

# Data Analysis of Phone Calls in Cloud Computing

Niranjan Satpute<sup>1</sup>, Priyanka Khatal<sup>2</sup>, Pradnya Katkar<sup>3</sup>

Dept. of Information Technology Dept. of Computer Engineering,  
Savitribai Phule Pune University, K L University  
AISSMS Institute Of Information Technology  
Kennedy Road, Near RTO, Pune-411001

**Abstract-** Customers want to phone a support agent for immediate help. In system, how we can add a SoftPhone, otherwise known as a call-control tools like transfer the call, conference the call, gives connection for new line. Ideally data analytics helps to eliminate much of the guesswork involved in trying to understand clients, instead systemically tracking data patterns to best construct business tactics and operations to minimize uncertainty. The efficient data analysis methods and model will help in business development. User can easily track product and their preferences using data analysis in cloud computing

**Keywords-** Cloud Computing , CTI System, Data analysis , Guesswork, Softphone.

## I. INTRODUCTION

A data deluge is currently drowning us. Data sources are everywhere, from Web 2.0 to large scientific experiments, from social networks to sensors. This huge amount of data is a mine hiding valuable information. Nowadays, the extraction of knowledge from such ocean is a competitive advantage for most companies. Cloud computing is the emerging technology for large scale data analysis, providing scalability to thousands of computers, in addition to fault tolerance and cost effectiveness. The current solutions in the yield of cloud computing provides transparent access to large scale computing and storage hardware by exposing two main functionalities to the users: distributed data management and distributed computing. Both the two layers provide simple access and an high level of fault tolerance. The system automatically handles multiple distributed copies and it monitors every running job possibly re-executing it in case of failure. On the flip side of the coin, the data level is designed for unstructured data and for parallel scans, while the computing level supports novel and completely different programming models. Much of the benefit from data analysis comes from its ability to recognize patterns in a set and make predictions regarding past experiences. For example in Virtual call center system SoftPhone with very limited capabilities. will have created a SoftPhone that lets agents make phone numbers clickable in Salesforce, and screen pop a contact record. Later, some resources that will help to build more robust call tools. Some background on CTI: Salesforce and its

partners offer many options to choose from, including Open CTI, which gives you the benefits of cloud architecture and less maintenance utilize sales data in hopes of equipping them to increase profit margins[2].

## II. DATA RETRIEVAL MODEL

It is hard to classify IR algorithms, and to draw a line between each type of application.

However, we can identify three main types of algorithms, which are described below. There are other algorithms used in IR that do not fall within our description, for example, user interface algorithms. The reason that they cannot be considered as IR algorithms is because they are inherent to any computer application.

We distinguish three main classes of algorithms. These are retrieval, indexing, and filtering algorithms [1].

### Retrieval Algorithms

The main class of algorithms in IR is retrieval algorithms, that is, to extract information from a textual database. We can distinguish two types of retrieval algorithms, according to how much extra memory we need:

- Sequential scanning of the text: extra memory is in the worst case a function of the query size, and not of the database size. On the other hand, the running time is at least proportional to the size of the text, for example, string searching .
- Indexed text: an "index" of the text is available, and can be used to speed up the search. The index size is usually proportional to the database size, and the search time is sub linear on the size of the text, for example, inverted files and signature files .

Formally, we can describe a generic searching problem as follows: Given a string  $t$  (the text), a regular expression  $q$  (the query), and information (optionally) obtained by pre-processing the pattern and/or the text, the problem consists of finding whether  $t \in \Sigma^* q \in \Sigma^*$  ( $q$  for short) and obtaining some or all of the following information:

1. The location where an occurrence (or specifically the first, the longest, etc.) of  $q$  exists. Formally, if  $t \in \Sigma^* q \in \Sigma^*$  find a position  $m \geq 0$  such that  $t \in \Sigma^m (from 0 to m) q \in \Sigma^*$ . For example,

the first occurrence is defined as the least  $m$  that fulfils this condition.

2. The number of occurrences of the pattern in the text. Formally, the number of all possible values of  $m$  in the previous category.

3. All the locations where the pattern occurs (the set of all possible values of  $m$ ). In general, the complexities of these problems are different. The efficiency of retrieval algorithms is very important, because we expect them to solve on-line queries with a short answer time. This need has triggered the implementation of retrieval algorithms in many different ways: by hardware, by parallel machines, and so on [3].

