# **Context-Aware Adaptive Data Stream Mining**

Pari Delir Haghighi<sup>a</sup>, Arkady Zaslavsky<sup>a</sup>, Shonali Krishnaswamy<sup>a</sup>, Mohamed Medhat Gaber<sup>b</sup>, and Seng Loke<sup>c</sup>

<sup>a</sup> Center for Distributed Systems and Software Engineering, Monash University, Australia
 <sup>b</sup> CSIRO ICT Center, Australia

<sup>c</sup> Department of Computer Science and Computer Engineering, La Trobe University, Australia

Corresponding author: Email: Pari.DelirHaghighi@infotech.monash.edu.au address: PO Box 7074, Faculty of Information Technology, Monash University 900 Dandenong Rd, Caulfield East, VIC - 3145, Australia Tel: +613 9903 1402 Fax +613 99031077

Abstract. In resource-constrained devices, adaptation of data stream processing to variations of data rates and availability of resources is crucial for consistency and continuity of running applications. However, to enhance and maximize the benefits of adaptation, there is a need to go beyond mere computational and device capabilities to encompass the full spectrum of context-awareness. This paper presents a general approach for context-aware adaptive mining of data streams that aims to dynamically and autonomously adjust data stream mining parameters according to changes in context and situations. We perform intelligent and real-time analysis of data streams generated from sensors that is under-pinned using context-aware adaptation. A prototype of the proposed architecture is implemented and evaluated in the paper through a real-world scenario in the area of healthcare monitoring.

Keywords: Data mining, context, fuzzy logic

#### 1. Introduction

There is a range of emerging applications that use mobile devices for data analysis and processing of data streams in ubiquitous computing environments such as mobile healthcare, supporting field workforce and Intelligent Transportation Systems (ITS). A very significant challenge for these applications is to process and analyze the vast amounts of data streams that are generated at very high rates with mobile devices such as a PDA in real-time and in a smart and cost-efficient way. Ubiquitous Data Stream Mining (UDM) [8] is the process of analyzing data originated from heterogeneous sources and sensors with mobile devices and has the potential to perform real-time analysis of sensory data onboard resource-constrained devices. However, to perform intelligent and cost-efficient analysis of data, there is a need to go beyond mere computational and device capabilities to encompass the full spectrum of context-awareness.

The review of the current state-of-the-art of data stream mining in resource-constrained environments indicates that there are data stream mining algorithms [12, 15, 16] that are able to function on devices such as PDAs but they have limited ability to cope with a multitude of changing contextual factors. A context-aware adaptation approach leverages the full potential of UDM and can provide continuity and consistency of the running application by adapting parameters of mining algorithms according to context changes.

In this paper we introduce a novel approach for context-aware adaptive data stream mining that provides real-time adaptation of data stream mining parameters according to current context/situation and performs intelligent analysis of data streams. The first contribution of this project is that it provides a general technique for modeling and reasoning about context at a higher level of abstraction called situations. We define the notion of a situation as a meta-level concept over context that is inferred from contextual information [18]. Situation-awareness provides applications with a more abstract and wider view of their environment rather than focusing on individual pieces of context [1].

We have based our context modeling and reasoning on the CS (Context Spaces) model [18-20] but extended this model using fuzzy logic, called the Fuzzy Situation Inference (FSI) Model. The FSI technique incorporates the CS model's underlying theoretical basis and reasoning formulas for supporting context-aware and pervasive computing environments while using fuzzy logic principles to address uncertainty associated with vague situations. Fuzzy logic has been applied to achieve context-awareness in [3, 4, 6, 17, 22] but without a supporting context model for sensory originated information.

The second contribution of this work is to integrate context-awareness into data stream mining and provide real-time gradual adaptation of data stream mining parameters according to context changes. Light weight algorithms such as LWC and RA-Cluster [9-11, 21] enable resource-aware data stream mining by adjusting certain parameters of the data stream mining algorithms according to resource availability. In our project we use these parameters that control the input, output and process of the mining operations for context-aware adaptation. This approach maximizes the benefits of adaptation and improves mining operations in an intelligent and proactive manner.

