UNDERSTANDING THE CLOUD COMPUTING ECOSYSTEM: A QUANTITATIVE ANALYSIS

Harsh Patel
Department of Computer Science,
Ranchi University, India

Prachi Shah
Department of Computer Science,
Ranchi University, India

ABSTRACT

A large number of companies are using cloud computing facilities to decrease budgets and rise the elasticity of their IT structure. This has sparked controversy among staff and investigators over the welfares and hazards of cloud computing. This study uses content analysis to test cloud computing systems. Analytical data includes high-quality investigation courses and staff-directed articles from journals and websites. We use n-gram and cluster algorithm k-methods for literature study. The input of this paper is twofold: first, it recognizes the key words and themes that are part of the cloud computing program that we have integrated with the entire model. Second, the paper reveals the sentiments of important issues expressed in both academic and educational contexts.

Keywords : Cloud Computing, Quantitative Content Analysis, Sentiment Analysis

INTRODUCTION

Over the years, cloud computing has evolved into a new computer model that aims to offer dependable, modified, advanced and robust computer services to eliminate users. In 2006, Amazon launched its fresh industry platform Amazon Web Services (AWS) and providing the foundation for this work-based learning. Cloud computing uses current machineries such as grid calculating and virtualization. However, technology used in cloud computing is still in its infancy. Also, definition features and features linked with computer compute carry on emerging over time.
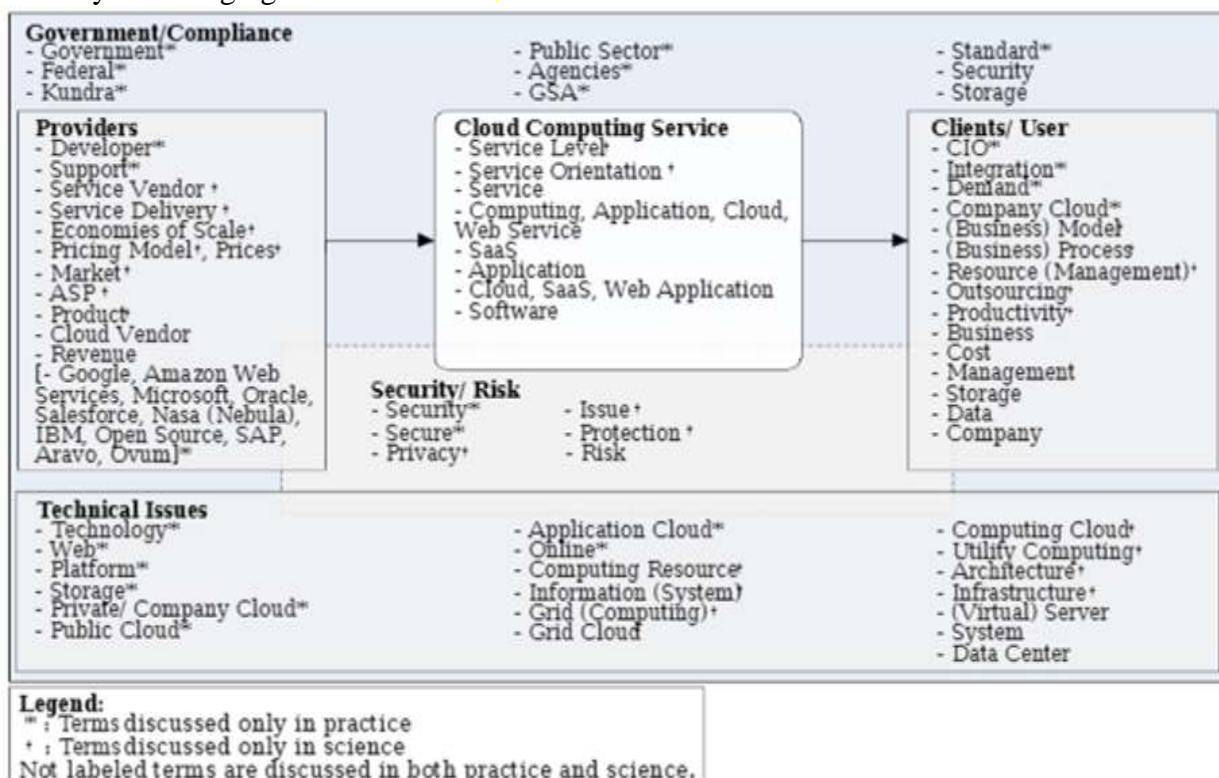

