

Research Article

Correlation between Triadic Closure and Homophily Formed over Location-Based Social Networks

Nauman Ali Khan ¹, Wuyang Zhou ¹, Mudassar Ali Khan ², Ahmad Almogren ³,
and Ikram Ud Din ²

¹Key Laboratory of Wireless-Optical Communication, University of Science and Technology of China, Hefei 230027, China

²Department of Information Technology, The University of Haripur, Haripur 22620, Pakistan

³Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11633, Saudi Arabia

Correspondence should be addressed to Wuyang Zhou; wyzhou@ustc.edu.cn

Received 15 January 2021; Revised 25 January 2021; Accepted 29 January 2021; Published 15 February 2021

Academic Editor: Habib Ullah Khan

Copyright © 2021 Nauman Ali Khan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Social Internet of Things (SIoT) is a variation of social networks that adopt the property of peer-to-peer networks, in which connections between the things and social actors are automatically established. SIoT is a part of various organizations that inherit the social interaction, and these organizations include industries, institutions, and other establishments. Triadic closure and homophily are the most commonly used measures to investigate social networks' formation and nature, where both measures are used exclusively or with statistical models. The triadic closure patterns are mapped for actors' communication behavior over a location-based social network, affecting the homophily. In this study, we investigate triads emergence in homophilic social networks. This evaluation is based on the empirical review of triads within social networks (SNs) formed on Big Data. We utilized a large location-based dataset for an in-depth analysis, the Chinese telecommunication-based anonymized call detail records (CDRs). Two other openly available datasets, Brightkite and Gowalla, were also studied. We identified and proposed three social triad classes in a homophilic network to feature the correlation between social triads and homophily. The study opened a promising research direction that relates the variation of homophily based on closure triads nature. The homophilic triads are further categorized into transitive and intransitive groups. As our concluding research objective, we examined the relative triadic throughput within a location-based social network for the given datasets. The research study attains significant results highlighting the positive connection between homophily and a specific social triad class.

1. Introduction

Homophily identifies the groups of individuals who are socially connected based on shared interests or behaviors. In the past decades, numerous sociologists premeditated clusters of people based on various sociocommunity parameters, including gender, religion, race, place of living, and work. These parameters were used to infer various relations like close friends, coworkers, life-partner, and other social associations. Based on these social parameters and their similarities, few broad applications include user mobility, influencing, and segregation. With the rapid growth of

communication networks, quantifying accurate homophily analysis is one of the most critical social network analysis (SNA) problems that is further subcategorized as triadic closure analysis and home location detection analysis. One of the fundamental challenges in detecting homophily is when a person with versatile personality features tends to change his behavioral pattern dynamically. Traditional techniques commonly use the clustering method to exploit and predict the reasons for a homophilic nature. For the scenario mentioned above, these techniques lack accuracy and precision when a social network accommodates diverse multiprofessional users having a dense structure.

Regarding detection applications of triadic closures and homophily, scientists also contributed to various application areas besides automation and network traffic management. These include refinement of recommendation systems, fake user identification, analysis of micro blogging, detection of natural disasters using real-time Twitter Big Data, business decision making, and healthcare systems [1–5]. Companies and businesses increase revenues and improve goodwill by maintaining their micro blogging systems. Machine learning algorithms extract meaningful information and help fetch the most related information, which helps in decision making [6]. In the literature, a great effort was made to gather the information related to a particular category of people on Facebook [7–9]. Aral and Walker identified the group of people on Facebook which were easier to influence. Their principal findings involve that young people are easier to influence in contrast to older generation people. Likewise, males have a more influential nature as compared to females. Similarly, other influential patterns were recognized in cross-gender comparisons. However, married people were categorized as the category which can get influenced [10].

A triadic closure in social networks can be interpreted as a communication group of precisely three individuals. Trio/triangle/triad is considered to be the necessary foundation of a social network. In literature, some modern research studies political campaigns, religious activities, organizational professionalism, web mining, and many more social networks based on such three-people subgraph [11]. Listing and counting of triads in a social network are considered triad census using the subgraph method of graph theory [7, 12, 13]. The clustering coefficient, a robust graph theory method, highlights the degree of nodes likely to be part of a cluster. A higher degree of the coefficient indicates a higher ratio of triads in a social network. One research also highlighted the positive correlation between the triads and community structures. Research findings reflect that community structures were coherent where the number of triads is remarkably high [14–17].