This paper is structured as follows: Section 2 reviews related works in the area of ubiquitous data mining. Section 3 provides a general view of our proposed architecture for context-aware adaptive data stream mining. Section 4 discusses our preliminary approach based on the Context Spaces model and its formula for finding similar situations. Section 5 describes our extended work that integrates fuzzy logic into the CS model and uses the results of fuzzy inference for adaptation of data stream mining parameters. Section 6 discusses the implementation and evaluation of the prototype system we have built. Finally section 7 concludes the paper and discusses the future work.

# 2. Related Work

Data streams generated in wireless sensor networks can be processed on sensor nodes called as in-network processing [5], on a high performance computing facility at a central

stationary site [13] or on mobile devices such as a PDA. Data stream mining algorithms and systems developed for mobile users/applications predominantly target devices such as handheld computers/PDAs.

Kargupta et al. [15] have developed a client/server data stream mining system called MobiMine which focuses on data stream mining applications for stock market data but the mining component is located at the server side rather than the PDA. In the Vehicle Data Stream Mining System (VEDAS) [16] the mining component is located at the PDA onboard the moving vehicle. A clustering technique has been used for analyzing and detecting the driver behavior. In [12], real-time mining of information from in-vehicle sensors to minimize driver distraction is proposed through adaptation of the instanceselection process based on changes to the data distribution. While this research recognizes the implicit need for adaptation, it is focused on intelligent sampling with little consideration for resource availability. In [24], context factors are used to filter datasets used for mining (rather than awareness of the application environment) and prune the user query to produce more accurate results. The final results are sent to a PDA but all the processing is performed on the server side.

Data streams are generated at very high rates that challenge computational resources at both the networked servers and on resource-constrained devices such as mobile phones. Adapting to data rates is an important factor in continuity and consistency of data streaming tasks. It has been experimentally demonstrated in [23] that even efficient data stream mining algorithms such as Very Fast K-Means (VFKM) can cause applications to fail when executed on devices such as PDAs without awareness to context such as variations in data rates and computational resource availability.

One of the innovative generic adaptation strategies developed for data stream mining on resource-constrained devices is Algorithm Output Granularity (AOG) [7] that adapts the output size of algorithms based on available memory. While UDM techniques (such as AOG) can provide adaptation of data stream mining operations based on resource availability, it is of paramount significance to the success of performing real-time processing of data streams to factor in contextual/situational information. A context-aware adaptive approach that could adjust input, output and processing parameters of data stream mining algorithms on mobile devices has not been introduced in the current stateof-the-art and is still an open issue.

### 3. An Architecture for Context-Aware Adaptive Data Stream Mining

The architecture for context-aware adaptive data stream mining consists of two main components as illustrated in Fig. 1. The first component is a situation manager that provides context-awareness and includes subcomponents for context/situation modeling and inference. The second component is the adaptation manager that is responsible for adjusting mining parameters based on current situations.

In our project, situation-awareness can be achieved by two different approaches. These two approaches are both based on the Context Spaces (CS) model [18] that is a formal and general context modeling and reasoning approach for pervasive environments. The first technique uses the CS model's basic concepts for context modeling and reasoning and applies the distance measure formula [20] for finding similar situations. The second approach is called the Fuzzy Situation Inference (FSI) model that complements the CS model using fuzzy logic to provide more flexibility for modeling vague situations.

First we discuss the situation and adaptation manager components using our initial approach based on the CS model. Then we describe these components applying fuzzy inference mechanisms.

## 3. Distance Measure-based Approach

This section describes the situation and adaptation manager components based on the CS model. In this approach, the situation manager applies the distance measure formula of the CS model for finding similar situations and the adaptation manager uses the results of distance measure for adaptation of data stream mining parameters.

#### 3. 1 Situation Manager

In this approach, the situation manager consists of three subcomponents: Situation Modeling, Situation repository and Situation Inference Engine.

