Mobile Data Mining for Intelligent Healthcare Support

Pari Delir Haghighi, Arkady Zaslavsky, Shonali Krishnaswamy, Mohamed Medhat Gaber
Center for Distributed Systems and Software Engineering
Monash University, Australia
{pari.delirhaghighi, arkady.zaslavsky, shonali.krishnaswamy,, mohamed.gaber}@infotech.monash.edu.au

Abstract

The growth in numbers and capacity of mobile devices such as mobile phones coupled with widespread availability of inexpensive range of biosensors presents an unprecedented opportunity for mobile healthcare applications. In this paper we propose a novel approach for Situation-Aware Adaptive Processing (SAAP) of data streams for smart and real-time analysis of data. The implementation and evaluation of the framework for a health monitoring application is described.

1. Introduction

Recently, innovations in mobile communications and low-cost of wireless biosensors have paved the way for development of mobile healthcare applications that provide a convenient, safe and constant way of monitoring of vital signs of patients. A key in the provision of mobile healthcare services is the issue of using technological innovation to support continuous monitoring of patient conditions, providing a degree of self-diagnosis and enabling effective real-time decision making to reduce fatalities. Ubiquitous Data Stream Mining (UDM) techniques [1] such as lightweight, one-pass data stream mining algorithms [2-3] can perform real-time analysis on-board small/mobile devices while considering available resources such as battery charge and available memory. However, to perform smart and intelligent analysis of data on mobile devices, it is imperative for adaptation strategies to factor in contextual information.

Contextual information can be related to a network, application, environment, process, user or device. As a meta-level concept over context we define the notion of a situation that is inferred from contextual information [4]. Situation-awareness provides applications with a more general and abstract view of their environment rather than focusing on individual pieces of context. Situation-aware adaptive data stream mining leverages the full potential of UDM by going beyond mere available resources and can enable, if not guarantee, the continuity and consistency of the running applications.

In real-world, situations evolve and change into other situations (e.g. 'healthy' changes to 'hypertension'). Changes that occur between situations are also good indicators of situations that may emerge – albeit with some vagueness and uncertainty. To enable situation-awareness in mobile healthcare applications, it is important for the situation modeling and reasoning approach to represent uncertainty and vagueness associated with health-related situations.

Reviewing recent works in mobile healthcare reveals that most of these projects [5-8] have mainly focused on using, enhancing or combining existing technologies and context-aware projects [9-13] mostly deal with a limited scope (i.e. not applicable to other context-aware scenarios). In mobile healthcare computing, a general approach for modeling and reasoning about uncertain, health situations and performing smart and cost-efficient analysis of data in real-time has not been introduced and is an open issue.

In this paper we propose situation-aware adaptive processing (SAAP) of data streams for mobile healthcare applications. The novelty and contribution of this project are as follows: i) situation-awareness is achieved by Fuzzy Situation Inference (FSI) that combines fuzzy logic principles with the Context Spaces (CS) model, a formal and general context modeling and reasoning for pervasive computing environments. The strengths of fuzzy logic for modeling of vague situations are combined with the CS model's underlying theoretical basis for supporting context-aware pervasive computing scenarios; ii) SAAP incorporates situation-awareness into data stream mining and provides gradual tuning of data streaming parameters according to occurring situations and available resources. This approach improves data stream mining operations in an intelligent and cost-efficient manner. The SAAP approach enables continuity and consistency of running operations that are of high important for health monitoring applications that deal with sensitive and critical data.

1.1. A Scenario

John has had a heart attack and is released from hospital but there are concerns that he might be susceptible to another heart attack and is also experiencing blood pressure fluctuations. Constant monitoring of his vital signs could help him to reduce his anxiety, decrease the need for routine visits to medical facilities, and also detect early warning features of a possible impending event. He has a smart phone with SAAP installed on it and is willing to wear biosensors to measure his vital signs. The data is wirelessly sent to his mobile where SAAP detects any changes not only in his vital signs but in any contextual information that is related to the application (e.g. the battery level of the mobile phone). SAAP uses this information to reason about situations in real-time and according to inferred situations, it performs intelligent and cost-efficient analysis of data. When fluctuations of vital signs are within a specified "acceptable" threshold, there is no need for frequent measurement and use of resources can be reduced and moderated. However, when these fluctuations are over the threshold, this "situation" warrants a closer monitoring by the system and more frequent measurements. This type of adaptation requires factoring in both available resources and criticality of health situations.

This paper is structured as follows: Section 2 discusses the related work. Section 3 presents the SAAP architecture. Section 4 describes the Fuzzy Situation Inference (FSI) that enables situation-awareness. Section 5 discusses the adaptation engine. Section 6 and 7 describes implementation and evaluation respectively. Finally section 8 concludes the paper and discusses the future work.

2. Related Work

Mobile healthcare computing is a new and evolving area of research that exploits the recent development in mobile networks and communications for health monitoring applications. EPI-MEDICS [5] is a large scale European project that provides personal monitoring of ECG signals for early detection of cardiac ischemia and arrhythmia and generating different levels of alarms. Another European project called the MobiHealth project [6] uses 2.5 (GPRS) and 3G (UMTS) technologies to integrate all the sensors and actuators into a wireless network called Body Area Network (BAN). The project of ubimon (Ubiquitous Monitoring Environment for Wearable and Implantable Sensors) [7] aims to provide continuous management of patients mainly focusing on sensors and wireless technology rather than data analysis techniques. Personalization is another area of focus in developing mobile health monitoring applications that has been studied in [8].

