Global Journal of Computer Science and Technology
Cloud and Distributed
Volume 13 Issue 3 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Soft Computing on Medical-Data (SCOM) for a Countrywide Medical System using Data Mining and Cloud Computing Features

By Kalyani Mali & Samayita Bhattacharya
University of Kalyani, India

Abstract - This paper focuses on hosting and analyzing medical diagnostic data using soft computing. This is a project proposal for medical database system using soft computing, using data mining and cloud computing features. Data mining features are added to gain more control over the system. Cloud computing is a general term for anything that involves delivering hosted services over the Internet. The proposed database system can provide new delivery models to make healthcare more efficient and effective, and at a lower cost to technology budgets.

Keywords : *medical data, soft computing, data mining, cloud computing, database.*

GJCST-B Classification : *H.2.8*

Strictly as per the compliance and regulations of:

Soft Computing on Medical-Data (SCOM) for a Countrywide Medical System using Data Mining and Cloud Computing Features

Kalyani Mali [α] & Samayita Bhattacharya [σ]

Abstract - This paper focuses on hosting and analyzing medical diagnostic data using soft computing. This is a project proposal for medical database system using soft computing, using data mining and cloud computing features. Data mining features are added to gain more control over the system. Cloud computing is a general term for anything that involves delivering hosted services over the Internet. The proposed database system can provide new delivery models to make healthcare more efficient and effective, and at a lower cost to technology budgets.

Keywords : medical data, soft computing, data mining, cloud computing, database.

I. Introduction

In this paper we are going to discuss about medical data handling and how a countrywide medical system can be made efficient and effective way using soft computing. Furthermore, we are going to discuss different dada mining features and how cloud computing features can be used to make this new model. This is a project proposal for a medical data handling system using soft computing, for a countrywide medical system, using data mining and cloud computing features.

In last two decades, the applications of soft-computing methods in medicine (Biomedicine, which is a highly multidisciplinary area connecting medicine, computer science, statistics and many others) became a popular and important issue.

The medical industry requires new engineering technologies, to assess information objectively. While recent developments in medical engineering have been achieved by state-of-the-art of intelligent computing techniques including computer-aided diagnosis, computer-aided radiography, developments in computational techniques including soft computing (SC), information processing and data mining hold new premises in this field. SC methods are becoming indispensable for to sport modern medical practice. SC combines fuzzy logic (FL), neural networks (NN), and genetic algorithms (GAs) methodologies.

Authors α σ : Department of Computer Science & Engineering, University of Kalyani, Kalyani, West Bengal, India.
E-mail : samayita.bhattacharya@gmail.com

II. Medical Data

The terms medical record, health record, and medical chart are used somewhat interchangeably to describe the systematic documentation of a single patient's medical history and care across time within one particular health care provider's jurisdiction. The medical record includes a variety of types of "notes" entered over time by health care professionals, recording observations and administration of drugs and therapies, orders for the administration of drugs and therapies, test results, x-rays, reports, etc [4], [5], [8]. The maintenance of complete and accurate medical records is a fundamental requirement of health care providers and is generally enforced as a licensing or certification prerequisite. The terms are used for both the physical folder that exists for each individual patient and for the body of information found therein. A good medical records management system could mean the difference between life and death for some individuals [7].

Medical records have traditionally been compiled and maintained by health care providers, but advances in online data storage have led to the development of personal health records (PHR) that are maintained by patients themselves, often on third-party websites.

III. Soft Computing

Conventional computing or often called as hard computing, requires a precisely stated analytical model and often a lot of computation time. Many analytical models are valid for ideal cases, and real world problems exist in a non-ideal environment.

Soft computing differs from conventional (hard) computing in that, unlike hard computing, it is tolerant of imprecision, un-certainty, partial truth, and approximation. In effect, the role model for soft computing is the human mind. The guiding principle of soft computing is: Exploit the tolerance for imprecision, uncertainty, partial truth, and approximation to achieve tractability, robustness and low solution cost. Soft computing may be viewed as a foundation component for the emerging field of conceptual intelligence.

Few soft computing tools are: Fuzzy Systems, Neural Networks, Evolutionary Computation, Machine Learning and Probabilistic Reasoning.

Different soft computing tools can be used in different phases of the proposed system. Noise is a big issue in data storing, signal processing and networking. Wavelets can be used in removing noise. For the hierarchical classification and clustering, wavelets, fuzzy logic and neural network can be used. For searching evolutionary algorithm and for decision making fuzzy system can be used [10].

If we can incorporate artificial intelligence by the means of soft computing, the system might be able to predict the possible cause of illness and diagnose in case of unavailability of experts, by comparing the symptoms, history checking; and also give the possible relief suggestions.

