Imep: Influence Maximization on Social Media with the Impact of E-Commerce Products

Amit Mittal
Chitkara Business School
Chitkara University
Punjab, India
amit.mittal@chitkara.edu.in

Meenal Arora
Chitkara Business School
Chitkara University
Punjab, India
meenal.bajaj20@gmail.com

Ajay Rana
Engineering and Technology
Amity University
Noida, Uttar Pradesh, India
ajay_rana@amity.edu

2022 5th International Conference on Contemporary Computing and Informatics (IC3I) | 979-8-3503-9826-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/IC3I56241.2022.10072595

***Abstract*—Friends and family have a large influence on consumer purchasing decisions. Furthermore, many internet users prefer to wait for early adopter reviews before making a purchase decision in order to reduce the risk of adopting a new product. E-commerce corporations actively construct web-based social networks that enable users to share their experiences by submitting evaluations, evaluating other people's assessments, and talking with genuine members. They act as a starting point for online customers and a means of directing them to other shopping sites. E-commerce companies have recently begun to collect data on consumer social interactions on their websites, with the potential goal of understanding and exploiting social influence in customer purchase decisions to improve customer relationship management and boost sales. This article proposed influence maximization on Social media with the impact of the E-Commerce Products (IMEP) framework. We collect dig and yelp datasets for finding the Social Influence (SI) for E-Commerce Products (EP). The IMEP framework utilized the United Community Recognition algorithm (UCR) to find influence maximization (IM) accuracy. The experimental results are discussed with greedy clustering and the performance measures are compared.**

Keywords— E-commerce, E-commerce products, IMEP, Social Media, IM, UCR

I. INTRODUCTION

WeChat, Weibo, and Snapchat are examples of social apps that create a substantial amount of network data due to improved network infrastructure and the spread of the Internet [1]. Network data analysis may result in long-term applications. Influence maximization (IM) has been a hot topic in recent years, and as a result, it has had a major impact in many different areas, including rumor suppression, network analysis, and content recommendation [2]. IM aims to discover the most important network seed nodes and their impact on other nodes while simultaneously expanding their influence [3]. Unfortunately, addressing the IM problem becomes more complicated as network capacity increases [4][5]. Real SNs often feature varying community organizations with strongly connected nodes. Notably, there are few connections between villages. Because communities are often smaller than the network as a whole, it is easy to discover seed nodes in each community [6-9]. Examine variations in community structure and distribution of influence to resolve IM in actual SNs, achieving a balance between influence spread and running duration [10]. In real-world applications, however, the expense of obtaining network data is sometimes prohibitive [11]. In many public health fields, an intervention's efficacy depends on information regarding interactions with community members [12]. The entity implementing the intervention must exert significant effort to collect data through in-person surveys.

Social networks significantly impact people's lives due to their immense influence, high reliance on recommendations, and influence from friends and family [17]. Today, these commonplace decisions are made online [13]. Online social structures expose Internet users to various products and services from which they can choose. They allow social network platforms to suggest things without requiring human participation. These innovations are anticipated to facilitate discoveries that would have otherwise gone unnoticed [14].

As businesses compete to find innovative viral marketing strategies to advertise their products with minimal effort and low budgets, the need for a practical strategy to maximize the number of people who purchase a targeted product grows [15]. People are frequently overwhelmed by the number of options and choices available when starting or running a business online [17]. This paper's main contribution is finding the IM using the UCR algorithm with the community detection method.

The remainder of this paper is structured as follows. Numerous authors address a variety of IM in SM strategies in Section 2. The IMEP model is shown in Section 3. Section 4 deals with results and discussion. Section 5 concludes with a discussion of the result and future work.

II. BACKGROUND STUDY

Abed, S. et al. [1] Despite creating social media solutions for small and medium-sized organizations (SMEs), companies and consumers were sluggish to embrace them. Social media may help SMEs build their brand before accessing foreign markets, accelerating internationalization. Social media helps SMEs engage employees and manage information. Social media helps attract new customers, listen to customer feedback, develop corporate networks, manage reputations, enhance client trust and loyalty, and collect marketing information. Competitors, specialists, and consumers impact SMM. SME marketing and customer communication affect business value. SME social media adoption and business values are difficult. Global corporations might benefit from examining social media's influence on SMEs, its relation to customer loyalty, and its efficacy for SME internationalization. This paper examined SMEs' social media e-commerce adoption.

