

Chapter 3

Analysis Report for Statistics in the Twitter Network

Parvathi R.

Vellore Institute of Technology, Chennai, India

Yamani Sai Asish

Vellore Institute of Technology, Chennai, India

Pattabiraman V.

 <https://orcid.org/0000-0001-8734-2203>

Vellore Institute of Technology, Chennai, India

ABSTRACT

Twitter is the most popular social networking service across the world. In Twitter, the messages are known as tweets. Tweets are mainly text-based posts that can be up to 140 characters long which can reach the author's subscribers. These subscribers are also known as followers. Such subscriptions form a direct connection. But these connections are not always symmetric. In this study, the authors have assumed that if two nodes are connected, then the tweet is propagated between them without any other conditions. But using sentiment analysis, the general opinion of people about various things can be figured. The Twitter data set analyzed includes almost 20k nodes and 33k edges, where the visualization is done with software called Gephi. Later a deep dive analysis is done by calculating some of the metrics such as degree centrality and closeness centrality for the obtained Twitter network. Using this analysis, it is easy to find the influencers in the Twitter network and also the various groups involved in the network.

INTRODUCTION

In the past few years people have faced a lot of problems due to the propagation of wrong information in the twitter network. A lot of problems are also faced because certain vital information is not being reached to people in time due to the modularity in the twitter network. Many of these problems can be solved just by analyzing the twitter network and drawing useful conclusions from it. In this analysis, tweet

DOI: 10.4018/978-1-7998-7728-8.ch003

Analysis Report for Statistics in the Twitter Network

propagation on Twitter social network with 23370 nodes and 33101 edges was analysed. Specific focus on the network influences and how one can use the knowledge about the network for various purposes is been studied through various analysis.

Twitter is the most popular social networking service across the whole world. In Twitter the message passing is known as tweets. Tweets are mainly text-based posts which can be up to 140 characters long which can reach the author's subscribers. These subscribers are also known as followers. Such subscriptions form a direct connection. But these connections are not always symmetric. In this study, we have assumed that if two nodes are connected, then the tweet is propagated between them without any other conditions. But using sentiment analysis the general opinion of people about various things can be figured. The Twitter data set analyzed include almost 20k nodes and 33k edges, where the visualization is done with software called Gephi. Later a dive deep analysis is done by calculating some of the metrics such as degree centrality and closeness centrality for the obtained twitter network. Using this analysis, it is easy to find the influencers in the twitter network and also the various groups involved in the network.

LITERATURE SURVEY

Petrovic (Petrovic et al., 2011) discussed how the tweets are propagated using the Twitter network. They analyzed three categories a) By following users b) By tweeting c) By re-tweeting. The authors argued that the re-tweeting is potentially the best, and proves why it is so by taking examples and by proving using algorithms. And he concludes as follows. The most important task in understanding retweeting is to say if a particular tweet is going to be Re-tweeted or not, which is the main motive of this work. Analysis was performed on humans which has shown a positive result. Machine Learning approach based on passive-aggressive algorithm have substantially outperformed the previous known baseline.

Carlos (Castillo et al., 2011) analyzed the Information credibility on Twitter by analysing the credibility of the twitter information is and how humans react to it.

Information Propagation(Speed and its Patterns) German Tweets has shown that the information travels through a certain pattern in the network. It has also proves that the friend relationships also significantly influence the information propagation speed on the Twitter network. Two approaches was used to evaluate technically as well as based on how a certain topic matches a pattern and how prominent the friends are compared to the rest of the users (Tareaf, et al., 2018).

Twitter presence of (HSS) journal, Titles which are obtained from mainstream citation indices, by analysing how the interaction and communication patterns in them (Raamkumar, et al., 2018). Their research utilizes webometric collection, descriptive analysis, and also the social network analysis. So the results indicate the presence of the HSS journals on Twitter. Sharing of general websites is seen as the key activity that is performed by (HSS) journals on Twitter. Among these the web content from the news portals and magazines are highly different. Sharing of articles and retweeting isn't the most common thing to be observed. Inter-journal communication is also witnessed, But it is less with journals from other indexes.

Shaoshi et al (Ye & Wu, 2010) have evaluated different ways of social influences by measuring their assessments, stabilities and their correlations. This study tells us about all the problems that were faced. The results have given important insights for the future OSN development and research. Bo et al (Wu & Shen, 2015) have used trace-driven experiments, where the data are validated and compared through different prediction model. This proposed model has better performance with respect to that of the re-

7 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/analysis-report-for-statistics-in-the-twitter-network/284202

Related Content

Fine-Grained Deep Feature Expansion Framework for Fashion Apparel Classification Using Transfer Learning

R. Regin, Pravin Kumar Sharma, Kunnal Singh, Y. V. Narendra, S. Rubin Bose and S. Suman Rajest (2024). *Advanced Applications of Generative AI and Natural Language Processing Models* (pp. 389-404). www.irma-international.org/chapter/fine-grained-deep-feature-expansion-framework-for-fashion-apparel-classification-using-transfer-learning/335848

Deep Learning Network: Deep Neural Networks

Bhanu Chander (2020). *Neural Networks for Natural Language Processing* (pp. 1-30). www.irma-international.org/chapter/deep-learning-network/245081

Named Entity Recognition (NER) in Low Resource Languages of Ho

Satya Ranjan Dash, Bikram Biruli, Yasobanta Das, Prosper Abel Mjimwa, Muhammed Abdur Rahmaan Kamaldeen and Aloka Fernando (2024). *Empowering Low-Resource Languages With NLP Solutions* (pp. 157-182). www.irma-international.org/chapter/named-entity-recognition-ner-in-low-resource-languages-of-ho/340505

Improving the K-Means Clustering Algorithm Oriented to Big Data Environments

Joaquín Pérez Ortega, Nelva Nely Almanza Ortega, Andrea Vega Villalobos, Marco A. Aguirre L., Crispín Zavala Díaz, Javier Ortiz Hernandez and Antonio Hernández Gómez (2021). *Handbook of Research on Natural Language Processing and Smart Service Systems* (pp. 289-308). www.irma-international.org/chapter/improving-the-k-means-clustering-algorithm-oriented-to-big-data-environments/263107

Estimating Importance From Web Reviews Through Textual Description and Metrics Extraction

Roney Lira de Sales Santos, Carlos Augusto de Sa, Rogerio Figueredo de Sousa, Rafael Torres Anchiêta, Ricardo de Andrade Lira Rabelo and Raimundo Santos Moura (2021). *Natural Language Processing for Global and Local Business* (pp. 154-179). www.irma-international.org/chapter/estimating-importance-from-web-reviews-through-textual-description-and-metrics-extraction/259788