International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 04 Issue: 12 | Dec-2017 www.irjet.net p-ISSN: 2395-0072

DISCOVERING INFLUENTIAL USER BY COUPLING MULTIPLEX HETEROGENEOUS OSN'S

Ms. A.Sivasankari[1], Ms.D.Janani [2], Mrs. K.Arulmozhi[3]

*[1]Head of the Department (CS), Dept of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India**

*[2]Research Scholar, Dept of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India**

*[3] Head of the Department (BCA), Dept of Computer Science and Applications, D.K.M College for Women (Autonomous), Vellore, Tamilnadu, India**

---***---

Abstract - *Recently, in online social network (OSNs), another heuristic plan for an influence maximization problem in social networks: how to monetarily choose a subset of people (seeds) to trigger an extensive course of further receptions of another conduct in light of a disease procedure. Nonetheless, nowadays users often join several OSNs such that information could be spread across different networks simultaneously. Therefore, in order to obtain the best set of seed users, it is crucial to consider the role of overlapping users under these circumstances. In this article, we propose a unified framework to represent and analyze the influence diffusion in multiplex networks. All the more particularly, we handle the LCI issue by mapping an arrangement of networks into a solitary one by means of lossless what's more, lossy coupling plans. The lossless coupling plan jelly all properties of unique networks to accomplish high caliber arrangements, while the lossy coupling plan offers an alluring elective when the running time and memory utilization are of essential concern. Different tests led on both genuine what's more, blended datasets have approved the adequacy of the coupling plans, which additionally give some intriguing bits of knowledge into the procedure of impact engendering in multiplex networks to research the issue in multiplex networks with heterogeneous dissemination models in which each system may have its own dissemination display. This model is capable of analyzing influential relationships in a heterogeneous social network in decentralized manners and identifying the influential users more efficiently than traditional seed selection algorithms.*

Key Words: **OSN, LCI, memory Utilization, Social Network, Cascade model**

1. INTRODUCTION

In the current decade, the ubiquity of online social networks, for example, Facebook, Google, MySpace and Twitter and so forth. has made another real correspondence medium and shaped a promising scene for data sharing and disclosure. All things considered, Facebook clients burn through 7 h and 45 min per individual every month on associating with their companions; 3.2 billion likes and remarks are posted each day on Facebook; 340 million tweets are conveyed ordinary on Twitter. Such engagement of online clients prepares the land for data proliferation to a degree which has never been accomplished in the broad communications. All the more vitally, online social networks (OSNs) additionally acquire one of the real properties of genuine social systems—the verbal impact, in which individual conclusion or choice can be reshaped or transformed through impact from companions and partners. As of late, roused by the noteworthy impact of viral promoting, OSNs have been the most appealing stages to build mark consciousness of new items and in addition fortify the connection between clients and organizations. All in all, a definitive objective is to locate the slightest promoting cost set of clients which can trigger a gigantic impact.

Alongside the quick advancement of all current OSNs, there have been a significant number of clients who keep up a few records all the while, which enable them to engender data crosswise over various systems. For instance, Jack, a client of both Twitter and Facebook, took in of another book from Twitter. In the wake of understanding it, he thought that it was exceptionally intriguing and shared this book with companions in Facebook and additionally Twitter. This should be possible by arranging both of the records to permit consequently posting crosswise over various informal organizations. As an outcome, the item data is presented to his companions and further spreads out on the two systems. On the off chance that we just concentrate on an individual system, the spread of the data is evaluated erroneously. Along these lines considering the impact just in one system neglects to recognize the most powerful clients, which rouses us to think about the issue in multiplex systems where the impact of clients is assessed in light of all OSNs in which they take an interest. Related works. Almost the majority of the current works examined unique variations of the slightest cost impact issue on a solitary arrange. first planned the impact boost issue which solicits to locate a set from clients who can augment the impact. The impact is spread in view of a stochastic process called Independent Cascade Model (IC) in which a client will impact his companions with likelihood corresponding to the quality of their kinship. The creator demonstrated that the issue is NP-hard and proposed a covetous calculation with estimate proportion of . From that point onward, an impressive number of works contemplated and planned new calculations for the issue variations on the same or expanded models such as. There are additionally chips away at the Linear Threshold edge (LT) show for impact proliferation in which a client will receive the new item when the aggregate impact of his companions outperform a few edge. The in approximability too as proposed productive calculations for this issue on an uncommon instance of LT display. In their model, the impact between clients is uniform and a client is affected if a specific portion of his companions are dynamic. As of late, specialists have begun to investigate multiplex systems with works studied the connection between offline and online networks.

