the patient's doctor may have full access to contents of MPEG-21-packaged files, that is, patients personal data, detailed medical data measurements taken by EMDs, time of measurements, and initial automated diagnosis. So the generated MPEG-21 container contains three different items; the first item is unprotected and contains a text data file with the patient's identification number, which can be used for storage of the file to a database for further handling, and some other data, like the time that the file was created at the user's PDA. The other two items are protected with the MPEG-21 IPMP components and are indented for use by the trained medical staff and the doctors, containing appropriate data. Each of these items contains a license of usage, in an encrypted form, which grants access to contents of file to specific user. The license also contains the key which has been used from the patient's PDA for the encryption of the medical data file. We assume that each patient and each of the doctors and trained medical staff have a prearranged set of public-private keys; those key pairs are assigned by the main hospital's server and can be revoked at any time since there is continuous communication between the hospital's server and the patient's (home) server. So the license for data usage can be encrypted using the public key of the entity with which it is related, for example, the doctor with whom the patient is related. In order to access the protected items' contents, the end-user must decrypt the license with his private key, retrieve the decryption key from the license, and decrypt the encrypted contents.

Following, in Figure 8, we depict a schematic view of the encapsulated MPEG-21-based data file which is transmitted from patient's home server toward hospital data servers.

```
<DIDL xmlns:xi="http://www.w3.org/2001/XInclude"
        xmlns:dii="urn:mpeg:mpeg21:2002:01-DII-NS"
        xmlns="urn:mpeg:mpeg21:2002:02-DIDL-NS"
        xmlns:ipmpdidl="urn:mpeg:mpeg21:2004:01-IPMPDIDL-NS"
        xmlns:ipmpinfo="urn:mpeg:mpeg21:2004:01-IPMPINFO-NS"
        xmlns:dip="urn:mpeg:mpeg21:2005:01-DIP-NS">
  <Container>
        <Item>
        <Descriptor id = "item_1">
        <Statement mimeType="text/plain">Patient's Identification Data</Statement>
        </Descriptor>
        <Component>
        <Resource xmlns:ns1="urn:example:01" mimeType="text/plain" ns1:localPath="patientID.txt"
        encoding="base64">cGF0aWVudCdzIElEIGlzOiBwYXRJRA0KDQpUaW1lIG9mIGRhdGEgZmlsZSBjcmVhdGlvbiBpczo
        gMjAwOC0wNC0xMlQyMzozNjo1Mg==</Resource>
        </Component>
        </Item>

        <ipmpdidl:Item>
                <ipmpdidl:Identifier>
                        <ax:ObjectIdentification xmlns:ax="urn:axmedis:01">
                                <dii:Identifier>urn:axmedis:00000:obj:d8c8e292-5d15-4cd9-b19b-71a98ba14b7e</dii:Identifier>
                        </ax:ObjectIdentification>
                </ipmpdidl:Identifier>
                <ipmpdidl:Info>
                        <ipmpinfo:IPMPInfoDescriptor>
                                <ipmpinfo:Tool>
                                        <ipmpinfo:ToolBaseDescription>
                                                <ipmpinfo:IPMPToolID>
                                                        urn:axmedis:ipmp:tool:id:0008
                                                </ipmpinfo:IPMPToolID>
                                        </ipmpinfo:ToolBaseDescription>
                                        <ipmpinfo:Remote ref="urn:IPMP_Tools:ToolServer:ToolSymEnc08"/>
                                </ipmpinfo:Tool>
                        </ipmpinfo:IPMPInfoDescriptor>
                </ipmpdidl:Info>
                <ipmpinfo:RightsDescriptor>
                        <ipmpinfo:License>
                                <r:license>
                                        <r:encryptedLicense
                                        Type="http://www.w3.org/2001/04/xmlenc#Content">
                                                <enc:EncryptionMethod
                                                Algorithm="http://www.w3.org/2001/04/xmlenc#rsa1_5"/>
                                                        <dsig:KeyInfo>
                                                                <dsig:KeyName>asymmetricKey</dsig:KeyName>
                                                        </dsig:KeyInfo>
                                                        <enc:CipherData>
                                                                <enc:CipherValue>
                                                                        Ktd63SDfkDWEoQ...
                                                                </enc:CipherValue>
                                                        </enc:CipherData>
                                        </r:encryptedLicense>
                                </r:license>
                        </ipmpinfo:License>
                </ipmpinfo:RightsDescriptor>
                <ipmpdidl:Contents>maObDtRqBAgm+9MshKHJZ+UBFO9XS7L3a2yU31h0/N19M/
                4rU1Lj412BGmaKQgJDvf2eiPD7sbm73ueNPSHZqNrmzztYIX/0I9J6YA3gYWX10PLrauA//
                41P4QO7yDGikRSkmiaeylc4YWmA9rwXIKc4nonLzrUNyGCPu7+qtvk06T2c+qaGkZkiL52j+At3EANCa+1g1/7skX2BwJZ8s..... </
                 ipmpdidl:Contents>
        </ipmpdidl:Item>

        <ipmpdidl:Item>
                <ipmpdidl:Identifier>
                        <ax:ObjectIdentification xmlns:ax="urn:axmedis:01">
                                <dii:Identifier>urn:axmedis:00000:obj:53d99252-d4ae-4bfd-8994-cab444514d88</dii:Identifier>
                        </ax:ObjectIdentification>
                <ipmpdidl:Info>
                        <ipmpinfo:IPMPInfoDescriptor>
                                <ipmpinfo:Tool>
                                        <ipmpinfo:ToolBaseDescription>
                                                <ipmpinfo:IPMPToolID>
                                                        urn:axmedis:ipmp:tool:id:0008
                                                </ipmpinfo:IPMPToolID>
                                        </ipmpinfo:ToolBaseDescription>
                                        <ipmpinfo:Remote ref="urn:IPMP_Tools:ToolServer:ToolSymEnc08"/>
                                </ipmpinfo:Tool>
                        </ipmpinfo:IPMPInfoDescriptor>
                </ipmpdidl:Info>
                <ipmpinfo:RightsDescriptor>
                        <ipmpinfo:License>
                                <r:license>
                                        <r:encryptedLicense
                                        Type="http://www.w3.org/2001/04/xmlenc#Content">
                                                <enc:EncryptionMethod
                                                Algorithm="http://www.w3.org/2001/04/xmlenc#rsa1_5"/>
                                                        t<dsig:KeyInfo>
                                                                <dsig:KeyName>asymmetricKey</dsig:KeyName>
                                                        </dsig:KeyInfo>
                                                        <enc:CipherData>
                                                                <enc:CipherValue>
                                                                        gyTE96639In0FZFY2/rMJ01EArmKZsR5VW3rwoPxw...
                                                                </enc:CipherValue>
                                                        </enc:CipherData>
                                        </r:encryptedLicense>
                                </r:license>
                        </ipmpinfo:License>
                </ipmpinfo:RightsDescriptor>
                <ipmpdidl:Contents>5UTBn6qHURmHMv3U5D9KB55HzJcTfIrAlrfxPVB/
                7DPDRPxg9XJsM9Dhezrkp2na6LTqOESjGmIeIzrtVYlT255V/
                TilLtxbqwPt+pPhCfa+q7DQ1mB7Yx0kiszJiTUT0oEmnlq+QhnkG+xrVNub9l6AL7euIFI8Ch2sAX/
                29ZhATZmDOh0aO0cMjxWFN9nPLccGQz0ZAvbsEm0b7KT+dBKmS08GGW6AIMjskk. . . . . .
                 </ipmpdidl:Contents>
        </ipmpdidl:Item>
  </Container>
</DIDL>
```

Figure 9: An example view of an MPEG-21 IPMP container which contains three items—one not protected and two protected with the use of a cryptographic tool that is remote referenced at the location *urn:IPMP_Tools:ToolServer:ToolSymEnc08*. Before the user decrypts the contents, he must decrypt the license that accompanies the content, which will provide to the user the decryption key and usage rules of the content. The license is binded to a specific user and it is included into the item encrypted using a public-key algorithm; in our implementation, we used RSA [11], and the license has been encrypted using the users' public key. For the creation of the aforementioned MPEG-21 IPMP-based file, we have used the AXMEDIS editor [12].

Finally, in Figure 9, we present an XML-view of the MPEG-21 IPMP-protected container that contains the three prementioned items. The contents of each item are protected (encrypted) with the use of a tool that implements a specific symmetric cryptographic algorithm, and it is remote referenced into the digital items' description. Even if one malicious user has access to the remote location of the tool, he needs some cryptographic sensitive information like, for example, a decryption key for decoding the protected contents; the key for the decryption is included into the encrypted license that accompanies each item, which furthermore defines the usage rules of the contents. Each item contains some information that identifies the user for whom this item is destined, that is, item 1 has to be processed by the IT stuff, item 2 has to be taken care by a nurse, and so forth. After the reception of each item, each relevant user initiates the procedure which will decrypt his/her license—with the use of his private key and then he/she will decrypt the protected contents—with the symmetric key that is embedded into his/her license, thus revealing the original packaged medical information.

5. Conclusions

We strongly believe that the outcome of this work contributes toward the establishment of a standard security environment in which parts of sensitive personal information can be viewed only by valid users even though many users can have access to them. The main characteristics of our approach, which directly affect medical safety and treatment efficacy, comprise standardization, minimalizm, and sophistication. Ultimately, affects reduction of total health care cost, thus leading to a better utilization of limited healthcare resources. Moreover, our endeavor demonstrates that no general-purpose home server is required for processing of medical data before sent toward end-users, that is, doctors, medical staff, thus making the proposed architecture a good candidate for portable devices in pervasive healthcare environments exposing limited resources. In our proposal, we use the MPEG-21 standards' IPMP components in order to enhance users' privacy and achieve security requirements that are applied in such environments. To our knowledge, there is no other security framework compliant with MPEG-21 that has been applied to protect such content, thus no comparison with other similar frameworks is provided. As a matter of fact, this aspect constitutes a future research activity. Furthermore, with respect to future work, we are planning to exploit some other elements of the MPEG-21 standard which, for example, would allow us to include (in line) into the IPMP-protected digital item the protection tools that are used to govern the protected item rather than remote referenced them, and also we could use some other elements of MPEG-21 IPMP that allow digitally authentication of the protection tools, that is, each tool carries a digital signature which validated each time the tool is used, before appliance of them to the protected contents. A lot of our research work is taken place investigating issues that are related with the security of our architecture under known attacks. Moreover, as the MPEG-21 is a quite new open standard that can be used for protection of sensitive data with the use of DRM, as a future work, we are investigating the identification of possible leaks and vulnerabilities of the standard, that may lead to attacks and a compromised system. We are also investigating issues that are related with the safe destruction of the medical data after their viewing; in that context, the use of trusted platform architectures [13] could be a direction that will lead to a solution. At this time, we assume that the hospital's server is a trusted one, and that the devices on which those data are viewed are also safe and trusted, and that there is no possibility for data storage or alteration.

References

[1] J. Gialelis, P. Foundas, A. Kalogeras, M. Georgoudakis, A. Kinalis, and S. Koubias, "Wireless wearable body area network supporting person centric health monitoring," in *Proceedings of the 7th IEEE International Workshop on Factory Communication Systems (WFCS '08)*, pp. 77–80, Dresden, Germany, May 2008.

[2] K. Venkatasubramanian and S. Gupta, "Security for pervasive healthcare," in *Security in Distributed, Grid, Mobile, and Pervasive Computing*, chapter 15, pp. 349–366, CRC Press, Boca Raton, Fla, USA, 2007.

[3] G. A. Brox, "MPEG-21 as an access control tool for the National Health Service Care Records Service," *Journal of Telemedicine and Telecare*, vol. 11, supplement 1, pp. 23–25, 2005.

[4] D. Halperin, T. Kohno, T. S. Heydt-Benjamin, K. Fu, and W. H. Maisel, "Security and privacy for implantable medical devices," *IEEE Pervasive Computing*, vol. 7, no. 1, pp. 30–39, 2008.

[5] D. Halperin, T. S. Heydt-Benjamin, B. Ransford, et al., "Pacemakers and implantable cardiac defibrillators: software radio attacks and zero-power defenses," in *Proceedings of IEEE Symposium on Security and Privacy (SP '08)*, pp. 129–142, Oakland, Calif, USA, May 2008.

[6] H. C. A. van Tilborg, Ed., *Encyclopedia of Cryptography and Security*, Springer, Berlin, Germany, 2005.

[7] R. J. Anderson, *Security Engineering: A Guide to Building Dependable Distributed Systems*, John Wiley & Sons, New York, NY, USA, 2001.

[8] I. S. Burnett and F. Pereira, "An introduction to MPEG-21," in *The MPEG-21 Book*, chapter 2, pp. 31–67, John Wiley & Sons, New York, NY, USA, 2006.

[9] ISO/IEC JTC1 SC29/WG11/N4681, "MPEG-21 Requirements v 1.0," Jeju, Korea, March 2002.

[10] ISO/IEC 21000-5:2004, "Information technology—multimedia framework (MPEG-21)—Part 5: Rights Expression Language".

[11] B. Schneier, *Applied Cryptography*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1996.

[12] "Automating Production of Cross Media Content for Multichannel Distribution," http://www.axmedis.org/com.

[13] Trusted Computing Group, https://www.trustedcomputinggroup.org/.

Hindawi Publishing Corporation
International Journal of Telemedicine and Applications
Volume 2009, Article ID 101382, 17 pages
doi:10.1155/2009/101382

Research Article

Agent-Oriented Privacy-Based Information Brokering Architecture for Healthcare Environments

Abdulmutalib Masaud-Wahaishi[1] and Hamada Ghenniwa[2]

[1] *College of Information Technology, United Arab Emirates University, P.O. Box 15551, Al-Ain, United Arab Emirates*
[2] *Department of Electrical & Computer Engineering, The University of Western Ontario, 1201 Western Road, London, ON, Canada N6G 1H1*

Correspondence should be addressed to Abdulmutalib Masaud-Wahaishi, amasaud@uaeu.ac.ae

Received 1 May 2008; Accepted 24 December 2008

Recommended by Yang Xiao

Healthcare industry is facing a major reform at all levels—locally, regionally, nationally, and internationally. Healthcare services and systems become very complex and comprise of a vast number of components (software systems, doctors, patients, etc.) that are characterized by shared, distributed and heterogeneous information sources with varieties of clinical and other settings. The challenge now faced with decision making, and management of care is to operate effectively in order to meet the information needs of healthcare personnel. Currently, researchers, developers, and systems engineers are working toward achieving better efficiency and quality of service in various sectors of healthcare, such as hospital management, patient care, and treatment. This paper presents a novel information brokering architecture that supports privacy-based information gathering in healthcare. Architecturally, the brokering is viewed as a layer of services where a brokering service is modeled as an agent with a specific architecture and interaction protocol that are appropriate to serve various requests. Within the context of brokering, we model privacy in terms of the entities ability to hide or reveal information related to its identities, requests, and/or capabilities. A prototype of the proposed architecture has been implemented to support information-gathering capabilities in healthcare environments using FIPA-complaint platform JADE.

1. Introduction

Healthcare systems are characterized by shared and distributed decision making and management of care. The distributed nature of the knowledge among different healthcare locations implies that a request may not be completely satisfied at a specific location or that one or more healthcare location may contain information similar to, though not exactly the same as, that required by the request.

Many initiatives and programs have been established to promote the development of less costly and more effective healthcare networks and systems at national and international scales. The objectives of these healthcare networks are to improve diagnosis through online access to medical specialists, online reservation of analysis and hospital services by practitioners extended on wide global scale, transplant matching, and so forth. A complete electronic medical patient-case file, which might be shared between specialists and can be interchanged between hospitals and with general practitioners (GPs), will be crucial in diagnosing diseases correctly, avoiding duplicative risky and expensive tests, and developing effective treatment plans.

However, medical patient-case files may contain some sensitive information about critical and vital topics such as abortions, emotional and psychiatric care, sexual behaviors, sexually transmitted diseases, HIV status, and genetic predisposition diseases. Privacy and the confidentiality of medical records have to be especially safeguarded. Without broad trust in medical privacy, patients may avoid crucial healthcare provision.

Healthcare professionals and care providers prefer to have the ability of controlling the collection, retention, and distribution of information about themselves. On the other hand, healthcare service providers need to effectively

manage and prevent any abuse of the information or service they provide in addition to the ability of protecting their identities. An important feature of the various healthcare sectors is that they share similar problems and are faced with challenges that can be characterized as follows.

(i) In open-distributed healthcare environments, it is no longer practical to expect healthcare clinicians, staff, care providers, and patients to determine and keep track of the information and services relevant to his/her requests and demands. For example, a patient will be ubiquitously able to access his/her medical record from anywhere at any time or may request medical services offered by available healthcare centers in a particular city without being aware of the distributed sources and irrespective of their locations. In addition, an application should be able to manage distributed data in a unified fashion. This involves several tasks, such as maintaining consistency and data integrity among distributed data sources, and auditing access.

(ii) The distributed nature of the knowledge among multiple healthcare locations may require collaboration for information gathering. For example, each unit in a hospital keeps its own information about patients' records.

(iii) The solution of specific medical problem includes complex activities and requires collaborative effort of different individuals who posses distinct roles and skills. For example, the provision of care to hospitalized patients involves various procedures and requires the coordinated interaction amongst various staff and medical members. It is essential that all the involved medical staff and professionals must coordinate their activities in a manner that will guarantee the best appropriate treatment that can be offered to the patient.

(iv) A recent survey shows that 67% of the American national respondents are concerned about the privacy of their personal medical records, 52% fear that their health insurance information might be used by employers to limit job opportunities, while only 30% are willing to share their personal health information with health professionals not directly involved in their case. As few as 27% are willing to share their medical records with drug companies [1].

To explore such issues, distributed healthcare systems need to have an access to a service that can enable collaboration between different healthcare service requesters and providers. Brokering facilitates achieving better coordination among various healthcare service requesters and providers, and permits healthcare personnel to get access to different services managed by various providers without having to be aware of the location, identities, access mechanisms, or the contents of these services.

The proactive health systems have the potential to improve healthcare access and management which significantly lower the associated incurred costs through efficiently controlled information flow between various physicians, patients, and medical personnel, yet threaten to facilitate data sharing beyond any privacy concerns. The high degree of collaborative work needed in healthcare environments implies that developers and researchers should think of other venues that can manage and automate this collaboration efficiently.

However, privacy concerns over inappropriate use of the information make it hard to successfully exploit and achieve the gains from sharing such information. This dilemma restricts the willingness of individuals and personnel to disseminate or publicize information that might lead to adverse outcomes. This paper presents an agent privacy-based information brokering architecture that supports ad hoc system configurations emphasizing the strategies for achieving privacy in healthcare environments. Within the context of brokering, we view privacy as "*the ability of entities to decide upon revealing or hiding information related to their identities, requests and capabilities in open distributed environments.*"

2. Related Work

Privacy concerns are key barriers to the growth of health-based systems. Legislation to protect personal medical information was proposed and put in effect to help building a mutual confidence between various participants in the healthcare domain.

Privacy-based brokering protocols were proposed in many application domain such as E-auctions [2], data mining [3], and E-commerce. Different techniques were used to enable collaboration among heterogeneous cooperative agents in distributed systems including brokering via middle agents. These middle agents differ from the role they play within the agent community [4–6]. The work in [7] has proposed an agent-based mediation approach, in which privacy has been treated as a base for classifying the various mediation architectures only for the initial state of the system. In another approach, agents capabilities and preferences are assumed to be common knowledge, which might violate the privacy requirements of the involved participants [8]. Other approaches such as in [9–11] have proposed frameworks to facilitate coordination between web services by providing semantic-based discovery and mediation services that utilize semantic description languages such as OWL-S [12] and RDF [13]. Another recent approach distinguishes a resource brokering architecture that manages and schedules different tasks on various distributed resources on the large-scale grid [14]. However, none of the above-mentioned approaches has treated privacy as an architectural element that facilitates the integration of various distributed systems of an enterprise.

Several approaches were proposed for integration of distributed information sources in healthcare [15]. In one approach [16], the focus was on providing management assistance to different teams across several hospitals by coordinating their access to distributed information. The brokering architecture is centralized around a mediator agent, which allocates the appropriate medical team to an available operating theatre in which the transplant operation may be performed. Other approach attempts to provide agent-based medical appointments scheduling [17, 18], in these approaches the architecture provides matchmaking mechanisms for the selection of appropriate recipient candidates whenever organs become available through a matchmaking agent that accesses a domain-specific ontology.

Other approaches proposed the use of privacy policies along with physical access means (such as smartcards), in which the access of private information is granted through the presence of another trusted authority that mediate between information requesters and information providers [19, 20]. A European IST project [21], TelemediaCare, Lincoln, UK, developed an agent-based framework to support patient-focused distant care and assistance, in the architecture composes two different types of agents, namely, stationary "static" and mobile agents. Web service-based tools were developed to enable patients to remotely schedule appointments, doctor visits, and to access medical data [22].

Different approaches had been suggested to protect the location privacy in open-distributed systems [23]. Location privacy is a particular type of information privacy that can be defined as "the ability to prevent other parties from learning one's current or past location". These approaches range from anonymity, pseudonymity, to cryptographic techniques. Some approaches focus on using anonymity by unlinking user personal information from their identity. One available tool is called anonymizer [24]. The service protects the Internet protocol (IP) address or the identity of the user who views web pages or submits information (including personal preferences) to a remote site. The solution uses anonymous proxies (gateways to the Internet) to route user's Internet traffic through the tool. However, this technique requires a trusted third party because the anonymizer servers (or the user's Internet service provider, ISP) can certainly identify the user. Other tools try not to rely on a trusted third party to achieve complete anonymity of the user's identity on the Internet, such as Crowds [25], Onion routing [26], and MIX networks [27].

Various programs and initiatives have proposed a set of guidelines for secure collection, transmission, and storage of patients' data. Some of these programs include the Initiative for Privacy Standardization in Europe (IPSE) and the Health Insurance Portability and Accountability Act (HIPAA) [28, 29]. Yet, these guidelines need the adoption of new technology for healthcare requester/provider interaction.

3. Brokering Requirements for Distributed Healthcare Systems

Brokering enables collaboration between different service requesters and providers, and allows the dynamic interpretation of requests for the determination of relevant service providers. For service providers, the brokering services permit dynamic creation of services' repositories after suitable assembly of service advertisements available from the various providers, or other additional activities. The major functional requirements of a brokering service include the following.

(i) *Provision of registration services*: the registration and naming service allows building up a knowledge base of the environment that can be utilized to facilitate locating and identifying the relevant existing service sources and their contents for serving a specific request. It is crucial to be able to identify the subset of relevant information at a source, and to combine partially relevant information across different sources; this requires the process of identification and retrieval of a subset of required service at any source. It is clear that in such environment, different sources would provide relevant information to a different extent. The most obvious choice of the source from which information will be retrieved is the one which returns most (or all) of the relevant request. In that case, the user will have to keep track of which source has the most relevant information.

(ii) *The acceptance of providers' service descriptions*: to enable the dynamic discovery of services, a mechanism is required to describe the capability aspects of services, such as the functional description of a service, the conditions and the constraints of the service, and the nature of the results.

(iii) *Receiving services' requests*: to enable requesters to define and describe the required parameters that are needed to represent a request.

(iv) *Interaction*: brokers may engage (on behalf of requesters) in the process of negotiation with various service providers to serve a request. The interaction requires a set of agreed messages, rules for actions based upon reception of various messages.

(v) *Communication*: the communication capability allows the entities to exchange messages with the other elements of the environment, including users, agents, and objects. In order to perform their tasks, these entities need to depend heavily on expressive communication with others not only to perform requests, but also to propagate their capabilities, advertise their own services, and explicitly delegate tasks or requests for assistance.

4. The Brokering Layer: Privacy-Based Agent-Orinted Architecture

Developing the brokering services comprises the automation of privacy to enhance the overall security of the system and accordingly entities should be able to define the desired degree of privacy. In fact, the brokering service permits entities to participate in the environment with different roles, and hence be capable of automating their privacy concerns and select a particular privacy. The challenge here is how to architect a service that could provide means and mechanisms by which entities would be able to interact with each other and determine any privacy degree that suits a particular situation. Such interaction is characterized by the nondeterministic aspect in addition to the dynamic nature of the environment, where these entities exist and operate for which they require to be able to change configurations to participate in different roles. These requirements could not be accomplished using traditional ways of manually configuring software.

We strongly believe that agent orientation is an appropriate design paradigm for providing coordination services and mechanisms in such settings. Indeed, such a paradigm is essential to modeling open, distributed, and heterogeneous environments in which an agent should be able to operate as a part of a community of cooperatively distributed systems environments, including human users. A key aspect of

agent orientation is the ability to design artifacts that are able to perceive, reason, interact, and act in a coordinated fashion. Here, we view agent orientation as a metaphorical conceptualization tool at a high level of abstraction (knowledge level) that captures supports and implements features that are useful for distributed computation in open environments. These features include cooperation, coordination, interaction, as well as intelligence, adaptability, economic and logical rationalities. We define an agent as an individual collection of primitive components that provide a focused and cohesive set of capabilities. We focus on the notion of agenthood as a metaphorical conceptualization tool at a high level of abstraction (knowledge level) that captures supports and implements features that are useful for distributed computation in open environments.

Architecturally, the brokering service is viewed as a layer of services and is modeled as an agent with a specific architecture and interaction protocol that are appropriate to carry the required privacy degree. The challenge in this context is how to architect the brokering layer with the appropriate set of services that enable cooperation across the different degrees of privacy. The interaction protocols represent both the message communication and the corresponding constraints on the content of messages. They describe the sequence of messages among agents, and illustrate various protocols that satisfy a desired privacy requirement. The focus for designing these patterns is to provide a mechanism to reduce the costs and risks that might be a result of violating privacy requirements. The patterns provide mechanisms allowing users (human/agents) to adjust the privacy attributes, and allowing these users to achieve and accomplish their tasks in addition to protecting their desired privacy attributes.

The agent interaction requires a set of agreed messages, rules and assumption of communication channels. These rules and constraints can be abstracted as agents' patterns that define various protocols for every possible privacy requirement. Using these protocols, agents would be able to protect the privacy aspects of the most concern. From the privacy standpoint, the brokering services are categorized into different roles that are classified according to the participants' (providers and requesters) desired degree of privacy. These degrees of privacy control the proper interaction patterns and will vary from a specific scenario to another. The brokering layer takes in consideration the protection of any privacy desires required by requesters, providers, or both.

Here, we define the degree of privacy in terms of three attributes: the entity identity, capability, and goals. Therefore, an agent can categorize its role under several privacy degrees. Formally, an agent can be represented as a 2-tuple $Ag \equiv \langle (RA : Id, G); (PA : Id, Cap) \rangle$, where RA and PA refer to the agent role as requester and provider while Id, G, and Cap, respectively, refer to the agent identity, goals, and capabilities, which might have a null value. For example, an agent might participate with a privacy degree that enables the hiding of its identity as a requester by setting the value of Id to null. Tables 1, 2 summarize the different scenarios and roles that might be played by the brokering layer categorized by the possible privacy concern of the requester (RA) and provider (PA) agents.

The layer permits various entities to participate in the environment with different roles, and hence be capable of automating their privacy concerns and select a particular degree. Each layer role is represented as a special broker with a specific architecture and interaction protocol that is appropriate to serve requests from various participants while maintaining the required privacy degree. An agent role is an abstract description of an entity with the specified functionalities. The brokering layer has the ability to interact, solicit help, and delegate services' requests from other available brokering agents who support different privacy degrees.

Responsibilities are separated and defined according to the roles played and the required degree of privacy. Within the layer two sets of brokering agents are available to service requesters and providers. The first set handles interactions with requesters according to the desired privacy degree that is appropriate to their preferences while the other set supports privacy degrees required by service providers.

Figure 1 shows a logical view of the brokering services and the relevant entities that are involved in any brokering scenario. Every brokering pattern is accomplished by the composition of the requester role, brokering agents, and the provider role, in which the interaction scenarios are produced automatically. A complete brokering session is divided into several stages, starting from requester-to-brokering layer interaction, brokering layer intra-interaction, and broker layer-to-provider interaction. Note that in the figure a negation on a specific privacy attribute variable exemplifies that the corresponding privacy attribute is hidden from the environment.

5. The Brokering Protocols: Privacy-Based Interaction Patterns

The brokering protocols describe a cooperative multibrokering system, which provides the solution for interaction among participants in a dynamic and heterogeneous environment of service providers and requesters. Each brokering entity performs basic brokering functionality, such as service discovery, dynamic service composition, and knowledge sharing with the community according to a required privacy degree. A brokering entity within the layer is called a broker hereafter.

Brokers within the layer might represent a set of services in which providers can advertise their service capability. The brokering protocols regulate and govern service knowledge discovery and sharing of acquired knowledge by defining interaction patterns that are composed of a set of messages that can be exchanged by other brokers within the layer or other registered entities that might benefit of the functionalities supported by the overall brokering service. The architecture permits the brokering agents to have various combinations with other brokering entities which support different privacy degrees. The following section describes the different interaction patterns supported by the brokering

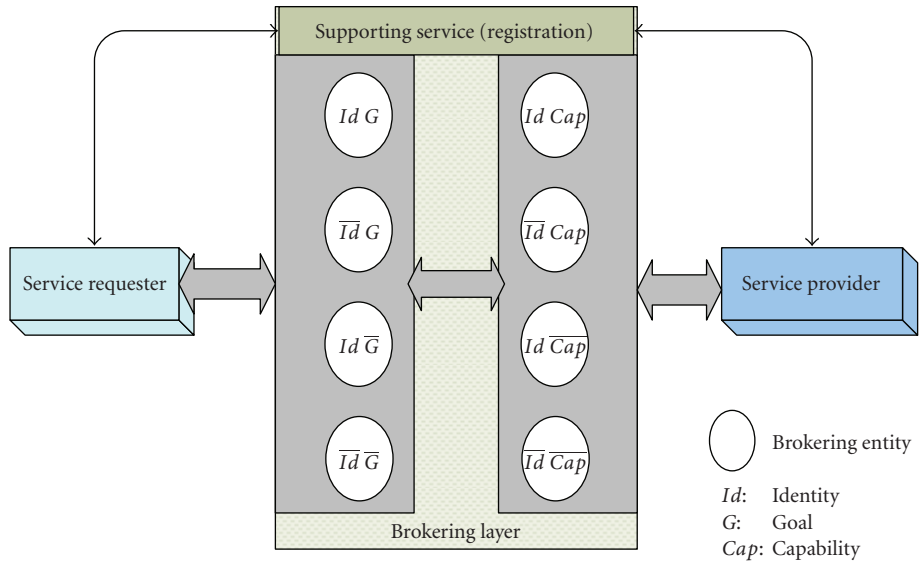


Figure 1: Logical view of the brokering service.

Table 1: The brokering layer interaction categorized by the privacy concern of service requesters.

Case	Privacy attributes		Interaction
	G	*Id*	
1	Revealed	Revealed	(i) Receive service request. (ii) Forward request to broker-provider side. (iii) Deliver result to requester.
2	Hidden	Revealed	(iv) Retrieve service request posted by a requester. (v) Forwards request to broker-provider side. (vi) Store result to be retrieved by requester.
3	Revealed	Hidden	(vii) Postservice request to service repository. (viii) Requester to search repository and request service. (ix) Retrieve a service request that was stored by a requester. (x) Forward request to available and capable providers. (xi) Store result to be retrieved by requester.
4	Hidden	Hidden	(xii) Requester to store service request. (xiii) Retrieve service request that was stored by a requester. (xiv) Forward request to available and capable providers. (xv) Store result to be retrieved by requester.

layer for entities that might play either a requester or a provider role.

5.1. The Requester-Brokering Layer Interaction

5.1.1. Requesters Revealing Identities and Goals. The broker protects the privacy of healthcare personnel, patients, or staff. It assists service requesters to achieve their goals without exposing their identities to the environment. For example, information about the number of patients who have Hepatitis B in a specific city and wanted by a doctor can be assessed by the broker agent without revealing, neither the doctor nor the patients identities. However, agents playing the role of requesters and wanting to benefit from such a service are required to reveal their identities and goals to the relevant broker within the layer. Note that each privacy degree is described in terms of two main interactions: an interaction amongst the various brokers within the brokering layer (intra-interaction) and the interaction between the domain (i.e., a requester or a provider) with the relevant broker that supports a particular privacy degree (inter-interaction).

Intra-Interaction. As shown in Figure 2, the broker might extend the pattern to include interaction with various brokers associated with supporting other privacy degrees of service providers, consequently the broker solicit help and forward request to all available provider-related brokers

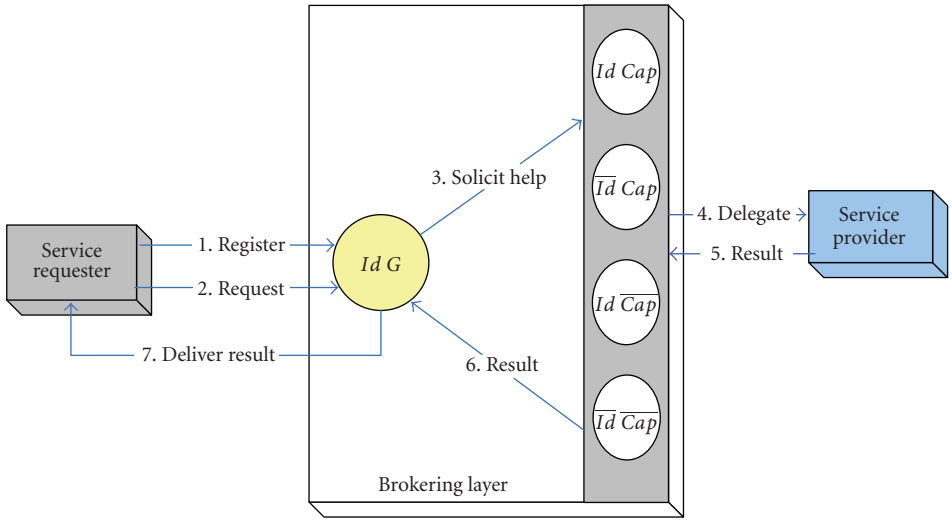


Figure 2: Interaction pattern for requesters revealing privacy attributes.

Table 2: The brokering layer interaction categorized by the privacy concern of service providers.

Case	Privacy attributes		Interaction
	Id	*Cap*	
1	Revealed	Revealed	(i) Search for capable provider. (ii) Forward request. (iii) Negotiate and assign a service request. (iv) Get service result and deliver result.
2	Hidden	Revealed	(v) Postservice request to service repository. (vi) Providers to access service repository. (vii) Providers to evaluate service parameters (viii) Store result. (ix) Brokering layer to retrieve and deliver result.
3	Revealed	Hidden	(x) Forward service request. (xi) Provider to evaluate request. (xii) Brokering layer to receive and deliver result back.
4	Hidden	Hidden	(xiii) Providers to access repository. (xiv) Provider to evaluate request. (xv) Provider to store service result. (xvi) Brokering layer to retrieve and deliver result back.

within the layer incorporating various interaction compositions. Note that for every potential composition, the provider-related brokers receive only a notification of a service request, and accordingly carry on its own interaction pattern to satisfy that request without exaggerating, overstressing, or overemphasizing any incurred rights or privileges (e.g., cost).

Inter-Interaction. The typical interaction pattern for this particular privacy degree comprises that the layer engages in performing the following: (1) accepting and interpreting service requests from pertinent requesters; (2) identifying and contacting a set of available providers, forwarding service requests, and controlling appropriate transactions to fulfill any required service request. These transactions should adhere to agreed appropriate interaction mechanism (e.g., auction, negotiation, etc.); (3) receives result of a service request and delivers it back to the relevant requester.

5.1.2. Requesters Hiding Identities. Requesters such as patients with fatal diseases may wish to access services or seek further assistance without revealing their identities. The brokering service dynamically identifies relevant service providers, and acts on behalf of those requesters to fulfill their goal(s). As shown in Figure 3, requesters will be responsible of checking the availability of the service result, which implies that requesters should be aware of a designated result location. The interaction imposes a significant effort on the performance and efficiency. System performance is clearly dependent on number of parameters, including the

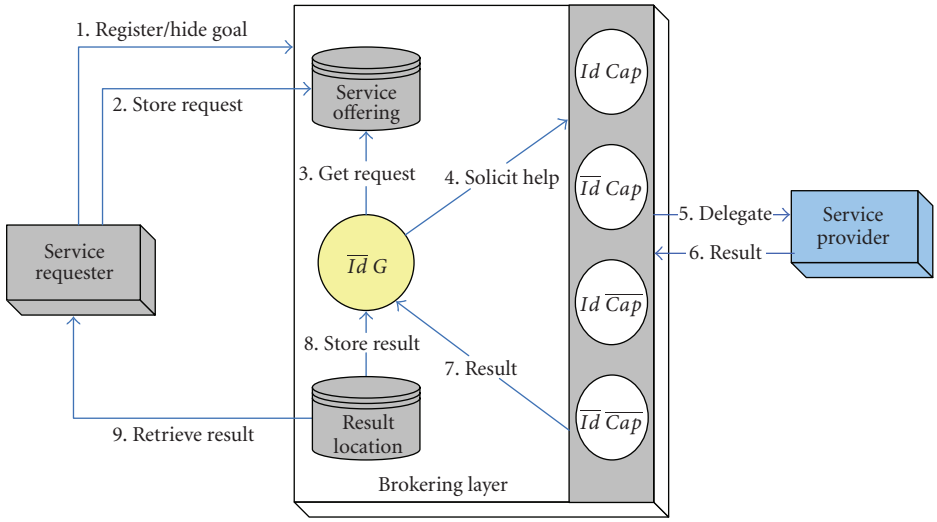


Figure 3: Interaction pattern for requesters hiding identity.

number of providers willing to carry out the request and the time needed by each provider to fulfill that request.

Intra-Interaction. As described in the previous case, the broker might extend its pattern to include an interaction composition with various brokers associated with supporting other privacy degrees for service providers. Upon receiving a service result, the broker stores the result in a dedicated repository (result repository) to be retrieved by the relevant requester.

Inter-Interaction. Requesters may wish to access services or seek further assistance without revealing their identities. The interaction pattern for this particular privacy degree is as follows: (1) requesters are required to store services requests in a predefined service repository along with preferred parameters. (2) As shown in Figure 3, requesters are responsible of checking the availability of the service result and hence retrieve it; this implies that requesters are able to link a service result to their own requests.

5.1.3. Requesters Hiding Goals. There might be certain situations where requesters prefer to hide their goals from the environment; the layer functionality entails the forwarding of every advertised service out to every registered requester with unknown preferences or interests. For example, clinician might benefit from variety of service advertisements regarding new medications, tools, medical equipments, and health-related notifications. The brokering service permits these clinicians to check a service repository for further information or to browse other service offerings that have been previously posted and accordingly determine an appropriate and interested service.

Intra-Interaction. Provider-related brokers representing providers with known capabilities will have the possibility to advertise existing service offerings to the broker which in turn promotes forwarding every received advertisement to the relevant requester. It is to be noted that whenever a requester decides on a particular service offering, the inter-interaction is not restricted only to contacting those who had offered such services, but might extend to all available provider-related brokers supporting other privacy degrees. For example, the same advertised service offering might be achieved by other providers in the environment who had the interest of hiding their own capabilities.

Inter-Interaction. They broker permits healthcare requesters to check a service repository for further information or to browse other service offerings that have been previously posted and accordingly determine an appropriate and interested service as shown in Figure 4. Once a requester selects a particular service advertisement and forwards that request to the broker, then it is the broker responsibility to determine the most suitable service provider that fulfills that request. Upon achieving the requester goal, the broker delivers back the service result to the requester. In an open environment, where many different services providers are in continual increase and with a competitive manner to sell their services, requesters would be flooded by a variety of service advertisements and notifications. Requesters have to determine whether the service advertised to them is of an interest or not. Clearly, this process implies that a significant time is required to assess every single-service notification. The broker sends the notifications along with any related parameters required for providing the service (such as name of the service, cost, and location).

