International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Revue d'Intelligence Artificielle
Vol. 37, No. 4, August, 2023, pp. 1047-1055
Journal homepage: http://iieta.org/journals/ria

A Scalable Approach for Strengthening Social Media Ties Using Multi-Dimensional Analysis

Deepjyoti Choudhury*, Tapodhir Acharjee

Department of Computer Science & Engineering, Triguna Sen School of Technology Assam University, Silchar 788011, India

Corresponding Author Email: deepjyotichoudhury05@gmail.com

https://doi.org/10.18280/ria.370427

Received: 16 June 2023
Revised: 23 July 2023
Accepted: 30 July 2023
Available online: 31 August 2023

Keywords:
social media, relationship, tie strength, networks, analogy

ABSTRACT

The assessment of relationship strength among interconnected users in online social networks remains a critical focal point in contemporary research. Despite a multitude of studies on strong tie identification, this constitutes an enduring challenge. In this work, a novel method is introduced that amalgamates factors such as user profile information, communication frequency, and network composition to ascertain the strength of ties among social media users. The proposed method encompasses computations involving three distinct variables: Analogy Profile (AP), Analogy Friendship (AF), and Analogy Reaction (AR). These variables collectively contribute to determining the overall quality of user relationships. Pearson's correlation, serving as AP, aids in identifying and quantifying user correlations' strength and orientation. Jaccard's coefficient offers a measure of user similarity, hence its use as AF. Lastly, the User Interconnection Potency Level, serving as AR, provides insights into user interaction dynamics and behaviour. For the purpose of experimental validation, ten different real-time social networks were considered. The performance of the proposed method was evaluated using Precision, Recall, and the Dice Similarity Coefficient (DSC) as evaluation matrices, on ten distinct real-world online social media datasets. Comparative analysis with two state-of-the-art methods, namely Trust Propagation-User Relationship Strength (TP-URS) and User Relationship Strength Fusing Multiple Factors (URSMF), demonstrated superior performance of our method. It achieved top scores of 92%, 98%, and 95% for Precision, Recall, and DSC, respectively. Overall, the proposed method outperforms TP-URS and URSMF in estimating relationship strengths among social media network users. These results underscore the utility of incorporating factors like profile information, communication patterns, and network composition when measuring tie strengths.

1. INTRODUCTION

The degree of connection or proximity between persons or entities inside a social network is referred to as relationship strength in social media. It can range from weak and superficial connections to strong and profound bonds. Measuring social media tie strength is significant for various reasons, including providing valuable insights to people, organizations, researchers, and the platforms themselves [1]. Understanding the strength of social media relationships between users may lead to more successful communication, marketing initiatives, and user experience improvements. Measuring the strength of connections in social networks is a difficult issue. This is due to the fact that the strength of a connection may be impacted by a variety of factors, including the frequency of interaction, emotional intimacy, and amount of trust between the two people. It is critical to accurately measure the strength of ties inside social networks for tasks such as identifying important users, forecasting information dissemination patterns, and enhancing recommendation systems. Quantitative evaluation enables more exact decision-making. That is why we have adopted a multi-dimensional approach to calculate tie strength in this paper.

So far in the real world of networks, only direct or mutual connections have been made among the users. Thus, conventional methods were used to ascertain the strong connections among users in social networking sites which are linked to one another. Methods for building networks and their associated tie structures have also been discussed so far. Full network methods [2], Snowball methods [3], and Ego-centric networks [4] are popular among them. In the context of measuring tie strength in social networks, these methods offer different approaches to data collection and analysis. While ego-centric networks and snowball approaches concentrate on certain subsets or individual views inside the network, full network methods offer a comprehensive picture of the whole network. Researchers may consider these techniques to learn more about tie strength and the importance of it in social networks, depending on the study aims and the resources at their disposal. In most cases, the information we collect comes from real-world networks where individuals have a variety of connections with one another due to their social activities. In addition, the acquired information must be pertinent to the connections between users. In a university, for instance, it makes no difference if student registration numbers are entered into the database instead of the number of student publications (the original requirement). Like one-way connections between individuals on social media, state-of-the-art algorithms to find strengths in different linkages do not work well with social media. Numerous efforts have been made for various user connections based on various metrics such as correlation, classification, clustered community, etc.

In addition, our paper emphasizes the importance of strengthening the robustness of existing network connections through the utilization of many, diverse connections.

The idea that the strength of ties between two networks can vary depending on how much they overlap was initially proposed by Granovetter [5]. Relationship strength, he mentioned, is the sum of many characteristics such as length of association, psychological impact, degree of user closeness, etc. Granovetter argues that strong ties are formed when social circles meet, while weak ties serve as a conduit for information between groups of friends who might not otherwise interact. For information to spread throughout a whole network, weak ties are therefore more important than strong ones.

One of the most debated areas of study currently is how to measure the quality of a network's connections. Finding a good way to measure relationship strength is a major obstacle to research in this area. Before tackling more complex issues in social networks, like Friend Recommendations in a Social Bookmarking System [6], diffusion-based similarity on tripartite graphs [7], etc., this fundamental problem must be solved. Having access to someone's connection strength with another user is helpful in many contexts, such as predictions, module recommendations, news items, and more [8]. The quality of certain public webwork aids may be raised based on the evaluation of correlation robustness by the public webwork contributor.