Figure 1: Cloud Computing Ecosystem

There are three major type of cloud computing services: software service (SaaS), which mentions to request services such as salesforce; As a platform service (PaaS), e. g., designer stages such as Google AppEngine; lastly, communications is a service (IaaS) that integrates primarily storage space services and computer influence services such as AWS [25,39]. The idea of cloud computing draws the concentration of teachers and students alike [18,23,25]. It attract investigators and causes from different fields (e.g., Economics vs. Technology) who move toward the subject in a variety of ways. In general, whole tendency is to increase interest in related areas such as cloud computing and IT rolling, network compute and virtualization. This view is established through the consequences of the Google Insight search results we made (cf. Figure 1). Comparing the information with others, it was done extensively on a scale of 0 and 100. every tip on the chart is separated into plurals or 100s. Since the third quarter of 2007 the numeral of look for query has been clear and the term "cloud computing" "has been steadily growing. , there has been a significant trend towards a large number of "virtualization" keyword search queries in line with the increase in cloud computing search queries in early 2010.

Figure 2: Search Queries for Cloud Computing and Related Concepts.

RELATED WORK

In view of the lack of standard computer definition, researchers have been focusing on the last few years in ahead a enhanced considerate of cloud computing and its many features. For example, Yusuf et al. Demonstrate an ontology that illuminates the applicable features of cloud computing and their relations. Investigators are too studying cloud computing with the intend of growing status of this explore topic in systematic group of people. Currently, little research is being done on motorists and performers of cloud computing systems, approval of cloud computing facilities, or the successes and risk linked with it. Instead, obtainable research has begun with the appearance of novel trade model and worth chains due to new technological advances.

Because cloud computing is very close to the technical point of view, Lemaster et al. The expanded concept includes a full sympathetic of commerce opportunity and commerce worth. They explain the emerging emergence of cloud computing from traditional IT development to computer computing. Customers and suppliers are key players in these growing cloud computing networks. From a customer perspective, Benlion discusses SaaS customer acquisition decisions based on transaction cost education.They acknowledged ecological hesitation and submission details as factors causative to SaaS acquisition. Kohler et al. acknowledged client options annotation for cloud computing services using optional integration statistics in solid search. They establish that the specific standing of a computer service contributor and the exercise of average information format were additional chief than monetary factors such as charge reductions or prices. Armbrost et al. It provides a catalog of ten cloud computing issues that are believed to affect access to the following three: business acquisition / continuity, data entry, data privacy and auditing. Although software distribution types and computer-related pricing types are believed to replace certain traditional products over time, they are not expected to be completely eliminated anytime soon.

QUANTITATIVE CONTENT ANALYSIS

Our advance to combining time frequency with aggregate analytics in the computer field. The general purpose of gratified examination is to examine, modify, and maintain a set of credentials to identify concealed skin and to collect data for ongoing work. Lizfert argues so as to satisfied examination show business an significant position in creating ideas in areas where basic education, for example, cloud computing is available. Corpus: We have selected two IT magazines, CIO Magazine and MIT Knowledge Review, as glowing as two online sheets, Silicon.com and InformationWeek.com, on the subject of the practice of always integrating the subject of cloud computing. . With the publication of online and printed goods we have tried to capture various topics. Thanks to the proven expertise of the authors we publish blog articles in our analysis and in its place rely on the skilled expertise of fortnightly and website publishing supervisor. Based on Google's search results, we selected the period from 2007 to August 2010. Although these articles tend to get a additional modest move toward to their article than peer-reviewed periodical trainings, they can dish up as helpful barometers in present presentation. And market response. On the other hand, we are systematically reviewing publications published in scientific journals and the cultural continuation of in order system. In our appraisal, we used keywords connected to cloud computing and conducted advanced and saved search of IS-based articles. Cloud computing and linked topic. We have investigated the policies of main international information systems conferences such as ICIS, ECIS, AMCIS and HICSS as fine as newspapers listed on 00 14.00 published by the Association of Information Systems (AIS) (cf Supplement: www.uwi.uos. de).) / Appendix 11.pdf). The articles for Cloud computing are listed in Table 1.

