



Analysis of Big Data in Social Network Marketing and Social Media

Shahida B^a

^aCSE department, PDIT, Hospet, India

DOI: <https://doi.org/10.55248/genmpi.2022.3.9.26>

ABSTRACT

Data is the foundation of every system. Data keep scaling at different types. The creation of an automated system in the field of digital media deals with the enormous amounts of data that accumulate over time on www. Social networks were used to update it globally every second. Social networks offer the chance to engage with others in a personal, professional, and entertaining way. This study tries to show how social networking and media recommendation use a tremendous quantity of data. There are many different social networking websites on www. The primary goal of this study is to examine how individuals interact with social media and how TV media use social media to rank channels and shows. Big data, big data analytics, social networks analytics, TV media social networks, emotional analysis. Every economy, production, organization, business function, and individual now depend on data in some way. The use of the internet, smartphones, social networks, the development of ubiquitous computing, and several other technical breakthroughs are all contributing to the world's expanding data volume. Cost savings, quicker, better decision-making, and the development of new goods and services are all benefits of big data. However, the three V's of big data—Velocity, Volume, and Variety—amplify security and privacy vulnerabilities. These variables include things like massive cloud infrastructures, variety in data sources and formats, acquisition of data occurring in streams, and a rising amount of inter-cloud migrations. We have examined how the Bloom filter (BF) can be used to overcome problems with sorting and decomposing huge amounts of data both in space and time.

Keywords: NoSQL database, MongoDB, CouchDB, Column Database, Key-value Database, Big Data, Big Data Analytics, RDBMS, and NoSQL

1. Introduction

Big data refers to a collection of data sets that are so numerous and intricate that they become challenging to process using typical data processing software or readily available database management tools. Capture, curation, storage, search, sharing, transfer, analysis, and visualization are among the difficulties [1]. Terabytes of data from a variety of information sources, including social media networks, research instruments, mobile devices, sensor technology, and networks, can be updated every second. For instance, assume a Market analysts continuously monitor the interactions and transactions of online users to develop new business trends. Big data analytics is the act of poring over vast quantities of data (big data) of various kinds to find undiscovered patterns, unknown relationships, and other valuable information. Such knowledge can give an organization a competitive edge over its rivals and have a positive impact on the company's bottom line through improved marketing and increased sales [2][18]. The software tools frequently utilized in advanced analytics disciplines like predictive analytics and data mining can be employed for big data analytics. However, traditional data warehouses might not be able to accommodate the unstructured data sources utilized for big data analytics. Big data analytics uses technologies like NoSQL databases, Hadoop, and MapReduce. Not Only SQL is another name for a NoSQL database. An alternative to the tabular relations used in relational databases, a NoSQL database offers a framework for the store and retrieval of data [3]. A software framework called MapReduce enables programmers to create applications that handle enormous volumes of unstructured data concurrently across standalone machines or distributed clusters of processors. For indexing Web pages, Google built it. Large data sets can be processed in a distributed computing environment using Hadoop, a free Java-based programming platform. It is a component of the Apache project, which the Apache Software Foundation sponsors. Applications can be executed on

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000.
E-mail address: shahida@pdit.ac.in

systems with thousands of nodes and thousands of terabytes of storage thanks to Hadoop [18]. Its distributed file system enables quick data transfers between nodes and enables the system to keep running even in the event of a node failure. This method reduces the possibility of a catastrophic system failure, even if a sizable portion of the nodes fail.

2. Review of Literature

Social Network Marketing and Media Analysis

Everything is interconnected, including people, information, events, and locations [4] [5]. This is especially true in the age of online social media. In virtual online communities and networks, interaction between people that involves information creation and/or exchange is referred to as social networking. It is crucial to understand the distinctive advantages of each person because this has become a fundamental component of all business. People are now using social networks to communicate with friends, family, and other types of people about their thoughts, feelings, and opinions. Their followers read the updates posted online by many famous persons and significant personalities [17][18]. More followers can be gained by comments and likes, but if anything unpleasant is said about someone, it breeds animosity for that person. Social networking site, or a user friend. It serves as a discussion forum for people to debate whether their opinions, thoughts, or other ideas are valid or unreliable while also providing an opportunity to spread information, increase knowledge, and discover novelties through visual and auditory media [20][21]. It is a many-to-many conversation rather than a one-to-one exchange that can be had with numerous people who have joined this link. There are many social media networking sites, including Facebook, Twitter, and LinkedIn. Even though TV media has a different blog for each station, they are nevertheless connected to social networks. It serves as an interface for communication with a wide range of people. Online social networking brings people together [6] [7] [8]. On YouTube, advertisements and lengthy TV shows are continuously viewed and even commented on. It benefits the populace by generating feedback, ideas, and increased popularity for their programs. The programs are rated by the viewers; if they give them high marks, the channel gets a higher rating; otherwise, the programs' issues need to be fixed. It offers the chance to learn about the channel program ratings of competitors [9]. It serves as a forum for people to easily share their likes and dislikes and learn about the perspectives of the general public.