Recent studies have focused specifically on developing methods for gauging the intensity of connections within social networks [9]. Predictions of a relationship's stability have been made using interaction data. However, the majority of the currently available methods only consider the strength of users' direct ties inside public internet networks. Alice and Bob may not be friends, but they do have a mutual friend. A similar challenge arises in social networks when trying to ascertain the quality of the connections between such individuals. As a result, we took an unconventional approach to determine the quality of connections mediated by third parties. In light of changing social network dynamics, it is necessary and potentially very advantageous to take an innovative approach to evaluating the quality of connections facilitated by third parties. It offers a more sophisticated view of connection strength and its implications for different sectors, including marketing, social science, and the architecture of online platforms. It also emphasizes the growing complexity of contemporary networks and the effect of intermediaries.

This paper mainly contributes to compute the tie strength based three different factors, namely Analogy Profile, Analogy Friendship and Analogy Reaction among social media users. While we combine the individual profile information, number of established connections with that particular individual and the frequency of sharing information or comments to its own network give the best possible way to find the tie strength among the users in that social media network. The novelty of the computation of tie strength in this manuscript provides the application of the three unique formula, like Pearson Correlation as AP, Jaccard's coefficient as AF, and User Interconnection Potency Level as AR together to fit on the three factors mentioned above. Finally, we compute the proposed formula for tie strength and implemented in different real time social media datasets to check the accuracy. Three influential factors α, β, and γ increase the significance of the outcome of our experiments. As all the three parameters are assigned almost equal weightages, it leads to a good score in precision, recall and DSC. In this way the probability of losing efficient connections will also be reduced and performs better to enhance relationship strength. After the execution of the experiments, our proposed method has achieved better results compared with the existing methods and it is well elaborated in the Results & Discussion section later.

The paper is structured as follows: the second section provides context for the topic by discussing relevant background information and previous research. In the third section, we discussed the enumerative structure of our suggested method in detail. Our suggested model's workflow, the methods used to calculate relationship strength, the assessment matrices, and the dataset are all discussed in detail here. The fourth section contains a summary and analysis of the experimental data. This paper concludes with a discussion of our evaluation results and directions for further research.

2. RELATED WORKS

A great deal of effort has gone into communicating methods for determining user strengths in various social media relationships. Even while the definition of friendship in available online social networks is broader than what is typically considered in sociological study, the information conveyed by the interactions is weaker [10]. In this paper, seven robust variables were established based on project data that can be used to predict the durability of ties. A technique was described by Gilbert et al. [11] that links publicly available network data with link stability. The published method uses a database of over 2,000 public network links and performs brilliantly, being able to tell the difference between robust and vulnerable connections with an accuracy of over 85%. In addition, they pinpointed the process by which social network connection resilience develops. Since Facebook's creation, academics have relied heavily on the platform's data to learn about users' actions. To capture this concept, Viswanath et al. [12] examined the development of activity among Facebook users. As the online public network ties have expanded, they found that there is a common weakening shift of endeavor in the linkages of the pursuit network that is about to appear and move fast over a period of time. To determine the quality of user connections based on their participation in a variety of network activities (sharing, labeling, etc.), an unsupervised model was developed [13]. In particular, a coordinate ascent optimization strategy and a link-based latent variable model for reasoning were developed. A method for assessing the closeness of connections in social networks was presented and evaluated by Srba and Bieliková [14]. The estimated end-user connections provide useful information about people who share similar interests, hobbies, and other characteristics. There are several ways in which the estimated relationship strengths might be put to use in order to provide consumers with better data.

Bookmarking services use a tag vocabulary to estimate the reliability of links between webpages and build the web of relationships between them [15]. At this point, the Bayes theorem was used to infer group strength from the strength of an individual. After performing a tag estimation, the network was built in a smaller footprint, with some non-principal linkages removed. Trustor and trustee each provided an estimate of the trust's value, and the total trust value was determined by a weighted average of these four assessments [16]. In addition, a fundamental architecture was developed that showed how to improve accuracy while decreasing coverage. Data mining techniques and methods were initially created by Adamic and Adar [17] to reveal social networks and the exogenous elements underlying network structure. An analysis revealed that some variables were more indicative of community ties than others, and that these indications varied among user profiles. An algorithm for classifying entities and determining the reliability of their connections was tested and reviewed by Bilgic et al. [18]. In a research article [19], a narrative-specified probabilistic model of network improvement is provided and to develop a clean gradual learning algorithm for such models, which we subsequently employ to foretell associations between nodes. This research also identified a topological method for defining features in a real-time network.

Khadangi et al. [20] made an effort to compute the tie strength based on information about a person's tasks and their side views. They'd come up with a Facebook program specifically for harvesting profile data. To learn the ebb and flow of public data beneath the noses of supplemental networks, a knowledge architecture with a categorized feature selection approach was presented [21]. In order to recommend friends in a social bookmarking system, Manca et al. [22] developed detailed design models and architectural techniques. The employment of probabilistic approaches in conjunction with traditionally popular user pursuits has gained popularity over time. To assess the likelihood of two nodes co-occurring, Wang et al. [23] developed a unique stochastic visual approach that can be scaled to huge networks. The strength of direct relationships among social media users can be quantified thanks to a method proposed by Lin et al. [24]. In this study, a novel method is introduced, Trust Propagation Strategy, for gauging the quality of a relationship. Twitter-like social networks provided a coordinated approach to user feedback. By banding together, data sparsity can be reduced, and efforts can be directed on discovering the latent qualities of groups rather than individuals [25]. By taking into account not only the collection of profile statistics but also interaction affairs and tasks areas, Zhao et al. [26] established a generic framework for evaluating the strength of ties between different users. Based on the individual's task area selection and interaction practice, a method of gauging relationship strength was presented [27]. In order to predict the neighborhood extends over, a characteristic often linked with strong links, Ureña-Carrion et al. [28] focused on a large cellular phone dataset and assessed a variety of get-through-to-time sequence characteristics for each link. Recently, Perikos and Michael [29] conducted a thorough assessment on the evolution of proactively foreseeing the strength of relationships in online social networks.