Table 1: Description of the Corpus

Publication	Publication Type	# of Articles per Year				Overall # of Articles
		2007	2008	2009	2010	
CIO Magazine	Magazine	5	5	9	11	30
MIT Technology Review	Magazine	3	8	21	4	36
Silicon.com	Internet Articles	0	38	38	16	92
InformationWeek.com	Internet Articles	6	99	133	49	287
AIS Journal Ranking with ≤14	Scientific Journals	0	1	3	1	5
ICIS, ECIS, AMCIS, HICSS	Scientific Conferences	0	5	9	5	19
Other	Cited in Scientific Articles	1	6	3	6	16
Total		15	162	217	92	485

Preliminary analysis: Prior to the commencement of the data processing process, we derivative the credentials to analyze the text and removed extra in order, such as lists of references and texts from scientific articles, to online advertising for work-related articles. In the original formulation of documents, we followed the widely accepted method of data acquisition and extraction method used by Sidorova et al. I have added the first additional operator that works mainly on the computer. One of the major evils in text investigation is presence of exploration stipulations with dissimilar spellings. Since cloud use is a growing, immature topic, this problem is significant.
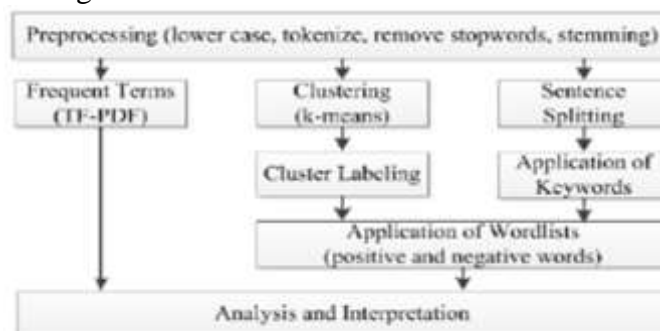
Figure 3: Process of Analysis

ANALYSIS

4.1 N-Gram Analyses

We analyze statistics from our two separate copies. They are transmitted by sound waves to a number vector. every place in the vector correspond to the word (unigram) in the body. We have selected 25 catchy names for each collection. We have seen a number of these terms used to represent current conversations in cloud computing. The consequences of Uni-gram analysis are exposed in Table 2. The top 10 to 15 names in these two lists are almost identical. However, if you look closely, there is a clear difference. In the case of traditional publishing, technological issue and marketplace actor can be seen as the the majority prominent features. Stipulations like "Technology", "Storage", "Server", "Software" and "Platform" often refer to conversations that focus on the use of cloud computing technology. Many deliberations also center on major retailers in cloud computing marketplace, such as EG. Microsoft, Google and Amazon. Additional important expression recognized in the investigation of staff booklets is safety.

Table 2: Top 25 Uni-Gram Ranked by TF-PDF

Practice		Science	
Term	TF-PDF	Term	TF-PDF
cloud	0.09990	service	0.06699
computing	0.04328	cloud	0.05800
company	0.04182	computing	0.04128
service	0.03938	customer	0.03478
application	0.03727	application	0.02930
customer	0.02521	resource	0.02587
data	0.02349	vendor	0.02354
business	0.01669	data	0.02294
software	0.01618	company	0.02017
vendor	0.01416	model	0.01992
server	0.01404	business	0.01635
system	0.01124	system	0.01506
technology	0.00858	software	0.01230
web	0.00857	management	0.01215
microsoft	0.00842	grid	0.01176
security	0.00797	server	0.01143
amazon	0.00774	cost	0.01134
google	0.00759	infrastructure	0.01019
center	0.00726	time	0.00873
infrastructure	0.00639	technology	0.00808
cost	0.00623	web	0.00800
management	0.00621	process	0.00764
platform	0.00512	information	0.00762
storage	0.00504	storage	0.00698
time	0.00496	saas	0.00594

