

Using Bayesian Networks to improve the Decision-Making Process in the Public Health System

Germanno Teles¹, Carina T. Oliveira², Luiz O. M. Andrade³, Mauro Oliveira⁴

¹State University of Ceará (UECE) - Fortaleza, Brazil

²National Research Network(RNP)

³Federal University of Ceará (UFC) - Fortaleza, Brazil

⁴ Federal Institute of Ceará (IFCE) - Fortaleza, Brazil

germanno@bnb.gov.br, carinatoliv@gmail.com, odorico@saude.gov.br, amauroboliveira@gmail.com

Abstract - This paper proposes the use of Bayesian networks to support the decision-making process in health systems governance. In particular, this paper presents LARIISA_Bay, a new component based on Bayesian networks that works together with LARIISA, a context-aware platform to support applications in public health systems. The main goal of the proposed component is to assist teams of health specialists in order to better diagnose diseases through data collected from users of LARIISA. As a case study, we focus on scenarios of dengue fever disease. We classify dengue cases into one of the following levels: emergency (i.e., dengue hemorrhagic fever), grave (i.e., dengue fever) or normal (i.e., absence of the disease). Based on this classification, teams of health specialists can accurately make decisions, for example, to alert a health care agent to visit locations with a high incidence of the disease, to send an ambulance when an dengue emergency case has occurred, as well as give technical instructions on how to deal with specific cases. We present a prototype of LARIISA_Bay and the corresponding interfaces to support the interactions of the patient, the health care agent and the specialist with the system.

Keywords—Bayesian Network, Health Systems Governance, Dengue Fever Disease

I. INTRODUCTION

Bayesian Networks are powerful and appropriate methodologies for the construction of systems that rely on probabilistic knowledge. The health system sector is an example for the use of probabilistic networks in uncertainty modeling, which this concept is used to obtain an approximated diagnosis of different diseases, such as Alzheimer's, heart disease, among others.

With the growing changes in medicine and computer science, it is noticed that became indispensable to the union of these two areas for the generation of more efficient systems and focused on the care and treatment of patients. According to Sigulem *et al.* [1], since the beginning of computing there was a significant enthusiasm regarding the use of the computer as a tool to aid medical diagnosis.

The need to systematize the variety of data produced separately, generating useful information and in a timely manner is a common problem for organizations, especially in the public sphere. In this health issue is even more suppressed given the complexity of the health / illness / care.

Considering that, most of the medical diagnosis models have a simple structure, composed of causal disease-symptom, the inference engine implemented takes advantage of the simplified model, alleviating the computational complexity.

With this mechanism in place, it is possible to make a correct identification of patients aims to optimize patient care and reduce the excessive number of unnecessary calls.

All things considered, this work presents a mechanism to assist specialists in the diagnosis of dengue cases, especially people who have symptoms and may be guided by health professionals in their own residence. Thus, it is important to accurately select the patients who must be treated immediately (dengue hemorrhagic fever) or directed to a specific hospital for further treatment and analyzes, or that can be treated in their homes in order to avoid severe problems.

From meetings with specialists this paper shows the main pathologies, the diagnostic tools and related factors related to worsening of dengue fever.

This paper presents an interface for three main actors in the system: the user, the agent and the health expert. The collected data feed the Bayesian Network of LARIISA_Bay.

This remainder of this paper is organized as follows: section 2 presents related works that make use of Bayesian networks in the health system. In Section 3, it is described the LARIISA platform, which was developed for the LARIISA_Bay. The section 3 presents the proposed model for dealing with probabilistic data in LARIISA. Section 4 shows aspects of implementation and testing with LARIISA_Bay. Finally, in section 5, concluding remarks are made about the work.

II. RELATED WORK

In this section, we present the related work that makes use of Bayesian networks in the health system context.

A. *Bayesian Agent of Support of Hospital Infection – SAVIH, in Portuguese*

This system uses a Bayesian network to support the manager of a hospital to assess the risk of a patient to a hospital infection. Shows the setting of the infection in different units and allowing forecasting the risk from an epidemiological characteristics and the patient's disease. Retrieves cases that are similar to a new

Comentado [1]: É mais indicado usar emails institucionais. Evitar gmail.

Comentado [5]: A frase está um pouco longa. Tem como dividir em 2?

Comentado [6]: Frase muito confusa. O que você queria dizer?

