Information Modeling for Real-Time Decision Support in Intensive Medicine

SANTOS, M.F.[1], PORTELA, F.[1], VILAS-BOAS, M.[1], MACHADO, J.[2], ABELHA, A.[2], NEVES, J.[2]
[1]Departamento de Sistemas de Informação
[2]Departamento de Informática
Universidade do Minho
Guimarães
PORTUGAL
[1]{mfs,cfp,mvb}@dsi.uminho.pt; [2]{jmac,abelha,jneves}@di.uminho.pt

SILVA, A., RUA, F.
Serviço de Cuidados Intensivos
Hospital de Santo António
Porto
PORTUGAL
moreirasilva@clix.pt, fernandorua.sci@hgsa.min-saude.pt

Abstract: - Intensive medicine is a very attractive field for applying knowledge discovery in databases due to the great amount of data that is gathered everyday in intensive care units. It is known that one can extract previously unknown knowledge from that data in order to create prediction and decision models. The challenge is to perform those tasks in real-time, in order to assist the doctors in the decision making process. Moreover, the models can be continuously assessed and optimized, if necessary, to maintain a certain accuracy. In this paper we present an adjustment to the INTCare system, an intelligent decision support system for intensive medicine and propose an information model to support it. We focus on the automatization of data acquisition avoiding human intervention, describing its steps and some requirements.

Key-Words: - Real-time data acquisition, knowledge discovery in databases, intensive care, INTCare, intelligent decision support systems, information models.

1 Introduction

Daily, a great amount of data related to patients' condition is collected. The databases that store that data may have precious but unknown information regarding the patients' prognosis. Critical patients' condition is so complex, that sometimes even doctors find it hard to decide about the most adequate procedure to provide them the best health care possible.

In order to improve the health care, allowing the physicians to take a pro-active attitude in the patients' best interest, a real-time and situated intelligent decision support system, called INTCare[1] is being developed [5], [18].

INTCare includes models induced by means of Data Mining techniques and is capable of predicting organ failure probability, the outcome of the patient for the next-day, as well as the best suited treatment to apply [1], [5], [6].

Due to the new fine-grained time response requirements, is very useful to have models to predict values for the next hour, which means that the system should be adapted to real-time data.

This paper relates to the information model needed to respond to those necessities.

This paper is organized as follows. In section 2 we present some background concepts and also present some related work. In section 3 we define some requirements for the INTCare system and in section 4 we present our proposal for the information model. There is some discussion in section 5 and in section 6 we conclude this paper and point future work.

2 Background and related work

2.1 Intelligent decision support systems

[1] The INTCare project is financially supported by FTC (PTDC/EIA/72819/2006).

According to Turban [2], a Decision Support System (DSS) is an interactive, flexible and adaptable information system, developed to support a problem solution and to improve the decision making. These systems usually use artificial intelligence techniques and are based on prediction and decision models that analyze a vast amount of variables to answer a question.

The decision making process can be divided in five phases: Intelligence, design, choice, implementation and monitoring [2]. Usually it is used in the development of rule based DSS [3]. However, these DSS are not adaptable to the environment in which they operate. To address this fault, Michalewicz [4] introduced the concept of adaptive business intelligence. The main difference between this and the regular decision support systems is that it includes optimization that enables adaptability. An Adaptive Business Intelligent (ABI) system can be defined as "the discipline of using prediction and optimization techniques to build self-learning 'decisioning' systems. ABI systems include elements of data mining, predictive modeling, forecasting, optimization, and adaptability, and are used to make better decisions." [4].

As it is known, predictive models' performance tends to degrade over time, so it is advantageous to include model re-evaluating on a regular basis so as to identify loss of accuracy [1] and enable their optimization.

There is a particular type of DSS, the real-time decision support systems. Ideally, the later includes adaptive behavior, supporting the decision making in real time.

To achieve real-time DSS, there is a need for a continuous data monitoring and acquisition systems. It should also be able to update the models in real time without human intervention [5]. In medicine, most systems only use data monitoring to support its activities.

2.2 Knowledge Discovery from Databases

Knowledge Discovery from Databases (KDD) is one of the approaches used in BI. According to Negash [9], BI systems combine data gathering, data storage, and knowledge management with analytical tools to present complex and competitive information to planners and decision makers. KDD is an interactive and nontrivial process of extracting implicit and previously unknown and potentially useful and understandable information from data [10].

The KDD process is divided in 5 steps: Selection, pre-processing, transformation, data mining, interpretation/evaluation [11]. This process starts with raw data and ends with knowledge.

