DISTRIBUTED RURAL HEALTHCARE DECISION SUPPORT SYSTEM USING CO-OPERATIVE MOBILE AND STATIC AGENTS

Kavitha P.T.[1], **Sasipraba T.**[*2]

[1]Research Scholar, Sathyabama University, Chennai
[2]Dean, Sathyabama University, Chennai
Email: [1]kavithapt@yahoo.com

Abstract

This paper focuses on a web based decision support system which is based on distributed data mining and designed for the healthcare industry. Innumerable Decision Support Systems are prevailing in the market to cater the needs of the various industries based on various purposes. The importance of any Decision Support System resides on the quality of information it provides compared to other information system which leads to take effective decision making. As most of the records in the healthcare industry are maintained manually, the use of Electronic Health Records is introduced recently which will enable the user to have a clear idea about the service he/she wants to utilize. Since it takes a lot of time to retrieve the information from the manually maintained records, use of Electronic Health Records plays a major role in the Healthcare industry. The present Decision Support Systems concentrate on the administration part of the Healthcare industry, their features support the physical practitioner, nurses, pharmacists, administrative personnel. The patients who are the main players in the healthcare industry are mainly concentrated for mapping their behaviour, and whether they will come back to the health care service provider again or not. The decision system which is presented here gives much importance for the patients and will be acting as a framework which can be included in the services provided by the hospitals. The feedback of the patients can rank the healthcare service provider and to improve the quality of their service.

Keywords—Decision Support Systems, Distributed Data Mining, Healthcare providers

I. INTRODUCTION

Data mining is doing data analysis or statistics on datasets that have been obtained from potentially many sources, as such the miner may not have control of the input data, but must rely on soruces that have gathered the data. Data mining can be defined in several ways, which differ primarily in their focus on different aspects of data mining. One of the earliest definitions is " the non trivial extraction of implicit, previously unknown and potentially useful information from data". [16]. Data mining has emerged as a means for identifying patterns and trends from a large amount of data. Data mining refers to extracting or mining knowledge from large amounts of data.[18]

DDM is widely used in industrial, scientific and commercial applications to analyze large data sets maintained over geographically distributed sites, which makes DDM a major research issued on today's data mining system. Nowadays, massive data collections in terabyte scale maintained over geographically distributed sites need to be used and analyzed in several scientific and commercial areas. So, mining data from distributed sites, which is called DDM, becoming increasingly essential.[20].

Analysis of past transaction data provide very valuable information for the decision making purpose [3]. The problem of extracting knowledge is a difficult task for large datasets due to their static nature and geographical distribution datasets [1]. Due to these properties, algorithms that handle large datasets cannot assume or control the partitioned structure, the sizes, and the location of the pieces of the datasets and must take account of the latencies and bandwidth required to move data among the places.

Mobile agent is a composition of software and data which can move from one system to another system autonomously in a distributed network. These programs are capable of suspending their execution on one platform and moving to another platform where they resume execution. Mobile Agent's application proves to be one of the best and robust methods to handle the distributed data and hence distributed data mining [4].

In this paper we have focused to increase the performance of existing algorithm by inducing parallel computing with the help of co-operative mobile and static agents. Here the static agent computes with the data in the local host automatically whenever it detects a change. Static agents do not move around the distributed databases but use embedded knowledge to assist in filtering and processing the volume of data. Later, the mobile agent interacts and retrieves required data from the static agents that are residing in the local systems.

In this paper, the patient agent operates as the mobile agent and the central system as the static agent. Patient agent is

The rural healthcare infrastructure in India has been developed as a three tier system with

Sub centre

Primary Health Centre (PHC)

Community Health Centre

being the three pillars of primary healthcare system.

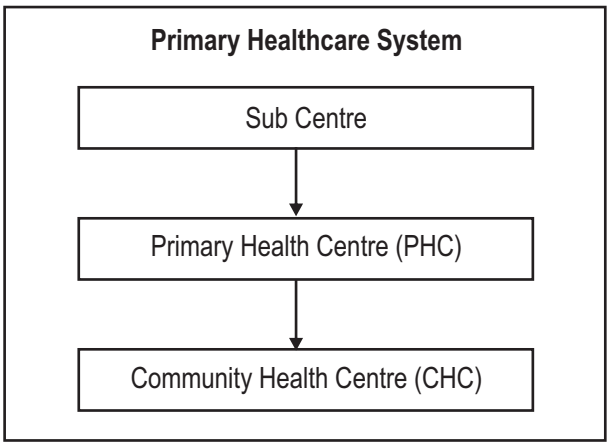

Fig. 1 Primary HealthCare System in India

At the end of the Sixth Plan (1981-1985), there were 84,376 sub centres, which increased to 1,30,1165 at the end of the seventh plan(1985-90), and to 1,45,272 at the end of Tenth plan (2002-2007). As on March 2010, 1,47,069 sub centres are functioning in the country [12].