In this article, we examine the minimum cost impact (LCI) issue which goes for finding an arrangement of clients with least cardinality to impact a specific part of clients in multiplex systems. Due to the intricate dissemination process in multiplex systems, it is troublesome to build up the arrangement by straightforwardly expanding past arrangements in a solitary system. Furthermore, contemplating the issue in multiplex systems presents a few new difficulties: 1) step by step instructions to precisely assess the impact of covering clients; 2) in which arrange, a client is simpler to be impacted; and 3) which arrange spreads the impact better. To answer these questions, we initially acquaint a model portrayal with represent how data proliferate in multiplex systems by means of coupling plans. By mapping numerous systems into one system, extraordinary coupling plans can safeguard fractional or full properties of the first systems. From that point onward, we can misuse existing arrangements on a solitary system to tackle the issue in multiplex systems. Besides, through exhaustive examinations, we have approved the viability of the coupling plans, and likewise give some intriguing bits of knowledge into the procedure of impact proliferation in multiplex systems. Our principle commitments are abridged as takes after.

• We propose a model portrayal by means of different coupling plans to diminish the issue in multiplex systems to a proportionate issue on a solitary system. The proposed coupling plans can be connected for most prevalent dissemination models including: direct edge demonstrate, stochastic edge model and free falling model.

• We give First, we unequivocally portray every client with two particular factors: the susceptibility of being influenced (SI) and persuasive power (IP) speaking to the capacity to effectively impact others and figure clients' SIs and IPs as per their social relations, and after that, an arched value request bend based model is used to legitimately change over every client's SI into influence cost (PC) speaking to the cost used to effectively influence the person to embrace another conduct. Besides, a novel financially savvy choice plan is proposed, which embraces both the value execution proportion (PC-IP proportion) what's more, client's IP as a coordinated determination paradigm and in the mean time expressly considers the covering impact; at long last, reproductions utilizing both falsely produced and genuine follow arrange information represent that, under similar spending plans, PPRank can accomplish bigger dissemination extend than other heuristic and beast drive insatiable plans without considering clients' influence costs. Specifically, the change factor scales up with the extent of the system which enables the calculation to keep running on expansive systems with a large number of hubs.

• We direct broad analyses on both genuine and orchestrated datasets. The outcomes demonstrate that thinking about multiplex organizes rather than a solitary system can adequately pick the most compelling clients.

2. MODEL AND PROBLEM DEFINITION

Definition 1: A Social networks we consider k networks G1,G2,.....Gk each of which is modeled as a weighted graph Gi = (Vi, E I,θ^{i}, W^{i}) with a clear topological structure, where V = {v1, v2, ..., vn} stands for the nodes (users) in the network, E = {eij |vi ∈ V ∧ vj ∈ V, vi ≠vj} denotes the edges (relationships) among nodes. A particular edge can be represented as a threetuple, i.e., eij = (vi , vj , wij), where wij is the weight of eij which represents the influence strength. Each node vi has a set of neighbours {vj |vj ∈ Γ(vi), eij ∈ E}. While, vi .q indicates the pheromone amount accumulated on corresponding node vi , which can be regarded as an attribute of vi . Similarly, since wij represents the weight of edge eij , it is denoted by using the notation eij .w in this paper.

Definition 2: an Agent is defined as an autonomous agent in the network G, which crawls across G based on the network topology. An agent can be represented as a three-tuple, i.e., am = (m, q_m^n, T_m^n), which means ant am carries pm pheromone in tour T_m^n (see Definition 3). There exist a number of ants, P= {p1, p2, ..., pn}, in the social network, and they keep crawling in the network. Moreover, they are capable of discovering and evaluating the amount of pheromone on the current node and the ones nearby. However, the ants cannot communicate directly with each other.