Context-awareness is one of the key requirements of health monitoring systems that enables autonomous operations without patient's intervention and enhances decision making of healthcare professionals on patient condition [9]. However, there are limited researches that have attempted to fully address the context-awareness or provide a general and formal representation of context [10-12]. One of the works in mobile healthcare that incorporates both context-awareness and adaptation is proposed in [13] but the paper does not provide the details of how and when the proposed adaptation strategies are applied. Studies in data stream processing [14-15] are very application-specific and focus on very limited areas of research. A general approach for smart and cost-efficient analysis of data for mobile healthcare systems has not been introduced in the current state-of-the-art and is still an open issue.

3. Situation-Aware Adaptive Processing (SAAP) of Data Streams

The architecture for Situation-Aware Adaptive Processing (SAAP) of data streams consists of three components of Fuzzy Situation Inference (FSI), Resource Monitor (RM) and Adaptation Engine (AE) as shown in Figure 1.

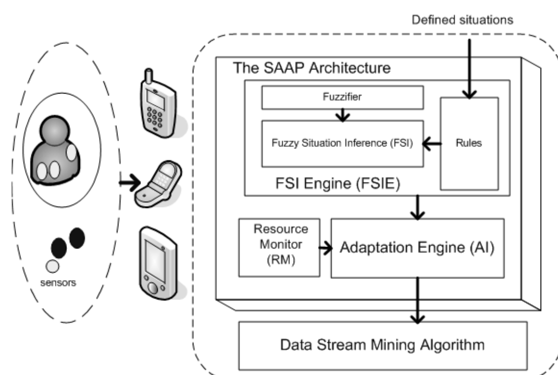

Figure 1. The architecture of SAAP (Situation-Aware Adaptive Processing) of Data Streams

The FSI engine enables situation-awareness using fuzzy logic principles. Resource Monitor (RM) is a software component that continuously monitors available resources such as available memory and battery usage and reports their availability to the adaptation engine. The Adaptation Engine (AE) is responsible for gradual tuning of data stream processing parameters in real-time according to the occurring situations and available resources. The

SAAP layer is built on the top of the data stream mining algorithms running on mobile devices and provides them with situation-aware adaptation. The next section discusses the FSI technique.

4. Fuzzy Situation Inference (FSI)

FSI is a situation modeling and reasoning approach that integrates fuzzy logic into the Context Spaces (CS) model [4]. FSI uses the benefits of the CS model for supporting pervasive computing environments while incorporating fuzzy logic to deal with uncertainty associated with vague and real-world situations.

4.1. The Context Spaces model

The Context Spaces model (hereafter CS) represents contextual information as geometrical objects in multidimensional space called situations [4]. The concept of a 'situation space' is characterized by a set of regions. Each 'region' is a set of acceptable values of a context attribute that satisfies a predicate. In addition to basic concepts and techniques for situation modeling and reasoning, the CS model provides heuristics developed specifically for addressing context-awareness under uncertainty. These heuristics are integrated into reasoning techniques that are utility-based data fusion algorithms and compute the confidence level in the occurrence of a situation [16]. The CS deals with uncertainty mainly associated with sensors' inaccuracies. Yet there is another aspect of uncertainty in human concepts and real-world situations that needs to be represented by the context model and reflected in the results of situation reasoning. Fuzzy logic uses multi-value logic and has the benefit of dealing with this level of uncertainty by assigning membership degrees to values.

4.2. Situation Modeling

FSI consists of three subcomponent including fuzzifier, rules and inference engine. Fuzzifier, as a software component, maps crisp input (i.e. values of context attributes) into fuzzy sets using trapezoidal membership functions. In a fuzzy set, membership of an item is gradual and is represented by a degree between 0 and 1 [17]. In FSI, situations of interest are defined using fuzzy rules by domain experts and stored in a rule repository. Each FSI rule consists of multiple conditions joined with the AND operator but a condition can itself be a disjunction of conditions [18]. To model the importance of conditions, we assign a weight w to each condition with a value ranging between 0 and 1. The sum of weights is 1 per rule. A weight represents the importance of its assigned condition relative to other conditions in defining a situation. An example of a FSI rule is as follows:

if Room-Temperature is 'hot' and Heart-Rate is 'fast' and (Age is 'middle-aged' or 'old) then situation is 'heat stroke'

The next subsection discusses situation reasoning.

4.3. Situation Reasoning

To reason about a situation, rules need to be evaluated to produce a single output that determines the membership degree of the consequent [19]. The conditions joined with the OR operator are evaluated using the maximum function. However, to evaluate the conditions joined with the AND operator, FSI provides four reasoning techniques as shown in Table 1.