In universal, researchers often use alike terms when discuss computer clouds. though, in its place of using physical conditions such as "server" and "storage" they like better extensions such as "resource" and "program". The word "grid" is often used because network compute is considered one of pioneers of cloud computing, and the two ideas are frequently straight compare. Also, service-related issues are common in computer-based courses, as evidenced by the recurrent employ of the conditions "service" and "sass". In addition, scientific articles on the terms "business" and "costs" suggest that they frequently talk about the property of cloud use on company. first investigate focuses on just one word. In the second step, we expanded our investigate to two grams per analysis and to each cell division. The goal is to increase a deeper considerate of compound words. Bi-gr has two consecutive names. The next consequences demonstrate subordinate TF-PDF principles than unigram analysis. This is since duplication of two identical names (e.g. "cloud_computing" and "cloud_service") is rarely compared to the same name (e.g. "Cloud").

4.2 Cluster Analyses

The major purpose of cluster investigation is to allocate individual credentials to copyright in regularly discuss topics. The three most important parameter of this algorithm are the maximum number of run and

the greatest performance category. The first limitation defines the numeral of run in the first arbitrary centroid start. The best performance rating set at 10 defines the number of repetitions made with a single run algorithm.We found the number of clusters per set in the method used for specification (cf. Equation 2) which resulted in 15 collections of corpus corpus (445 pages) and 5 collections (40 pages) of the academic corporation. Due to the nature of K- i.e. heuristic algorithm, there is little difference in the combined text. As a result, we analyzed more than a few times. Two author of this document humbly determined that the greatest impact would serve as the basis for this analysis. By submitting and discussing the results, we have maintained a series of articles related to the numeral of text they have provided. Collections are written with great reviews. To improve the label label, two authors of this paper participated in the writing of test results and independent collection collections.

Table 3: Results of the Cluster Analysis (Practice)

#	Cluster	# of Documents (Percentage)	Positive Words	Negative Words
1	General Topics	92 (20.7%)	58.9%	41.1%
2	Technical Topics	64 (14.4%)	41.1%	58.9%
3	Company Perspective (Cloud Computing)	54 (12.1%)	48.9%	51.1%
4	Company Perspective (IT Outsourcing)	37 (8.3%)	56.4%	43.6%
5	SaaS (Provider)	31 (7.0%)	65.6%	34.4%
6	Microsoft Azure	31 (7.0%)	38.1%	61.9%
7	Vendors	27 (6.1%)	56.3%	43.8%
8	SaaS (Business/ Management)	23 (5.2%)	41.1%	58.9%
9	Government	21 (4.7%)	72.9%	27.1%
10	Open Source/ Standards	20 (4.5%)	57.3%	42.7%
11	Amazon Web Services	20 (4.5%)	70.5%	29.5%
12	Security/ Risk	10 (2.2%)	54.3%	45.7%
13	Nasa Nebula	7 (1.6%)	31.0%	69.0%
14	Oracle Fusion	5 (1.1%)	56.0%	44.0%
15	IT Outsourcing/ Cloud Computing Market	3 (0.7%)	46.4%	53.6%
	Overall	445 (100.0%)	48.7%	51.3%

Table 4: Results of the Cluster Analysis (Science)

#	Cluster	# of Documents (Percentage)	Positive Words	Negative Words
1	General Topics	16 (40.0%)	42.5%	57.2%
2	Resource Management	8 (20.0%)	46.5%	53.5%
3	Grid vs. Cloud Computing	8 (20.0%)	67.2%	32.8%
4	SaaS/ IT Outsourcing	4 (10.0%)	55.1%	44.9%
5	Business/ Management	4 (10.0%)	68.8%	31.2%
	Overall	40 (100.00%)	52.4%	47.6%