5.1.4. Requesters Hiding Identities and Goals. Requesters would have the possibility to hide their identities and goals from the entire environment; as shown in Figure 5, they have the option either to post their want ads to the layer service repository directly, or might check for any services that

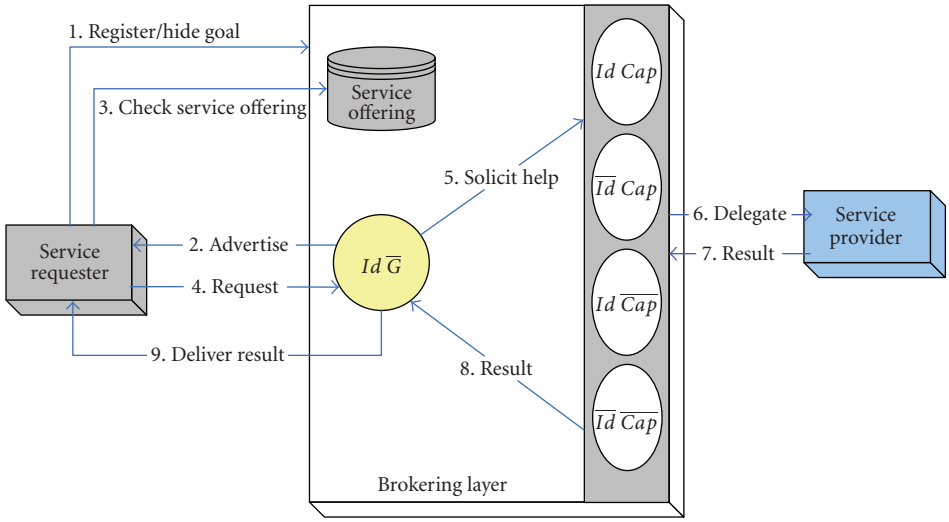


Figure 4: Interaction pattern for requesters hiding goals.

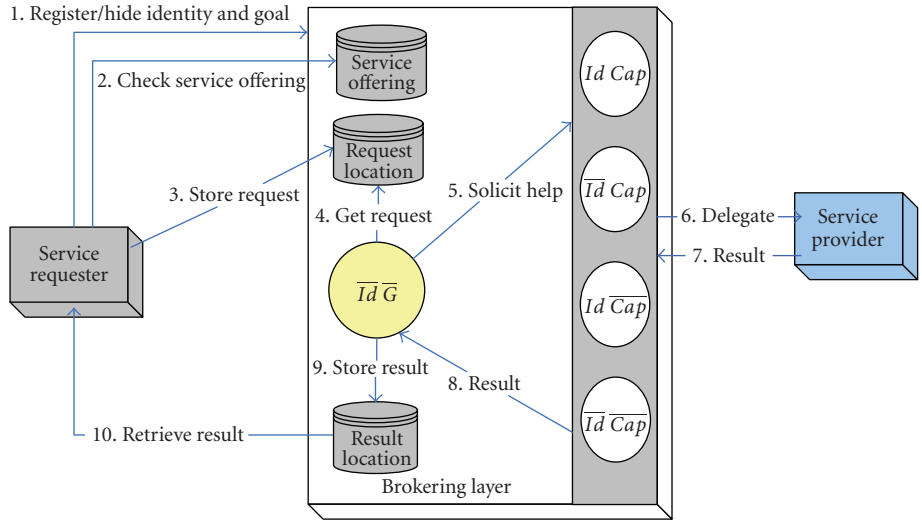


Figure 5: Interaction pattern for requesters hiding privacy attributes.

would be of an interest. For example, patients with narcotic-related problems (such as drug or alcohol addiction) can seek services that provide information about rehabilitation centers, specialized psychiatrists, or programs that will help overcoming a particular critical situation without revealing either their identities nor the desired information.

Inter-Interaction. Requesters will have the option to either post their want ads to a service repository directly, or might check for any service offerings that would be of an interest. In both cases, requesters will be permitted to store their service requests and retrieve services results. The broker identifies and interprets the required requests, and accordingly will determine the applicable provider which is capable of achieving and fulfilling the requester goal. Note that, for this degree of privacy, it is the requester responsibility to check for the availability of the service result, and hence retrieve it.

5.2. The Provider-Brokering Layer Interaction

5.2.1. Providers Revealing Identities and Capabilities. Providers with this degree of privacy will have the ability to register their presence along with the capability of the service they offer. Although providers with this privacy degree are required to reveal their privacy attributes to the relevant broker, the protocol will suppress any other entity from knowing the provider attributes.

Intra-Interaction. The interaction between the broker and other requester-related brokers is accomplished through

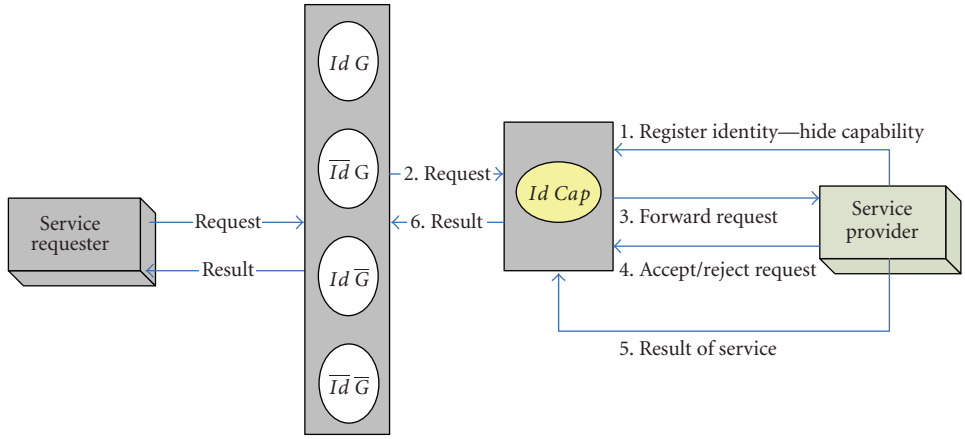


Figure 6: Interaction pattern for providers revealing privacy attributes.

sending and receiving messages related to service proposals, service offerings, and services results.

Inter-Interaction. As shown in Figure 6, a service provider registers itself with the broking service, along with the description of its service capabilities which is stored as an advertisement in a repository maintained by the broker and contains all available service descriptions. Assigning requests to providers with known capabilities and identities can be based on either broadcasting or focusing, however, the interaction is neither restricted to specific service providers nor committed to a fixed number of them. This ability is particularly useful in which a brokering agent acts in a dynamic environment in which entities may continually enter and leave the society unpredictably. For every received service request, the broker matches the most applicable providers that are appropriate to fulfill that request, and thus maintains a pertinent queue that contains the capable providers along with their identities.

5.2.2. Providers Hiding Identities. Healthcare providers can have the option to hide their identities from the environment and advertise their service offerings to the relevant brokering agent. Protection for the core identity prevents service abuses that impact availability of service and hence improving the ability to consistently deliver reliable access. Since the service capabilities are known to the broker, service requests that are believed to be fulfilled by such providers will be posted to a dedicated repository for which providers will have the possibility to browse such requests and select whichever of an interest.

Intra-Interaction. The broker interacts with other entities in the layer to engage in receiving and sending messages related to service requests and offerings. The broker task includes (1) receiving service requests; (2) determining whether these requests are within the provider capabilities; (3) storing service requests to be browsed by authorized registered providers (providers hiding identities); (4) retrieving and delivering back service result. A broker supporting this privacy case will have the ability to advertise registered provider capabilities, and hence engage in various interaction patterns of available requester-related brokers.

Inter-Interaction. A provider can participate in any interaction mechanism and may respond to call-for-proposal requests by proposing service offerings that are stored in a queue-structured repository. Upon assigning and delegating a service request to a provider with this degree of privacy, it is the provider responsibility to store pertinent service result to be retrieved by the broker, and thus delivered to the proper destination as shown in Figure 7.

5.2.3. Providers Hiding Capabilities. The brokering services allow providers that do not wish to reveal their own capabilities to participate in fulfilling a service request. After receiving a request, the brokering interaction protocol exemplifies the forming out of requests to every registered provider with unknown capability. It is noteworthy that, for every advertised request, providers have to determine whether the request is within their capabilities and/or of an interest. Clearly, such an interaction implies that a considerable elapsed time will be spent on evaluating every single request. Therefore (under the assumption of an open dynamic environment), providers would be deluged by a variety of service requests, which significantly impact performance and efficiency. Figure 8 shows the interaction pattern.

Intra-Interaction. The broker interacts with other entities in the layer to engage in receiving and sending messages related to service requests and offerings. The broker task includes (1) receiving service requests from requester-related brokers; (2) receiving service proposals; (3) delivering back service result.

Inter-Interaction. After receiving a service request, the broker sends out requests in the form of broadcasting to every registered provider with unknown capabilities. Figure 8 shows the interaction pattern. Once a provider selects a particular service request, it forwards a service proposal to

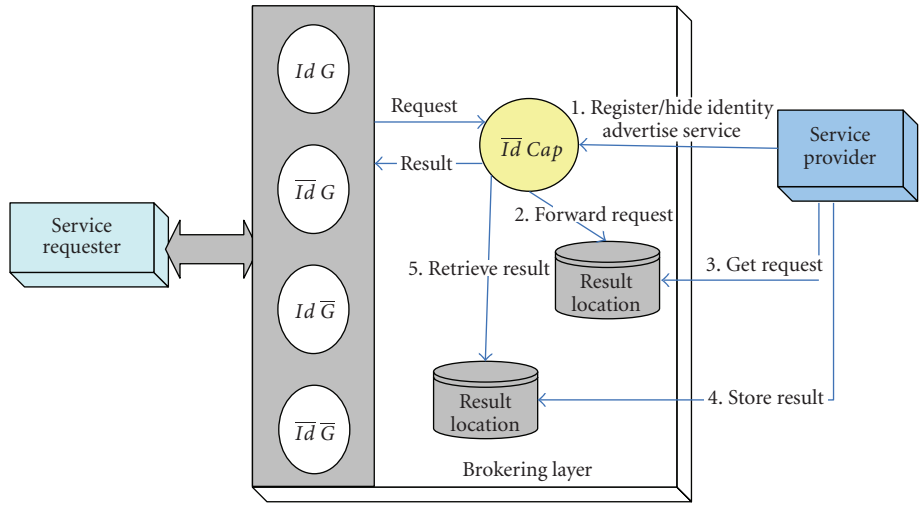


Figure 7: Interaction pattern for providers hiding identity.

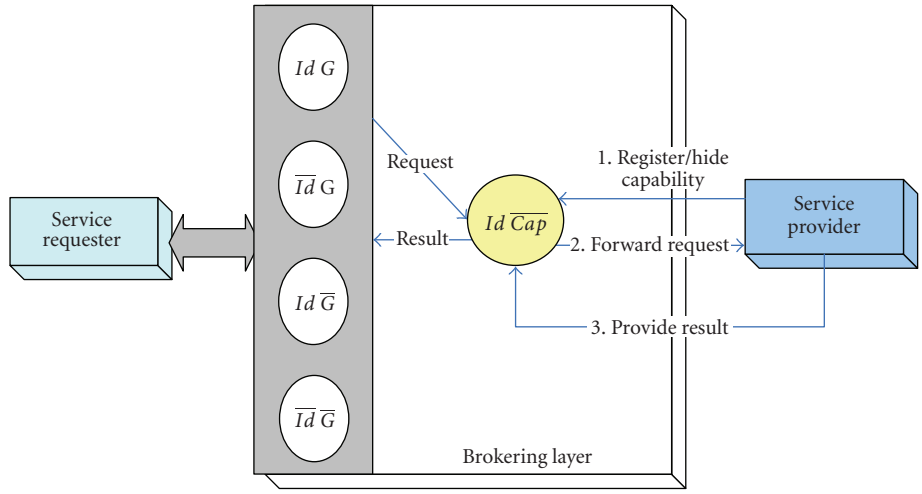


Figure 8: Interaction pattern for provider hiding capability.

the broker who controls the remaining transaction according the appropriate negotiation mechanisms similar to what has been described in the former patterns.

5.2.4. Providers Hiding Identities and Capabilities. Providers will have the ability to browse a special request repository and consequently determine the relevant requests that might be of an interest and within their capabilities. As shown in Figure 9, the broker-provider side agent responds back with the service result (a result location within the layer has to be identified to the provider upon registration within the brokering layer).

Intra-Interaction. The broker intra-interaction comprises the following: (1) receiving service requests from requester-related brokers; (2) storing service requests; (3) accessing and evaluating service proposals; (4) retrieving and delivering back service result.

Inter-Interaction. In this protocol, the brokering functionality is mainly seen as a directory service, in which the broker maintains a repository of service requests along with any required preferences. Providers will have the ability to browse this repository to determine applicable relevant requests that might be fulfilled. As shown in Figure 9, providers with this degree of privacy have to take in consideration linking the result of the service to the request.

6. Design and Implementation

6.1. Modelling Healthcare-Distributed Systems. It is clear that the development of coordination solutions in open and distributed healthcare environments requires a new design

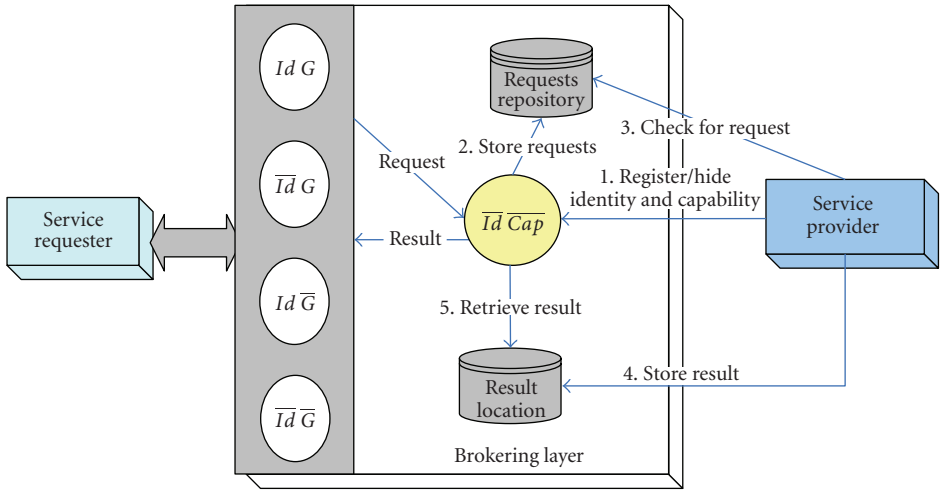


Figure 9: Interaction pattern for provider hiding privacy attributes.

paradigm, improved integration architectures and services. A cooperative distributed systems (CDSs) approach is an ideal and appropriate design paradigm which allows the various healthcare entities to exercise some degree of authority in sharing their information and capabilities.

The architecture must describe the organization and the interconnection among the software entities. In this architecture, the environment can be envisioned as a cooperative distributed system (CDS) comprised of a collection of economically motivated software agents that interact competitively or cooperatively, find and process information, and disseminate it to humans and other agents. It also enables common services that facilitate the coordination and the cooperation activities amongst various domain entities and support ad hoc and automated configurations.

In our proposed model, a CDS is conceptualized as a dynamic community of agent and nonagent entities that contribute with different services. Based on the above view, an agent might play different roles and be able to coordinate cooperatively or competitively with other agents, including humans. Therefore, healthcare CDS entities are mapped as follows.

(i) *Service requester*: is a domain specific entity that can interact with the environment and request services.

(ii) *Service provider*: a domain entity that provide application-specific services.

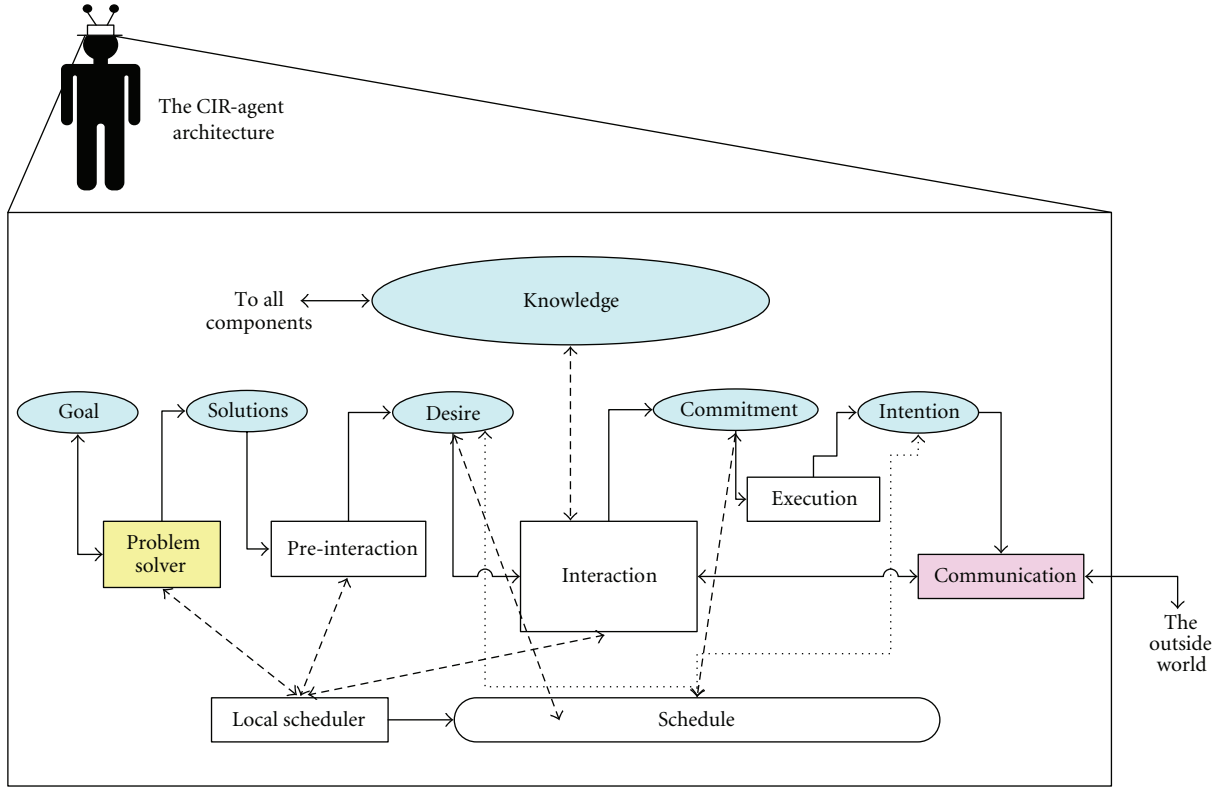
(iii) *Brokering entity*: is an agent that provides common coordination services, and facilities for the generic cooperative distributed systems environment.

6.2. The Coordinated Intelligent Rational Agent (CIR-Agent) Model. The representative agents of domain and brokering entities within the context of healthcare-based CDS are built on the foundation of CIR-agent architecture with focuses on utilizing the model to capture the participants' individual behavior toward achieving a desirable goal while maintaining a required privacy degree.

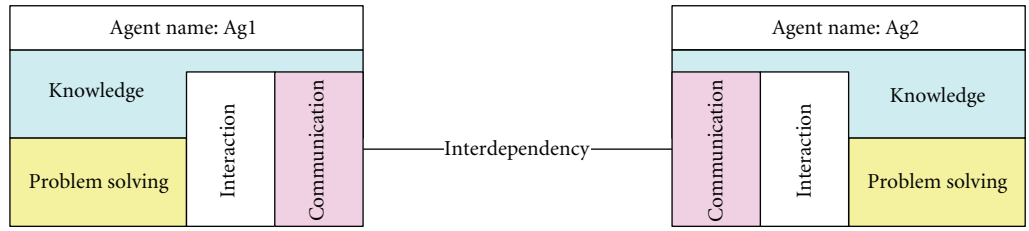
The CIR-agent is an individual collection of primitive components that provide a focused and cohesive set of capabilities. The basic components include problem-solving, interaction, and communication components, as shown in Figure 10(b). A particular arrangement (or interconnection) of components is required to constitute an agent. This arrangement reflects the pattern of the agent mental state as related to its reasoning about achieving a goal. However, no specific assumptions need to be made on the detailed design of the agent components. Therefore, the internal structure of the components can be designed and implemented using object oriented or another technology, provided that the developer conceptualizes the specified architecture of the agent as described in Figure 10.

Basically, each agent consists of knowledge and capability components. Each of which is tailored according to the agent specific role. The agent knowledge contains information about the environment and the expected world. The knowledge includes the agent self-model, other agents' model, goals that need to be satisfied, possible solutions generated to satisfy each goal, and the local history of the world that consists of all possible local views for an agent at any given time. The agent knowledge also includes the agent desires, commitments, and intentions toward achieving each goal. The capability package includes the reasoning component; the domain actions component which contains the possible set of domain actions that when executed, the state of the world will be changed; the communication component where the agent sends and receives messages to and from other agents and the outside world.

The problem solver component represents the particular role of the agent and provides the agent with the capability of reasoning about its knowledge to generate appropriate solutions directed to satisfy its goal. During the interaction processes, the agents engage with each other while resolving problems that are related to different types of interdependencies. The coordination mechanisms are meant to reduce



(a) Detailed architecture of CIR agent



(b) Logical architecture of CIR agent

Figure 10: The CIR agent architecture.

and resolve the problems associated with interdependencies. Interdependencies are goal-relevant interrelationships between actions performed by various agents.

As argued in [30], the agent interaction module identifies the type of interdependencies that may exist in a particular domain. Consequently, agents select an appropriate interaction device that is suitable to resolve a particular interdependency. (Interaction device is an agent component by which it interacts with the other elements of the environment through a communication device. A device is a piece or a component with software characteristics that is designed to service a special purpose or perform a special function). These devices are categorized as follows.

(i) Contract based includes the assignment device.

(ii) Negotiation based includes resource scheduling, conflict resolution, synchronization, and redundancy avoidance devices.

Within the context of brokering, the interdependency problem is classified as capability interdependency, and the interaction device is the "assignment". The basic characteristics of the assignment device are problem specifications, evaluation parameters, and the subprocesses. The problem specifications might include, for example, the request, the desired satisfying time, and the expiration time. A collection of basic components comprises the structure of the agent model and represents its capabilities. The agents architectures are based on the CIR-agent model as shown in Figure 11. A brokering session mainly recognizes two types of agents, namely, domain agent (requester or provider) and brokering agent (ReqBroker or ProvBroker). The architecture of each agent type is described in detail below.

6.2.1. The Domain Agent: Service Providers and Requesters. Service providers and requesters are modeled as domain

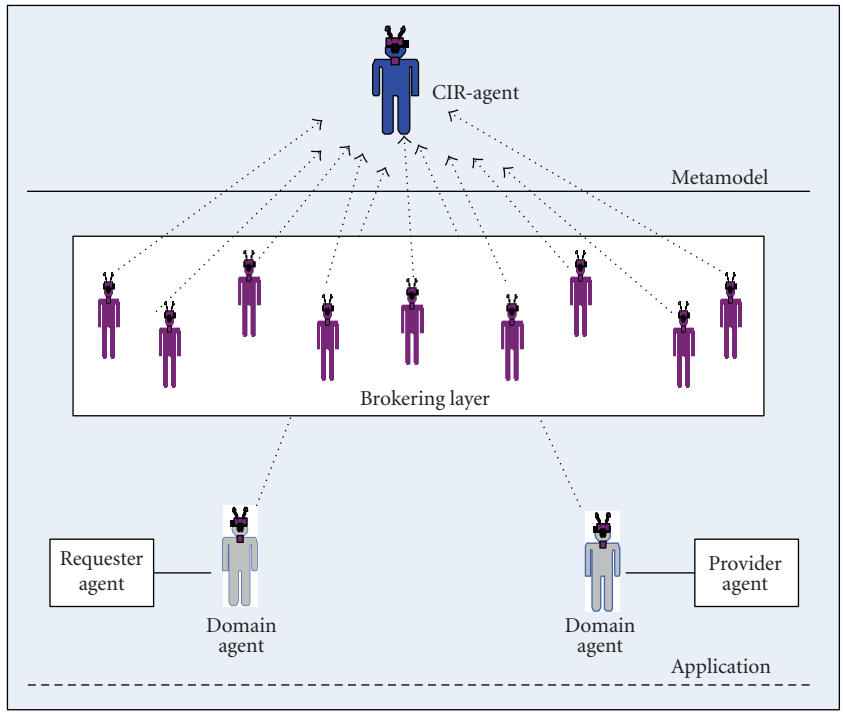


Figure 11: The overall system model.

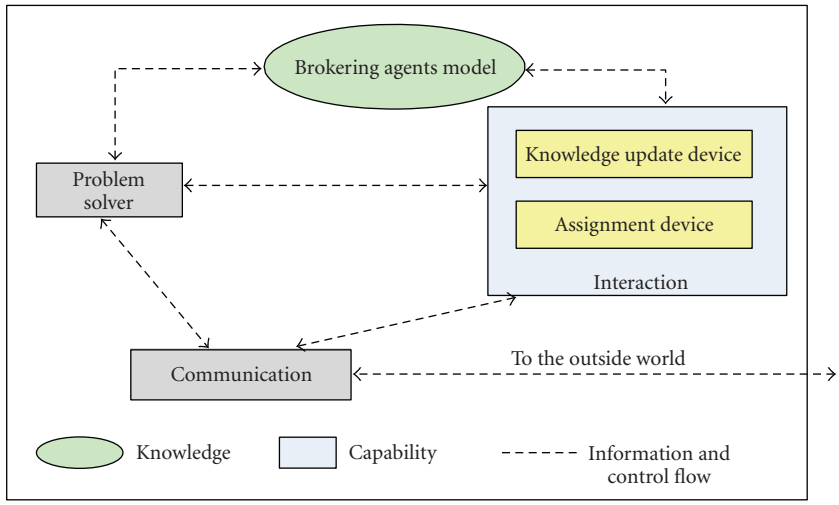


Figure 12: The domain agent architecture.

agents as shown in Figure 12. The requester agent can participate with various privacy degrees and request services from the brokering layer. A requester delegates the service request(s) to the relevant brokering agent according to the interaction protocol of the selected privacy degree. The domain agent possesses knowledge and capability. The knowledge includes the model of the brokering agents in terms of the supported privacy degree, self-model, and the local history. The capability is categorized into three components: reasoning that includes problem-solving and coordination, communication, and a set of domain actions.

A domain agent playing the role of a service provider can select the appropriate privacy degree, and thus participate in providing the capability that meets the needs of another domain entity. The problem solver of the domain agent hiding any of the privacy attributes encompasses the accessing of different storage repositories. For example, the problem solver of a requester includes functionalities related to formulating service requests, checks for available service offerings, and accesses various storage repositories to store requests or to retrieve service results. On the other hand, the problem solver of a provider hiding its identity and capability

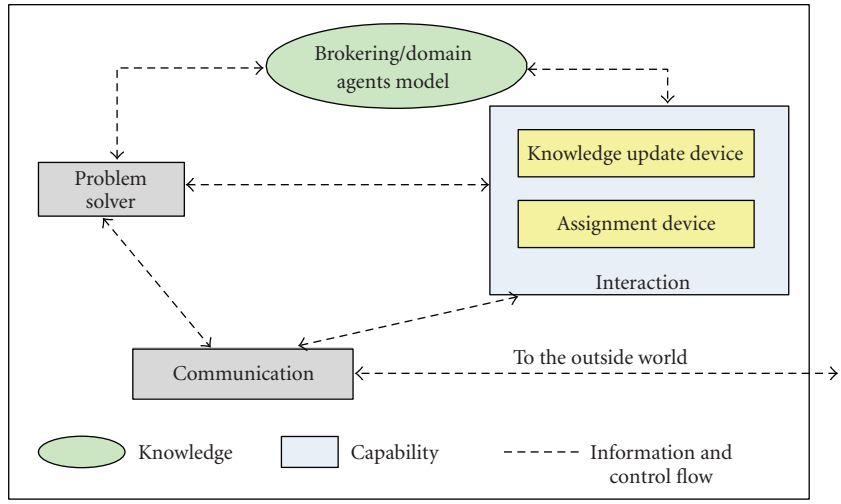


Figure 13: The brokering agent architecture.

attributes consists of modules related to accessing storage repositories to check for stored service requests that might be fulfilled and hence participating in storing service proposals and service results.

The coordination component of a requester comprises the interaction device which entails soliciting service from the relevant ReqBroker agent. The interaction device of the provider agent manages the coordination activities which involve proposing services to specific CFP messages and engage in bidding processes.

6.2.2. The Brokering Agents: ReqBrokers and ProvBrokers. A brokering agent is composed of two components, namely, the knowledge and capability. The knowledge component contains the information in the agent memory about the environment and the expected world. As shown in Figure 13, this includes the agent self-model, models of the domain agents in terms of their roles (requester/provider) and/or capabilities, and the local history of the world. The knowledge includes all possible local views for an agent at any given time (such as the knowledge of physical repositories, available services requests, services offerings, and service results).

6.2.3. Implementation Example: Agent-Oriented Privacy Brokering for Healthcare CDS. In this section, we show an example of our proposed model applied to healthcare environments to support information-gathering capabilities. We describe the implementation of one pattern associated with an information requester hiding identities and goals and with three information providers; one is revealing privacy attributes, the second hiding its identity, while third is hiding its own privacy attributes (identities and capabilities). The broker agent (called ReqBroker henceforth) protects the privacy of requesters, understands the preferences, routes requests, and replies appropriately. All the inter-interactions utilize the FIPA Contract Net Protocol [13] as a negotiation mechanism. Consider an online three information providers, E-VirtualMedInfo Inc., E-VirtualDiagnosis Inc., and Future-DocAssistants Inc. (names are fictitious), each is represented by an agent.

The three providers offer medical-related information, healthcare guidelines, and clinical diagnosis procedures that can be supplied to various medical students, clinicians, staff, doctors, and physicians in various formats (online delivery, hard copies, or access to online medical repositories). All the three companies decided to register and subscribe to the brokering service and make use of the various privacy degrees. *E-VirtualMedInfo* registered with the brokering service while revealing it privacy attributes, *E-VirtualDiagnosis* comprises diagnosis capabilities jointly derived retired medical doctors and had selected hiding its identity, whereas *FutureDocAssistants*, a company that can also provide various online samples of medical exams and virtual evaluation assessments for different medical specialties, decided to hide both the identity and the capabilities. Upon registration, a dedicated brokering agent (ProvBroker) will be assigned to each company.

Alice, a four-year medical student, is conducting a research on the most top fatal diseases in Canada, the mortality and death rates of each and the possible diagnosis and prevention procedures that would help a trainee-student in examining and diagnosing patients with such diseases. Deciding to hide her own identity, Alice anonymously requests this information by posting the required information in special repository dedicated to such privacy degree.

After storing the request, Alice's assigned broker (ReqBroker) interacts with various ProvBrokers associated with supporting other privacy degrees of service providers (including the three mentioned companies) and consequently (acts as a manager) issues, and announces a call-for-proposals (CFPs) to those ProvBrokers (act as potential contractors) informing them of Alice's request specifications (note that Alice's identity is anonymous to each participant including its own supporting ReqBroker).

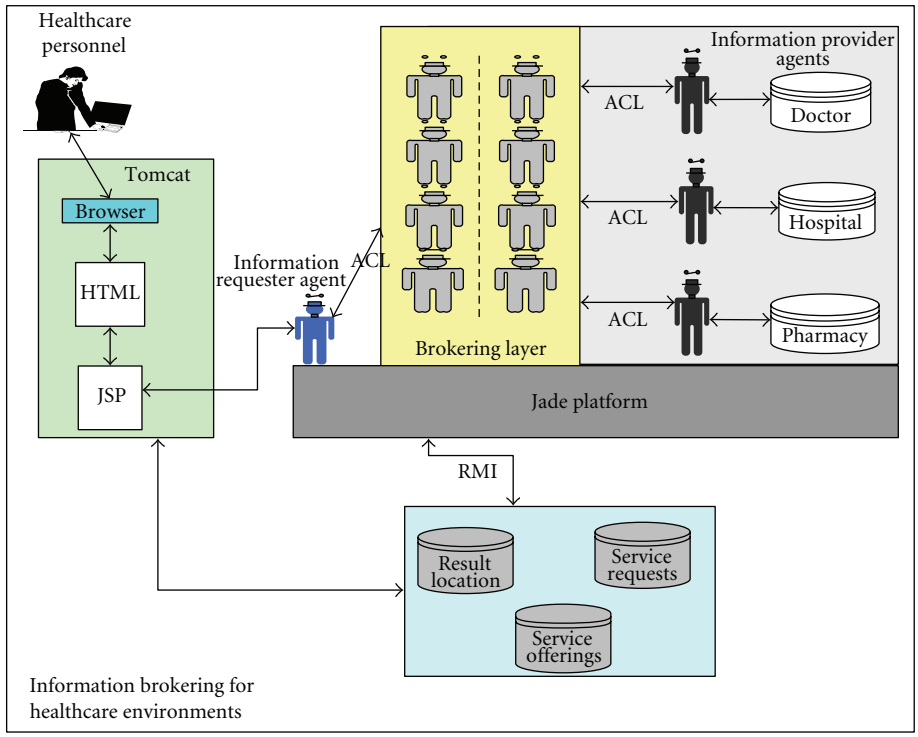


Figure 14: Privacy based brokering prototype for information gathering in healthcare.

The announcement includes task abstraction, a brief description of the required information; bid specification, a description of the expected format of the information; expiration time, a specified time interval during which the required information is valid.

Each ProvBroker working on behalf of each company contacts the registered company agent and sends the request. Note that for the *FutureDocAssistants* company, the request is dispatched in special dedicate storing repository allowing its own agent to browse this repository and retrieve the request (if interested).

Every company (through its representing agent) determines the evaluation parameters (such as information quality, expiration time, and cost) and accordingly submits a bid along with the offer parameters (such as quality, cost, and availability). The *E-VirtualMedInfo* and *E-VirtualDiagnosis* agents will send the bids directly to their assigned ProvBrokers, while the *FutureDocAssistants* agent stores the bid in a repository that will be retrieved by the relevant ProvBroker.

Alice's dedicated ReqBroker receives those bids from every ProvBroker and carries on the evaluation process and accordingly determines the most bid (or bids) that fulfills Alice's request for all the interested, and sends back an acceptance-proposal message to the potential companies (winners) and a rejection-message to the bids that do not meet the evaluation parameters. After receiving the information that Alice was requesting, the ReqBroker stores it in a special repository for which she has a valid access to retrieve it without having to reveal her own identity or being exposed to the identity and the capabilities of the three companies which had participated in fulfilling her request.

A web-based prototype of the proposed system has been implemented using Jade [31], an FIPA [32] compliant, and Java Web Services Development Pack (JWSDP) platform to support and provide information-gathering capabilities to different participants in healthcare environments, where the accessibility of private information is a desirable feature to various categories of the healthcare personnel, patients, and clinicians.

The proposed architecture has been implemented using coordinated intelligent, rational agent (CIR-agent). As shown in Figure 14, three relational databases represent various medical data for three distributed locations, each being managed by a dedicated agent that can play both roles of an information requester as well as a provider.

Upon the required privacy degree and the role desired, the A Web interface is available for healthcare participants to select and register their desired privacy degree along with any information capability they might posses (medical data, patient diagnosis and treatment reports, pharmaceutical data reports, etc.). Based on the privacy degree required by the both the requester and information provider, a dedicated broker agent within the brokering layer will handle all the interaction required to fulfill an information request.

7. Discussion and Conclusion

Current advances in nowadays technologies coupled with the rapidly evolving healthcare paradigms allow us to foresee novel applications and services for improving the quality of daily life health activities. The increasing demand and

dependency on information in healthcare organizations has brought the issues of privacy to every aspect of the healthcare environments. It is expected and with no doubt that medical data such as genome information, medical records, and other critical personal information must be respected and treated with a great concern. As awareness of the threats that organizations face becomes more well understood, the need for additional privacy specifications for open, distributed, and heterogeneous systems grows clear. Tremendous efforts have been devoted to privacy and security issues in distributed systems for the last few decades to find technological means of guaranteeing privacy by employing state-of-the-art encryption and anonymization technology. The proposed architecture provides feasible solution to privacy protection in open environments, and presents myriad of additional privacy and security opportunities without negative impact to the utilization of these services.

Architecturally, the proposed model is viewed as a layer of services, where different roles can be played by the various entities (requesters, brokers, and providers). While existing approaches provide traditional information brokering by incorporating agent-based solutions to make healthcare information more accessible to individuals, the proposed architecture classifies the brokering role into several subroles based on the attributes designated to describe the privacy-desired degree of both the information provider and the information requester. Each role is modeled as an agent with a specific architecture and interaction protocol that are appropriate to support a required privacy degree.

Within the layer, two sets of brokering entities are available to service requesters and providers. The first set handles interactions with requesters according to the desired privacy degree that is appropriate to their preferences, while the other set supports privacy degrees required by service providers. A brokering pattern is realized by the different roles played by the domain entities and their corresponding brokering agent. A complete brokering scenario is accomplished by performing different levels of interaction, namely, (1) requester-to-broker interaction, (2) broker-to-broker interaction, and (3) broker-to-provider interaction. Different combinations within the layer can take place to support the interbrokering interactions. The proposed layered architecture provides an appropriate separation of responsibilities, allowing developers and programmers to focus on modeling solutions and solving their particular application problems in a manner and semantics most suitable to the local perspective. Agent technology has been viewed as one of the key technologies for supporting information brokering in heterogeneous open environments. The use of agent technology provides high degree of decentralization of capabilities, which is the key to system scalability and extensibility.

Another important aspect of the model is that it treats the privacy as a design issue that has to be taken into consideration in developing healthcare information brokering systems. In healthcare environments, the proposed model provides feasible solution to the problem of information overload and privacy concerns. It enables transparent integration amongst different participants of healthcare CDS, and provides querying ability and coordination solutions that enhance the overall connectivity of distributed, autonomous, and possibly heterogeneous information sources (databases) of different healthcare sectors. It can efficiently govern different types of health-oriented information and critical medical data such as genetic, HIV, mental health, and pharmacy records from not distributed, disseminated, or abused. Based on the level and the amount of information that can be released, patients, clinicians, service providers, and medical staff members can securely translate their privacy policies to an applicable-related privacy case in the proposed model.

The proposed approach is innovative in the sense that it treats the privacy as a design issue for information brokering systems, and it supports ad hoc and automated configurations among distributed, possibly autonomous, and heterogeneous entities with various degrees of privacy requirements. The multilayer architecture minimizes the architecture complexity encountered in direct-interaction architectures (where interactions between agents often take more complex processes for encompassing series of message exchange and forming a single point of failure), and makes it less vulnerable to failures. The proposed layered architecture provides an appropriate separation of responsibilities, letting developers and programmers focus on solving their particular application problems in a manner and semantics most suitable to the local perspective.

References

[1] The Association of American Physicians and Surgeons, "New Poll: Doctors Lie To Protect Patient Privacy," Poll Survey, http://www.aapsonline.org/press/nrnewpoll.htm.

[2] O. Baudron and J. Stern, "Non-interactive private auctions," in *Proceedings of the 5th International Conference on Financial Cryptography (FC '01)*, vol. 2339 of *Lecture Notes in Computer Science*, pp. 364–378, Springer, Grand Cayman, British West Indies, February 2001.

[3] Y. Lindell and B. Pinkas, "Privacy preserving data mining," in *Proceedings of the 20th Annual International Cryptology Conference on Advances in Cryptology (CRYPTO '00)*, vol. 1880 of *Lecture Notes in Computer Science*, pp. 36–54, Springer, Santa Barbara, Calif, USA, August 2000.

[4] R. B. Bayardo, W. Bohrer, R. Brice, et al., "InfoSleuth: agent-based semantic integration of information in open and dynamic environments," MCC Technical Report MCC-INSL-0088-96, Microelectronics and Computer Technology Corporation, Austin, Tex, USA, October 1996.

[5] K. Decker, V. Lesser, M. V. Nagendra Prasad, and T. Wagner, "MACRON: an architecture for multi-agent cooperative information gathering," in *Proceedings of the Workshop on Conference on Information and Knowledge Management (CIKM '95)*, Baltimore, Md, USA, December 1995.

[6] M. R. Genesereth, A. M. Keller, and O. M. Duschka, "Infomaster: an information integration system," in *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '97)*, pp. 539–542, Tucson, Ariz, USA, May 1997.

[7] K. Decker, K. Sycara, and M. Williamson, "Middle agents for the internet," in *Proceedings of 15th International Joint Conference on Artificial Intelligence (IJCAI '97)*, pp. 578–583, Nagoya, Japan, January 1997.

[8] D. Kuokka and L. Harada, "On using KQML for matchmaking," in *Proceedings of the 1st International Conference on Multi-Agent Systems (ICMAS '95)*, pp. 239–245, AAAI Press, San Francisco, Calif, USA, June 1995.

[9] E. Motta, J. Domingue, L. Cabral, and M. Gaspari, "IRS-II: a framework and infrastructure for semantic web services," in *Proceedings of the International Semantic Web Conference (ISWC '03)*, vol. 2870 of *Lecture Notes in Computer Science*, pp. 306–318, Sanibel Island, Fla, USA, October 2003.