D. S. Alorini et al. [3] This article analyses a user's impact and how to stop it to avoid erroneous information on Twitter utilizing followers, retweets, likes, and mentions. Spammers create innovative tricks to deceive social media users. They monitor key users to propagate false information. Competing groups trying to reduce social media misinformation found influence blocking difficult. Twitter's influence blocking maximized remarks' impact. When a tweet is unpleasant, restricting its circulation may reduce its reach. Large, well- known firms with a blue checkmark have

979-8-3503-9826-7/22/$31.00 ©2022 IEEE

verified accounts. Users may authenticate non-spamming accounts and trust their material. Twitter recommended checking the verification mark before applying to guarantee performance. This article's greedy algorithm showed Twitter's detrimental influence. Several Twitter influencers' data was reviewed to measure impact.

J. Tong et al. [5] This article focuses on maximizing information in a competitive economy. The authors used social network data and HITS to compile the data. The strategy for optimising the impact took into account the user's preferences and interests, as well as the interactional content between nodes. Experiments on real-world Twitter datasets indicate the author's methodology may enhance accuracy and computation time over existing methods.

Kim, Y. A., & Srivastava, J. [7] Social influence influences e-commerce decision-making, however until recently, social interaction data wasn't gathered frequently, therefore few studies have explored social effect in an e-commerce decision support system. Online shoppers care more about connections than purchases. Unfamiliar shoppers might depend on user evaluations on social networks. Not manufacturers or recommendation algorithms influenced consumer selections. Positive reviews may inspire others to try your items. E- commerce companies employed consumer recommendations and reviews as a decision-making tool. Long-term, online retailers may increase sales and save advertising costs. Social impact in E-commerce decision-making may benefit companies, says author.

Singh, A. K., & Kailasam, L. [11] In this research, the author offers a link prediction-based effective node tracking approach for discovering seed nodes in a dynamic social network. In the graph snapshot, this seed set was utilised for influence spread. This strategy promotes the propagation of influence in highly dynamic social networks by increasing the number of influenced nodes.

Yersinia, S. et al. [15] The impact of review comments on product sales was assessed when they were included on a Facebook page. To compute the overall rating, a mixed influence model was used, which took into account both the author's point of view and the impact of previous review comments.

A. Problem Specification

The ultimate goal is to select a seed node group that will maximize information diffusion. Examine information transmission in a network using the commonly used Independent Cascade Model (ICM). The ICM only uses active or dormant nodes. Except for the first seed collecting, the system is initially inactive. The technique is divided into carefully timed segments. Each freshly activated node initiates contact with every nearby node and succeeds a given number of times with probability p. At the end of the procedure, no new nodes are present at the final stage

III. SYSTEM MODEL

Before This article discusses E-commerce product aspects of this community detection and influence maximization algorithm in SMI. To maximize the influence and radiation of this node in a network by combining previously explored methods such as community detection and influence weight calculation. To hypothesize that by analyzing social network users, one can identify the social groups of users who influence their peers' behavior. This framework's first step is identifying communities, such as social networks.

An improved approach for Community-Based Influence Maximization was provided to capitalize on social influence (CTIM). Although assembling communities is an important stage, it is not required. Because of social influence in communities, these rung influence the success or failure of SN behavior adoption. A well-organized community will aid in the diffusion of influence. As a result, this initial step may considerably impact the mechanism for spreading influence. Similar users have similar behaviors, and identifying the major nodes in each Community makes it easier to spot these communities in a graph. Figure 1 depicts the dispersion of influence.

A. Data Set

The proposed model is implemented in Java on NetBeansIDE8.2 platform. Here use two real-world and openly accessible datasets: Yelp dataset challenge 20141 and The Digg dataset. 2 The Yelp dataset challenge 2014 has 366,715 users, 2,949,285 links, and 61,184 items, whereas the Digg dataset has 30,358 users, 99,846 directed arcs, and 7,100 items. To estimate that the time the review was placed on the site was when the consumer purchased the goods for the Yelp Dataset Challenge 2014.