Zhang et al. [30] explained structural equation modeling based on the S-S-O (Stressor-Strain-Outcome) theory framework to investigate the causes of social media burnout among Chinese WeChat users. According to the findings, factors including data deluge, privacy worries, and time commitment have favorable effects on the likelihood of experiencing social media burnout, whereas obsessive use has the opposite effect. Liu et al. [31] adopted meta-analysis and investigated potential factors that modify links between social media users and organizational performance in order to develop a more complete map of the relationship between the two. This study uses data from 65 empirical investigations with a sample size of 24,576 firms to try to break down SMU into its component parts, which include social marketing, social listening and monitoring, social communication, and social networking and cooperation. Ghorbanzadeh et al. [32] had analyzed the case study among hotel employees with respect to the connections between social media engagement, social capital, and job performance in Iran. In a recent survey, Khan et al. [33] polled 475 social media users to see if there is a correlation between privacy skepticism and dissatisfaction with and distrust of social media. Users' happiness with social media was found to be highly impacted by their skepticism about data privacy.

A lot of work has been contributed so far to measure the tie strength among online social media users. Different parameters, like activities performed by the individual, trust-based connections, sentiment of the shared information have been considered in the existing literature to generate the best possible way to enhance the tie strength. Despite of numerous existing literature and established theory to amplify tie strength, there is lacking behind of considering the combined factors like details of individual in a social media, the greatest number of connections having with those individual and weakly ties among the users, and the composition of the network etc. These salient factors have motivated us to construct a new idea and fruitful execution to compute the best possible relationship strength among social media users. In the next section, our proposed method has been elaborated in an efficient manner.

3. METHODOLOGY

Figure 1 depicts the workflow of our proposed approach. In our model, we have considered mainly three factors to compute the tie strength. We have considered ten numbers of different real time social media networks for our experiment. All the dataset consists of the profile values, strongly and weakly connected component and the structure of the network. We must first gather profile information, circles, and ego networks from all networks. One circle represents the real-time network's interconnectivity ('friend list').

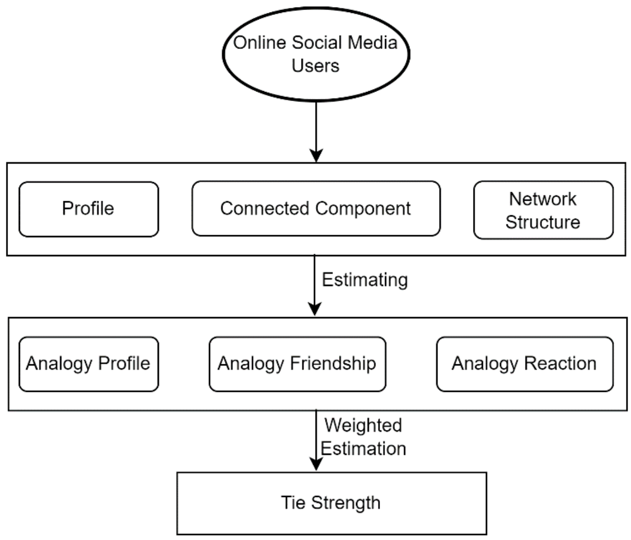

Figure 1. Work flow of the proposed model

3.1 Computation of tie strength

In our work, we used three parameters to calculate relationship strength among Facebook users: Analogy Profile, Analogy Friendship, and Analogy Reaction. Pearson Correlation was chosen as the Analogy Profile, Jaccard's coefficient as the Analogy Friendship, and User Interconnection Potency Level as the Analogy Reaction. Finally, we compute our relationship strength for Facebook users. Let's go through all three elements briefly in order to calculate the tie strength:

(1) Analogy Profile:

In our innovative method, we employed modified Pearson Correlation [34] as an Analogy Profile. In order to evaluate the chance of knot formation between the vertices v_i & v_j, the unification neighbourhood set, Uni_{ij} is defined as:

$$Uni_{ij}=\{p|(A_i[p]>null)otherwise(A_j[p]>null)\} \quad (1)$$

An appreciable association connecting the unification neighborhood set, Uni_{ij}, A_i and A_j specifies the greater constructional analogy between vertices i and j. To determine the association in the midst of two vertices, the association coefficient in the midst of the unification neighborhood set of vectors is determined. In our paper, the Pearson correlation coefficient is used as an analogy profile to calculate the correlation strength of two users on the social media network. The association in the midst of the unification neighborhood vectors set A_i and A_j is estimated as:

$$\mathrm{CR}_{ij} = \frac{\sum_{p\in Uni_{ij}}(A_i[p]-\overline{A_i})(A_j[p]-\overline{A_j})}{\sqrt{\sum_{p\in U_{ij}}(A_i[p]-\overline{A_i})^2}\sqrt{\sum_{p\in U_{ij}}(A_j[p]-\overline{A_j})^2}} \quad (2)$$

$\overline{A_i}$ is the mean standards in the unification neighbourhood vector set A_i and it can be estimated as: $\overline{A_i} = \frac{\sum_{p\in Uni_{ij}}(A_i[p])}{Uni_{ij}}$. Even if two nodes have no shared neighbours in our approach, they may have considerable structural similarities. As a result, a link may be identified by comparing their neighbours.

