ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 3, Issue 11 , November 2016

Optimized communication by Static Analysis and Mobile Agent for Distributed Data Mining

Najat Rafalia , Jaafar Abouchabaka

PH Professor LARIT, Faculty of Sciences, IbnTofail University, Kenitra, Morocco
PES Professor LARIT, Faculty of Sciences, IbnTofail University, Kenitra, Morocco

ABSTRACT: We present a new strategy of messages communication in a Multi Agent Distributed Data Mining. This strategy is based on static analysis and mobile agent. The static analysis compiles all messages "send(A,m) to send(A,m,L), where **L** is the necessary portion of message which will be sent to distant site. In the case where the number of messages "send(A,m,L)" is important, a mobile agent migrates to this site in order to accomplish all tasks locally, thus the send(A,m,L) will not be executed and their cost will be minimized. The agents we consider manipulate data who's structure corresponds to a tree. They also integrates a pattern-matching mechanism in order to recognize the messages. The agent behavior is an expression of the l_{AGD} -calculus [1], extended by the primitives to create an agent, to send a message, to change behavior and those for the pattern matching and the construction of the terms.

KEYWORDS: Multi-Agent System, Distributed Data mining, Agent, Mobile Agent , Communication

I. INTRODUCTION

Distributed Data Mining (DDM) is the extraction of knowledge from several databases (Data Mining) regardless of their physical location; it allows the partial analysis of the data extracted from individual distributed sites, and then send the different partial results to other sites to form the final result. The need for such a feature comes from the fact that the data generated locally at each site cannot often be transferred over the network due to the excessive amount of data and privacy issues. Nowadays, DDM becomes increasingly a key element of knowing the systems, because such a decentralized architecture can reach all network-related businesses.

The Multi agent systems (MAS) that deal with complex distributed applications require a distributed resolution of the problems. In many applications, the individual and the collective behavior of agents depends on the observed data originated from distributed sources. In a typical distributed environment, the analysis of distributed data is a real challenge due to several constraints such as limited bandwidth (wireless networks ...), the sensitivity of confidential data, distributed computing resources. As the MAS are simultaneously distributed systems, their combination with the DDM for data-intensive applications is attractive. A number of DDM are available using various techniques like distributed association rules, distributed clustering, classification and compression, but only o few of them use intelligent agents. The combination of the tow systems involves to work in a new system named MADDM (Multi Agents Distributed Data Mining).

After an overview of the integration of multi-agent systems and distributed data mining known as MADDM (Multi Agent-based Distributed Data Mining) [3], we find that communication between sites in DDM is expensive and it is performed only by message. The cost of transferring large blocks of data can be prohibitive and lead to very inefficient implementations. The main objective of a many research in DDM methods is to minimize the number of sent messages.

In This work, we apply a new strategy of message communication between sites in MADDM system. This strategy has been tested in a distributed Multi Agents System where the agents manipulate data having structure like a tree and has a behavior which can filter all arrived messages by patterns having the same structure [2]. This strategy is based on static analysis and mobile agent. The static analysis compiles all messages "send(A,m) to send(A,m,L), where **L** is the necessary portion of message which will be sent to distant site. In the case where the number of messages "send(A,m,L)" is important, a mobile agent migrates to this site in order to accomplish all tasks locally, thus the send(A,m,L) will not be executed and their cost will be minimized.

Other sections of the paper are organized as follows. The 2nd and 3rd section will describe respectively the "Agent" and the "Based on the distributed data mining agents." In section 4 we present based notions on distributed data mining.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 3, Issue 11 , November 2016

In section 5, we clarify the combination of MAS and DDM. The 6th section will propose our strategy of message communication between agents in a MADDM system and finally we perceptual issues research".

