Survey on Social Network Analysis

Manali Mahadev Wadekar, Prashant Itankar, Ankit Sanghavi

M. E Student, Dept. of Computer Engineering, ARMIET, Shahapur, Mumbai University, India

Assistant Professor, Dept. of Computer Engineering, DMCE, Airoli, Mumbai University, India

Head of Department, Dept of Computer Engineering, ARMIET, Shahapur, Mumbai University, India

ABSTRACT: Earlier during the times of conventional print media, there used to be one-way information dissemination which was limited to geological boundaries having restricted span and reach. With the beginning of online social media, the process of information sharing has changed significantly. It is the rapid medium for communication purpose. Online Social Networks like Facebook, Twitter have revolutionized the interpersonal communication by providing a platform for folks to communicate themselves at a worldwide , beyond their immediate geography. Many research in this field are focused on analysis of diffusion of information. Our aim is to study diffusion dynamics of specific real world events, discussed on Twitter.
The diffusion phenomenon has a great impact on Online Social Networks (OSNs). Gathering dissemination data over these huge networks encounters many challenges which can be alleviated by adopting a suitable sampling approach. The contribution of this paper is to analyse the diffusion phenomenon on twitter .

KEYWORDS: Information diffusion analysis, social network, Link diffusion.

I. INTRODUCTION

Information diffusion as a new era of research has a incredible effect on social networks [1]. In recent years, large Online Social Networks such as Facebook, Twitter and LinkedIn have been the source of information dissemination in diverse formats such as posts, tweets, and videos. The development of networks results in large information networks. Twitter provides an open online data-sharing stage that allows average people to transmit their real-time information at an unprecedented scale. For example, every second on average about 6,000 tweets are tweeted on twitter. Which corresponds to over 350,000 tweets sent per minute,500 millions tweets per day [2]. Therefore, collecting the diffusion data over large amount OSNs is frequently infeasible in many applications. This challenges the need for manipulating the diffusion data in an efficient means to examine the diffusion process performance. This paper surveys diffusion phenomena, triad formation and diffusion strength of different links.

II.APPROACHES FOR DIFFUSION ANALYSIS

Hamza Loucif, Abdelhak Boubetra, Samir Akrouf work on "An Experimental Study of the Diffusion Process in the Twitter Platform Using the ICM, LTM, and Bass Models"[1]. They present the overview of Linear Threshold Model (LTM), the Independent Cascade Model (ICM), and the Weighted Cascade Model (WCM) a variant of ICM. The main findings of this study reveal that we can propose a new model as Following Cascade Model. An initial set of nodes to start the diffusion process is the important factor that decides the success or failure of the diffusion.

The author in [2] discuss the spatial and temporal diffusion analysis of events using retweet interaction networks between users. Therefore, spatio-temporal analysis of distribution of event related tweets on Twitter gives us an insight into the characteristics exhibited by different event discussion phenomenon in social media networks.

[3] "Maximizing the Spread of Influence through a Social Network" propose that approximation algorithm can be used to influence the maximization. From approximation algorithm you will learn how to apply the maximum likelihood estimation (MLE) approach to solve the truth discovery problem on Twitter and analyze the integrity of tweets from a real-world case study.

Erick Stattner, Reynald Eugenie, Martine Collard, May 13-15 2015 Athens [4] in this paper we Understand how the information spread through the network. More particularly, identify in which conditions a person forwards information. One of the key challenges is to understand how information diffuse through social networks and what can positively or negatively affect it.

There are both unsupervised and supervised methods for link prediction. Liben-Nowell and Kleinberg [5] survey the unsupervised methods, including preferential attachment , random walk with restart. Unsupervised learning approach provides the various methods for link prediction based on node neighbourhood, ensemble of all paths and other higher level approaches. Finally, link prediction mainly focuses on predicting whether a link will be formed or not without caring about how links are diffused under certain effects, so we can study diffusion mechanism of links.

Dong Wang, Lance Kaplan, Hieu Le, Tarek Abdelzaher [6] proposed a paper describing a maximum likelihood estimation approach to accurately discover the truth in social sensing applications. Build a MLE based tweet credibility analysis tool using Expectation Maximization to analyze the credibility of reported tweets.

Learning to predict reciprocity and triadic closure in social networks, T. Lou, J. Tang, J. E. Hopcroft, Z. Fang, and X. Ding [8] authors proposed a how one way and two way link form in social network. Paper gives unsupervised as well as semisupervised learning approach to predicting reciprocity and triadic closure in social networks and a learning framework to formulate the problems of predicting reciprocity and triadic closure into a graphical model.

Supervised Random Walks: Predicting and Recommending Links in Social Networks Lars Backstrom edu [9] propose Supervised Random Walks, an innovative approach for link prediction and link recommendation by using learning algorithm. By utilizing node and edge attribute information our system guides the random walks towards the desired target nodes & to estimate the strength of social links. Supervised Random Walks are inadequate to link prediction, and can be applied to many other problems that require learning to rank nodes in a graph.

