Exploiting need of Heterogeneous Data Mining Environment based on Mobile Computing Environments

Smriti Pandey
Dept. of Information Technology
T.I.T,Bhopal,India

Prof. Illyas khan
Dept. of Information Technology
T.I.T,Bhopal,India

Prof. Vijay K. Chaudhari
Dept. of Information Technology
T.I.T,Bhopal,India

ABSTRACT

Today the concept of Data Mining services is not alone sufficient. Data mining services play an important role in the field of Communication industry. Data mining is also called knowledge discovery in several database including mobile databases and for heterogeneous environment. In this paper, we discuss and analyze the consumptive behavior based on data mining technology. We discuss and analyze different aspects of data mining techniques and their behavior in mobile devices. We also analyze the better method or rule of data mining services which is more suitable for mobile devices. In this paper, we survey several aspects of open service framework based on grid structure which provides the heterogeneous environment for data mining on mobile computing environments.

KEYWORDS

Data Mining, Heterogeneous, CLDC, MIDP, Grid

I. INTRODUCTION

The rapidly expanding demand for digital mobile communication services, along with the wide popularity of mobile devices, has given rise to the development of new dimensions and requirements for future mobile communication systems. The innovations in computer science have made it possible to acquire and store enormous amounts of data digitally in databases, currently giga or terabytes in a single database and even more in the future. Many fields and systems of human activity have become increasingly dependent on collected, stored, and processed information. But today is also the requirement of heterogeneous environment where we accept different data sets according to the user and apply data mining on them. However, the abundance of the collected data makes it laborious to find essential information in it for a specific purpose. In the late 1980's, the disciplines of knowledge discovery and data mining emerged to help survey the information content of data. It is also use in mobile devices with the use of MIDLET and CLDC component of J2ME. In few years back, mobile extensions to Grid systems have been increasingly proposed in order to support ubiquitous access and selection to the Grid and to include mobile devices as additional Grid resources [1, 2]. In today's scenario mobile devices, such as mobile phones, PDAs, notebook and others, provide a basic building block [3][4][5][6].

Finding prevalent mobile user patterns and behavior in a heterogeneous environment has been one of the major problems in the area of mobile data mining. Particularly, the algorithms of discovering frequent user's behavior patterns in the mobile agent system have been studied extensively in recent years. The key feature in most of these algorithms is that they use a dataset and frequent Item-Sets visited by the customers. In this case, some problems occur because they do not consider that mobile user's behavior patterns are dynamically variable as time passes. In this paper we discuss some of the data mining service which are use in different areas and then apply those services to mobile devices and then apply those DMS services in mobile computing and exploiting the need of DMS in mobile computing environments using CLDC and MIDP components. The Connected, Limited Device Configuration (CLDC) and the Mobile Information Device Profile (MIDP) have emerged as J2ME standards for mobile phone applications development which are used with DMS services. The role of CLDC and MIDP component is to apply Data Mining Services in mobile. Data Mining Services are useful in several sectors including Mobile, WWW, HealthCare scenario etc. We discuss several aspects step-by-step in this paper and analyze those aspects and approaches sequentially. Discuss their advantages and disadvantages and conclude with new concept.

The remaining of this paper is organized as follows. We discuss CLDC and J2Me in Section 2. In Section 3 we discuss about MIDP.The Grid Structure and Evolution in section 4. In section 5 we discuss about the Challenges. The conclusions and future directions are given in Section 6. Finally references are given.