### 3.1.1 Situation Modeling

The core of the Context Spaces (CS) Model is the concept of situations. The CS model represents contextual information as geometrical objects in multidimensional space called situations [18]. A situation space is a tuple of regions of attribute values related to a situation. Each region is a set of accepted values for an attribute based on a pre-defined predicate and each context state a collection of values of context attributes at the given time. For a health monitoring application, context attributes could include: systolic blood pressure (SBP)<sub> $a_1$ </sub>, diastolic blood pressure (DBP)<sub> $a_2$ </sub> and heart rate (HR)<sub> $a_1$ </sub>.

Weights are values from 0 to 1 assigned to every context attribute and they represent the importance of each context attribute in a situation [19]. Weights could have the total value of 1 per situation. For example, for the situation 'hypertension', the predicate and weight of context attributes could include respectively: (SBP: >130 mm Hg, 0.4), (DBP: >95 mm Hg, 0.45) and (HR: >90 bpm, 0.15).

Using the CS model, a situation occurs if every sensed context attribute value satisfies the predicate of the region set for the same type of attribute. We consider these situations as 'known' situations. For example the context state of (SBP:142 mm Hg, DBP:97 mm Hg, HR:112 bpm) is contained in the situation space of 'hypertension' and suggests that 'hypertension' is occurring. However, if the current context state does not match any of the pre-defined situations, it indicates the occurrence of an 'unknown' situation. Any unknown situation could be similar/dissimilar to the situations already defined.

3.1.2 Situation Repository

Situation repository contains a set of pre-defined situations that specify the most important regions of context attributes for the application. We have used XML schema for defining our context model and XML documents for defining situations. Applications could express their domain-related situations as an XML document in a simple way without requiring the knowledge of the underlying situation model.

### 3.1.3 Situation Inference Engine

Situation inference is here referred to as discovering which pre-defined situation matches the current context state (values of a collection of context attributes). Situation inference finds a perfect match for known situations and the closest match for unknown situations.

1. Let  $R_{situations} = \{S_1, S_2, ..., S_n\}$  be a set of pre-defined situations in the Situation Repository.

2. Let context state  $C = \{a_1^t, a_2^t, ..., a_m^t\}$  be a set of context attributes' values collected at time t.

3. We use the function perfectMatch  $(C', R_{situations})$  to find a pre-defined situation that matches the current context state C'.

4. If there is not a perfect match, we find the most similar situation  $S_{similar}$  using the State-Space difference  $\Delta_{state}^{space}$  measures [29-30] as follows.

$$\Delta_{state}^{space} = \sum_{i=1}^{n} w_i \cdot f(a_i^s, a_i^r, \hat{a}_i)$$
<sup>(1)</sup>

where  $w_i$  denotes the weight of the context attribute,  $a_i^s$  represents the value of context attribute,  $a_i^r$  defines the closest value in the corresponding region to the  $a_i^s$ , and  $\hat{a}_i$  denotes the accepted region's absolute size. The function *f* computes the distance between the context attribute and the region of accepted values. The state-space difference function calculates similarity/dissimilarity between the context state and a situation based on the weighted distance of context attribute values and their corresponding region elements.

5. After computing the state-space difference for all the pre-defined situations, the situation with the least difference value is selected as the most similar situation. The distance difference between the current context state and the closest situation is used for adjusting data stream parameters. The next subsection discusses adaptation of data stream mining parameters based on the context changes.

### 3. 2 Adaptation manager

Adaptation manager is responsible for gradual tuning of data stream processing parameters according to the occurring situation/s in real time. Light weight algorithms such as LWC and RA-Cluster [9-10, 21] are the data stream mining algorithms that their certain parameters can be adjusted according to memory availability, battery charge or CPU utilization. These parameters control output, input or the process of the algorithm. The LWC (LightWeight Clustering) algorithm is based on the AOG (Algorithm Output Granularity) [7] approach that controls the output of the data stream mining according to the available memory.

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 (the 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.