Definition 3: A Path T_m^n =< v_1, v_2, ..., v_n > is defined as the path that ant am walks through in the n round. Specifically, agent am randomly selects a starting point. Next, it crawls from one node to the adjacent neighbours and eventually ceases when reaches the end point ve, where Γ(ve) ⊂ T_m^n), ∪ |Γ(ve)| = 1.

Definition 4: Heterogeneous Network represents the information and heuristics passed by an agent to the peers based on its experience. q_m^n, denotes the total amount of artificial pheromone carried by ant am in the n round, which will be distributed to each node of T_m^n, after am completes the Path.

Problem Definition In this paper, we address the key issue of viral showcasing in multiplex systems: the LCI issue. The issue solicits to discover a seed set from least cardinality which impacts an expansive division of clients.

Definition 5 (LCI Problem) : Given a system of k networks G1…k with the set of users U, a positive integer d, and 0 , the LCI problem asks S ⊆ U to find a seed set of minimum cardinality such that the number of active users after hops according to the LT model is at least β fraction of users, i.e.,|Ad(G1…k ,S)|≥ β |U|.

Whenever k=1, we have the variation of the issue on a solitary system which is NP-difficult to settle inapproximability and proposed a calculation for a unique situation when the impact between clients is uniform and a client is enacted if a specific portion of his companions are dynamic. In the accompanying segments, we will display diverse coupling systems to decrease the issue in multiplex systems to the issue in a solitary system keeping in mind the end goal to use the calculation plan.

3. ALGORITHMS

Algorithm 1 Path Formation Algorithm

Input: : A system of networks G1…k ,am, n Output: T_m^n, T_m^n ⊆ Vi

1: Initialize am and random select a starting point vs, vs ∈ V i

2: for G (1…K)

3: Initialize a tour list T_m^n := ∅

4: while ∃Γ(vs) ∧ Γ(vs) ⊂ T_m^ndo

5: Initialize candidate list Vc := ∅

6: for ∀vi ∈ Γ(vs) ∧ vi ∈/ T_m^ndo

7: Compute the probability psi using Equation 1.

8: if psi > σ then

9: Vc := Vc ∪ {vi}

10: end if

11: end for

12: if Vc ≠∅ then

13: Determine the next node vn ∈ Vc using Equation 1.

14: T_m^n:= T_m^n∪ {vn}

15: vs := vn

16: else

17: vs := null

18: end if

19: end while

20 :end if

Algorithm 2 User (Seed) Selection Algorithm

Input: n, k, λ, Δt, Gi = (Vi, Ei)

Output: Vs

1: Initialize Agent set A := {p1, p2, ..., pn} which contains n Agents.

2: Initialize seed set Vs := ∅

3: All the n agent start to crawl in combined network G in the distributed servers.

4: while !convergence do

5: Compute EQ.

6: for vi ∈ V do

7: vi.q := vi.q − EQ

8: end for

9: Sleep for Δt

10: end while

11: Sort V order by q descend

12: for ∀vi ∈ V do

13: if |Vs| < k then

14: Vs := Vs ∪ {vi}

15: end if

16: end for

Advantages of Using Coupled Networks

To comprehend the advantage of taking thought of covering clients and coupled system, in this part, we are going to contrast the seed measure and/out utilizing coupled system. Specifically, we complete two correlations on: 1) affecting a small amount of the hubs in all systems by choosing seeds from each system what's more, taking the union to contrast and seeds accomplished from lossless coupling plan; 2) affecting a small amount of the hubs in a specific system by just picking seeds from that system contrast with the seeds acquired from the lossless coupling conspire.

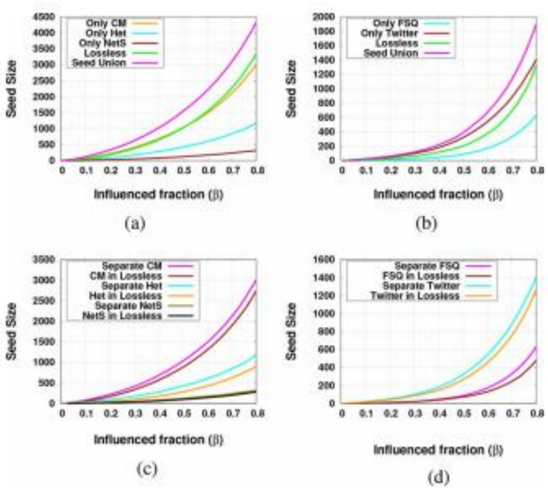

Fig 1. Quality of seed sets with and without using the coupled network. (a) Coauthor networks. (b) FSQ and Twitter. (c) Coauthor networks. (d) FSQ and Twitter.