The sub centres are provided with basic drugs for minor ailments needed for taking care of essential health needs of men, women and children. The Primary Health Centres (PHC) were envisaged to provide an integrated curative and preventive health care to the rural population with emphasis on preventive and promotive aspects of healthcare. Community Health Care Centres (CHC) are being established and maintained by the State Governments. As per minimum norms, a CHC is required to be manned by four medical specialists i.e. Surgeon, Physician, Gynecologist and Pediatrician supported by 21 paramedical ad other staff. An existing facility (district hospital, community health centre etc) can be declared a fully operational First Referral Unit (FRU) only if it is equipped to provide round the clock services for Emergency Obstetric and New Born Care, in addition to all emergencies that any hospital is required to provide [12].

This paper introduces a Decision Supportive System for supporting the patients in the rural area of India to take the right decision regarding the selection of the healthcare providers based on many criteria. And it tries to improve access of rural people, especially poor women and children to equitable, affordable, accountable and effective primary healthcare.

II. MOTIVATION AND RELATED WORK

An Urban HealthCare System includes Community Health Systems(CHS) and Medical Delivery System (MDS). More co-operation between these two systems could greatly improve overall healthcare: making medical service more convenient and cost effective. The Urban HCS is a complex system extensively influenced by human behavior. Using Agent Based Modeling and Simulation (ABMS) as an innovative tool, we build an Artificial HCS as a platform on which to study medical co-operation, such as sharing beds, sharing doctors and cost accommodation. A referral appointment system based on linear programming and present its benefits and capabilities [13].

A previous work that explores the capabilities of mobile agents to build an appropriate framework and an algorithm that better suits the Distributed Data Mining applications. This work aims at providing solutions to the issue of knowledge, consolidation with less communication overhead due to minimum information exchange by overlapped operations thereby improving the efficiency of the Distributed Data Mining. Percentage of saving in processing time increases with

the distance between the central site and distributed business sites.[1].

In current healthcare, information is conveyed from one healthcare professional to another through paper notes or personal communication. A distributed e healthcare system is developed based on Service Oriented Architecture as a basis for designing, implementing, deploying, invoking and managing healthcare services. It supports the physicians, nurses, pharmacists and other healthcare professionals as well as for patients and medical devices used to monitor patients [14].

Many hospital information systems are designed to support patient billing, inventory management and generation of simple statistics. Some hospitals use decision support systems, but are largely limited. They can answer simple queries like "What is the average age of patients who have heart disease?" , "How many surgeries had resulted in hospital stays longer than 10 days?", "Identify the female patients who are single, above 30 years old, and who have been treated for cancer." However they cannot answer complex queries like "Given patient records, predict the probability of patients getting a heart disease." Clinical decisions are often made based on doctors' intuition and experience rather than on the knowledge rich data hidden in the database [15].

A common approach for mining distributed data systems is the centralized one, in which all data is moved to a single central location and then mined. Another approach is the local one, where models are built locally at each site, and then moved to a common location where they are combined [2].

These existing approaches for Distributed Data Mining suffer from one or more of the following disadvantages.

- Lack of coordination among the distributed sites while generating the local knowledge, affects the quality of global knowledge [1].
- Non-flexibility in addition of new algorithm to its knowledge base [1]
- Capability to dynamically discover data sites based on user requirements [1]
- Network overhead, communication overhead and time overhead.

Existing decision support systems focuses on the hospital administration, billing, collection of money, insurance coverage , tracking of patient etc. Our Decision Support System focuses on the patients and their preferences. The cost of the healthcare service, the distance are considered as the important factors for our healthcare decision support system.

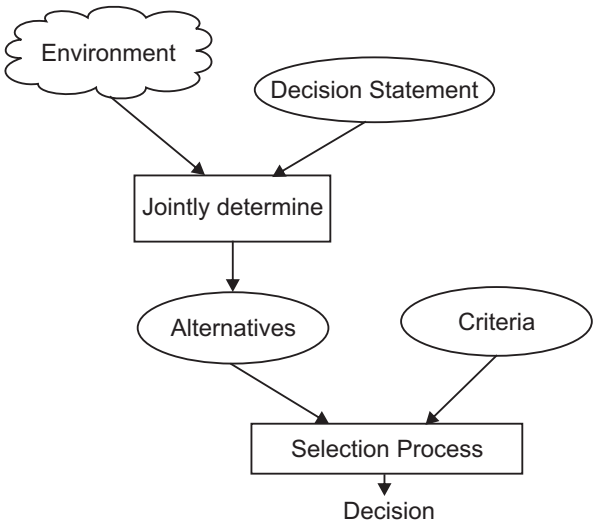

Fig. 2 Decision Support System

Some of the characteristics of Decision Support System are given below:[17]

1. They tend to be aimed at the less well structured, underspecified problems that upper level managers typically face.
2. They attempt to combine theuse of models or analytic techniques with traditional data access and retrieval functions
3. They specifically focus on features which make them easy to use by non computer people in a interactive mode
4. They emphasize flexibility and adaptability to accommodate changers in the environment and the decision making approach of the user

In our proposed work, with the use of the co-operative agents we are trying to minimize the above flaws prevailing in the existing works.