LWC considers a threshold distance measure for clustering of data. Increasing this threshold discourages forming of new clusters and in turn reduces resource consumption. We have borrowed the concepts of AOG for situation-aware adaptation but rather than

adjusting data stream mining parameters according to resource availability we tune them based on occurring situations and changes of context.

Since we have certain number of pre-defined situations, the values of parameters need to be provided for each situation as the starting point. At run-time, if the current situation is a 'known' situation, the pre-set values are used. However, if the inferred situation is 'unknown', the adaptation manager uses results of the state-space difference measure (i.e. the least distance value) to adjust the threshold value. The adaptation function used to adjust the threshold is as follows.

Adjusted value= set value for the most similar situation – (distance value \* constant (e.g. 10))

Adaptation of data stream parameters based on the distance measure results provides real-time and context-aware adaptation; however, it only considers the closest pre-defined situation disregarding the results of situation inference for other situations. For example, if the results show that the distance of the current context state is 0.04 from situation S1 and 0.03 from S2 and 0.12 from S3, the adaptation process will only use the state-space difference of S2 to adjust the parameter of interest and overlooks other situations. This limitation has been addressed in our extended work that is discussed in the following section. In our extended work, we have integrated fuzzy logic into the CS model and used the fuzzy inference results for adaptation of data stream mining parameters.

### 4. Fuzzy Approach

Fuzziness is defined in [26] as uncertainty and vagueness related to description of the semantic meaning of events and phenomena. Using fuzzy inference mechanisms addresses uncertainty in reasoning about situations as situation inference is not a dual-valued concept that it always returns precisely true or false. For example, fuzzy situation reasoning about the situations of 'normal' and 'hypertension' could also address uncertain and real-world situations such as 'very normal' or 'more or less hypertension'.

Fuzzy Situation Inference (FSI) integrates fuzzy logic principles into the main concepts and reasoning formulas of the CS model [18]. Fuzzy logic has the benefit of representing multi-value logic and can deal with uncertain context [14, 26]. Although the CS model addresses the uncertainty associated with sensors' inaccuracies and characteristics of context, it does not focus on dealing with the concept of vague situations. Next subsection discusses the situation and adaptation managers using a fuzzy approach.

#### 3.1Situation Manager

With a fuzzy approach, the situation manager consists of three subcomponents of fuzzifier, rules and situation inference engine.

## 3.1.1 Fuzzifier

In FSI, crisp inputs are context attribute values such as temperature degree or light level that are obtainable by the application. Fuzzifier, as a software component, uses membership functions to map crisp inputs (i.e. context attribute values) into fuzzy sets. Prior to fuzzification, we need to define linguistic variables and their terms and then select appropriate membership functions for mapping crisp inputs into fuzzy sets.

Linguistic variables [25] are defined for each context attribute that is used in the situation reasoning process (e.g. heart rate and systolic and diastolic blood pressure). Then linguistic variables are divided into terms or fuzzy variables. For example, the terms of the linguistic variable 'heart rate' could be {'slow', 'normal' and 'fast'}. Each term of a linguistic variable is characterized by a fuzzy set. An input *x* is related to a fuzzy set *A* by a membership function  $\mu$  and the relation is denoted as  $\mu_A(x)$ . A membership function maps the input *x* to a membership grade between 0 and 1 [14, 26].

# 3.1.2 Rules

In fuzzy logic, rules represent situations of interest and each rule consists of multiple conditions/antecedents joined with the AND operator but a condition can itself be a disjunction of conditions [2]. An example of a FSI rule for a health-related situation can be expressed as follows:

Rule1: if SBP is 'high' and DBP is 'high' and HR is 'fast' then situation is 'hypertension'

In many cases, there are some fuzzy variables that are more important than others in describing a situation. For example, low blood pressure is a strong indication of 'hypotension' in a person while heart rate may not be equally important. To model the importance of fuzzy variables and conditions, we assign a pre-defined 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.