Comentado [7]: ??? o que você queria dizer nesse parágrafo?

Comentado [2]: A introdução dever ter: (1) contexto e motivação do trabalho (cenário geral); (2) apontar um problema nesse contexto; (3) detalhar como você abordará o problema apontado; (4) como você valida sua solução. Você consegue identificar esses 4 pontos na sua introdução?

Comentado [8]: Não é melhor usar especialista ao invés de expert?

Comentado [9]: Até agora na Introdução, você não falou do LARIISA_Bay. Só falou rapidamente no abstract. Eu acho que esse final da introdução pode ser melhorando deixando mais claro o objetivo do trabalho e o que será apresentado (resultados?).

Comentado [10]: Atualizar/revisar no final.

Comentado [3]: São vários autores: et al.

Comentado [4]: Não entendi essa frase, como seria em português?

patient in the unit. It was developed using the shell Netica modeling [2].

B. Support System for Differential Diagnosis for Cephalgia

It is a medical expert system that provides support to general practitioners, emergency room physicians or residents in the differential diagnosis for Cephalgia (headaches).

This system uses the Bayesian network approach. The base of the knowledge was build considering the International Headache Society Ranking Criteria, taking in consideration the patient's symptoms and also the estimated values of probabilities provided by experts who participated in the project. It was developed using the shell Nética modeling. This system was evaluated when comparing the experts and the system answers, from a set of medical records of patients with headache specialists randomly selected by the project. The experimental results indicate that the system was able to provide the same diagnoses that experts project in 95% of the cases. Also, general practitioners evaluated the same clinical cases and they only had 53% of correct answers. Therefore, it is clear that the developed system shows very good performance when giving differential diagnosis for headaches[3].

These studies are demonstrating the relevance of the study, in other words the relationship of the area of computing, specifically the line of research of computational intelligence with the treatment of uncertainty in support of diagnosis domains healthcare.

III. PROJETO LARIISA

LARIISA was specified taking into account specific requirements of five governance fields: Knowledge Management, Systemic Normative, Clinical and Epidemiological, Administrative and Shared Management[4]. Therefore, the system proposed in this paper provides a context-aware diagnosis based on geolocation to LARIISA, applying it to the scenario of decision-making for local and global contexts[5][13].

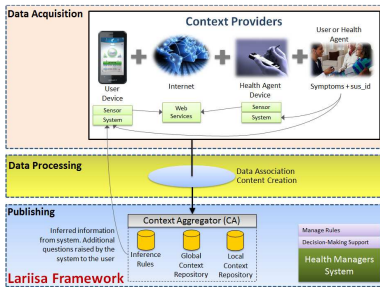


Fig. 1. Clariisa Architecture

LARIISA is centered on the concept of health context information. Based on Dey's definition of context[6], we consider health context as any information that can be used to

characterize the situation of an entity in a health system. LARIISA is able to perceive the status of emergency epidemiological and adapt itself in real time to a risk situation.

Aiming in assisting users in their day-to-day tasks, context-aware applications have been using elements of ubiquitous systems to obtain user context information. A Yesple example is the use of sensors that detect the presence of people and automatically trigger lighting to an environment, according to the people location and time.

Figure 1 presents an overview of the proposed system, which is divided in three main parts: *Data Acquisition*, *Data Processing* and *Publishing*. *Data Acquisition* concerns the sensor application, information (e.g. User ID and symptoms) added by the user, and the acquired data. After that, all acquired data will be processed in the *Data Processing* step. In this part, the system uses the raw data in order to capture all necessary diagnosis data.

IV. NOISY-OR BAYESIAN NETWORKS

Bayesian Networks (BN) are directed and acyclic graphs, that allow the representation of the joint distribution of probabilities of a set of random variables. This paper will focus on BN with discrete variables. BN with discrete variables satisfy the Markov condition[10], which states that any node in a Bayesian Network is conditionally independent of its nondescendants, given its parents.

An important aspect of a BN is its structure (topology of the graph), which enables the representation of complex relationships between graphically and intuitively variables. The graphical structure of a BN facilitates the understanding of relationships between variables in its domain, and allows the combined use of information obtained from expert knowledge with historical data to obtain the joint probability distribution of the network.

A. Noisy-or classifier

The BN can be used in classification problems in a clear and straightforward way and BN used in data sorting problems are called Bayesian Classifiers.