IV. Data Mining

Data mining constitutes one or more of the following functions, namely classification, regression, clustering, summarization, image retrieval, discovering association rules and functional dependencies, rule extraction, etc [11].

There exist several domains where large volumes of data are stored in centralized or distributed database. Some of the important issues in data mining include the identification of applications for existing techniques and developing new techniques for traditional as well as new application domains like Web, bioinformatics.

There are many areas we can envisage, where data mining can be applied. Some of these areas which we can associate with our proposed system are:

- Medicine (determine disease outcome and effectiveness of treatment by analyzing patient diseases history to find some relationship between diseases).
- Molecular or pharmaceutical (identify new drugs and their effectiveness).
- Historical data analysis (search and access of historical data on similar case histories).
- Scientific data analysis (identify new medicines by searching for sub-clusters).

Other related services are:

- Website or web-store design and promotion, and layout modification (of the medical system for more user friendly interface).
- Marketing (of new found drugs can also be associated).
- Insurance (health insurances).
- Geological/geographical studies (studies related to disease and drug effects depending on the geographical location).
- Planning.

A particular data mining algorithm is usually an instantiation of the model-preference-search components. The models that can be useful for medical data handling are:

- Classification
- Regression
- Clustering
- Rule generation
- Summarization or condensation.

V. Cloud Computing

Cloud computing is a technology that uses the internet and central remote servers to maintain data and applications. Cloud computing allows consumers and businesses to use applications without installation and access their personal files at any computer with internet access. This technology allows for much more efficient computing by centralizing storage, memory, processing and bandwidth.

A simple example of cloud computing is Yahoo email, Gmail, or Hotmail etc [3]. One doesn't need a software or a server to use them. All a consumer would need is just an internet connection and you can start sending emails. The server and email management software is all on the cloud (internet) and is totally managed by the cloud service provider Yahoo, Google etc. The consumer gets to use the software alone and enjoy the benefits.

Figure 1 shows the cloud computing system. In a cloud computing system, there's a significant workload shift. Local computers no longer have to do all the heavy lifting when it comes to running applications. The network of computers that make up the cloud handles them instead. Hardware and software demands on the user's side decrease. The only thing the user's computer needs to be able to run is the cloud computing system's interface software, which can be as simple as a Web browser, and the cloud's network takes care of the rest.

Cloud computing is a general term for anything that involves delivering hosted services over the Internet. These services are broadly divided into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). The name cloud computing was inspired by the cloud symbol that's often used to represent the Internet in flowcharts and diagrams [2]. Cloud computing providers offer their services according to these fundamental models, where IaaS is the most basic and each higher model abstracts from the details of the lower models [1]. Figure 2 shows the layers of cloud computing.

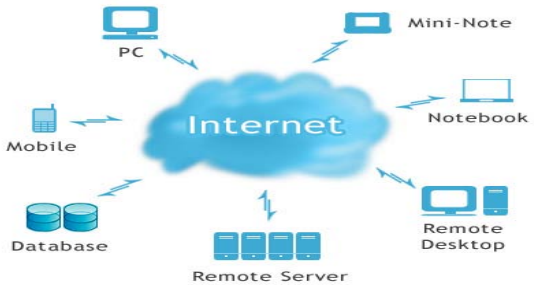

Figure 1 : Cloud Computing

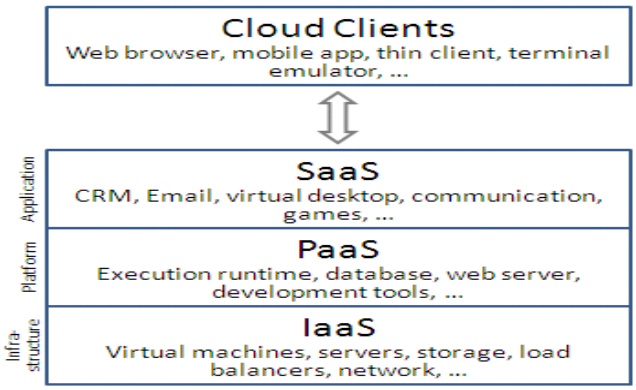

Figure 2 : Layers of cloud computing

Cloud computing is the delivery of computing as a service rather than a product, whereby shared resources, software, and information are provided to computers and other devices as a metered service over a network cloud (typically the Internet).

Cloud computing is a marketing term for technologies that provide computation, software, data access, and cloud services that don't require end-user knowledge of the physical location and configuration of the cloud that delivers the services.

Cloud computing providers deliver applications via the internet cloud, which are accessed from web browsers and desktop and mobile apps, while the business software and data clouds are stored on servers at a remote location.