II. CLDC and J2ME

The J2ME architecture is described in general before the components in the J2ME technology are introduced.J2ME applications are also discussed in general, and it is explained how they are made available to end users.J2ME is a highly optimized Java runtime environment. J2ME is aimed at the consumer and embedded devices market. This includes devices such as cellular telephones, Personal Digital Assistants (PDAs) and other small devices. Fig 1 shows the J2ME architecture. Java 2 Standard Edition (J2SE) developers
should be familiar with Java Virtual Machines (JVMs) and at least one host operating System (OS).
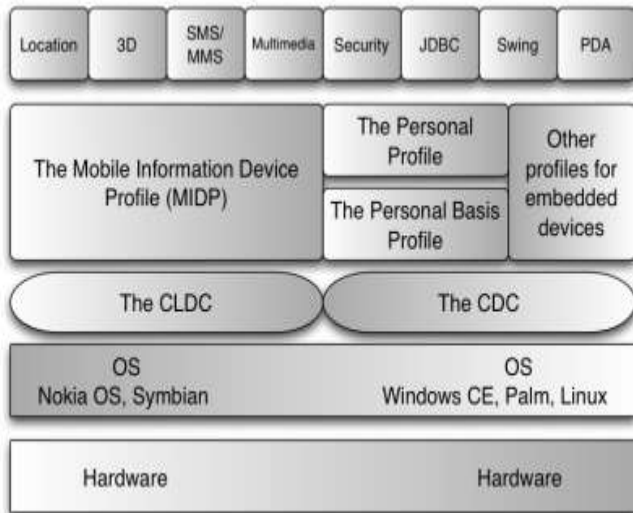
Fig1 J2ME Architecture

The fundamental branches of the J2ME plstform are configurations. A configuration is a specification that describes a Java Virtual Machine and some set of APIs that are targeted at a specific class of device.

The Connected, Limited Device Configuration is one such specification. The CLDC specifies the APIs for devices with less than 512 KB of RAM available for the Java system and an intermittent (limited) network connection. It specifies a stripped-down Java virtual machine, called the KVM, as well as several APIs for fundamental application services. Three packages are minimalist versions of the J2SE java.lang, java.io, and java.util packages. A fourth package, javax.microedition.io, implements the Generic Connection Framework, a generalized API for making network connections.

Many J2ME games already exist and enjoy great popularity especially among young generation. Java comes with the immense requirement of the object-oriented programming language for developers to implement new mobile applications [7]. Configurations provide core functionality and a way to provide greater flexibility but no services for managing the application life-cycle, for driving the user interface, for maintaining and updating persistent data on the device or for secure access to information stored on a network server [8]. Fig 2 shows the CLDC position in J2ME Architecture.
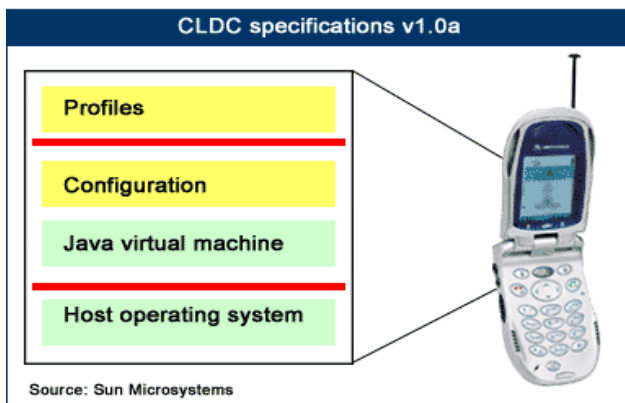

Fig 2 CLDC in J2ME Architecture

Several networks have conducted a survey on users' watching behavior [9] which reflects that user behavior pattern recognition is not so easy task , we can achieve this by CLDC and MIDP component. Instead of replacing existing TV service, mobile services should be complementary [10], and offer more interactive means for users to watch their chosen content.

III. MIDP

The Mobile Information Device Profile is a specification for a J2ME profile. It is layered on top of CLDC and adds APIs for application life cycle, user interface, networking, and persistent storage [11]. An application written for MIDP is called a MIDlet. MIDlet applications are subclasses of the javax.microedition.midlet.MIDlet class that is defined by MIDP.

MIDlets are packaged and distributed as MIDlet suites. A MIDlet suite can contain one or more MIDlets. The MIDlet suite consists of two files:

- Java Application Descriptor (.jad) file

  The Java Application Descriptor file lists the archive file name, the names and class names for each MIDlet in the suite, and other information. This file is used by the mobile device to ensure that device has the minimum requirements to run the application.

- A Java Archive file (.jar) file.