4.3 Sentiment Analysis

Collection examination has help to make out key themes in computer use in the cloud, while emotional analysis reveals the initial trends of ideas in the field. However, there is still no deep understanding of positive and negative feelings. Therefore, a subsequent sentencing analysis that included multiple processing steps was performed. First, sentences are punctuated by punctuation. In these sentences we are looking for keywords that cover specific topics with cloud computing drivers. lastly, we noticeable both positive and negative words to name. Roberts believes that content analysis results should always be understood in the context of the research site in order to obtain a complete definition of a particular word. Collection label selection also requires conditional support. Therefore, in order to systematically test cloud computing, we have compiled a list of drivers and building materials based on scientific literature provided by the first scientific organization (General Elements) a number of of these topics have open deliberations concerning

cloud computing and proposed investigate agenda, divided into six input topic describe in Table 6. Also, we have additional concept that can be used as common analytical terms.

Table 5: Results of the Sentiment Analyses

Topic	Practice			Science		
	TF-PDF	Positive Words	Negative Words	TF-PDF	Positive Words	Negative Words
Technology	0.0659	54.4%	45.6%	0.0822	58.8%	41.2%
Costs	0.0186	51.7%	48.3%	0.0262	57.1%	42.9%
Personnel	0.0177	46.7%	53.3%	0.0225	49.9%	50.1%
Security	0.0143	30.2%	69.8%	0.0095	29.9%	70.1%
Quality	0.0087	53.7%	46.3%	0.0198	48.7%	51.3%
Compliance	0.0056	44.3%	55.7%	0.0049	41.6%	58.4%

DISCUSSION OF RESULTS

explore ecosystem computers from a variety of sources provides excellent information about the gap flanked by discipline and live out. For example, n-gram analysis and integration exposed a burly center on active computer provider. Naturally, the company that use it are involved in novel cloud computing facilities and goods. Well-known and long-established provider (such as AWS and Salesforce) have a good standing as the chief distributors of cloud computing. In dissimilarity, Microsoft's Azure growth stage was ruthlessly discussed (61.9% with bad names). Conversations in data centers (sometimes called clouds) indicate another difference and should be discuss. It is clear from Table 3 that doctors often talk about the theme of "there are clouds". In universal, cloud concept is divided into secluded (interior), public (external) and cross (integrated cloud types mentioned above). Scientifically speaking, this topic has not been discussed in detail. For example, Wlodarczyk et al. providing initial insights into supporting these assessments and building a solution for companies in the middle to address security issues.

CONCLUSIONS

In this paper, we examine the computer model in the cloud from a scientific and technological point of view using content analysis. The paper's contribution is twofold: first, it identify key words and topic that are fraction of the present computer-based conversation of carry out and study. We cover key main beliefs and elements in a computer program representation. This model shows the complete effect as a complete computer service procedure (see Figure 3). Second, the book stimulates the emotions of the main points that will be featured in the articles in both copies. Here, the biggest quest is for what the clouds look like in general. There are just a few things that test how bad shops and educators are. The results of emotional analysis differ between work and discipline. It is significant to retain in attention that this investigate method has its confines. though, we have try to reduce discrimination by next a well-established investigate move toward. We believe that our collection offers the highest height of excellence and deserves the division between work and discipline.

REFERENCES

1) Benlian, A., 2009. A transaction cost theoretical analysis of software-as-a-service (SAAS)-based sourcing in SMBs and enterprises.
2) Huang, K.W. and Wang, M., 2009. Firm-level productivity analysis for software as a service companies. ICIS 2009 proceedings, p.21.
3) MacQueen, J., 1967, June. Some methods for classification and analysis of multivariate observations. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability (Vol. 1, No. 14, pp. 281-297).
4) Martens, B., Poeppelbuss, J. and Teuteberg, F., 2011. Understanding the Cloud Computing Ecosystem: Results from a Quantitative Content Analysis. Wirtschaftsinformatik, 16, p.2011.