II. DATA MINING

A. Definition of Data mining

There are several definitions of DM:
According to P.CABENA and al, The Data Mining is an interdisciplinary field that uses at the same time automatic learning techniques, pattern recognition, statistics of data bases and visualization to identify the ways to extract information from huge data bases [5].
The Data Mining is the analysis of large observational datasets, to discover new relations between them and reformulate these relations to make them more usable by their owners. [6]

B. Data mining tasks

The nature of the tasks exercised by Data mining depends on the use of data mining results. Many intellectual, economic or even commercial issues can be expressed as these tasks which are classified as [7]:

[1] **Exploratory Data Analysis**: It is simply exploring the data without any clear ideas of what we are looking for. These techniques are interactive and visual.

[2] **Descriptive Modelling:** It describes all the data, It includes models for overall probability distribution of the data, partitioning of the dimensional space into groups and models describing the relationships between the variables.

[3] **Predictive Modelling:** This model permits the value of one variable to be predicted from the known values of other variables.

The first three tasks are examples of Data Mining supervised whose goal is to use the available data to create a model describing a particular variable considered as goal in terms of these data.

- **Discovering Patterns and Rules:** It concerns with pattern detection, the aim is spotting fraudulent behavior by detecting regions of the space defining the different types of transactions where the data points significantly different from the rest.
- **Retrieval by Content:** It is finding pattern similar to the pattern of interest in the data set. This task is most commonly used for text and image data sets.

C. Process steps of Data mining

As discussed in [8] [9] [10] [11] [12], we can define data mining process steps as follow:
- **Data collection:** the combination of multiple data sources, often heterogeneous, in a database.
- **Data cleaning** (normalization): the elimination of noise (attributes with invalid or no values).
- **Data Selection:** Select the useful database attributes for a particular data mining task.
- **Data transformation:** the transformation process of attributes structures to be adequate to the information extraction procedure.
- **Extracting information** (Data Mining): the application of some Data Mining algorithms on the generated data by the previous step (Knowledge Discovery in Databases, or KDD).
- **Data Visualization:** Using visualization techniques (histogram, camembert, tree, 3D visualization) for interactive data exploration (discovery data models).
- **Evaluation of models**: identifying strictly interesting models based on data measurements.

D. Data Mining techniques

To perform the tasks of Data Mining there are several techniques from different scientific disciplines (statistics, artificial intelligence, databases) to show hidden correlations in data repositories to build models from these data.

Below the most mentioned data techniques in different documents:

Neural networks, the decision trees, Genetic algorithms, the association rules, Fraud detection, the k-Nearest neighbours, the k-means algorithm (K-Means), clustering algorithm ….

The data mining techniques presented above represent some of the existing techniques for data mining, and the reason why there are so many techniques is that these tasks don't have
the same object, and none of them can be optimal in all cases, they complement each other when they are combined intelligently (obtaining very significant performance gains building the so-called meta-models or models).

E. Categorization of data mining systems

Data mining systems can be categorized according to several criteria. Among the existing categorizations we note:

- **Classification based on the type of data to be explored**: in this classification data mining systems are grouped according to the type of data they handle such as spatial data, time series data, textual data and the World Wide Web,etc.
- **Classification by advanced data models:**
This classification categorizes data mining systems on the basis of advanced data models such as relational databases, the object oriented databases, data warehouses, transactional databases, etc.
- **Classification by type of knowledge to discover:** this classification categorizes data mining systems based on the type of knowledge to be discovered or data mining tasks such as classification, estimation, prediction, etc.
- **Classification by the exploration techniques**
used: this classification categorizes data mining systems following the data analysis using the pattern recognition approach, neural networks, genetic algorithms, statistics, visualization, database-oriented or-oriented data warehouse, etc. [8].

F. Data Mining application fields

The Data Mining technology has a great significance thanks to the possibilities that it offers to optimize the management of human and material resources.
Banking, Bio-informatics and Biotechnology, Fraud detection. The management of scientific data, The insurance sector, Telecommunications, Medicine and Pharmacy, E-commerce and the World Wide Web, The stock market and investment, Analysis of supply chain.