In the called Bayesian Classifiers with discrete variables $\{A_1, A_2, \dots, A_n, C\}$, one of them, C , is the variable class (response variable) and the others, $\{A_1, A_2, \dots, A_n\}$ are the attributes (predictor variables).

The excessive number of probabilities required for the quantification of a network is one of the greatest difficulties in the practical application of the Bayesian networks. In a node X with k categories the number of probabilities to be specified is

$$\text{Numbers of elements of the table}(X) = P(x_i) \prod_{pa(X)} n_{category\ pa(X)} \quad (1)$$

However, for certain types of nodes, these probabilities can be calculated from other instead of being specified directly. The *noisy-or* model [11] allows

such a calculation, with the restriction that the parents of the node may contribute independently to the probability of the node on which it is applied the noisy-or model, and that the combined effect of several parents contribute cumulatively on the

probability of the node. In this paper, it is also necessary that the node in question is binary, ie, has only two categories, with one category to represent true and the other one false.

Using appropriate names for nodes in networks shaping medical problems, there is one node D with category d and \bar{d} representing a disease, which causes are R_1, R_2, \dots, R_n , the probabilities for D are given by the combined table of conditional probabilities $P(D/R_1, R_2, \dots, R_n)$.

The noisy-or model allows the calculation of the joint table of conditional probabilities from the probability $P(D/R_i)$ for each parent node R_i , whilst respecting the constraint towards parents put earlier.

In the case of having only binary nodes R_i , with stating r_i (true) and \bar{r}_i (false), these conditional probabilities are known as sensitivity $P(d|r_i)$ and specificity $P(\bar{d}|\bar{r}_i)$ that often are available in public studies. However, the joint probabilities $P(D/R_1, R_2, \dots, R_n)$ are more difficult to obtain from experts or bibliographic information because they involve a high number of combinations of conditions.

The expression that represents the probability of each factor causing the disease independently is calculated as:

$$p_i = P(d|r_i \text{ alone}) = P(d|\bar{r}_1, \bar{r}_2, \dots, r_i, \dots, \bar{r}_n) \quad (2)$$

$$P(d|H) = 1 - \left(\prod_{R_i \in H^+} [1 - p_i] \right) \quad (3)$$

Where H is a configuration of (R_1, R_2, \dots, R_n) and H^+ Subset of nodes is set as true.

V. LARIISA_BAY

The main idea of the LARIISA project consists of probabilistic inferences about the data collected from the user and data held in their records, identified by SUS_ID from the expert knowledge of the area represented in the health system.

In this evolving context of the LARIISA project, this paper is concerned with how the data of the probabilistic information is treated. Therefore, this work relates both to the representation of context-sensitive information (data collected) as to the knowledge of experts, representing the ontology for LARIISA

As it can be verified on Figure 2, the component LARIISA_Bay consist the following features:

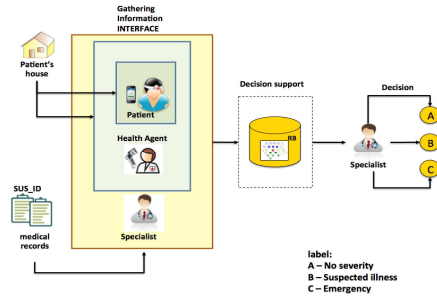


Fig. 2. - LARIISA_Bay Interface

- Interfaces with three decision makers of LARIISA: the user, the agent and the health expert. It is noticed that the interface contains the Specialist Health Agent interface, which contains the user interface (Patient)
- A broker Type, a Decision Making Support System, works as a support to help the medical team in decision making. This broker, based on Bayesian Networks, has two purposes:
 - Support the diagnosis of the medical staff, filtering probable cases of dengue in three levels of classification:
 - Normal patient (no gravity)
 - Patient with suspected dengue fever
 - Patient with suspected dengue hemorrhagic
 - Supporting the diagnosis of outbreaks/epidemics in certain areas of a municipality
- A team of health specialists able to better evaluate the selections made by the broker. This team has the responsibility of the diagnosis that would result in one of the following three procedures:
 - (A) Sending guiding procedures to the user (patient);
 - (B) Sending a Health Agent to home of the user, primarily;
 - (C) Sending an emergency team to the home of the user.