  The archive file contains the MIDlet classes and resource files

Java Studio Mobility is an integrated development environment (IDE), based on the Net Beans development platform, that enables you to use J2ME technologies and add special tools that enable you to code and test J2ME applications, such as emulators and obfuscators.

A MIDlet is a J2ME application designed to operate on an small computing device. A MIDlet is defined with at least a single class that is derived from the javax.micoedition.midlet.MIDlet abstract class.

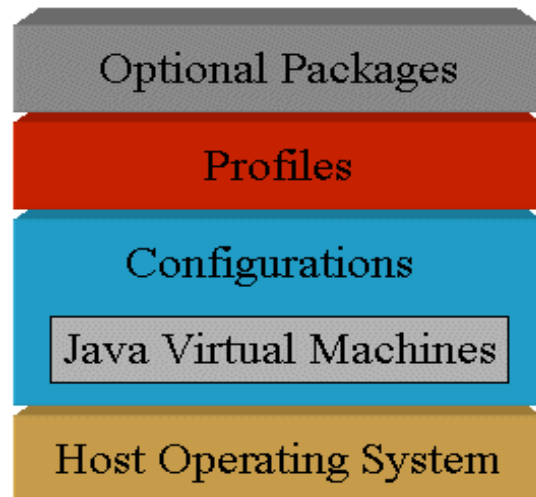The Position of MIDP is shown in Fig 3.

Fig 3 MIDP in J2ME Architecture

IV. Grid Structure and Evolution

Grid computing is a term referring to the combination of computer resources from multiple administrative domains to reach a common goal. The Grid can be thought of as a distributed system with non-interactive workloads that involve a large number of files. What distinguishes grid computing from conventional high performance computing systems such as cluster computing is that grids tend to be more loosely coupled, heterogeneous, and geographically dispersed. Although a grid can be dedicated to a specialized application, it is more common that a single grid will be used for a variety of different purposes. Grids are often constructed with the aid of general-purpose grid software libraries known as middleware.

Grid size can vary by a considerable amount. Grids are a form of distributed computing whereby a "super virtual computer" is composed of many networked loosely coupled computers acting together to perform very large tasks. Furthermore, "Distributed" or "grid" computing in general is a special type of parallel computing that relies on complete computers (with onboard CPUs, storage, power supplies, network interfaces, etc.) connected to a network (private, public or the Internet) by a conventional network interface, such as Ethernet. This is in contrast to the traditional notion of a supercomputer, which has many processors connected by a local high-speed computer bus.Fig4 shows the mechanism of grid computing.
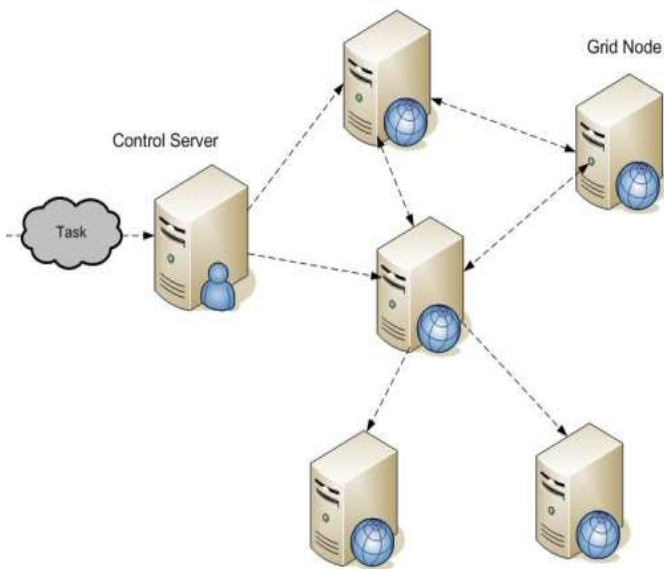