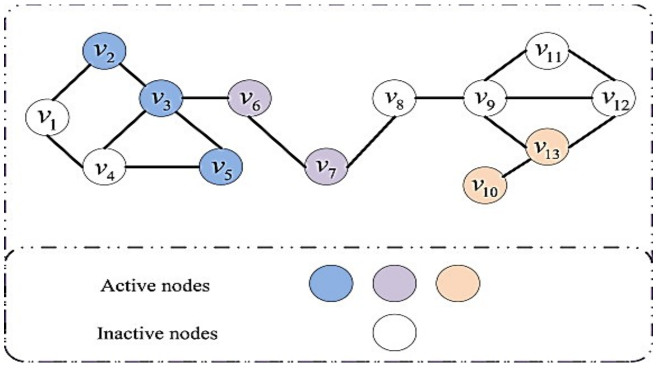

Fig.1. Illustration of influence spread

B. Explicit Separation Between Individual's Ip & Si Using United Community Recognition Algorithm: Esucr

The Viral marketing, as a successful marketing method for increasing the SN's impact, targets a limited number of initial consumers centred on their shared network of friends or family. This chapter suggests a fresh approach to viral marketing, a new community-based influence maximization (CTIM) technique to boost viral marketing effectiveness. Create a comprehensive latent variables model that captures each user's community-level subject interest, item-topic relevance, and distribution of community membership. The model is then put to the test through Gibbs sampling. The level of community-to-community influence is then inferred using topic-irrelevant influence and interest in the community issue, and the strength of user-to-user influence is inferred using the distribution of community-to-community influence across each user's community membership. To improve quality and efficiency, ESUCR is a community-based heuristic algorithm that will employ a divisive strategy to identify influential nodes that are both subject-aware and community-relevant. The IM is analyzed based on community structure and distribution disparities in influence to increase the effectiveness of the Greedy technique.

Because of the power-law distribution of SN node effect, the IM solution has a candidate phase and a greedy phase that seeks the ideal solution. Critical nodes inside and outside of each community's borders are identified using the heuristic technique during the preliminary stages of design. First, we use the Greedy technique, which is based on the greedy submodular property, to locate seed nodes within the candidate set, and then we constrain the effect of those nodes on their neighbors. SN G. is a politically divided society, as seen in Figure 2. Nodes may be allocated once the IM issue has been resolved. The percentage option is to postpone the decline in influence spread caused by an increase in the number of highly influential nodes. The number of people willing to work as candidate nodes is directly related to the population size, as the distribution of power follows the same pattern. Let Cs be the partition of a network G whose communities do not overlap for a particular community partitioning.

$Cs=(C1, C2,\ldots,Cr)$

Let U represent the collection of potential nodes. Then there are two scenarios for the distribution of the number of U.

Case 1: If $r < |U|$, then the number of candidate nodes |UCi

| concerning Community Ci is calculated as:

$|Uci| = \mu \times (K \times |Ci|)/n \quad i \in \{1,2,\cdots,r\}$,

In this formula, n represents the total number of nodes and k represents the number of seed nodes, respectively, and represents the number of candidate nodes mined from each community.

As shown in Figure 2, the preceding study implies that nodes that are much more essential to the network would have higher Community Detection scores. Then, depending on the candidate nodes' impact rankings, candidate nodes with a big influence are picked to create candidate nodes inside the Community. Figure 3 depicts the suggested architecture.

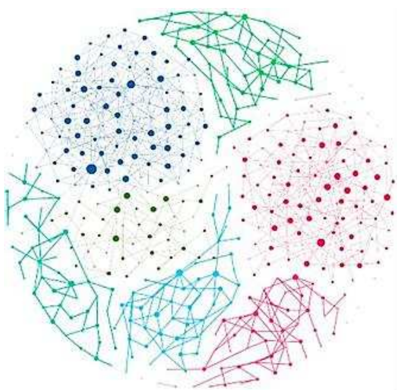

Fig.2. Community Detection

C. United Community Recognition Algorithm

Input: Social Graph G(V, E).

