

Pseudo Activity Development for Effective Information Diffusion in Online Social Network

Reena Pagare
Research Scholar
Pacific University
Udaipur India

Akhil Khare, PhD
Professor
MVSR Engineering College,
Osmania University, India

ABSTRACT

Online social networks are dynamic social interaction platforms for billions of users worldwide. Information and ideas are rapidly disseminated among these users through online social interactions. The online interactions among online social network users generate a huge volume of data that provides the opportunity to study human behavioral patterns. A social network is generally assumed to be viewed as graphs, where vertices denote users and edges represent relationships among users. The importance of a user in a network can be calculated by using metrics imported from graph theory. Social influence is recognized as a key factor that governs human behavior. It indicates the attitude of certain individuals to be affected by other subjects' actions and decisions. A social network, in reality, is evolving dynamically and continuously. Such evolution is coupled with the spread of information on top of the network: the network topology affects the channels of information diffusion; the birth and death of connections in the network are, in turn, triggered by the traffic. As Information spread on networks, it can be observed that the cascades which ensue as agents can get infected. The diffusion process is affected by both the actions of agents and the underlying network structure. This paper focuses on existing research in this domain and possible research methodology for ongoing research.

General Terms

Information diffusion, social network, Influence maximization

Keywords

Trend detection, online social network, network topology, machine learning, graph theory, influential nodes

1. INTRODUCTION

Online social networks play a significant role in the propagation of information among the society. Plenty of attempts are already made to understand this trend, including well-liked topic recognition to information diffusion modeling, as well as important spreaders recognition [1]. One of the primary tasks when understanding information diffusion is to build intelligent methods to offer a global view of the topics which are popular with time or can become popular, and also animate the community. This requires extracting useful information, to sum up, conversations, promoting popular topics to end users, or forecasting upcoming popular topics [2]. Conventional topic detection methods designed to evaluate static corpora are certainly not adapted to information streams produced by OSNs [3]. To be able to proficiently detect topics in textual streams, it's been recommended to spotlight bursts.

The work in information diffusion has four main components the actors or users, content, network structure and the diffusion Process. (Refer Fig 1)

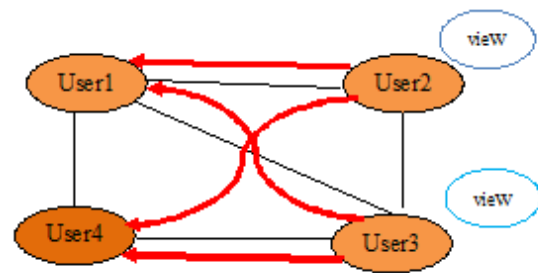


Fig 1: The study of information diffusion incorporates four important factors: Actors/Users, Content (shown as view), Structure (Graph of users as nodes and relationships as black edges) and diffusion (shown in red color lines)

In a network as in Fig 1 when a decision of an actor here user1 is influenced by decision of his friend or colleague (user3), we say the user1 is influenced by user3

Hence, information diffusion includes: popular topic detection, understanding information diffusion, Influential spreaders identification. This model will be elaborated for future development.

2. IDENTIFICATION OF INFLUENTIAL USER USING MACHINE LEARNING

All User impact is calculated dependent on numerous components and by making use of several approaches. This section is targeted on state-of-the-art strategies on the impact measurement and prognosis of influential end users within the OSN framework [4]. The category of such recognition approaches is dependent on their particular functioning guidelines.

Learning methods make use of machine learning algorithms to estimate influential end users and the most frequent which is supervised learning. Naïve Bayes [5], support vector machine [6], and also decision tree is the most typical supervised powerful set of capabilities that may provide the significant discriminative capacity to superior anticipate results. Learning solutions need adequate datasets in training and assessment of the machine learning model [7].

In another research, numerous capabilities are extracted to train a support vector machine (SVM). Characteristics utilize three different ways of aggregation, such as score and SVM-based aggregation. Another method fuses end user position in a network, individual opinion polarity, and tweet quality to acquire a merged impact score [8].