For marketers, this presents an opportunity to better comprehend, target, approach, and sell to customers, prospects, and suspects. It also presents an opportunity to innovate, set oneself apart from the competition, and create a competitive edge. Products and tweets that highlight a brand's popularity draw customers to it and help manufacturers promote sales [10]. By examining the social ties, the business will undoubtedly find that the individuals are partially working outside of the boundaries of their department or group, circumventing the current organizational structure. The company may be able to increase productivity by redesigning their internal organizational structure to match the connections they have made on social networks by using the connection information. If you distribute your blog post by email, Twitter, Facebook, or another social network, be sure your company name is available on those networks first. For seeing the information and tracking user visitor activity, we may follow the content and evaluate the blog through a number of channels (Email, Facebook, Google Plus, LinkedIn, and Twitter) [11]. Additionally, it offers the choice of looking up competition rankings.

Because they are major determinants of our behaviors, opinions play a central role in almost all aspects of human activity [12]. We seek out other people's perspectives whenever we need to make a decision. In the real world, companies and organizations are constantly interested in the public's or consumers' perceptions of their goods and services. Before making a purchase decision on a product and before casting their vote in a political election, individual customers want to hear what other people think about the product's current users as well as the political candidates. When someone needed advice in the past, they typically turned to their friends and relatives. Surveys, opinion polls, and focus groups were conducted by organizations and businesses when they needed the opinions of the general public or customers [14]. It has long been a significant business for marketing, public relations, and political campaign firms to gather consumer and public opinion. People and organizations are using the content in social media—such as reviews, forum discussions, blogs, microblogs, Twitter, comments, and postings on social networking sites—for decision-making more frequently now than ever before [2][15].

Enterprises should think about what data to gather and how to use the collected data to better their internal operations as the organization has a much smaller number of internal social network users than the public social networks like Facebook [18]. Business information can be obtained via the analysis of internal social networks. Google Analytics is a fantastic tool for tracking the effectiveness of your blog, but you might not use it enough. Setting up a weekly report for email distribution to you is one approach to get around this.

3. Discussion

Big data refers to statistical units that can be so large or intricate that conventional statistical processing software is unable to handle them. The challenges include data collection, evaluation, curation, storage, switching, querying, updating, and privacy concerns. The phrase "big facts" is frequently used to refer specifically to the usage of advanced information analytics techniques that extract value from data, such as predictive analytics, person behavior analytics, and others, rather than to a particular size of data set [16]. "There is no question that the amounts of data that are currently available are enormous, but it's no longer the most applicable function of this new information environment." In order to "identify company patterns, save you from illnesses, and fight crime, among other things," analysis of fact sets can uncover new correlations. In fields such as Internet search, finance, city informatics, and commercial company informatics, scientists, business executives, practitioners of medicine, marketers, and governments frequently encounter issues with large information devices. In eScience art, which includes meteorology, genomics, connectomes, challenging physics simulations,

biology, and environmental research, scientists bridge barriers [8]. The majority of computations in distributed systems only have one level of security, which is not necessarily recommended. Since NoSQL databases are constantly changing, it is difficult for security solutions to keep up with demand [18]. Automated data transfer necessitates additional security measures, many of which may not be available. This exercise doesn't always get up; when a machine receives a huge amount of data, it needs to be checked to remain honest and accurate. However, unethical IT specialists engaged in records mining can compile private data without alerting or asking users for their consent. Access controls for connections protection and encryption may become outdated and unusable for IT professionals who rely on them. Some firms are unable to implement access controls to split the level of secrecy within the company or choose not to do so. Big Data does not automatically undergo recommended favorable audits due to the sheer volume of statistics involved. Big Data's scale prevents constant monitoring and tracking of its origins [9].

4. Summary

As a result, the survey paper presents an analysis of big data in social networking and media interaction from various angles. Every second, it can undergo a significant update in the form of big data. Ultimately, it was determined that social networks had both benefits and drawbacks. This document provides an overview of the methods employed for managing and safeguarding enormous amounts of data. We have carried out research analysis related to strategies and measures for big data security and coping with, and we are providing evaluation of these papers. In those papers, we find that BF may be used to overcome inefficiencies in both the ordering and decomposition of large amounts of data. Experts in cloud computing agree that the ongoing expansion of the antivirus provider is the most cost-effective solution to improve the security of Big Data. A wider range of antivirus providers, each offering a variety of solutions, provides stronger defense against threats to Big Data security. It's encouraging that the antivirus company is frequently praised for being transparent. Antivirus software application providers freely share information about current Big Data protection dangers, and company executives frequently collaborate to combat fresh malware attacks, resulting in the greatest possible improvements in Big Data security.