III. DISTRIBUTED DATA MINING

DDM is derived from data mining field that focuses on the distribution of resources on the network. In a distributed scenario where the DDM is applied, the system consists of several independent sites, these sites of data and calculation communicate by sending messages, and this communication is often expensive in terms of power autonomy, privacy and security.

A. Architecture and Features of a DDM system

DDM is planned to perform partial analysis of data at individual sites and then send the partial results to other sites where it may need to be aggregated to the overall result. Thus individual nodes communicate with a rich centralized node resources and neighboring nodes via messages on an asynchronous network to accomplish their tasks. Hence, the majority of DDM methods manage more abstract architecture which includes several Sites with an important storage capacity and independent computing power. The local computation is performed on each site and a central site that communicates with each location distributed to calculate global models using a peer-to-peer architecture.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 3, Issue 11 , November 2016

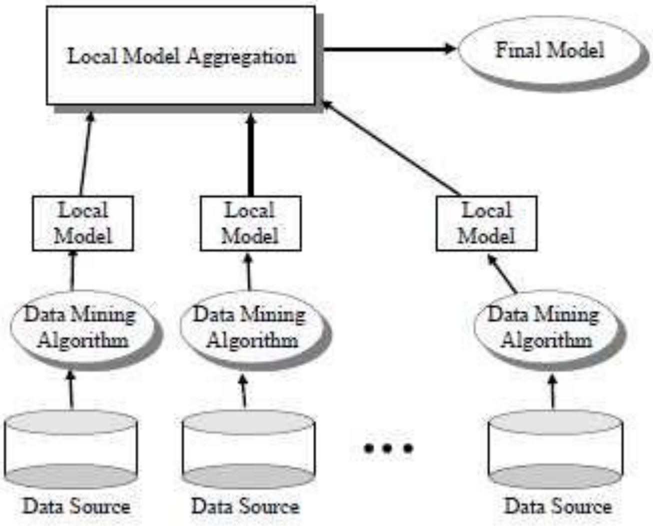

Figure 1: Architecture of DDM System

B. DDM methods

Several systems have been developed for distributed data mining. These systems can be classified according to their strategy in three types, Central Learning, Meta-Learning and Hybrid Learning. More details about these strategies is given in [3].

C. Application fields of DDM

The distributed Data Mining applications include detecting fraud credit card system, detection intrusion system, health insurance, the distributed clustering, segmentation, sensor networks the customer profiling, assessment of retail promotions, analysis of credit risk, etc.

D. Issues and Challenges of DDM

The main criticism of these systems is that it is not always possible to obtain an accurate final result, knowing that the global knowledge model obtained may be different from the one obtained by applying a model approach (if possible) to the same data.

The approximate results are not always a major concern, but it is important to be aware of this. In addition, the use of material resources of these systems is not optimized. In a distributed environment analysing distributed data is a significant problem because of many constraints such as: Communication, Resources and distributed computing nodes, Confidentiality, Scalability.

So for the **communication** which is the object of this paper , we present some limits

- The Limit of the overall performance of the system, since it is supposed to be carried out by message passing
- The communication between sites is expensive and it is performed only by message.
- The cost of transferring large blocks of data can be prohibitive and lead to very inefficient implementations.
- The limited bandwidth (eg wireless networks)
- The main objective of a many research and

DDM methods is to minimize the number of sent messages).

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 11 , November 2016

These and some other peculiarities require the development of new approaches and technologies of data mining to identify patterns in distributed data. Multi-agent technology is a solution to the above challenge. In this paper we will propose in section 6 an hybrid strategy of message communication based on static analysis and optimal mobility of the agents. For these reasons, the characteristics of the agents and of the Multi Agents System are important to clarify.

IV. MULTI-AGENT SYSTEMS:

Before starting the contribution of multi-agents in the field of data mining systems, it would be interesting to understand a few concepts like, distributed artificial intelligence, the evolution of the individual aspect (the behavior of a single agent) towards the collective aspect (behavior of agents in a society).