Table 1. Reasoning techniques

Heuristic: weights and contribution level	
CS	$Confidence = \sum_{i=1}^{n} w_i c_i$
FSI	$Confidence = \sum_{i=1}^{n} w_i \mu(x_i)$
Heuristic: sensors' inaccuracy	
CS	$Confidence = \sum_{i=1}^{n} w_i . \Pr(\hat{a}_i^t \in A_i)$
FSI	$Confidence = \sum_{i=1}^{n} w_i \mu(f(x_i, e_i))$
Heuristic: symmetric and asymmetric attributes	
CS	$Confidence = \sum_{i=1}^{n} \hat{w}_i . \Pr(\hat{a}_i^t \in A_i)$ where $a_i \in CA_S \cup CA_A$
FSI	$Confidence = \sum_{i=1}^{n} \hat{w}_i \mu(f(x_i, e_i))$ where $x_i \in FS$ and $FS \in LV_S \cup LV_A$
Heuristic: partial and complete containment	
CS	$Confidence = q_1 \sum_{i=1}^{n} \hat{w}_i . p(\hat{a}_i^t \in A_i) + q_2 \prod_{k=1}^{m} p(\hat{a}_k^t \in A_k)$ where $q_1 + q_2 = 1$ and $a_i \in CA_S \cup CA_A, a_k \in CA_S$
FSI	$Confidence = q_1 \sum_{i=1}^{n} \hat{w}_i . \mu(f(x_i, e_i) + q_2 \prod_{k=1}^{m} \mu(f(x_k, e_k))$ where $q_1 + q_2 = 1$ and $x_i \in FS$, $FS \in LV_S \cup LV_A$ $x_k \in FS$, $FS \in LV_S$

These techniques integrate fuzzy logic into the CS reasoning methods to provide another aspect of uncertainty (i.e. uncertainty of situations and delta changes of context) in the computation of confidence value for the occurrence of a situation.

The situation reasoning techniques of CS are based on four heuristics that are introduced to manage uncertainty in pervasive computing environments. These heuristics are as follows: i) relative weights of context attributes and confidence level of values; ii) sensors' inaccuracy; iii) symmetric and asymmetric context attributes; iv) and partial and complete containment of symmetric context attributes. Table 1 depicts reasoning methods of CS, their FSI equivalent that are combined with fuzzy logic and their underlying heuristics and theoretical concepts.

The next subsections discuss each heuristic and reasoning technique in more detail.

4.3.1. Weights and contribution level. The first reasoning technique of CS is based on the weights of context attributes and the level of confidence of attributes' values. Weights are assigned to context attributes and represent relative importance of each context attribute for inferring a situation. Level of confidence is assigned to each element and reflects how that element relates to the modeled situation. In this heuristic, the contribution function that computes the contribution level is proposed at a conceptual level and its implementation is later introduced in the second reasoning technique based on sensors' inaccuracy.

In FSI, the concept of weights is associated with linguistic variables (i.e. context attributes). The concept of contribution level is similar to the membership degree of elements in a fuzzy set but they are implemented using membership functions. The result of $w_i \mu(x_i)$ represents a weighted membership degree of x_i and *n* represents the number of conditions in a rule (1≤i≤n).

4.3.2. Sensors' inaccuracy. To provide automatic computation of the contribution level at run-time, the second reasoning method of CS uses the impact of sensor inaccuracies and unreliability as a determining factor to compute the contribution level. This method computes the probability of a context attribute correct value $\hat{a}_i^t$ being contained in the region A_i. To compute the probability value based on the reliability of a sensor, the reliability of reading (e.g. 95%) is used to represent the probability value (i.e. =0.95).

The second option to compute the probability value is to integrate the sensors' inaccuracy of reading rather than the reliability of reading. Using this option, the probability value is calculated in the following format:

$$1)\ \Pr(e_j \le a_i^t - \min(A_i^j)) - \Pr(e_j \le a_i^t - \max(A_i^j)).$$

where a_i^t denotes the sensed value of the context attribute, e_j represents the sensor reading error (i.e. a_i^t - $\hat{a}_i^t$) and $\min(A_i^j)$ and $\max(A_i^j)$ represent minimum and maximum values of the region.

The second reasoning method of CS deals with uncertainty factoring in inaccuracies of sensors however this equation does not reflect delta changes of values in the equation and is not adequate to reason about vague situations. The FSI equivalent technique not only incorporates the contribution level associated with sensors' inaccuracy but includes the membership of the values as another factor affecting the contribution level. In the FSI model, we first calculate the correct value based on the reliability or error rate and then pass it to the membership function. The function *f* calculates the correct value of the context based on the inaccuracy value e_i. If e_i is a reliability rate, the sensed value is multiplied by it and if it is an error rate (i.e. ±) it is added to the sensed value.

4.3.3. Symmetric and asymmetric context attributes. The third reasoning technique of CS introduces the concepts of symmetric context attribute CA_S and asymmetric context attribute CA_A. A symmetric context attribute increases the confidence in inferring a situation if its value is within the corresponding region and decreases the confidence if it is outside that region (e.g. reasoning about the 'hypertension' situation based on 'blood pressure'). An asymmetric context attribute increases the confidence in inferring a situation if its value is within the corresponding region but would not decrease the confidence if it is outside that region (e.g. reasoning about the 'heat stroke' situation based on 'age').