As a metaphor for the Internet, "the cloud" is a familiar cliché, but when combined with "computing," the meaning gets bigger and fuzzier. Some analysts and vendors define cloud computing narrowly as an updated version of utility computing: basically virtual servers available over the Internet.

Cloud computing is at an early stage, with a motley crew of providers large and small delivering a slew of cloud-based services, from full-blown applications to storage services to spam filtering.

The applications of cloud computing are practically limitless. With the right middleware, a cloud computing system could execute all the programs a normal computer could run. Potentially, everything from generic word processing software to customized computer programs designed for a specific company could work on a cloud computing system.

We can use cloud computing for medical data storage and diagnosis.

VI. Medical Data and Diagnosis using Soft Computing

We are proposing a countrywide (or preferably worldwide) medical database system using soft computing, using data mining and cloud computing features [9].

a) *Features of Soft Computing to be used in the System*

Different soft computing tools can be used in different phases of Medical Data Handling Using Soft Computing. i.e. denoising, classification, clustering, filtering, searching, matching, customized searching and filtering.

Soft computing (SC) is not a new term; we have gotten used to reading and hearing about it daily. Nowadays, the term is used often in computer science and information technology. It is possible to define SC in different ways. Nonetheless, SC is a consortium of methodologies which works synergistically and provides, in one form or another, flexible information processing capability for handling real life ambiguous situations. Its aim is to exploit the tolerance for imprecision, uncertainty, approximate reasoning and partial truth in order to achieve tractability, robustness and low-cost solutions. SC includes fuzzy logic (FL), neural networks (NNs), and genetic algorithm (GA) methodologies. SC combines these methodologies as FL and NN (FL-NN), NN and GA (NN-GA) and FL and GA (FL-GA). Recent years have witnessed the phenomenal growth of bio-informatics and medical informatics by using computational techniques for interpretation and analysis of biological and medical data. Among the large number of computational techniques used, SC, which incorporates neural networks, evolutionary computation, and fuzzy systems, provides unmatched utility because of its demonstrated strength in handling imprecise information and providing novel solutions to hard problems.

Search results put the case clearly that FL-GA methodology has not applied well enough to medicine yet. Undeniable interest in studying SC methodologies in genetics, physiology, radiology, cardiology, and neurology disciplines proves that studying SC is very fruitful in these disciplines and it is expected that future researches in medicine will use SC more than it is used today to solve more complex problems.

i. *Wavelets*

A wavelet is a wave-like oscillation with amplitude that starts out at zero, increases, and then decreases back to zero. It can typically be visualized as

a "brief oscillation" like one might see recorded by a seismograph or heart monitor. Generally, wavelets are purposefully crafted to have specific properties that make them useful for signal processing. Wavelets can be combined, using a "shift, multiply and sum" technique called convolution, with portions of an unknown signal to extract information from the unknown signal.

As wavelets are a mathematical tool they can be used to extract information from many different kinds of data, including, but not limited to, audio signals and images. Sets of wavelets are generally needed to analyze data fully.

A use of wavelet is that of smoothing/denoising data based on wavelet coefficient thresholding, also called wavelet shrinkage. By adaptively thresholding the wavelet coefficients that correspond to undesired frequency components smoothing and/or denoising operations can be performed.

Role of wavelets in different aspects of data mining is gaining significant importance. Today it has become a very powerful signal processing tool in different application area such as image processing, compression, digital libraries, image clustering and databases.

Spatial data mining aims to handle the huge amount of spatial data obtained from medical equipments. The objective is to automate the process of understanding spatial data by concise representation and reorganization to accommodate data semantics.

Wavelet transform is a signal processing technique that decomposes a signal or image into different frequency sub-bands at number of levels and multiple resolutions. This property led to application of medical image fusion.

Wavelets are used for clustering too [11].

ii. *Neural Network*

The term neural network was traditionally used to refer to a network or circuit of biological neurons. The modern usage of the term often refers to artificial neural networks, which are composed of artificial neurons or nodes. Artificial neural networks are composed of interconnecting artificial neurons (programming constructs that mimic the properties of biological neurons).

Artificial neural networks may either be used to gain an understanding of biological neural networks, or for solving artificial intelligence problems without necessarily creating a model of a real biological system. The real, biological nervous system is highly complex and includes some features that may seem superfluous based on an understanding of artificial networks.

Real life applications and the tasks to which artificial neural networks provide solutions include classification; including pattern and sequence recognition, novelty detection and sequential decision making.