A. Agent concept :

An agent can be defined as an object (as defined object languages) whose behavior is described by a "script" (principal function main), with its own calculation ways, and can move from place to place (a place that can be a remote computer site from the original site of the agent) to communicate with other agents.
showed is limited when it comes to solving complex problems. In order to address these limitations the researchers felt the need to move from individual behavior to collective behavior and the need to distribute intelligence across multiple entities.
All studies covering collective behavior constitute the field of Distributed Artificial Intelligence (DAI).
Technically, the DAI is a branch of AI, which proposes to replace the centralized software by software based on the interaction of basic software components. It is based on the principle of 'divide and rule' that facilitates the development and test of problem-solving systems and allow better use of components, and it is structured around three axes:
With its "script", the agent is able to follow a life behavior that will be instilled at the time of implementation and that will allow him to have as main feature to be fully autonomous.
One of the discriminate characteristics of the agents is the representation and reasoning on the environment (the external world and other agents), based on this feature, we find two different classes.

- **Cognitive agents**
  A cognitive agent is an agent that has an explicit representation of its purpose and its environment. The actions it performs to achieve its goal are the result from a reasoning on the state of the environment. Usually a cognitive system includes a small number of agents; each is similar to a more or less complex expert system. [2]
  In this case we speak of high granularity agent.
- **Reactive agents**
  A reactive agent is an agent whose behavior responds only to the stimulus/ share law, the stimulus is an element of the environment (action, message, location, etc.). Typically a reactive system has a large number of low granularity agents. These agents do not necessarily have an explicit goal to obtain. By cons, they can implement a complex reasoning on their internal state to perform their actions [19].

B. Mobile Agent

A mobile agent is a running program that can move from host to host in a network which created a new paradigm for data exchange and resource sharing in rapidly growing and continually changing computer network.
It is capable of migrating autonomously and intelligently in various target nodes through network to perform computation in response to changing conditions in the network environment.
The agent's dispatcher purposes are to achieve and fulfil
l user's objective on behalf of user. It is used for information retrieval, searching information, filtering, and intrusion recognition in networks. Mobile agent suspends its execution, Transport itself from one host connected to the network to another, and continue its execution on the new host.[17] [Biblio2].

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 3, Issue 11 , November 2016

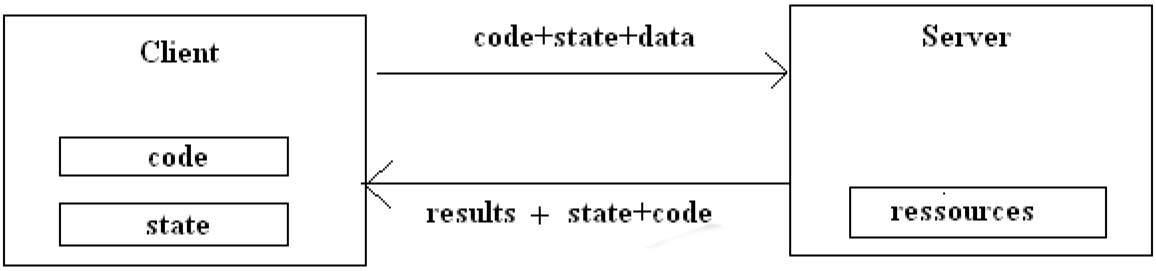

Figure2: Mobile Agent

C. Multi Agents system

1. Definition

A multi-agent system is a system with a set of agents that interact with the communication protocols that are capable of acting on their surroundings. Different agents have different spheres of influence, because they control (or at least can influence) on different parts of the environment. These spheres of influence may overlap in some cases; the fact that they coincide may cause dependencies reports between agents.
The MAS can be used in many application areas such as electronic commerce, economic systems, distributed information systems, organizations. [18] [biblio3]

2. When using a MAS?

When the problem is too complex but can be divided.