The automation of the knowledge acquisition process is desirable and it is achieved by using methods of several areas of expertise, like machine learning [6]. The knowledge acquisition takes advantage of KDD techniques, simplifying the process of decision support [1].

2.3 Intensive medicine

Intensive medicine can be defined as a multidisciplinary field of the medical sciences that deals with prevention, diagnosis and treatment of acute situations potentially reversible, in patients with failure of one or more vital functions [12]. These can be grouped into six organic systems: Liver, respiratory, cardiovascular, coagulation, central nervous and renal [13].

Intensive care units (ICU) are hospital services whose main goal is to provide health care to patients in critical situations and whose survival depends on the intensive care [14], [15]. In the ICU, the patients vital signs are continuously monitored and its vital functions can be supported by medication or mechanical devices, until the patient is able to do it autonomously [14].

Clinical intervention is based on the degree of severity scores that allow the evaluation of the patient's condition according to a predefined set of values [16].

The assessment of these severity scores are based on several medical data acquired from bedside monitors, clinical analysis and clinical records.

2.4 INTCare system

INTCare is an intelligent decision support system for intensive medicine that is being developed in the ICU of the Hospital Santo António in Porto, Portugal. It makes use of intelligent agents [5] that are capable of autonomous actions in order to meet its goals [17] [18].

In Fig. 1 we can see the conceptual design of the INTCare system, which is divided into four subsystems [5]: data entry, knowledge management, inference and interface.

The data entry sub-system is responsible for the activities related to data acquisition that will gather all required data into a data warehouse. The later will be used by the agents in charge of the knowledge maintenance.

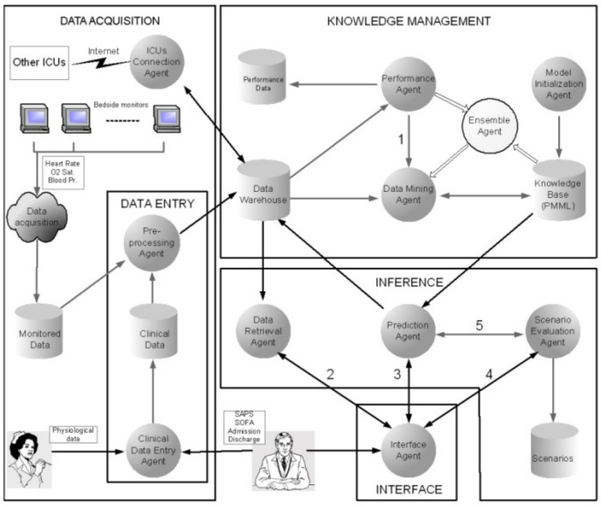

Fig. 1 – INTCare System

A data acquisition sub-system was added and it is responsible for the data gathering that will feed the data entry sub-system (mainly, the data warehouse).

The knowledge management sub-system maintains the prediction models used by the inference sub-system, assessing and updating the models when necessary. At last, the interface sub-system is the responsible for the interaction between the doctors and the system [5].

This system will predict, in real-time, organ failure and mortality assessment and, according to these predictions, it will suggest therapeutic treatment. In this paper we will focus on the Data Acquisition module, because now we are using real-time data acquisition, which has three different types of data sources and therefore, some changes need to be made.

3 Information requirements

In order to model information for KDD processing, there are some requirements that should be met:

Online Learning - The system must to act online, i.e., the models should be induced using online data in opposition of an offline approach, where the data is gathered and after processed;

Real-Time - The system must to act in real-time, the data acquisition and storing must be in made immediately after the events. By other hand, the decisions must be taken whenever an event occurs;

Adaptability - The system needs to have the capacity to, automatically, optimize the models with the new data. This information is obtained from the analysis of the results archived;

Data mining models - The success of IDSS depends, among others, on the acuity of the data mining models, i.e., the models of prediction must be reliable. Those models make it possible to predict events and avert some clinical complications to the patients;

Decision models - The achievement of the best solutions depends heavily on the decision models created. Those are based in factors like differentiation and decision that are applied on prediction models and can help the doctors to choose the better solution on the decision making process;

Optimization – The data mining models need to be optimized over time. With this, the models are in continuous training so that increasingly accurate and reliable solutions are returned, improving the models acuity;

Intelligent agents - This type of agents makes the system work through autonomous actions that execute some essential tasks. Those tasks support some modules of system: Data Acquisition, Knowledge Management, Inference and Interface. The flexibility and efficiency of this kind of system emerges from the agents and their interaction [18].