Fig 4 Grid Computing

In 2001, I. Foster, C. Kesselman and S. Tuecke [12] proposed Grid computing which is the combination of computer resources from multiple administrative domains applied to a common task, usually to a scientific, technical or business problem that requires a great number of computer processing cycles or the need to process large amounts of data.One of the main strategies of grid computing is using software to divide and apportion pieces of a program among several computers, sometimes up to many thousands. Grid computing is distributed, large-scale cluster computing, as well as a form of network-distributed parallel processing

In 2004, Ian Foster and Carl Kesselman [13] proposed about the characteristics of spatial information, including large scale, distribution of data resource and user, and high computation which demand a new, parallel and distributed knowledge discovery and sharing platform. Grid is a novel network platform which succeeded

by World Wide Web, it aims to provide an infrastructure for various resources sharing to achieve all the resources, including storage resource, computing resource, information resource and correspondence resource etc. connection and sharing. The development of grid technology provides a credibility network environment for resources unitization and sharing.

In 2004, H.Zhuge et.al [14] Proposed that Knowledge Grid is a sustainable human machine interconnection environment that enables people or agents to effectively generate, capture, publish, share, manage and promote knowledge, to process any type of resource through machines, and to transform resources from one form to another.

In 2007, A. Congiusta, D. Talia, and P. Trunfio [15] Proposed that the Knowledge Grid services are organized in two hierarchic layers which are (1) Core K-grid layer - offers services directly implemented on the top of generic grid services.2) the High level K-grid layer - is used to describe, develop and execute distributed knowledge discovery computations

In 2005, D. D. Roure, N. R. Jennings, and N. R. Shadbolt [16] proposed that because of the coarseness of the description and expression of data and ambiguousness of expression of data semantics and internal logic, integration and analysis of data become very difficult. The Semantic Grid is an extension of the current Grid in which information and services are given well-defined meaning, better enabling computers and people to work in cooperation.

In 2008, Rong Zhang, Koji Zettsu and Yutaka Kidawara [17] proposed that SIKA a grid-based overlay network for knowledge grid system. It groups the nodes into domain sensitive communities, but by defining new inter community organization mechanism. It successfully avoids using super peers. The main characteristics of this architecture are highlighted by its convenience for community analysis, promising scalability, its search efficiency, as well as its robustness.

In 2009, Domenico Talia and Paolo Trunifo [18] presented that a grid-based architecture and services for distributed and pervasive knowledge discovery. They proposed about grid frameworks, which can be developed, as a collection of grid services and it is used for distributed data analysis task and knowledge discovery process.

In 2009, Xiao and Li Jianhua [19] proposed that semantic grid is an extension part of the current grid. They flow on the scalability and adaptation.

In 2009 Wu Yueliang et.al [20] proposed a concept for Web services which is used to deal with the interaction between the server end and the customer end and supports the issue of backstage service end program.

In 2009, Huimin Wang, Guihua Nie and Kui Fu [21] proposed that Knowledge acquisition from distributed data resources to support decision-making is receiving an increasing attention. The paper proposes distributed knowledge acquisition architecture and puts forward any new solve algorithms for knowledge acquisition from large-scale distributed and heterogeneous data resources. Semantic web technology is used to explicitly define data semantics. By the semantic mapping between data resources and ontology, the query and acquisition of user-demanded knowledge can be realized by distributed data mining and semantic reasoning.

In 2010, Ashutosh Dubey et al. [22] proposed a novel data mining algorithm named J2ME-based Mobile Progressive Pattern Mine (J2MPP-Mine) for effective mobile computing. It is used for incremental mining in open framework.

V. CHALLENGES

A number of constraints and technical difficulties faced by researchers, which are discussed in this section. These general problems must be considered for further research in this area to propose new technologies for making mobile computing easier. Some of these are:

- The screen size of the mobile is a big limitation. The screen size can affect the approximate visualization of complex results representing the discovered model.

- Mobile navigation facility is also a big task to achieve and implement.

- The overhead due to the communication between MIDLET and Data Mining service should not affect the execution time.

- The experiments on system performance depend almost entirely on the computing power of the server on which data mining task is executed.