- When there is no general solution or when it is too expensive in CPU.
- For modeling purposes (populations, molecular structures, sand piles…)
- When we can parallelize the problem (saving time).
- When we want certain robustness (redundancy).
- When the expert comes from different sources.
- When the data, controls, resources are distributed.
- When the system needs to be adaptive. [22]

V. MULTI AGENT DISTRIBUTED DATA MINING SYSTEM (MADDM)

Generally the construction of a MADDM system involves three essential features: interoperability, dynamic system configuration and performance aspects, discussed as follows.

A. Interoperability

The concerns about interoperability are not only the cooperation of agents in the system, but also the external interaction that allows new agents to enter into the system seamlessly. The system architecture must be open and flexible in order to support interaction, including communication protocol, integration policy, and the service directory.

B. Dynamic System Configuration

The issue is further discussed in relation to the characteristic of interoperability, which tends to manage the dynamic system configuration; it is a matter of challenge because of the complexity of planning and exploration algorithms. A Mining group can have multiple data sources and agents which are configured to equip an algorithm and process the provided data sets. The change in the data affects the Mining task, as an agent that may still be running the algorithm.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 3, Issue 11 , November 2016

C. Performance

Finally, the performance can be improved or altered because the distribution of data is a major constraint. In a distributed environment, tasks can be executed in parallel and in exchange the competition issues arise. The performance of service quality control for DM perspectives and data system is desired, but it may be derived from both data mining and field agents.

D. Architecture of MADDM

Most MADDM Frameworks adapt to similar architectures (see Fig .3.) And provide common structural components [24], [25], [23]. They use KQML or FIPA-ALC, which are standard agent communication languages that facilitate interactions between agents.
Each agent in multi-agent system generally contains interface module, process module and knowledge module. The interface module is responsible for communicating with other agent or with the environment. The knowledge module provides necessary knowledge to be proactive or reactive in various scenarios. The process module does necessary processing to make decisions. Different types of agents in multi- agent based spatiotemporal data mining system are Interface agent, Facilitator agent, Broker agent, Data agent, Pre-processing agent, Data mining agent, Result validation agent. These agents are briefly described in [6].

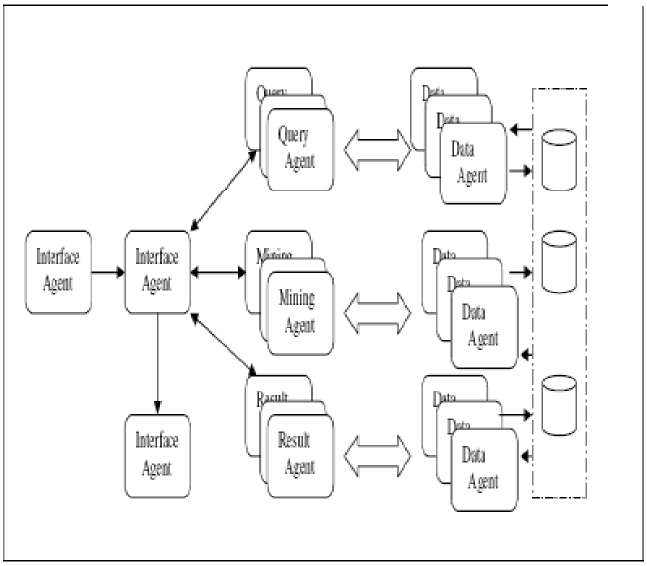

Figure3: Architecture of MADDM

E. MADDM task planning and movement of system operations:

DM task scheduling is performed through negotiations between the facilitator agent and mining agents via message passing mechanism. Suppose a user agent sends a request to the facilitator agent to inform that it would like to do data mining with other agents in the organization. The user agent also needs to give information of model definition (dependent and independent attributes, attribute type (numeric or categorical), model type (linear or nonlinear) with its request. When the facilitator receives the request from the user agent, it negotiates with the broker agent to determine which agents to launch for this task. For example, If the user wants to find all items Y which have statistical significance for a given X, the mining agent must ask only the agents that have information about X. Finally, if the user specifies an X and an Y and asks for the level of support and confidence between the two, the agent must only ask the agents that have information about both X and Y. The mining agent is then responsible for completing the task, while

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 3, Issue 11 , November 2016

the facilitator agent continues to plan future DM requests. When the mining agent completes its work it returns the results and the facilitator agent passes them onto the user agent. [25].