When the mobile agent acts agent patient agent, the central system acts as the static agent. The patient agent determines the selection of the hospital or healthcare provider based on the cost and distance

factor. The patient agent determines the selection based on the factors mentioned above.

III. PROPOSED WORK

To improve the accuracy of the knowledge obtained from the local model, strong association rules between the item sets are established. The decisions are derived based on the association rules. Static agents residing in the local system compute the local models, and are brought to the central system for combining them to generate global model at the central system based on which global knowledge is mined[11].

Knowledge from distributed sites is extracted in the form of association rules. Based on the types of values, the association rules can be classified into two categories [1]:

Example:

Boolean Association Rules: Keyboard $\rightarrow$ Mouse [Support $= 6\%$, Confidence $= 70\%$]

Quantitative Association Rules: (Age $= 26 \ldots 30) \rightarrow$ (Cars $= 1, 2$) [Support 3%, Confidence $= 36\%$]

A. Basic Concepts

Generally the Distributed Data Mining process consists of the following steps [1]:

- Generate the knowledge locally at each distributed sites by means of a static agent.
- Integrate the local distributed knowledge model to get global knowledge model
- Analyze to check the quality of the global model

The terminologies used in this algorithm proposed are as below:

DB	$\rightarrow$	Database
D	$\rightarrow$	Number of Transactions
n	$\rightarrow$	Number of Distributed systems (S_1, S_2, ... S_n)
DBi	$\rightarrow$	Distributed Data sets at Si, DB U DBi, i=1 to n
XSup	$\rightarrow$	Support count of a X at DB – Global
$XSup_i$	$\rightarrow$	Support count of a X at DBi – Local
Minsup	$\rightarrow$	Minimum support threshold
GFI	$\rightarrow$	Global Frequent Item set
CGFI	$\rightarrow$	Candidate Global Frequent Item set
X	$\rightarrow$	Global Frequent Item set if XSup=Minsup * D
LFi	$\rightarrow$	Local Frequent Item set at system i
PGFI	$\rightarrow$	Possible Global Frequent Item sets (not part of LFIi, but by adding these counts at Central system converts CGFI to GFI)
SAi	$\rightarrow$	Static Agent at each distributed system from $I = 1$ to n

B. Proposed Algorithm

Input: Distributed dataset DBi, i=1to n, Minsup

Output: Global Frequent Item set (GFI)

1. Sending mining agent to all sites

```
For I = 1 to n do
{
MA.send (Location = I, S=Support, Addresses of all
     Distributed sites);

}
```

2. (a) Each static agent (SAi) computes LFi in parallel

(b) Each cooperative agent gets LFIi from SAi and sends it to neighbors and the central system

3. (a) Compute GFI and CGFI at central system

GFI$\cap$ nLFIi, $I = 1$ to n; CGFI = ULFIi - $\cap$LFI, $i = 1$to n

(b) Calculate PGFI and their count at each distributed system and send it to the central system PGFIj = All Item sets at site $j \cap$ LFIi, $i = 1$to n, $i <> j$

4. Calculate GFI at central system using PGFI count

```
For all X ε CGFI do
{
If X.Sup=ÓX.Supi, i=1to n =MinSup*D then
{
GFI = GFI U {X}
}
}
```