Whenever an asymmetric attribute is not contained within its region, the redistribution method assigns 0 to the weight of the attribute and recalculates the relative weights for the remaining attributes as follows.

$$2)\ \hat{w}_i = w_i / \sum_{i=1}^{n} w_i$$

The concept of symmetric and asymmetric attributes and its corresponding reasoning technique is applied into FSI (as shown in Table 1). However, since values that linguistic variables take are not numeric (i.e. these values are called terms that represent fuzzy sets), the concept of symmetric and asymmetric concepts are applied to the values of fuzzy sets associated with linguistic variables.

4.3.4. Partial and complete containment. The fourth heuristic deals with the fact that the value of an important context attribute should affect the result of the situation inference more than the other attributes (i.e. less important ones) and when several attributes are significant for the evaluation of a situation we may want to ensure that all of them are contained in their regions. This heuristic has been integrated into the fourth reasoning technique that aims to address the trade-off between complete containment of all symmetric context attributes (i.e. when all values of symmetric attributes are contained in their corresponding regions) and their individual contribution using the third reasoning technique. This heuristic does not apply to asymmetric attributes because they do not decrease the confidence for the occurrence of a situation.

To address the trade-off between complete and partial containment, the fourth reasoning technique presents each aspect of containment with a dimension using utility weights (i.e q_1 and q_2) and combines them towards inferring the occurrence of a situation. The utility weights of two dimensions determine which aspect of containment is more important (i.e. complete or partial).

The concept of partial and complete containment and its reasoning technique are applied to FSI. Similar to the third reasoning method, FSI maps values of symmetric context attributes into the values of fuzzy sets corresponding to symmetric linguistic variables.

Results of situation reasoning using the techniques discussed earlier suggests the degree of confidence in the occurrence of a situation. In FSI, if the output of a rule evaluation for the 'hypertension' situation yields a degree of 0.885, we can suggest that the level of confidence in the occurrence of 'hypertension' is 0.885. This value can be compared to a confidence threshold ε between 0 and 1 (i.e. predefined by the application's designers) to determine whether a situation is occurring.

The next section discusses the component of the AE (Adaptation Engine).

5. Adaptation Engine (AE)

The AE (Adaptation Engine) is responsible for gradual tuning of data stream processing parameters according to the occurring situation/s and available resources in real time. Lightweight data stream mining techniques such LWC, LWCLass, RA-Cluster, ERA-Cluster, and DRA-Cluster [2-3, 20-22] are adaptive to availability of resources via adjusting the algorithm parameters. These parameters control output, input and/or the process of the algorithm. In these algorithms, the adaptation process is done through Algorithm Granularity (AG) approach.

AG has three different variations of AOG (Algorithm Output Granularity), AIG (Algorithm Input Granularity) and APG (Algorithm Processing Granularity) [21-22]. AOG controls the algorithm output rate based on the availability of memory via changing the data stream mining algorithm parameters to encourage or discourage the creation of new output structures. Similarly AIG and APG [22] control the input rate and consumption of processing power according to the battery level and CPU usage respectively.

We have inspired by the concepts of AG and developed three different adaptation strategies. These strategies include resource-aware, situation-aware and hybrid strategies as shown in Figure 2.

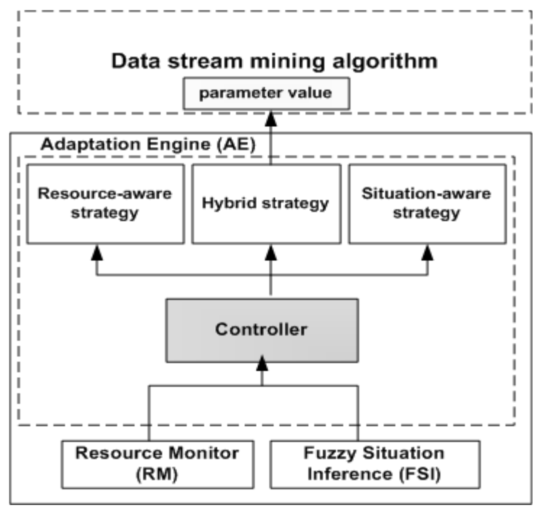

Figure 2. Adaptation of data stream mining

AE constantly monitors occurring situations that are inferred by FSI and availability of resources reported by RM.

Each pre-defined situation needs to be assigned a criticality value (i.e. a value between 0 and 1) that indicates their importance. For both situations (S) and computational resources (R), there are two thresholds (i.e. lower and upper bounds), a value between 0 and 1, which indicate safe, medium and critical levels. The higher the value is, the higher the situation importance and resource usage is. Based on these levels of criticality for situations and resources, there can be nine possible variations (cases) of adaptation at run time. Controller that is a subcomponent of AE makes decisions on which strategy needs to be performed according to these thresholds. These nine cases are presented in Table 2. We have allocated the adaptation strategies according to these nine cases. When resources are critical it means that the mobile device can not continue the mining operations and the adaptation strategies that we provide are not adequate to address the issue. Therefore other strategies such as migration of the data or the process need to be performed which are out of the scope of this project.