The features described above conducted to following requirements:

R1 - Fault tolerance capacities;
R2 - Processing to remove null and noisy data;
R3 - Continuous data acquisition process;
R4 - Time restrictions for the data acquisition and storage;
R5 - Online learning mode;
R6 - Digital data archive in order to promote the dematerialization of paper based processes (e.g., nursing records);
R7 - Database extension to accommodate the new data structures;
R8 - Correct usage of the equipment that collects the vital signs.

4 Information model

In order to follow the requirements enumerated above, an information model was drawn Fig. 2 presents the data acquisition module including three types of information sources:

- Bedside monitors;
- Clinical analysis; and
- Nursing records.

All sources can produce information to the INTCare system and that information can be used to develop predicting models in Intensive Care (knowledge).

The first type of sources relates to data acquisition from bedside monitors. This acquisition is in real-time, the data is received through a gateway, and it is stored on a table by an agent.

At the moment, this information is collected approximately once per minute. This method collects information from the HL7 messages (Sending Application, Sending Facility, Receiving Application, Message Control ID), patient information (Process ID, Name, Location), and patient observations and results (Observation Identifier, Observation Sub-ID, Observation Value, Units).

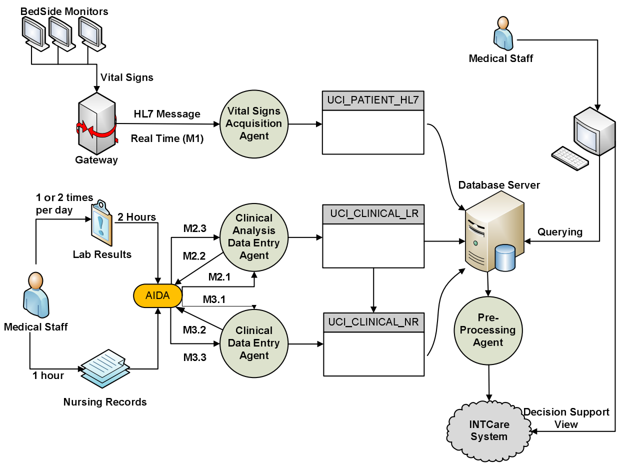

Fig. 2 – Data Acquisition Sub-system

With this type of system it is possible to obtain the results of mean, Blood Pressure (BP), Heart Rate (HR), Systolic and Diastolic, for the variables: Central Venous Pressure (CVP), Invasive Blood Pressure (IBP), Intracranial Pressure (ICP), Noninvasive Blood Pressure (NIBP), Positive Airway Pressure (PAP) and other special types of measurement.

Additionally, it receives Respiration Rate (RR), Heart Rate, Arrhythmia (ARH) and ST level from ECG, Saturation of Peripheral Oxygen (SPO_2), SPO2 HR, Cardiac Output (CO), P*ulmonary capillary wedge pressure* (*PCWP*), Cardiac Output & Wedge Pressure and Temperature.

The second type of sources is the one that contains the less frequent observations, because the patient normally does this type of analysis once or twice per day, except in extraordinary situations. With this method we can collect the data related with some clinical analysis, such as: number of Blood Platelets, Creatinine, Billirubin, SOFA, Partial Pressure of Oxygen in Arterial Blood (PaO_2) and Fraction of Inspired Oxygen (FiO_2). The medical staff usually does the clinical analysis once or twice per day, and in the next two hours the results will be available to be stored in UCI_CLINICAL table through CADE agent.

The third type is a particular one because some of the values measured by the monitors are registered in the nursing records. However, in the nursing records, there are observations and results that are required, like Urine Output, Amines dosage, PaO2/FiO2, Glasgow and Patient Output. Normally, these records are filled in a once per hour rate, in paper format. The CDEA agent will be responsible for the storage of that data into the database.

Next, the tables contained in the database are defined in terms of their attributes:

UCI_PATIENT_HL7 ⊆ {Monitored Data}
UCI_CLINICAL_LR ⊆ {Creatinine, Billirubin, Blood Platelets, PaO_2, FiO_2, SOFA}
UCI_CLINICAL_NR ⊆ {PaO_2, FiO_2, Glasgow, Urine Output, Amines, Patient Output}
UCI_DATABASE ∪ {UCI_PATIENTES_HL7, UCI_CLINICAL_LR, UCI_CLINICAL_NR}

4.1 Data Acquisition Agents

INTCare system should be rewritten in order to incorporate the new requirements. A set of new agents has been conceived. These agents will be in charge of the tasks associated to the data acquisition. Formally, INTCare can be defined as a tuple:

$$\Xi \equiv < C_{INTCare}, \Delta_{INTCare}, \boldsymbol{a}_{\boldsymbol{gat}}, \boldsymbol{a}_{\boldsymbol{vsa}}, \boldsymbol{a}_{\boldsymbol{cade}}, \boldsymbol{a}_{\boldsymbol{ada}}, a_{pp}, a_{cde}, a_{dm}, a_{pf}, a_{mi}, a_{dr}, a_{pd}, a_{sc}, a_{int}, a_{ic} >$$

The new agents are explained in the following lines.