### 3.1.3 Situation Inference Engine

To interpret a rule we need to evaluate its antecedents and produce a single output that determines the membership degree of the consequent. In FSI, we use the function maximum for the conditions joined with the OR operator; however, we apply the weighted sum function [21] to evaluate the conditions joined with the AND operator as follows.

weighted sum = 
$$\sum_{i=1}^{n} \mu(x_i) w_i$$
 (2)

where  $\mu(x_i)$  denotes the membership degree of the element  $x_i$  and  $w_i$  represents a weight assigned to a condition (discussed in subsection 3.1.2). The result of  $\mu(x_i)w_i$  represents a weighted membership degree of  $x_i$  for the condition. Table 1 shows the use of weighted sum in the evaluation of rule 1 (defined in subsection 3.1.2) for input values of SBP, DBP and HR at time *t*.

The result of rule evaluation is applied to the consequent and the output of the weighted sum function determines a situation's membership degree  $\mu_{s_i}(x)$  that suggests the level of confidence in the occurrence of the situation. The level of confidence can be compared to a confidence threshold  $\varepsilon$  between 0 and 1 (i.e. predefined by the application's designers) to determine whether a situation is occurring [20-21] as follows.

$$\mu_{s_i}(x) \ge \mathcal{E} \tag{3}$$

Results of situation inference are fuzzy situations are used as an input to the adaptation manager for adjustments of data stream processing parameters. The following section describes the adaptation manager.

## 3.2 Adaptation manager

The adaptation manager is responsible for adjusting data stream processing parameters according to the occurring situation/s. Situation-aware adaptation is performed based on the results of situation inference from the situation manager. These results are multiple situations with different level of certainty.

To provide a fine-grained adaptation and reflecting the level of confidence of each situation in the adaptation process, we use a weighted average function for adjusting the parameter value. This function is similar to the weighted average function used in the Sugeno's fuzzy inference model [14, 26]. The weighted average function enables reflecting all the results of situation inference in the adaptation of parameter values. The weighted average function format is as follows.

$$\hat{p}_{j} = \sum_{i=1}^{n} \mu_{i} p_{j} / \sum_{i=1}^{n} \mu_{i}$$
(5)

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

For data stream mining, we have used AOG-based algorithms such as LWC that considers a threshold distance measure for clustering of data. The threshold parameter controls the output of the data mining. Situation-aware adaptation gradually tunes the threshold value based on the results of weighted average function that factors in the level of confidence of all situations and addresses the limitation of our previous work. Next section discusses the implementation and evaluation of the context-aware adaptive data stream mining architecture.

## 5. Implementation and Evaluation

A prototype based on our proposed architecture is implemented in J2ME and tested on the Nokia N95 mobile phone to represent a real-world scenario in the area of mobile healthcare for monitoring patients suffering from blood pressure fluctuations (as shown in Fig. 2). The pre-defined situations provided as fuzzy rules include 'normal', 'prehypotension', 'hypotension', 'pre-hypertension' and 'hypertension'. Contextual information used as crisp input for situation modeling and reasoning include SBP, DBP and HR that are mapped into fuzzy sets using membership functions.

Data mining algorithm that we have used for evaluation is LightWeight Clustering (LWC) [9-11] (discussed in 3.2 subsection). The threshold distance parameter of the LWC algorithm determines the distance between the center of a cluster and a new incoming data record. In our evaluation, we have set the threshold value for the situation of 'normal' to 42, 'pre-hypotension' to 36, 'hypotension' to 26, 'pre-hypertension' to 18 and 'hypertension' to 10. For critical situations the threshold needs to be decreased and for normal situations the threshold needs to be increased. This is because 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.

The purpose of the evaluation is to demonstrate that 1) the results of situation inference reflect context changes and vagueness of real-life situations; 2) the threshold parameter is gradually tuned in real time according to context changes and occurring situations. To analyze the results of situation reasoning and adaptation for all the situations, we have simulated a 5-day scenario for a patient that experiences fluctuations of blood pressure.

We have generated our test data such that the values of SBP, DBP and HR increase and decrease similar to a real life scenario. The scenario is described in Table 2.