Table 2. Adaptation Cases

Cases	Adaptation strategy
1- if R at safe level and S at safe level	Situation-aware strategy
2- if R at safe level and S at medium level	Situation-aware strategy
3- if R at safe level and S at critical level	Situation-aware strategy
4- if R at medium level and S at safe level	Resource-aware strategy
5- if R at medium level and S at medium level	Hybrid strategy
6- if R at medium level and S at critical level	Hybrid strategy
7- if R at critical level and S at safe level 8- if R at critical level and S at medium level 9- if R at critical level and S at critical level	Other strategies e.g. migration

5.1. Resource-aware Adaptation Strategy

Resource-aware adaptation strategy occurs when the situation is at safe level but resource availability is at medium level. This is because normal situations do not require frequent monitoring and the results of resource-aware adaptation do not contradict the requirements of normal situations. Resource-awareness is inspired by the AG approach. One of the AOG-based clustering algorithms is called LightWeight Clustering (LWC) [29]. LWC considers a threshold distance measure for clustering of data. Increasing this threshold discourages forming of new clusters and in turn reduces resource consumption.

AOG is a three-stage, resource-aware distance-based mining data streams approach. The process of mining data streams using AOG starts with a mining phase. In this step, a value of threshold distance measure is determined. This threshold has the ability to control the output rate of the running mining algorithm.

The second stage in AOG-mining approach is the adaptation phase. In this phase, the threshold value is adjusted to cope with the data rate of the incoming stream, available memory, and time constraints to fill the memory with generated knowledge (data mining output).

The last stage in AOG approach is the knowledge integration phase. This stage represents the merging of generated results when the memory is full. This integration allows the continuity of the mining process on resource-constrained devices.

The next subsection discusses situation-aware adaptation strategy based on the results of the FSI.

5.2. Situation-aware Adaptation Strategy

Situation-aware adaptation in AE is performed when resources are available and at safe level. Situation-aware adaptation occurs based on occurring situations inferred by FSI. These results are multiple situations with different level of confidence. To provide a fine-grained adaptation and reflecting the level of confidence of each situation in the adaptation phase, we compute weighted average of the data mining parameter value based on confidence values of situations and the pre-set value of the parameter for each situation. The pre-set values of parameters are automatically calculated based on the importance values of the situations that will be discussed further in the evaluation section. The situation-aware adaptation enables reflecting all the results of situation inference in the adaptation of parameter values and is represented as follows:

3) $\hat{p}_j = \sum_{i=1}^{n} \mu_i p_j / \sum_{i=1}^{n} \mu_i$

where p_j represents the set value of a parameter for a pre-defined situation S_i, μ_i denotes the membership degree of situation S_i where 1≤i≤n and n represents the number of pre-defined situations, and $\hat{p}_j$ represents aggregated value of the parameter.

Situation-aware adaptation itself results in cost-efficiency because when a situation has a lower importance value, the computed set value for the threshold will be a higher value. This decreases the output of the LWC algorithm and reduces the memory consumption.

The next subsection describes hybrid adaptation strategy.

5.3. Hybrid Adaptation Strategy

When resources are at medium level and situations are at medium or critical level (i.e. cases 5 and 6 in Table 2), the controller applies the hybrid adaptation strategy. When the adaptation cases 5 or 6 occurs, resource-aware and situation-aware adaptation strategies each compute different values according to resource availability and occurring situations respectively. Therefore there is a trade-off between the results of these two strategies. Hybrid adaptation strategy addresses this issue by computing the average value of parameter based on the results of the two strategies and criticality values of the situation and resource as follows:

$$4)\ \hat{p}_I = \frac{(\hat{p}_R.criticality_R)+(\hat{p}_S.criticality_S)}{criticality_R + criticality_S}$$

Having discussed the theoretical framework of our work, the following section presents the implementation and evaluations we have performed.

6. Implementation

We have implemented a prototype of health monitoring application based on FSI in J2ME and deployed it on a Nokia N95 (shown in Figure 3). The prototype reasons about situations of 'normal', 'pre-hypotension', 'hypotension', 'pre-hypertension' and 'hypertension'. This application can be used by patients who suffer from blood pressure fluctuations. A trapezoidal membership function is used to compute membership degree of context values. Contextual information used includes systolic and diastolic blood pressure (SBP and DBP) and heart rate (HR).

Figure 3. The prototype of SAAP-based health monitoring application with an ECG biosensor

To capture the patient's heart rate, we have used a two lead ECG biosensor from Alive Technologies [23] that transmits ECG signals using Bluetooth to the mobile phone. For the blood pressure, we have used randomly generated data that simulates blood pressure fluctuations. The health monitoring application performs situation reasoning and situation-aware adaptation in real-time on the mobile device using the LWC algorithm. Status bars on the mobile phone displays the level of certainty and confidence in the occurrence of each situation.

The evaluation of FSI and adaptation engine is presented in the next section.

7. Evaluation

For evaluation of SAAP, we have performed two evaluations. First evaluation is a comparative evaluation of FSI, CS and Dempster-Shafer and second evaluation focuses on the adaptation of threshold parameter of LWC according to occurring situations.

