To Study The Role Of Data Mining And Cloud Computing

Tejaswini V. Ramekar[#] , Ranu R. Tuteja[*]
[#]ME Scholar, Department of Computer Science & Engineering, PRMIT&R, Badnera
[*]Assistant Professor, Department of Computer Science & Engineering, PRMIT&R, Badnera
tejaswiniramekar1@gmail.com
rrtuteja@mitra.ac.in

Abstract: **Data Mining is nothing but a process of extracting potentially useful information from raw Data, so as to improve the quality of the information service. Data mining has been considered as an essential component in business domain. We have recently seen an increase in data mining techniques targeted to such applications as fraud detection, identifying criminal suspects, and prediction of potential terrorists. Cloud Computing has become a main source for the data processing, storage and distribution. The storage of the data is simple and free to use. In data mining the data which is used as data security in a parallel computing platform.**

KEYWORD :
 Data Mining and cloud Computing.

I.   INTRODUCTION :

Data mining has been an effective tool for analyze data from different fields and getting useful information from data. Data mining techniques have evolved and became more useful and discovering knowledge in databases becoming increasingly in various sectors i.e business, medicine, science and engineering, spatial data etc. Cloud computing is a distributed computing paradigm which enables large datasets to be sliced and assigned to available computer nodes where the data can be processed locally, avoiding network transfer delays. The most revolutionary concept of recent year is Cloud Computing. Many companies are choosing as an alternative to building their own IT infrastructure .The use of cloud computing is gaining popularity due to its mobility, huge availability and low cost. Cloud infrastructure can be effectively used for demanding operations with data that is typical for processes of data mining. We believe that cloud computing can be an effective platform for data mining. Cloud computing not only provide users with a common parallel programming model and big data

processing capacity, but also provide users with an open computing services.

Literature Review:

In 1960's-1990's **John McCarthy** who has opined the modern-day characteristics of cloud computing, later in 1990's telecommunication companies who previously offered primarily dedicated point to point data circuit, began offering virtual private network(VPN) services with comparable quality of service, but at a lower cost. By switching traffic as they saw fit to balance server use, they could use overall network bandwidth more effectively. They
began to use the cloud symbols to donate the demarcation point between what the provider was responsible for cloud computing extends this boundary to cover server as well as the network infrastructure.

As computer become more prevalent, scientists and technologists explored ways to make large-scale computing power available to more users through time sharing, experimenting with algorithms to provide the optimal use of the infrastructure, platform and application with prioritized access to CPU and efficiency for the end users.

Since 2000 the dot-com bubble, Amazon played a key role in all the developing in cloud computing by modernizing their data center which like the most computer network, were using as little as 10% of their capacity at any one time, Just to leave room for occasional spikes having found that the new cloud architecture resulted in significant interval
efficiency improvements by whereby small, Fast moving "two-pizza terms" cloud added new features faster and more easier, Amazon indented a new product development effort to provide cloud computing to external customers, and launched Amazon web services (AWS) on utility computing basis in 2006.

In early 2008, Eucalyptus became the first open source the federation of the clouds in the same year efforts where focused on providing quality of service guarantees to cloud based infrastructure, in the frame works of the IRMOS European commission funded project resulting to a real time cloud environment by mid 2008, Gartner saw an opportunity for cloud computing to shape the relationship among consumers of IT services, projected shift to computing will result in dramatic growth in IT products in some areas and significant reduction in other areas. On march 1, 2011, IBM announced the IBM Smart Cloud framework to support smarter Planet, Among the various components of the smarter computing foundation, cloud computing is a critical piece.

II. CLOUD COMPUTING:

Cloud computing works in both the software and the hardware delivered as services over the Internet. NIST (National Institute of Standards and Technology) defines Cloud Computing as a model that provides a simple and on-demand network access to a shared set of resources. Cloud Computing can be defined as a new way, new technology and various distribution channel. Figure I. Below shows the six distinct phases of computing:

2.1. *Phase 1:* Many users use terminals to join to mainframes share by people.

2.2. *Phase 2*: Strength personal computers has enough to fulfill users every day work.

2.3. *Phase 3:*By the use of computer networks we has possible to connect computers to each other.

2.4. *Phase 4*: local networks could get connected to other local networks.

2.5. *Phase 5:*The electronic grid   used to share computing power and storage resources.