We analyze the results of situation inference by mapping to context changes to find out their accuracy in reflecting the changes and vagueness of the situation. These results are presented in Fig. 3. The top graph shows the changes of context attribute values for each day and the bottom graph illustrates the corresponding fuzzy situations. Fig. 3 shows as the values of SBP, DBP and HR decrease (i.e. day 2), the membership degrees of 'hypotension' and 'pre-hypotension' situations increase and as these values increase (i.e. day 4), the membership degrees of 'hypertension' and pre-hypertension' increase. Furthermore, the results of situation inference for day 2 and 4 also reflect vague situations when the patient is recovering or moving towards a situation (e.g. 'more or less normal' and 'slightly pre-hypertension').

The bottom graph in Fig. 3. also shows how the results of situation inference are used for tuning the values of the distance threshold. The value of the threshold is dynamically adjusted according to the fuzziness (i.e. membership degree) of each situation at run-time. Decreasing the threshold value increases the number and accuracy of the output (clusters) that is required for closer monitoring of more critical situations such as hypertension. Alternatively, increasing the threshold value for normal situations decreases the mining output and also provides cost-efficiency of resources.

### 7. Conclusion and Future Work

Ubiquitous applications such as healthcare monitoring applications need to analyze and process data streams that are generated at very high rates in real-time with mobile devices. Therefore it is of great importance for data stream mining techniques to be equipped with adapting strategies to promote the continuity and consistency of the running application.

In this paper we proposed a general approach for context-aware adaptive data mining that incorporates context-awareness into ubiquitous data stream mining and enables realtime analysis of data onboard mobile devices in a smart and cost-efficient manner. Context-awareness is achieved using Fuzzy Situation Inference (FSI) that integrates fuzzy logic into the CS model [18], a formal context modeling and reasoning approach for supporting pervasive computing environments. The results of situation inference are used for gradual tuning of parameters of data stream mining algorithms. We implemented and tested the architecture with a 5-day scenario to validate the adaptation process.

For future work, we are including a resource-monitor in our architecture and extending the adaptation manager to compute the parameter value according to the occurring situations as well as availability of resources. We also intend to provide strategies of migration of the data and/or process to the nearby devices when availability of resources and occurring situation are both critical.

### References

- C.B. Anagnostopoulos, Y. Ntarladimas and S. Hadjiefthymiades, Situational Computing: An Innovative Architecture with Imprecise Reasoning, the Journal of Systems and Software 80 (2007), 1993-2014.
- [2] G. Bruce, B.G. Buchanan and E.D. Shortliffe, Rule-based expert systems : the MYCIN experiments of the Stanford Heuristic Programming Project. Reading, Mass: Addison-Wesley, 1984.
- [3] H. Byun and C. Keith, Supporting Proactive 'Intelligent' Behaviour: the Problem of Uncertainty, Proceedings of the UM03 Workshop on User Modeling for Ubiquitous Computing, Johnstown, PA, 2003, 17–25.
- [4] J. Cao, N. Xing, A. Chan, Y. Feng and B. Jin, Service Adaptation Using Fuzzy Theory in Context-aware Mobile Computing Middleware, In Proceedings of the 11<sup>th</sup> IEEE Conference on Embedded and Real-time Computing Systems and Applications (RTCSA'05), 2005.
- [5] Y. Chen, H. V. Leong, M. Xu, J. Cao, K. Chan, A. Chan, In-network Data Processing for Wireless Sensor Networks, Proceedings of the 7th International Conference on Mobile Data Management (MDM'06), 2006.

