

A Software Model for Mobile base Data Mining System (MDMS)

Ashish Kumar¹, Dr. Kavita Choudhary²

ABSTRACT

Data mining is a computer-assisted process of digging and sifting through data warehouse to obtain meaning of the data. A specialized tool for data mining is required to extract trends to predict the future values to take knowledge base decisions. Now a day, every organization is exploring and using the latest technology. People are using the internet technology on hand-held devices, like mobile or tablet. Just to sit in front of a single machine (desktop) in the office or using laptop anywhere but tedious to use and heavy to carry; and use internet to access the resources, tools for data mining purpose supports the fix environment. In era of smart technology everyone wants to access any information from anywhere, anytime. Hand-Held devices are available in the market having the supporting platform on which we can think about an application for the data mining purpose. But no system or model is available that is platform independent or on which we can implement data mining application for mobile devices for any domain respectively. This paper is the proposal of a model that is used to develop a platform independent application for hand-held devices for the data mining related to any domain called as Mobile Data Mining System (MDMS).

KEYWORDS

data mining; data mining on hand held devices; platform independent model for data mining

INTRODUCTION

Mobile software engineering is the future solution to accessing and fetching the information on handheld devices for enterprise level solutions or using these as intermediate solution to access from data warehouses for end customer level decision making. After submission of data at the last entry point manually or automatically, the decision makers have to use that for analysis purpose. In varies domain the intermediate layers between the data leaf and the analyzer may depend on the structure of the organization. There may be a number of decision levels in order to top level decision. In this context if we have to extract data; in fact we should have a software tool that is used for data mining as per the need of that level. The solutions available in the market are totally desktop based and need heavy platform to run. Such software's are very smart and efficient to generate the future trends and give data for prediction that is helpful in decision making. But there are few cons with these kind of software;

like location dependency, one should sit at a specific and fix location to work on these software for data mining, it may be portable but not necessary that it provides mobility, if it is being installed on laptop or high configuration notebook then it is a tedious job to carry a laptop all-time and everywhere and necessary that every software is developed with the platform independent supporting technology or language.

We can see some real scenario where such desktop based software have failed. Let's take an example of area sales manager, he is in the market (at any sales point of his territory), and wants to analyze the footfall of the customer on the given date and time to define the strategy; now in this situation he unable to get this data easily, may be he is having the laptop, in that condition he has to carry that all-time and it will take enough time to connect it with the internet and then used to installed software which will retrieve data information from the main server. He will be fed up and exhausted with the high frequency of this set of steps. Now in the next example, if on the regular check-up visit doctor want to see any patient a process graph of his condition since the time he admitted in the hospital then same will be done by doctor to sit in front of a desktop and get the details, but not a routine check-up visit that is not possible for the doctor to sit in the office or to installed a desktop with every patient is not an economical feasible solution. One more example of a health club, the gym-trainer gives the suggestions and provides consultancy to improve their health, the exercise routine, and which new exercise has to add and which has to stop on the base of history, here the trainer will face the same problem of location dependency, he has to sit with each person with the desktop and then only he can given consultancy. During the exercise he can't suggest them anything. There are a lot of examples where we are feeling the lack of use of mobility supporting technologies and applications.

The data mining solution on the mobile devices (hand held devices like tablet, smart phones, PDAs) is the only solution for the above situations. But still there is a problem that, there is not any standard Platform Independent Model (PIM) that can be used to implement such application for mobile devices in the area of data mining. It will be based on client-server architecture, and flexible to adopt by the developers to develop server side and client side application to give mobility and platform independency. This paper is organized as follows. Section II discuss the few available system belong to the same group. Section III discusses the architecture of proposed model MDMS used to implement software as the

¹ Research Scholar, Department of Computer Science & Engineering, Jagannath University, Jaipur(Rajasthan),India

²Asst. Professor, Department of Computer Science & Engineering, Jagannath University, Jaipur(Rajasthan),India

Email: ¹erashishkumarsaxena@gmail.com

Email: ²kavita.yogen@gmail.com

base architecture with the work flow and work of different layer components. Future work is briefly defined in the Section IV and Section V is the conclusion of the paper.

EXISTING MOBILE BASED DATA MINING SYSTEM

There are several number of mobile based data mining tools are available in the market belong to different domains, e.g., health club, health care, departmental store, automobile etc, few name are MobiMine[2], VEDAS[3], Genie of the Net[4], PDM Framework[5].