Output: Community Structure C = {C1, C2, C3, . . . , Cm}.

1: C ← ϕ ;

2: CTemp ← {V1 , V2 , . . . , VN} ;

3: for i=1 to |V |−1 do

4: Cj ← Select a random community from C ; 5: density ← 0;

6: CMerge ← ϕ ;

7: for each community Ci ∈ CT emp do 8: if Ci = Cj then

9: temp ← ∆Ci ,Cj ;

10: if temp > density, then

11: density = temp;

12: CMerge ← Ci ;

13: else if temp = density and CMerge = φ then

14: if GetCommunity(Cj , Ci) > GetCommunity(Cj

, CMerge) then

15: CMerge = Ci ;

16: end if

17: end if

18: end if

19: end for

20: if CMerge = φ then

21: CMerge ← Merge (CMerge, Cj); 22: else

23: C ← C ∪ Cj ;

24: end if

25: CTemp ← CTemp \ Cj ; 26: end for

27: return C = {C1 , C2 , C3 , . . . , Cm };

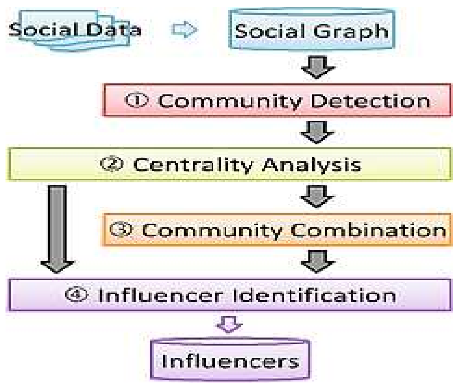

Fig.3. Architecture Diagram

D. Analysis

The approach for recognizing group membership in a network may allow for several types of communication. To aid in identifying persons, the problem researcher must locate individuals with similar interests or ties. Research on the benefits and drawbacks of each available detection technology was conducted. The report examines and evaluates several strategies used to compare and contrast various aspects of the procedures.

Choosing seed nodes from the set of candidate nodes using the sub-modular property-based UCR algorithm instead of choosing the inside and the communities' boundary.

This structure is a probabilistic modelling strategy for the impact of user attributes, user edges, and purchase logs. To consider that influence logs are not as dense as friend relationship links, this analysis integrates influence logs and friend relationships into the model. Thus, to use a network of friendships as a supplement to influence diffusion logs to imitate the spread process. Three core components are focused on this generative process: item properties, user edges, and possible impact logs are generated Figure 4 depicts the probabilistic visual representation of the generative model, and the UCR Algorithm explains its generating process.

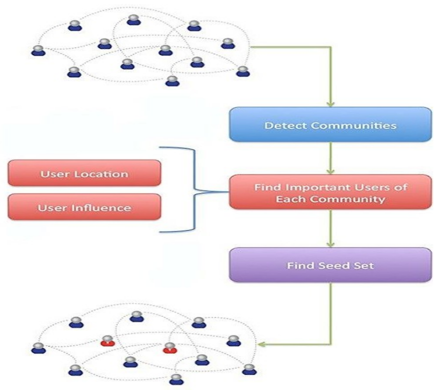

Fig.4. Communities Aware Influence Maximization

IV. RESULTS AND DISCUSSION

The IMEP framework has been implemented by using a java programming language with a yelp and dig dataset. By comparing the greedy algorithm and the proposed IMEP framework.

Figure 5 represents the accuracy comparison experiments on Yelp dataset with the influence spread. The X-axis represents the influence spread, and the Y-axis represents the influence spread accuracy.