[10] M. Paolucci, J. Soudry, N. Srinivasan, and K. Sycara, "A broker for OWL-S web services," in *Proceedings of AAAI Spring Symposium on Semantic Web Services*, vol. 6, pp. 92–99, Stanford, Calif, USA, March 2004.

[11] L. Li and I. Horrocks, "A software framework for matchmaking based on semantic web technology," in *Proceedings of the 12th International Conference on World Wide Web (WWW '03)*, pp. 331–339, Budapest, Hungary, May 2003.

[12] The OWL Services, "Semantic Markup for Web Services (OWL-S)," December 2008, http://www.daml.org/services/owl-s/1.0.

[13] "RDF: The Resource Description Framework," December 2008, http://www.w3.org/RDF/.

[14] J. Brook and D. Fellows, "An architecture for distributed brokering on the grid," in *Proceedings of the 11th International Euro-Par Parallel Processing*, Lisbon, Portugal, August-September 2005.

[15] V. Shankaraman, V. Amorosiadou, and B. Robinson, "Agents in medical informatics," in *Proceedings of the 18th IASTED International Conference on Applied Informatics*, Innsbruck, Austria, February 2000.

[16] A. Moreno, A. Valls, and J. Bocio, "Management of hospital teams for organ transplants using multi-agent systems," in *Proceedings of the 8th Conference on Artificial Intelligence in Medicine in Europe (AIME '01)*, vol. 2101 of *Lecture Notes in Computer Science*, pp. 374–383, Springer, Cascais, Portugal, July 2001.

[17] A. Aldea, B. López, A. Moreno, D. Riaño, and A. Valls, "A multi-agent systems for organ transplant co-ordination," in *Proceedings of the 8th Conference on Artificial Intelligence in Medicine in Europe (AIME '01)*, vol. 2101 of *Lecture Notes in Computer Science*, pp. 413–416, Springer, Cascais, Portugal, July 2001.

[18] A. Moreno and D. Isern, "Accessing distributed health-care services through smart agents," in *Proceedings of the 4th IEEE International Workshop on Enterprise Networking and Computing in the Health Care Industry (HealthCom '02)*, pp. 34–41, Nancy, France, June 2002.

[19] R. Clarke, "Identification, Anonymity and Pseudonymity in Consumer Transactions: A Vital Systems Design and Public Policy Issue," December 2008, http://www.anu.edu.au/people/Roger.Clarke/DV/AnonPsPol.html.

[20] G. Yee, L. Korba, and R. Song, "Ensuring privacy for e-health services," in *Proceedings of the 1st International Conference on Availability, Reliability and Security (ARES '06)*, pp. 1–8, Vienna, Austria, April 2006.

[21] L. M. Camarinha-Matos and H. Afsarmanesh, "Virtual communities and elderly support," in *Advances in Automation, Multimedia and Video Systems, and Modern Computer Science*, pp. 279–284, WSES Press, New York, NY, USA, September 2001.

[22] B. G. Silverman, C. Andonyadis, and A. Morales, "Web-based health care agents; the case of reminders and todos, too (R2Do2)," *Artificial Intelligence in Medicine*, vol. 14, no. 3, pp. 295–316, 1998.

[23] A. Beresford and R. Stajano, "Location privacy in pervasive computing," *IEEE Pervasive Computing*, vol. 2, no. 1, pp. 46–55, 2003.

[24] Anonymizer Tool, December 2008, http://www.anonymizer.com.

[25] M. K. Reiter and A. D. Rubin, "Crowds: anonymity for Web transactions," *ACM Transactions on Information and System Security*, vol. 1, no. 1, pp. 66–92, 1998.

[26] D. Goldschlag, M. Reed, and P. Syverson, "Onion routing," *Communications of the ACM*, vol. 42, no. 2, pp. 39–41, 1999.

[27] H. Cheng, D. Zhang, and J. Tan, "Protection of privacy in pervasive computing environments," in *Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC '05)*, vol. 2, pp. 242–247, Las Vegas, Nev, USA, April 2005.

[28] "Health Insurance Portability and Accountability Act (HIPAA)," http://www.intellimark-it.com/privacysecurity/hipaa.asp.

[29] "Initiative for Privacy Standardization in Europe (IPSE)," http://www.hi-europe.info/files/2002/9963.htm.

[30] H. Ghenniwa and M. Kamel, "Interaction devices for coordinating cooperative distributed systems," *Intelligent Automation and Soft Computing*, vol. 6, no. 3, pp. 173–184, 2000.

[31] "Java Agent Development Framework: Jade," December 2008, http://jade.cselt.it/.

[32] "FIPA Agent Software Integration Specification," December 2008, http://www.fipa.org/specs/fipa00079/.

Hindawi Publishing Corporation
International Journal of Telemedicine and Applications
Volume 2009, Article ID 279091, 13 pages
doi:10.1155/2009/279091

Research Article

Enhancing E-Health Information Systems with Agent Technology

Minh Tuan Nguyen, Patrik Fuhrer, and Jacques Pasquier-Rocha

Department of Computer Science, University of Fribourg, 1700 Fribourg, Switzerland

Correspondence should be addressed to Patrik Fuhrer, patrik.fuhrer@unifr.ch

Received 30 April 2008; Accepted 1 September 2008

Recommended by Yang Xiao

Agent Technology is an emerging and promising research area in software technology, which increasingly contributes to the development of value-added information systems for large healthcare organizations. Through the MediMAS prototype, resulting from a case study conducted at a local Swiss hospital, this paper aims at presenting the advantages of reinforcing such a complex E-health man-machine information organization with software agents. The latter will work on behalf of human agents, taking care of routine tasks, and thus increasing the speed, the systematic, and ultimately the reliability of the information exchanges. We further claim that the modeling of the software agent layer can be methodically derived from the actual "classical" laboratory organization and practices, as well as seamlessly integrated with the existing information system.

1. Introduction

The business of today's complex organizations such as hospitals in a healthcare network relies on sophisticated information systems which often inherit many weaknesses from the past. For instance, due to its lack of flexibility, a legacy information system cannot integrate the ever-increasing requirements in order to assist the users or to free them from many routine tasks. (A legacy information system represents a massive, long-term investment in the past [1], with poor system quality, design, and architecture. It is costly to adapt to rapidly changing business requirements.) This weakness of legacy information systems is one of many aspects of the "automation gap." Another major weakness relates to the increasing physical mobility of users. Many legacy information systems are designed for users working at fixed client workstations in fixed offices. They do not take into account recent advances in mobile technology such as PDAs, mobile phones, and smartphones. In many legacy information systems, the information flow still requires human interaction between actors either face-to-face or through the plain old telephone communication system to get things done (information delivery, alert sending, people search, feedback, etc.). Automation gap, lack of mobility, and direct human interaction result in an inefficient information flow and data processing:

(i) nonautomated information search and retrieval are time-consuming;

(ii) errors may occur in data transmission by humans;

(iii) users must be physically present at either end of the communication link to successfully establish a conversation (i.e., only synchronous interaction);

(iv) the lack of a systematic activity log makes it difficult to determine the responsibilities of actors when problems or errors occur during a business process.

This research aims at applying a systematic agent technology approach to overcome these weaknesses. The design of a software agent layer on top of a legacy information system offers many advantages to users:

(i) it adds interesting properties to the information system: ubiquitousness, intelligence, scalability, systematic management, logging of the information flows, and so forth;

(ii) it helps humans to interact efficiently among themselves and with the information system. Indeed, human effort and time can be saved by transferring routine tasks from humans to software.

After this first introductory part, Section 2 provides background information on software agents, agents platforms, and development methodologies in general.

Section 3 presents a case study conducted at the HCF Laboratory (HCF is the French acronym for Hospital of the state of Fibourg, Switzerland) . This section is further divided as follows:

(i) the mission and the information system of the HCF Laboratory are presented;

(ii) the weaknesses and potential problems of the current information system are identified;

(iii) finally, a software agent-based solution to enhance the system is proposed.

Section 4 focuses on the medical multiagent system (MediMAS) prototype, which represents our first implementation of the proposed agent-based solution. It simulates an end user's (lab personnel, physician) point of view by considering software agents as personal assistants and by showing them in action.

Section 5 shows how it was possible to define the requirements and to sketch the architecture of the prototype using a well-defined and systematic approach, and this section also briefly describes its main components.

Finally, Section 6 concludes this paper by summarizing the main achievements of our work and by discussing some extensions and improvements planned for the future.

2. Background

It is out of the scope of this paper to offer full background information on software agents and their related technologies. Therefore, the three next subsections only provide a short introduction to the domain and refer the interested reader to the abundant literature for further details.

2.1. What is an Agent? The term "agent" appears in a wide spectrum of research areas such as economics, physics, biology, mathematics, artificial intelligence, and software engineering. Therefore, a unified notion of agent is difficult to extract from the research literature. In this section, we do not aim to coin a new definition, but to highlight the fundamental properties of agents from two published definitions.

Definition 1. An autonomous agent is a system situated within and a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to effect what it senses in the future [2].

Definition 2. An agent is a small, autonomous, or semi-autonomous software program that performs a set of specialized functions to meet a specific set of goals, and then provides its results to a customer (e.g., human end-user, another program) in a format readily acceptable by that customer [3].

The first definition proposes the most general notion of agent which may be a person, a robot, a piece of software, and so forth. The second definition focuses on agents in the software domain which is of interest to us. Both definitions exhibit the following basic properties of software agents:

(i) autonomy: agents have some degree of control over their actions and can work without intervention of humans;

(ii) social ability: agents can coordinate their actions and cooperate with other agents to achieve their goals, using a common language to communicate with each other;

(iii) reactivity: agents can perceive their environment and respond to environmental changes;

(iv) proactiveness: agents can act on their own initiative to achieve their goals instead of simply reacting with the environment.

For our research purposes, we further characterize a software agent as *a running program object, capable to initiate, receive, execute, or reject a message autonomously to attain its goals during its life cycle*.

2.2. Agent Platforms. An agent platform is a software environment in which agents are incarnated and operate to achieve their goals. The agent platform must provide the following minimum set of functionalities [4, 5]:

(i) agent management (creating, starting, removing, migrating agents, etc.),

(ii) agent communication,

(iii) supervision of agents, error notification,

(iv) security mechanism.

Today, several platforms have been developed (e.g., JADE [6], JACK [7], AgentBuilder [8], Aglet [9], etc.) and researches are being conducted to define new platforms for building agent systems. JADE was selected based on two criteria:

(i) the selected platform is well-proven;

(ii) it is scalable for our research and experimental purposes.

Java Agent DEvelopment Framework (JADE) is a software framework fully implemented in the Java language. It simplifies the implementation of multiagent systems through a middleware that complies with the Foundation For Intelligent Physical Agents (FIPA) specifications and through a set of graphical tools that supports the debugging and deployment phases. (FIPA is an IEEE computer society standards organization that promotes agent-based technology and the interoperability of its standards with other technologies [10].) This agent platform can be distributed across machines (which do not even need to share the same OS) and the configuration can be controlled via a remote

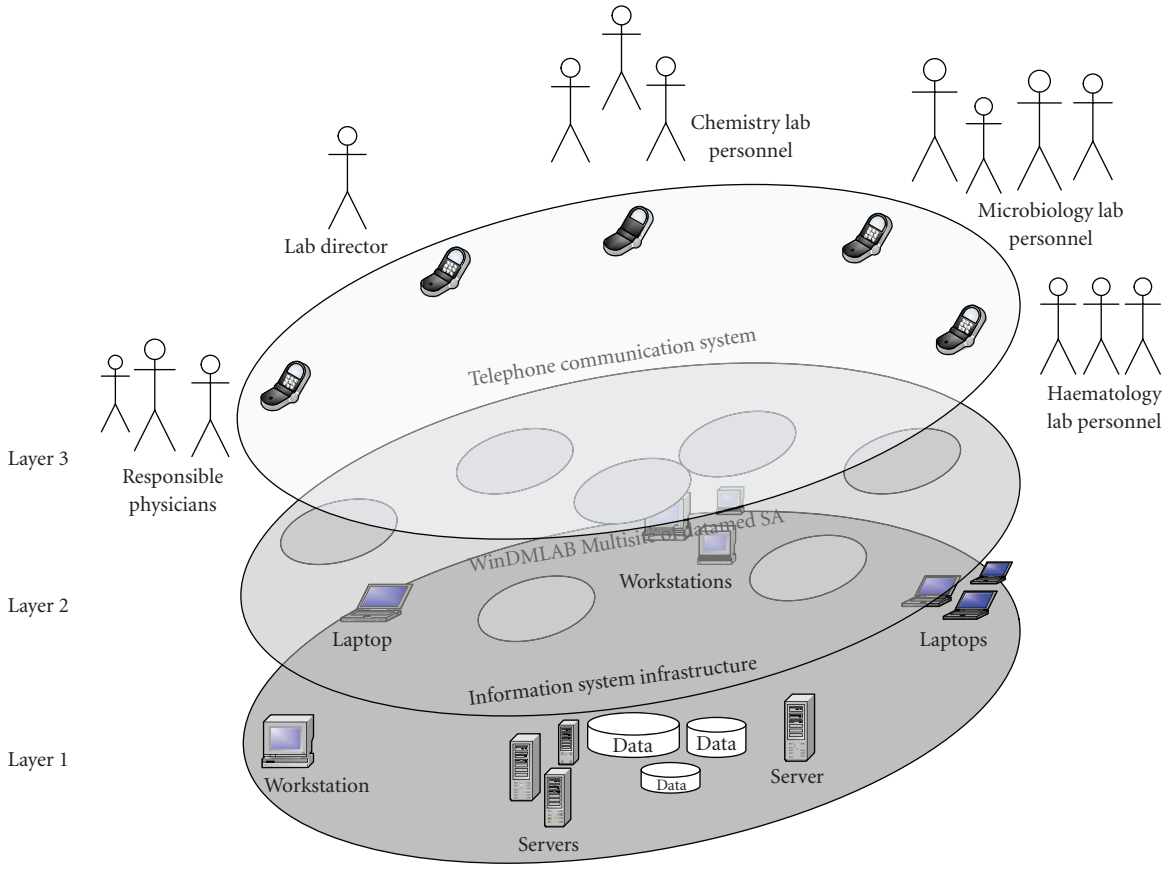


Figure 1: Layers of the current laboratory information system.

GUI. JADE has been developed by the Telecom Italia Lab [6], and the Agent and Object Technology Lab at the University of Parma [11]. It is open-source, cost-free, and offers the developer complete control over the framework. We refer the interested reader to [12] for a good introduction to the JADE agent platform.

2.3. Agent-Oriented Methodologies. The concept of agents was first introduced in the 1970's. However, the development of agent-based systems is a relatively new domain of software engineering. Today, several agent-oriented methodologies have been developed (e.g., Gaia [13], MaSE [14], and MAS-CommonKADS [15]). They are based on different theoretical foundations [16]: artificial intelligence (AI), object-oriented Programming (OOP), combination of AI and OO, as well as *i** organization modeling framework (Tropos) [17].

These methodologies contribute significantly to the rigorous and systematic development of agent-based systems. The JADE_Methodology [18] is a new agent-oriented methodology that supports the ontology approach. It encompasses the analysis and design phases to develop software agents on the JADE platform. This methodology proposes to build the ontology at the end of the design phase in order to share the knowledge between software agents.

3. HCF Laboratory: Current Organization and Software Agent Solution

3.1. The HCF Laboratory. The HCF Laboratory [19] provides medical analysis ordered by hospitals in the state. The laboratory is located on several sites with different domains: haematology, immuno-haematology, chemistry, and microbiology. It receives daily hundreds of orders with specimens, analyzes the specimens, then delivers final results to the requesters (doctors, hospital departments, etc.). The method of transmission of test results depends on their urgency level.

Besides the lab equipment for carrying out medical analysis, the personnel of the HCF Laboratory are supported in their daily tasks by the WinDMLAB Multisite laboratory information system [20], coupled with a traditional telephone communication system. They constitute two major components of the current HCF Laboratory Information System (cLIS).

cLIS ensures the availability of medical results in a centralized database and their transmission:

(i) between departments and sites of the laboratory,

(ii) between the laboratory and the HCF,

(iii) between the laboratory and other requesters in the province of Fribourg.

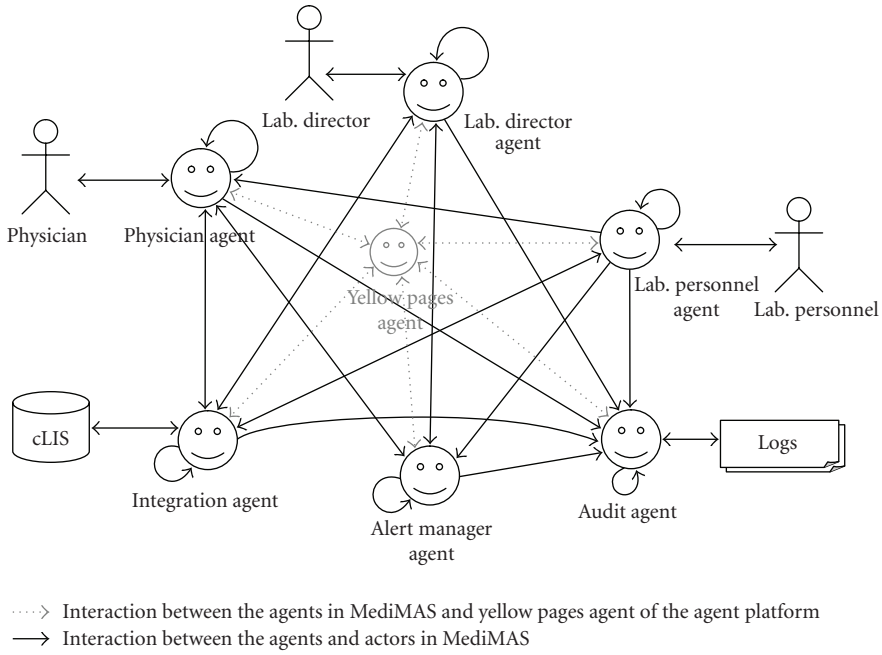


Figure 2: MediMAS overview.

Figure 3: Patrik's laboratory personnel agent GUI.

Each requester (doctors, hospital departments, etc.) can access and review the test reports on their patients at any level of detail.

The WinDMLAB Multisite system and the traditional telephone communication system must coexist to achieve all the functionalities as cLIS was initially designed for. Indeed, several scenarios still require the telephone communication system to get things done, for example, in the following circumstances:

(i) a lab technologist calls a physician to transmit patient's test results;

(ii) a physician calls the laboratory to obtain by phone the test results;

(iii) a lab technologist asks, by phone, his director to make a decision in an emergency situation, and so forth.

Figure 1 illustrates cLIS as a three-layer system in which both the laboratory information system and the telephone communication system coexist:

(i) the first layer defines the information system infrastructure, which is composed of servers running different operating systems and application software in a computer network;

(ii) the second layer is the WinDMLAB Multisite system;

(iii) the third layer provides the telephone communication system which allows requesters and laboratory staff to exchange test results via voice and fax.

One can notice that human actors interact with each other directly or indirectly through the second and third layers.

3.2. Potential Problems. cLIS raises numerous potential problems [21]:

(i) even though the major part of results (80%) are transferred through automats and WinDMLAB Multisite system, the quality of services provided by cLIS depends to a more or less extent on human factors, for example, any mistake of a lab technologist in transferring test results to a doctor may cause dramatic consequences on patients;

(ii) cLIS does not allow the requesters to know when results become available;

(iii) the processes which take place in the telephone communication system (layer 3) cannot be logged automatically in cLIS for monitoring and tracking purposes;

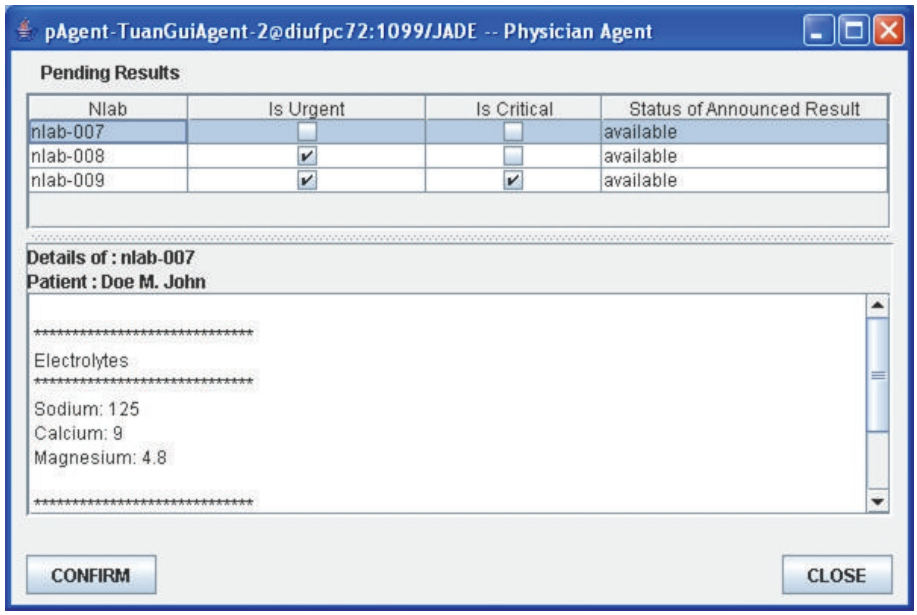


Figure 4: Tuan's physician agent GUI.

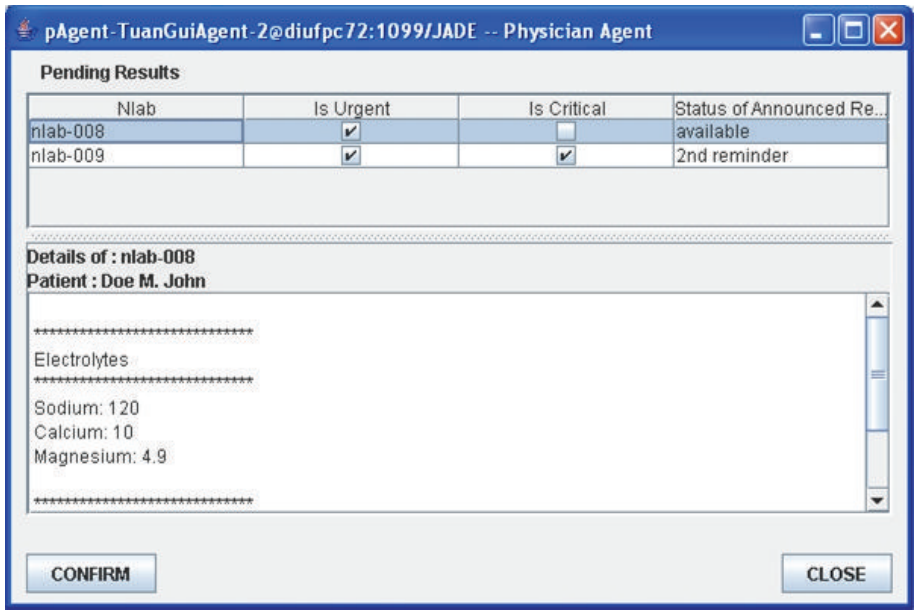


Figure 5: Tuan's physician agent GUI—after nlab-007 has been confirmed.

(iv) physicians who use cLIS spend a lot of time searching, retrieving, consulting, and interchanging the test results;

(v) to establish a successful phone communication, two actors must be present, therefore, time is wasted if either one cannot reach the other when needed;

(vi) because of the time-consuming use of cLIS in many scenarios, physicians and laboratory personnel have less time for their real medical activities.

The above-identified problems, caused by human operations, often prevent information to flow smoothly from cLIS to actors and vice versa. These problems illustrate the so-called "automation gap" [22, 23]. What is needed is a systematic, strategic approach that automates error-prone human processes.

3.3. A Software Agent Solution. The "automation gap" may be filled using different software technologies, for example, JavaSpaces with SMS message technology, Web services technology, multiagent technology, and so forth. It is out of the scope of this paper to compare these technologies. Our purpose is to propose a methodology for allowing us to migrate from the legacy human agent-centered cLIS

Table 1: The three simulated specimens.

Criticality/priority	None	Urgent
Non-critical	nlab-007	nlab-008
Critical	—	nlab-009

toward an enhanced software agent-based system. In cLIS, actors (laboratory personnel, laboratory director, physicians, etc.) are human agents. A human agent is a professional characterized by experience, skills, intelligence, reactiveness, proactiveness, and ability to work autonomously and to cooperate with other human agents. They also have weaknesses inherent to human beings. Our proposal aims at designing software agents which will work on behalf of human agents with similar characteristics. In other words, our solution delegates daily routine tasks performed by human agents to software agents. In this new approach, each actor is assigned a personalized software agent which acts as his personal assistant. We also say that the actor is an assistant's owner. When talking about these personal assistants, we could also use the "virtual twin" metaphor [24] or consider them as avatars representing humans like in virtual worlds. The assistant receives a list of things to do from its owner, performs the assigned tasks in close cooperation with other software agents, and delivers the final result to the owner. In our solution, the software agents are designed on Layer 3, shifting the telephone communication system up to the fourth layer (cf. Figure 1). The software agent solution offers significant advantages for cLIS:

(i) the features and functionalities of WinDMLAB Multisite are maintained, preserving the investment in this legacy laboratory application;

(ii) in the new software agent-based cLIS, the delegation of routine tasks from human to software agents (personal assistants) allows human actors to focus their attention on specimen analysis, test result interpretation, medical decision making, and so forth;

(iii) the new software agent-based cLIS, coupled with mobile devices (PDAs, mobile phones, smartphones, etc.), allows the actors to view the test results transmitted by personal assistants anywhere and anytime;

(iv) all events and actions are systematically logged and centralized to support auditing of the system. Traceability and exception investigation, for example, to answer a patient's complaint, is also improved.

4. The MediMAS Prototype

The MediMAS prototype [21] is the first experimental implementation of the proposed agent-based solution. A case study was conducted at the HCF Laboratory to test it in the real world, and to explore different practical aspects.

4.1. Agents as Personal Assistants. MediMAS has six agent categories:

(i) physician agents,
(ii) lab personnel agents,
(iii) lab director agents,
(iv) alert manager agent,
(v) integration agent, and
(vi) audit agent.

Figure 2 depicts their organization in which the agents assist different categories of humans in their daily tasks. This figure also shows the social ability of agents to cooperate with each other in order to automate the information flow between the actors themselves, as well as between the actors and the cLIS.

4.2. Software Agents in Action

4.2.1. Environment Setup. In the environment of our MediMAS prototype, the integration agent (riAgent) plays a central role. Therefore, it is launched first with the JADE platform before starting any other agent. When the setup is complete, the agents are attached to the MediMAS's containers (a JADE container is a runtime environment for agents [25]):

(i) riAgent is the integration agent,
(ii) amAgent is the alert manager agent,
(iii) adAgent is the audit agent,
(iv) pAgents are the physician agents,
(v) lpAgents are the lab personnel agents,
(vi) ldAgents are the lab director agents.

In the MediMAS system, each human actor (physician, lab personnel, lab director) is assigned an Agent, and simultaneously, one or more GuiAgents. For example, a single agent pAgent TuanAgent and two GuiAgents are assigned to the physician Tuan.

We now setup our sample WinDMLAB database by feeding it with the fictitious test results of specimens nlab-007, nlab-008, and nlab-009 in order to simulate the three test results which are recorded into the database by the lab analysers, and validated by the lab technologist. (Our sample WinDMLAB database was developed using SQLite RDBMS [26].)

Let us introduce the actors who will play different roles in our scenario:

(i) Tuan is a physician in the HCF and is assigned the ID 3;
(ii) Jacques is the lab director;
(iii) Patrik is a lab technologist in HCF Laboratory: he is working on the specimens: nlab-007, nlab-008, and nlab-009, ordered by a caregiver Tuan.

In the following scenario, starting with the notification of results availability, we study in finer detail the human actors, their assigned personal assistant agents, and their interactions.

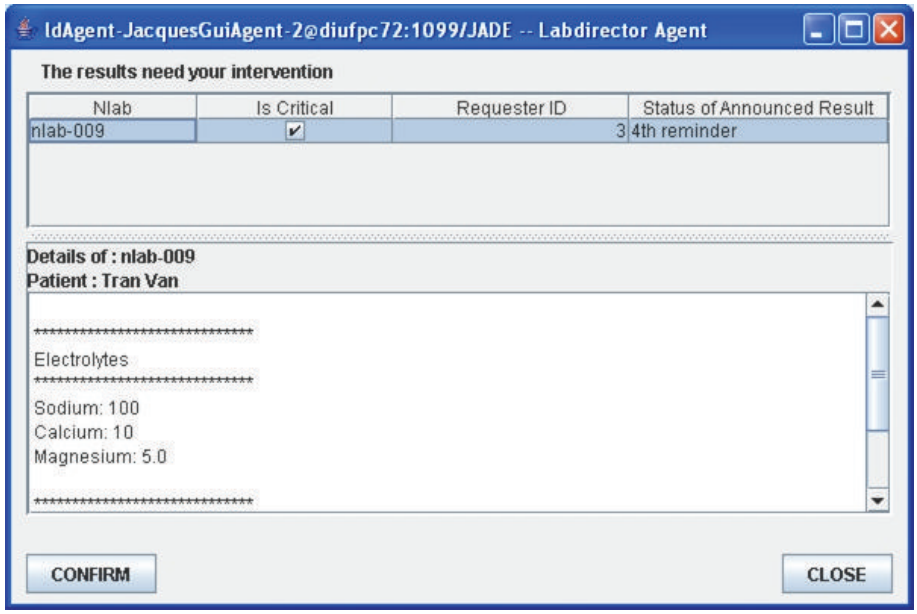


Figure 6: Jacques' lab director agent GUI.

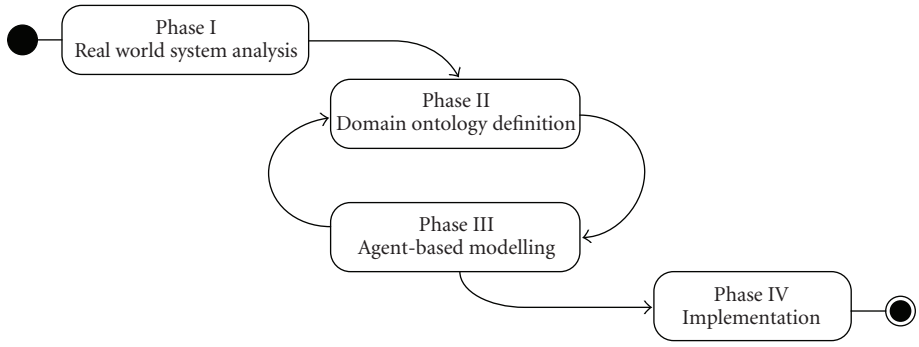


Figure 7: The phases of the methodology.

4.2.2. Notification of Results Availability.

(i) Patrik has finished the analysis of all specimens. The three test results are recorded into WinDMLAB database. Table 1 shows the priority of the specimens and their degree of criticality. (The priority of an analysis is set by its requester and the degree of criticality depends on its result and is set by the lab technologist.)

(ii) At completion of the nlab-007 analysis, Patrik observes that the test results are noncritical (see Table 1). In order to notify the availability of the test results to Tuan (requester ID = 3), Patrik enters nlab-007 and clicks on the button beside the NLAB field to automatically fill in the other fields (Figure 3). Finally, Patrik clicks the "notify result" action button to direct his lpAgent to announce the availability of test results to the requester.

(iii) Patrik further treats the other results in the same manner.

(iv) Patrik's lpAgent sends the announcements of the results to Tuan's pAgent.

(v) It also sends these announcements to amAgent which records the announcements and starts to monitor closely the read/unread status of the new test results.

4.2.3. Acknowledgments of Notification Receipt.

(i) Concurrently with amAgent, Tuan's pAgent receives the announcements and refreshes the list of pending results in the upper pane of its window by adding the new announcements of nlab-007, nlab-008, and nlab-009 test results, flagged as "available" in the status of announced Result column (Figure 4).

(ii) Tuan clicks on the received announcement nlab-007 in the list of pending results in order to preview the details of the test results. Tuan's pAgent requests riAgent to retrieve the contents of the nlab-007 test results and displays the contents of the nlab-007 test results in the lower pane of its window (Figure 4).

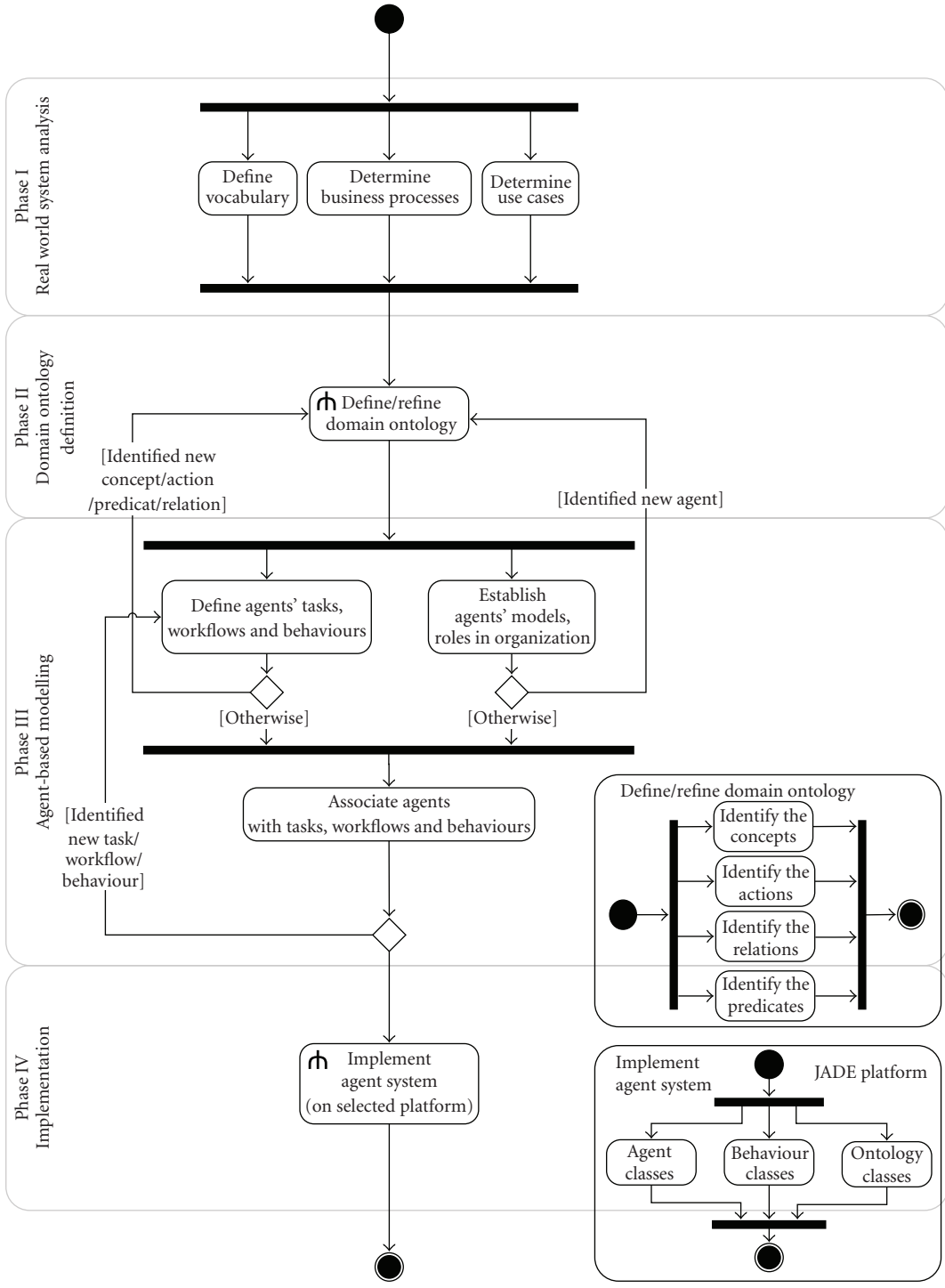


Figure 8: Development methodology.

(iii) Tuan clicks the "confirm" button to acknowledge receipt of the notified announcement of nlab-007 and thus directs his pAgent to send this acknowledgement to amAgent.

(iv) amAgent updates the status of nlab-007 as "read" and removes the nlab-007 announcement from his own internal list. This terminates the monitoring of nlab-007 by amAgent.

(v) Once the announcement is flagged as "read," Tuan's pAgent removes nlab-007 from the list of pending results (Figure 5).

(vi) Tuan further acknowledges the nlab-008 result.

One notices that, in the pAgent's window, each announcement is first flagged as "available" during a predefined time interval, for example, 20 minutes for normal test

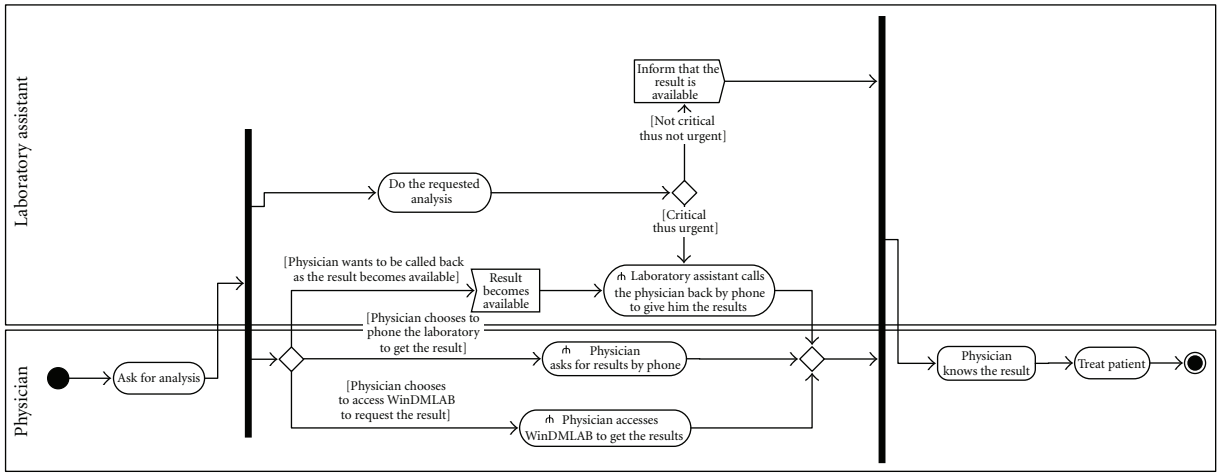


Figure 9: The business processes of the HCF laboratory.

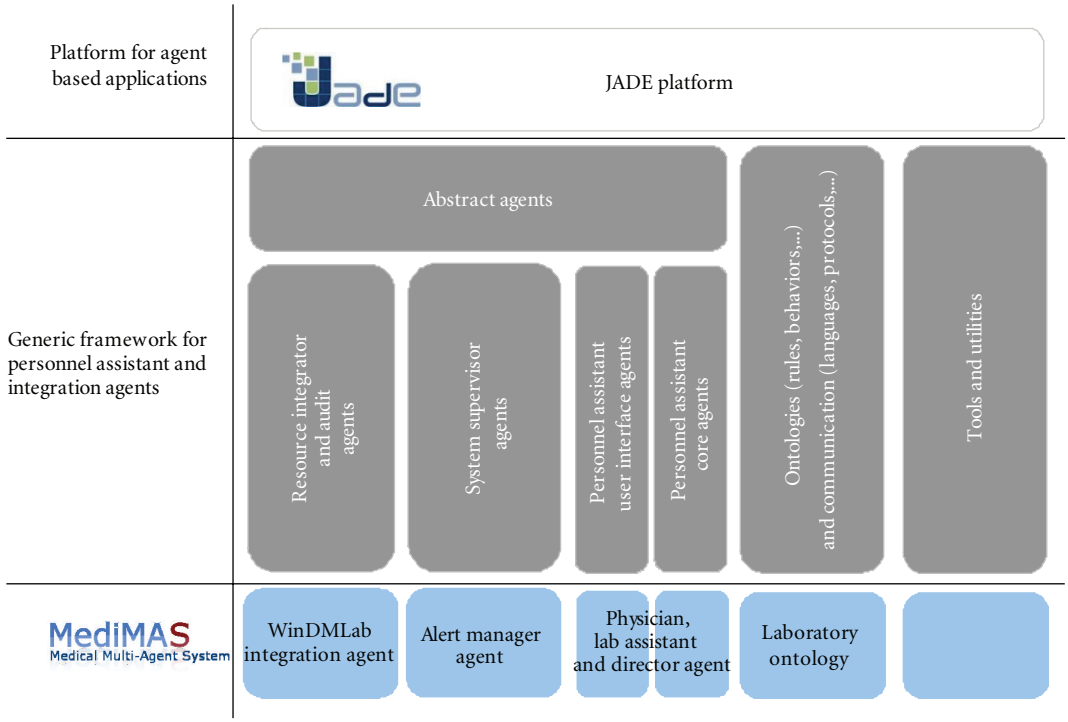


Figure 10: Overview of the software architecture.

results; and 10 minutes for critical ones. Thanks to the close monitoring of pending announcements, amAgent alerts pAgent as soon as an announcement has not been confirmed within the predefined time interval. pAgent immediately flags the alerted announcement as "1st reminder," then "2nd reminder," and so on in the Status of Announced Result column.