(2) Analogy Friendship:

Analogy Friendship is calculated using Jaccard's coefficient [35] in this paper. In general, two individuals in online social networks are more likely to be connected if they have the greatest number of common buddies. Relationship strength may therefore be measured based on mutual connections in a network. It is a typical similarity metric in data recovery that assesses whether or not both *p* and *q* create a feature *f* (arbitrarily picked feature). This approach yields the following measurement:

$$value(p;q) := \frac{|\gamma(p)\cap\gamma(q)|}{|\gamma(p)\cup\gamma(q)|} \quad (3)$$

(3) Analogy Reaction:

Here, we have used the User Interconnection Potency Level [36] as Analogy Reaction (AR) in our paper. Mostly two users on social media have the strongest association if they share useful information most of the time on a priority basis between themselves. Relationship strength may also be computed by measuring the interconnection potency level. Suppose two users x and y have dissimilar interconnection levels as A, B, C and D to constitute 'communicate', 'like', 'opinion', and 'onward' respectively. In both ways, communication is established between x and y. Then the analogy reaction (AR) is calculated as:

$$AR(x,y) = \frac{\min(A_{ij},A_{ji})+\min(B_{ij},B_{ji}) + min(C_{ij},C_{ji})+min(D_{ij},D_{ji})}{\max(A_{ij},A_{ji})+\max(B_{ij},B_{ji}) + max(C_{ij},C_{ji})+max(D_{ij},D_{ji})} \quad (4)$$

The logical principles span is [0,1] for Analogy Reaction (x, y), where 0 indicates fragile interconnection potency and 1 indicates secure interconnection potency.

(4) Computation of Tie Strength:

All three variables have been measured at this point. By giving more value to connections between people who are otherwise very different, we can determine the relative strength of the relationships in our real-world network. The desired tie strength can be found as shown below:

$$Tie\ Strength = \alpha * Analogy\_Profile + \beta * Analogy\_Friendship + \gamma * Analogy\_Reaction \quad (5)$$

Here, α, β, γ are weighted parameters and the logical principles span is [0,1] and the summation of α, β, and γ is 1.

If we consider about the limitations of Pearson Correlation, Jaccard's coefficient and User Interconnection Potency Level, it can briefly be described as:

- Pearson Correlation suggests that variables have a linear connection. Because links in social networks are frequently complex and nonlinear, Pearson's correlation may not completely convey the intricacies of tie strength.
- With binary data, where connections are either present or missing, Jaccard's coefficient works effectively. It is possible that it may not account for differences in tie strength within binary relationships.
- As a notion, User Interconnection Potency Level can be subjective and context dependent. Different individuals might assess tie strength differently, therefore correctly quantifying user perceptions may be difficult.

3.2 Evaluation matrices

Strong relationships are less prevalent in many social networks than weak ties, resulting in skewed data. When considering tie strength data, it is critical to account for this imbalance. Tie strength may be defined as the degree of overlap between people's interactions, interests, or behaviors. Our unconventional approach considers this overlap, making it crucial to assess how well it captures the shared characteristics between individuals. In measuring tie strength, we take an innovative technique that takes into account several characteristics such as interaction frequency, reciprocity, and the effect of reciprocal connections. We are looking for a complete review that takes into account the overall quality of tie strength forecasts.

Due to the above-mentioned rationale, Precision (P) [37], Recall (R), and the Dice Similarity Coefficient (DSC) were chosen as our evaluation metrics because of their relevance to tie strength measurement, suitability for dealing with imbalanced data, ability to assess overlap, ability to provide a holistic evaluation, and ease of interpretation. These measures provide a thorough evaluation of the accuracy and comprehensiveness of our novel technique to evaluate tie strength in social networks.

We have also compared our obtained P, R and DSC values with the other existing algorithms to check the performance of our proposed approach. The formulas are stated below:

$$Precision, P = \frac{|Accurate\ Pragmatic|}{|Accurate\ Pragmatic| + |Fake\ Pragmatic|} \quad (6)$$

$$Recall, R = \frac{|Accurate\ Pragmatic|}{|Accurate\ Pragmatic| + |Fake\ Contradiction|} \quad (7)$$

$$Dice\ Similarity\ Coefficient, DSC = 2 * \frac{P * R}{P + R} \quad (8)$$

3.3 Dataset

In this paper, the real time social media networks have been collected from SNAP [38] library. SNAP is renowned for its efficiency and scalability in handling large-scale network datasets. As our research involves the analysis of social network data, which can be extensive and complex, SNAP's ability to efficiently process and manipulate such data is invaluable. SNAP has extensive popularity and a thriving user and development community. This guarantees that it receives frequent updates, bug corrections, and access to a multitude of materials, making it a dependable and well-supported tool for our research requirements. The specific functions provided by SNAP are closely aligned with the objectives of our research. Its skills for network visualization, statistical analysis, and tie strength assessment are critical to fulfilling our research objectives. The SNAP library was chosen for data collecting and analysis because of its efficiency, scalability, broad algorithmic support, community support, and compatibility with the study aims. These benefits ensure that we may undertake thorough and extensive network analysis to draw significant insights and successfully contribute to the area.