We can use neural network for classification.

iii. *Fuzzy Logic*

Fuzzy logic is a form of many-valued logic derived from fuzzy set theory to deal with reasoning that is fluid or approximate rather than fixed and exact.

In contrast with "crisp logic", where binary sets have two-valued logic, fuzzy logic variables may have a truth value that ranges in degree between 0 and 1.

In simple words we can say fuzzy logic is a super set of conventional (boolean) logic that has been extended to handle the concept of partial truth--the truth values between completely true and completely false. Furthermore, when linguistic variables are used, these degrees may be managed by specific functions.

For implementing the proposed expert system, fuzzy logic is needed. Using fuzzy logic, the probability of a certain case to fall in any of the cluster can be calculated, and there after decision can be made based upon the value.

iv. *Evolutionary Algorithm*

In artificial intelligence, an evolutionary algorithm (EA) is a subset of evolutionary computation, a generic population-based meta-heuristic optimization algorithm.

Evolutionary optimization methods can be used for searching the database.

b) Fields of Data Mining to be used in the System

- Medical imagery: Large volume of data generated (image reports like x-ray, scans) and stored in large centralized and distributed databases in medical management systems. Automatic mining of these data is important to the medical community.
- Health care: Other than image data-type, there are many important documents associated with medical management system, such as health insurance information, personal care physician's information, specialist information, medical history etc. These data are stored and several diagnostic information are also stored by hospital management system for ready reference and research.
- Investment
- Manufacturing and Production
- Business and Marketing
- Scientific Domain [11].

c) Benefits of using Cloud Computing

One of the greatest benefits of using cloud computing technology in medical office environment is ease of access to data as opposed to library rows of physical files.

One way it's beneficial is access to patient records and office management systems from any platform or device without having to worry about maintaining and managing on-premise hardware and software.

It can also be especially useful, not only to the medical staff, but office administrators as well.

Cloud storage solutions can be used to deliver records and applications to tablets and desktop computers in the medical office, allowing patients and office administrators to move away from handwritten forms, duplicate data entry, and error prone record copying.

Cost savings is another advantage of moving to cloud computing. One can manage infrastructure a lot more effectively and lower the cost needed to support these applications by managing a shared infrastructure mode.

d) Database Related Issues

A database is a structured collection of data. The data are typically organized to model relevant aspects of reality (for example, the availability of beds in hospitals), in a way that supports processes requiring this information (for example, finding a hospital within a fixed distance with available bed for a patient).

The term database system implies that the data are managed to some level of quality (measured in terms of accuracy, availability, usability, and resilience) and this in turn often implies the use of a general-purpose database management system (DBMS). A general-purpose DBMS is typically a complex software system that meets many usage requirements to properly maintain its databases which are often large and complex. This is specially the case with client-server, near-real time transactional systems, in which multiple users have access to data. Figure 3 shows a simplified view of a database system.

There are two main types of database; flat-file and relational. Which is the best one to use for a particular job will depend on factors such as the type and the amount of data to be processed; not to mention how frequently it will be used.

One major advantage of the relational model is that, if a database is designed efficiently, there should be no duplication of any data; helping to maintain database integrity. This can also represent a huge saving in file size, which is important when dealing with large volumes of data. The "relation" comes from the fact that the tables can be linked to each other, for example the author of a book could be cross-referenced with the authors table (assuming there was one) to provide more information about the author. These kind of relations can be quite complex in nature, and would be hard to replicate in the standard flat-file format.

In most cases, we would want your database to support various types of relations; such databases, particularly if designed correctly, can dramatically improve the speed of data retrieval as well as being easier to maintain.

Ideally, we will want to avoid the replication of data within a database to keep a high level of integrity, otherwise changes to one field will have to be made manually to those that are related.

This proposed system needs robust searching and filtering. Also classification is needed during the query.

Few things to be taken good care of are:

- Accuracy
- Availability
- Security
- Authorization
- Integrity

We are going to use cloud computing so the database related issues are to be dealt using the cloud which is often a third party database system. And in our system it's essentially a 3rd party system. And as we are proposing a worldwide system it would be consisting of more than one cloud.

Figure 3 : Simplified view of a database system

e) Our Proposal

- Make a countrywide (or preferably worldwide) database.
- Given every individual has or need own personal, health care data, give every citizen a unique MEDICAL-ID.
- Whenever anyone goes for any medical test, or any other medical supervision, update the data in the database for the particular ID.
- Every diagnostic centre, hospital, doctor's chamber, medical shop, and also every individual should have access to that database.

f) Features of the Proposed System

- This way anyone will not need to carry any medical data with them, all the data will be stored in the database and can be accessed from anywhere in the country (or worldwide).
- This way we not only can save data (like x-ray report) from being lost or mutilated, also will get access to data anywhere and whenever we needed.
- When a new test is conducted, the reports will be updated in the database, and when a doctors or panel of doctors' advice is needed doctors only

needed to check the database, and update their views about the patient's case. And this way we can also see the others doctors views about a patient case.