7.1. Evaluation of FSI

To evaluate the FSI model, we have compared the FSI situation reasoning technique to the CS and Dempster-Shafer (hereafter DS) reasoning approaches. The purpose of this evaluation is first to validate the FSI model against a well-known reasoning technique such as DS and a context model developed for pervasive computing environments such as CS. The second objective of the evaluation is to highlight the benefits of the FSI to deal with uncertain situations.

In this evaluation, we have considered situations of 'hypotension', 'normal' and 'hypertension'. These situations are defined using context attributes of systolic blood pressure (SBP) and diastolic blood pressure (DBP) with the scale of 40-170 and 20-150 mm Hg and heart rate (HR) with the range of 20-150 bpm.

Table 3 depicts modeling of the three situations in the CS model including the weights of attributes and their corresponding regions of values. Assigned weights are 0.4 for SBP and DBP and 0.2 for HR.

Table 3. Situation definitions in CS

Situation	Context attribute	Region of values
Hypotension	1=SBP	≤85
	2=DBP	≤60
	3=HR	≤45
Normal	1=SBP	>85 and ≤135
	2=DBP	>60 and ≤110
	3=HR	>45 and ≤85
Hypertension	1=SBP	>135
	2=DBP	>110
	3=HR	>85

The modeling of the three situations in the FSI model is presented in Table 4. Weights of conditions for the FSI rules conform to the weights used in CS.

Table 4. Situation definitions in FSI

Situation	Linguistic Variable	Terms
represented	1=SBP	low, normal, high
below via FSI	2=DBP	low, normal, high
rules	3=HR	slow, normal, fast
Rule1: if SBP is low and DBP is low and HR is low then situation is hypotension		
Rule2: if SBP is normal and DBP is normal and HR is normal then situation is normal		
Rule3: if SBP is high and DBP is high and HR is high then situation is hypertension		

To apply the DS algorithm for reasoning about situations, we use the Dempster's rule of combination. The normalized version of the combination rule is as follows.

$$5)\quad m(R) = \frac{\sum_{P \cap Q = R} m_i(P).m_j(Q)}{1 - \sum_{P \cap Q = \phi} m_i(P).m_j(Q)}$$

where m(R) denotes the mass value computed for a proposition R given the evidences i and j. If R represents a situation, considering all existing propositions, the intersection of some of these propositions denoted as P and Q results in the proposition R (i.e.) and the intersection of other combinations of propositions results in an empty set.

To model the three situations of Hypotension (L), Normal (N) and Hypertension (H) with DS, we first need to define propositions and events. Since all three situations are incompatible we include a proposition of Unknown (U) that would consist of three situations. Then we identify the events and mass values that reflect the association of an event with the occurrences of each proposition as depicted in Table 5.

Table 5. Definitions of events and mass values

Event	N	L	H	U
SBPLow (40-85)	0	0.7	0	0.3
SBPMed(86-135)	0.7	0	0	0.3
SBPHigh(136-180)	0	0	0.7	0.3
DBPLow(20-60)	0	0.7	0	0.3
DBPMed(61-110)	0.7	0	0	0.3
DBPHigh(110-130)	0	0	0.7	0.3
HRSlow(20-45)	0.2	0.4	0	0.4
HRMed(46-85)	0.4	0.2	0.2	0.2
HRFast(86-130)	0.2	0	0.4	0.4

Mass values are assigned in a way that they reflect to what degree each event indicates a situation. Since we have based our situations on three context attributes, we define three mass functions corresponding. Then we apply DS combination over all propositions and evidence.

The dataset used for the evaluation consists of 131 context states and their scales contribute to the occurrence of each pre-defined situation as well as the uncertain situations that occurs when situations evolve.

Figure 4 presents the results of the evaluation of CS, DS and FSI for situation reasoning about 'hypotension', 'normal' and 'hypertension'.

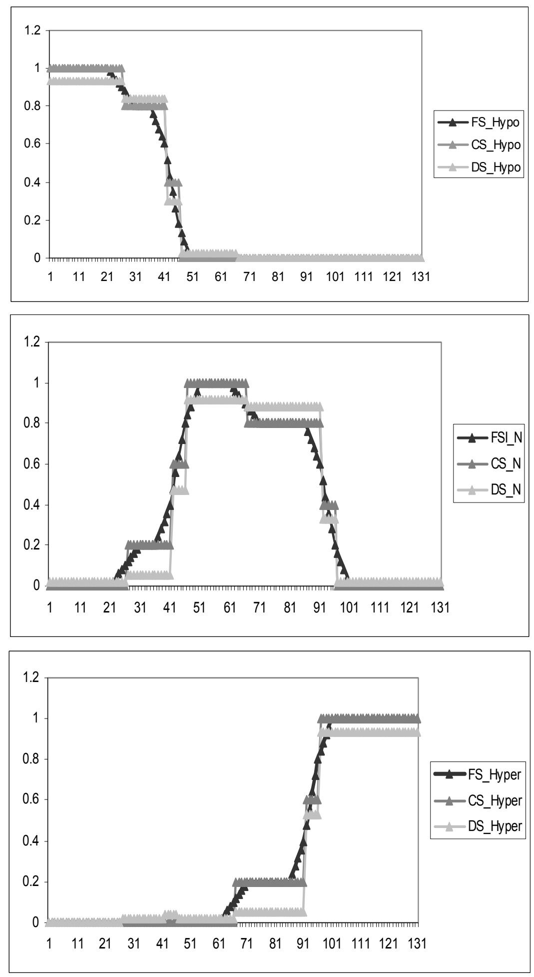

Figure 4. Results of the evaluation

Figure 4 shows three approaches of CS, DS and FSI have a relatively similar trend according to context changes. When the data corresponds to a pre-defined situation the results of three approaches almost overlap. However, when changes of data indicate the occurrence of an unknown and uncertain situation, differences of reasoning results between CS, DS and FSI are more apparent.