2.6. *Phase 6*: It gives advantage of exploitation of all present resources available on the Internet in a scalable and simple way.
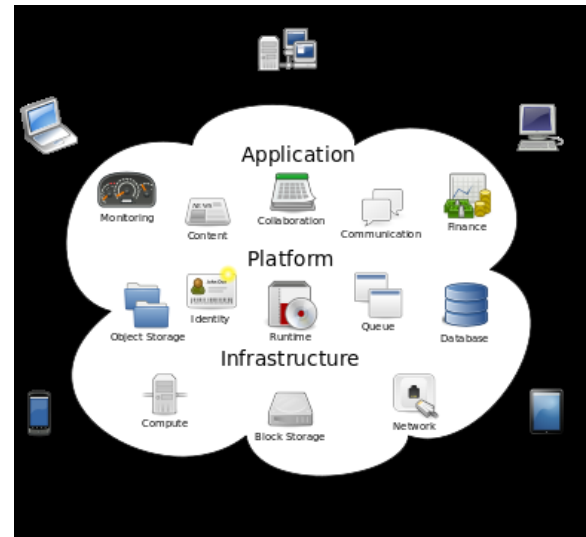

fig-1: Transferring data from one server to another server through the data mining.

Cloud computing model is buildup of five essential characteristics, three service models, and four deployment models."
The essential characteristics of cloud computing are on-demand self-service, broad network access, resource pooling, rapid elasticity and measured service.
The service models that compose cloud computing are Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS).
Regardless of the type of service delivery models (SaaS, PaaS, or IaaS),

2.1.There are four basic models of implementing Cloud Computing services, including:

2.1.1.*Public Cloud* - cloud available and open to the public for single person or organization.

2.1.2. *Private Cloud* - cloud accessible to only one organization. It can be managed by the organization itself.

2.1.3.*Community Cloud* - model of implementation that provides the ability for more organizations to share the same structure.

2.1.4.*Hybrid Cloud*-Model, which consists of two or more previously, discussed types of the establishment of cloud computing structure which remainunique and independent entities, but witha certain kind of reciprocal link, in order to achieve mobility of data between them.

Various characterics for cloud computing:

1.*Client server model* : Client server computing refers broadly to any distributed application that distinguishes between
service providers and service requesters.

2.*Grid computing*: A form of all distribute
and parallel computing, where by super and virtual composed by clusters on network
loosed coupled computers acting in concert to perform very large tasks.

3.*Mainframe computer*: Powerful computers where used mainly by large origination for critical applications ,typically
bulk data processing such as censes, industry costumers statistics ,police and secret agent service. Enterprises resources
planning and financial transaction processing.

4.*Utility computing*: The package of computing resource, such as computing and storage, as a mattered service similar
to traditional public utility such as electricity.

5.*Peer to peer*: It means the distributed architecture without the need of central coordination participants are both suppliers and consumers of resources in contrast other model of client server model.

DATA MINING:

Data mining finding useful patterns or trends through
large amounts of data.Data mining is carried out over large volumes ofdata in order to extract new information out of them. Data mining is a complex and challenging activity,
or set of activities, whose implementation requires experts from different fields.

Data mining components include*:*
1)Association- Association rule mining is an important segment of data mining of discovering in data. It is may be the most critical model created and examined by databases and information mining community.
2)Sequence analysis– Running across examples where one occasion prompts an alternate ensuing occasion.
3)Classification-Classification is an information mining strategy used to anticipate bunch for information examples.

4)Clustering- Clustering is a mechanized methodology to aggregation related records together on the groundwork of having comparative qualities for characteristics values for attributes.
5) Forecasting- Running across examples in information that can prompt sensible forecasts about what's to come. This range of information mining is reputed to be prescient investigation. There are numerous provisions in which the information mining term is exceptionally helpful. Information mining offers various provisions as Hospital, Student Management, Airline Reservation, Forecasting, Biometrics, Geographical, Web Mining and so forth. From wide region, we will be conclude on distributed computing, and will be depicting how information mining is utilized within distributed computing.
Data mining is broadly used multidisciplinary field, which has its roots in statistics, mathematics, information theory,
artificial intelligence, machine learning theory, data bases and in the whole series of other related fields.
Data mining is defined as a "type of database analysis that attempts to discover useful patterns or relationships in a group of data.