We are attempting to implement knowledge discovery applications. Techniques and tools can also be implemented in DMS as decentralized and interoperable services that enable the development of complex system such as distributed knowledge discovery suits.

VI. CONCLUSIONS AND FUTURE DIRECTIONS

Along with the rapid development of information technology, executing advanced technologies through mobile handset is the prime direction of development. Implementation of intelligent modules on mobile devices through the combination of J2ME and related computing will be the base to introduce data mining features in Mobile Computing.

In this paper, we discuss about several aspects of grid computing in open framework and also discuss about heterogeneous environment. In future we also work on the limitations that were faced by the researchers.

VII. REFERENCES

[1] M. Migliardi, M. Maheswaran, B. Maniymaran, and P. Mobile Interfaces to Computational, Data, and Service Grid Systems. ACM SIGMOBILE Mobile Computing and Communications Review, 6(4), 2004.

[2] S. Wesner, T. Dimitrakos, and K. Jeffrey. Akogrimo – The Grid goes Mobile. ERCIM, October(59), 2004.

[3] A. Arcelus, M.H. Jones, R. Goubran, and F. Knoefel. Integration of smart home technologies in a health monitoring system for the elderly. Advanced Information Networking and Applications Workshops, 2007, AINAW '07. 21st International Conference on, 2:820–825, May 2007.

[4] Raphael M. Bahati and Michael A. Bauer. Adapting to runtime changes in policies driving autonomic management. In ICAS '08: Proceedings of the Fourth International Conference on Autonomic and Autonomous Systems (ICAS'08), pages 88–93, Washington, DC, USA, 2008. IEEE Computer Society.

[5 Michael Beetz, Jan Bandouch, Alexandra Kirsch, Alexis Maldonado, Armin M¨uller, and Radu Bogdan Rusu. The assistive kitchen a demonstration scenario for cognitive technical systems. (HAM), 2007.

[6] Ralph Bergmann. Ambient intelligence for decision making in fire service organizations. In AmI, pages 73–90, 2007.

[7] Isakow, A. and Shi, H. "Review of J2ME and J2MEbased Mobile Applications", International Journal of Communication and Network Security, Vol. 8 No. 2, pp. 189-198.

[8] Ortiz, 2004a, A Survey of J2ME Today, Sun Developer Network (SDN), viewed 13 August 2007, <http://developers.sun.com/mobility/ getstart/articles/survey/>

[9] http://www.3g.co.uk/PR/Sept2005/1943.htm

[10] http://www.cellular-news.com/story/18707.php

[11] http://docs.sun.com.

[12] I. Foster, C. Kesselman, and S. Tuecke, Organizations International Journal of High Performance Computing Applications, 2001.

[13] Julio Sanchez, Maria P. Canton, "JAVA Programming forEngineers", CRC Press LLC, Printed in the United States of America, 2002.

[14] H.Zhuge. "Semantics, Resource and Grid". Future Generation Computer Systems, 2004.

[15] A. Congiusta, D. Talia, and P. Trunfio. Distributed data mining services leveraging wsrf. Future Generation Comp. Syst., 2007.

[16] D. D. Roure, N. R. Jennings, and N.l R. Shadbolt, "The Semantic Grid: Past, Present and Future", The Proceedingof IEEE, 2005.

[17] Rong Zhang, Koji Zettsu and Yutaka Kidawara," SIKA: DC, USA, 2008.

[18] Domenico Talia and Paolo Trunifo "Distributed data mining services leveraging WSRF", Future Generation Computer Systems, ACM 2009.

[19] Xiao and Li Jianhua "Semantic Grid Security Solutions Based on Knowledge Base", Proceedings of the Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services, 2009

[20] Wu Yueliang" Mobile Business Application based on J2ME and Web Services", Eighth International Conference on Mobile Business, IEEE 2009.

[21] Huimin Wang, Guihua Nie and Kui Fu "Distributed Data Mining Based on Semantic Web and Grid" in International Conference on Computational Intelligence and Natural Computing, IEEE, 2009

[22] Ashutosh Dubey and Shihir Shandilya "ICT 2010",Springer LNCS-2010.