Compared to FSI, the results of situation reasoning by the CS and DS methods show sudden rises and falls with sharp edges when situations change which do not match the real-life situations. This is because the DS and CS approaches do not deal with delta changes of the values and are not able to reflect the gradual evolution of one situation to another situation. When the value of context attributes decreases or increases, its membership degree also increases and decreases accordingly and gradually. This enables FSI to provide more accurate situation reasoning results in terms of reflecting very minor changes of context.

The evaluation validates the accuracy of the FSI model for situation modeling and reasoning and it also shows that FSI is able to reflect very minor changes of context in situation inference and represent changes in a more gradual and smooth manner. The evaluation shows that the FSI model is more appropriate approach for representation of human concepts and for reasoning about the real-world situations that are defined by continuous values. Health-related situations are examples of these types of scenarios where FSI can prove to be more fitting approach compared to the DS and CS reasoning approaches.

7.2. Evaluation of Situation-Aware Adaptation

In the implementation of the SAAP we have used the LightWeight Clustering (LWC) [29] algorithm as the data stream mining algorithm. This algorithm is one-pass and operates using the AOG principals as discussed earlier in the paper. The LWC algorithm provides adaptability by adjusting the parameter of threshold distance measure according to the available memory on a device such as a PDA. In the evaluation of situation-aware adaptation, we have adjusted the parameter of threshold of LWC according to the confidence level of the occurring situations. The values of LWC threshold for each situation are computed based on the importance value of each situation and the minimum and maximum values of the threshold (i.e. 6 and 45 respectively) using the following formula:

threshold=minValue+(maxValue-minValue)(1-importance)*

Using the above formula, if we assign the situations of ‘normal’, ‘hypertension’ and ‘hypotension’ the importance values of 0.1, 0.9 and 0.5, the computed threshold values of each situation will be 42, 10 and 26 respectively. These values are acceptable given a variation of 12 (i.e. 42 divided by 3) for any of the context attributes of SBP, DBP and HR has no significant impact on a healthy individual while a variation of 3 for ‘hypertension’ can be significant.

To evaluate the situation-aware adaptation, we have used the same 131 context states used for the first evaluation. Figure 5 shows that the threshold value is adjusted according to the confidence value of each situation. Decreasing the threshold value increases the number of the output (clusters) that is required for closer monitoring of more critical situations.

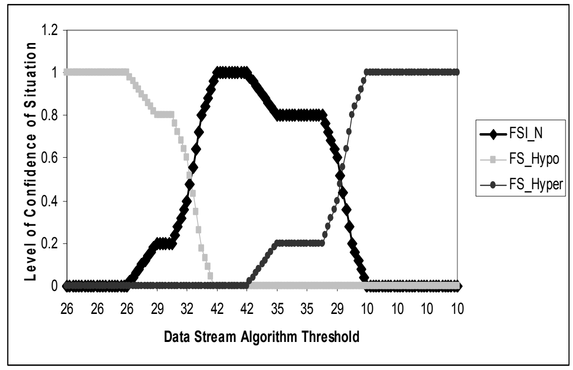

Figure 5. Situation-Aware Adaptation Results

The next section concludes the paper and discusses future work.

8. Conclusion and Future Work

In this paper we proposed and validated a general approach for situation-aware adaptive processing (SAAP) of data streams that incorporates situation-awareness into data stream processing using fuzzy logic. The fuzzy situation inference model allows modeling and reasoning about real-world and health-related situations. The SAAP architecture enables real-time analysis of data emanating from multiple sensors including bio-sensors onboard mobile devices while factoring in contextual/situational information and resource availability. This approach significantly enhances a range of mobile healthcare applications.

There are several directions in which we are extending this work. We are currently finalizing implementation and evaluation of hybrid adaptation using RA-Cluster [22] that enables adaptation of the parameters of radius threshold, randomization factor and sampling rate according to the memory, CPU and battery usage respectively. Furthermore, we are working on extensive testing of our prototype in real-world situation in conjunction with relevant healthcare professionals and domain experts in order to develop an understanding of high risk situations for the monitoring of patients and identifying what information is required from bio-sensors.

10. References

[1] M.M. Gaber, S. Krishnaswamy, and A. Zaslavsky, “Ubiquitous Data Stream Mining”, Current Research and Future Directions Workshop Proceedings held in conjunction

with The Eighth Pacific-Asia Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 2004.

[2] M.M. Gaber, S. Krishnaswamy, and A. Zaslavsky, “Resource-Aware Mining of Data Streams”, Journal of Universal Computer Science. 11(8), 2005, pp.1440—1453.