REFERENCES

- [1] Abramova, V., & Bernardino, J. (2013, July). NoSQL databases: MongoDB vs Cassandra. In Proceedings of the international C* conference on computer science and software engineering (pp. 14-22).
- [2] Ali, W., Shafique, M. U., Majeed, M. A., & Raza, A. (2019). Comparison between SQL and NoSQL Databases and Their Relationship with Big Data Analytics. *Asian Journal of Research in Computer Science*, 4(2), 1-10
- [3] Becker, M. Y., & Sewell, P. (2004, June). Cassandra: Flexible trust management, applied to electronic health records. In Proceedings. 17th IEEE Computer Security Foundations Workshop, 2004. (pp. 139-154). IEEE.
- [4] Berg, K. L., Seymour, T., & Goel, R. (2013). History of databases. *International Journal of Management & Information Systems (IJMIS)*, 17(1), 29-36.
- [5] Bjeladinovic, S., Marjanovic, Z., & Babarogic, S. (2020). A proposal of architecture for integration and uniform use of hybrid SQL/NoSQL database components. *Journal of Systems and Software*, 168, 110633.
- [6] Chandra, D. G. (2015). BASE analysis of NoSQL database. *Future Generation Computer Systems*, 52, 13-21.
- [7] Chen, J. K., & Lee, W. Z. (2019). An introduction of NoSQL databases based on their categories and application industries. *Algorithms*, 12(5), 106.
- [8] Cuzzocrea, A., & Shahriar, H. (2017, December). Data masking techniques for NoSQL database security: A systematic review. In 2017 IEEE International Conference on Big Data (Big Data) (pp. 4467-4473). IEEE.
- [9] de Oliveira, V. F., Pessoa, M. A. D. O., Junqueira, F., & Miyagi, P. E. (2021). SQL and NoSQL Databases in the Context of Industry 4.0. *Machines*, 10(1), 20.
- [10] Deka, G. C. (2013). A survey of cloud database systems. *It Professional*, 16(2), 50-57. IEEE.
- [11] Di Martino, S., Fiadone, L., Peron, A., Riccabone, A., & Vitale, V. N. (2019, June). Industrial Internet of Things: Persistence for Time Series with NoSQL Databases. In 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE) (pp. 340-345). IEEE.
- [12] dos Santos Ferreira, G., Calil, A., & dos Santos Mello, R. (2013, December). On providing DDL support for a relational layer over a document NoSQL database. In Proceedings of International Conference on Information Integration and Web-based Applications & Services (pp. 125-132).
- [13] Gessert, F., Wingerath, W., Friedrich, S., & Ritter, N. (2017). NoSQL database systems: a survey and decision guidance. *Computer Science-Research and Development*, 32(3), 353-365.
- [14] Guimaraes, V., Hondo, F., Almeida, R., Vera, H., Holanda, M., Araujo, A., ... & Lifschitz, S. (2015, November). A study of genomic data provenance in NoSQL document-oriented database systems. In 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 1525-1531). IEEE.
- [15] Rodriguez, K. M., Reddy, R. S., Barreiros, A. Q., & Zehtab, M. (2012, June). Optimizing Program Operations: Creating a Web-Based Application to Assign and Monitor Patient Outcomes, Educator Productivity and Service Reimbursement. In DIABETES (Vol. 61, pp. A631-A631). 1701 N BEAUREGARD ST, ALEXANDRIA, VA 22311-1717 USA: AMER DIABETES ASSOC.

-
- [16] Kwon, D., Reddy, R., & Reis, I. M. (2021). ABCMETAapp: R shiny application for simulation-based estimation of mean and standard deviation for meta-analysis via approximate Bayesian computation. *Research synthesis methods*, 12(6), 842–848. <https://doi.org/10.1002/jrsm.1505>
- [17] Reddy, H. B. S., Reddy, R. R. S., Jonnalagadda, R., Singh, P., & Gogineni, A. (2022). Usability Evaluation of an Unpopular Restaurant Recommender Web Application Zomato. *Asian Journal of Research in Computer Science*, 13(4), 12-33.
- [18] Reddy, H. B. S., Reddy, R. R. S., Jonnalagadda, R., Singh, P., & Gogineni, A. (2022). Analysis of the Unexplored Security Issues Common to All Types of NoSQL Databases. *Asian Journal of Research in Computer Science*, 14(1), 1-12.
- [19] Singh, P., Williams, K., Jonnalagadda, R., Gogineni, A., & Reddy, R. R. (2022). International students: What's missing and what matters. *Open Journal of Social Sciences*, 10(02),
- [20] Jonnalagadda, R., Singh, P., Gogineni, A., Reddy, R. R., & Reddy, H. B. (2022). Developing, implementing and evaluating training for online graduate teaching assistants based on Addie Model. *Asian Journal of Education and Social Studies*, 1-10.
- [21] Sarmiento, J. M., Gogineni, A., Bernstein, J. N., Lee, C., Lineen, E. B., Pust, G. D., & Byers, P. M. (2020). Alcohol/illicit substance use in fatal motorcycle crashes. *Journal of surgical research*, 256, 243-250.
- [22] Brown, M. E., Rizzuto, T., & Singh, P. (2019). Strategic compatibility, collaboration and collective impact for community change. *Leadership & Organization Development Journal*.
- [23] Sprague-Jones, J., Singh, P., Rousseau, M., Counts, J., & Firman, C. (2020). The Protective Factors Survey: Establishing validity and reliability of a self-report measure of protective factors against child maltreatment. *Children and Youth Services Review*, 111, 104868