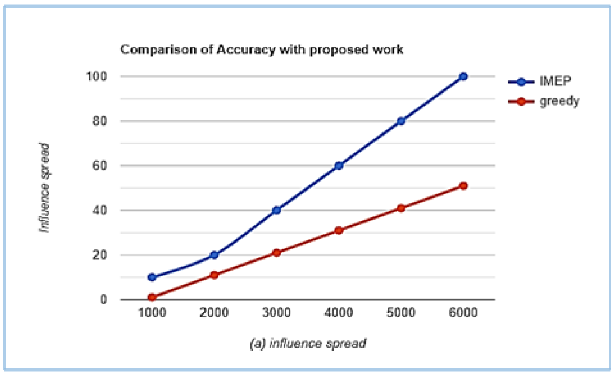

Fig.5. Accuracy comparison chart for yelp dataset

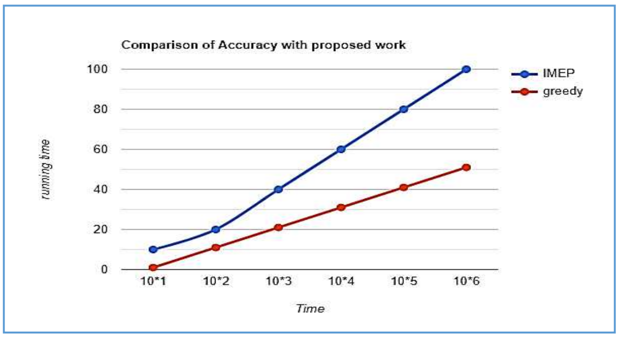

Fig.6. Running time comparison chart for yelp dataset

Figure 6 represents the yelp dataset influence spread running time comparison chart. In X-axis denotes the time and y-axis denotes the running time.

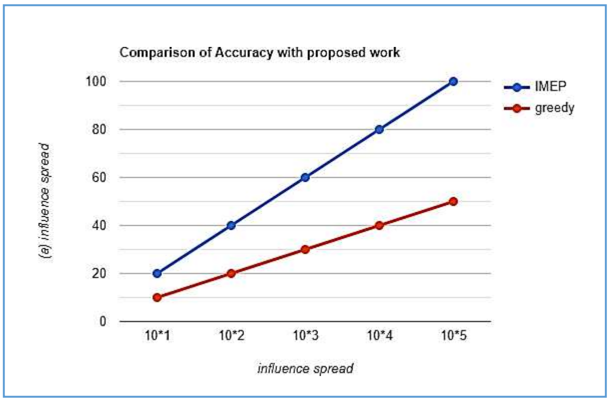

Fig.7. Accuracy comparison chart for digg dataset

Figure 7 represents the accuracy comparison experiments on digg dataset with the influence spread. The X-axis represents the influence spread, while the Y-axis represents the influence spread accuracy.

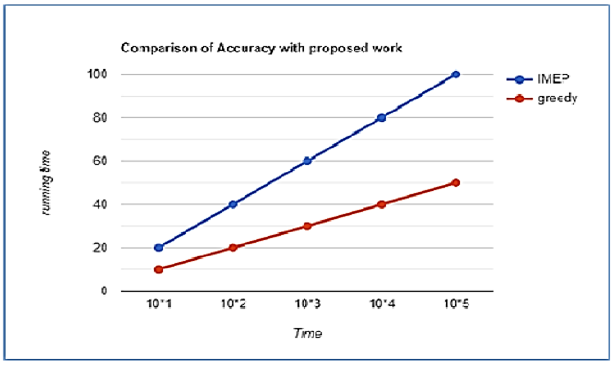

Fig.8. Running time comparison chart for digg dataset

Figure 8 represents the digg dataset influence spread running time comparison chart. The X-axis represents time, while the Y- axis represents running time. Finally overall comparison the proposed framework is better than an existing greedy algorithm.

V. CONCLUSION

Viral marketing, which derives from word-of-mouth marketing and applies the theory of information diffusion, is a popular topic of study because of the value it provides to businesses. Viral marketing may increase the spread of influence in the SN by carefully identifying a sample of first users who are most like the target customer [16]. We proposed IMEP framework model for IM for E-commerce products. The ESUCR model is used to propose a new IP and SI method. These essential components were combined separately by each component of this framework to create this complete model for maximizing effects in OSNs. The experimental results are compared with the greedy algorithm with the use of yelp and dig dataset. Finally, the accuracy has achieved 99%. Further, we consider more datasets with machine learning algorithms.