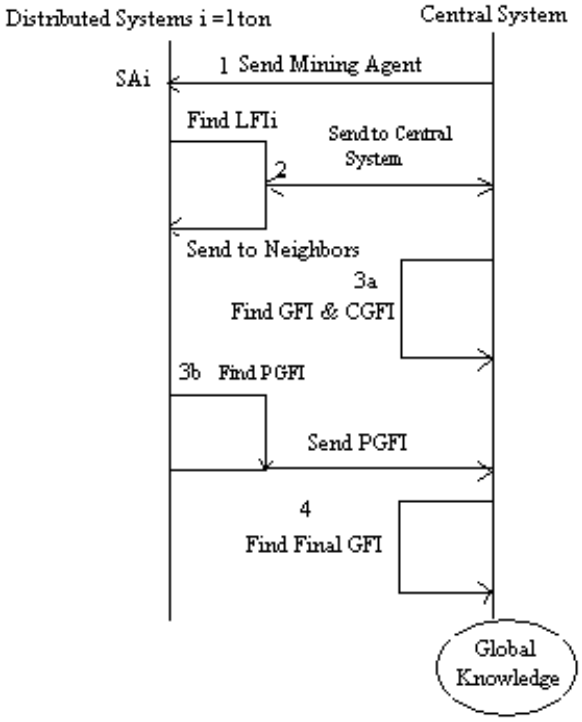

Fig. 3. Finding Final GFI using Static and Mobile agents

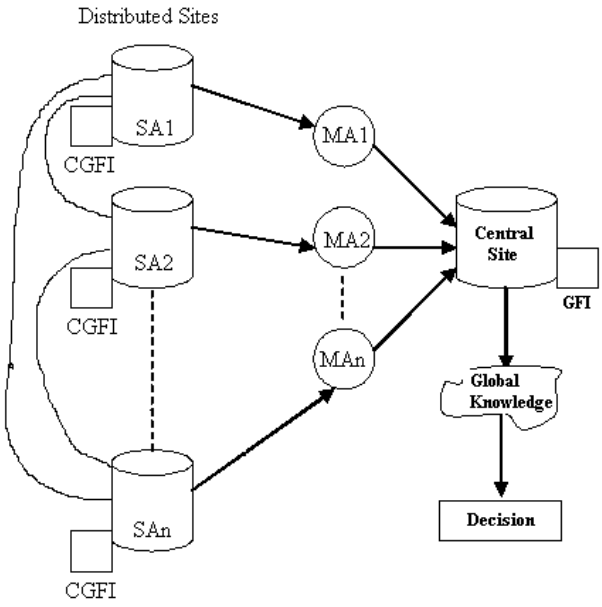

Fig. 4. Data mining in Distributed Communication Systems

The above Fig.2 depicts the functioning of our proposed algorithm. Mobile agents are directed towards