VI. LAZY STRATEGY OF COMMUNICATION

The execution of a transmission send(a,m), consists of the valuation of the receiver agent a and that of the message m. If the valuation of the agent a, detects that this later is in a distant site, then, the transmission of the message address is not sufficient because distant agents could not get to each other site memory. In this case we opt for a lazy transmission between distant sites.
The general problem is presented as follow:
Given a message C(C 1 ,;C 2 ;..., C n) destined to an agent a, what are in this message the necessary levels to accomplish the pattern-matching and the treatment of the message by the agent a.
Our lazy communication strategy consists of two phases:
Static analysis phase: it"s accomplished at the time of the compilation. It allows determining the parts of the message, that are necessaries for the pattern-matching and the treatment of this message. These parts are expressed in the level number of the tree which represents the message.
Dynamic transmission phase: because the static analysis is not always informative, several parts of the message are not detected necessary at the time of the compilation. So, this phase allows completing these needs in the execution.

A. Static Analysis

The static analysis concerns in fact, all the patterns of the application behaviors. The analysis of an initial behavior involves, through the become primitive, to analysis the replacement behaviors starting by this initial behavior. It consists of four principal steps:

I. **Marking:** through each pattern, the marking phase marks the necessary parts in a message which will be filtered by this pattern and treated by the action corresponding to the successful pattern-matching.

An action corresponds to a sequence of several send and several create, ended by a become.

II. **Flattening** : marking is don behavior by behavior. A part can be necessary in a behavior and not necessary in other one. The flattening step allows to "flatten" the results of marking concerning the same pattern which appears in an initial behavior and in other replacement behaviors from this initial behavior. The ended set of replacement behaviors can be determined through the application code. This warrants the termination of our algorithms.
III. **Compilation** of the patterns: the necessary in a pattern is expressed by a number of levels. This phase consists of associating to each pattern the number of its necessary levels.
IV. **Compilation of the send:** the ad-equation between the patterns and the messages m figuring in the send(a,m), and the use of the precedent phase results, allow to compile the transmissions send(a,m) into send(a,m,L), where L is the number of m levels which are necessaries to accomplish the pattern-matching and the treatment of m by a. The detail of the static analysis phases is presented in [1].

The static analysis determines for each class of agents having the same initial behavior bi , the number of necessary levels in a message m which can be filtered and treated by the behavior b i .
We group the different values of NecLev in a table called table of needs (see table 1). In this table, each value L=NecLev(b i i,[m]) corresponds to the number of levels in the message m, which, each agent having the initial behavior bi, needs in order to value (b i m). m is the message filtered by the pattern [m].
If the message m is not filtered neither by bi nor by any replacement behavior obtained from b i then NecLev(b i ,[m])=0.

B. MOBILE IMPLEMENTATION OF OUR COMMUNICATION APPROACH

The agent model provides a unit of encapsulation for a thread of control along with internal state. An agent is either unblocked or blocked. It is unblocked if it is processing a message or has messages in its message box, and it is blocked otherwise. Communication between agents is purely asynchronous: non-blocking and non-First-In-First-Out (non-FIFO).