4.2.4. Problem Detection and Alert.

(i) For the nlab-009, amAgent has not yet received an acknowledgment message from Tuan's pAgent within the preset time interval. After three unsuccessful warnings, amAgent escalates up the organizational hierarchy by sending an alert to Jacques' ldAgent.

(ii) Jacques' ldAgent receives the nlab-009 alert from amAgent and displays it in the ldAgent's window (Figure 6).

(iii) Jacques clicks on the nlab-009 alert in order to preview it. Jacques's ldAgent requests riAgent to retrieve the contents of the nlab-009 test result and displays the contents of the nlab-009 test results in the lower pane of its window.

(iv) Jacques contacts Tuan to manually transmit the test results to him.

Table 2: Tasks performed by agent categories.

Agent categories	Tasks
Physician agent	Receives notification of test results availability from the lab personnel agents.
	Receives alerts of unread available test results from the alert manager agent.
	Notifies the physician that test results are available.
	Queries the integration agent for test results according to search criteria determined by the physician.
	Receives test results data from the integration agent.
	Displays test results data to the physician.
	Informs the alert manager agent about the read/unread status of the test results sent to the physician.
	Informs the audit agent before and after each action.
Lab personnel agent	Notifies the alert manager agent that test results are available.
	Notifies the physician agents that results are available.
	Informs the audit agent before and after each action.
Lab director agent	Receives alerts from the alert manager agent signaling the abnormal unread status of a test result.
	Reports alert to the lab director.
	Acknowledges the alert manager agent that the lab director read the alert sent to him.
	Informs the audit agent before and after each action.
Alert manager agent	Alerts the lab director agent as soon and as the abnormal unread status of a given test result is detected.
	Receives test results from the lab personnel agent.
	Receives from the physician agent the status "test results have been read by physician."
	Receives from the lab director agent the status "alert message has been acknowledged by the lab director."
	Informs the audit agent before and after every action.
Integration agent	Retrieves test results from cLIS, based on the query issued by the physician agent or the lab director agent.
	Delivers extracted test results to the requester agent.
	Informs the audit agent before and after every action.
Audit agent	Receives the actual action start/end notifications and log them with their date and time.

(v) Jacques clicks the "confirm" button to acknowledge receipt of the nlab-009 alert and thus directs his ldAgent to send this acknowledgment to amAgent.

(vi) AmAgent updates the status of nlab-009 as "read," and removes the nlab-009 announcement from his own internal list. This terminates the monitoring of nlab-009 by amAgent.

(vii) Once the announcement is flagged as "read," Jacques's ldAgent and Tuan's pAgent remove nlab-009 from their respective windows.

(viii) Throughout the above-simulated scenario, each agent sends to the audit agent (adAgent) the start and stop times of every performed task along with its relevant information (date and time, involved actors, action, etc.).

We have simulated some specimens to demonstrate the working of assistant agents in the MediMAS prototype and the benefits of a software agent approach to enhance a legacy information system. In order to fully grasp the power of our solution, one however must consider the real laboratory, where hundred of specimen analysis are ordered everyday by dozen of physicians. After a rather simple configuration process, each human actor will be able to transparently rely on his software counterpart to be reminded what he has to do next with respect to the hospital regulations. Furthermore, all communication exchanges and reminder warnings will be coordinated, timely delivered to all the appropriate actors, and properly logged for further references.

At this stage, the attentive reader has certainly noticed that we used a very high level approach in order to describe the concrete run-time working of the MediMAS prototype. It is, however, very important for her to understand that MediMAS components are not just plain objects, but they are, indeed, software agents in the sense of the definition given at the end of Section 2.1. Because of that, the use of agent technologies in general and of an agent platform in particular is a necessity if one does not want to reinvent the wheel by implementing from scratch many low-level services such as naming and yellow pages services, code mobility support, debugging and monitoring/management facilities, security mechanism, agent communication, or resource control. For example, the alert manager agent, amAgent introduced above, is a running program object, with its own thread of control (i.e., having its own autonomy), which

(i) reacts to physician and lab personnel agents messages by updating its test results pending list;

(ii) has an aim to timely detect and to act upon test results with abnormal unread status;

(iii) acts autonomously (i.e., without the necessity of a special external event or method call) in order to fulfill its goal. It does so by constantly monitoring its test results pending list and by sending warning messages to the appropriate agents (physician and lab director ones) according to the hospital regulations.

Messages are based on the FIPA ACL Message standard [10], and the behaviors or agent "intelligence" are programmed in Java classes using either plain procedural code or declarative rules with the help of the Jess to JADE Toolkit developed by our research group [27]. Note that with the latter technology, it is even possible to change the agent behavior by modifying rules at run-time (e.g., escalating up the organizational hierarchy after two instead of three unsuccessful warnings or warning another physician in the same group if available instead of the lab director).

5. Development Methodology

We have designed our own "in-house" methodology, inspired by the theoretical foundations mentioned in Section 2.3. More precisely, we adapted the JADE_Methodology [18] to our own purposes by integrating the ontology in the earlier phases of the modelling process. Our strategy has been applied to develop the MediMAS prototype. The next paragraphs present it in four phases (see Figure 7), while Figure 8 summarizes it and put in evidence the relationships between its different phases.

5.1. Phase I: Real-World System Analysis. The analyst perceives the current system in order to understand its goals, problems, and its future requirements. This phase aims at defining a common vocabulary and describing the current organization of entities (actors, human agents), use cases, and/or business processes of the system. The deliverables of Phase I consist in a well-defined set of goals and requirements, the common vocabulary describing the entities with their organization, a set of identified use cases, and business processes. In our case study, the outputs of our real-world system analysis are the three-layer information system structure of the HCF Laboratory (Figure 1), and UML activity diagrams of its business processes (Figure 9).

5.2. Phase II: Domain Ontology Definition. The Domain Ontology Definition phase takes the deliverables of Phase I as input and aims at defining the domain or application terminology standards and semantics. To this end, the analyst focuses on concepts, actions, predicates and relations between concepts. In MediMAS, we adopt the following guidelines to build the ontology:

(i) Concepts are substantives (e.g., doctor, patient, analysis, etc.).

(ii) Actions are verbs or verbal phrases (e.g., SendResult, Alert, SendAvailableList, etc.).

(iii) Predicates are expressions that make statements about something, which can be evaluated true, false or indeterminate (e.g., isTestResultCritical, isResultComfirmed, etc.).

(iv) Relations are expressions that establish the relationship between concepts.

The output of this phase is the domain or application ontology, that actors will use to understand each other in their communications.

In software engineering, ontology development tools, such as Protégé [28], TopBraidComposer [29], etc., have been developed in order to assist the ontologists to build the domain or application ontology efficiently. The interested reader is refered to [30] for a graphical overview of the ontology we defined using the Protégé suite of tools.

5.3. Phase III: Agent-Based Modelling. The modelling phase consists in the following set of tasks using the deliverables of Phase I and II as inputs:

(i) identify and create eligible software agents which will be assigned to actors;

(ii) determine the tasks (also called the responsibilities) of each agent;

(iii) specify the workflow of elementary operations in each task and the agent's operational behavior;

(iv) assign tasks, workflows, and behaviors to agents according to their roles in the organization.

Figures 7 and 8 draw our attention to the iterative nature of the tasks within Phase III on one side, and between Phases II and III on the other side. Indeed, successive refinement steps are required in order to enrich the domain ontology as new concepts, actions, predicates, and relations between concepts are identified.

The deliverables of this phase are the documents:

(i) describing the agents in different categories, and

(ii) specifying all the tasks, workflows, and behaviours, and their assignment to agents.

The agent categories and their assigned tasks in MediMAS are summarized in Table 2.

5.4. Phase IV: Implementation. The previous phases are platform-independent. In Phase IV, the selection of a platform closely impacts the implementation process. In our case study, the JADE platform was selected to implement the MediMAS prototype.

This phase involves the programmer team to implement and test the agent-based system according to the model specifications. To this end, the programmers use the deliverables of the previous phases as inputs, and then translate them into system components which are extensions of the existing classes in JADE, namely:

(i) designed agents are translated into classes of agents according to the terminology used in JADE;

(ii) designed tasks, workflows, and behaviours are converted into classes of behaviours in the sense of JADE.

The domain ontology must also be implemented as extensions of the existing ontology in JADE. This task is achieved:

(i) either by manually coding vocabulary, bean classes, ConceptSchema, AgentActionSchema, PredicateSchema, and so forth, or

(ii) through the bean generator plug-in for Protégé [31].

The completion of phase IV results in a multiagent system that fulfils the defined user goals and requirements and operates on the selected platform. It would be out of the scope of this article to fully describe the software architecture of the MediMAS prototype. It is nevertheless worth giving an overview of its main software components. (The interested reader can find the class diagram of MediMAS as implemented on the JADE platform in [30] and its complete source code is available at [32].)

A typical layered approach has been adopted (see Figure 10): the upper layer is an abstract layer providing the basic classes, interfaces, and agent types, and it directly extends the JADE platform. The second layer offers the main functionalities and default behaviors for each kind of identified agent type: resource integration for seamless interfacing with legacy systems, audit agent for addressing logging issues, and alike system supervisor agents which enhance the system with some new services and the personnel assistant agents which embody the "virtual twin" paradigm. Note that this latter category is split into core and user interface agents. This separation allows for a one-to-many relationship between a personal agent's core part and several user interface agents which are deployed on the humans' computing devices (desktops/laptops and/or smartphones and/or web browsers, etc.).

These agent families form the main vertical blocks of our architecture. Eventually, the lowest layer is dedicated for application specific implementations of the agents. In the case of MediMAS, this layer contains

(i) the WinDMLAB integration agent,

(ii) the alert manager agent,

(iii) the lab assistant, lab director and physician personnel assistant agent.

Beside these blocks, there are two further components (rightmost on Figure 10): one for ontology related issues and one for miscellaneous tools and utilities.

This layered architecture actually provides a general framework that could be used for other application domains than our medical laboratory use case. In order to reuse the framework, one could simply inject a new ontology, attach the according behaviours to the personnel assistant agents, and implement the business logic of the system supervisor agents.

6. Conclusion

This research paper discusses major features and benefits of our agent-based approach to enhance a hospital laboratory legacy information system. Such approach preserves the investment in the legacy system and allows developers to seamlessly add new features, which aim at filling the automation gap, satisfying the needs of growing user mobility, and providing intelligent assistance to users. Finally, a methodology to systematically adopt and implement such a solution is proposed and it is validated with the implementation of the concrete MediMAS prototype.

6.1. Achievements. The current version of the MediMAS prototype provides physicians, lab personnel, and lab director with software agents running on desktop computers. (The whole source code and related documentation are available for download from [32].) These agents act as personal assistants to free the actors from tedious and routine work so that they can really concentrate on their medical activities.

6.2. Work in Progress

6.2.1. Mobile MediMAS. Our research will extend the model to allow software agents to run on mobile devices (e.g., PDAs, mobile phones, smartphones, etc.). The agents that work for the same owner on different devices must collaborate and synchronize their tasks to efficiently assist the owner who may work anywhere and anytime. A first prototypal version of this extended model is already available [32, 33], but still needs some fine tuning.

6.2.2. MediMAS Simulation Tool. The development of a simulation tool for MediMAS is another topic of our research. The tool offers the HealthCare experts the opportunity to visualize the working of MediMAS prototype by simulation, and to get an insight in the properties of an agent-based system in the HealthCare domain (ubiquitousness, intelligence, reactiveness, proactiveness, scalability, etc.). A first version of the tool is now available [34] and has been extensively used in order to debug and test the MediMAS prototype.

6.2.3. Adaptive MediMAS Agents. Withing another project, we developed the Jess to JADE (J2J) toolkit [27], which allows JADE agents to seamlessly use the Jess rule engine [35] in order to perform appropriate behavior. This solution has been tested on our alert manager agent and it allowed us to declaratively define and modify the agent behavior at runtime.

6.2.4. Methodology Enhancement. The light in-house agent-based system design methodology has been defined, and applied in the MediMAS experimental project in HealthCare domain. Future extensions will enhance the methodology with additional modelling possibilities to design more complex real-world systems.

References

[1] J. Bisbal, D. Lawless, B. Wu, and J. Grimson, "Legacy information systems: issues and directions," *IEEE Software*, vol. 16, no. 5, pp. 103–111, 1999.

[2] S. Franklin and A. Graesser, "Is It an agent, or just a program?: a taxonomy for autonomous agents," in *Proceedings of the Intelligent Agents III Agent Theories, Architectures and Languages (ATAL '96)*, vol. 1193, pp. 21–35, Springer, Budapest, Hungary, August 1996.

[3] Daniel H. Wagner Associates, Inc., "Software Agents," http://www.wagner.com/technologies/softwareagents/softwareagents.html.

[4] F. Bellifemine, A. Poggi, and G. Rimassa, "Developing multi-agent systems with a FIPA-compliant agent framework," *Software: Practice and Experience*, vol. 31, no. 2, pp. 103–128, 2000.

[5] F. Bellifemine, G. Caire, A. Poggi, and G. Rimassa, JADE—A White Paper, http://jade.tilab.com/papers/2003/WhitePaperJADEEXP.pdf.

[6] Telecom Italia Lab, "Java Agent DEvelopment Framework," http://jade.tilab.com/.

[7] Agent Oriented Software Limited (AOS), "JACK Documentation," http://www.aosgrp.com/products/jack/documentation_and_instructi/jack_documentation.html.

[8] Acronymics, Inc., "AgentBuilder," http://agentbuilder.com/.

[9] IBM, "Aglets," http://www.trl.ibm.com/aglets/.

[10] FIPA, "IEEE Foundation for Intelligent Physical Agents," FIPA, http://www.fipa.org/.

[11] The Agent and Object Technology Lab at University of Parma, University of Parma, Italy, April 2008, http://aot.ce.unipr.it/.

[12] F. L. Bellifemine, G. Caire, and D. Greenwood, *Developing Multi-Agent Systems with JADE*, John Wiley & Sons, New York, NY, USA, 2007.

[13] F. Zambonelli, N. Jennings, and M. Wooldridge, "Multi-agent systems as computational organizations: the Gaia methodology," in *Agent-Oriented Methodologies*, chapter 6, pp. 163–171, Idea Group Publishing, Hershey, Pa, USA, 2005.

[14] S. A. DeLoach and M. Kumar, "Multi-agent systems engineering: an overview and case study," in *Agent-Oriented Methodologies*, chapter 11, pp. 317–340, Idea Group Publishing, Hershey, Pa, USA, 2005.

[15] C. A. Iglesias and M. Garijo, "The agent-oriented methodology MAS-CommonKADS," in *Agent-Oriented Methodologies*, chapter 3, pp. 46–78, Idea Group Publishing, Hershey, Pa, USA, 2005.

[16] B. Henderson-Sellers and P. Giorgini, Eds., *Agent-Oriented Methodologies*, Idea Group Publishing, Hershey, Pa, USA, 2005.

[17] P. Giorgini, M. Kolp, J. Mylopoulos, and J. Castro, "Tropos: a requirements-driven methodology for agent-oriented software," in *Agent-Oriented Methodologies*, chapter 2, pp. 46–78, Idea Group Publishing, Hershey, Pa, USA, 2005.

[18] M. Nikraz, G. Caire, and P. A. Bahri, "A methodology for the development of multi-agent systems using the JADE platform," *Computer Systems Science and Engineering*, vol. 21, no. 2, pp. 99–116, 2006.

[19] Hôpital cantonal de Fribourg, Kantonspital Freiburg, http://www.hopcantfr.ch/index.html.

[20] Datamed SA: Informatique médicale et scientifique: WinDM-LAB, http://www.datamed.ch/cms/front_content.php?changelang=1&idcat=73.

[21] A. Ruppen, *Systèmes multi agents—MediMAS: etude de cas dans le domaine du E-health care*, Bachelor thesis, Department of Informatics, University of Fribourg, Fribourg, Switzerland, September 2007, http://diuf.unifr.ch/softeng/student-projects/completed/ruppen/.

[22] S. W. Gozdan, "How big is your process automation gap?" *Mortgage Banking*, vol. 68, no. 1, p. 176, 2007.

[23] A. Rasmussen, "Closing the IT Process Automation Gap," Ptak, Noel & Associates LLC—White Paper, May 2007, http://www.opalis.com/upload/whitepapers/Closing_the_IT_Process_Automation_Gap.pdf.

[24] A. Gachet and P. Haettenschwiler, "The virtual twin: a socialization agent for peer-to-peer networks," *International Journal of Intelligent Information Technologies*, vol. 1, no. 2, pp. 56–67, 2005.

[25] D. R. A. de Groot and F. M. T. Brazier, "Identity management in agent systems," in *Proceedings of the 1st International Workshop on Privacy and Security in Agent-Based Collaborative Environments (PSACE) at the 5th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS '06)*, N. Foukia, J. Seigneur, and M. Purvis, Eds., pp. 23–34, Future University, Hakodate, Japan, May 2006.

[26] SQLite, http://www.sqlite.org/.

[27] J. Vogt, *Jess to JADE toolkit (J2J)—a rule-based solution supporting intelligent and adaptive agents*, M.S. thesis, Department of Informatics, University of Fribourg, Fribourg, Switzerland, August 2008.

[28] Stanford Center for Biomedical Informatics Research, "The Protégé Ontology Editor and Knowledge Acquisition System," http://protege.stanford.edu.

[29] TopQuadrant, Inc., "TopBraid Composer," http://topbraidcomposer.com/.

[30] M. T. Nguyen, P. Fuhrer, and J. Pasquier-Rocha, "Enhancing legacy information systems with agent technology: the case of a hospital medical laboratory," Internal Working Paper 07-11, Department of Informatics, University of Fribourg, Fribourg, Switzerland, December 2007.

[31] C. V. Aart, Ontology Bean Generator, http://protege.cim3.net/cgi-bin/wiki.pl?OntologyBeanGenerator.

[32] MediMAS, "Medical Multi-Agent Systems," http://diuf.unifr.ch/softeng/projects/medimas/.

[33] J. Schaeppi, *Extension de MediMAS: développement et déploiement d'agents JADE sur des supports mobiles*, Bachelor thesis, Department of Informatics, University of Fribourg, Fribourg, Switzerland, September 2008, http://diuf.unifr.ch/softeng/student-projects/completed/schaeppi/.

[34] B. Pointet, *MediMASim: a test and simulation toolkit for the medimas application*, Bachelor thesis, Department of Informatics, University of Fribourg, Fribourg, Switzerland, June 2008, http://diuf.unifr.ch/softeng/student-projects/completed/pointet/.

[35] Jess, "Jess, the Rule Engine for the Java Platform," http://herzberg.ca.sandia.gov/.

Hindawi Publishing Corporation
International Journal of Telemedicine and Applications
Volume 2009, Article ID 519417, 12 pages
doi:10.1155/2009/519417

Research Article

An Adaptive Source-Channel Coding with Feedback for Progressive Transmission of Medical Images

Jen-Lung Lo,[1] Saeid Sanei,[1] and Kianoush Nazarpour[1,2]

[1] *Centre of Digital Signal Processing, School of Engineering, Cardiff University, Cardiff CF24 3AA, UK*
[2] *Behavioural Brain Sciences, School of Psychology, University of Birmingham, Birmingham B15 2TT, UK*

Correspondence should be addressed to Jen-Lung Lo, loj@cf.ac.uk

Received 31 December 2007; Revised 11 August 2008; Accepted 6 November 2008

Recommended by Hui Chen

A novel adaptive source-channel coding with feedback for progressive transmission of medical images is proposed here. In the source coding part, the transmission starts from the region of interest (RoI). The parity length in the channel code varies with respect to both the proximity of the image subblock to the RoI and the channel noise, which is iteratively estimated in the receiver. The overall transmitted data can be controlled by the user (clinician). In the case of medical data transmission, it is vital to keep the distortion level under control as in most of the cases certain clinically important regions have to be transmitted without any visible error. The proposed system significantly reduces the transmission time and error. Moreover, the system is very user friendly since the selection of the RoI, its size, overall code rate, and a number of test features such as noise level can be set by the users in both ends. A MATLAB-based TCP/IP connection has been established to demonstrate the proposed interactive and adaptive progressive transmission system. The proposed system is simulated for both binary symmetric channel (BSC) and Rayleigh channel. The experimental results verify the effectiveness of the design.

1. Introduction

Transmission of medical images aiming at medical consultation, diagnosis, and treatment, or for training purposes demands highly reliable and high-speed communication systems. Since, typically, medical images contain large amounts of crucial clinical data, no visible encoding error is tolerated in the clinically important regions. Hence, advanced methodologies for gaining distortion-less reconstructed images, in fairly acceptable transmission rate, make it possible to transmit a medical image over a noisy channel. However, even by utilizing current high-speed broadband connections, the transmission of medical images has not been entirely successful; although using more parity bits enables a higher protection of the data against channel noise, the transmission time and the complexity of the channel coding considerably increase. By prioritizing the image regions based on their clinical information contents in one hand, and establishing a variable parity length with regard to the state of the communication channel on the other hand a higher data speed and immunity against noise can be achieved.

The concept of image coding utilizing discrete wavelet transform (DWT) and embedded zerotree wavelet (EZW) has been initially proposed in [1] and various methods have been developed based on this idea [2–8]. In [2], an effective scheme for image compression has been proposed where the spatial-spectral features of the image have been taken into account in order to show that wavelet transform is particularly well suited for progressive transmission. In [3], the concepts of DWT, EZW, progressive image transmission, and RoI have been utilized. These techniques have been effectively employed in our work. In [4], a method for progressive image transmission using EZW is presented where a number of constraints are imposed to provide variable bitrates for each frequency band. In [5], a compression method using linked significant tree (LST) has been utilized to encode all the significant coefficients in order to reduce the number of symbols and increase the compression efficiency. The authors in [6] have developed an image coding algorithm

that incorporates the features of zerotree and zero-block based algorithms. The main contribution in this recent algorithm is partitioning the wavelet-transformed image into coefficient blocks and to generate roots in top-most subband by using a block tree. In [7], another progressive image transmission method has been proposed based on a quadtree segmentation procedure in order to provide fairly good quality transmitted images while keeping the computational cost low. The authors in [8] have developed strategies to exploit the wavelet coefficients in different subbands for designing different vector quantization (VQ) coding to achieve a fast and efficient progressive transmission.

Shannon's information theory states that the performance of transmission schemes can be optimized in source and channel coding separately. However, the result holds with infinite block size, infinite coding complexity, and stationary channels. Such conditions are difficult to meet in practice. Hence, joint source-channel coding (JSCC) scheme attracts the interest of many. The JSCC scheme consists of a quantizer, an entropy, and channel coders to meet the target source rate, to achieve the required robustness in channel coding, and to find an optimal bit allocation between source and channel coding systems. Several methods have focused on designing adaptive joint source-channel coding (JSCC) schemes and introducing the properties of unequal error protection (UEP) and rate allocation to achieve efficient transmission [9–20]. In [9], authors surveyed several algorithms in JSCC system design and fast source-channel bit allocation for various transmission channels. They also provided a local search strategy to improve the initial low-complexity and rate-optimal scheme to achieve a distortion-optimal solution. An image transmission system was proposed in connection with JPEG2000 [10, 11]. The JSCC and UEP schemes in this approach try to optimize the rate-distortion function to achieve an efficient transmission. JPEG2000 provides high compression efficiency and introduces various features not included in other compression algorithms. This method has also adopted a standard component for medical image compression by digital imaging and communications in medicine (DICOM). JPEG2000 adopts embedded block coding with optimal truncation (EBCOT) as a source coder. The EBCOT algorithm has high compression efficiency with a high complexity price, which needs a powerful central processing unit (CPU) for support. The compression part of the proposed scheme is similar to that of JPEG2000's. However, unlike JPEG2000's our scheme provides variable parity length for an adaptive channel coding, which replaces EBCOT with EZW for compression and progressive image transmission. In [12], a framework for solving the end-to-end problem of progressive transmission of images over noisy channels has been presented which allows finding the optimal length of parity codes for each fixed length package to have minimum distortion in the decoded images. In [13], a joint source-channel decoder-based method for data transmission over noisy channels has been introduced. Multipath fading for communication over a noisy channel has become a complex problem. The authors in [13] have used the maximum a posteriori (MAP) method to design a joint source-channel coding system. An adaptive source-channel coding scheme based on subband coding has been used in [14]. In this approach, a suitable source and channel coding rates for each subblock has been proposed to minimize the total distortion. In [15], a low-complexity technique for the transmission of medical images has been proposed whereby the channel noise information has been effectively exploited. The main idea in [15] is to segment the bitstream into consecutive subblocks of variable lengths and consider a tradeoff between the levels of source and channel coding systems. The authors in [16] proposed a JSCC method for transmission of images over fading channels and demonstrated the application of rate-compatible low-density parity-check (RC-LDPC) codes constructed by the progressive edge-growth algorithm, and used the UEP to protect the images. In [17], a parametric methodology in progressive source-channel coding for rate allocation has been developed. The channel code rate is chosen based on the properties of source coder and the conditions of channel. The UEP strategies for efficient progressive transmission are proposed in [18]. Under the condition of a target transmission rate, the JSCC algorithm computes a UEP scheme that maximizes the number of corrected bits over a noisy channel. In [19], authors have used a concatenation of rate-compatible punctured convolution code and cyclic redundancy check (CRC) code to form a UEP scheme and find the optimal rate allocation solutions for progressive image transmission.

Many communication systems allow two-way communication implying that the signals are back from receiver to transmitter to adjust the system parameters and obtain better system performance. The authors in [20–23] have utilized the concept called hybrid automatic repeat request (HARQ) to ask for retransmission of erroneously received data and tradeoff allocation between the source and channel codes according to a rate-distortion optimization policy. Many researches on tradeoff allocation bits between source codes and channel codes assume that the noise-level in the channel is known in advance. Therefore, the feedback signal is figured out based on the known noise levels and the constraints set by the user. In the proposed algorithm, this has been modified since the parity lengths change according to the noise level in the received data whereby the amount of detected incorrect data is used to predict the conditions of the practical transmission channel. For medical image transmission, the quality of the reconstructed images (especially in the RoI) should be acceptable. This can be set as the constraint for the quantizer and the compression algorithm in advance. Therefore, the quality of the reconstructed image is only affected by the channel state and the proximity to the image RoI. The feedback signal in the proposed scheme updates the parity length without the need for retransmission of the data or adding any extra overhead.

The principal idea behind all these methods is that in a progressive transmission framework, the receiver reconstructs the transmitted image at various bit rates, which makes the fast and reliable retrieval of large images possible. In other words, the quality of the reconstructed image totally

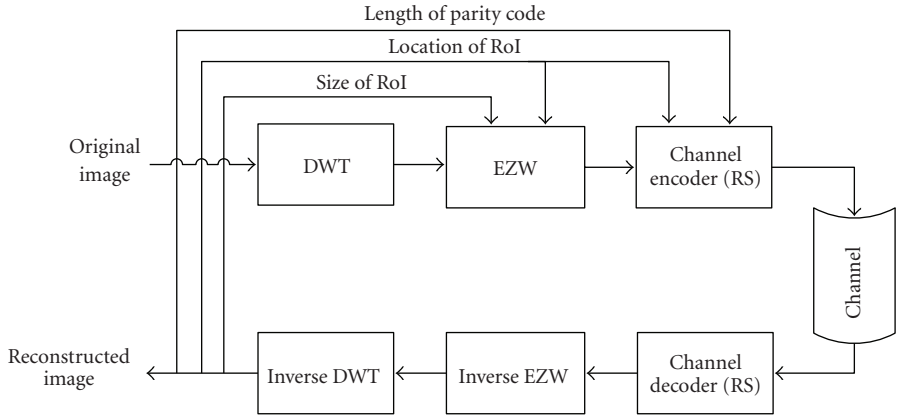


Figure 1: The overall system block diagram.

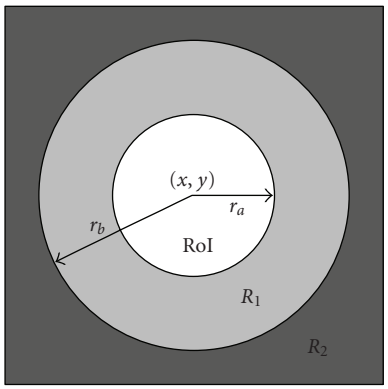


Figure 2: Areas of different priorities in an image centered at the center of RoI.

depends upon the volume of the received data, and the images can be reconstructed in any bitrate. Furthermore, the image subblocks are coded separately. Due to the high sensitivity to transmission noise, progressive transmission of images over noisy channels has to be accompanied by an appropriate channel coding, or a joint source-channel coding scheme [12]. The noise in the current communication systems can be due to the electronic components, fading, Doppler shift for mobile systems, bad weather, interferences, attenuations, and so forth.

The Reed-Solomon (RS) codes utilized here are block-based error correcting codes and are widely used for channel coding. The RS(p, q) codes correct the symbol error and not the bit error; lengths in terms of symbols. Thus, RS is suitable for burst error detection and correction [23].

2. Joint Source-Channel Coding

In this paper, we propose a novel interactive and adaptive joint source-channel coding with feedback algorithm for progressive transmission of medical images. This approach benefits from the idea of the JSCC, RoI, UEP, and feedback technique together as follows.

(1) The conventional RS channel coding has been used.

(2) The variability of the parity code corresponds to both the proximity to the center of RoI and the state of the practical transmission channel at the same time; this makes an efficient source-channel coding possible.

(3) The selectivity of the RoI is totally interactive and can be defined by the user in the receiver. This makes the method favorable to be used by clinicians who require fast access to the RoI in the image.

(4) An algorithm for detection of the blocks in error is developed to detect and recover the corrupted data, estimate the noise level in the practical transmission channel, and feedback the information of the noisy channel to the transmitter to control the error rate in the reconstructed images in the subsequent transmission.

By utilizing our flexible system, a minimum distortion of the transmitted images in a fairly shorter transmission time is achieved. As the main contribution of this research, we adaptively control the lengths of parity code streams simultaneously with respect to the selected region (i.e., longer lengths correspond to the areas closer to the center of RoI) and the amount of corrupted received data in the receiver. The system block diagram is shown in Figure 1.

This paper is organized as follows. Section 3 briefly describes the concepts of DWT and EZW. Following that, we provide the details of RoI selection. In Section 2, application of RS channel coding in a variable-parity length scheme will be explained. In Section 4, an algorithm for detection of the blocks in error is developed to evaluate the amount of incorrect received data in the receiver. Simulation results are subsequently reported in Section 5 followed by concluding remarks in Section 6.

3. Discrete Wavelet Transform and Embedded Zerotree Wavelet

Wireless transmission of medical images involves construction of an effective joint source-channel coding to not

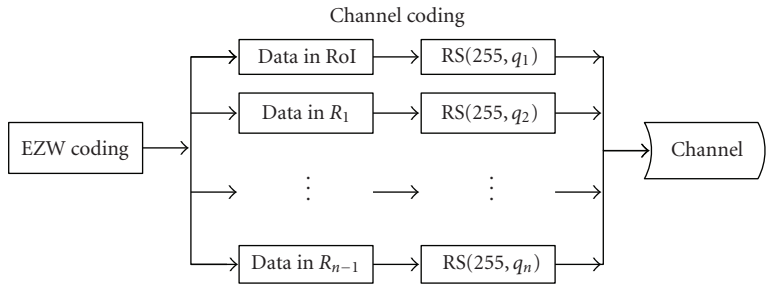


Figure 3: The transmitter including the proposed channel coding block diagram.

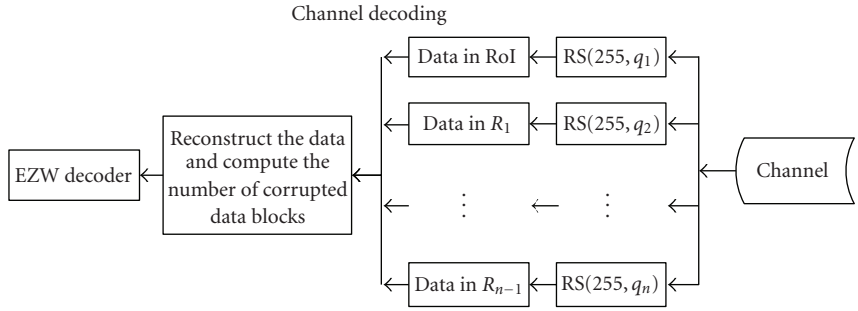


Figure 4: The block diagram for the receiver.

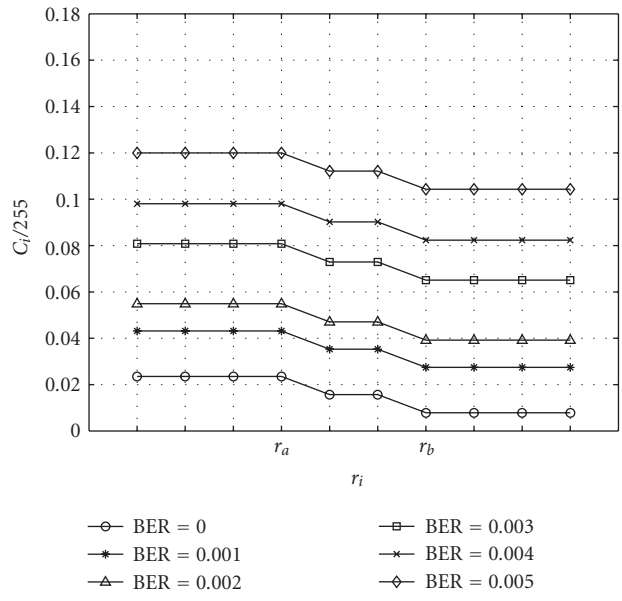


Figure 5: The ratios between the lengths of parity code C_i and overall codeword of 255 symbols in different regions at various fixed noise levels.

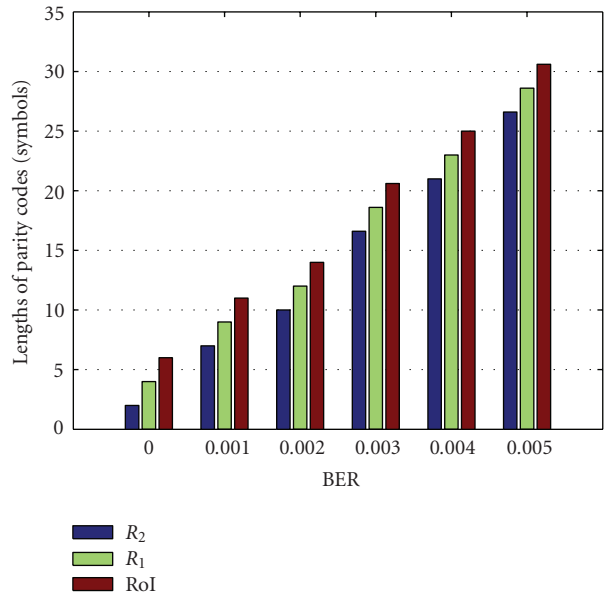


Figure 6: Lengths of the parity codes based on various channel noise levels.

only preserve the diagnostic information but also to enable progressive streaming of the data from the host to the receiver. EZW is a simple, efficient, and flexible compression algorithm for low bitrate image coding. The properties of DWT and EZW allow us to code and compress the data blocks individually and also compress it at any bitrate. Therefore, based on progressive encoding, we can compress a block into a bitstream with increasing accuracy. Traditionally, the input images are decomposed into many subblocks each to be coded, compressed, and transmitted individually. Therefore, the input image is segmented into a number of subblocks firstly. And then wavelet transform (WT) decomposes each subblock into different time-frequency components.

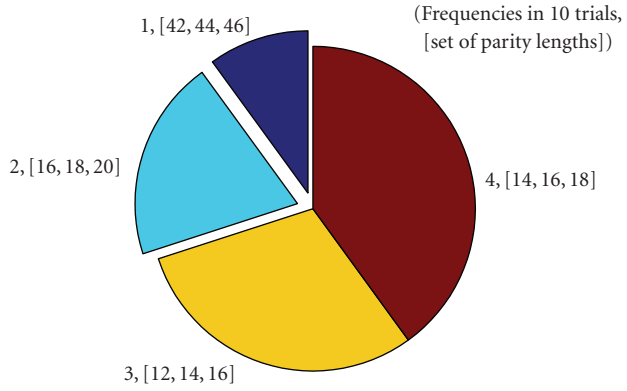


Figure 7: Average distribution (frequency) of the set of parity lengths in 10 trials for BER = 0.003.

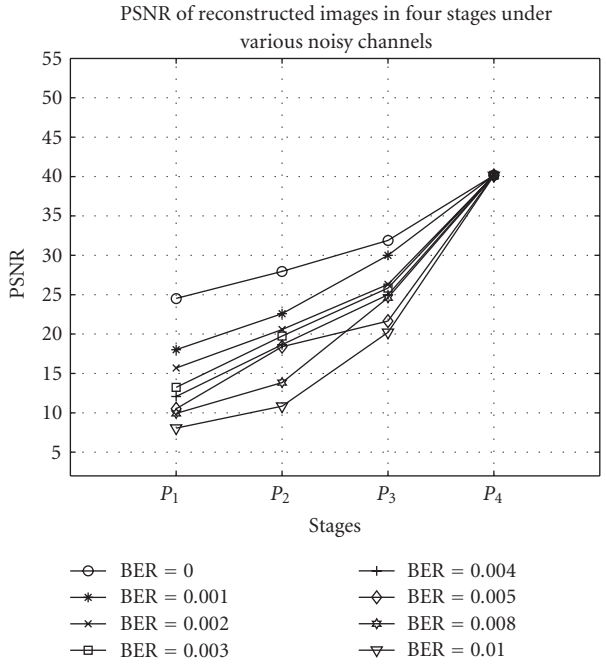


Figure 8: PSNR for successive transmission of four stages at different BERs under various noise-level conditions.

As it is detailed in [24], EZW codes the image into streams of six symbols, namely, $p, n, z, t, 0$, and 1. In mathematical terms, considering the image amplitude at location (x, y) is denoted by $\gamma(x, y)$, and t_n is the threshold in nth iteration, the definitions of the symbols are given in Table 1. EZW suits progressive data transmission since it enables hierarchical encoding and decoding.