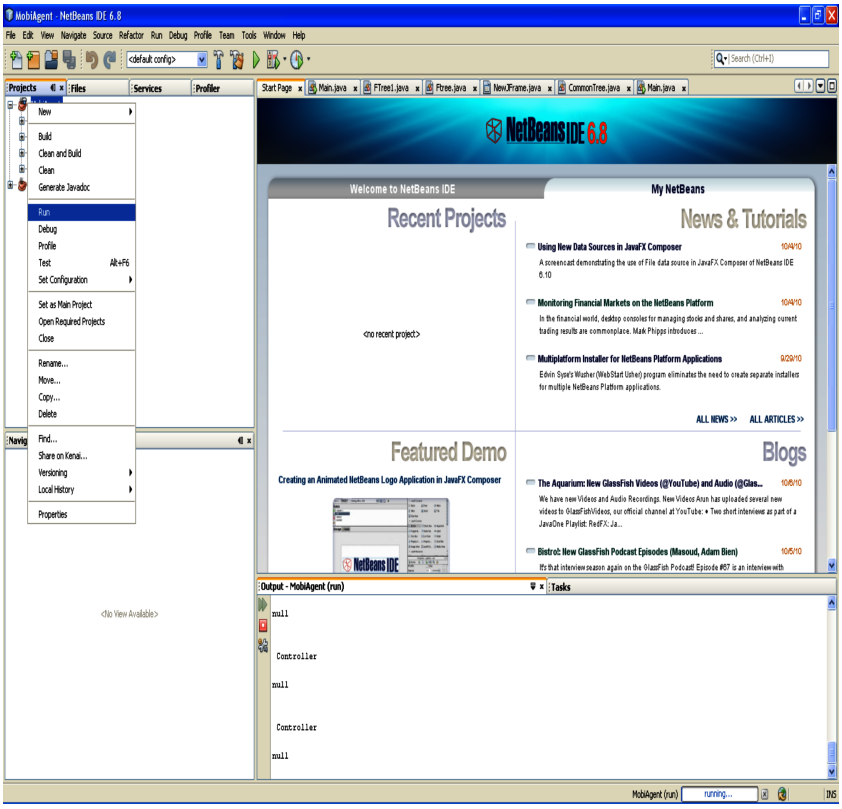

Fig. 5. Agent Initialization

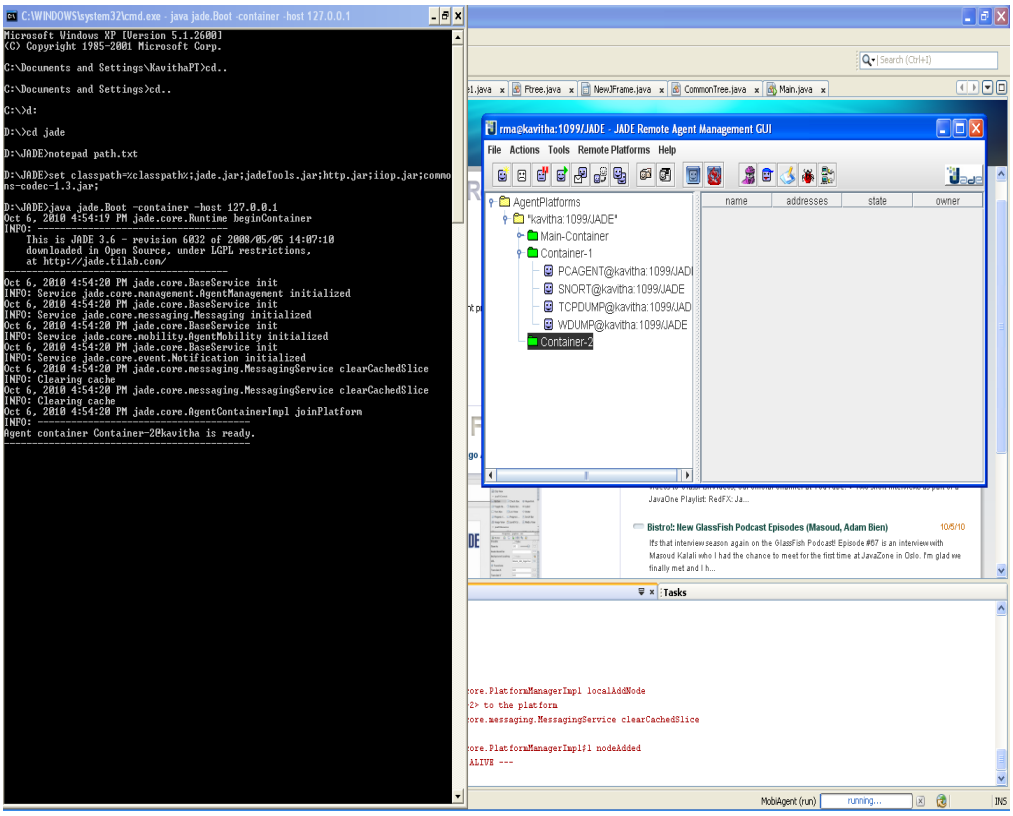

Fig. 6. Agent Implementation

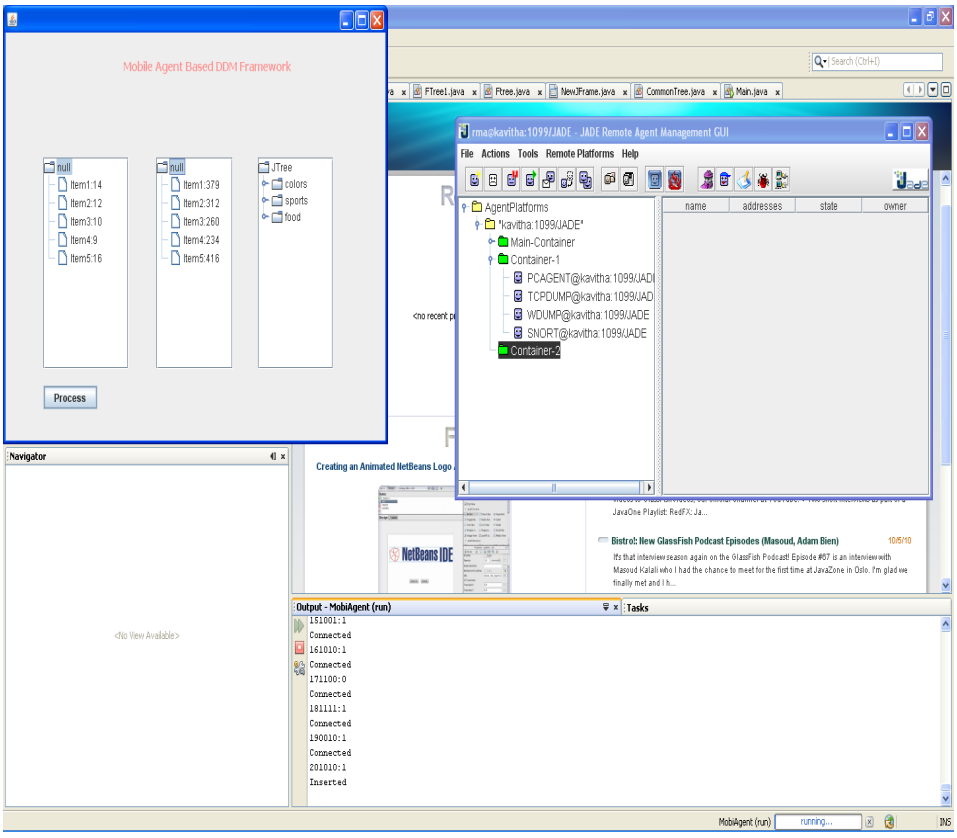

Fig. 7. Information Retrieval

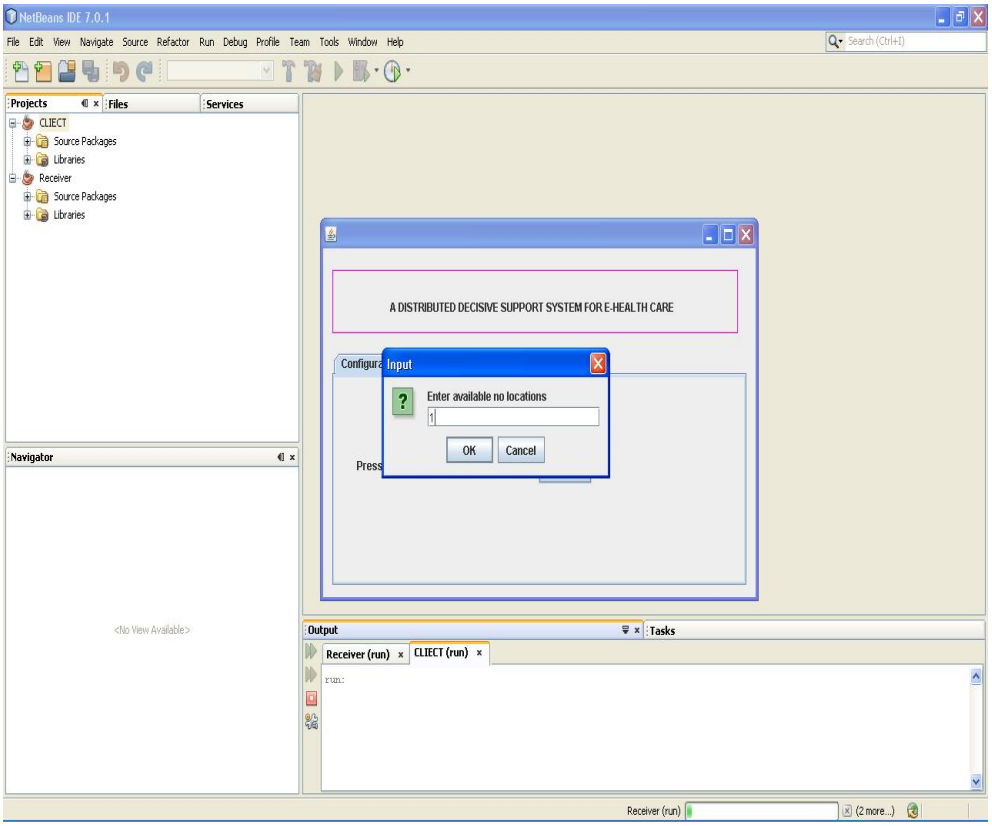

Fig. 8. Location Entry

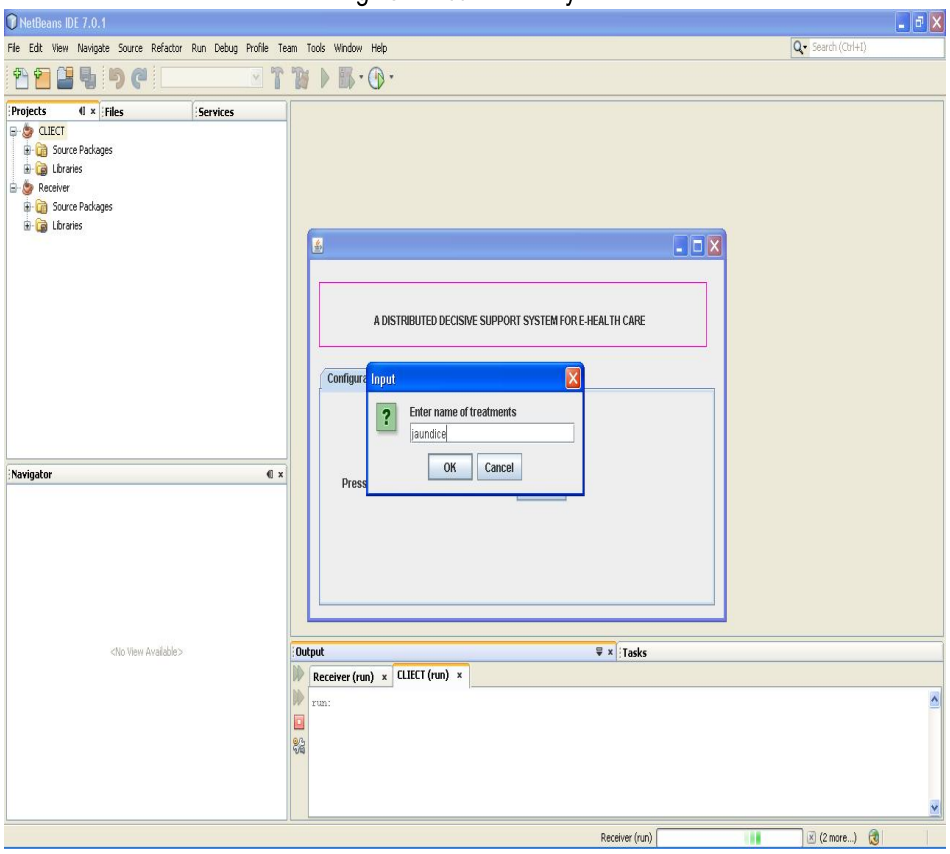

Fig. 9. Diagnosis of Treatments

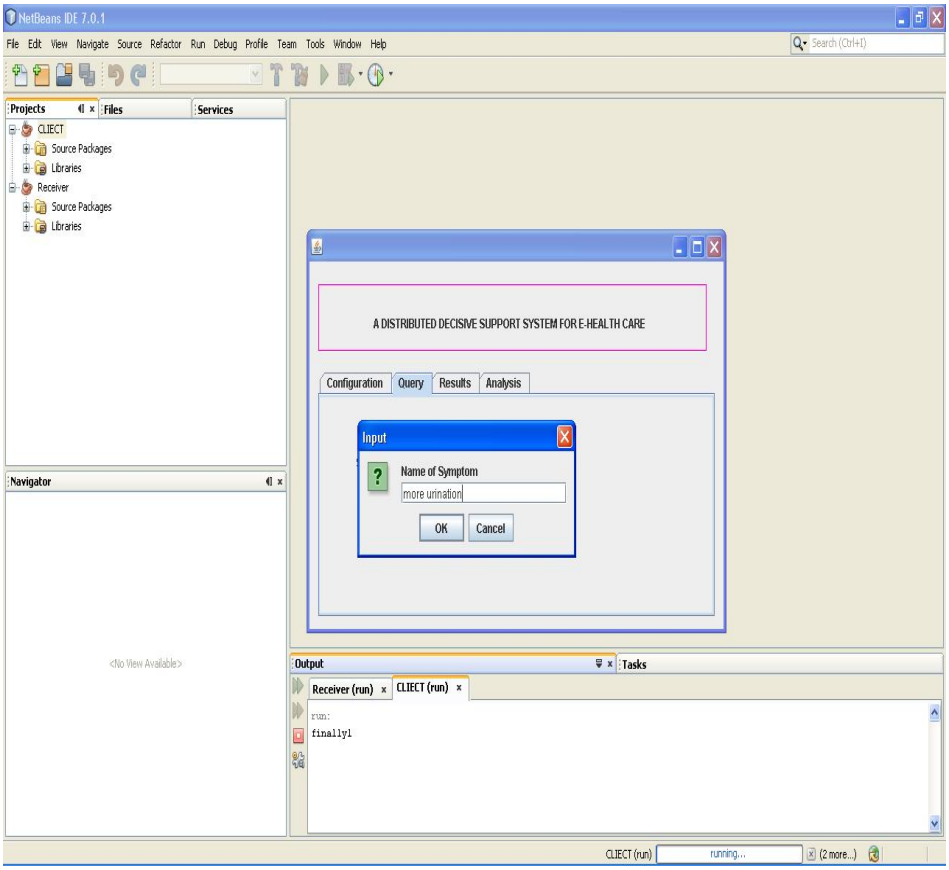

Fig. 10. Diagnosing the Diseases

the distributed systems with the static agents and pass the query.[11].

Agent technology is best able to cope with them in terms of autonomy, interaction, dynamic selection and gathering, scalability, multistrategy and collaboration [19].