Kargupta et al [6], [7], [8] have developed the first ubiquitous data stream mining system termed MobiMine[2]. It is distributed data mining application for financial data streams based on PDA-based client-server architecture [1]. It is a mobile based monitoring application for the stock market. This application provides the facility to store related data of selected stock portfolio on the mobile device, and facilitate to user to in determining trends and behavior of stock market. It's a client-server architecture support data mining application. The core functionalities of this application at client side are Portfolio Management and Stock Tickers, maintain WatchList and providing a Reporting Module too. On the server side it provides the functionalities for data collection from different related web resources, applying data mining techniques online, work on most active stocks, etc.

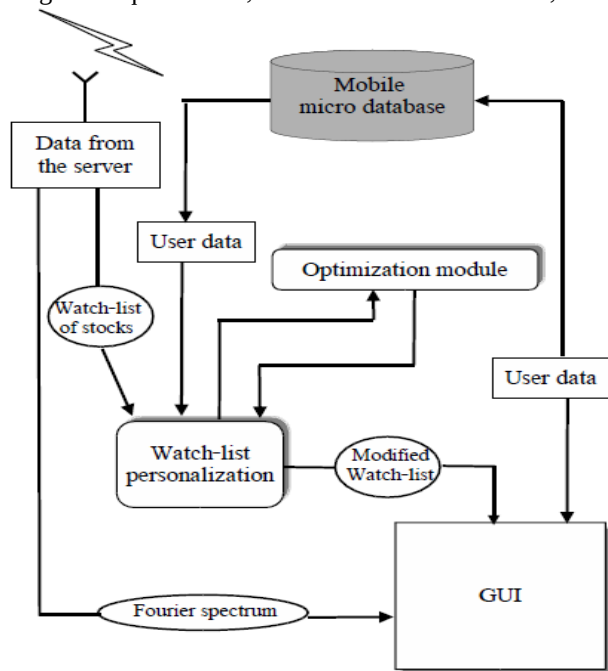


Figure 1. Data flow in the client system [2, Figure 2]

Kargupta et al [3] have developed Vehicle Data Stream Mining System (VEDAS)[5]. It is also a ubiquitous data stream mining system, using it user can continuously monitor the pattern of data streams generated on-board a moving vehicle^[1]. The component that works to perform mining on data is located on the PDA. A commercial version of VEDAS termed as MineFleet has been successfully deployed [9], [10]

[3]. VEDAS uses to real time vehicle-health and driver characterization monitoring^[1]. The Figure 2 shows the conceptual diagram of VEDAS.

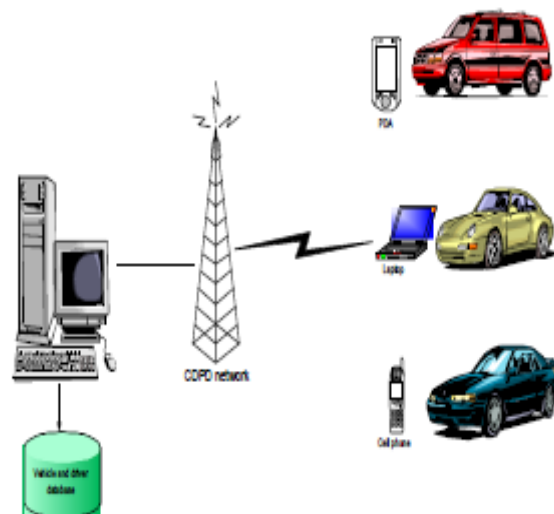


Figure 2. Conceptual overview of the VEDAS system. [3, Figure 1]

VEDAS is specially designed for vehicle monitoring through a PDA using wireless networks^[1]. It is helpful in resolving the problems of real time scenario, like a driver is drunk and driving a heavy vehicle like truck or school bus. For all these there is need of an efficient on-board monitoring and mining system. The unique characteristics of VEDAS^[1]:

- Using the PDAs or others hand-held mobile computing devices, it supports data mining and on-board data stream management.
- It supports distributed mining with less inclusion of centralization of data.
- In context of design, there are following some important constraints: (i) Reduce data communication over wireless network, (ii) minimize power consumption, (iii) reducing on-board storage; (iv) reducing time and space complexity.
- Privacy of data is also a major constraints

Pirttikangas et al [4] have implemented a mobile agent based ubiquitous data mining for a health club for cyclists^[1]. The system is called Genie of the Net[5]. The core purpose of this application development is to give the consultancy and guidance to the cyclists in a health club. Cyclist information is stored throughout the exercise such as heart rate. Later this information will be used for the analysis purpose by performing the data mining techniques, and useful to advice the user after the exercise. This system provides the services related to exercises, planning and scheduling, instruction creation for exercising, and also instructing to cyclist during

exercise^[1]. This application is developed on the base of Genie architecture shown in Figure 3. The user can use the Genie for the different task before the exercise and after the exercise : exercise schedule, instruction summary, exercise calendar and right or wrong during exercise, suggestions for improvement respectively.