- This system can be used for expert's suggestion for a particular patient, just by checking the database by the expert from anywhere of the world without the patient being needed to be present in front of the expert.
- Every individual also should have access that database, so that everyone can check about the current status.

g) Added Features of the System

- ✓ Interaction between medical experts: To gain maximum flexibility about diagnosis and patient follow-ups by experts.
- ✓ Interaction between patients: (group of similar people interaction) to know about the life style and care method of other patients diagnosed with the same diseases.

h) An Expert System

In case of emergency and if any expert is not available, this system night be used as an expert system to give temporary relief to the patient, by providing the symptoms; and the system might be able to give suggestion for instant and /or temporary relief. Even if a medical expert wants to take advise/support from the expert system he can access it; to take a suggestion, from the system to get the probable cause of illness, or supposed remedy. An expert can take full support from the system or take partial support, or run the system and stop at a certain point and use own method in earlier or later phase of medication.

As the database has all the medicine names stored in it along with its usage, a medical expert can rely on this system rather than depending on memory.

VII. Barriers for Medical Offices to Overcome

There are still some barriers for some medical practices. It all starts with change, it's going to change the way people fundamentally understand how things. Also, it's a matter of familiarity, as people become more familiar with it and sees the advantages; they are going to start accepting it more.

Cost and security concerns are other barriers for many practices. Cloud computing requires some upfront investment, training and education, patient data is extremely sensitive, so there are security and privacy concerns that give pause to new technology adoption. The first thing people worry about is security. Security is always the most prominent concern when dealing with sensitive information such as patient information. They worry about relinquishing control because it's no longer in their office or in their data center.

VIII. Conclusion

Soft computing features can help us to make the medical data handling system more reliable. With the help of artificial intelligence we can make an automated expert system.

Soft computing features in different phases of the system can help to make the system more effective and efficient.

Cloud computing is revolutionizing the healthcare industry. Many current desktop applications that medical practices now use require large processing power on the client side, but there is a shift towards cloud computing. It can provide new delivery models to make healthcare more efficient and effective, and at a lower cost to technology budgets.

Prevalence of new and innovative applications that utilize the cloud will make it easier for doctors to reference patient history, submit referrals, process prescriptions, and interact with the patient community.

One of the things that need to going to start to see moving forward is more solutions offered in the cloud. The reason for that is in healthcare, we have to balance several things at the same time. We have to balance the fact that infrastructure is getting more complicated moving forward, everywhere from the servers to storage to everything else, and also have a struggle with privacy and security.

There is going to be a challenge with how will healthcare organizations make all of this data electronic as well as how are they also going to make it private and secure. On top of that, we are seeing organizations and individuals who need access to this information anytime, anywhere, so the ability, whether they access this information on their mobile phone, tablet device, laptop, is becoming increasingly important.

References Références Referencias

1. Cloud computing http://en.wikipedia.org/wiki/Cloud_computing
2. What is Cloud Computing http://searchcloudcomputing.techtarget.com/definition/cloud-computing
3. Concept: Cloud Computing http://www.wikinvest.com/concept/Cloud_Computing
4. Medical record http://en.wikipedia.org/wiki/Medical_record
5. Medical diagnosis http://en.wikipedia.org/wiki/Medical_diagnosis
6. Data, Tools and Statistics-U.S. National Library of Medicine http://www.nlm.nih.gov/hsrinfo/datasites.html
7. Medical Records Management: Conquering Piles of Paper http://www.smead.com/hot-topics/medical-records-management-1374.asp
8. Health care http://en.wikipedia.org/wiki/Health_care

9. Samayita Bhattacharya and Kalyani Mali, "Medical Data Handling Using Cloud Computing And A Proposal for Countrywide Medical System", i-manager's Journal on Software Engineering (JSE), July - September 2012 issue, ISSN Print: 0973-5151, ISSN Online: 2230-7168.
10. Kalyani Mali and Samayita Bhattacharya, "Soft Computing Approach towards Fingerprint Recognition", Proc of National Conference on Recent Trends in Soft Computing (NCRTSC 2011), ISBN 978-93-81006-03-0, pp 95-99.
11. Sushmita Mitra, Tinku Acharya, "Data Mining Multimedia, Soft Computing and Bioinformatics", Wiley-interscience, ISBN: 9812-53-063-0.

This page is intentionally left blank