[3] M.M. Gaber, A. Zaslavsky and S. Krishnaswamy, “A Cost-Efficient Model for Ubiquitous Data Stream Mining”, Proceedings of the Tenth International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Perugia Italy, 2004.

[4] A. Padovitz, S. Loke, and A. Zaslavsky, “Towards a Theory of Context Spaces”, In Proceedings of the 2nd IEEE Annual Conference on Pervasive Computing and Communications, Workshop on Context Modeling and Reasoning (CoMoRea), Orlando, Florida: IEEE Computer Society, 2004.

[5] D. Konstantas, V. Jones, and R. Bults, “MobiHealth-Innovative 2.5/3G Mobile Services and Applications for Healthcare”, 2007

[6] P. Rubel, J. Fayn, G. Nollo, D. Assanelli, B. Li, L. Restier, S. Adami, S. Arod, H. Atoui, M. Ohlsson, L Simon-Chautemps, D. Te´lisson, C. Malossi, G. Ziliani, A. Galassi, L. Edenbrandt, and P. Chevalier, “Toward Personal eHealth in Cardiology: Results from the EPI-MEDICS Telemedicine Project”, Journal of Electrocardiology 2005, 38: pp. 100-106.

[7] K. Laerhoven, L. Benny, J. Ng J., S. Thiemjarus, R. King, S. Kwan, H. Gellersen, M. Sloman, O. Wells, P. Needham, N. Peters, A. Darzi, C. Toumazou, G. Yang, "Medical healthcare monitoring with wearable and implantable sensors", UbiHealth 2004: The 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare Applicationsm, 2004.

[8] P. Leijidekkers, and V. Gay, “Personal Heart Monitoring and Rehabilitation System using Smart Phones”, In Proceedings of the International Conference on Mobile Business (ICMB'05), 2005.

[9] P. Vajirkar, S. Singh, Y. Lee, “Context-Aware Data Mining Framework for Wireless Medical Application”, Lecture Notes in Computer Science (LNCS), Volume 2736, Springer-Verlag. ISBN 3-540-40806-1, 2003, pp. 381 – 391.

[10] B. Jansen, and R. Deklerek, “Context-Aware Inactivity Recognition for Visual Fall Detection”, In Pervasive Health Conference and Workshops, 2006.

[11] Kim, H., and Jo, H. “A Context-Aware Traveler Healthcare Service (THS) System”, In Pervasive Health Conference and Workshops. November-December 2006.

[12] M. Becker, E. Werkman, M. Anastasopoulos, and T. Kleinberger, “Approaching Ambient Intelligent Home Care Systems”, In Pervasive Health Conference and Workshops, 2006.

[13] H. Mei, I. Widya, A. Halteren and B. Erfianto, “A Flexible Vital Sign Representation Framework for Mobile Healthcare”, 2007.

[14] G. Brettlecker, H. Schuldt, and R. Schatz, “Hyperdatabases for Peer-to-Peer Data Stream Processing”, In Proceedings of the IEEE International Conference on Web Services (ICWS'04). San Diego, California, 2004..

[15] C. Chen, H. Agrawal, M. Cochinwala and D. Rosenbluth, “Stream query processing for healthcare bio-sensor applications”, In Proceedings of the 20th International Conference on Data Engineering (ICDE'04). Boston, MA, USA: IEEE Computer Society, 2004.

[16] A. Padovitz, A. Zaslavsky and S. Loke, “A Unifying Model for Representing and Reasoning About Context under Uncertainty”, 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU), Paris, France), 2006.

[17] Jang, J.R., Sun, C. and Mizutani E., Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Prentice-Hall: Upper Saddle River, NJ, 1997.

[18] Bruce, G. Buchanan B.G. and Shortliffe, E.D., Rule-based expert systems : the MYCIN experiments of the Stanford Heuristic Programming Project. Reading, Mass: Addison-Wesley, 1984.

[19] Zimmermann, H., Fuzzy Set Theory - and Its Applications. Kluwer Academic Publishers, Norwell, Massachusetts, 1996.

[20] M.M. Gaber, S. Krishnaswamy, and A. Zaslavsky, “On-board Mining of Data Streams in Sensor Networks”, A Book Chapter in Advanced Methods of Knowledge Discovery from Complex Data, (Eds.) S. Badhyopadhyay, U. Maulik, L. Holder and D. Cook, Springer Verlag.

[21] M.M. Gaber, S. Krishnaswamy, and A. Zaslavsky, “Adaptive Mining Techniques for Data Streams Using Algorithm Output Granularity”, The Australasian Data Mining Workshop (AusDM 2003), Held in conjunction with the 2003 Congress on Evolutionary Computation (CEC 2003), Canberra, Australia, Springer Verlag, Lecture Notes in Computer Science (LNCS), 2003.

[22] N. Phung, M.M. Gaber and U. Roehm, “Resource-aware Distributed Online Data Mining for Wireless Sensor Networks”, Proceedings of the International Workshop on Knowledge Discovery from Ubiquitous Data Streams (IWKDUDS07), in conjunction with ECML and PKDD 2007, Warsaw, Poland, 2007.

[23] Alive Technologies, http://www.alivetec.com