- [6] R. Cheung, An Adaptive Middleware Infrastructure Incorporating Fuzzy Logic for Mobile computing, In Proceedings of the International Conference on Next Generation Web Services Practices (NWeSP'05), 2005.
- [7] 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.
- [8] 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.
- M. M. Gaber, S. Krishnaswamy, and A. Zaslavsky, Resource-Aware Mining of Data Streams. Journal of Universal Computer Science. 11(8), 2005, 1440—1453.
- [10] 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.
- [11] 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., 2005.
- [12] M. Galan, H., Liu, K., Torkkola, Intelligent Instance Selection of Data Streams for Smart Sensor Applications. SPIE Defense and Security Symposium, Intelligent Computing: Theory and Applications III., 2005, 108-119.
- [13] A. Hossain, An intelligent sensor network system coupled with statistical process model for predicting machinery health and failure, Sensors for Industry Conference, 2002.
- [14] J.R. Jang, Ch. Sun and E. Mizutani, Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Prentice-Hall: Upper Saddle River, NJ, 1997.
- [15] H. Kargupta, B. Park, S. Pittie, L. Liu, D. Kushraj, K. Sarkar, MobiMine: Monitoring the Stock Market from a PDA. SIGKDD Explorations 3(2) (2002), 37-46.
- [16] H. Kargupta, R., Bhargava, K., Liu, M., Powers, P., Blair, S., Bushra, J., Dull, K., Sarkar, M.,Klein, M., Vasa and D., Handy, VEDAS: A Mobile and Distributed Data Stream Mining

System for Real-Time Vehicle Monitoring. In: Proceedings of the SIAM International Data Mining Conference (SDM 2004).

- [17] J. Mäntyjärvi and T. Seppanen, Adapting Applications in Mobile Terminals Using Fuzzy Context Information, In the Proceedings of 4th International Symposium on Mobile HCI2002, Italy, 2002, 95-107.
- [18] 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.
- [19] A. Padovitz, S.W., Loke, A., Zaslavsky, B., Burg, and C., Bartolini, An Approach to Data Fusion for Context-Awareness. Fifth International Conference on Modeling and Using Context, CONTEXT'05, 2005.
- [20] 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.
- [21] 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.
- [22] A. Ranganathan, J. Al-Muhtadi and R.H. Campbell, Reasoning about Uncertain Contexts in Pervasive Computing Environments. In IEEE Pervasive Computing, 3(2), 2004, 62-70.
- [23] R. Shah, S. Krishnaswamy and M. M. Gaber, Resource-Aware Very Fast K-Means for Ubiquitous Data Stream Mining, Proceedings of Second International Workshop on Knowledge Discovery in Data Streams, ECML 2005 and PKDD, 2005.
- [24] P. VajirkarST, S. SinghST, 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, 381 – 391.
- [25] L. Zadeh. The Concept of a Linguistic Variable and Its Application to Approximate Reasoning. Information Systems, 1975, 199-249.
- [26] H. Zimmermann, Fuzzy Set Theory and Its Applications. Kluwer Academic Publishers, Norwell, Massachusetts, 1996.

# Tables

	Ŭ	5
Antecedent	Input Value	Weighted Membership Degree
1: SBP is 'high'	129 mm Hg	0.9*0.35=0.315
2: DBP is 'high'	93 mm Hg	0.8*0.4=0.32
3: HR is 'fast'	102 bpm	1*0.25=0.25
	WeightedSum = 0.315 + 0.32 +	-0.25=0.885

Table 1: Evaluation of rule 1 using the weighted sum function

Day	Patient's health state
1	Vital signs are normal and RT is mild
2	Morning: RT is gradually rising and SBP, DBP and HR are decreasing
	Afternoon: SBP, DBP and HR are very low (necessary medical treatment provided)
	Evening: SBP, DBP and HR are gradually increasing and RT is dropping
3	Vital signs are normal and RT is mild
4	Morning: SBP, DBP and HR are increasing
	Afternoon: SBP, DBP and HR are very high (necessary medical treatment provided)
	Evening: SBP, DBP and HR are gradually decreasing
5	Vital signs are normal and RT is mild

# Table 2: Description of the evaluation scenario

# **Figure captions**

Fig. 1 A general architecture for context-aware adaptive data mining

Fig. 2 The prototype of the context-aware adaptive data mining

Fig. 3 Evaluation of situation and adaptation manager based on a 5-day Scenario

# Figures



Fig. 1 A general architecture for context-aware adaptive data stream mining



Fig. 2 The prototype of the context-aware adaptive data mining