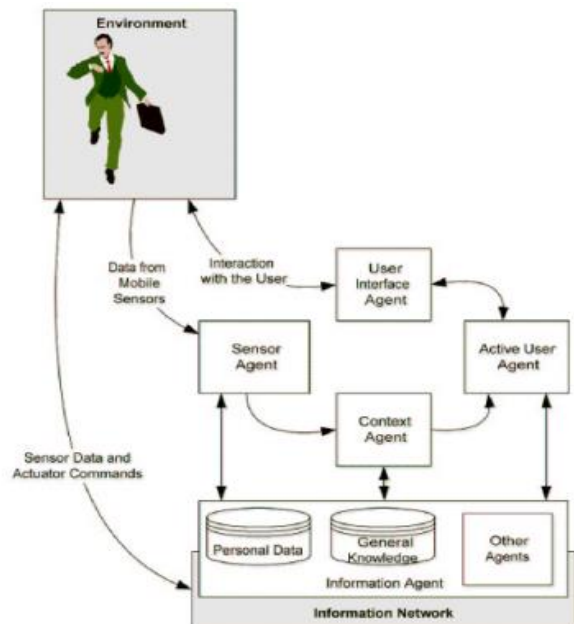


Figure 3. Genie of the Net architecture [4, Figure 1]

ARCHITECTURE OF MDMS

The implemented application discussed in the Section II, are not supporting the features like mobility, platform independency. The architecture of the proposed model is shown in Figure 4. The model is proposing the software development with platform independency.

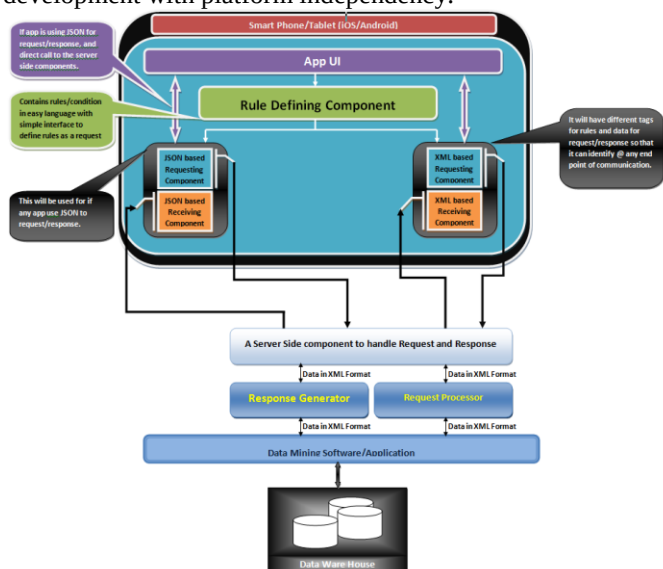


Figure 4. Architecture of Mobile base Data Mining System

The developer has to check the availability and compatibility of technology with each other those are in use during implementation. The architecture is divided in four layers: Client Side, Server Side, Data Mining Software and Data Ware House.

A. Client Side Application and Components

Client can have the any type of hardware and operating system e.g. iOS, android , Java, Microsoft Windows 8 etc. As per the model client side will have one app through which the user will interact with the server side data mining software. This app is collection of the following components:

i. Application User Interface (App UI)

ii. Rule Defining Component

iii. Requesting Component

o JSON Based

o XML Based

iv. Receiving Component

o JSON Based

o XML Based

The App UI will provide the graphical interface to the user. Through this interface user can generate the request to extract data, select the defined rules. It is used to select defined rule and defining new rules. These rules will be applied on requested data set by the user.

To overcome from the platform dependency problem here the use of JSON and XML is suggesting. The two components one for generate the request and one for receive the response to from the server respectively, both using the JSON or XML. Because selection of any of these technologies is totally depend on the other selected technologies at the server-end. Now a day all latest technologies support the JSON and XML for transformation and exchange to data in a very secure manner on the public channel. Although XML is the better solution, the number of APIs are available almost the in all technologies, but gradually JSON is also in used, it is becoming the substitution of XML in the area of data transformation.