VI. NETWORK MODELING

Obtaining structural model of the network, using the knowledge of physicians and bibliographic sources, sought to identify what information regarding medical problem could be represented as a variable in the network, as well as the causal relationships between these variables. In order to perform this task specific medical information of the problem was obtained, which later were placed in the initial network structure. The iterative refinement of the initial structure, which involved the insertion and removal of nodes and arcs and other modifications led to final structure for quantification.

Using physicians knowledge and bibliographic sources, it is suspected dengue case each patient presenting acute febrile illness with a maximum duration of seven days, accompanied by

at least two of the signs or symptoms as headaches¹, retro-orbital pain, myalgia², arthralgia³, prostration⁴ or rash⁵, with or without the presence of bleeding or bleeding with positive epidemiological history, having been in the last 15 days in an dengue transmission area or has the presence of Aedes aegypti mosquito[12].

Other pathological symptoms observed are:

- Severe and continuous abdominal pain

The patient does not support superficial abdominal palpation and may prevent movement of sitting and walking.

- Orthostatic hypotension and / or syncope⁶;

- Persistent Vomiting;

The patient does not tolerate the ingestion of any substance including water.

- Painful hepatomegaly;

The patient has an increased liver size.

- Bleeding from mucosa;

The patient has bleeding in respiratory mucosa and / or digestive system

- Drowsiness and / or irritability;

The patient with difficulty in maintaining alertness and psychomotor agitation presents

- Decreased diuresis;

Due to the dehydration caused by vomiting and / or hypotension

- Hypothermia;

Difficulty in maintaining normal body temperature.

- Sudden increase in hematocrit;

Proportional imbalance of water and blood cells.

- Abrupt platelet decrease;
- Respiratory Distress

The patient has difficulty breathing, which may be due to bleeding or lung edema.

A. Risk Factors Related to the Disease

The occurrence of dengue this directly connected to some factors risks where the main are:

- places without adequate sanitation
- poor/bad garbage collection system places

¹ Medical term for pain in the head
² Term used to explain muscle pain in any part of the body
³ It means joint pain
⁴ It is complete physical and mental exhaustion, total immobility and lack of reaction of external requests.

B. Diagnostic Tools

The accurate diagnosis of dengue is carried by serology. However, this test is performed in laboratories and hospitals. Therefore, rapid diagnosis is important in order to referring patients for treatment. The main tests are:

- Serology;
- Blood Count.

C. Evolution of built models

The network construction was completed in two steps in an iterative manner, obtained two models as results described below.

The first model built is shown in Figure 3, this model contains only part of dengue fever direct pathologies.

The main purpose of this model is pre-select which network nodes and show that such nodes can indicate another symptom. It follows the model and decides to withdraw the network node peritonitis, since there is no added value in the discovery of diagnostic

Starting from this study, we worked on improving the network seeking to identify and classify patients with symptoms of dengue fever or dengue hemorrhagic and it is demonstrated in Figure 4

The separation of nodes that are informed from the users, health workers and experts is also clear in the figure below, showing the two steps to be taken in obtaining the information.

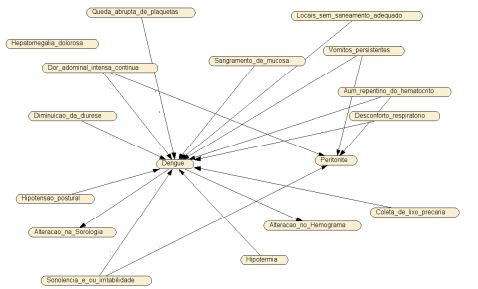


Fig. 3. First model of the network created from Netica

⁵ Skin Eruption usually red that happens because of the dilation of blood vessels
⁶ Act of standing with his foot on the floor and legs extended

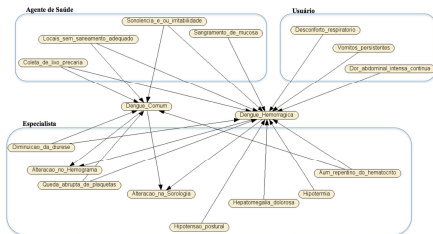


Fig. 4. Second model of the network created from Netica

Observed in the models created, some symptoms could indicate the severity of dengue. In relation to categories, all nodes were defined only two: Yes and No. This representation is appropriate when there is a disease with many factors or causes a disease. For example, for node Dengue Hemorrhagic Fever network constructed in this work, the symptoms are:

Abrupt_platelet_decrease, Sudden_increase_of_hematocrit, Hypothermia, Respiratory_Distress, Decreased_diuresis, irritability_and_drowsiness, Painful_hepatomegaly, Persistent_Vomiting, Orthostatic_hypotension, poor_garbage_collection_system_places, Severe_and_continuous_abdominal pain, places_without_adequate_sanitation and Bleeding_from_mucosa. In this paper, the model was used noisy-or nodes for dengue_hemorrhagic and Dengue_fever.