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 3, Issue 11 , November 2016

However, communication is guaranteed: all messages are eventually and fairly delivered. In response to an incoming message, an agent can use its thread of control to 1) modify its encapsulated internal state, 2) send messages to other agents, 3) create agents, or 4) migrate to another host (picture 2).
The create primitive (create(X 0)) allocates a unique mail address to the newly created agent, and creates a thread which represents the computation potency of this created agent. The system makes an adequacy between the agent and its mail queue. So, the name of an agent and its mail queue address become synonymous.
Each new agent is created in the host having the minimal number of process. It‟s important to allow the programmer to ignore the details concerning the physical location of the agents in different processors which constitute the network of the program execution. To that effect, every site has a monitor agent. This agent manages the distribution of the agents and the communication between sites.
The mobile implementation of communication strategy witch we present in this paper is based on mobile agents. In the distributed agent system, each host has its mobiles agents witch can migrate to the other hosts with the parts of messages, that must be transmitted outside.
Mobil agent is a soft entity that can migrate during the execution from host to other host. It migrates with its code, its execution state and its data. Mobil agent can receive messages in his transit, so the mobility don't handicap the it's communication process. The mobility is controlled by the application not by the execution system. The object of the migration is to reach distant data and resources, to do the treatment locally and to move just necessary data.

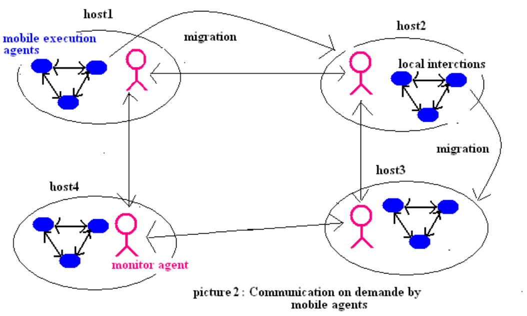

Figure 4 : Communication on demand and by mobile agent

So, mobile agent migrates to the hosts which have a lot of interactions with his host. In the case of the low interactions, this approach become less efficient, it‟s penalized by the size of the mobile agent. In this second case, the messages towards distant sites are addressed to the server agent of the sender site in order to treat and send them to their destination. In the receiver site, the server agent treats the external messages that arrive and transmit them to their receivers.
Using mobile agent involves the optimization of the traffic because exchanged messages become locals and discharge the network. Mobil agents can use the resources of distant hosts, so, it discharges its initial host. The size of the mobile agent adds a surcharge when it‟s migrating but this charge is compensated by local interactions (picture 3).
In each host, the static analysis is done by some local calculator agents. The results of the static analysis are deposits in the table of need as explained above.
Our approach also allows remote code invocation when distant site dispose of needed service.

VII.CONCLUSION AND PERSPECTIVE

Multi-agent systems and data mining are among other domains that are the most active areas in the field of information technology. The synergy between the two technologies offers great potential and opportunities for more sophisticated applications.
We have presented an hybrid strategy of messages communication in a Multi Agent Distributed Data Mining System. This strategy consists of a static analysis which detect the necessary part of message which will be sent. A mobile implementation of this strategy is proposed. We have used a mobile agent which migrates to distant sites having massive interactions with its site. This approach minimizes the traffic in the distributed system. We are testing it on some consistent benchmarks cost concerning the execution time.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 3, Issue 11 , November 2016

REFERENCES

[1] Najat Rafalia and Jaafar Abouchabaka, "A New Approach of Communication Between Agents in a Distributed Environment" International Journal of Emerging Technology and Advanced Engineering. Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 3, March 2013)

[2] Najat Rafalia, Jaafar Abouchabaka, "Multi Agent System and Distributed Applications Optimized Communication using Mobile Agents", International Journal of Advanced Research in Computer Science and Software Engineering. Volume 5, Issue 10, October-2015 ISSN: 2277 128X. Research Paper Available online at: www.ijarcsse.com.

[3]AHMEDAMINE FARIZ, 2 JAAFAR ABOUCHABAKA, 3 NAJAT RAFALIA "USING MULTI-AGENTS SYSTEMS IN DISTRIBUTED DATA MINING: A SURVEY", Journal of Theoretical and Applied Information Technology 31 st March 2015. Vol.73 No.3, ISSN: 1992-8645, www.jatit.org, E-ISSN: 1817-3195

[4] G. Piatetsky-Shapiro, "Data Mining and Knowledge Discovery 1996 to 2005", 15(1), 99-105.