Table 1. Statistics of the facebook network

Nodes	4039
Edges	88234
Nodes in largest WCC	4039 (1.000)
Edges in largest WCC	88234 (1.000)
Nodes in largest SCC	4039 (1.000)
Edges in largest SCC	88234 (1.000)
Average clustering coefficient	0.6055
Number of triangles	1612010
Fraction of closed triangles	0.2647
Diameter (longest shortest path)	8
90-percentile effective diameter	4.7

Table 2. Statistics of the Twitter network

Nodes	81306
Edges	1768149
Nodes in largest WCC	81306 (1.000)
Edges in largest WCC	1768149 (1.000)
Nodes in largest SCC	68413 (0.841)
Edges in largest SCC	1685163 (0.953)
Average clustering coefficient	0.5653
Number of triangles	13082506
Fraction of closed triangles	0.06415
Diameter (longest shortest path)	7
90-percentile effective diameter	4.5

Table 3. Statistics of the Google+ network

Nodes	107614
Edges	13673453
Nodes in largest WCC	107614 (1.000)
Edges in largest WCC	13673453 (1.000)
Nodes in largest SCC	69501 (0.646)
Edges in largest SCC	9168660 (0.671)
Average clustering coefficient	0.4901
Number of triangles	1073677742
Fraction of closed triangles	0.6552
Diameter (longest shortest path)	6
90-percentile effective diameter	3

Table 4. Statistics of the epinions network

Nodes	75879
Edges	508837
Nodes in largest WCC	75877 (1.000)
Edges in largest WCC	508836 (1.000)
Nodes in largest SCC	32223 (0.425)
Edges in largest SCC	443506 (0.872)
Average clustering coefficient	0.1378
Number of triangles	1624481
Fraction of closed triangles	0.0229
Diameter (longest shortest path)	14
90-percentile effective diameter	5

Table 5. Statistics of the wiki-vote network

Nodes	7115
Edges	103689
Nodes in largest WCC	7066 (0.993)
Edges in largest WCC	103663 (1.000)
Nodes in largest SCC	1300 (0.183)
Edges in largest SCC	39456 (0.381)
Average clustering coefficient	0.1409
Number of triangles	608389
Fraction of closed triangles	0.04564
Diameter (longest shortest path)	7
90-percentile effective diameter	3.8

Table 6. Statistics of the google web graph

Nodes	875713
Edges	5105039
Nodes in largest WCC	855802 (0.977)
Edges in largest WCC	5066842 (0.993)
Nodes in largest SCC	434818 (0.497)
Edges in largest SCC	3419124 (0.670)
Average clustering coefficient	0.5143
Number of triangles	13391903
Fraction of closed triangles	0.01911
Diameter (longest shortest path)	21
90-percentile effective diameter	8.1

Table 7. Statistics of the astro physics collaboration

Nodes	18772
Edges	198110
Nodes in largest WCC	17903 (0.954)
Edges in largest WCC	197031 (0.995)
Nodes in largest SCC	17903 (0.954)
Edges in largest SCC	197031 (0.995)
Average clustering coefficient	0.6306
Number of triangles	1351441
Fraction of closed triangles	0.1345
Diameter (longest shortest path)	14
90-percentile effective diameter	5

Table 8. Statistics of the amazon product network

Nodes	262111
Edges	1234877
Nodes in largest WCC	262111 (1.000)
Edges in largest WCC	1234877 (1.000)
Nodes in largest SCC	241761 (0.922)
Edges in largest SCC	1131217 (0.916)
Average clustering coefficient	0.4198
Number of triangles	717719
Fraction of closed triangles	0.09339
Diameter (longest shortest path)	32
90-percentile effective diameter	11

Table 9. Statistics of the livejournal social network

Nodes	4847571
Edges	68993773
Nodes in largest WCC	4843953 (0.999)
Edges in largest WCC	68983820 (1.000)
Nodes in largest SCC	3828682 (0.790)
Edges in largest SCC	65825429 (0.954)
Average clustering coefficient	0.2742
Number of triangles	285730264
Fraction of closed triangles	0.04266
Diameter (longest shortest path)	16
90-percentile effective diameter	6.5

Table 10. Statistics of the stanford web network

Nodes	281903
Edges	2312497
Nodes in largest WCC	255265 (0.906)
Edges in largest WCC	2234572 (0.966)
Nodes in largest SCC	150532 (0.534)
Edges in largest SCC	1576314 (0.682)
Average clustering coefficient	0.5976
Number of triangles	11329473
Fraction of closed triangles	0.002889
Diameter (longest shortest path)	674
90-percentile effective diameter	9.7

We have considered mainly ten different real time networks, namely Facebook, Twitter, Google+, Epinions, Wikipedia vote network, Google web graph, Astro Physics collaboration network, Amazon product co-purchasing network, LiveJournal social network, and Stanford web graph. The statistics of the nodes as well as edges present in the largest strongly connected components (SCC) and weakly connected components (WCC) are also available. The diversified values of the clustering coefficient can be seen among all the networks. All these real-world social media networks consist of 'friends list' that is indicated by circles. Survey participants provided all of the existing data in the networks. Profile information, 'circles', and ego networks are used to depict the datasets. By reinstating an individual user's internal id with the current merit, the existing data in these networks is identifiable. Furthermore, when feature vectors from datasets are provided, their meaning is obscured. The statistics of these networks are shown in Tables 1-10.

4. RESULTS AND DISCUSSION

The laboratory assessments were done using Python programming language in Jupyter notebook. The machine has a Windows 10 operating system and 8 gigabytes of random access memory.