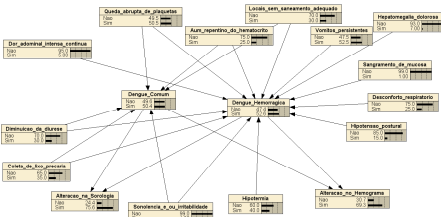


Fig. 5. Final model of the network created from Netica

With the proceeding of the meeting were able to create the final model of the network, Figure 5, which was used as the basis of inferences used in LARIISA_Bay

C. Obtening Probabilities

For the quantification of the Bayesian Network presented in this paper was used the direct description of probabilities, using tables that be expert physician filled in, from their own professional experience and bibliographic data. Such method the facilitated the quick acquiring of the probabilities, followed by the correction of some of the value of these probabilities that were incoherent to the evaluation performed.

TABLE I. PRIOR PROBABILITY

Node	Yes	No
Abrupt_platelet_decrease	1%	99%

Node	Yes	No
Painful_hepatomegaly	7%	93%
Dor_adominal_intensa_continua	5%	95%
Decreased_diuresis	30%	70%
Orthostatic_hypotension	15%	85%
Irritability_and_drowsiness	1%	99%
Hypothermia	40%	60%
Poor_garbage_collection_system_places	35%	65%
Respiratory_Distress	25%	75%
Sudden_increase_of_hematocrit	25%	75%
Persistent_Vomiting	5%	95%
Places_without_adequate_sanitation	30%	70%

Only three nodes have conditional probability related respectively to the parents in the LARIISA_Bay network. For the detailing of this probability the nodes were divided in two groups: diagnosed exams and disease symptoms. The nodes of the group disease symptoms represent the disease characteristics that, even though it is difficult to measure, it is indispensable for the modeling. Architecture

TABLE II. CONDICIONAL PROBABILITY

Variable Condition Categories	Blood_Count_Change	
	Yes	No
Dengue = Yes	87%	13%
Dengue = No	62%	38%

TABLE III. CONDICIONAL PROBABILITY

Variable Condition Categories	Serology_Change	
	Yes	No
Dengue = Yes	99%	1%
Dengue = No	9%	91%

VII. MODELING THE LARIISA AND LARIISA_BAY

LARIISA_Bay is a mechanism to support decision making based on probabilistic data for dengue fever disease. This is an aggregate component to the architecture of LARIISA project that aims to support experts in the diagnosis of dengue in tort.

The contributions of this work was the creation of a new architecture for LARIISA. As shown in Figure 8, this architecture consists of the following components:

- Entry System;
 - Interface composed by the patient, health worker and specialist;
 - Sensors the patient's vital signs;
 - Other providers of the context;

- Dialogue with the patient;
- Triggering the health care agent;

Fig. 6. Proposed System Architecture

1. Decision Specialist: scenario that considers the existence of a team of experts able to better diagnose the injury of dengue from information received from the Interface Module Decision LARISSA Bay. From this information, the result of the inference module features the team would take the most appropriate decision in dealing with a particular patient;
2. Validation Specialist: intermediate scenario between the two mentioned above. In this scenario it is estimated that a shortage of professionals for decision making. However, it is considered that the information provided by the Interface Module Decision can be filtered / validated rather than analyzed as in the first scenario;
3. "Pass Thought": scenery diametrically opposed to earlier regarding availability of specialized staff for

It is noteworthy that in any of the above scenarios, the component acts LARIISA_Bay with a reducing the sample space.

Fig. 7. Functional architecture of LARISSA-BAY

Also on figure 6, the user interface is a subset of the health agent interface. This last one is a subset of the expert interface. Figure 7 shows the screens and reports associated respectively to the user, health agent and expert profiles.