The static agents then process the query to generate LFIi accordingly. The calculated LFIi is then sent to the neighboring systems and to the central system. Based on the received LFIi, the central system calculates GFI and CGFI based on the Minsup.

Now the infrequent item sets are analyzed and a PGFI count is computed by static agents and sent to the central system. With the PGFI count, the central site calculates the final GFI which is then updated in the global knowledge. This makes the process of decision making more efficient. In this analysis we have utilized the features of the JADE mobile agent [7].

The Decision Supportive System described here focuses on the feedback of the patients.

The decision making process makes the system to take a decision by learning from distributed datasets. For this association rule mining with supervised learning technique is used for machine learning by

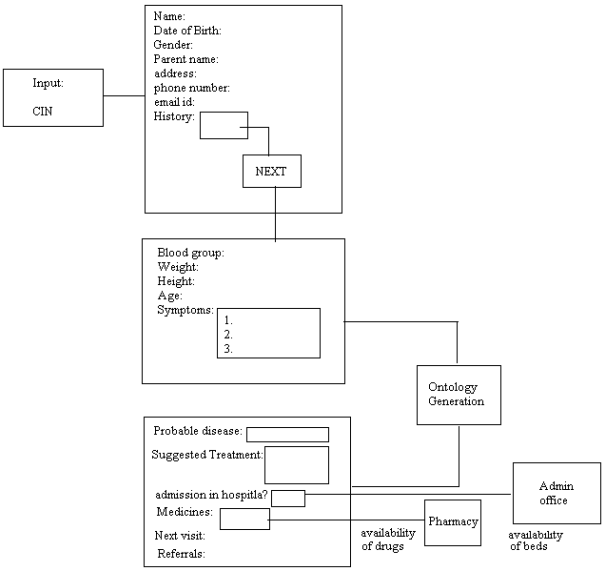

Fig. 11. Details of the Input

creating a function from training data. The global knowledge will be useful for the central system to effectively control the distributed systems with effective communication using the agents [11].