In the proposed system, we chose a 3-level Haar wavelet transform (HWT) to perform the DWT for each subblock due to its simplicity and being faster and easier to implement in comparison with other DWT methods [21]. The coefficients in the lowest frequency subbands show the background information of the subblocks. The coefficients in the higher frequency subbands represent the details and edges. After computing the HWT, we compress the coded data according to a variable thresholding mechanism governed by the EZW. Hence, a suitable approach is to use a variable threshold and transmit only those coefficients to the decoder that are larger than the threshold. The first step in the EZW algorithm is to determine the initial threshold level t_0 and then repeatedly lowering the threshold by half at a time until the threshold has become smaller than the smallest coefficient to be transmitted; or the iteration is stopped by request. The initial threshold t_0 is set as follows:

$$t_0 = 2^N, \quad N = \log_2 \max\left(\left|\gamma(x, y)\right|\right), \tag{1}$$

where $\max(\cdot)$ refers to the maximum value. The final threshold level determines the length of the bitstream output through the EZW process, the compression ratio of the input images, and the resolution of the reconstructed image. The length of the output bitstream M_i is related to the number of times the threshold is halved as

$$M_i = \sum_{k=0}^{n_{Ti}} B\left(\frac{t_{0i}}{2^k}\right), \tag{2}$$

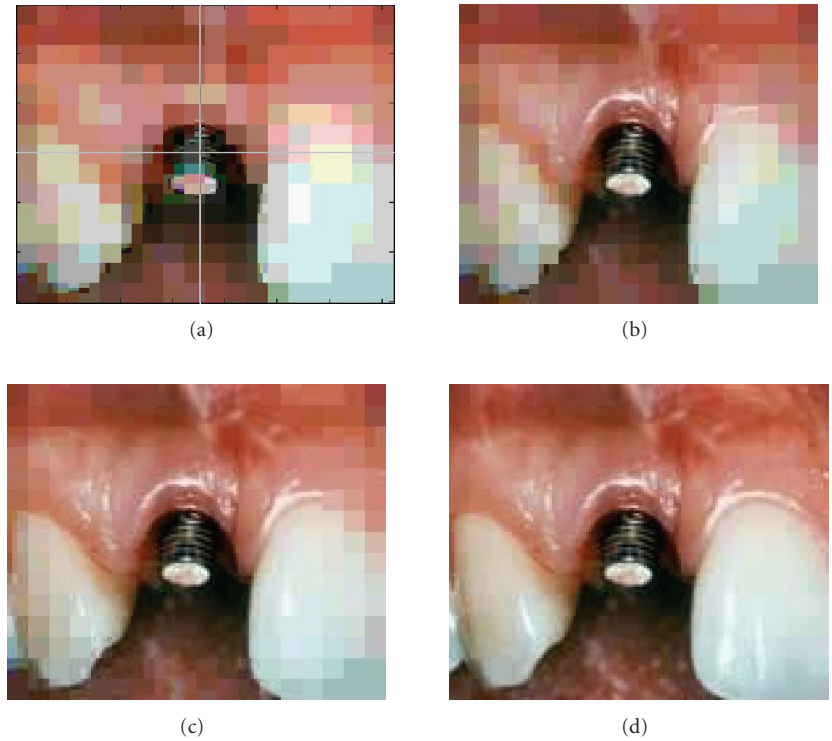
where $B(t)$ is the output bitstream of EZW based on the threshold t. t_{0i} is the initial threshold in the ith subblock, and n_{Ti} is the number of times the threshold is halved in the ith subblock. Therefore, potentially, we can achieve any resolution in the reconstructed images through setting the initial threshold, and the number of times the threshold is halved. As an example, Figure 2 shows the three regions of RoI, R_1, R_2 centered at point (x, y). In Figure 2, the resolution in area R_2 is the lowest and that of RoI is the highest. Therefore, the quality of the reconstructed subblocks and consequently the compression rate depends on the size of the embedded zerotree. This is set based on the distance from the center of RoI. Based on the assigned parameters for EZW, the data in each subblock would be compressed with different rates depending on the location of the subblock. Often, the physician is only interested in a particular part of the image. Therefore, the system is designed in such a way to enable changing the location and size of the RoI without any emphasis on the other regions. In this example, the RoI, R_1, and R_2 may be defined as

for RoI:

$$\sqrt{(i - x)^2 + (j - y)^2} \leq r_a, \tag{3}$$

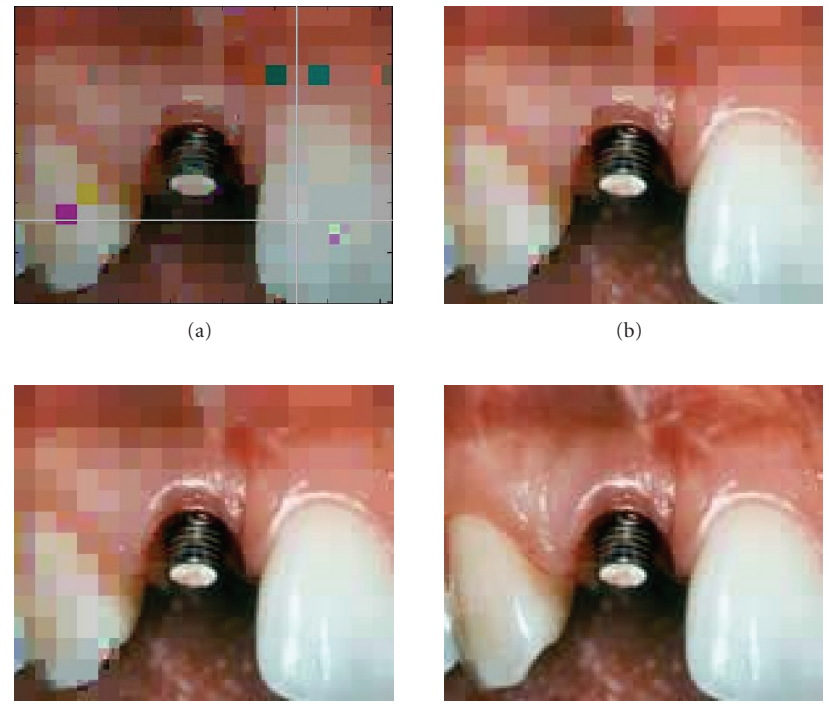
Table 1: Definition of symbols.

Symbols	Description
p	For $\gamma(x, y) \geq t_n$, called *significant coefficients* at threshold t_n.
n	For $\gamma(x, y) < 0$ & $\lvert\gamma(x, y)\rvert \geq t_n$, called *negative significant*.
z	For $\gamma(x, y) < t_n$, but some of its descendants have a value greater than t_n, called *isolated zero*.
t	For $\gamma(x, y) < t_n$, and all its descendants have magnitudes less than t_n, called *zerotree zero*.
1,0	Refinement bits for reconstructing image.



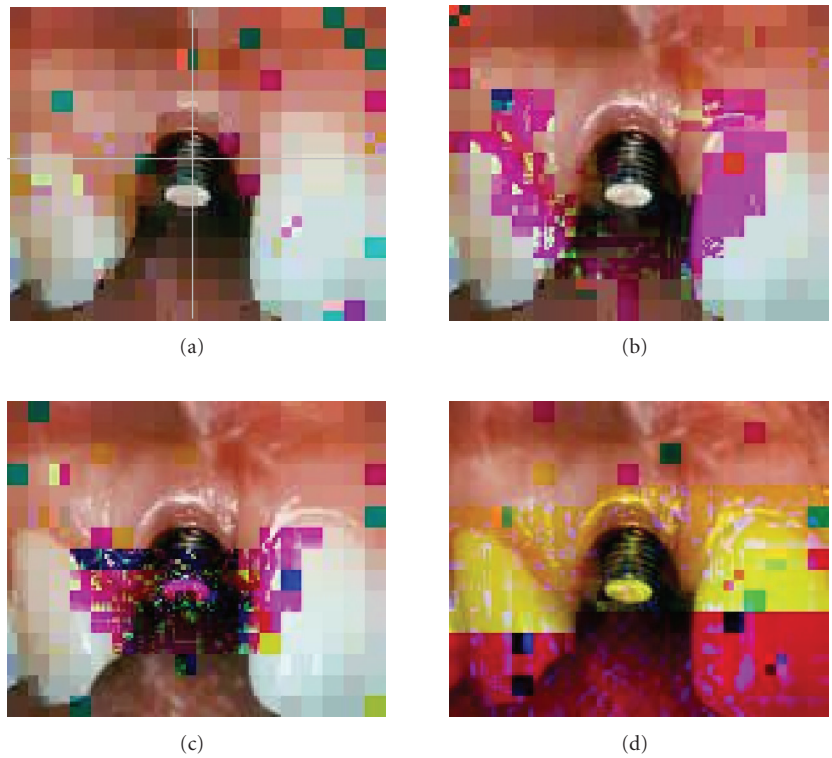
(a) (b) (c) (d)

Figure 9: The transmitted image over the low noise: (a) the background image at stage P_1 and the location of RoI in the center of the image, (b) the transmitted image after stage P_2, (c) the transmitted image after stage P_3, and (d) shows the completely decoded image after stage P_4.



(a) (b) (c) (d)

Figure 10: Similar results as in Figure 9 when the RoI is selected in the corner of the image.



(a) (b) (c) (d)

Figure 11: An example of quality of the decoded image with fixed length parity: (a) background image with the location of RoI in the center, (b) the image reconstructed at stage P_2; several subblocks in error in the area of RoI and some error subblocks in the vicinity of RoI, (c) the number of subblocks in error increases when the volume of the data in the receiver increases, that is, the resolution of the higher-priority regions increases, and (d) the complete transmitted image.

for R_1:

$$r_a < \sqrt{(i-x)^2 + (j-y)^2} \leq r_b, \tag{4}$$

for R_2:

$$\sqrt{(i-x)^2 + (j-y)^2} > r_b, \tag{5}$$

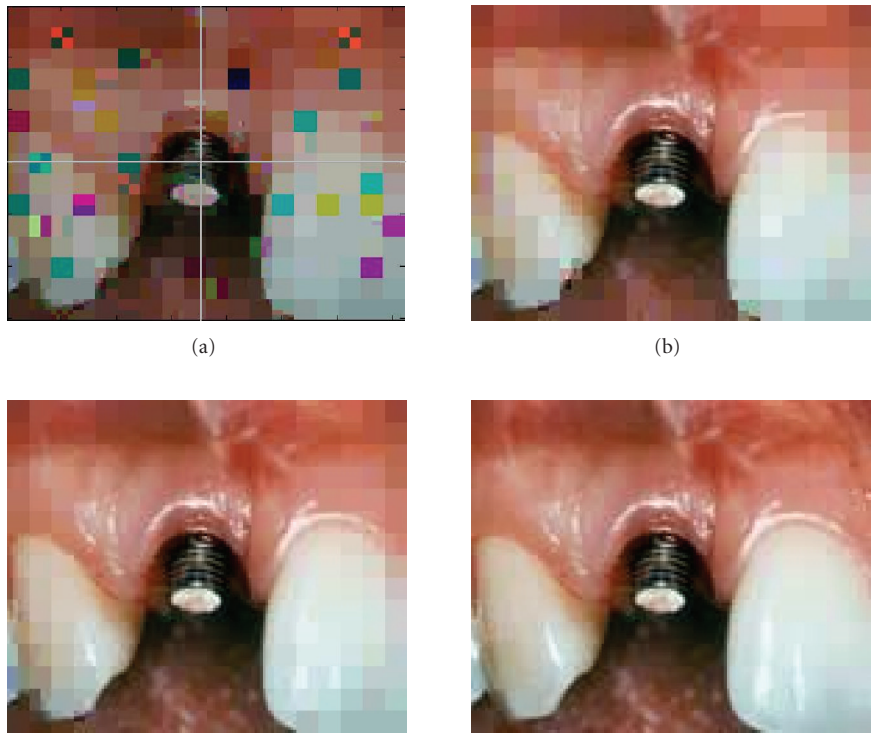
where x and y are the coordinates of the center of RoI assigned by the physician using a mouse click, thereafter both values are sent back to the host. The values of r_a and r_b are the radius of RoI and R_1, and i and j express the $x - y$ image pixels' coordinates. In the successive progressive stages, the values of r_a and r_b gradually expand in the next progressive transmission to create the reconstructed image of higher quality. In our proposed algorithm, the first transmitted image is the background low-resolution image. Then, the reconstruction is progressively continued regarding the highest priority of RoI.

The progressive bitstream is heavily affected by the channel noise; a single bit error may make the reconstruction of the image impossible [10]. The adaptive joint source-channel coding and the developed blind technique for evaluation of errors in the receiver are the effective ways to address this problem. The channel coding variables can be adjusted by using the feedback information about the channel state and the image region content. The RS code is a subset of BCH codes; they are linear block codes and are efficient for bursty-type transmission channels. The RS codes are constructed by considering a polynomial for the input information and then use the polynomial coefficients instead of the original data for transmission. In an RS(p, q) code, $p = 2^m - 1$ is the number of symbols in a codeword, and m is the number of bits in each symbol. In the proposed algorithm, the data in different image regions, as denoted in Figure 2 for three regions, are protected by variable length parity codes as for the UEP. The data in the RoI is treated as the most important information and protected by longer length parity codes. The rest of the data is protected by shorter parity codes.

Figure 3 illustrates the channel coding strategy and Figure 4 shows the receiver. The overall channel code length remains fixed and the length of message k, q_k; $k = 1, 2, \ldots, n$, and the parity length C are variable. For RS codes, $(255 - qn)/2$ indicates the error-correction capability of RS coder. Here, the RS codes, RS$(255, q)$, have 255 symbols in length. According to the UEP, the ratio of parity to overall code length for the n regions should follow

$$C_{\text{RoI}} > C_{R_1} > C_{R_2} \cdots > C_{R_{n-1}}, \tag{6}$$

where C_{RoI}, the length of parity code, is for the RoI and so on. Furthermore, the length of parity code is also affected by the noise in the channel, that is, $C_{\text{region}} \sim (r, N)$, where r is the



(a) (b) (c) (d)

Figure 12: The decoded image with variable length of parity codes over the noisy channel. (a) A background image and the location of RoI selected in the center of image reconstructed after stage P_1, (b) the reconstructed image after stage P_2, no error subblocks are found in the reconstructed image because the lengths of parity codes are adjusted automatically based on the previous volume of incorrect data in the receiver, (c) the reconstructed image after stage P_3; the lengths of parity bits in stage 3 are as same as in stage 2 because no incorrect data was found in the reconstructed image after stage 2, (d) the complete transmitted image with no error subblocks.

distance from the center of RoI, and N expresses the noise in the practical transmission channel.

An adaptive variable parity allocation requires the error between the transmitted image $I(x, y)$ and the received image $\hat{I}(x, y)$ to be minimized under the desired conditions. Suppose that the error is defined as

$$\varepsilon = ||I(x, y) - \hat{I}(x, y)||_2, \tag{7}$$

where $\|\cdot\|_2$ denotes the I_2-norm. Generally, we wish to have the optimum parity length such that

$$C_{\text{opt}} = \min_{C} \varepsilon \text{ subject to } \varepsilon \leq \varepsilon_T, \tag{8}$$

where ε_T is an acceptable error level in the receiver. According to the above discussion, the parity length can be defined as

$$C = g\left(r, \frac{S}{N}\right) = f(r, \text{BER}), \tag{9}$$

where S/N and BER stand, respectively, for signal-to-noise ratio and bit-error rate (the BER here represents the noise situation in the channel and does not refer to the output bit-error rate). The functions g and f are generally nonlinear functions, which can be defined empirically based on a number of trials. From Figure 5, $f \sim (\alpha_0 - \alpha r)$ for a fixed BER, where r is measured with respect to the number of pixels, and from Figure 6, it can be concluded that $f \sim (\beta\,\text{BER} + \beta_0)$ for fixed proximity distance r. In more general applications, a more accurate and possibly complicated function may be adopted. Therefore, a reasonably accurate function can be modeled as

$$f(r, \text{BER}) = (\alpha_0 - \alpha r)(\beta\,\text{BER} + \beta_0), \tag{10}$$

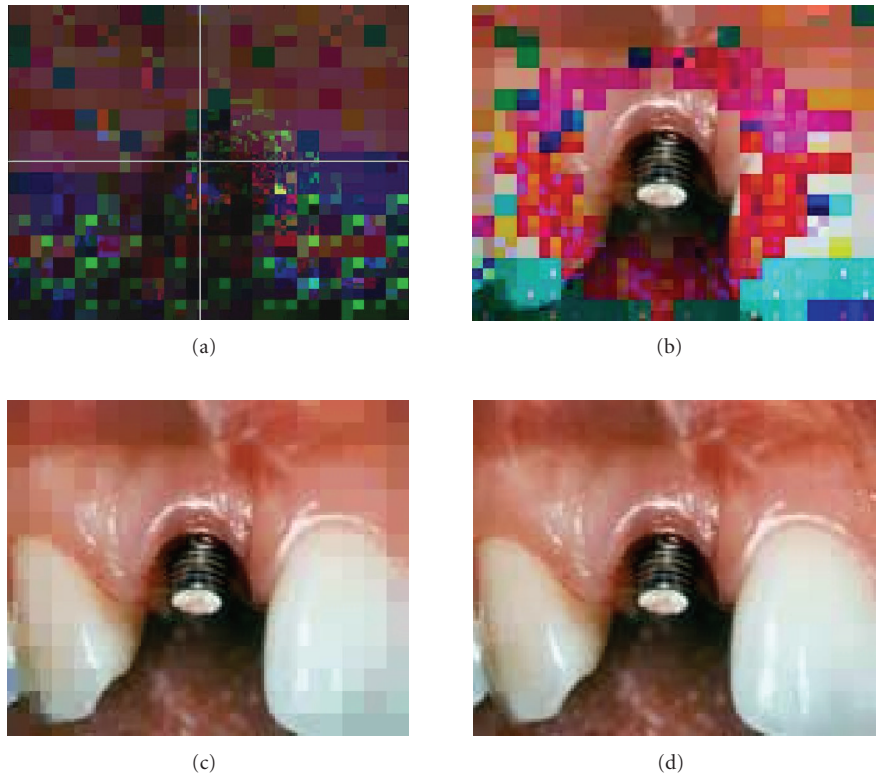
or

$$f(r, \text{BER}) = \mu_3 \text{BER} - \mu_2 r \cdot \text{BER} - \mu_1 r + \mu_0, \tag{11}$$

where μ_is are constant coefficients and can be easily found based on Figures 5 and 6.

4. Detection of the Blocks in Error

Embedded coding has many advantages, especially for progressive image/video transmission, because the reconstructed images can be decoded at any bitrate. However, it is highly sensitive to transmission noise and frequently collapses if even a single transmitted information bit is incorrectly decoded. In most cases, the receiver sends back a signal called HARQ to the transmitter requesting for retransmission under the new constraints. Here, the channel noise level has to be somehow estimated in the receiver. This



(a) (b) (c) (d)

Figure 13: A decoded image with variable length of parity codes over noisy channel: (a) several error subblocks are detected in stage 1, (b) several error subblocks still are found in stage 2 indicating that the feedback message is incorrect or the channel condition becomes noisier. However, RoI is still error-free based on the UEP. (c) No error subblock is detected in stage 3 because the length of parity is adjusted again according to previous channel state, although there can still be some error. (d) The complete transmitted image with no error subblocks.

will enable the change in the parity length in the following transmissions. The estimation process is blind since there is no a priori information about the channel. In some practical cases however, a test image can be transmitted occasionally and evaluated in the receiver.

In the proposed scheme, we have developed an algorithm to detect the corrupted data in the receiver. This algorithm detects the address in which the number of symbols for each subblock is indicated. The algorithm reassigns the number of symbols to each subblock according to a built-in decision making criterion (policy) when the number of symbols within the received data is determined incorrectly by the receiver. Since the header information provides the number of symbols per each data package, an extra check can be carried out to ensure that the header is divisible by 4 and is not greater than 64 (to address each pixel) for the 8×8 subblocks. For more accuracy, the algorithm checks the next address of the number of symbols to ensure that the current data is correct. If successive error data is detected, the algorithm is able to determine the number of subblocks and reassign the number of symbols for each subblock in the set of detected incorrect data. Although, the proposed algorithm does not have capability of recovery of the current corrupted package, but it can conjecture an appropriate number to replace the corrupted data based on the built-in decision making criterion (policy) to avoid the reconstructed image to collapse. This adjusts the system for the transmission of the subsequent package and prevents propagation of the error. This not only enhances the quality of the reconstructed image but also provides a feedback for the adjustment of parity length.

The performance of the system in different noise-levels is evaluated using the peak signal-to-noise ratio (PSNR) defined as

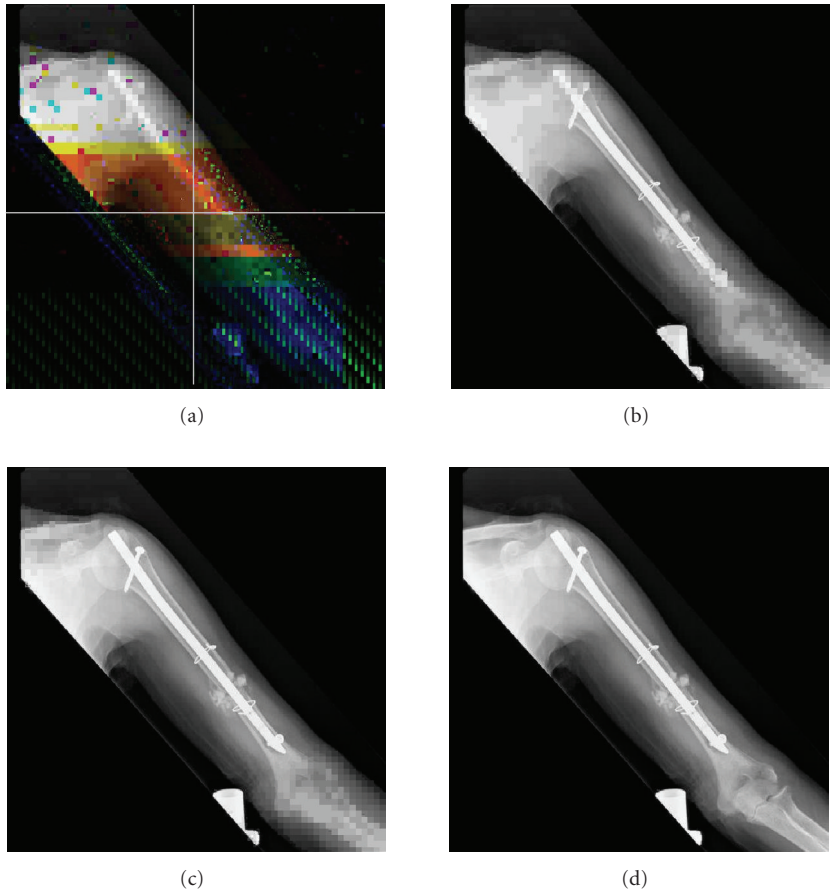
$$\mathrm{PSNR(dB)} = 10 \log_{10} \frac{A^2}{\mathrm{MSE}}, \tag{12}$$

where A is the maximum image amplitude, and MSE is defined as

$$\mathrm{MSE} = \frac{\|\hat{I} - I\|_2^2}{\|I\|_2^2}, \tag{13}$$

where $\hat{I}$ represents the reconstructed subblock of the image, and I is the subblock of the original input image.

Finally, the parallel structure of the channel coding and decoding scheme is very suitable in hardware implementation of the system. A number of parallel boards can be easily used in order to speed up the overall process.



(a) (b) (c) (d)

Figure 14: As another example, a decoded image with variable length parity, which is a 508×512 pixels monochrome X-ray bone image: (a) a background image and the location of RoI selected in the center of image reconstructed after stage P_1; many subblocks are in error in the background image, (b) the reconstructed image after stage P_2, no error subblock is found in the reconstructed image. (c) The reconstructed image after stage P_3, and (d) the complete transmitted image having no error subblocks.

5. Simulation Results

In this work, the proposed system is simulated in both binary symmetric channel (BSC) and flat-fading Rayleigh channel. The BSC is the simplest channel model, only zeros and ones are conveyed in the channel. Therefore, we can simplify the analysis and enable a fast software implementation. But the wireless mobile communication channels are often considered to be with flat-fading Rayleigh noise. In this paper, we simulated both BSC and flat-fading channel models and tested the performance of the proposed techniques against both models. RS$(255, q)$ is used in the proposed scheme. The RS codes correct the symbol error and not the bit error. The noise in the simulated channel has been considered in such a way to set a BER of 0.01 in the received end. For when the errors are uniformly distributed, the average parity length is 42 for a 255 length code length. This recovers the RoI perfectly when either BSC or Rayleigh channel is considered and the channel noise is equivalent to BER = 0.01.

The developed algorithm has been tested for a number of images, two of which are analyzed here. The first image is a 150×123 pixels color dental implant image, and the second image is a 508×512 pixels monochrome X-ray bone image. Both are to be transmitted via TCP/IP. The proposed system follows the diagram in Figure 1. Each noisy channel involves a binary symmetric channel (BSC) and flat-fading Rayleigh channel with a certain BER. In this simulation, the error bits are generated by randomly inverting a certain percentage of bits in the EZW bitstream. To verify the effectiveness of the system, the image regions are progressively transmitted through four stages of P_1, P_2, P_3, and P_4. During P_1, the background image is transmitted. P_2 and P_3 are the second and third stages, both for transmission of RoI, R_1, and R_2. P_4 is the fourth stage mainly to transmit the details of the RoI (and the rest of the image if necessary). In the approach presented here, firstly, the user (physician) specifies the address of the transmitted medical image in the transmitter within the dialogue-based software. After receiving the low-resolution background image, the user identifies the center of RoI by a mouse and the radius of RoI by entering a value within the dialogue-based software. Then, the algorithm adjusts the length of the parity codes initially proportional to the proximity of the image regions to the center of RoI as C_0, $C_0 - 2$, $C_0 - 4$ for RoI, R_1, and R_2, respectively.

Table 2: The information percentage transmitted for each area during the stages P_1 to P_4 images.

Percentage of area	P_1	P_2	P_3	P_4
RoI	0%	18.15%	41.86%	100%
R_1	0%	37.04%	44.07%	0%
R_2	100%	44.81%	14.07%	0%

Table 3: The average compression ratio for various regions for the four stages of the progressive image transmission.

Compression ratios	P_1	P_2	P_3	P_4
Overall image	3.063	1.069	0.716	0.405
RoI	0	0.353	0.401	0.405
R_1	0	1.232	1.814	0
R_2	3.063	3.746	3.387	0

Accordingly, the receiver detects and counts the packages in error by estimating the status of the channel. The parity lengths remain the same if the distortion in the reconstructed image is acceptable. Otherwise, the codes will be adjusted automatically. Typically, the ratio of parity code to the overall code length is larger for the clinically higher priority areas, that is, the areas closer to the center of RoI as stated in (6). Figure 5 indicates the ratios of parity length and the overall codeword according to the experimental results.

Table 2 gives an example of the percentage of information for the regions R_1, R_2, and R_2 for fixed proximity levels of r_a and r_b, as in Figure 9. Table 3 indicates the average compression ratios for various regions before the channel coding during the four successive transmission stages. The compression ratio is defined as the ratio between the data volume of the coded data and the original data. However, by changing one of RoI coordinates, or r_a and r_b, data in Tables 2 and 3 are also changed.

In Figure 6, the parity lengths are found by averaging the results of 10 trials under various noise levels. These are estimated by the algorithm developed for detection of the blocks in error in the receiver. Data in RoI is the most important data in the overall image; therefore, the length of parity codes is longer than that in R_1 and R_2. In the proposed system, the codeword length of RS codes is 255, and the number of error bit is generated at random.

Figure 7 shows the frequency of the set of parity lengths in 10 trials for when the channel noise is set by BER = 0.003 equivalent to the occurrence of 7 errors. As long as the error in the receiver remains higher than a threshold t_h, the length of parity increases. Consequently, if the error falls below a level $t_l < t_h$ the parity length increases. These threshold levels are empirically selected by following the constraint in (8). In these cases, the parameters in (10) and (11) are approximately $\alpha_0 = 0.08$, $\alpha = 2 \times 10^{-4}$, $\beta_0 = 5$, $\beta = 5 \times 10^3$ and accordingly $\mu_0 = \alpha_0\beta_0 = 0.4$, $\mu_1 = \alpha\beta_0 = 10^{-3}$, $\mu_2 = \alpha\beta = 1$, and $\mu_3 = \alpha_0\beta = 4 \times 10^{-2}$.

Figure 8 shows the PSNRs for successive transmission of four stages under various noise-level conditions. Figure 9(a) shows the background image sent during P_1 stage. Figure 9(b) is progressively reconstructed image after stage P_2 in which the RoI, R_1, and R_2 are reconstructed. At this stage, the center of RoI is denoted by the user via mouse click. Figure 9(c) represents the reconstructed image at stage P_3 during which the regions RoI, R_1, and R_2 are reconstructed. The regions of RoI and R_1 are gradually increased in resolution. Figure 9(d) is the final and complete image after stage P_4. The same procedure can be followed for encoding and transmission of any other medical image. However, the coordinates of the center of RoI as well as the size of RoI maybe adjusted according to the requirement by the user. For example, in Figure 10, the RoI is selected in the corner. Figure 11 demonstrates that a fixed-size parity code is not suitable for an efficient transmission system. However, the system has been modified based on the proposed method in Sections 2 and 4 to allow variable lengths of parity. Figures 12 and 13 show no error in the RoI stating that the overall system has been remarkably improved. In Figure 14, another example of a decoded image (a 508×512 monochrome X-ray bone image) is given, and the variable length parity has been examined. The background image suffers from heavy noise. However, the transmission can continue until the last stage during which a complete error-free image is reconstructed.

6. Conclusion

In this paper, we presented a new adaptive source-channel coding with feedback for the progressive transmission of medical images. The system is adaptive to both image content and channel specifications. However, this application is merely for wireless channels (generally narrowband). The capability of data error detection and correction with automatic adjustment, low image transmission time, and efficient communication are the key features in this proposed system. Therefore, the length of parity codes can be adjusted automatically based on the location of the image subblocks and the practical characteristics of the communication channel to provide an adequate data protection. The overall code length for the channel encoder/decoder is fixed. This makes it easy for hardware implementation. A wide range of fluctuations in the channel characteristics (mainly noise level) can be tolerated in the system. The algorithm of detection of subblock in errors can detect the packages in error in the receiver and feed it back to the transmitter for adjustment of the parity length. The proposed system has also been tested for the communication channels with different capacities and noise levels. The presented results verify the effectiveness of the system in terms of both adaptivity and flexibility of interaction. A MATLAB-based TCP/IP connection has been established to demonstrate the proposed interactive and adaptive progressive transmission system. Although some theoretical results are comparable to those of other new techniques such as the UEP or rate allocation approaches, this contribution provides a practical, flexible, and interactive method which suits medical applications.

References

[1] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, 1993.

[2] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions of Image Processing*, vol. 1, no. 2, pp. 205–220, 1992.

[3] C. Chapman, S. Sanei, R. Dilmaghani, and F. Said, "Progressive transmission of medical images using embedded Zertree wavelet encoding," in *Proceedings of IEEE International Workshop on Biomedical Circuits and Systems (BioCAS '04)*, pp. S3.3-9–S3.3-12, Singapore, December 2004.

[4] R. S. Dilmaghani, A. Ahmadian, M. Ghavami, and A. H. Aghvami, "Progressive medical image transmission and compression," *IEEE Signal Processing Letters*, vol. 11, no. 10, pp. 806–809, 2004.

[5] T. Muzaffar and T.-S. Choi, "Embedded linked significant tree wavelet image coder," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, vol. 2, pp. 461–464, Toulouse, France, May 2006.

[6] A. A. Moinuddin and E. Khan, "Wavelet based embedded image coding using unified zero-block-zero-tree approach," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, vol. 2, pp. 453–456, Toulouse, France, May 2006.

[7] Y.-C. Hu and J.-H. Jiang, "Low-complexity progressive image transmission scheme baesd on quadtree segmentation," *Real-Time Imaging*, vol. 11, no. 1, pp. 59–70, 2005.

[8] M.-N. Wu, C.-C. Lin, and C.-C. Chang, "A fast wavelet-based image progressive transmission method based on vector quantization," in *Proceedings of the 1st International Conference on Innovative Computing, Information and Control (ICICIC '06)*, vol. 2, pp. 677–680, Beijing, China, August-September 2006.

[9] R. Hamzaoui, V. Stanković, and Z. Xiong, "Optimized error protection of scalable image bit streams [advances in joint source-channel coding for images]," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 91–107, 2005.

[10] Z. Wu, A. Bilgin, and M. W. Marcellin, "Joint source/channel coding for image transmission with JPEG2000 over memoryless channels," *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1020–1032, 2005.

[11] W. Liu and D. G. Daut, "Progressive image transmission based on joint source-channel decoding using adaptive sum-product algorithm," *EURASIP Journal on Image and Video Processing*, vol. 2007, Article ID 69805, 9 pages, 2007.

[12] H. Yousefi'zadeh, H. Jafarkhani, and F. Etemadi, "Distortion-optimal transmission of progressive images over channels with random bit errors and packet erasures," in *Proceedings of Data Compression Conference (DCC '04)*, pp. 132–141, Snowbird, Utah, USA, March 2004.

[13] J. Chebil, B. Boashash, and M. Deriche, "Combined source channel decoding for image transmission over fading channels," in *Proceedings of the 9th Asia-Pacific Conference on Communications (APCC '03)*, vol. 1, pp. 297–301, Penang, Malaysia, September 2003.

[14] M. Srinivasan and R. Chellappa, "Joint source-channel subband coding of images," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '97)*, vol. 4, pp. 2925–2928, Munich, Germany, April 1997.

[15] P. G. Sherwood and K. Zeger, "Progressive image coding on noisy channels," in *Proceedings of Data Compression Conference (DCC '97)*, pp. 72–81, Snowbird, Utah, USA, March 1997.

[16] X. Pan, A. H. Banishashemi, and A. Cuhadar, "Progressive transmission of images over fading channels using rate-compatible LDPC codes," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3627–3635, 2006.

[17] A. Nosratinia, J. Lu, and B. Aazhang, "Source-channel rate allocation for progressive transmission of images," *IEEE Transactions on Communications*, vol. 51, no. 2, pp. 186–196, 2003.

[18] V. M. Stanković, R. Hamzaoui, Y. Charfi, and Z. Xiong, "Real-time unequal error protection algorithms for progressive image transmission," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, pp. 1526–1535, 2003.

[19] Z. Liu, M. Zhao, and Z. Xiong, "Efficient rate allocation for progressive image transmission via unequal error protection over finite-state Markov channels," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4330–4338, 2005.

[20] V. Chande, N. Farvardin, and H. Jafarkhani, "Image communication over noisy channels with feedback," in *Proceedings of IEEE International Conference on Image Processing (ICIP '99)*, vol. 2, pp. 540–544, Kobe, Japan, October 1999.

[21] J. Lu, A. Nosratinia, and B. Aazhang, "Progressive joint source-channel coding in feedback channels," in *Proceedings of Data Compression Conference (DCC '99)*, pp. 140–148, Snowbird, Utah, USA, March 1999.

[22] V. Chande, N. Farvardin, and H. Jafarkhani, "Image communication over noisy channels with feedback," in *Proceedings of IEEE International Conference on Image Processing (ICIP '99)*, vol. 2, pp. 540–544, Kobe, Japan, October 1999.

[23] F. Zhai, Y. Eisenberg, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Rate-distortion optimized hybrid error control for real-time packetized video transmission," *IEEE Transactions on Image Processing*, vol. 15, no. 1, pp. 40–53, 2006.

[24] C. Valens, "EZW encoding," http://pagesperso-orange.fr/polyvalens/clemens/ezw/ezw.html.

Hindawi Publishing Corporation
International Journal of Telemedicine and Applications
Volume 2009, Article ID 987046, 13 pages
doi:10.1155/2009/987046

Research Article

Delay Analysis of GTS Bridging between IEEE 802.15.4 and IEEE 802.11 Networks for Healthcare Applications

Jelena Mišić[1] and Xuemin (Sherman) Shen[2]

[1] *Department of Computer Science, University of Manitoba, Winnipeg, MB, Canada R3T 2N2*
[2] *Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1*

Correspondence should be addressed to Jelena Mišić, jmisic@cs.umanitoba.ca

Received 16 June 2008; Accepted 18 September 2008

Recommended by Yang Xiao

We consider interconnection of IEEE 802.15.4 beacon-enabled network cluster with IEEE 802.11b network. This scenario is important in healthcare applications where IEEE 802.15.4 nodes comprise patient's body area network (BAN) and are involved in sensing some health-related data. BAN nodes have very short communication range in order to avoid harming patient's health and save energy. Sensed data needs to be transmitted to an access point in the ward room using wireless technology with higher transmission range and rate such as IEEE 802.11b. We model the interconnected network where IEEE 802.15.4-based BAN operates in guaranteed time slot (GTS) mode, and IEEE 802.11b part of the bridge conveys GTS superframe to the 802.11b access point. We then analyze the network delays. Performance analysis is performed using EKG traffic from continuous telemetry, and we discuss the delays of communication due the increasing number of patients.

1. Introduction

Wireless sensor networks in healthcare applications require small lightweight devices with sensing, computational, and communication features to be unobtrusively placed on patient's body. They need to communicate results of sensing of healthcare data periodically over very short range to the devices which can be carried by the patient or mounted on patient's bed. We refer to the wireless sensor network on patient's body as body area network (BAN). Low power, that is, short communication range of wireless sensors is needed for health, interference, and security reasons. Devices which collect the results of measurements need to provide some limited data processing and aggregation, add security/privacy functions, and communicate aggregated data to the LAN access point in patient's ward room as shown in Figure 1. We refer to this device as bridge.

In this paper, we consider interconnection of IEEE 802.15.4 beacon-enabled network cluster with IEEE 802.11b network. The IEEE 802.15.4 nodes comprise patient's body area network (BAN) and are involved in sensing some health-related data which will be transmitted to the access point in the ward room using wireless technology such as IEEE 802.11b. It is clear that the network performance depends on the characteristics of the interconnection device. Therefore, we model the interconnected network where IEEE 802.15.4-based BAN operates in guaranteed time slot (GTS) mode, and IEEE 802.11b part of the bridge conveys GTS superframe to the 802.11b access point. We then analyze the impact of important parameters such as acceptable load ranges and the delays of communication due the increasing number of patients. Since real-time operation of the bridges is necessary for many measurements in hospitals besides intensive care units (ICUs), we will study communications through the bridge for simple case of EKG continuous telemetry.

The remainder of the paper is organized as follows. We review the networking aspect of healthcare wireless sensor networks with the emphasis on continuous electrocardiogram telemetry in Section 2. In Section 3, we review the properties of 802.15.4 beacon-enabled MAC and IEEE 802.11b MAC related to the operation of bridge. In Section 4, we present the concepts of bridge design between IEEE 802.15.4 BAN and IEEE 802.11b ward LAN. Analytical model of bridge where BAN operates in GTS mode is presented

in Section 5. In Section 6, we present performance results for interconnected clusters under various modes of bridge access. Finally, Section 7 concludes the paper.

2. Wireless Sensor Networks for Healthcare

In today's hospitals, there is an urgent need for timely monitoring the health status of many patients, especially those with respiratory and cardiac problems. The need for continuous fetal heart rate monitoring as well as monitoring for movements of patients suffered from stroke or Parkinson's disease is also high. Although these applications require different kinds of sensors to measure levels of oxygen in blood, heart rate, or motion of body parts, there is unified need for sensors to be unobtrusively attached to the patient's body and for measured data to be transmitted in a reliable and secure way in order to be recorded on monitoring devices in real time. Most of the health variables are periodic and have to be periodically sampled and digitized. The sampling period has to be at least twice as large as the highest frequency of the healthcare variables.

Recently, several medical telemetry applications have been prototyped so far and moved to production phase such as pulse oximeters, EKG devices, and motion analysis systems [1]. Wireless transmission is currently implemented using Bluetooth IEEE 802.15.1 and IEEE 802.11b technologies although IEEE 802.15.4 emerges as most suitable for medical applications due to its low-power and low-bandwidth requirements. Some initial work in this area has been reported in [1–4]. However, current reports on wireless healthcare products are focused on reliable hardware and software designs of sensor modules while the wireless transmission has been considered in testing phase and for single device only. There is an area of research involving coordination and real-time transmission of large number of healthcare measurements, which has not attracted sufficient attention. In this paper, we will address the following problems.

(1) Interconnection of low-power IEEE 802.15.4 motes (which are convenient for attachment on patient's body) with IEEE 802.11b network which has larger bandwidth and larger transmission range. We will design bridge between IEEE 802.15.4 wireless communication interface(s) residing at patient's body and IEEE 802.11b residing at bedside or carried in patient's pocket. More specifically, we will use TDMA feature of IEEE 802.15.4 called guaranteed time slots to fill it with digitized samples and pass it to the interface of IEEE 802.11b for further transmission to the hospital room's access point.

(2) Analysis of delay of real-time health measurement data incurred by transmission technologies. We will use measurement data for electrocardiogram and analyze traffic when the number of patients increases.

Security is another important issue in wireless healthcare sensor networks. In this paper, we will also address data integrity issue by assuming that there is a shared secret key between 802.15.4 sensor mote and bridge device. Secret key will be used to provide message authentication code in each 802.15.4 superframe (packet) using HMAC function [5, 6].

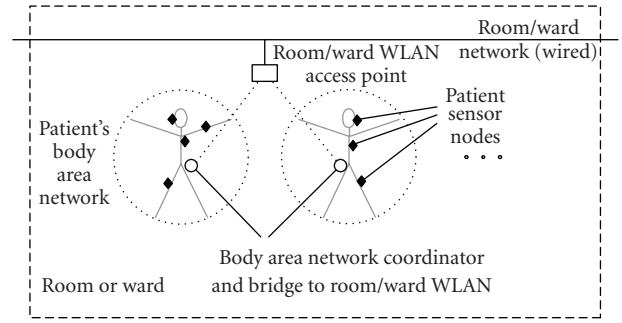


Figure 1: Networking structure of the ward room with BANs, bridges, and ward LAN.