Gateway (a_{gat}) is responsible to capture the vital signal data from bedside monitors. This data is packed into HL7 messages and sent to the Vital Signs Acquisition Agent;

Vital Signs Acquisition (a_{vsa}) is an AIDA process that parses the HL7 messages, extracts information blocks and stores them in the database tables: **UCI_PATIENT_HL7.**

The HL7 message starts with the header "MSH" and it is separated with "|" and "^". This agent needs to split the message into individual data information. The data is verified and, if the information is correct, the agent performs the next steps. For the PID, PV1, OBR and OBX variables, it reads the information from the gateway, splits the hl7 message and gets the required data to database. For optimization purpose, if more than one message is received within one minute, an algorithm is applied so that only one message per minute is stored in the database. When the number of messages received per minute is greater than 1, it calculates the average of each one of the variables and inserts them in the database as a single record:

MSG_Data_Average algorithm
If count(msg_per_min)>1 Then
avg(msg_obx_result)
Insert into database (avg_msg_obx_result)
Else
Insert into database (msg_obx_result)

Clinical Analysis Data Entry (a_{cade}) is responsible for capturing the clinical data from the lab results that are done in the hospital;

Clinical Data Entry (a_{cde}) is responsible for capturing the clinical data from the medical and nursing staff [18], especially from nursing records;

AIDA (a_{ada}) is an agency to archive and to disseminate medical exams and results. This agency will supply the lab results and nursing records through the clinical analysis data entry agent and clinical data entry agents.

4.2 Agents' messaging

The system has various agents responsible for the necessary tasks related to the data acquisition process.

The a_{vsa} agent processes the monitored data. When the gateway receives the vital signs from the monitors, sends an HL7 message (*M1*) to the vital signs acquisition agent.

Next, we can see an example of a HL7 message:

MSH|^~\&|DHV |***h2***|***h3***|***h4***|||ORU^R01|***h1***|P|2.3.1
PID|1||***d1***||***d2***
PV1|1|U|***v1***
OBR|1|||DHV|||***r1***|
OBX|***x2***|NM|***x3***^***x4***^^^***x5***||***x6***|***x7***|||||R||||***x1***^ ***v1***||

Table 1 presents the meaning of each variable involved in the exchange of messages between the agents a_{gat} and a_{vsa}.

Table 1 - HL7 message variables

h1	Version ID
h2	Sending Facility
h3	Receiving Application
h4	Receiving Facility
d1	Patient ID (Internal ID)
d2	Patient Name
v1	Assigned Patient Location
r1	Observation Date/Time
x1	Producer's ID
x2	Value Type
x3	Observation Identifier (cod)
x4	Observation Identifier (cod2)
x5	Observation Identifier (descp)
x6	Observation Value
x7	Units

The a_{ada} agent exchanges messages with the a_{cad} agent. When the a_{ada} agent receives lab results, it sends a message (*M2.1*) notifying that new data is available. The a_{cad} agent reads the message and sends one (*M2.2*) with the requested variables (Table 2, column 4). Finally, the a_{ada} agent sends the message (*M2.3*) with the required data.

When the nursing records are filled in, a_{ada} agent sends a message (*M3.1*) to the a_{cde} agent informing about the new data (Table 2, column 2). A_{cde} agent sends a message to the a_{ada} agent with the requested data and a_{ada} agent sends back a message with the required data.

Table 2 – Clinical Variables

e1	Urine Output	c1	Billirubin
e2	Glasgow	c2	Creatinine
e3	Amines	c3	Blood Platelets
e4	SOFA		

Figure 3 summarizes the agents' messaging process described above.