REFERENCES

[1] S. S. Abed, Y. K. Dwivedi, and M. D. Williams, "Social media as a bridge to e-commerce adoption in SMEs: A systematic literature review," Mark. Rev., vol. 15, no. 1, pp. 39–57, 2015.

[2] F. Amato, V. Moscato, A. Picariello, and G. Sperlí, "Influence maximization in social media networks using hypergraphs," in Green, Pervasive, and Cloud Computing, Cham: Springer International Publishing, 2017, pp. 207–221.

[3] D. S. Alorini, D. B. Rawat and G. S. Alorini, "On the Influence Blocking Maximization for Minimizing the Spreading of Fake information in Social Media," 2020 Spring Simulation Conference (SpringSim), 2020, pp. 1-10, doi: 10.22360/SpringSim.2020.CSE.005.

[4] J. Ge, L. -L. Shi, Y. Wu and J. Liu, "Human-Driven Dynamic Community Influence Maximization in Social Media Data Streams," in IEEE Access, vol. 8, pp. 162238-162251, 2020, doi: 10.1109/ACCESS.2020.3022096.

[5] J. Tong, L. Shi, L. Liu, J. Panneerselvam and Z. Han, "A novel influence maximization algorithm for a competitive environment based on social media data analytics," in Big Data Mining and Analytics, vol. 5, no. 2, pp. 130-139, June 2022, doi: 10.26599/BDMA.2021.9020024.

[6] Kianian, S., & Rostamnia, M. (2020). An Efficient Path-Based Approach for Influence Maximization in Social Networks. Expert Systems with Applications, 114168. doi:10.1016/j.eswa.2020.114168

[7] Y. A. Kim and J. Srivastava, "Impact of social influence in e-commerce decision making," in Proceedings of the ninth international conference on Electronic commerce - ICEC '07, 2007.

[8] M. Sha, Y. Li, Y. Wang, W. Guo, and K.-L. Tan, "River: A real-time influence monitoring system on social media streams," in 2018 IEEE International Conference on Data Mining Workshops (ICDMW), 2018.

[9] J. S. More and C. Lingam, "A SI model for social media influencer maximization," Appl. Comput. Inform., vol. 15, no. 2, pp. 102–108, 2019.

[10] N. Vafaei, M. R. Keyvanpour and S. Vahab Shojaedini, "Influence Maximization in Social Media: Network Embedding for Extracting Structural Feature Vector," 2021 7th International Conference on Web Research (ICWR), 2021, pp. 35-40, doi: 10.1109/ICWR51868.2021.9443150.

[11] A.K. Singh and L. Kailasam (2021). Link prediction-based influence maximization in online social networks. Neurocomputing, 453, 151–163. doi:10.1016/j.neucom.2021.04.084

[12] C. Stubb and J. Colliander, "'This is not sponsored content' – The effects of impartiality disclosure and e-commerce landing pages on consumer responses to social media influencer posts," Comput. Human Behav., vol. 98, pp. 210–222, 2019.

[13] E. Turban, N. Bolloju, and T.-P. Liang, "Social commerce: An e-commerce perspective," in Proceedings of the 12th International Conference on Electronic Commerce Roadmap for the Future of Electronic Business - ICEC '10, 2010.

[14] Yadav, M., & Rahman, Z. (2018). The influence of social media marketing activities on customer loyalty: a study of e-commerce industry. Benchmarking: An International Journal, 00–00. doi:10.1108/bij-05-2017- 0092

[15] Yerasani, S., Appam, D., Sarma, M., & Tiwari, M. K. (2019). Estimation and maximization of user influence in social networks. International Journal of Information Management, 47, 44–51. doi:10.1016/j.ijinfomgt.2018.12.016

[16] Tandon, U., Mittal, A., & Manohar, S. (2020). Examining the impact of intangible product features and e-commerce institutional mechanics on consumer trust and repurchase intention. Electronic Markets, 1-20.

[17] Mittal, A., Bhandari, H., & Chand, P. K. (2021). Anticipated positive evaluation of social media posts: social return, revisit intention, recommend intention and mediating role of memorable tourism experience. International Journal of Culture, Tourism and Hospitality Research.