2.1. EKG Measurement. Electrocardiogram (ECG or EKG) is a surface measurement of the electrical potential generated by the electrical activity, which controls pumping action of cardiac muscle fibers. These electrical impulses generate voltage, which further generates current flow in the torso and potential differences on the skin. Standard EKG monitoring involves short-term (≤30 seconds) monitoring of heart pulses using 12 skin electrodes (called leads) placed at designated locations on the patient's body including chest, arms, and legs. Each pair of leads measures voltage which gives one aspect of the heart's activity. An EKG picture produced by 12 leads allows diagnosis of wide range of heart problems. However, this measurement is short-term and requires wired connection between the patient and electrocardiograph.

In many cases, however, it is necessary to have continuous and tetherless measurements of heart rate. For such application, only three leads placed at patient's upper and lower chest can trace a wide range of cardiac arrhythmias. One node of these three collects signals, amplifies the signal difference, samples the amplified analog signal, and digitizes it. Standard clinical EKG application has the bandwidth of 0.05 Hz to 100 Hz. For pacemaker detection, upper frequency can be up to 1 kHz. There are many design issues out of scope of this work related to noise suppression and filtering frequencies from power line and respiration. In this paper, we assume that upper frequency of the EKG signal is 100 Hz. EKG signal is sampled with 200 Hz, and each sample is digitized with 12 bits [3]. Therefore, basic bandwidth of EKG signal in standard continuous telemetry is only 2400 bps.

3. Basic Properties of IEEE Std 802.15.4 and IEEE 802.11b MACs

3.1. Basic Properties of IEEE Std 802.15.4 MAC. In beacon-enabled networks, the personal area network (PAN) coordinator divides its channel time into superframes [7]. Each superframe begins with the transmission of a network beacon, followed by an active portion and an optional inactive

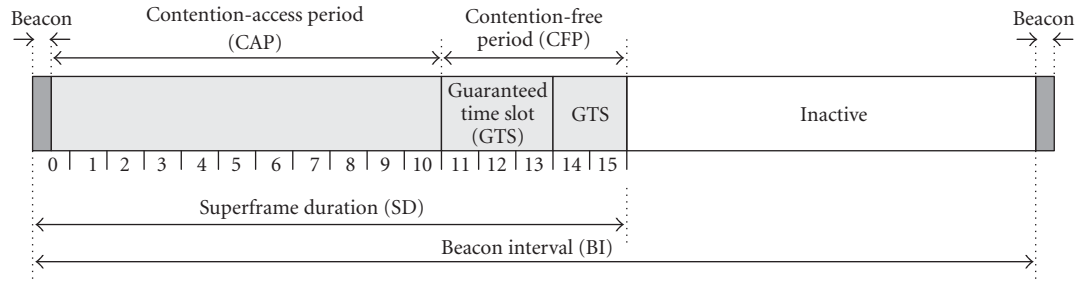


Figure 2: The composition of the superframe under IEEE Std 802.15.4 (adopted from [7]).

portion, as shown in Figure 2. The coordinator interacts with its PAN during the active portion of the superframe, and may enter a low-power mode during the inactive portion. Raw data rate in industrial, scientific, and m edical (ISM) band is 250 Kbps. Basic time unit in the standard is backoff period which contains 10 bytes. Duration of active and inactive parts of the superframe is regulated with MAC parameters $SO = 0,\ldots,14$ which is known as *macSuperframeOrder* and $BO = 0,\ldots,14$, also called *macBeaconOrder*. Active superframe part is divided into 16 slots. Each slot consists of $3\cdot 2^{SO}$ backoff periods, which gives the shortest active superframe duration *aBaseSuperframeDuration* of 48 backoff periods when $SO = 0$. Duration of an active superframe part is denoted as $SD = aBaseSuperframeDuration \cdot 2^{SO}$ (superframe duration). The time interval between successive beacons is equal to $BI = aBaseSuperframeDuration * 2^{BO}$. The duration of the inactive period of the superframe can be determined as $I = aBaseSuperframeDuration * (2^{BO} - 2^{SO})$. Period between the beacons is equal to the active superframe duration only if there is no inactive period in the cluster time, and, otherwise, it is larger than active superframe part, that is, $BO \geq SO$.

An active superframe part consists of contention part and TDMA, that is, guaranteed time slot (GTS) part. GTS bandwidth must be requested by the node using the MAC command frame. Coordinator allocates the GTS bandwidth in multiples of slots. One slot contains $3*2^{SO}$ backoff periods. Data transfer from a node to PAN coordinator can be done in GTS slots or using slotted CSMA-CA access described below. Slotted CSMA-CA algorithm consists of backoff activity, two clear channel assessments (CCAs), packet transmission, and optionally receipt of the acknowledgment. Backoff value is uniformly chosen in the range $(0, w_{15} - 1)$ which is called contention window. During backoff countdown node does not listen to the medium, and checks the activity on the medium only twice when backoff count is finished. By default, the node can have $m_{15} + 1 = 5$ transmission attempts with backoff window sizes $W_{15,0} = 8\,W_{15,1} = 16$, $W_{15,2} = 32$, $W_{15,3} = 32$, and $W_{15,4} = 32$. To avoid confusion, we will use subscript 15 to label MAC parameters which have their counterparts in the IEEE 802.11b standard.

3.2. Basic Properties of IEEE 802.11b Needed for Bridging. IEEE 802.11 has much more sophisticated CSMA-CA scheme at the MAC layer. In this protocol (opposite to IEEE 802.15.4), a station having a packet to transmit must initially listen to the channel to check if another station is transmitting. If there is no transmission in distributed interframe space (DIFS) time interval, the transmission can proceed. If medium is busy, the station has to wait until current transmission has finished. Then, station will wait for DIFS time period and then generate a random backoff time before transmitting its frame. This backoff time is uniformly chosen in the range $(0, w_{11} - 1)$. Backoff counter will be decremented after each *slot time* given that transmission medium is free, otherwise, its value will be frozen until medium becomes free again for DIFS time units (slot time is derived from the propagation delay time to switch from receiving to transmitting mode and time to pass the information about the physical channel state to MAC layer. It actually corresponds to backoff period from IEEE 802.15.4). Station will transmit when its backoff counter reaches zero value. When the packet is received, receiver replies with acknowledgment (ACK) packet after short interframe space (SIFS) time interval. Whenever packet collision occurs, acknowledgment will not be received within SIFS + ACK time, and transmission has to be reattempted with doubled contention window. If starting window size is $w_{11} = W_{11,\min}$ after m_{11} retransmissions, its maximal value will become $W_{11,\max} = 2^{m_{11}} W_{11,\min}$ (in order to distinguish between similar variables in two standards, we use subscript 11). In our model, we assume that if packet experiences more than m_{11} collisions, last backoff stage will be entered for every subsequent retransmission until frame is successfully transmitted. In order to limit packet collision time, and guard against hidden terminal problem, the standard allows small reservation packets request to send (RTS) and clear to send (CTS) sent using CSMA-CA. After transmission of RTS packet, receiver replies with CTS after short interframe space(SIFS) time. Due to sensitivity of healthcare applications, we will assume that RTS/CTS scheme is used to protect packets with measurement data of health variables.

The IEEE 802.11b standard is mostly deployed in current implementations of healthcare wireless sensor networks [1]. The IEEE 802.11b has higher-speed physical layer than original IEEE 802.11 and allows transmissions at 1, 2, 5.5, and 11 Mbps, while IEEE 802.11 supports transmission at 2 Mbps. However, physical layer header is transmitted at 1 Mbps, MAC layer header and payload can be transmitted

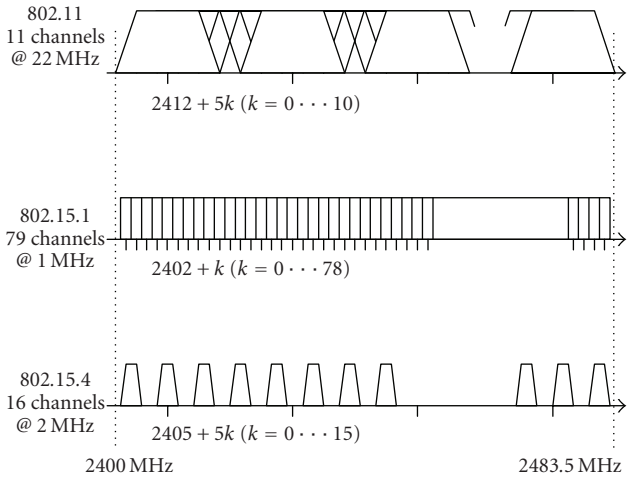


Figure 3: Used spectra of wireless LAN and PAN technologies in ISM band.

at 1, 2, 5.5, or 11 Mbps while control frames RTS, CTS, and ACK are transmitted at 1 or 2 Mbps. It is also worth noting that since IEEE 802.11b adapters transmit at a constant power, distances covered with transmission speeds of 1, 2, 5.5, and 11 Mbps are 120, 90, 70, and 30 m, respectively [8].

Starting with seminal work in [9], performance of IEEE 802.11 has been extensively studied first for saturated case and later for unsaturated case [10–13]. Standard extension IEEE 802.11e enhances CSMA-CA access by introducing different interframe spaces and different backoff window ranges for different traffic classes. This scheme has been modeled using similar approach (although more complex) as basic 802.11 scheme in [14–16]. However, given the fact that all BANs have the same priority and short packet sizes with sensing information, we believe that IEEE 802.11b can serve the purpose and that deployment of IEEE 802.11e at this point may not be necessary.

4. GTS Bridge Design

Bridge consists of IEEE 802.15.4 PAN coordinator and ordinary IEEE 802.11b interface. These two components are interconnected through a buffer which is filled by the PAN coordinator and emptied by IEEE 802.11b packet transmission facility. IEEE 802.15.4 PAN coordinator interface and IEEE 802.11b interface have their wireless transmit/receive antennas. Both networks operate in ISM band as shown in Figure 3, and there is a need to coordinate operation of bridge's interfaces either in TDMA or FDMA manner. From Figure 2 and discussion in Section 3.1, we observe that it is possible to achieve TDMA coordination between the interfaces using the fact that WLAN bridge interface can operate during silent BAN periods. However, in the presence of multiple BANs within WLAN coverage, this approach requires synchronization of BAN beacons (we assume that interference among BANs is avoided by separation in space or by allocating separate channels as shown in Figure 3). Also bandwidth allocation through *SO* and *BO* parameters must be achieved such that bridged traffic from all BANs can be delivered during (common) inactive superframe part.

It is much more convenient if operation of BAN and WLAN is separated in frequency domain because BAN beacons do not need to be synchronized and more bandwidth is allocated to the bridge. BAN channels should be chosen in such way that they do not overlap with ward WLAN channel. From Figure 3, we see that each WLAN channel overlaps with four BAN channels. Therefore, for each ward WLAN channel, 12 out of 16 BAN channels should be used in order to avoid interference.

Duration of beacon interval *BI* is tuned according to period of sensed health variable. Duration of active period *SD* is chosen in order to

(1) achieve data transmission with high success probability in the case of CSMA-CA MAC,

(2) in case of GTS transmission of sensed data, GTS bandwidth has to match necessary size of a group of GTS packets and acknowledgment lanes where group size corresponds to the number of IEEE 802.15.4 nodes in BAN. However, some small contention periods must be reserved in the superframe in order to communicate command frames between sensing nodes and PAN coordinator.

Since bandwidth of IEEE 802.11b is much larger than the bandwidth of IEEE 802.15.4 BAN, transmission of one IEEE 802.11b packet will take short time, and rest of active period can be used to exchange some command data between the bridge and access point of the ward room. Between two transmission phases, sensing nodes and bridge are idle and can turnoff their transmitters and receivers. Duration of this sleep time is $BI - SD$. We make an important remark about time scales in IEEE 802.15.4 and IEEE 802.11b. Duration of backoff period in 802.15.4 is 320 microseconds, and with raw data rate of 250 Kbps, one backoff period carries 10 bytes of data. IEEE 802.15.4 slot duration is $3*2^{SO}$ backoff periods. On the other hand, IEEE 802.11b backoff counter is decremented after slot time which is equal to 20 microseconds. Therefore, time-scale translation is needed between two networks.

We assume that traffic intensity due to sensing of health variables in BAN will be light given that the number of nodes is lower than 16, and that offered load per node is lower than few Kbps. This is well below the rate of 250 Kbps supported by IEEE 802.15.4.

Bridge's buffer will be served by IEEE 802.11b CSMA-CA MAC for which the Markov chain model is presented in Figure 4. In our modeling, we will assume that data buffers at BAN nodes and at the bridge are infinite. Although, this assumption may appear unrealistic for sensor nodes (and we have always modeled small finite buffers in our work [17]); offered load to the node is low so that node's buffer is empty most of the time. Therefore, both the queuing model with finite and infinite buffers will give similar results. Given the lower computational complexity of the model with infinite buffer, we use it in our analysis although use of finite buffer model is straightforward. Assumption of infinite buffer at

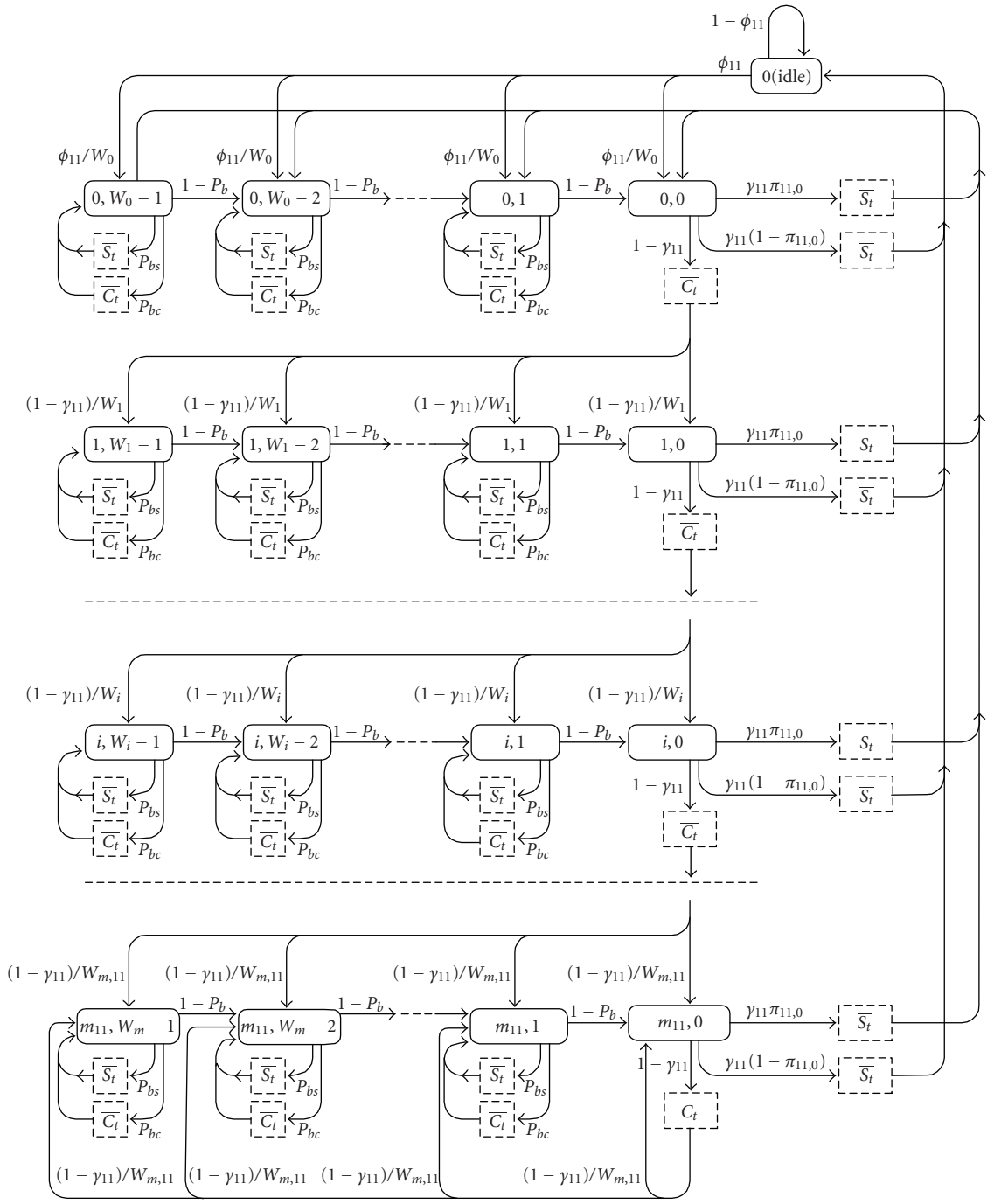


Figure 4: Markov chain for IEEE 802.11b coupled with device's queue.

the bridge is reasonable since it may contain more complex hardware and software, and offered load per bridge is low-to-moderate, depending on the number of patients in vicinity of access point.

5. Analytical Model of GTS Bridging

In general case of GTS, bridge sensors may report different medical variables, such as heart rate, level of oxygen in blood, and temperature. Each sensor is allocated a number of slots in uplink direction to carry uplink sensing data and a slot in downlink direction to carry acknowledgment. We will refer to the number of slots in uplink direction as *packet lane*.

PAN coordinator receives data from stations in separate GTS packet lanes, generates acknowledgments, and passes aggregated packets to IEEE 802.11b buffer.

Assume that each sensor needs to be served with data rate D_s bps. This results in allocation of d_s packet lanes such

Table 1: Comparison of MAC parameters between beacon-enabled IEEE 802.15.4 and IEEE 802.11b.

CSMA-CA parameter	IEEE 802.15.4	IEEE 802.11b
Backoff period size	320 μs	20 μs
Listen to the medium during backoff count	no	yes
Freeze the backoff ctr. when medium is busy	no	yes
Listen to the medium immediately before the trans.	yes	no
Action when medium is sensed busy	New backoff phase	Freeze the backoff ctr.
Typical size of minimum backoff window	8	32
Typical number of backoff phases m	5	5
Raw data rate	250 Kbps	1, 2, 5.5 or 11 Mbps
Typical physical + MAC layer header size	48 + 72 bits	192 + 272 bits
RTS and CTS	No in beacon enabled version	Yes

that

$$D_s = \frac{d_s * 3 * 2^{SO} * 80}{48 * 2^{BO} * 0.00032}, \quad (1)$$

and that period between the beacons $BI = 48 * 2^{BO} * 0.00032$ matches the sampling period of health variable. Assuming that there are $n_{15} < 16$ sensors in the BAN, each sensor will have one GTS slot, and the last slot will contain aggregated acknowledgments for all sensors. For larger number of sensors where $15(k-1) < n_{15} \leq 15k$ period between the beacons has to be decreased k times, such that each superframe carries readings from at most 15 sensors (and acknowledgments in the last GTS slot).

For simplicity, let $n_{15} < 16$ and the payload of the IEEE 802.15.4 superframe becomes payload of IEEE 802.11b frame consisting of $n_{15} \cdot d_s \cdot 3 \cdot 2^{SO} \cdot 10$ bytes. IEEE 802.11b frame length in bytes has to be augmented with headers from physical and MAC 802.11b layer which is 50 bytes. Finally, we have to find frame size in 802.11b slots (backoff periods), since each slot carries number of bits s_{11} equal to the product of raw data rate and slot duration, and its value is equal to $l_{11} = (50 + n_{15} \cdot d_s \cdot 3 \cdot 2^{SO} \cdot 10)/s_{11}$.

Duration of RTS, CTS, and ACK frames expressed in slots will be denoted as *rts*, *cts*, and ack_{11}, respectively (we will use subscript 11 to denote IEEE 802.11b whenever potential ambiguity may arise between two standards). Duration of DIFS and SIFS periods in slots will be denoted as *difs* and *sifs*.

Probability generating function (PGF) for the successful packet transmission time is equal to [12]

$$St(z) = z^{rts+cts+l_{11}+3sifs+difs+ack_{11}}, \quad (2)$$

with average value $\overline{St} = St'(1) = rts + cts + l_{11} + 3sifs + difs + ack_{11}$. In the case of collision of RTS packets activity on medium has PGF:

$$Ct(z) = z^{rts+cts+sifs+difs}, \quad (3)$$

with average value $\overline{Ct} = Ct'(1) = rts + cts + sifs + difs$.

Assume that there are n_{11} bridges attached to the IEEE 802.11b access point. Each bridge communicates the same kind of sensing traffic towards the access point. Using the assumption from previous work [2–5, 9, 14, 15] that probability of successful transmission is independent of the backoff stage, we will denote it as γ_{11} while collision probability is $1-\gamma_{11}$. Access probability is also independent of the backoff stage and is denoted as τ_{11}. Relationship between these two probabilities is

$$\gamma_{11} = (1 - \tau_{11})^{(n_{11}-1)}. \quad (4)$$

Probability that medium will be active during the backoff countdown of one station has two components. First one is the probability that station sensed the medium busy due to successful transmission of one among $n_{11} - 1$ stations, and it has the value $p_{bs} = (n_{11} - 1)\tau_{11}(1 - \tau_{11})^{(n_{11}-2)}$. The other component is the probability that station senses the medium busy due to collision among some of $n_{11} - 2$ other stations and has the value $p_{bc} = 1 - (1 - \tau_{11})^{(n_{11}-1)} - p_{bs}$. Their sum is the probability that medium is busy during the backoff countdown, and that backoff counter is frozen $p_b = p_{bs} + p_{bc} = 1 - \gamma_{11}$. The PGF for the duration of time between two successive backoff countdowns is represented with the following equation [12]:

$$Hd(z) = z\gamma_{11} + (p_{bc}Ct(z) + p_{bs}St(z))Hd(z). \quad (5)$$

At this point, we note that duration of period between two successive decrements of backoff counter is limited to the maximum packet size. After transmission is finished and DIFS period passes any station doing the backoff, countdown has to decrement its backoff counter at least once before packet transmission:

$$B_{11,i}(z) = \sum_{k=0}^{W_{11,i}-1} \frac{1}{W_{11,i}} H_d^k(z) = \frac{H_d^{W_{11,i}}(z) - 1}{W_{11,i}(H_d(z) - 1)}, \quad (6)$$

where $W_{11,i} = 2^i W_{11,0}$ for $i \leq m_{11}$, and $W_{11,i} = 2^{m_{11}} W_{11,0}$ for $i > m_{11}$.

Assuming that packet will be retransmitted until valid acknowledgement is received, the PGF for the packet service time becomes

$$T_{11}(z) = \sum_{i=1}^{m_{11}+1} \left(\prod_{j=0}^{i-1} B_{11,j}(z) \right) (1-\gamma_{11})^{(i-1)} Ct(z)^{(i-1)} \gamma_{11} St(z) + \left(\prod_{j=0}^{m_{11}} B_{11,j}(z) \right) \sum_{i=m_{11}+1}^{\infty} (B_{11,m_{11}}(z))^{(i-m_{11})} \times (1-\gamma_{11})^{(i-1)} Ct(z)^{(i-1)} \gamma_{11} St(z), \tag{7}$$

and its average value is obtained as $\overline{T_{11}} = T'_{11}(1)$.

5.1. Markov Chain Model and Queuing Model for the GTS Bridge's Output. In the derivations above, we have derived probability of successful transmission and probability of freezing the backoff period using the variable which represented access probability per 802.11b slot τ_{11}. Access probability, on the other hand, has to be derived using two modeling components. First one is the Markov chain which represents conditional activities within the CSMA-CA process. Second component indicates the probability that bridge will be idle and that it will not perform backoff count and attempt transmission. This happens only when the bridge's packet buffer is empty. In order to find this probability, we must deploy queuing theory and we have to know the arrival process to the bridge's queue, the size of the bridge's queue, and the probability distribution packet service time by which packets depart from the queue. Therefore, Markov chain model and queuing model of the bridge are coupled and have to be modeled and solved simultaneously. We will first solve the Markov chain for CSMA-CA MAC using the variable $\pi_{11,0}$ which represents the probability that bridge's buffer is empty after the packet departure.

Since similar Markov chain models have been solved with detailed steps of setting transition probabilities in the past in [9–12], we will just state the most important steps in derivation of access probability. Markov chain $\{s(t), b(t)\}$ is discrete as transitions are observed at ends of slot times. It is bidimensional (given that γ_{11} and p_b are independent of backoff stage) where $s(t)$ represents backoff stage and $b(t)$ represents value of backoff counter. Corresponding states of the Markov chain will have the state probabilities $y_{i,j}$, $i = 0, \ldots, m_{11}$, $j = 0, \ldots, W_{11,i} - 1$. Since packet sizes in this application are relatively small, we have included the states when transmission of RTS/CTS and data packets is going on. Access probability is equal to $\tau_{11} = \sum_{i=0}^{m_{11}} y_{i,0}$.

Probability of idle state when bridge's buffer is empty is equal to

$$P_{\text{idle}} = \frac{\tau_{11}\gamma_{11}\pi_{11,0}}{\phi_{11}}, \tag{8}$$

where ϕ_{11} denotes average packet arrival rate of the arrival process to the bridge (note that packet arrival process is not Poisson when packets arrive from IEEE 802.15.4 BAN to the bridge).

By inserting all necessary transition probabilities, we obtain

$$\tau_{11} = \left(\frac{\gamma_{11}\pi_{11,0}}{\phi_{11}} + \sum_{j=0}^{m_{11}} \gamma_{11}(1-\gamma_{11})^j + \sum_{j=0}^{m_{11}} \frac{(W_{11,j}-1)\gamma_{11}(1-\gamma_{11})^j}{2(1-p_b)} (p_{bc}\overline{Ct} + p_{bs}\overline{St}) + \frac{(W_{11,m_{11}}-1)(1-\gamma_{11})^{(m_{11}+1)}}{2(1-p_b)} (p_{bc}\overline{Ct} + p_{bs}\overline{St}) + \gamma_{11}\overline{St} + (1-\gamma_{11})\overline{Ct} \right)^{-1}. \tag{9}$$

5.2. Derivation of Probability Distribution of Occupancy of Bridge's Buffer. As we mentioned, we will assume that bridge's buffer has an infinite capacity. Rationale behind that is that bridge device indeed can have much larger memory than sensing device and that utilization of the bridge is expected to be light-to-moderate. Therefore, bridge is not expected to work in the regime close to its stability limit where it can reject packets due to finite buffer.

In GTS-based bridge period between arrivals of IEEE 802.15.4 superframes to the bridge is constant, and bridge can be modeled as D/G/1 queuing system modeled at Markov points of packet departures from the bridge. Let us denote period of arrival of sensing information from all the sensors to bridge as $\Phi_{11} = 1/\phi_{11}$. As mentioned earlier, if the number of sensors $n_{15} < 16$, then $\Phi_{11} = BI$, and if $15(k-1) < n_{15} \leq 15k$, then $\Phi_{15} = kBI$. Assume that PGF for the bridge's packet service time can be expressed as the series as $T_{11}(z) = \sum_{k=0}^{\infty} t_{11,k} z^k$. Probability of l, $l = 0, 1, \ldots$ arrivals of GTS superframes during packet service time of the bridge has the value

$$a_{11,l} = \text{Prob}[l\Phi_{11} \leq T_{11} < (l+1)\Phi_{11}] = \sum_{j=l\Phi_{11}}^{(l+1)\Phi_{11}-1} t_{11,j}. \tag{10}$$

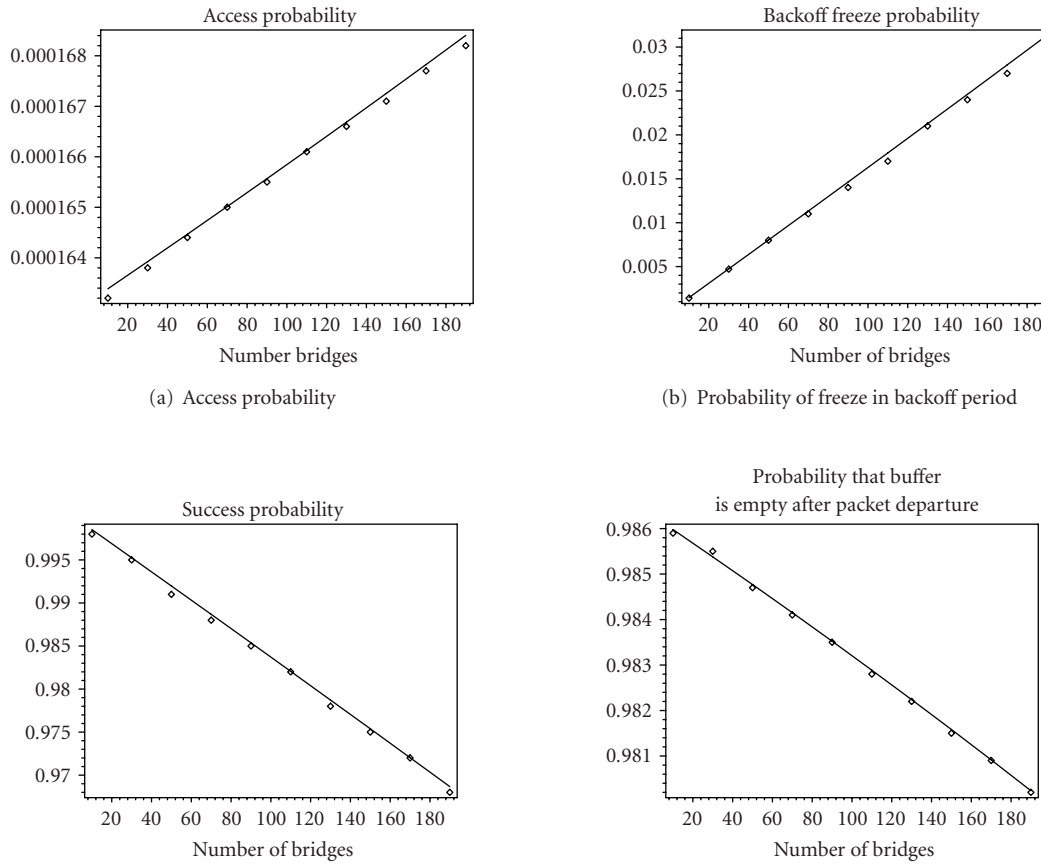
PGF for the probability distribution of the number of 802.15.4 superframe arrivals during packet service time by the 802.11b interface is $A_{11}(z) = \sum_{l=0}^{\infty} a_{11,l} z^l$. An equation which shows number of packets in bridge's buffer left after the departure of the packet has the form

$$\pi_{11,l} = \pi_{11,0} a_{11,l} + \sum_{j=1}^{l+1} \pi_{11,j} a_{11,l-j+1}. \tag{11}$$

By multiplying both left-hand and right-hand side of (11) with z^l and summing over $l = 0, \ldots, \infty$, we obtain PGF for the number of packets left in the bridge's queue after the departing packet $\Pi_{11}(z) = \sum_{l=0}^{\infty} \pi_{11,l} z^l$ as

$$\Pi_{11}(z) = \frac{A_{11}(z)(1-\rho_{11})(1-z)}{A_{11}(z) - z}, \tag{12}$$

where $\rho_{11} = \phi_{11}\overline{T_{11}}$ presents offered load to the bridge.



(a) Access probability

(b) Probability of freeze in backoff period

(c) Probability of successful transmission

(d) Probability that buffer is empty after departure

Figure 5: Access probability, probability that backoff count will be frozen, probability of successful transmission, and probability that buffer is empty after departing packet. Analytical results are shown as lines and simulation results are shown as points.

Table 2: Tradeoff between packetization delay and number of samples carried in each superframe.

BO value	Period between beacons	Number of EKG samples in the superframe
6	0.983s	196
5	0.492s	98
4	0.245s	49
3	0.123s	24

Table 3: Parameters used in the analytical modeling.

Number of bridges	20–200
SO (802.15.4)	0
BO (802.15.4)	3
Raw data rate (802.15.4)	250 Kbps
Superframe size (802.15.4)	480 bytes
MAC and payload data rate for 802.11b	2 Mbps
Payload size of 802.11b packet	10 slots at 2 Mbps
IEEE 802.11 physical + MAC header size	16.4 slots

5.3. Output Process and Throughput. Output process from bridge has the PGF for packet interdeparture times as

$$\Delta_{11}(z) = (1 - \pi_{11,0})\, T_{11}(z) + \pi_{11,0} z^{\Phi_{11}} T_{11}(z). \tag{13}$$

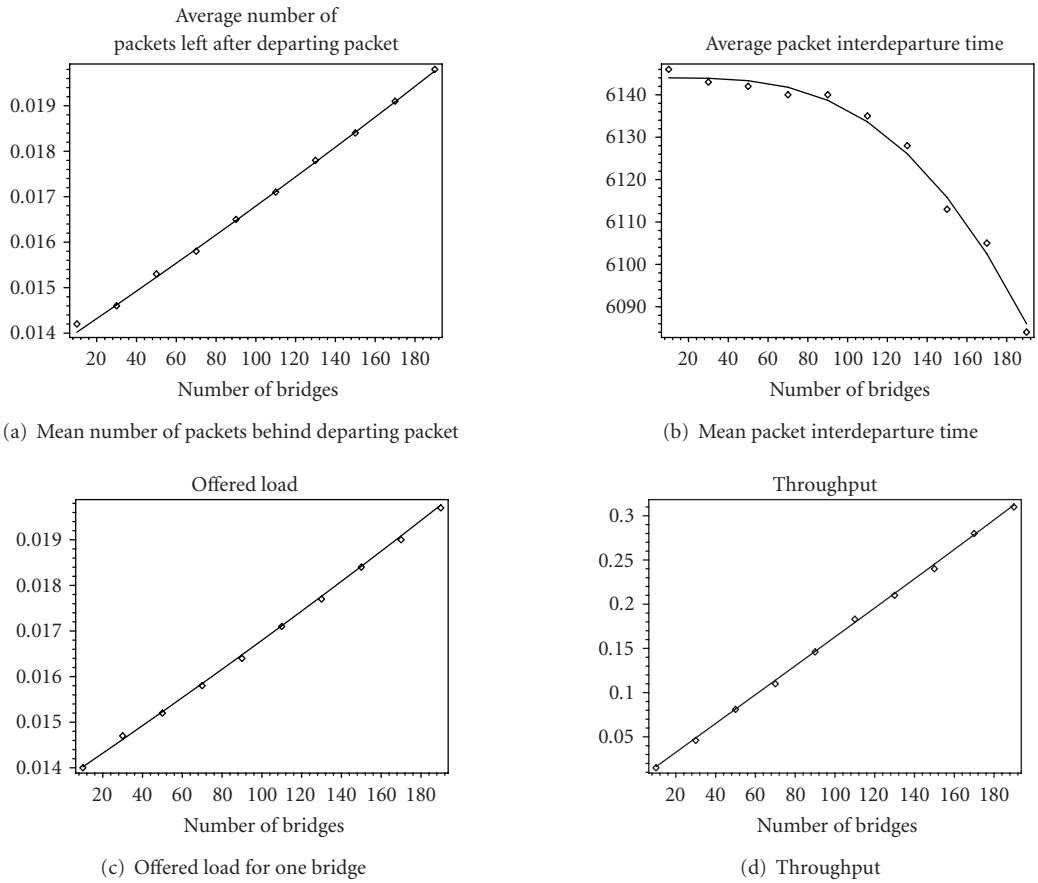
Assuming that there are n_{11} bridges communicating with access point throughput in 802.11b LAN can be then presented with expression

$$\Theta_{11} = n_{11} \frac{l_{11} - 2}{\overline{\Delta_{11}}}, \tag{14}$$

where 2 slots correspond to header information.

5.4. Distribution of Packet Waiting Time in Bridge's Buffer. Assuming FIFO service discipline, packet arriving at the bridge has to wait for the currently transmitted packet to depart and for complete service time of all packets already queued. According to renewal theory, remaining service time of the packet has PGF [18]

$$T_{11}^{+}(z) = \frac{(1 - T_{11}(z))}{\overline{T_{11}}(1 - z)}. \tag{15}$$



(a) Mean number of packets behind departing packet

(b) Mean packet interdeparture time

(c) Offered load for one bridge

(d) Throughput

Figure 6: Mean number of packets left in the bridge's queue after departing packet, mean packet interdeparture time, bridge offered load, and throughput. Analytical results are shown as lines, and simulation results are shown as points.

Then, PGF for the waiting time of the packet has the form

$$\begin{aligned} W_{11}(z) &= \pi_{11,0} + \pi_{11,1}T_{11}^{+}(z) + \pi_{11,2}T_{11}^{+}(z)T_{11}(z) \\ &\quad + \pi_{11,3}T_{11}^{+}(z)\left(T_{11}(z)\right)^2 \cdots \\ &= \pi_{11,0} + T_{11}^{+}(z)\left(\pi_{11,1} + \pi_{11,2}T_{11}(z) + \pi_{11,3}\left(T_{11}(z)\right)^2 + \cdots\right) \\ &= \pi_{11,0} + T_{11}^{+}(z)\frac{\Pi_{11}\left(T_{11}(z)\right) - \pi_{11,0}}{T_{11}(z)} \\ &= \pi_{11,0}\left(1 + T_{11}^{+}(z)\frac{1 - A_{11}\left(T_{11}(z)\right)}{A_{11}\left(T_{11}(z)\right) - T_{11}(z)}\right). \end{aligned} \tag{16}$$

Average delay can be obtained by differentiating (16) and applying L'Hospital's rule:

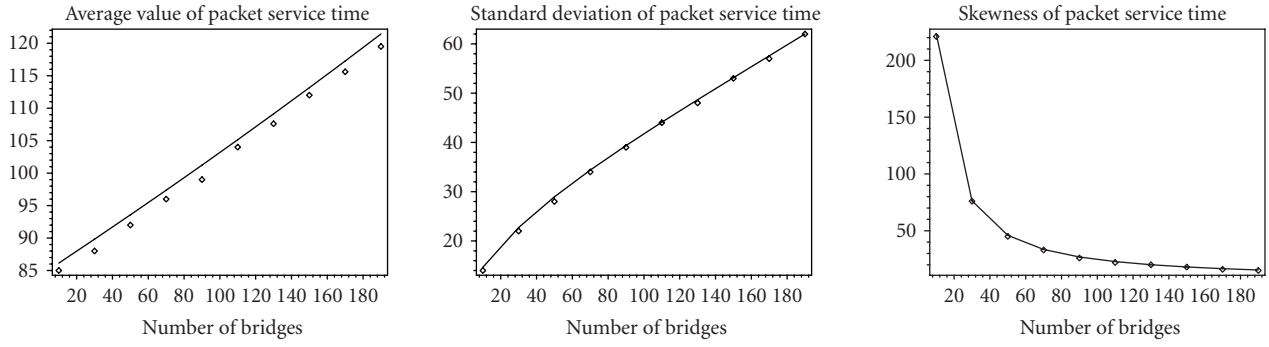
$$\overline{W_{11}} = \frac{\phi_{11}T_{11}^{(2)}}{2(1-\rho_{11})}, \tag{17}$$

where $T_{11}^{(2)}$ denotes second moment of packet service time. This result matches Pollaczek-Khinchin mean value formula [19].

The complete access time which includes waiting time and service time of the target packet is then equal to $\overline{S_{11}} = \overline{W_{11}} + \overline{T_{11}}$.

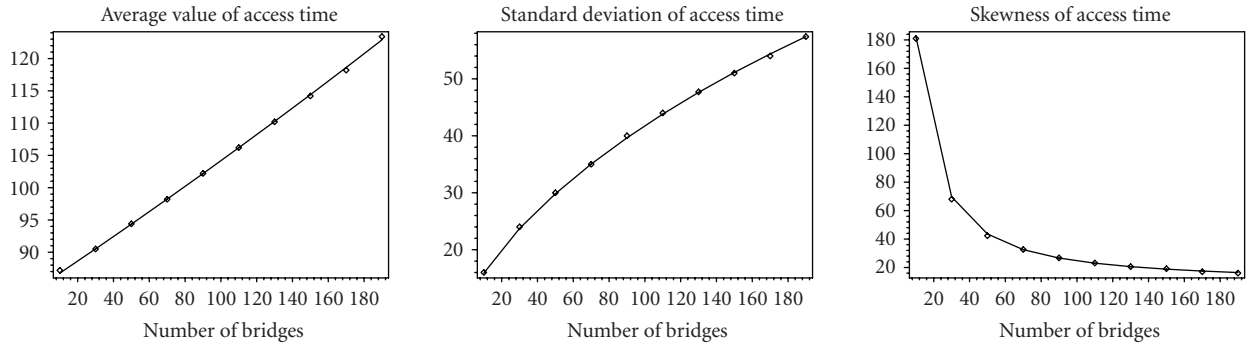
6. Performance Evaluation of GTS Bridge in Continuous EKG Telemetry

In this section, we present performance results for GTS bridge between IEEE 802.15.4 and IEEE 802.11b deployed in continuous EKG telemetry. In design of GTS bridge for EKG telemetry, we assume that superframe will contain only three GTS parts. First one is management slot used for control communication between IEEE 802.15.4 mote and the bridge. Second part is used to carry digitized EKG samples, and the third part should contain acknowledgment from bridge to mote. Duration of these parts depends on the duration of superframe and time distance between the beacons. For example, if $SO = 0$, then superframe including the beacon frame contains 16 slots with three backoff periods each. Duration of beacon frame is 30 bytes (3 backoff periods) since beacon can carry acknowledgment information for previous superframe. Minimal duration of management slot is three backoff periods (30 bytes). However, 14 slots are then left to carry samples and packet authentication code.