The flow of the request is given in the Figure 4, the user will select the data through the App UI, then will select the rules or define the new rules. Then this data will be passed to the requesting component which ever will be selected as per need. This component will perform the transformation and pass the data or generate the request to server side component. Later, on when the response is received by selected receiving component then it will again transform into the app implemented language. The most important rules are also added with the requested data attributes into the JSON format or XML format.

B. Server Side Components

Server side technologies are working on enterprise level. So as such till now there is no research has been done till now that any up-gradation required on the server side technology. The server side application is collection of the following components:

- i. Server side component to handle Request and Response
- ii. Request Generator
- iii. Response Generator
- iv. Data Mining Software
- v. Data Ware House

The client side requesting component generated request will not be directly reach to the data ware house. First it will be handled by the server side component that is responsible for the request and response too. After receiving the request this component will forward this request to the Request Processor into XML format, and then Request Processor forward the request to the Data Mining Software in XML format because almost all server side technologies and data mining software support XML.

This request contains the rules that have to use for mining and the data attribute. The data mining software interact with the data ware house. After performing the data mining process on the requested information the software will send the response to the Response Generator in XML format. The Response Generator forwards this to server side Request/Response handler. After all these processes the server returns to the client app with response. Now the client will perform the operations on the received data as per the predefined and configured guidelines.

CONCLUSION

The model is not hundred percent perfect but a core and a pioneer in the area of mobile application development for data mining. Such models can be globally accepted and may become standard for this area of application development.

FUTURE WORK

The model proposed in this paper is a fundamental and core frame of architecture and a skeleton on which new components can be added. The implementation is not yet started, so as the implementation will get place then the need of other layers or other component may be required to add at client or server end. The applications are developed or developing not considering platform independency as major issue. The components will created for varies platform available in the market.

REFERENCES

- [1] A. Kumar, K. Choudhary., A survey: Data Mining System for Mobile Devices, International Journal of Enhanced Research in Science Technology &

- Engineering, ISSN: 2319-7463 Vol. 3 Issue 2, February-2014, pp: (293-300)
- [2] H. Kargupta, B. Park, S. Pittie, L. Liu, D. Kushraj, and K. Sarkar., MobiMine: Monitoring the Stock Market from a PDA. ACM SIGKDD Explorations Volume 3, Issue 2. January 2002, Pp. 37–46. ACM Press.
- [3] 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. Proceedings of the Fourth SIAM International Conference on Data Mining, Lake Buena Vista, Florida, USA, April 22-24, 2004, pp 300-311,
- [4] S. Pirttikangas, J. Riekk, J. Kaartinen, J. Miettinen, S. Nissila, and J. Roning. Genie of The Net: A New Approach For A Context-Aware Health Club. In Proceedings of Joint 12th ECML'01 and 5th European Conference on PKDD'01. September 3-7, 2001, Freiburg, Germany
- [5] Frederic Stahl, Mohamed Medhat Gaber and Max Bramer, "Pocket Data Mining: Towards Collaborative Data Mining in Mobile Computing Environments", Tools with Artificial Intelligence (ICTAI), 2010 22nd IEEE International Conference on (Volume:2), pp. 323 - 330, ISSN 1082-3409, Print ISBN: 978-1-4244-8817-9
- [6] H. Kargupta, B. Park, S. Pittie, L. Liu, D. Kushraj, and K. Sarkar (2002). MobiMine: Monitoring the Stock Market from a PDA. ACM SIGKDD Explorations. January 2002. Volume 3, Issue 2. Pp. 37–46. ACM Press.
- [7] H. Kargupta, K. Sivakumar, and S. Ghosh. (2002). Dependency Detection in MobiMine and Random Matrices. Proceedings of the 6th European Conference on Principles and Practice of Knowledge Discovery in Databases, Pp. 250–262. Helsinki, Finland .
- [8] S. Pittie, H. Kargupta, and B. Park. (2003). Dependency Detection in MobiMine: A Systems Perspective. Information Sciences Journal. Volume 155, Issues 3-4, pp. 227-243, Elsevier
- [9] H. Kargupta, V. Puttagunta, M. Klein, K. Sarkar, On-board Vehicle Data Stream Monitoring using MineFleet and Fast Resource Constrained Monitoring of Correlation Matrices. Next Generation Computing. Invited submission for special issue on learning from data streams, volume 25, no. 1, pp. 5–32, 2007
- [10] Agnik, MineFleet Description, <http://www.agnik.com/minefleet.html>