The factors α, β, and γ have a considerable influence on the computation of Tie Strength, as well as the friend recommendation from the 'circles' of the real time networks mentioned above. There is a simple and effective method for adjusting its value to a tolerable range. Then the value 1 is respectively assigned to α, β, or γ while the other two parameters are assigned as 0 and compute the outcome into 3 categories. The outcome for Facebook network is shown in Table 11 below:

Table 11. Experimental results of precision, recall and DSC

α	β	γ	Precision	Recall	DSC
1	0	0	0.88	0.93	0.90
0	1	0	0.90	0.95	0.92
0	0	1	0.86	0.90	0.88

As per the values, we have achieved for DSC, α, β, and γ are recalculated as:

$$\alpha = \frac{0.90}{0.90 + 0.92 + 0.88} = 0.33 \tag{9}$$

$$\beta = \frac{0.92}{0.90 + 0.92 + 0.88} = 0.34 \tag{10}$$

$$\gamma = \frac{0.88}{0.90 + 0.92 + 0.88} = 0.33 \tag{11}$$

After applying the new value to the trial, the following change results in all evaluation indicators showing a significant improvement, as indicated in the table below. We tested our algorithm against TP-URS [24] and URSMF [8] to prove that the approach is effective. The tensile strength of synthetic ties has been calculated in TP-URS. The outcome in TP- URS is calculated by considering both direct and indirect links. TP-URS is based on trust propagation strategy and weight coefficients are assigned based on the number of direct and indirect connections. The limitations involved in this method include only two weight coefficients α and β that lead to computing the relationship strength. If the quantity of the flow of direct connections is much larger than the undirected connections, then it will impact the score of relationship strength. On the other hand, the quantity of undirected connections will increase the enhancement of relationship strength. Instead of evaluating real-world networks, the TP-URS experiment was run on a simplified weighted social network graph.

Relationship strength was calculated using URSMF's triangulation of three user-specific variables: degree of profile similarity, degree of friendship network structure similarity, and intensity of user interaction. The data was collected from the scholarly online social network SCHOLAT. Although it has been demonstrated that this approach examined using three factors but have performance limitations in terms of data loss and noise. Because the weightages assigned in the factors have the significant differences that will affect the score of relationship strength.

Our suggested method, on the other hand, evaluates three separate parameters α, β, and γ concurrently in a compatible manner and almost equal weights have been assigned to all the three factors after recalculation of DSC values these factors. We achieved the initial DSC values based on the precision and recall scores. In this way, data noises and losses may be drastically suppressed to provide a more accurate connection strength value, hence increasing the performance of the relationship strength among users in online social media. We were inspired to compare our results to the other two approaches because the criteria used are so close to those of our proposed methodology. Our results are tabulated in Table 12.

In the result shown in Table 12, we have noticed that our proposed method achieves efficient results in all the networks than TP-URS as well as URSMF. As only two factors impact to compute relationship strength in TP-URS, precision score is obtained lesser than the URSMF and our proposed approach. Also, it will affect to calculate the DSC value. Due to the insignificant distribution of influence factors, URSMF provides the lower precision score than our proposed approach. And it will affect to enhance the relationship strength score as it is based on DSC value. On the other hand, almost equal distribution of weightages in influence factors, our proposed approach executed efficient precision and recall values. That will lead to better DSC values. But parameter tuning is also important to balance the relationship strength among users. And we have considered that approach as our future work.

As we have considered ten different real time online social media, our proposed approach executed the best value in Google web graph dataset. Our proposed method has also performed better and generates efficient DSC values in other social media datasets. Our proposed approach provides the highest 92% precision value, 98% as the recall value and 95% as the DSC value in Google web graph network. Also, it has performed better than TP-URS at precision value as 81%, recall value 85% and DSC value as 83% in the same network. In URSMF, the highest precision value as 83% in Epinions and Google web graph, recall value as 88% in Epinions, and DSC value as 85% in both Epinions and Google web graph have been achieved. Overall, our proposed method has done well and outperformed other methods in our experiment.

Table 12. Experimental results after adjustments of parameters

Sl No.	Network Name	α	β	γ	TP-URS			URSMF			Proposed Method		
					Precision	Recall	DSC	Precision	Recall	DSC	Precision	Recall	DSC
1	Facebook	0.33	0.34	0.33	0.78	0.83	0.80	0.80	0.85	0.82	0.89	0.95	0.92
2	Twitter	0.35	0.33	0.32	0.80	0.84	0.82	0.81	0.87	0.84	0.88	0.93	0.90
3	Google+	0.32	0.34	0.34	0.75	0.81	0.78	0.79	0.84	0.81	0.91	0.97	0.94
4	Epinions	0.33	0.35	0.32	0.77	0.83	0.79	0.83	0.88	0.85	0.87	0.92	0.89
5	Wiki-Vote	0.34	0.33	0.33	0.73	0.80	0.76	0.81	0.87	0.83	0.90	0.96	0.93
6	Google web graph	0.32	0.33	0.35	0.81	0.85	0.83	0.83	0.87	0.85	0.92	0.98	0.95
7	Astro Physics collaboration network	0.31	0.34	0.35	0.79	0.83	0.81	0.82	0.86	0.84	0.91	0.96	0.93
8	Amazon product co-purchasing network	0.33	0.33	0.34	0.78	0.82	0.80	0.81	0.86	0.83	0.88	0.92	0.89
9	LiveJournal	0.34	0.32	0.34	0.80	0.85	0.82	0.80	0.84	0.82	0.89	0.93	0.91
10	Stanford web graph	0.35	0.32	0.33	0.77	0.81	0.79	0.79	0.84	0.81	0.88	0.93	0.90