(a) Mean packet service time (b) Standard deviation of packet service time (c) Coefficient of skewness of packet service time

Figure 7: Moments of packet service time. Analytical results are shown as lines and simulation results are shown as points.



(a) Mean system time (b) Standard deviation of system time (c) Coefficient of skewness of system time

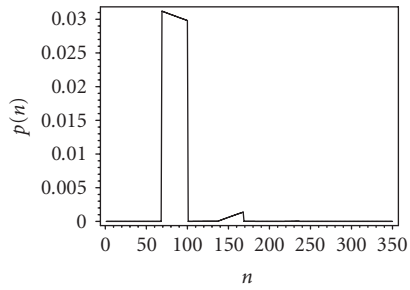
Figure 8: Moments of system (access) time. Analytical results are shown as lines, and simulation results are shown as points.

We assume that HMAC function adopted is constructed from secure hash algorithm (SHA-1 hash [6, 20]) which is a widely used cryptographic hash function with a message digest output of 160 bits. Therefore, 400 bytes are left in the superframe which covers at least 200 digitized samples, that is, measurement period of 1 second. This confirms that $SO = 0$ is a correct choice as long as the superframes are sent with the period less than a second. Choice of the *BO* parameter, that is, the period between the beacons is result of contradicting requirements (Table 2 outlines our design options). First requirement is related to low power consumption and asks that *BO* is chosen to be as large as possible, but still able to carry all the samples generated during beacon interval (preferably close to 1 second). Second requirement is related to the packetization delay and reliability. Low packetization delay requires small amount of data in the packet. Second, both IEEE 802.15.4 and IEEE 802.11b (and even IEEE 802.15 1 Bluetooth if it happens to operate in vicinity) operate in ISM band and cause interference to each other. Interference can corrupt the whole superframe, and, therefore, shorter superframe sizes are preferable.

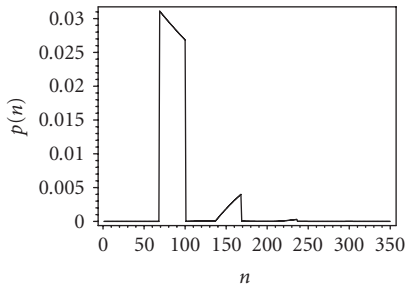
We consider an ideal wireless channel (without noise and fading). MAC and physical layer parameters are given in Table 3 resulting in a period between the superframes of 0.123 seconds.

We have numerically solved the overall system of equations under varying number of bridge devices. Analytical processing was done using Maple 11 from Maplesoft. We have also implemented the simulation model using Petri Net simulation engine Artifex [21]. Figures 5 and 6 show values of basic network parameters when the number of bridges is varying between 10 and 190. Delays are shown in numbers of IEEE 802.11b slots (20 microseconds). Although total network load is light, we observe that an increase of the number of devices under constant load per device causes linear increase of access probability, freeze probability, and throughput; while at the same time, transmission success probability, probability that buffer is empty after departure experiences linear decrease. We also observe that analytical (shown as line) and simulation results (shown as points) are close.

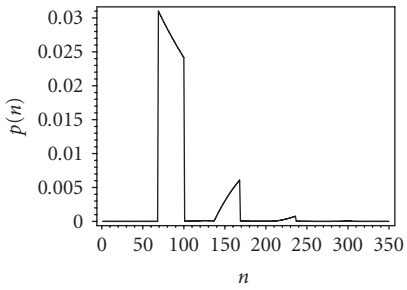
Figures 7 and 8 show first three moments of the packet service time and packet access time which includes waiting time in the queue and packet service time. We observe that for an increase of the average packet service time of 30%, standard deviation has increased three times. We present coefficient of skewness which is derived as the ratio of the third moment of the probability distribution and third power of standard deviation [22] which indicates symmetry of the probability distribution around the mean value. This coefficient is close to zero for distributions



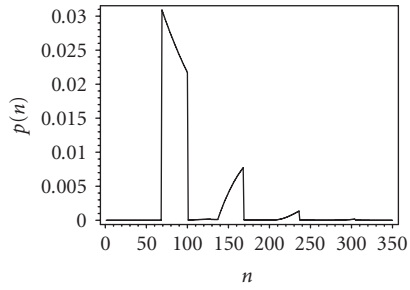
(a) Probability distribution of packet service time, 10 bridges



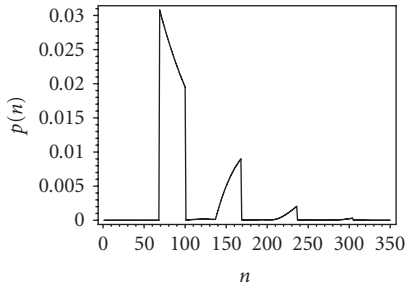
(b) Probability distribution of packet service time, 30 bridges



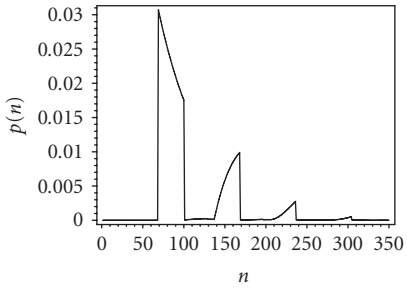
(c) Probability distribution of packet service time, 50 bridges



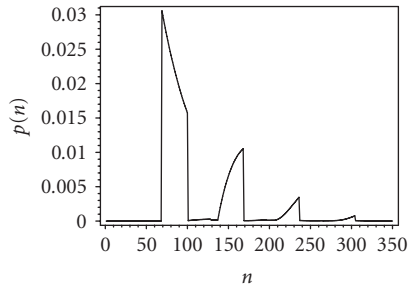
(d) Probability distribution of packet service time, 70 bridges



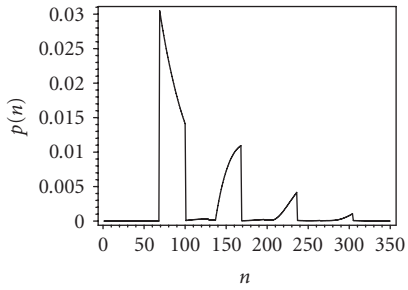
(e) Probability distribution of packet service time, 90 bridges



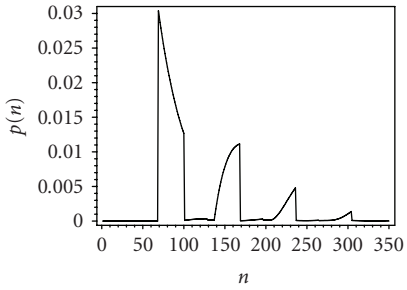
(f) Probability distribution of packet service time, 110 bridges



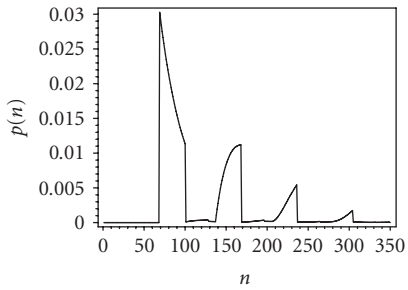
(g) Probability distribution of packet service time, 130 bridges



(h) Probability distribution of packet service time, 150 bridges



(i) Probability distribution of packet service time, 170 bridges

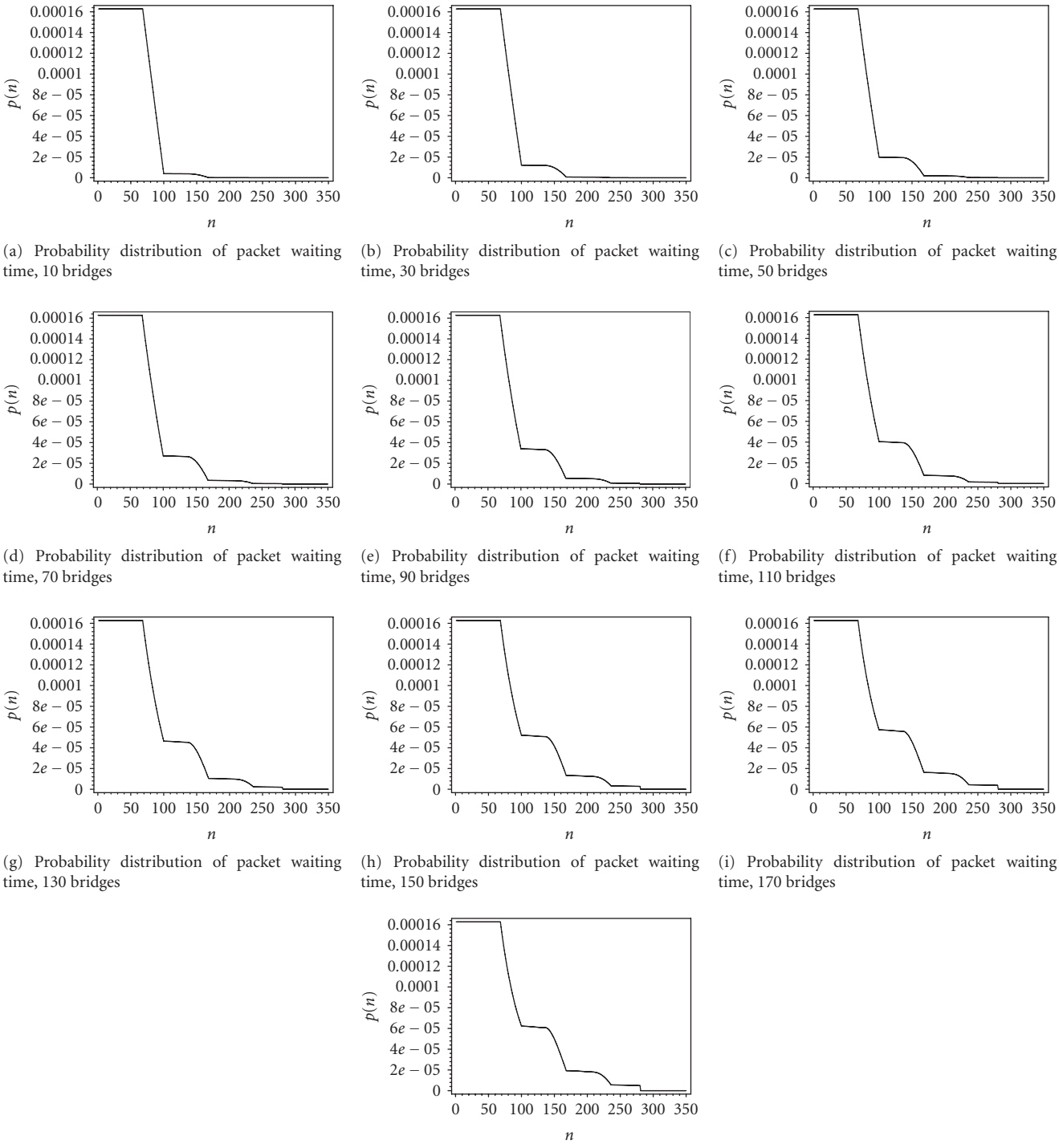


(j) Probability distribution of packet service time, 190 bridges

Figure 9: Probability distribution of packet service time.

which are symmetric around their mean values. However, calculated values of skewness parameter indicate high level of asymmetry, and we have been motivated to calculate complete probability distributions and discuss them.

Figures 9 and 10 show probability distributions of packet service time and packet waiting time obtained analytically for cases when the number of bridges increases from 10 to 190 in steps of 20. For packet service time, each peak on the graph corresponds to one backoff attempt. We notice that for small number of bridges, almost all packets are served in first backoff attempt. The beginning of the first peak is determined by the time to complete single packet



(a) Probability distribution of packet waiting time, 10 bridges

(b) Probability distribution of packet waiting time, 30 bridges

(c) Probability distribution of packet waiting time, 50 bridges

(d) Probability distribution of packet waiting time, 70 bridges

(e) Probability distribution of packet waiting time, 90 bridges

(f) Probability distribution of packet waiting time, 110 bridges

(g) Probability distribution of packet waiting time, 130 bridges

(h) Probability distribution of packet waiting time, 150 bridges

(i) Probability distribution of packet waiting time, 170 bridges

(j) Probability distribution of packet waiting time, 190 bridges

Figure 10: Probability distribution of packet waiting time in bridges' buffer.

transmission without the backoff count $\overline{St} = 69$ slots. The width of the first peak corresponds to the size of initial backoff window enlarged by freezing. After first backoff, transmission is either successful or collided where collision lasts for $\overline{Ct} = 29$ slots (which corresponds to the distance between the peaks). As the number of bridges increases, the number and intensity of higher-order backoff attempts increases and higher-order peaks become more pronounced. Situation is similar with packet waiting time in the sense that its distribution becomes wider as the number of bridges increases. The first peak of this distribution comes from the residual packet service time (of the packet currently being

transmitted) without the collision, and it further widens with increasing number of backoff phases. Flat parts after the first one are due to packet retransmissions after collision.

Both distributions show that packet delay in ward network is random with large variance as the number of bridges increases. This is bad news for real-time payload transmitted in the IEEE 802.11 packets since EKG samples have to be displayed in constant time periods. The results shown are useful to determine the amount of playback buffering for EKG data in order to ensure intelligible display of data.

7. Conclusion

In this paper, we have presented the design issues and performance evaluation of the bridge between the BAN implemented using beacon-enabled IEEE 802.15.4 network and IEEE 802.11b wireless LAN. Bridge has been implemented using GTS feature of IEEE 802.15.4. Performance results show that for small offered load and very small packet sizes which carry EKG data (with basic bandwidth of 2400 bps), large number of devices generates very wide probability distribution of the packet access time. Given that EKG data has to be displayed in real time, accurate estimation of access delay is necessary in order to dimension buffering at the receiver. We have shown that probability distributions of packet service time and packet waiting time cannot be characterized using first two moments, instead the whole probability distributions are needed in order to accurately estimate buffering delays at the receiver.

Acknowledgment

This research is supported by the NSERC Strategic grant.

References

[1] V. Shnayder, B. Chen, K. Lorincz, T. R. F. Fulford-Jones, and M. Welsh, "Sensor networks for medical care," Tech. Rep. TR-08-05, Harvard University, Cambridge, Mass, USA, 2005.

[2] T. R. F. Fulford-Jones, G.-Y. Wei, and M. Welsh, "A portable, low-power, wireless two-lead EKG system," in *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology (EMBC '04)*, vol. 3, pp. 2141–2144, San Francisco, Calif, USA, September 2004.

[3] E. Hartmann, "ECG front-end design is simplified with microconverter," *Analog Dialogue*, vol. 37, no. 4, pp. 1–5, 2003.

[4] P. O. Bobbie, H. Chaudhari, C. Z. Arif, and S. Pujari, "Electrocardiogram (EKG) data acquisition and wireless transmission," *WSEAS Transactions on Systems*, vol. 3, no. 8, pp. 2665–2672, 2004.

[5] M. Bishop, *Computer Security: Art and Science*, Pearson Education, Boston, Mass, USA, 2003.

[6] W. Stallings, *Cryptography and Network Security: Principles and Practice*, Prentice Hall, Upper Saddle River, NJ, USA, 3rd edition, 2003.

[7] "Standard for part 15.4: wireless MAC and PHY specifications for low rate WPAN," IEEE Std 802.15.4, IEEE, New York, NY, USA, October 2003.

[8] G. Anastasi, E. Borgia, M. Conti, and E. Gregori, "IEEE 802.11b ad hoc networks: performance measurements," *Cluster Computing*, vol. 8, no. 2-3, pp. 135–145, 2005.

[9] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 2000.

[10] E. Ziouva and T. Antonakopoulos, "CSMA/CA performance under high traffic conditions: throughput and delay analysis," *Computer Communications*, vol. 25, no. 3, pp. 313–321, 2002.

[11] Y. Xiao and J. Rosdahl, "Throughput and delay limits of IEEE 802.11," *IEEE Communications Letters*, vol. 6, no. 8, pp. 355–357, 2002.

[12] H. Zhai, Y. Kwon, and Y. Fang, "Performance analysis of IEEE 802.11 MAC protocols in wireless LANs," *Wireless Communications and Mobile Computing*, vol. 4, no. 8, pp. 917–931, 2004.

[13] H. Zhai, X. Chen, and Y. Fang, "How well can the IEEE 802.11 wireless LAN support quality of service?" *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 3084–3094, 2005.

[14] J. Hui and M. Devetsikiotis, "A unified model for the performance analysis of IEEE 802.11e EDCA," *IEEE Transactions on Communications*, vol. 53, no. 9, pp. 1498–1510, 2005.

[15] X. Chen, H. Zhai, X. Tian, and Y. Fang, "Supporting QoS in IEEE 802.11e wireless LANs," *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, pp. 2217–2227, 2006.

[16] Y. Xiao, "Performance analysis of priority schemes for IEEE 802.11 and IEEE 802.11e wireless LANs," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1506–1515, 2005.

[17] J. Mišić and V. B. Mišić, *Wireles Personal Area Networks: Performance Interconnections and Security with IEEE 802.15.4.*, John Wiley & Sons, New York, NY, USA, 2008.

[18] H. Takagi, *Queueing Analysis. Volume 1: Vacation and Priority Systems*, North-Holland, Amsterdam, The Netherlands, 1991.

[19] L. J. Kleinrock, *Queueing Systems. Volume 1: Theory*, John Wiley & Sons, New York, NY, USA, 1972.

[20] B. Schneier, *Applied Cryptography*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1996.

[21] RSoft Design, *Artifex v.4.4.2*, RSoft Design Group, Inc., San Jose, Calif, USA, 2003.

[22] P. Z. Pebbles Jr., *Probability, Random Variables, and Random Signal Principles*, McGraw-Hill, New York, NY, USA, 1993.

Hindawi Publishing Corporation
International Journal of Telemedicine and Applications
Volume 2009, Article ID 181842, 8 pages
doi:10.1155/2009/181842

Research Article

Using Discharge Abstracts to Evaluate a Regional Perinatal Network: Assessment of the Linkage Procedure of Anonymous Data

Catherine Quantin,[1,2] Béatrice Gouyon,[1,2] Paul Avillach,[3,4] Cyril Ferdynus,[1,2] Paul Sagot,[5] and Jean-Bernard Gouyon[6]

[1] *INSERM, UMR 866, University of Burgundy, 21000 Dijon, France*
[2] *Division of Medical Informatics, University Hospital, 21000 Dijon, France*
[3] *Laboratoire d'Enseignement et de Recherche sur le Traitement de l'Information Médicale (LERTIM), Faculté de Médecine, Université de la Méditerranée, 13385 Marseille, France*
[4] *Institut de Santé Publique, d'Epidémiologie et de Développement (ISPED), Université Victor Segalen Bordeaux II, 33076 Bordeaux, France*
[5] *Division of Obstetrics and Gynaecology, University Hospital, 21079 Dijon, France*
[6] *Division of Pediatrics, University Hospital, 21079 Dijon, France*

Correspondence should be addressed to Catherine Quantin, catherine.quantin@chu-dijon.fr

Received 25 March 2008; Revised 3 September 2008; Accepted 6 November 2008

Recommended by Hui Chen

To assess the Burgundy perinatal network (18 obstetrical units; 18 500 births per year), discharge abstracts and additional data were collected for all mothers and newborns. In accordance with French law, data were rendered anonymous before statistical analysis, and were linked to patients using a specific procedure. This procedure allowed data concerning each mother to be linked to those for her newborn(s). This study showed that all mothers and newborns were included in the regional database; the data for all mothers were linked to those for their infant(s) in all cases. Additional data (gestational age) were obtained for 99.9% of newborns.

1. Introduction

The Burgundy perinatal network was created in 1992 to improve the quality of perinatal care in Burgundy, a French region with 1 800 000 inhabitants and 18 500 births annually. This network gradually included the 18 private and public hospitals of Burgundy. The medical steering committee, made up of medical representatives of each hospital (one paediatrician and one obstetrician for each) has progressively reorganized perinatal care in the region. In particular, medical conventions between hospitals were signed to improve coordination of admissions, transfers of patients, and the use of technical facilities. Moreover, a half-yearly systematic analysis of perinatal deaths was initiated.

The impact of actions promoted in the perinatal network has been assessed yearly since 1998.

A multidisciplinary working group previously chose and precisely defined specific indicators (25 for each mother and 17 for each newborn) [1, 2]. These indicators correspond to maternal medical history, psychosocial risk factors, maternal and neonatal medical data such as hospitalisations during pregnancy, delivery method, postpartum period, and the different hospitalisations of the newborns. Most of these items could be obtained from the mandatory discharge abstracts for each hospitalised patient within the Programme de Médicalisation du Système d'Information (PMSI). The PMSI collection system was used as it is compulsory in both public and private hospitals in France. Each hospital provides its own perinatal data on a voluntary basis. These data are

rendered anonymous before being sent from each hospital to the committee in charge of the assessment of the perinatal network's performance (regional audit committee).

Optimal assessment of perinatal care obviously requires effective linkage between maternal and neonatal data. Indeed, in order to study the mechanisms of neonatal diseases it is essential to link (1) the abstracts of a mother to those of her newborn even if they were not cared for in the same hospital and (2) all discharge abstracts for the same patient (mother and/or newborns), who may have several discharge abstracts from different hospitals. In our case, this happens when an antenatal or postnatal transfer from one hospital to another is decided, according to the written rules for good medical practice in the perinatal network.

This study intends to assess the performance of the file-linkage process of maternal and neonatal data used for the evaluation of the Burgundy perinatal network.

2. Materials and Methods

2.1. Population. All deliveries and births, whether live or not, are included in the perinatal network if the gestational age is at least 22 weeks and/or if the birth weight is greater than 500 g. For the purpose of this study, we took into account the population included from 1998, the year data collection began, to 2001. In 1998, only nine hospitals participated in the data collection, whereas in 2001 all of the 18 hospitals involved in the regional perinatal care network were involved, thus providing data for 100% of the 18 500 annual births.

2.2. Data Collection. Discharge abstracts for all mothers (social and medical data about gestation, childbirth and the postnatal period) and all newborns (hospitalisations in maternity hospitals, paediatric wards, neonatal units, and neonatal intensive care units) collected within the PMSI were provided by each hospital in Burgundy. The collection of discharge abstracts is mandatory and does not require the agreement of the patient. The usual data collection system of the PMSI was expanded to include the collection of birth weight and gestational age. Moreover, six identification items (maiden names, first names and birth dates of mothers, first names and birth dates of newborns, zip codes of the main residence of mothers) were recorded for both mothers and newborns to allow linkage between their discharge abstracts.

2.3. Anonymity of Information. The Directive 95/46/CE of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data and the Act n°78-17 of 6 January 1978 on Data Processing, Data Files and Individual Liberties (amended by the Act of 6 August 2004 relating to the protection of individuals with regard to the processing of personal data) are the European and French legislations on data privacy and security. Following these recommendations, the data were rendered anonymous in each hospital using 3 ANONYMAT software [3–5]. The principle of this software based on the Standard Hash Algorithm is to ensure the irreversible transformation of independent fields, which prevents the identification of an individual. The aim is to obtain a strictly anonymous code, but always the same one for a given couple of a mother and her newborn(s). The use of ANONYMAT software has been authorized by the French Department for Information System Security (SCSSI), and in 1998 the French data protection commission (Commission Nationale de l'Information et des Libertés—CNIL) specifically authorized the management of perinatal data in Burgundy. The collection of the data does not require the specific consent of the patient, who only has to be informed that the information contained in discharge abstracts will be used for epidemiological purposes and not only for budgetary purposes. An information leaflet explaining the use of the data is given to the patient on admission to the hospital.

In order to reduce the impact of typing errors, spelling transformation was introduced into the anonymity process before hash coding. The principle is to transform the spelling of names according to phonetic rules. Among the different treatments, SOUNDEX is most widely used. However, in the perinatal network, a treatment specifically adapted to the particularities of the French language was used [6].

2.4. File Linkage. Once rendered anonymous, the data were transmitted to the regional audit committee located in the University Hospital of Dijon (France) and included in a regional database (cf. Figure 1). The first part of the linkage process is to gather all discharge abstracts concerning the same patient (mother or newborns) from the same or different hospitals. The second part is devoted to linking the abstracts of a mother to those of her newborn even if they were not cared for in the same hospital. In order to respect the confidentiality of medical files, linkage is carried out on files that are rendered anonymous and not directly on nominative data. Technically, there is no difference between using original nominative data or one-way hash-coded data for record linkage, as the correspondence between these data is almost one-to-one (low collision rate). In fact, the spelling processing used in the anonymity procedure reduces the impact of typing errors and increases the efficiency of the linkage. However, two types of linkage errors are of concern: erroneous links of notifications (information) from two distinct patients, also called homonym errors, and failure to link multiple notifications (information) on the same patient, also called synonym errors.

These errors would be reduced if we could use the French Identification Number which is unique for each individual. However, this number is not recorded by hospitals, which only use the Social Security Number (SSN), which is not unique for each individual as it is assigned to each insured person, who may register his family under the same SSN. Thus, the SSN of a woman may change, for example, after her marriage. Moreover, communication of the SSN is not allowed by French legislation except for the transmission of information from private-sector hospitals to French health-insurance companies.

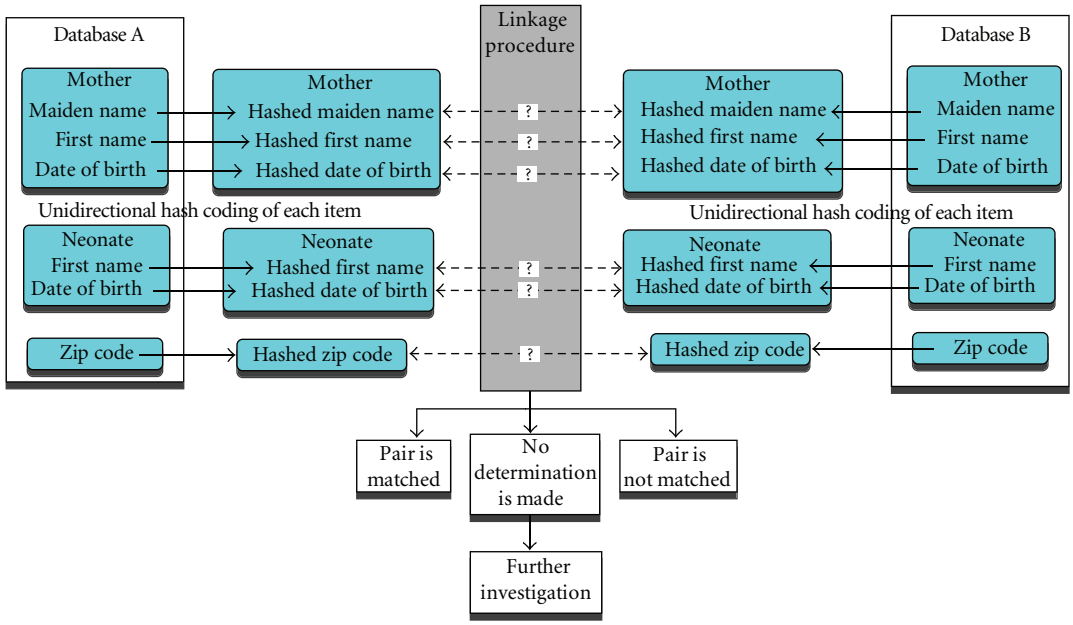


Figure 1: File linkage.

Therefore, we have developed a record linkage system which is based on the method proposed by Jaro [7] and takes into account the 6 identification variables.

In this procedure, like in other probabilistic methods [8–15], a weight is associated with each variable according to the reliability of the information provided. The weight given to two identical dates of birth is then higher than that given to two identical zip codes, as date of birth is more discriminating than zip code. It is then possible to compute the weight corresponding to each pair of records by summing the positive weights obtained for concordant variables and negative weights for discordant variables. Two threshold values can then be computed, from which three sets of possible decisions can be determined as follows: (1) the pair is matched; (2) no determination is made; (3) the pair is not matched. In case 2, no determination, other information is explored.

2.5. Reliability of the Linkage Procedure. It was assumed that a link between a mother and a noncorresponding newborn was impossible in the case of perfect agreement between two records on the 6 identification items. Indeed, it is highly unlikely that in the same maternity hospital, two women having the same zip code, the same maiden name, the same first name, the same birth date would give birth on the same day to babies with the same first name. Additionally, the risk for a homonym error was very low (10^{-48}) with the standard hash algorithm [16].

In a perinatal network, it is obvious that every newborn has a mother and vice versa. When a mother abandons her infant at birth and has the right to remain anonymous, the maiden name, the first name, and the date of birth of the mother were filled with random numbers and thus rendered anonymous in the hospital file before hash coding, both in the mother's and in the newborn's files. As the linkage items for the mother and her newborn are the same (filled in with the same numbers), linkage is possible.

So, the fact that no link was found between a newborn and a mother would indicate either a linkage error or the lack of the mother's record. Linkage errors were identified by again using the linkage method on the basis of only five items or even four. To verify these potential links, each hospital was asked to verify the identification items of the records corresponding to the given anonymous numbers. The corrected data were rendered anonymous before being returned from the hospitals to the regional database.

Finally, the linkage was only performed on the basis of all 6 identification items in the final database.

2.6. Data Validation. Before linkage, the exhaustiveness of both medical items and linkage items was monitored as follows: (1) exhaustiveness for the 6 linkage items was verified for each discharge abstract; (2) exhaustiveness for gestational age and birth weight was verified for each neonatal discharge abstract.

After linkage, the exhaustiveness for the number of mothers and newborns registered in the regional database was assessed from hand-written notebooks which are used in each hospital for the registration of births and/or admissions of sick newborns to units caring for these infants. The law requires these hand-written notebooks to be completed.

Regarding data quality, computerized procedures have also been developed to reveal discrepancies between medical data or between dates of exit and admission for successive hospitalisations. These inconsistencies were validated by a paediatrician. All erroneous data were then corrected in the nominative files in each hospital, before being rendered anonymous and sent back to the regional database. As a consequence, after validation, the number of mothers and newborns agreed with hand-written notebooks, and the

Table 1: Assessment of the linkage procedure: sensitivity, specificity, rates of true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP), for each year.

Years	Records no.	Sensitivity (%)	Specificity (%)	TP (%)	TN (%)	FN (%)	FP (%)
1998	17865	89.04	95.25	85.68	3.59	10.55	0.18
1999	15769	97.53	97.82	89.04	8.52	2.25	0.19
2000	20941	97.24	97.77	86.35	10.50	2.45	0.70
2001	25070	93.03	87.47	78.99	13.20	5.92	1.89
2002	25489	94.17	89.74	79.72	13.76	4.94	1.57
2003	25264	93.82	86.84	78.64	14.05	5.18	2.13
2004	25976	88.26	87.31	69.14	18.91	9.20	2.75
2005	24440	93.13	86.02	70.57	20.84	5.20	3.39
2006	27238	97.09	77.74	73.97	18.51	2.21	5.30

medical or administrative inconsistencies between discharge abstracts were corrected. We then considered data after validation as a gold standard for the assessment of the linkage procedure (on data before validation).

3. Results

This collection started in 1998. The exhaustiveness of data collection and the linkage rate have improved with time.

During validation after linkage, the number of mothers and newborns registered in the regional database was assessed from the hand-written notebooks which are used in each hospital for the registration of births and/or admissions of sick newborns to units caring for these infants. At the beginning of the collection in 1998 or 1999, we detected some discrepancies. We thus had to modify the selection criteria for data collection. From 2000, the consistency between these numbers was perfect.

(1) Exhaustiveness of Discharge Abstracts. In 1998, 9 hospitals were involved in the collection of discharge abstracts. Abstracts for 84.1% and 99.1% of all eligible mothers were retrieved in the regional database, respectively, before and after the validation procedure; for newborns, the percentages were 100% and 98.7%. Overestimation of the hospitalisation rate in some neonatal files in 1998 was explained by undue inclusion of hospitalisation beyond the neonatal period.

From 2001 to 2005, 18 hospitals were involved in the collection of discharge abstracts and the overall exhaustiveness for both mothers and newborns reached 99.9% after the validation procedure. The exhaustiveness of discharge abstracts was 100% but some mothers may not have been hospitalised if the delivery occurred at home.

(2) Exhaustiveness of the Data in the Discharge Abstracts. In 1998, the six items used for the linkage procedure were recorded in 80% of discharge abstracts before validation and in 99% after validation. From 2001 to 2005, the exhaustiveness of these items was between 91.9% and 98.5% before and 100% after validation.

(3) Linkage Assessment. Before validation, the percentage of newborns linked to their corresponding mothers on the basis of the 6 identification items was 71% in 1998, 92.9% in 2001, and 93.4% in 2005.

Different types of errors were found during the validation procedure after linkage. The most frequent one corresponded to errors in spelling surnames and first names [6] leading to phonetic changes that were not subsequently corrected by the spelling processing included in the ANONYMAT software. The inversion of the married name, and the maiden name, during data collection was also a source of linkage failure, and was corrected through the validation procedure. After validation, 99.98% of newborns were linked to their mothers whatever the year concerned. Only newborns who were transferred from a hospital of the Burgundy Region to a hospital of another region were missed (0.2%).

The results regarding the assessment of the linkage procedure, using validated data as a gold standard, are presented in Table 1, for each year.

Sensitivity is most often higher than 90% (97.1% in 2006). Of course the figure regarding the first year is lower. The year 2004 may be considered as an outlier: several hospitals had to change their software and faced many difficulties in collecting data. Specificity is nearly always higher than 85%, except in 2006 (77.7%).

The estimation of the false positive and false negative rates, given in Table 1, also indicates an unexpected higher rate of false positives in 2006 (5.3%). The analysis of the false positives is presented in Table 2, which provides, in case of false positives, the percentage of missing data for each identification item. We can observe that the main reason for false positives appears to be related to missing data; this was particularly the case in 2006. Among false positives, 85.11% were due to perfect concordance between the record of the baby and the record of the mother on the birth dates of the mother and on the zip codes of the main residence of the mother, while the other items were nearly all missing.

Of course, in these cases, agreement between the two records on the 6 identification items cannot be considered perfect. In fact, in case of missing data on an identification item, the two records are considered discordant on this item.

Table 2: Percentages of missing data for each identification item, among false positives, for each year.

Years	Maiden name (%)	Mother's first name (%)	Mother's date of birth (%)	Infant's first name (%)	Infant's date of birth (%)	Zip code (%)
1998	43.75	18.75	18.75	46.88	40.63	18.75
1999	33.33	33.33	20.00	56.67	53.33	20.00
2000	3.71	1.68	1.89	22.46	16.72	1.33
2001	89.87	90.08	0.21	97.89	97.68	0.21
2002	78.30	79.05	29.68	97.01	96.51	0.50
2003	84.94	84.39	0.37	98.14	98.14	0.37
2004	69.33	56.30	0.14	86.41	95.38	0.00
2005	82.49	58.82	8.70	74.76	79.11	8.82
2006	75.21	71.05	0.07	97.02	97.85	0.14

Table 3: Percentages of missing data for each identification item, among false negatives, for each year.

Years	Maiden name (%)	Mother's first name (%)	Mother's date of birth (%)	Infant's first name (%)	Infant's date of birth (%)	Zip code (%)
1998	95.33	97.40	97.40	91.14	90.98	0.32
1999	55.49	94.08	94.08	12.68	5.63	2.54
2000	62.62	67.91	69.28	20.94	16.05	25.24
2001	45.04	40.53	23.74	27.92	27.11	2.63
2002	58.22	27.40	29.47	34.47	19.30	1.03
2003	33.10	32.42	39.53	59.56	24.92	1.91
2004	27.46	25.66	11.85	38.80	30.85	0.63
2005	88.36	87.66	19.65	87.19	81.92	3.07
2006	74.96	74.46	71.97	47.43	32.34	1.00

For false negatives, the same analysis is presented in Table 3. Here again, missing data is the main cause. In particular, because of missing data on the maiden and first names of the mother, the two records were often considered discordant on these items, which prevented linkage of these records. Hopefully these errors have been corrected thanks to the validation procedure.

4. Discussion

4.1. Discussion of the Results. Optimal assessment of perinatal care needs a linkage procedure between successive files for the same patient and between files for the mothers and their corresponding newborn(s). The latter linkage was found to be essential in the assessment of the postnatal consequences of antenatal risk factors and maternal diseases. For instance, in 1999, one maternity hospital showed significantly higher rates of both Caesarean section and neonatal hospitalisation compared with regional and national rates. The linkage procedure revealed that the excess in neonatal hospitalisations was related to an excess in the use of caesarean section in a population of mothers similar to that admitted to other maternity hospitals in Burgundy. This study was performed after adjustment for gestational age and maternal diseases. This discrepancy was particularly noticeable for full-term infants. This finding was a strong argument for reorganising perinatal care leading to a significant decrease in the use of Caesarean section in this hospital.

The fact that each mother corresponded to a newborn and vice versa was particularly helpful in testing the linkage procedure. Indeed, this study demonstrated that coupling the direct linkage of anonymous data files with the validation procedure that takes into account the potential links revealed by the statistical procedure generated very satisfactory results on a regional scale.

This linkage method could also be used to link the siblings (having the same mother). The linkage of

(1) *full brothers or sisters* by sorting the dates of birth of all the individuals with the same parents,

(2) *half brothers and sisters* by sorting the dates of birth of all the individuals with the same father or the same mother,

would require the implementation of a new identifier, such as the family-based identifier (also including the names and date of birth of the father) [17, 18].

The assessment method used in Burgundy relied on the PMSI system. The main advantages of using the PMSI information system were that this data collection system is mandatory in all French hospitals and that the classification used in this program (i.e., the International Disease classification—10th revision) contained most of the items necessary for an audit of the perinatal network. Moreover, the number of additional items included was limited. Although PMSI was not designed for the assessment of care networks, using this information system made it unnecessary to set up another data collection system in each participating hospital and thus avoided duplication of work.

However, the collection of additional data and the need to extract items from PMSI required extension of the existing software in each hospital. Moreover, these changes involved several companies that were more or less interested in modifying their software. These changes have thus been implemented slowly over a 3-year period.

Overall, the collection and analysis of perinatal data had a substantial cost. Each hospital spent 1000 euros for changes in the PMSI software. Moreover, health professionals and a computer engineer were needed for validation procedures, management of the regional database, statistical analyses, transmitting regional data to the medical steering committee, drafting reports, and helping the hospitals to improve item recording.

Our results showed that satisfactory collection of the linkage items was more difficult to obtain than was the collection of medical items. This is easily explained by the fact that physicians are less motivated to collect administrative data. Therefore, the validation procedure for the identification items was made a cornerstone in the quality process.

The results regarding the assessment of the linkage procedure revealed that sensitivity and specificity are quite high (resp., higher than 90% and 85%). The decrease in sensitivity in 2004 is related to software changes in several hospitals. The decrease in specificity in 2006, as well as the overall false positive rate among the whole period, seemed to be the consequence of missing data. These observations led us to propose harmonization of the procedures (software and rules for collecting data) in all hospitals. This initiative, called the EXTRANAT project, is described in the next section.

4.2. Perspectives for the Development of the Perinatal Network. The Burgundy perinatal network was created in 1992 because of the high perinatal morbidity and mortality in France (7th highest among OCDE countries) notably in Burgundy (14th highest in France).

Since 1995, it has included the following.

(i) A hierarchal interestablishment network of all of the maternity clinics (1 type III maternity clinic, 2 type II maternity clinics, 8 type I maternity clinics, and 4 local perinatal care centers). The missions of each establishment and the criteria and procedures for the transfer of the mother (before or after delivery) and the newborn have been defined.

(ii) A common paper-based record document.

(iii) An original and exhaustive evaluation system, based on the PMSI, for all of the 18 000 annual pregnancies, with anonymization of data in each establishment after discharge, and linkage between the mother and her infant.