### Filtering Algorithms

This class of algorithms is such that the text is the input and a processed or filtered version of the text is the output. This is atypical transformation in IR, for example to reduce the size of a text, and/or standardize it to simplify searching.

The most common filtering/processing operations are:

- Common words removed using a list of stop words
- Uppercase letters transformed to lowercase letters
- Special symbols removed and sequences of multiple spaces reduced to one space;
- Numbers and dates transformed to a standard format
- Word stemming (removing suffixes and/or prefixes)
- Automatic keyword extraction;
- Word ranking.

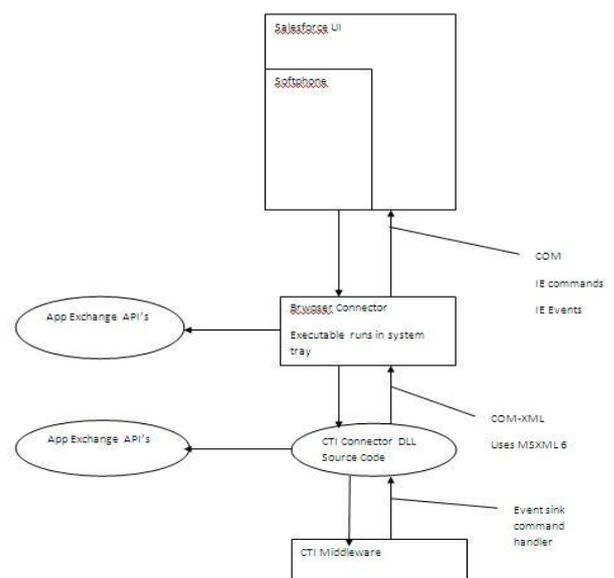
Unfortunately, these filtering operations may also have some disadvantages. Any query, before consulting the database, must be filtered as is the text; and, it is not possible to search for common words, special symbols, or uppercase letters, nor to distinguish text fragments that have been mapped to the same internal form [4].

### Indexing Algorithms

The usual meaning of indexing is to build a data structure that will allow quick searching of the text, as we mentioned previously. There are many classes of indices, based on different retrieval approaches. For example, we have inverted files, signature files, tries, and so on. Almost all type of indices are based on some kind of tree or hashing. Perhaps the main exceptions are clustered data structures (this kind of indexing is called clustering), which is covered in further laboratories, and the Direct Acyclic Word Graph (DAWG) of the text, which represents all possible sub words of the text using a linear amount of space and is based on finite automata theory. Usually, before indexing, the text is filtered.

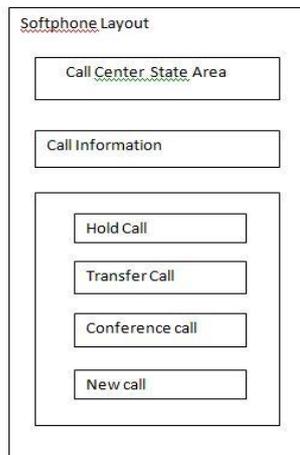
The pre-processing time needed to build the index is amortized by using it in searches. For example, if building the index requires  $O(n \log n)$  time, we would expect to query the database at least  $O(n)$  times to amortize the pre-processing cost. In that case, we add  $O(\log n)$  reprocessing time to the total query time (that may also be logarithmic) [5].

**CTI System:** A set of telephony hardware and software that supports integration with other computer systems. The CTI system provides the base framework for the calls that are made and received through a SoftPhone CTI adapter. A lightweight software program that controls the appearance and behavior of a Salesforce SoftPhone. The adapter acts as an intermediary between a third-party computer telephony integration (CTI) system, Salesforce, and a Salesforce CRM Call Center user. It must be installed on any machine that needs access to Salesforce CRM Call Center functionality. The source of call-related data and SoftPhone layout instructions. In addition to providing access to records that are related to an incoming call, Salesforce stores call center connection information and directories, SoftPhone layouts, and automatic call logs.[9]



**Softphone:** Users can manually trigger CTI system events, such as an incoming call with an automatic number identification (ANI) value, an incoming call without an ANI value, an incoming conference call, and an incoming transfer request. The SoftPhone operates as if these events had been triggered by a valid CTI system and allows the user to view associated records and generate call logs as if they had been

real. As a developer, you can use the demo adapter to quickly prototype new SoftPhone functionality, or as a template for a new CTI connector[9].