DATA MININIG AND CLOUD COMPUTING:

Cloud computing provide users with not only a common parallel programming model and big data processing capacity, but also provide users with an open computing services platform. As cloud computing is penetrating more and more in all ranges of business and scientific computing, it becomes a good area to be focused by data mining. Data mining with cloud computing offers tremendous.

potential for analyzing and extracting the (useful) information in various sectors of human activities i.e finance, banking, medicine, genetics, biology, pharmacy, marketing, etc. Data mining in cloud computing is the process of extracting important information from unstructured or semi-structured web data sources. Cloud provides technology that can "handle" huge amounts of data, which cannot be processed efficiently and at reasonable cost using standard technologies. Analyzing data which courses to social networks, pattern recognition, processing of large-scale images, encryption and description and, of course, data mining is just one of the examples of the tasks that are ideal for implementation in the Cloud. The implementation of data mining

techniques through Cloud computing will allow the users to retrieve meaningful information from virtually integrated data warehouse that reduces the costs of infrastructure and storage. Data mining in Cloud computing is, from a technical point of view, a very tedious process that requires a special infrastructure based on application of new storage technologies, handling and processing. Big Data is the latest hype in the field of data processing. The data mining in Cloud Computing allows organizations to centralize the management of software and data storage, with assurance of efficient, reliable and secure services for their users. Using data mining through Cloud computing reduces the barriers that keep small companies from benefiting of the data mining instruments. Data mining is one of the fastest growing fields in computer industry that deals

with discovering patterns from large data sets. It is a part of knowledge discovery process and is used to extract human

understandable information Mining is preferably used for a large amount of data and related algorithms often require

large data sets to create quality models .The relationship between data mining and cloud is worth to discuss.The system provides a variety of parallel data conversion rules and parallel data mining algorithms, the full support of the production, sales, marketing, financial management, corporate decision-making activities in the field, has broad application prospects.

The main effects of data mining tools being delivered by the Cloud are:

1.the customer only pays for the data mining tools that he needs – that reduces his costs since he doesn't have to pay for complex data mining suites that he is not using exhaustive;

2. the customer doesn't have to maintain a hardware infrastructure, as he can apply data mining through a browser – this means that he has to pay only the costs that are generated by using Cloud computing.
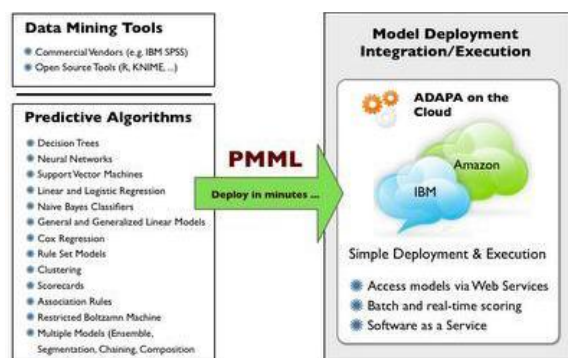

Fig:Data mining and cloud computing.

V. CONCLUSION:

The data that is processed does not originate only from multiple information system of companies, giant amount of it comes from "on-line" environment, with a variety of services that users use for both commercial and private purposes.

Data mining technologies provided through Cloud computing is an absolutely necessary characteristic for today's businesses to make proactive, knowledge driven decisions, as it helps them have future trends and behaviors predicted.

VI. REFERENCES:

[1] IEEE standard for information technology - portable operating system interface (POSIX) - part 1: System Application program interface (API) - amendment 2: Threads Extension, 1995.

[2]C. Ranger, R. Raghuraman, A. Penmetsa, G. Bradski, and C. Kozyrakis, Evaluatingmapreduce for multi-core and multiprocessor systems, IEEE13thInternational Symposium on High Performance Computer Architecture. (2007)13--24.

[3]Parikshit Prasad, BadrinathOjha, Rajeev Ranjanshahi, RatanLal, 3 Dimensional Security in Cloud Computing, IEEE, 2011.

[4] Uma Somani, KanikaLakhani, Manish Mundra, Implementing Digital Signature with RSA Encryption to Enhance Data Security of Cloud in Cloud Computing, IEEE.2010.

[5]Merriam-Webster Dictionary, "Definition of data mining",Link:http://www.merriam-webster.com/dictionary/data%20mining.

[6]ZhidongShen, Qiang Tong, "The Security of Cloud Computing System enabled by Trusted Computing Technology", 2010

[7]http://searchsqlserver.techtarget.com/definition/data-mining

[8]Wikipedia. Amazon elastic compute cloud — Wikipedia, the free encyclopedia, 2012. [Online; accessed 10-May-2011].

.