Thanks to the combination of these three elements and to the perseverance of those involved in perinatal care and the Burgundy DIM, since 1999, our region has climbed to be among the top three in France with regard to low maternal and perinatal mortality and morbidity.

However, though retrospective analysis of records on perinatal mortality showed a considerable improvement in the quality of intra- and interhospital care between 1996 and 2003 (in-utero transfers for extremely premature births, high-risk patients, antenatal corticotherapy, pediatric care at the maternity clinic), it also revealed a significant

deterioration in preclinic conditions, showing the need to set up an "extranet" tool for all of those involved in perinatal care (GPs, specialists, midwives), whether they are based in public hospitals, in private practice or in Mother and Child Care Centres (Protection Maternelle et Infantile).

This is the rationale behind the "Extranat Bourgogne" project, a computer system to facilitate

(i) the communication (between professionals),

(ii) the training and the evaluation of professional practices (discussions about protocols),

(iii) the implementation of shared interactive computerized medical records, with the medical and organizational references of the network.

Thanks to Extranat, it will be possible to have real-time communication, and regular meetings in the form of videoconferences (telemedicine) to study problem cases. The telemedicine system is already up and running and Extranat is currently being set up.

The aim of this project is to fulfill the three missions of the perinatal network as follows:

(i) improvement of the quality of care by enabling the health-care professional to input initial medical data (no longer final data at discharge from the establishment). This is a criterion in the improvement in information collection and in the sharing of medical data with complementary specialists (specialists in women's health, specialists in neonatal health for infants that may be transferred to other medical departments or other establishments),

(ii) evaluation of professional practices, which requires identification and anonymization of the professionals concerned,

(iii) evaluation of the medical and organizational aspects of the Burgundy perinatal network (anonymized database with mother-child linkage), already operational and presented in this article.

It is therefore essential for the new Extranat system to use the same patient identification and anonymization system (mothers, newborns, fathers) as the perinatal network, so as to

(i) provide access (limited and controlled) to medical data for professionals who are treating the mother or newborn (data identified at the physician's practice and/or the medical establishment),

(ii) respect the anonymization requirements for data for the patients (as requested by the CNIL), and for the professionals, in the regional data base.

5. Conclusion

It is possible to set up a continuous and exhaustive recording system for linked perinatal data to assess the quality of care on a regional scale. A pre-existing database like PMSI may be valuable, but may have to be extended if necessary. The linkage of anonymous files may greatly enhance the accuracy of the assessment procedure. These principles of assessment of a perinatal network could be extended to other medical domains.

Acknowledgments

This research was supported by the "Fonds National de Prévention, d'Education et d'Informations en Santé (URCAM, CRAM)".

References

[1] B. Cornet, J. B. Gouyon, C. Binquet, et al., "Using discharge abstracts as a tool to assess a regional perinatal network," *Revue d'Epidemiologie et de Sante Publique*, vol. 49, no. 6, pp. 583–593, 2001 (French).

[2] C. Quantin, C. Binquet, K. Bourquard, et al., "Which are the best identifiers for record linkage?" *Medical Informatics and the Internet in Medicine*, vol. 29, no. 3-4, pp. 221–227, 2004.

[3] C. Quantin, H. Bouzelat, F. A. Allaert, A. M. Benhamiche, J. Faivre, and L. Dusserre, "Automatic record hash coding and linkage for epidemiological follow-up data confidentiality," *Methods of Information in Medicine*, vol. 37, no. 3, pp. 271–277, 1998.

[4] C. Quantin, H. Bouzelat, F. A. Allaert, A. M. Benhamiche, J. Faivre, and L. Dusserre, "How to ensure data security of an epidemiological follow-up: quality assessment of an anonymous record linkage procedure," *International Journal of Medical Informatics*, vol. 49, no. 1, pp. 117–122, 1998.

[5] C. Quantin, B. Gouyon, F. A. Allaert, and O. Cohen, "Methodology for chaining sensitive data while preserving anonymity: an application to the monitoring of medical information," *Courrier des Statistiques*, no. 12, pp. 31–42, 2006.

[6] C. Quantin, C. Binquet, F. A. Allaert, et al., "Decision analysis for the assessment of a record linkage procedure: application to a perinatal network," *Methods of Information in Medicine*, vol. 44, no. 1, pp. 72–79, 2005.

[7] M. A. Jaro, "Probabilistic linkage of large public health data files," *Statistics in Medicine*, vol. 14, no. 5–7, pp. 491–498, 1995.

[8] M. Fair, M. Cyr, A. C. Allen, S. W. Wen, G. Guyon, and R. C. MacDonald, "An assessment of the validity of a computer system for probabilistic record linkage of birth and infant death records in Canada. The Fetal and Infant Health Study Group," *Chronic Diseases in Canada*, vol. 21, no. 1, pp. 8–13, 2000.

[9] L. E. Gill, "OX-LINK: the Oxford medical record linkage system 1997," in *Proceedings of an International Workshop and Exposition*, pp. 15–33, Arlington, Va, USA, March 1997.

[10] L. Dal Maso, C. Braga, and S. Franceschi, "Methodology used for "software for automated linkage in Italy" (SALI)," *Journal of Biomedical Informatics*, vol. 34, no. 6, pp. 387–395, 2001.

[11] S. Kendrick and J. Clarke, "The Scottish record linkage system," *Health Bulletin*, vol. 51, no. 2, pp. 72–79, 1993.

[12] S. Gomatam, R. Carter, M. Ariet, and G. Mitchell, "An empirical comparison of record linkage procedures," *Statistics in Medicine*, vol. 21, no. 10, pp. 1485–1496, 2002.

[13] J. D. Langley and J. L. Botha, "Use of record linkage techniques to maintain the leicestershire diabetes register," *Computer Methods and Programs in Biomedicine*, vol. 41, no. 3-4, pp. 287–295, 1994.

[14] M. G. Elfeky, V. S. Verykios, and A. K. Elmagarmid, "TAILOR: a record linkage toolbox," in *Proceedings of the 18th International Conference on Data Engineering*, pp. 17–28, San Jose, Calif, USA, February-March 2002.

[15] D. Nitsch, S. Morton, B. L. DeStavola, H. Clark, and D. A. Leon, "How good is probabilistic record linkage to reconstruct reproductive histories? Results from the Aberdeen children of the 1950s study," *BMC Medical Research Methodology*, vol. 6, article 15, pp. 1–9, 2006.

[16] C. Quantin, F. A. Allaert, H. Bouzelat, et al., "Security of healthcare data networks used for epidemiological studies," *Revue d'Epidemiologie et de Sante Publique*, vol. 48, no. 1, pp. 89–99, 2000.

[17] C. Quantin, F. A. Allaert, B. Gouyon, and O. Cohen, "Proposal for the creation of a European healthcare identifier," *Studies in Health Technology and Informatics*, vol. 116, pp. 949–954, 2005.

[18] C. Quantin, O. Cohen, B. Riandey, and F. A. Allaert, "Unique patient concept: a key choice for European epidemiology," *International Journal of Medical Informatics*, vol. 76, no. 5-6, pp. 419–426, 2007 (French).

Hindawi Publishing Corporation
International Journal of Telemedicine and Applications
Volume 2009, Article ID 767145, 8 pages
doi:10.1155/2009/767145

Research Article

Evaluation of a Telemedicine System for the Transmission of Morpho/Immunological Data Aiming at the Inclusion of Patients in a Therapeutic Trial

Jean-François Lesesve and Richard Garand

Centre Hospitalier Universitaire de Nancy et de Nantes (CHU Nancy et Nantes), 54000 Nancy, France

Correspondence should be addressed to Jean-François Lesesve, jf.lesesve@chu-nancy.fr

Received 29 May 2008; Accepted 6 November 2008

Recommended by Yang Xiao

Due to their high levels of achievement and efficiency, image digitalization and teletransmission tools are more and more frequently used. Applied to cellular haematology, these tools often contribute to diagnosis confrontation, sometimes within the framework of therapeutic trials. We present one of the first approaches of the use of telehaematology for the inclusion of patients in the GOELAMS chronic lymphocytic leukaemia 98 trial. The advantages were (1) the creation of a unique, protected, stable data bank that could be remotely consulted, (2) the use of digitized pictures which made expertise on identical documents possible, (3) the facility of computer exchanges between experts, in terms of reception as well as replying time delays. We were able to set out new standards of image sampling for CLL, solve the semantic divergences, and point out interobserver variability as regards morphology. The limiting factors were the important need for expert investment, but they more importantly concerned the first line morphologists who should benefit from adequate tools, in terms of computer equipment as well as members of staff, so as to apprehend this second reading system as a quality control procedure.

1. Introduction

"Expert" reviewing of microscopic data is a well-known procedure in medical practice, within the framework of cooperative studies with therapeutic, epidemiological, or scientific purposes. In the current state of things, this notion, however, mostly remains a theoretical one owing to persisting practical difficulties in its implementation [1]. Telehaematology consists in sending pictures of camera-digitized cells from one computer to another via the internet network. It appears as something easy which is more and more resorted to. Teletransmission of microscopic images enables us to overcome the usual obstacles usually met with traditional methods of smear reviewing (transporting delays, glass slide breakages) and offers new theoretical advantages (above all standardization of the observed cells) [2]. In practice, this teletransmission system remains under-used in multicentric studies. Three GOELAMS protocols (Groupe Ouest-Est d'Etude des Leucémies Aiguës et autres Maladies du Sang—Western/Eastern group for the study of acute leukaemias and other blood pathologies) have been completed: our study on chronic lymphocytic leukaemia (CLL) and two other ones on acute myeloid leukaemias. The GOELAMS CLL 98 protocol is being achieved; our goal was to develop the advantages and drawbacks of telehaematology for patient inclusion.

2. Materials and Methods

2.1. Patients. The GOELAMS CLL 98 study is a randomized multicentric study that compared the effectiveness and tolerance of an intensive treatment with autologous bone marrow transplantation versus CHOP Binet treatment as first-line treatment in patients under 60 years of age, with stage B or C CLL [3]. 86 patients were included on the following criteria: blood lymphocytosis $> 15 \times 10^9$/L or $> 5 \times 10^9$/L for at least 3 months, cytologic and histologic medullary infiltrate $\geq 30\%$, stage B and C, between 18 and

60 years of age, no preliminary treatment or chlorambucil only for less than 6 months.

2.2. Methods. May Gründwald Giemsa stained blood smears were sent to an haematological expert located either in Nancy (JFL) or in Nantes (RG) according to the geographic location of the center where the patients had been recruited, France having been for that purpose arbitrarily divided into two parts (East and West). The first expert captured the digital images of the cells and then sent those pictures to the second expert via a teletransmission device. The experts were supposed to have at their disposal the complete blood count and the immunophenotypes performed by the recruiting center and sent to them either by post or via the internet network (after having been scanned) in an attachment. The aim was to obtain a fast second reading of the results and provide the first expert with feedback, the consensus eventually being transmitted to the recruiting center. The morphologic documents were saved in a digitized visual-data bank, the smears could thus quickly be sent back to the labs they initially came from.

The digital images were captured using an optical photonic microscope at ×1000 magnification, an analogic tri CCD camera, a computer connected to the Internet network and to a secure web site (where the digitized pictures could be collected by the first expert and where the second one could receive the files) (TRIBVN and CRIHAN systems). The pictures of the lymphoid cells frequently corresponded to an almost continuous sampling, supposedly representative of the blood smears (ghost cells and poor quality pictures having been removed). Each and every cell was described, different percentages could thus be obtained (mature cells, cleaved cells, lymphoplasmocytoid variants, etc.) and, following from this, morphologic classification of the CLL (common or atypical).

The recruiting center's opinion as well as that of the first and second experts as regards cytology and the interpretation of the immunophenotypes by flow cytometry were codified thanks to a thesaurus which enabled standardization of the vocabulary used by the different specialists as well as improved interpretation of the results (additional thesaurus for haematology of the ADICAP code; Association pour le Développement de l'Informatique en Cytologie et en Anatomopathologie—Association for the Development of Computer science in Cytology and Anatomopathology), regarding the morphology: typical CLL H400; atypical CLL H 401 (mixed prolymphocytic), H402 (mixed pleomorphic), H403 (other cytology, plasmocytoid).

3. Results

3.1. Workable Files. The duration of the protocol was approximately 7 years (1st review: 29/November/1999; last review: 05/January/2006). 86 patients were included but we could only work on 79-patient data. Some files were indeed not provided by the laboratories, either because they had failed to send us the requested documents (5 patients) or because data had been lost owing to laboratory relocation (2 patients). 17 files were incomplete—with some parameters missing such as the date of validation, the lymphocytosis, the cytologic/immunological data. 56 files were digitized in Nancy and 23 in Nantes.

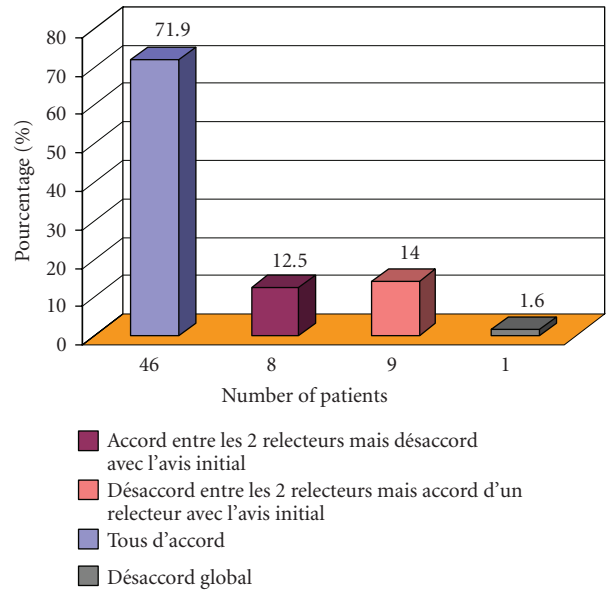


Figure 1: Cytologic concordance.

3.2. Agreements between Experts

3.2.1. Morphology (Figure 1). Overall agreement was obtained for 72% of the files (42 cases of typical CLL and 4 atypical). In one case, the three morphologists agreed to exclude the patient (non-Hodgkin lymphoma). Disagreements were reported in 28% of the cases. The diagnosis itself was not challenged, only the exact morphologic classification being at stake (CLL subtypes, "minor" disagreements). Disagreements between the opinions of the recruiter and both experts mainly consisted of reclassification of atypical towards typical CLL (five cases out of eight) or differences regarding the morphologic subtype of atypical CLL (two cases). In one case, the experts changed the CLL subclass from atypical to typical. Disagreements between the experts themselves occurred in 14% of the cases, the first one having classified 6 CLL out of 9 as typical whereas the second had identified them as atypical. In 3 out of 9 CLLs, the two experts did not agree on the morphologic subtype of atypical CLL. In another case, utter disagreement (three diverging opinions) concerned a file that had initially been identified as typical CLL and then changed to atypical, the two experts disagreeing on its morphologic subtype.

As a conclusion, all three agreed on a majority of files and disagreed on minor aspects of a minority of cases. Morphologic agreement is thus possible to achieve. Cytologic classification can consequently definitely be regarded as a reliable and repeatable tool. Last but not least, as for "difficult" cytologies, the method opened the door to discussion,

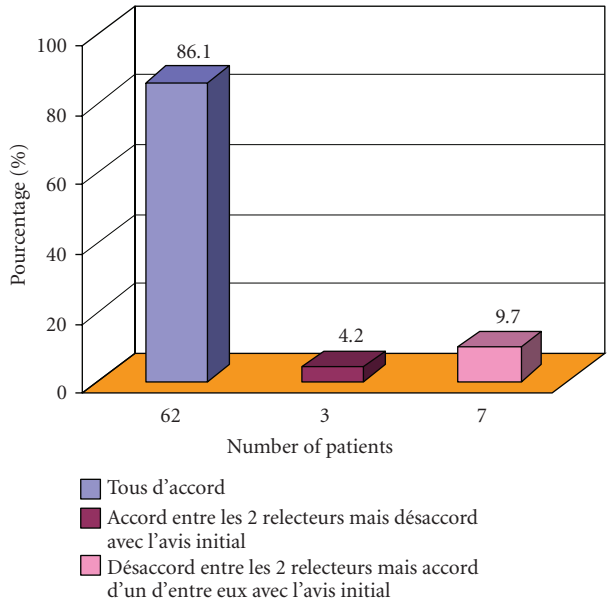


Figure 2: Flow cytometry concordance (Matutes scores).

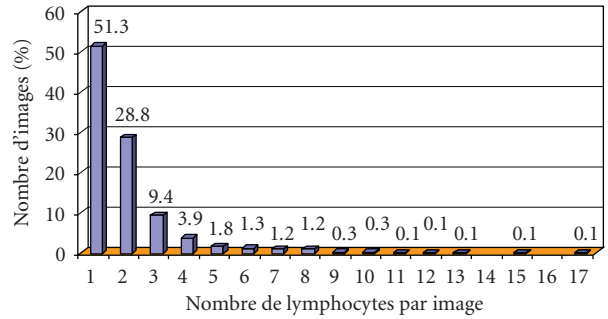


Figure 3: Distribution of the number of lymphocytes per image.

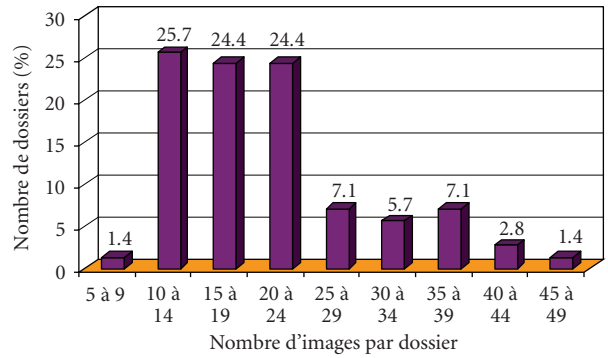


Figure 4: Distribution of the number of images per file.

which had not really been the case up to then (the specialist being isolated in their centre and the expert difficult to reach timely).

3.2.2. Immunophenotype (Figure 2). The results of the immunophenotypes were sent to the first expert by the recruiting centre along with the blood smears. Everything (the immunophenotypes and the digitized pictures) was then transferred to the second expert, either by e-mail in an attachment or by post, with a one-day time-lag. The histograms were re-interpreted, and Matutes scores were calculated. All patients had a Matutes score of 4 or 5, except for 2 cases with a score of 3 and one case with a score of 1 (not considered as CLL). Overall agreement between the 3 observers was 86%, including the 3 patients with a score ≤3. Disagreements dawned in 14% of the files. In 4% of the cases, the recruiter's opinion was different from the experts' point of view whereas in 10% of cases, the experts disagreed, one of them however concurring to the recruiter's viewpoint. Disagreements were related to the calculation of the score, between 4 (2 observers) and 5 (1 observer) or vice-versa. The CLL diagnosis was thus never questioned, whatever the morphology. There was no utter disagreement (three diverging opinions). We could thus conclude that as far as immunophenotypical diagnosis is involved, global consensus is not out of reach.

3.3. Methodology. We aimed to assess its feasibility (in terms of efficiency, practical side, etc.).

3.3.1. Number of Digitized Cells per Digitized Pictures and per Files (Figures 3–5). This criterium is essential for the feasibility of this method (review is quicker when cell concentration is higher) and quite a number of cells have to be analyzed before giving one's opinion. The total number of images acquired in this protocol was 1460, consisting of 2938 lymphocytes, which (theoretically) represents an average of 2 cells per picture. The minimum number of lymphocytes captured per image was 1 and the maximum 17, which corresponded to a 508 × 10^9/L lymphocytosis, namely the highest concentration that could be found in our series (Figure 3). The minimum number of pictures per file was 9 and the maximum 48 (Figure 4). The minimum number of captured lymphocytes for one file was 15 and the maximum 121 (corresponding to the 508 G/L lymphocytosis, Figure 5), the median being 38. 42 lymphocytes per file was the theoretical average (2938 photographed lymphocytes/70 files). Consequently, photographing around 40 lymphocytes per patient seemed appropriate to us as regards this type of lymphoproliferative syndrome. This figure enables the observer taking the pictures to provide the others with a sampling representative of the blood smear and the expert can thus reach a relevant diagnosis.

3.3.2. Number of Lymphocytes Photographed and Complexity of the Morphologic Diagnosis (Figure 6). The CLLs were classified either as typical (code ADICAP H400) or atypical morphology (H401, H402, H403). We chose to focus only on the cases where both experts had a similar cytologic diagnosis ($N = 53$). The average number of lymphocytes captured per file was 40 (±21, $N = 40$) for typical CLL versus 46 (±13, $N = 13$) for atypical CLL. The difference in averages between the typical and atypical CLLs was not significant

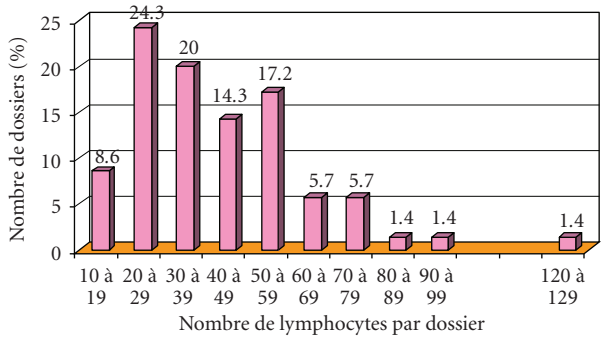


FIGURE 5: Distribution of the number of lymphocytes per file.

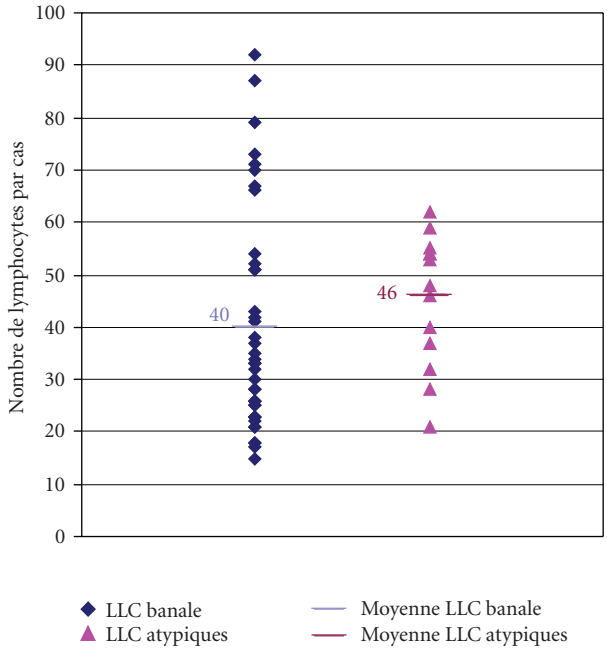


FIGURE 6: Average numbers of captured lymphocytes functions of the morphologic type of CLL.

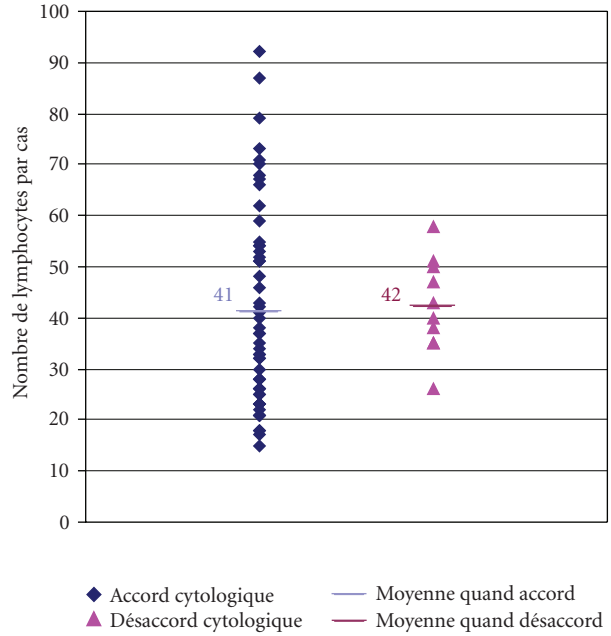


FIGURE 7: Average numbers of captured lymphocytes functions of cytological concordance.

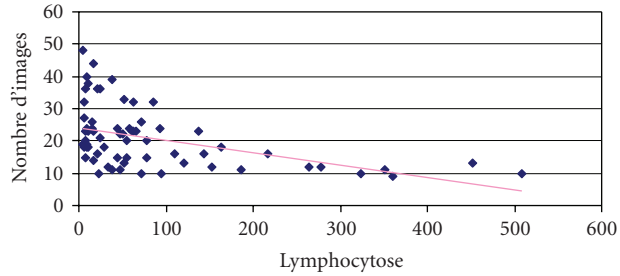


FIGURE 8: Number of pictures taken functions of the lymphocytosis.

($P > .2$, student's t-test). The number of digitized lymphocytes is thus not significantly higher when the morphology is atypical.

3.3.3. Cytologic Agreement/Disagreement and Number of Lymphocytes Photographed (Figure 7). The average number of lymphocytes captured per file was 41 (± 19, $N = 53$) when both experts were in agreement and 42 (± 10, $N = 10$) when they disagreed. The difference between the averages of captured lymphocytes was not relevant ($P > 0.2$, student's t-test). The number of captured cells is consequently of no influence on cytologic agreement.

3.3.4. Lymphocytosis and Number of Photographed Images (Figure 8). A link between the lymphocytosis and the number of digitized images could be established (linear Pearson coefficient, -0.46). The higher the lymphocytosis, the less images were taken. This is an important aspect as regards the practical side, since less time was necessary to save and maybe review the documents.

3.3.5. Delays of File Reception (Figure 9). In most cases, the cytologic and immunological documents were sent to one of the experts by post, who then looked after their digitization. Two cases were transmitted directly by e-mail (contrary to what was mentioned in the protocol). In 30% of the cases, the files were received within one month. For the remaining 70%, the time delay was several months and sometimes even reached years! The delay between the two experts using the teletransmission device was less than two weeks in 70% of the cases and 84% of the files were validated within one month of receipt. However, 16% of the files were validated after more than one month (16 weeks being the maximum). Telehaematology enables validation to be completed approximately 12 times faster than the traditional way of review (by post).

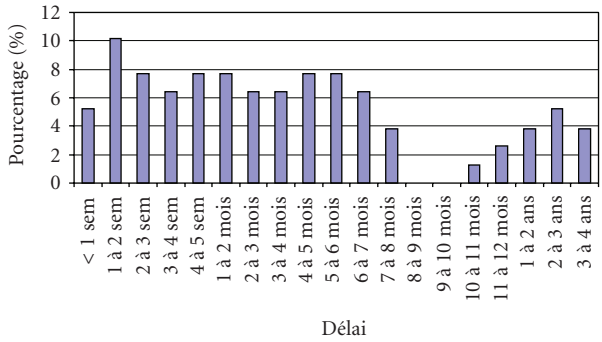


Figure 9: Time delay between initial diagnosis and first review.

4. Discussion

This study aimed at putting forward one of the first approaches of the use of telehaematology for the quality control of diagnosis [4]. The GOELAMS CLL 98 trial was chosen to assess this second reading system (blood smears and immunophenotypes). Above all, we wanted to question the efficiency of the method in use since second reading through resort to digitized pictures is, contrary to what might be assumed, a technique which remains largely underused [5].

As regards the morphologic classification, we noticed unanimous agreement in more than 2/3 of the cases. When the morphology was typical, agreement on the cytologic conclusion was almost systematic. 8 cases of disagreement have been reported—the experts having an opinion diverging from the recruiting centre. This can be due to the limited experience of the latter in the field, which justifies the request for a second expert opinion for the inclusion of patients. Furthermore, the experts themselves disagreed on 9 cases. This might seem weird in so far as each digitized cell was analyzed by both experts to classify the CLL. It proves that the same cell can be classified differently by two different specialists. Cytology is above all a matter of interpretation, thus implying a potentially important bias. On the whole, the three specialists agreed on most cases; and the experts disagreed on minor aspects of a minority of cases (the CLL diagnosis was never challenged). Immunophenotypical profile of the CLLs is the second important means to classify the illness. Matutes scores were really helpful for the diagnosis. In our study, almost all the patients scored 4 or 5, as expected. One patient scored 1, which enabled us to leave CLL aside, since it corresponded to the morphology of a leukaemic phase of lymphoma. The experts' conclusions had no impact on the score which had been predefined by the recruiting centre. Indeed, despite global disagreement in 13% of the cases, the only evolution of Matutes scores was between 4 and 5, both indicating CLL. These differences could be accounted for by the different threshold levels used for the isotypic controls, which modified the percentage of cells considered positive as well as the fluorescence intensity, thus affecting data interpretation. On the whole, global agreement between the three specialists prevailed. The differences of interpretation were not significant. We can thus conclude on the reliability of the laboratories where the immunophenotypes had been performed and say that they are the most capable of interpreting their histograms. Agreement was more frequent with the immunophenotypes than with cytology, which can be explained by the fact that immunophenotypical interpretation is more objective (charts are provided) compared to the relatively subjective dimension of cytologic analysis ("individual" morphologic interpretation).

So as to increase the feasibility of this type of inclusion protocol, we suggest sending the digitized pictures only accompanied with the data related to the Matutes scores as sufficient to validate patient inclusion.

Although no minimal standards of sampling had been pre-established, we found around 40 lymphocytes per file to be appropriate for this type of lymphoproliferative syndrome. This figure is in agreement with the minimum threshold of 30 lymphocytes previously established in another study, which consisted in requesting for specific opinion on various haematological disorders [1]. That figure can be used for both typical and atypical CLL since we have shown that there is no significant difference between the average numbers of captured cells in relation to the complexity of the diagnosis. Moreover, (dis)agreement about the CLL codification between the two experts was not related to the number of captured lymphocytes. Time is another parameter that has to be taken into account in the evaluation of the feasibility of the method. 60 to 90 minutes were necessary to digitize, classify and send the file. We wanted to find out whether there was a connection between the number of pictures taken (and thus the time spent on each smear) and the lymphocytosis. This indeed appears to be true to some extent, and although the tendency was not very clear ("visual" analysis of the slope of the trendline), a statistical link could yet be established. The number of images taken and the lymphocytosis were inversely correlated. Nevertheless, the specialist had to use the maximal magnification available ($\times 100$) since analysis of the cellular detail (shape and structure of the cell and its nucleus) is crucial for morphologic subclassification. However, with such a magnification, finding more than one cell in each image proved unfrequent, most images containing only one cell, whatever the cellular density on the smear may have been (high lymphoytoses $> 100 \times 10^9$/L being included). What is important was to manage to get a sampling which was representative of the smear's morphologic variability (diagnostic criterium). An average of around 40 lymphocytes, whatever the lymphocytosis, seemed to make this possible (practical criterium). A line sampling enabled us to meet these two requirements in most cases.

The time delay between the recruiting centre and the first expert for files sent by post was around one month for one third of the files, which can arbitrarily be defined as "acceptable". But for the remaining 2/3, the delay was tantamount to several months and even years. One of the causes that could be put forward is that the centres which are requested to participate in cooperative studies are often reluctant to part from their records and archives. The

delays and uncertainties caused by sending fragile documents by post indeed represent a serious drawback: the archive materials are frequently deteriorated (broken blood smears) and are not available for quite long periods of time. It was precisely one of the aims of the second reading protocol—namely to avoid parcel sending (at least from an expert to the other) after having certified that the original documents would be returned to the recruiting center once the digitized file was ready. Teletransmission clearly improved the second reading system. However, it is not yet completely satisfying since 16% of the files had still not been validated after a month. That can be explained by a two-week maintenance of the CRIHAN secure website and by the expert's unavailability or material incapability, and so forth. However, electronic data interchange has proved much faster than traditional mail (by post). This is obviously due to a greater speed of transmission (a few seconds instead of several days, or even more). We, however, believe that "motivation" plays a key role and is even more crucial. A direct email enclosing a request for a second reading is probably a greater incentive than more anonymous form letters. The time taken for the expert to reply was, for the most part, compatible with that of clinical decision-making. The expert's role was mainly to give a second opinion (validation). As for traditional second reading (slides sent by post or seminars such as "Forum Workshop"), the slowness of the whole thing prevented the results from being returned on time and, as such, they often came too late to influence in any way the therapeutic options that had already been chosen (for CLLs, a one-month time delay remains acceptable). Thanks to telehaematology, we can thus move on from a hypothetical and dubious retroactive assessment to an upstream quality control of the therapeutic decisions. The notion of time delay could even be eliminated thanks to software that enable two users to establish a direct connection and thus to hold a real-time dialogue and comment upon an image simultaneously thanks to a mobile pointer [5]. For example, that system allowed an expert to guide a nonspecialist technician and thus confirm the diagnosis. But however tempting this solution may appear, it should definitely not be privileged for financial reasons, since this equipment is far more expensive [1, 6]. Indeed, the expert (from where he is) is supposed to perform the tasks that the local morphologist would normally be doing; the expert thus replaces the morphologist instead of merely assisting them. The experts' availability being one of the major parameters accounting for the slowness of telepathology's development, asking them to become substitutes, does not seem very realistic to us [1].

Last but not least, one of the major gains was the creation of an objective database that was available for remote consultation (geographically and temporally). Indeed, one of the aims of the biological protocol was to set up a morphologic database. All the digitized pictures were saved in the CRIHAN secure website and were thus freely available to users, especially in case of dispute over the conclusions. A first step has been taken to solve the semantic inconsistencies related to the use of classifications [7]. The diversity of interpretations now remains to be tackled since, as our study showed, two experts do not necessarily classify the same cell in the same category, even though the differences are almost negligible. The creation of this database also enabled us to keep track of the initial blood smear (which is at present not mandatory in the majority of therapeutic protocols) and thus to keep an eye on disease evolution. As every time information is exchanged, respect of medical secrecy has to be ensured. Using a specifically dedicated talk-back secure website enabled us to guarantee the safety of sent and received messages since passwords were required. All messages posted through the secure website were saved (including the content and the replies) and both document history and traceability were preserved. Consequently, it is essential for good practice to use a dedicated secure website, as we did, rather than traditional emails and the protocol did not call this into question.

We sometimes noticed the difficulties encountered by some recruiting centres to provide us with the files. The time delays sometimes reached months, and even years. Several hypotheses can be put forward: (1) the lack of experience of small-size laboratories in taking part in protocols and the fact that they also keep less track of patient files, (2) the innovative dimension of our protocol, which consisted in assessing the validity of the initial diagnosis, (3) the main (therapeutic) objective of the protocol involved that the specialist was supposed to hand over data to "their" expert so that transmission of the biological files could be properly achieved (the need for which might have been overlooked by the specialist, especially for such a common pathology as CLL). We thus once again would like to stress the importance of establishing a dialogue between the specialists and the experts, from which transmission of the information about the biological and therapeutic protocols would follow. What can also (especially) be called into question is the current chronic underequipment of the laboratories, both in terms of specific hardware and members of staff. The current implementation of teletransmission devices will allow on the premises digitization of the files, which will consequently simplify the procedure and lead to better compliance and efficiency of such types of second reading. Moreover, the conditions regulating resort to "experts" and to "telehaematology expert networks" remain vague in so far as this activity has not yet been officially taken into account by the healthcare authorities and consequently does not benefit from any specific status or any guidelines as far as the financial aspect is involved [8]. There are unfortunately very few documented experiments that we know of dealing with that kind of cytologic activities [4, 9]. We could thus hardly find any data to compare our conclusions with. However, several initiatives implying telehaematology have recently been or are currently being implemented in France (Table 1). Moreover, databases associated with second reading protocols via teletransmission devices have been set up. However, the conclusions of such second readings are not awaited to introduce treatment. City/hospital networks have been created [10]. To finish, there are a few websites, such as Médecin'images or the one of the Collège de Hôpitaux Généraux thanks to which high quality exercises can be performed via teletransmission [11].

Table 1

Pathology	Protocol	Chair(wo)man	Starting date
French protocols using telemedecine for mandatory morphological consensus whorkshop (needed for inclusion)			
Chronic lymphocytic leukemias	GOELAMS LLC 98	JF Lesesve (Nancy) and R Garand (Nantes)	Sept-1999 (achieved)
Acute myeloblastic leukemias (adults)	GOELAMS 3	S Daliphard (Reims) and V Leymarie (Strasbourg)	March-2002 (achieved)
Acute myeloblastic leukemias (adults)	LAM-SA 2002	P Mossuz (Grenoble)	April-2004 (in progress)
Acute myeloblastic leukemias (pediatric)	ELAM02	O Fenneteau (Paris Robert Debré)	March-2005 (in progress)
French morphological data banks using telemedecine			
Myelomas	IFM	M Zandecki (Angers)	2000 (achieved)
Acute myeloblastic leukemias	Matchslide	G Flandrin (Paris Necker)	01/11/2001 (achieved)
Red blood cells	Teleslide	G Flandrin (Paris Necker), JF Lesesve (Nancy), O Fenneteau (Paris R Debré), T Cynober (Kremlin Bicêtre)	01/09/2003 (achieved)
Myelodysplastic syndromes	GFMDS	F Picard (Paris Cochin)	Dec-04 (in progress)
French morphological quality-control tests			
All topics	Medecin'image	JX Corberand (Toulouse)	
All topics	Collège des Hopitaux Généraux/teleslide	D Lusina, JM Martelli (Aulnay sous bois)	January-2003 (in progress)
Forum whorkshops (congresses of the "Groupe Français d'Hématologie Cellulaire") open to discussion			
Myeloproliferative diseases	Congress GFHC SMP2005, Nantes	R Garand (Nantes)	2005.05.17
B-cell lymphoproliferative syndromes	Congress GFHC SLP2007, Lyon	R Garand (Nantes)	2007.05.22
Acute leukemias and myelodysplastic syndromes	Meeting GFHC, Paris	S Daliphard (Reims)	2006.11.09
Thrombopenia et thombopathies (excluding malignancies)	Meeting GFHC, Paris	S Daliphard (Reims)	2007.11.28

Acknowledgment

Jean-François Lesesve and Richard Garand are for the Groupe Ouest-Est d'Etude des Leucémies Aiguës et autres Maladies du Sang (GOELAMS).

References

[1] G. Flandrin and L. Benattar, "Télépathologie: avantages et perspectives, contraintes et limites. Exemple de l'hématologie," *Feuillets de Biologie*, vol. 42, no. 242, pp. 5–9, 2001.

[2] J. M. Girardel, N. Chioukh, and N. Cabanel, "Numérisation d'images hématologiques, confrontation à distance et démarche qualité," *Spectra Biologie*, vol. 141, pp. 45–50, 2004.

[3] A. Brion, B. Mahé, B. Kolb, et al., "Preliminary results of a prospective randomized study comparing first line conventional therapy versus high dose therapy with autologous CD34+ purified progenitor cell support in B CLL patients with B and C Binet stages: the GOELAMS LLC98 study," *Blood*, vol. 106, no. 11, abstract 48a, 2005.

[4] V. Leymarie, G. Flandrin, M. E. Noguera, F. Leymarie, B. Lioure, and S. Daliphard, "Telehematology: a pilot experience of cytological diagnosis of acute myeloid leukemia via the Internet. A GOELAMS study," *Haematologica*, vol. 91, no. 9, pp. 1285–1286, 2006.

[5] U. Luethi, L. Risch, W. Korte, M. Bader, and A. R. Huber, "Telehematology: critical determinants for successful implementation," *Blood*, vol. 103, no. 2, pp. 486–488, 2004.

[6] Z. Agha, R. S. Weinstein, and B. E. Dunn, "Cost minimization analysis of telepathology," *American Journal of Clinical Pathology*, vol. 112, no. 4, pp. 470–478, 1999.
[7] G. Flandrin, "Télépathologie et banque d'images en hématologie," *Hematologie*, vol. 3, no. 3, pp. 256–259, 1997.
[8] G. Flandrin, "Télédiagnostic et banque de données en hématologie. Une étape préalable: les réunions de consensus avec microscope et vidéo," *Nouvel Hospitalier*, vol. 1, pp. 32–33, 1995.
[9] T. Dennis, R. D. Start, and S. S. Cross, "The use of digital imaging, video conferencing, and telepathology in histopathology: a national survey," *Journal of Clinical Pathology*, vol. 58, no. 3, pp. 254–258, 2005.
[10] F. Trimoreau, N. Gachard, M. Donnard, et al., "Réseau ville / hôpital de télétransmission en cytologie: l'expérience de la région Limousin," *Hematologie*, vol. 10, 2004, résumé 13-03.
[11] J. X. Corberand, "HEMATimage: logiciel de formation continue en hématologie cellulaire avec transfert d'images numériques," *Feuillets de Biologie*, vol. 40, no. 231, pp. 49–54, 1999.