- This component assists a team of specialists to better diagnose dengue cases as collects data from system users, classifying them as: emergency (dengue hemorrhagic fever), severe (presence of dengue fever) or normal (absence of the disease). From this classification,

the team of experts can make the decision more accurately, for example, the immediate action of sending an ambulance or to ask a health care agent to visit the patient or, simply, giving instructions on how to deal with the case, accordingly to the classification

- The creation of structure of the Bayesian Network and the quantification of probabilities were done with the help of experts. They were obtained based on bibliographic data and through meetings with specified physician. The probabilities were obtained through technical clarification, questionnaires and subsequent correction of the initial values.
- The proposed model of this interface involves starting from mobile interfaces to the placement of metadata that are able to feed the table of probabilities to Bayesian Network created from consults to various expert professionals of dengue fever worsening. Three interfaces developed (user, health agent and expert) are related to three different metadata that translate the different levels of access to information and decision ability. That is, even though the final decision (dengue fever diagnosed) is the expert's, the health agent and the user are also decision-makers in restricted cases.

Finally, it is important to highlight that the same solution proposed in the paper can be analyzed and evaluated taking as reference an ontological domain model built to meet the strategic goals of the study area. The research of hybrid mechanism, using Bayesian Network and ontology, makes the prediction more refined and enable indirect inferences that are difficult to obtain without a modeling based on ontology. In this case, the Bayesian Network result could feed the ontological model, e.g. its data base, and from there, generate the resulting inference to the LARIISA ontology, using semantics research strategies in the base information of modeling with ontology.

This work tries to developing a design registered for the Information Technology Department of SUS (DATASUS, in Portuguese) for management governance in health.

REFERENCES

- [1] Sigulem D., Anção M. S., Ramos M. P., Leão B. F. Sistema de Apoio à Decisão em Medicina. 1998 [Accessed 28/04/2014]; Available in: http://www.virtual.epm.br/material/tis/curr-med/sad_html/sistema.htm
- [2] Venicius, M. Agente Bayesiano de Apoio à Vigilância de Infecção Hospitalar - SAVIH. Master's Thesis, Graduate Program in Computer Science, UFSC. 2002.
- [3] Toledo, R. V. Sistema de Apoio ao Diagnóstico Diferencial de Cefaléia. Master's Thesis, Graduate Program in Computer Science, UFSC. 2000.
- [4] Oliveira, M., Andrade O. M., Hairon C. G., Moura R. C., Fernandes S., Bringel J., Gensel J., Martin H., Sicotte C., Denis J. L.. A ContextAware Framework for Health Care Governance Decision-Making Systems: A model based on the Brazilian Digital TV. Second IEEE Workshop on Interdisciplinary Research on E-health Services and Systems (IREHSS).
- [5] Gardini, L. M. et al. Clariisa, a Context-Aware Framework Based on Geolocation for a Health Care Governance System. 5TH International Conference on e-Health Networking, Application & Services. 2013.
- [6] Dey, A. K.; Abowd, G. D.. Towards a Better Understanding of Context and Context-Awareness. In: Workshop on the what, who, where, when and how of context-awareness, CHI, Abril 2000.
- [7] Sistema Único de Saúde do Brasil. [Accessed 22/03/2014]; Available in: http://portal.saude.gov.br/portal/saude/Gestor/area.cfm?id_area=944##.
- [8] Breaga, R. B., Costa, S. M. M., Viana, W., Andrade, R. M. C., Martin, H.. "Captain: A context-aware system based on personal tracking", in: The 17th International Conference on Distributed Multimedia Systems/ DMS 2011, KSI, Florence, 2011.
- [9] W3C. 2010. OWL-Guide. World Wide Web Consortium Homepage. [Accessed 09/17/2013]; Available from: <http://www.w3.org/TR/owl-guide/>
- [10] Neapolitan, R. E. Learning Bayesian Networks. New Jersey , 2004. 674 p.
- [11] Heckerman, D.; Breese, J. A new look at causal independence. roceedings of the 10th Conference on Uncertainty in Artificial Intelligence.: Morgan Kaufmann. 1994. p. 226-286.
- [12] MINISTÉRIO DA SAÚDE. Programa Nacional de Controle da Dengue. Portal da Saúde, [Accessed 09/04/2013]; Available from: http://portal.saude.gov.br/portal/saude/visualizar_texto.cfm?idtxt=23614
- [13] OLIVEIRA, M., Andrade O. M., et al. Applying Ontology And Context Awareness Concepts On Health Management System: A Dengue Crisis Study Case, IADIS International Conference e-Health 2013, 24-26 July, Prague, Czech Republic