IV. CONCLUSION AND FUTURE WORK

In this paper we have focused the problems of mining frequent item sets on dynamic and distributed data sets in different parallel and distributed systems using mobile and static agents. We proposed an approach to minimize the response time and increase the accuracy of knowledge mining for the global set of frequent item sets, as well as to find frequent item patterns in infrequent item sets. The patient agent and system agents communicate and cooperate among themselves to make the right decision of selecting a healthcare service in rural areas.

The proposed framework cannot replace the doctor in anyway. It cannot derive the diagnosis for the patients but it can give useful suggestions and advices regarding the decisions patients make. The decision support system is designed to enhance and support the human.

Further development of the proposed Decision Support System should focus on the following:

(a) Identification and distribution of best practices

(b) Research about bench marking of DSS

(c) Questions of Security, data protection and confidentiality

(d) Quality and authenticity information

(e) Standardisation and inter operability of DSS

(f) Co-operation with other health care providers

Decision Support Systems are susceptible to security attacks. They always contain sensitive information. Medical Ethics should be followed in designing Decision Support Systems.

REFERENCES

[1] U.P.Kulkarni, P.D.Desai, Tan tanveer Ahmed, J.V.Vadavi, A.P.Yardi, " Mobile Agent based Distributed Data Mining", International Conference on Computational Intelligence and Multimedia Applications, 2007

[2] Matthew Eric Otey, Srinivasan Parthasarathy, member, IEEE, Chao Wang, Adriano Veloso, and Wagner Meira, "*Parallel and Distributed Methods for Incremental Frequent Intem set Mining*", IEEE transactions on systems, man and cybernetics Vol 34, No.6 December 2004

[3] Yun-Lan Wang, Zeng-Zhi Li, Hai-Ping Zhu, "Mobile agent based distributed and Incremental techniques for Association Rules", Proceedings of the Second International Conference on Machine Learning and Cybernetics, Wan, 2-5 November 2003.

[4] U.P.Kulkarni, K.K.Tangod, S.R.Mangalwede, A.R.Yardi, "Exploring the Capabilities of Mobile Agents in Distributed Data Mining", 10th International Database Engineering and Applications Symposium (IDEAS'06), 2006

[5] Wu-Shan Jijang, Ji-Hui Yu, "Distributed Data mining on the Grid", Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 18-21 August 2005

[6] M.Vigilson Prem, S.Swamynathan, "Group Communication based approach for Reliable Mobile Agent in Information Retrieval Applications", International Conference on Advances in Computing, Control and Telecommunication Technologies, 2009 IEEE

[7] F.Bellifemine, G.Caire, A.Poggi, G.Rimassa, "JADE- A White Paper", September 2003

[8] Raquel Trillo, Sergio Ilarri and Eduardo Mena, "Comparison and Performance Evaluation of Mobile Agent Platforms"

[9] Matthias Klusch, Stefao Lodi, Gianluca Moro, "Issues of Agent –Based Distributed Data Mining" AAMAS'03 Melbourne ACM I-58113-683-8/03/0007

[10] Vuda Sreenivasa Rao, "Multi Agent based Distributed Data Mining: An Overview", International Journal of Reviews in Computing 2009-2010 IJRIC & LLS, ISSN 2076-3328

[11] P.T.Kavitha, Dr.T.Sasipraba, " Co-Operative Mobile Agents for Distributed Communication Systems for Data Mining Applications", Emerging Trends Robotics and Communication Technologies (INTERACT), 2010.

[12] Ministry of Health and Family Welfare, " Rural Healthcare System in India", http://nrhm-mis.nic.in/

[13] Xuvan Xu and Lefei Li, " An Artificial Urban Health Care System and Applications", Intelligent Systems, IEEE, Issue 3, Volume 25, Page 63-73

[14] Firat Kart, Gengxin Miao, L.E.Moser,P.M Melliar-Smith, "A Distributed e-Healthcare System based on the Service Oriented Architecture", IEEE International Conference on Services Computing, 2007

[15] G.Subbalakshmi, K.Ramesh, M.ChinnaRao, " Decision Support in Heart Disease Prediction System using

Naïve Bayes", Indian Journal of Computer Science and Engineering(IJCSE), Vol. 2 No. 2 Apr-May 2011

[16] Robert Nisbet, John Elder, Gary Miner, " Handbook of Statistical Analysis and Data Mining and Applications", Elsevier 2007

[17] Efrem G.Mallach, " Decision Support and Data warehouse systems" Tata McGraw-Hill Edition, 2008

[18] J.Kan, M.Kamber, " Data Mining Concepts and Techniques", Morgan Kaufmann Publishers, pp.1-35, 2001

[19] Longbing Cao, Vladimir Gorodetsky and Peicles A.Mitkas, " Agent Mining: The Synergy of Agents and Data Mining", IEEE Intelligent Systems, 2009

[20] Wu-Shan Jiang, Ji Hui Yu, "Distributed Data Mining on the Grid", Fourth International Conference on Machine Learning and Cybernetics, IEEE, 2005