The outcomes for the principal situation are appeared in Fig. 1. The seed acquired by the lossless coupling technique beats different techniques. The span of the union set is roughly 30% what's more, 47% bigger than lossless coupling technique in coauthor and FSQ-Twitter, separately. This demonstrates covering clients do engender data through a few systems and subsequently successfully help lessen the general seed estimate.

At long last, our paper portrays the impact proliferation in an interpersonal organization demonstrated as a static, coordinated and weighted chart in heterogeneous multiplex system Model, which, it might be said, could be viewed as one depiction of social chart over a gathered period. In any case, due to the following contemplations, conduct dissemination in informal communities ought to normally consolidate the measurement of time. Initially, as a rule, individuals' social connections may differ after some time: New connections show up, old connections are erased, and interface quality expands/diminishes. Subsequently, we will normally get time serial information of social diagrams; second, the impact of focal clients tends to spread more quickly all through a system than the impact of fringe clients does. Additionally, it more often than not sets aside less opportunity to achieve focal clients than fringe clients when data streams in social systems. In this way, it is intriguing and testing to altogether examine how to successfully choose introductory seeds in the previously mentioned dynamic condition in which time-differing social connections and dissemination rate are unavoidable.

4. CONCLUSIONS

In this paper, we contemplate the LCI issue in multiplex systems. To handle the issue, we presented novel coupling plans to diminish the issue to a rendition on a solitary system. At that point, we planned another metric to measure the stream of impact inside and between systems in light of the coupled system. Comprehensive tests give new bits of knowledge to the data dispersion in multiplex systems. In future, we can design a lightweight and distributed protocol to identify influential users through fixed length random walks.

REFERENCES

[1] "216 social media and internet statistics," [Online]. Available: http://thesocialskinny.com/216-social-media-and-internet-statistics-september-2012/

[2] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in Proc. KDD, 2003, pp. 137–146.

[3] W. Chen, C. Wang, and Y. Wang, "Scalable influence maximization for prevalent viral marketing in large-scale social networks," in Proc. SIGKDD, 2010, pp. 1029–1038.

[4] H. Zhang, T. N. Dinh, and M. T. Thai, "Maximizing the spread of positive influence in online social networks," in Proc. IEEE Int. Conf. Distrib. Computing Syst., 2013, pp. 317–326.

[5] J. M. Kleinberg, "Cascading behavior in social and economic networks," in Proc. 14th ACM Conf. EC, 2013, pp. 1–4.

[6] W. Chen, L. V. S. Lakshmanan, and C. Castillo, "Information and influence propagation in social networks," in Synthesis Lectures on DataManagement. San Rafael, CA, USA: Morgan & Claypool Publishers,2013.

[7] P. Domingos and M. Richardson, “Mining the network value of customers,” in Proc. 7th KDD, 2001, pp. 57–66.

[8] D. Kempe, J. M. Kleinberg, and É. Tardos, “Maximizing the spread of influence through a social network,” in Proc. 9th KDD, 2003,pp. 137–146.

[9] J. Leskovec et al., “Cost-effective outbreak detection in networks,” in Proc. of 13th KDD, 2007, pp. 420–429.

[10] A. Goyal, W. Lu, and L. V. S. Lakshmanan, “CELF++: Optimizing the greedy algorithm for influence maximization in social networks,” in Proc. of 20th WWW, 2011, pp. 47–48

BIOGRAPHIES

Ms. A.SIVASANKARI
Head of the Department (CS),
Assistant Professor,
Dept of Computer Science and Applications,
D.K.M College for Women (Autonomous),
Vellore, Tamilnadu, India

Ms. D.JANANI
Research Scholar,
Dept of Computer Science and Applications,
D.K.M College for Women (Autonomous),
Vellore, Tamilnadu, India

Ms. B.ARUL MOZHI
Head of the Department (BCA),
Dept of Computer Science and Applications,
D.K.M College for Women (Autonomous),
Vellore, Tamilnadu, India