Additionally, a logistic regression evaluation was given to determine important features for forecasting individual influence. This kind of features had been employed to train and also evaluate 4 machine learning algorithms. The ACQR composition extracts some characteristics that are thought to

be discriminatory features in discovering productive end users in OSNs [9]. Characteristics are produced by 4 elements, particularly, activeness, centrality, quality of write-up, and popularity. The 4 characteristics are consequently employed to train the SVM.

For reference, author developed a method of inferring the model parameters (refer fig 2) for a given network and the tract of a single contagion. Author fitted the model to each contagion separately. Algorithm calculates the network and the infection times for each node that got infected with the contagion under consideration [12].

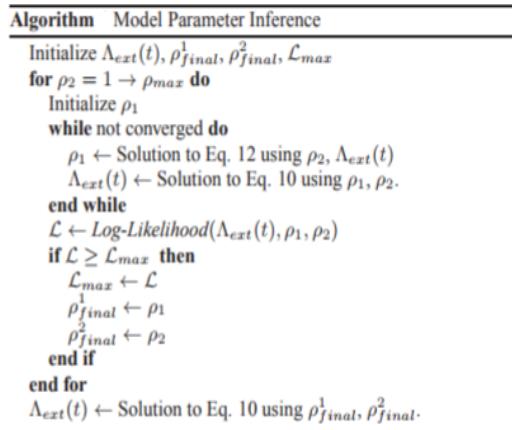


Fig 2: Reference MPI Algorithm [12]

A diffusion route of a graphic encodes the collection of end users exactly who propagated it. The group of all feasible diffusion routes depends upon both the end users in the social network along with the system framework. Author mapped specific end users to a preset number of canonical prototype end users. The unavailability of comprehensive system data differentiates online social networks from numerous intricate systems and inhibits the immediate validation and evaluations of the efficiencies of end user impact computing approaches.

3. IMPACT OF NETWORK TOPOLOGY

It is crucial to investigate the way the network topology influences the distribution of information and the way the targeted traffic movement, consequently, patterns the network progression. It is identified that the combination of these kinds of rivalry and social network topology on its own can describe the noticed worldwide behavior without perhaps presuming the several implicit values amongst suggestions.

Based on the connection between the social networking and also the topic area, most of us assign each and every end user or information together with a range of topic and evaluate relevant diversity. The author discovered that higher topic range is an effective predictor of the attractiveness of communications, whilst low variety gains individuals for amassing social impact [10].

The framework of a social network is normally presented as a graph of the area. A community is a group of densely linked nodes in a graph. The community framework is believed to be one crucial property of social support systems, recommending that a social networking might be partitioned into numerous clusters to ensure that nodes a single cluster are generally densely hooked up internally although not externally; these kinds of clustering may possibly gain through popular interests, family members subscriptions, or geographical categories.

4. TOPIC DETECTION

Data flows can be mapped to the topic area exactly where each societal site is manifested as a node and clusters of societal sites form topics [11]. We are able to, therefore, discover topics by discovering online communities in the societal site co-occurrence community. In line with the relationship between the social networking and also the topic space, it is crucial to designate each end user or information with various topics and measure topical diversity. As excessive topic variety is an effective predictor of recognition of communications, although minimal multiplicity gains individuals for amassing social impact.

Well-liked topics, for example, vacation greetings, are similarly well-known amongst powerful and poor connections. For this reason, suggested effects will verify and develop preceding theoretical along with survey-based work, representing that robust ties are seen as a variety of pursuits.

Social networking is increasing dynamically and constantly. Such development is in conjunction with the spread of information in addition to the network: the system topology has effects on the routes of information diffusion; the particular procedure of connections in the system is brought about by the site visitors. The proposed Pseudo Activity flow for information diffusion in online social network is shown in fig 3 below. This will be base for further development of proposed system.