5. CONCLUSION AND FUTURE WORK

A novel approach to identify relationship strength based on three authenticate factors has been presented in this paper. Ten different real time social media networks have been used for the experimental result. We have compared our findings with two popular methods, namely TP-URS and URSMF, in this paper. According to the obtained results, our proposed approach delivers efficient performances in all the datasets. Our proposed strategy, on the other hand, performed efficient accuracy and recall values due to almost equal distribution of weightages in influence factors. This resulted in higher DSC levels. This is how our approach has outperformed the other two methods. However, parameter adjustment is also necessary to balance the intensity of user relationships which is considered as the current limitations in our method. Our method simultaneously evaluates three distinct properties in a compatible manner, resulting in significantly reduced data noises and losses, a more accurate connection strength rating, and improved buddy recommendation system performance. The results showed that the proposed method excelled in practically all real-time social media datasets. The results provided by our proposed approach open a new window to enhance the relationship strength among the users in social media platforms. Although it is very difficult to compute relationship strength in a dynamic network running in real time as the numbers of users are increasing or decreasing in non-linear way. Our findings could be the potential impact on large social media datasets to compute relationship strength for any dynamic network. In this way, the real time connections in social media may be calculated and the monitoring activities of the administrators will also be efficient. However, our proposed approach requires parameter adjustment for improved performance. As a consequence, we wish to make the algorithm adaptive so that we may change the parameters automatically in future work. Applications of machine learning and deep learning algorithms may enhance the better result in case of computing relationship strength among online social media users. In the near future, we will try to apply our strategy on additional real-world networks to see how it compares to existing methods.

ACKNOWLEDGMENT

The work is carried out in the computer labs, Department of CSE, Assam University. We would like to thank all faculties and staff members for providing the opportunity to accomplish our work.

REFERENCES

[1] Hanneman, R.A., Riddle, M. (2005). Introduction to Social Network Methods. https://wiki.gonzaga.edu/dpls707/images/6/6e/Introduction_to_Social_Network_Methods.pdf.

[2] Hasan, M.A. (2016). Methods and applications of network sampling. In book: Optimization Challenges in Complex, Networked and Risky Systems, pp. 115-139.

https://doi.org/10.1287/educ.2016.0147

[3] Naderifar, M., Goli, H., Ghaljaie, F. (2017). Snowball sampling: A purposeful method of sampling in qualitative research. Strides in development of medical education, 14(3): 1-6. https://doi.org/10.5812/sdme.67670

[4] Herz, A., Petermann, S. (2017). Beyond interviewer effects in the standardized measurement of ego-centric networks. Social Networks, 50: 70-82. https://doi.org/10.1016/j.socnet.2017.01.003

[5] Granovetter, M. (1973). The strength of weak ties. American Journal of Sociology, 78(6): 1360-1380. https://www.jstor.org/stable/2776392.

[6] Manca, M., Boratto, L., Carta, S. (2015). Using behavioral data mining to produce friend recommendations in a social bookmarking system. In Data Management Technologies and Applications: Third International Conference, DATA 2014, Vienna, Austria, pp. 99-116. https://doi.org/10.1007/978-3-319-25936-9_7

[7] Shang, M.S., Zhang, Z.K., Zhou, T., Zhang, Y.C. (2010). Collaborative filtering with diffusion-based similarity on tripartite graphs. Physica A: Statistical Mechanics and its Applications, 389(6): 1259-1264. https://doi.org/10.1016/j.physa.2009.11.041

[8] Tang, F. (2017). Link-prediction and its application in online social networks. Doctoral dissertation, Victoria University. https://vuir.vu.edu.au/id/eprint/35048.

[9] Brauer, K., Sendatzki, R., Gander, F., Ruch, W., Proyer, R.T. (2022). Profile similarities among romantic partners' character strengths and their associations with relationship-and life satisfaction. Journal of Research in Personality, 99: 104248. https://doi.org/10.1016/j.jrp.2022.104248

[10] Kahanda, I., Neville, J. (2009). Using transactional information to predict link strength in online social networks. In Proceedings of the International AAAI Conference on Web and Social Media, 3(1): 74-81. https://doi.org/10.1609/icwsm.v3i1.13957

[11] Gilbert, E., Karahalios, K. (2009). Predicting tie strength with social media. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 211-220. https://doi.org/10.1145/1518701.1518736

[12] Viswanath, B., Mislove, A., Cha, M., Gummadi, K.P. (2009). On the evolution of user interaction in facebook. In Proceedings of the 2nd ACM Workshop on Online Social Networks, pp. 37-42. https://doi.org/10.1145/1592665.1592675

[13] Xiang, R., Neville, J., Rogati, M. (2010). Modeling relationship strength in online social networks. In Proceedings of the 19th International Conference on World wide web, pp. 981-990. https://doi.org/10.1145/1772690.1772790

[14] Srba, I., Bieliková, M. (2010). Tracing strength of relationships in social networks. In 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Toronto, ON, Canada, 3: 13-16. https://doi.org/10.1109/WI-IAT.2010.241

[15] Yanagimoto, H., Yoshioka, M. (2012). Relationship strength estimation for social media using Folksonomy and network analysis. In 2012 IEEE International Conference on Fuzzy Systems, Brisbane, QLD, Australia, pp. 1-8. https://doi.org/10.1109/FUZZ-IEEE.2012.6251238

[16] Nasir, S.U., Kim, T.H. (2020). Trust computation in online social networks using co-citation and transpose trust propagation. IEEE Access, 8: 41362-41371. https://doi.org/10.1109/ACCESS.2020.2975782