### III. APPLICATIONS

1. Social Media: A popular use for cloud data analytics is compounding and interpreting social media activity. Before cloud drives became practical, it was difficult processing activity across various social media sites, especially if the data was stored on different servers. Cloud drives allow for the simultaneous examination of social media site data so results can be quickly quantified and time and attention allocated accordingly [7].

2. Tracking Products: Long thought of as one of the kings of efficiency and forethought, it is no surprise Amazon cosmozes data analytics on cloud drives to track products across their series warehouses and ship items anywhere as needed, regardless of items proximity to customers. Alongside Amazon's use of cloud drives and remote analysis, they are also a leader in big data analysis services thanks to their Redshift initiative. Red shift gives smaller organizations many of the same analysis tools and storage capabilities as Amazon and acts as an information warehouse, preventing smaller businesses from having to spend money on extensive hardware [6].

3. Tracking Preference: Over the last decade or so, Netflix has received a lot of attention for its DVD deliver service and the collection of movies hosted on their website. One of the highlights of their website is its movie recommendations, which tracks the movies users watch and recommends others they might enjoy, providing a service to clients while supporting the use of their product. All user information is remotely stored on cloud drives so users' preferences do not change from computer to computer. Because Netflix retained all their users' preferences and tastes in movies and television, they were able to create a television show that statistically appealed to a large portion of their audience based on their demonstrated taste. Thus in 2013, Netflix's House of Cards

became the most successful internet-television series ever, all thanks to their data analysis and information stored on clouds [6].

### IV. CONCLUSION

Automate account details on your incoming calls. Our "Screen Pops" instantly pulls up the record of your contact. As soon as you answer the phone, you have all the information needed to make your next sale. Alongside analytics, cloud computing is also helping make business more effective and the consolidation of both clouds and analytics could help businesses store, interpret, and process their big data to better meet their clients' needs.

### ACKNOWLEDGMENT

We take this opportunity to thank Mrunal Pathak for her valuable guidance and for providing all the necessary facilities and her suggestions.

### REFERENCES

- [1] HadoopDB: An Architectural Hybrid of MapReduce and DBMS Technologies for Analytical Workloads. A. Abouzeid, K. Bajda-Pawlikowski, D. Abadi, A. Silberschatz, and A. Rasin.
- [2] Linticum, David. InfoWorld. <http://www.infoworld.com/d/cloud-computing/amazons-redshift-big-data-analytics-the-pros-and-cons-213049.pdf>
- [3] A Novel Technique for Information Retrieval based on Cloud Computing: [https://www.ijceronline.com/papers/vol2\\_issue5/AF025133611364.pdf](https://www.ijceronline.com/papers/vol2_issue5/AF025133611364.pdf)
- [4] IEEE, Data Retrieval TechiquesIn Cloud, ZhigangZhou Sch. of Comput. Sci. &Eng., Harbin InstofTechnol., Harbin, China HongliZhang ; Xiaojiang Du ; Panpan Li ; Xiangzhan Yu , 14-19 April 2013, Pages 2643 – 2651.
- [5] Data Analysis and Cloud Computing , Varalakshmi, P. Dept. of Inf. Technol., Anna Univ., Chennai, India Thangavel, M. ; Nithya, K. ; Priya, T. ; Sakthya, D., 18-20 Dec. 2013, Pages 413 – 418.
- [6] Toward Secure Multikeyword Top-k Retrieval over Encrypted Cloud Data. Jiadi Yu, Member, IEEE, Peng Lu, Yanmin Zhu, Member, IEEE, GuangtaoXue, Member, IEEE Computer Society, and Minglu Li, IEEE transactions on dependable and secure computing, vol. 10, no. 4, july/august 2013
- [7] SK-IR: Secured keyword based retrieval of sensor data in cloud: Sumalatha, M.R. Dept. of Inf. Technol., Anna Univ., Chennai, India Praveenraj, K. ; Selvakumar, C. , 25-27 July 2013 Pages-341 – 346.
- [8] iPEKS: Fast and Secure Cloud Data Retrieval from the Public-Key Encryption with Keyword Search Fu-KuoTseng, Dept. of Comput. Sci., Nat. Chiao-Tung

Univ., Hsinchu, Taiwan. 16-18 July 2013. Pages 452 – 458

- [9] Sheng-Lin Chou et al. “Computer Telephony Integration And Its Applications” IEEE network ,vol 3 no 1,First Quarter 2000,pp1-11.