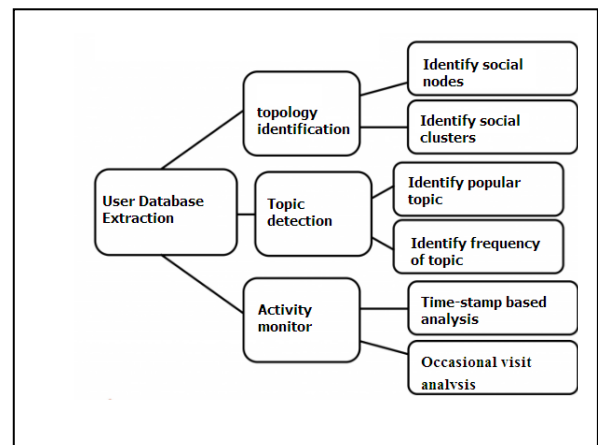


Fig 3: Pseudo Activity flow for information diffusion in online social network

To develop the strategic framework and algorithm pseudo code for proposed research individual database/datasets from LinkedIn, MANET community, IIARD dataset, and Pinterest are going to be considered for information extraction. The flow diagram shown in Fig1 offers research execution measures for the proposed research. Further, proposed algorithm will be tested using Apache JMeter and HIVE portal.

5. CONCLUSION

The proposed system need and methodology is discussed in this paper. The pseudo flow of information diffusion in the online social network is provided, where LinkedIn, MANET system, and Pinterest will be considered as a key element for analysis. This system will provide significant contribution to identify security elements, marketing elements for e-commerce and can also be useful for various domain user analyses.

6. REFERENCES

- [1] Kim, Jooho, and Makarand Hastak. "Social network analysis." *International Journal of Information Management: The Journal for Information Professionals* 38.1 (2018): 86-96.
- [2] Arnaboldi, Valerio, et al. "Online social networks and information diffusion: The role of ego networks." *Online Social Networks and Media* 1 (2017): 44-55.
- [3] Wang, Ru, et al. "Modeling of large-scale social network services based on mechanisms of information diffusion: Sina weibo as a case study." *Future Generation Computer Systems* 74 (2017): 291-301.
- [4] Yang, Yile, et al. "Noncooperative Information Diffusion in Online Social Networks Under the Independent Cascade Model." *IEEE Transactions on Computational Social Systems* 4.3 (2017): 150-162.
- [5] Kumar, Nagendra, et al. "Toward maximizing the visibility of content in social media brand pages: a temporal analysis." *Social Network Analysis and Mining* 8.1 (2018): 11.
- [6] Cetto, Alexandra, et al. "'Thanks for Sharing'—Identifying Users' Roles Based on Knowledge Contribution in Enterprise Social Networks." *Computer Networks* (2018).
- [7] Cui, Laizhong, et al. "DDSE: A novel evolutionary algorithm based on degree-descending search strategy for influence maximization in social networks." *Journal of Network and Computer Applications* 103 (2018): 119-130.
- [8] Kohout, Jan, et al. "Learning communication patterns for malware discovery in HTTPs data." *Expert Systems with Applications* (2018).
- [9] Razis, Gerasimos, and Ioannis Anagnostopoulos. "Modeling Influence with Semantics in Social Networks: a Survey." *arXiv preprint arXiv:1801.09961* (2018).
- [10] Liu, Yezheng, et al. "Identifying impact of intrinsic factors on topic preferences in online social media: A nonparametric hierarchical Bayesian approach." *Information Sciences* 423 (2018): 219-234.
- [11] Corbellini, Alejandro, et al. "DPM: A novel distributed large-scale social graph processing framework for link prediction algorithms." *Future Generation Computer Systems* 78 (2018): 474-480.
- [12] Myers, Seth A., Chenguang Zhu, and Jure Leskovec. "Information diffusion and external influence in networks." *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012.