[5] P. Cabena, P. Hadjinian, R. Stadler, J. Verhees and A. Zanasi, "Discovering Data Mining: From Concept to Implementation", Prentice Hall, Upper Saddle River, Nj, 1998.

[6] D. Hand, H. Mannila and P. Smyth, "Principles of Data Mining", Mit Press, Cambridge, Ma, 2001.

[7] S. P. Deshpande1 And V. M. Thakare International Journal of Distributed and Parallel Systems (IJDPS) Vol.1, No.1, September 2010

[8] M. J. Berry, G. S. Linoff, "Data Mining Techniques for Marketing, Sales, and Customer Relationship, Management", Second Edition, 2004.

[9] O. R. Zaïane, "Principles of Knowledge Discovery in Databases", Cmput690, University Of Alberta, 1999.

[10] Ph. Preux, « Fouille De Données : Notes De Cours », Université De Lille 3, 9 Octobre 2008.

[11] G. Dong, J. Pei, 'Sequence Data Mining', Springer Edition, 2007.

[12] S. Prabhu, N. Venkatesan, "Data Mining And Warehousing", New Age International (P) Ltd., Publishers, New Delhi, 2007.

[13] J. Dasilva, C. Giannella, R. Bhargava, H.Kargupta, and M. Klusch. "Distributed Data Mining and Agents. Engineering Applications of Artificial Intelligence", 18(7):791–807, October 2005.

[14] S. Datta, K. Bhaduri, C. Giannella, R. Wolff, and H. Kargupta. "Distributed Data Mining In Peer-To-Peer Networks'. Internet Computing, IEEE, 10(4):18–26, 2006. Intelligent Agent technology (IAT 2003), Pages 211–217, 2003.

[15] W. Davies and P. Edwards. "Distributed Learning: An Agent-Based Approach To Data-Mining". In Proceedings Of Machine Learning 95 Workshops On Agents That Learn From Other Agents, 1995.

[16] U. Fayyad, R. Uthurusamy, and Others. "Data Mining and Knowledge Discovery In Databases". Communications of the ACM, 39(11):24–26, 1996.

[17] Vuda Srinivasa Rao Et. Al. / International
Journal Of Engineering Science And Technology Vol. 2(6), 2010, 1758-1766

[18] K. P. Sycara, Multi agent Systems, American Association for Artificial Intelligence, 1998.

[19] J. Ferber, « Les Systèmes Multi Agents: Vers Une Intelligence Collective »,1995.

[20] S. Sathia kumar Et Al. (Eds.), "Proceedings Of International Conference On Internet Computing And Information Communications, Advances In Intelligent Systems And Computing" 216, Doi: 10.1007/978-81-322-1299-7_39, _ Springer India 2014.

[21] S. Maalal; M. addou (IJACSA) International Journal Of Advanced Computer Science And Applications, Vol. 2, No. 11, 2011

[22] F. Sandakly, « Contribution À La Mise En Œuvre D'une Architecture À Base De Connaissances Pour L'interprétation De Scènes » 2d Et 3d, Thèse Doctorat, Université De Nice-Sophia Antipolis, 1995.

[23] Xining Li and Jingbo Ni. Deploying Mobile Agents in Distributed Data Mining. Lecture Notes in Computer Science, 4819:322–331, 2007.

[24] M. Klusch, S. Lodi, and G. Moro. "Agent Based Distributed Data Mining: The KDEC Scheme". Agent link, Springer Lecture Notes In Computer Science, 2586, 2003.

[25] H.Kargupta, I. Hamza oglu and B. Stafford "Scalable, Distributed Data Mining An Agent Based Application". Proceedings Of Knowledge Discovery And Data Mining,August, 1997.