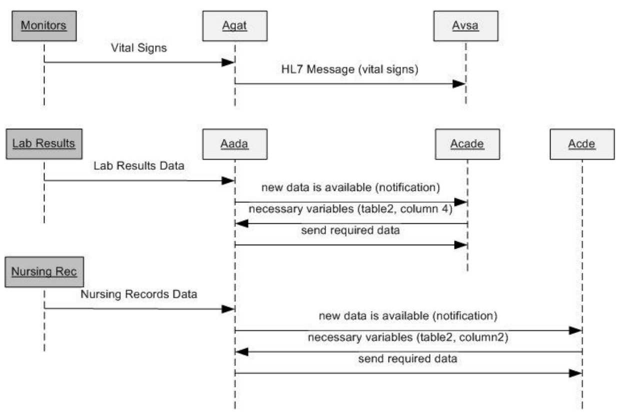

Figure 3 – Sequence diagram of the messages

5 Discussion

The INTCare system has been redesigned taking into account its new challenges. Some requirements have already been met, while others are being tested and/or implemented, as discussed below:

R1 - A local repository has been configured to work as a buffer in order to avoid information loss;

R2 – The method implemented in MSG_Data_Average Algorithm reduces in 25% the size of the table maintaining the accuracy of the monitored data;

R3 – The Gateway gathers the monitored data in a continuous mode;

R4 – The processing times of monitored data and response times for decision support have been evaluated being values in acceptable levels;

R5 – Real-time data is assured for online learning in order to evaluate and optimize the data mining models;

R6, R7- In development;

R8 - Commitment of the medical staff.

6 Conclusions and further work

This paper presented the information model necessary to adapt the system INTCare – Intelligent Decision Support System for Intensive Medicine - to real-time and online data acquisition and processing requirements. The new approach implements all the data acquisition steps in an automatic and continuous way and it is being tested in a real-world environment in the ICU of Hospital Santo António, in Oporto, Portugal. The solution encompassed the deployment of a gateway, a set of agents and data structures in order to give response to a set of requirements (R1 to R8). A discussion about how the solution met the requirements has been included.

Further work includes the dematerialization of paper based processes (e.g., nursing records) and database extension in order to accommodate the new data structures (R6 and R7 requirements).

References:

[1] P. Gago and M. F. Santos, "Towards an Intelligent Decision Support System for Intensive Care Units," presented at the 18th European Conference on Artificial Intelligence, Greece, 2008.

[2] E. Turban, *et al.*, *Decision Support Systems and Intelligent Systems*, 7 ed.: Prentice Hall, 2005.

[3] D. Arnott and G. Pervan, "A critical analysis of decision support systems research," in *Conference on Decision Support Systems*, Prato, ITALY, 2004, pp. 67-87.

[4] Z. Michalewicz, *et al.*, *Adaptive Business Intelligence*: Springer, 2007.

[5] M. F. Santos, *et al.*, "Intelligent decision support in Intensive Care Medicine," in *2nd International Conference on Knowledge Engineering and Decision Support*, Lisbon, Portugal, 2006, pp. 401-405.

[6] P. Gago, *et al.*, "Adaptive decision support for intensive care," in *13th Portuguese Conference on Artificial Intelligence*, Guimaraes, PORTUGAL, 2007, pp. 415-425.

[7] Á. Silva, *et al.*, "Organ failure prediction based on clinical adverse events: a cluster model approach," 3th International Conference on Artificial Intelligence and Applications, 2003.

[8] Á. Silva, *et al.*, "Multiple organ failure diagnosis using adverse events and neural networks," in *6 th International Conference on Enterprise Information Systems*, 2004, pp. 401-408.

[9] S. Negash and P. Gray, "Business Intelligence," *Communications of the Association for Information Systems,* vol. 13, pp. 177-195, 2004.

[10] W. J. Frawley, *et al.*, "Knowledge Discovery in Databases: An Overview," *AI Magazine,* vol. 13, pp. 57-70, 1992.

[11] U. M. Fayyad, *et al.*, "From data mining to knowledge discovery: an overview," 1996.

[12] Á. Silva, "Modelos de Inteligência Artificial na análise da monitorização de eventos clínicos adversos, Disfunção/Falência de órgãos e prognóstico do doente critico," Doutoramento, Ciências Médicas, Universidade do Porto, Porto, 2007.

[13] J. B. Hall, *et al.*, *Principles of Critical Care*: McGraw-Hill's AccessMedicine, 2005.

[14] J. Ramon, *et al.*, "Mining data from intensive care patients," *Advanced Engineering Informatics,* vol. 21, pp. 243-256, 2007.

[15] S. M. Rao and S. T., "Organization of intensive care unit and predicting outcome of critical illness," *Indian J. Anaesth,* vol. 47 (5), pp. 328-337, 2003.

[16] J. Pereira, "Modelos de data mining para multi-previsão: Aplicação à medicina intensiva," Mestrado, Universidade do Minho, Guimarães, 2004.

[17] N. R. Jennings, "On agent-based software engineering," *Artificial Intelligence,* vol. 117, pp. 277-296, 2000.

[18] P. Gago, *et al.*, "INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine," *Journal of Decision Systems,* 2006.