[17] Adamic, L.A., Adar, E. (2003). Friends and neighbors on the web. Social Networks, 25(3): 211-230. https://doi.org/10.1016/S0378-8733(03)00009-1

[18] Bilgic, M., Namata, G.M., Getoor, L. (2007). Combining collective classification and link prediction. In Seventh IEEE International Conference on Data Mining Workshops (ICDMW 2007), Omaha, NE, USA, pp. 381-386. https://doi.org/10.1109/ICDMW.2007.35

[19] Kashima, H., Abe, N. (2006). A parameterized probabilistic model of network evolution for supervised link prediction. In Sixth International Conference on Data Mining (ICDM'06) Hong Kong, China, pp. 340-349. https://doi.org/10.1109/ICDM.2006.8

[20] Khadangi, E., Zarean, A., Bagheri, A., Jafarabadi, A. B. (2013). Measuring relationship strength in online social networks based on users' activities and profile information. In ICCKE 2013, Mashhad, Iran, pp. 461-465. https://doi.org/10.1109/ICCKE.2013.6682863

[21] Lu, Z., Savas, B., Tang, W., Dhillon, I.S. (2010). Supervised link prediction using multiple sources. In 2010 IEEE International Conference on data Mining, Sydney, NSW, Australia, pp. 923-928. https://doi.org/10.1109/ICDM.2010.112

[22] Manca, M., Boratto, L., Carta, S. (2015). Friend recommendation in a social bookmarking system: Design and architecture guidelines. In Intelligent Systems in Science and Information 2014: Extended and Selected Results from the Science and Information Conference 2014, London, UK, pp. 227-242. https://doi.org/10.1007/978-3-319-14654-6

[23] Wang, C., Satuluri, V., Parthasarathy, S. (2007). Local probabilistic models for link prediction. In Seventh IEEE international conference on data mining (ICDM 2007) Omaha, NE, USA, pp. 322-331. https://doi.org/10.1109/ICDM.2007.108

[24] Lin, X., Shang, T., Liu, J. (2014). An estimation method for relationship strength in weighted social network graphs. Journal of Computer and Communications, 2(4): 82. http://dx.doi.org/10.4236/jcc.2014.24012

[25] Zhao, G., Lee, M.L., Hsu, W., Chen, W., Hu, H. (2013). Community-based user recommendation in uni-directional social networks. In Proceedings of the 22nd ACM international conference on Information & Knowledge Management, pp. 189-198. https://doi.org/10.1145/2505515.2505533

[26] Zhao, X., Yuan, J., Li, G., Chen, X., Li, Z. (2012). Relationship strength estimation for online social networks with the study on Facebook. Neurocomputing, 95: 89-97. https://doi.org/10.1016/j.neucom.2011.06.036

[27] Tao, W., Ju, C., Xu, C. (2020). Research on relationship strength under personalized recommendation service. Sustainability, 12(4): 1459. https://doi.org/10.3390/su12041459

[28] Ureña-Carrion, J., Saramäki, J., Kivelä, M. (2020). Estimating tie strength in social networks using temporal communication data. EPJ Data Science, 9(1): 37. https://doi.org/10.1140/epjds/s13688-020-00256-5

[29] Perikos, I., Michael, L. (2022). A Survey on Tie Strength Estimation Methods in Online Social Networks.

ICAART (3): 484-491. https://doi.org/10.5220/0010845100003116
[30] Zhang, Y., He, W., Peng, L. (2022). How perceived pressure affects users' social media fatigue behavior: A case on WeChat. Journal of Computer Information Systems, 62(2): 337-348. https://doi.org/10.1080/08874417.2020.1824596
[31] Liu, Z., Geng, R., Tse, Y.K.M., Han, S. (2023). Mapping the relationship between social media usage and organizational performance: A meta-analysis. Technological Forecasting and Social Change, 187: 122253. https://doi.org/10.1016/j.techfore.2022.122253
[32] Ghorbanzadeh, D., Khoruzhy, V.I., Safonova, I.V., Morozov, I.V. (2023). Relationships between social media usage, social capital and job performance: the case of hotel employees in Iran. Information Development, 39(1): 6-18. https://doi.org/10.1177/0266666921103055
[33] Khan, M.I., Loh, J.M., Hossain, A., Talukder, M.J.H. (2023). Cynicism as strength: Privacy cynicism, satisfaction and trust among social media users. Computers in Human Behavior, 142: 107638. https://doi.org/10.1016/j.chb.2022.107638
[34] David, N. (2014). Chapter 6—Selection of Variables and Factor Derivation. Commercial data Mining: Processing, Analysis and Modeling for Predictive Analytics Projects, 79-104. https://doi.org/10.1016/B978-0-12-416602-8.00006-6
[35] Liben-Nowell, D., Kleinberg, J. (2003). The link prediction problem for social networks. In Proceedings of the Twelfth International Conference on Information and Knowledge Management, pp. 556-559. https://doi.org/10.1145/956863.956972
[36] Llorena, J.M., Roldán, X.A. (2007). Skc: Measuring the Users Interaction Intensity. Computers and Education: E-learning. From Theory to Practice, pp. 123-132. https://doi.org/10.1007/978-1-4020-4914-9_11
[37] Goutte, C., Gaussier, E. (2005). A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In European Conference on Information Retrieval, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 345-359. https://doi.org/10.1007/978-3-540-31865-1_25
[38] Leskovec, J. (2010). Stanford network analysis project. http://snap.stanford.edu/.