III.PROPOSED SYSTEM

We address the problem of the diffusion of information by using Twitter. Twitter is a good case study: Wide variety of topics, 310 millions monthly active users (business, professionals, politicians, etc.) 500 millions of messages posted each day. So we are going to analyse twitter network with its diffusion phenomena, formation of different links.
 We propose "following" link cascade model and learn diffusion strength in different triadic structure by maximizing an objective function based on the "following" link cascade model. We consider two specific influence maximization applications, follower maximization and followee maximization to reveal usefulness of the proposed model.
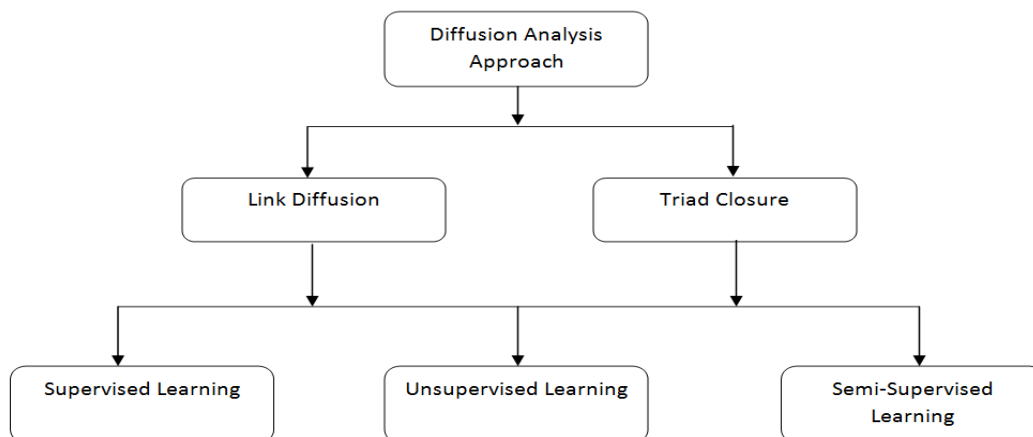

Fig. 1 Diffusion Analysis Techniques

So, I propose a diffusion model using Diffusion Techniques shown in fig1. Which uses Supervised, Unsupervised and Semi-Supervised algorithms to enhance the follower and followee maximization. This Proposed Model will try to give the best results as far as possible and provide better performance to the netizen.

IV.CONCLUSION AND FUTURE WORK

Most of the people in recent days use social network to express their views and opinions, share their emotions. So the social media is the place where huge amount of information is published. As a result social media has become an important source of people's opinions. In this paper we studied the diffusion mechanism of links in micro blogging networks. we demonstrate that "following" links propagate according to the triadic structures with different diffusion strength.

For future work, it is intriguing to study other triadic structures in addition to already existing. Designing and implementing randomized controlled experiments would also be an important direction to validate the casual relationship in the formation of links

V.ACKNOWLEDGEMENT

I owe a great many thanks to all who helped and supported me during development of this paper. I sincerely thank my project guide Prof. Prashant Itankar sir for his expert guidance and valuable support throughout the course of this project work. I also acknowledge with deep sense of gratitude and inspiration received from my head of computer department and principle of my college.

REFRENCES

1. Hamza Loucif, Abdelhak Boubetra , Samir Akrouf, "An Experimental Study of the Diffusion Process in the Twitter Platform Using the ICM, LTM, and Bass Models", ResearchGate Conference paper December 2014.
2. Purva Pruthi, Anu Yadav, Farheen Abbasi and Durga Toshniwal, "How has Twitter changed the Event Discussion Scenario?: A Spatio-Temporal Diffusion Analysis" , 2015 IEEE International Congress on Big Data, 978-1-4673-7278-7/15.
3. D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network". Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'03), p.137–146, (2003).
4. Erick Stattner, Reynald Eugenie and Martine Collard, "How do we Spread on Twitter?", May 13-15 2015. Athens, Greece.
5. D. Liben-Nowell and J. M. Kleinberg, "The link-prediction problem for social networks" JASIST, 58(7):1019–1031, 2007.
6. Dong Wang, Lance Kaplan, Hieu Le, Tarek Abdelzaher "On Truth Discovery in Social Sensing: A Maximum Likelihood Estimation Approach" ,IPSN12 Research Paper, ACM 978-1-4503-1227-1/12/04.
7. .W. Kermack, A. McKendrick, "A contribution to the mathematical theory of epidemics. Proc. Roy. Soc.", A 115:700–721, 1927.
8. T. Lou, J. Tang, J. E. Hopcroft, Z. Fang, and X. Ding, " Learning to predict reciprocity and triadic closure in social networks", TKDD, 7(2), 2013.
9. Lars Backstrom, Jure Leskovec, "Supervised Random Walks: Predicting and Recommending Links in Social Networks" WSDM 978-1-4503-0493-1/11/02 ,February 9–12, 2011,
10. W. Chen, W. Lu, and N. Zhang, "Time-critical influence maximization in social networks with time-delayed diffusion process", In AAAI'12, 2012.
11. M. Kimura, K. Saito, K. Ohara, and H. Motoda, " Learning information diffusion model in a social network for predicting influence of nodes", Intell. Data Anal., 15:633–652, 2011.
12. M. Y. Li, J. R. Craef, L. C. Wang, and J. Karsai, "Global synamics of an seir model with a varying total population size",Math. Biosci., 160:191–213, 1999.
13. D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In KDD'03, pages 137–146,2003

BIOGRAPHY

Ms. Manali Mahadev Wadekar has received B.E. in Computer Engineering from Mumbai University, Mumbai, India in 2011. Currently she is pursuing M.E. in Computer Engineering from Mumbai University, Mumbai, India.