Social triad analysis in a multicluster environment helps to overcome the mentioned problem. Origins of dyads and triads in the social network encourage exploiting the homophilic nature further, specifically when the triad nodes belong to two different groups [18]. Generically, a triad is a group of three socially connected individuals in a social network, also referred to as the smallest group of that social network.

Triadic closure and homophily are two separate social network analysis evaluation measures. Applications of triadic closure and homophily involve friend recommendation systems, online social blogging services, community influence systems, and structural and informational construction systems. It further enhances learning systems, improves competition, and also increases work performance [19]. Previously, these evaluation measures were used individually to assist the above-mentioned issues and areas. In this research, we found a strong association between these measures and proposed a technique which uses these measures together.

In our research, we also explored the patterns of homophily in the multidomain social network. We took a

sample of the call detail records (CDRs) dataset and constructed a social network graph. In large-scale dataset of CDRs, each record is represented in the following format.

Caller ID	Call type	Callee ID	Time
Duration	LAC ID	CELL ID	

We constructed a social network using telecommunication-based anonymized call detail records and two openly available location-based social network datasets, similar to the work of Brightkite and Gowalla represented in [20]. Distinct caller ID is considered a distinct social network's user, and communication between two callers is considered a social tie. For every user, one home location is selected from various locations depending upon the maximum number of incoming and outgoing calls. Furthermore, we have identified the users' triads, in which all users belong to a shared home location. Figure 1 illustrates social triads' formation by variant home location of individuals in a social network. According to Figure 1, a standard social network is illustrated; each node is represented with v_v while each home location is represented with HL_{hl} . There is a scenario in which several triad nodes belong to a shared home location, such as v_2 , v_3 , and v_4 triad belonging to HL_2 . Our research identifies the origins of triadic closure in a homophilic network and proposes a classification model that creates subclasses into three groups.

In this study, our contribution relates to the proper classification of the triads, which is discussed as follows:

- (i) We first studied the user mobility patterns and their diversity by observing the entropy. We developed a social network graph of users and identified home location using home detection algorithm from the datasets.
- (ii) Based on home locations, we grouped users and critically observe their interconnections. Furthermore, we identified the homophilic patterns formed inside the social network.
- (iii) We investigated the origins of social triads in detail and examined the formation of triads. Based on the analysis, we categorized social triads and compared their behaviors within the homophilic social network. Interestingly, we found positive correlations between the homophily coefficient and a subset of social triads discussed in the relevant section.
- (iv) In the later part of the research, we organized homophilic triads into transitive and intransitive groups, and we examined the effect of categorized triads with the network's throughput.

The rest of the article is organized as follows: Section 2 describes the literature review. Section 3 presents the problem formulation and evaluation measures. Section 4 introduces the triadic closure in the homophilic environment and its effect on homophily. Section 5 describes the datasets and observations. Section 6 explains the results and their discussion. Section 7 concludes with future recommendations.

2. Literature Review

A social network is generally composed of three artifacts, i.e., user description, social connection direction, and communication contents exchanged over the social network [21]. The user-based artifact study explores the user's behavior in different scenarios and environments [22]. Individual personal networks are the social network sub-graphs that identify all the communication behavior of a single entity [23]. Individual personal networks have a transitive tendency, i.e., a friend of a friend is also a friend, as discussed by [24]. Transitivity is the propensity that two people, who are not direct friends to each other but have a familiar mutual friend, may also become friends over time [16, 25]. Researchers analyze the reason for triads' formation, why a dyad converts to a triad with time, and how, in a three-person small network, all the users want to reduce the hesitation discrepancies [18, 26]. In an unbalanced triad social network, where two different people like one person, but these two people do not like each other, this creates emotional tension between them, forcing the relationship to be complete and consistent, or discourages the triad formation [27]. According to a comprehensive survey, it was consistently observed that transitivity exists in about 70% to 80% of various small groups [28–30]. In another research study, the effect of gender was highlighted, and it was revealed that the formation of triads in boys is more common than in girls [31]. One other study compared homogeneous behavior of users with heterogeneous environment actors, and it was concluded that heterogeneous actors are less transitive concerning religion, race, and education than homogeneous actors [32, 33]. A study highlights the baseline of triads forming; trust plays a vital role in making the relationships more robust and balanced [34]. While establishing and building new ties, people may have hidden or apparent interests such as knowledge sharing and a social relationship like friendship, educational purpose, and scientific collaboration [35]. Moreover, an existing study shows the positive correlation between authorship sharing and research-based relationship building that spreads over time [36].

Online location-based social networking applications enable the users to build social ties based on location [37–39]. In addition to social connection details, a social network formed over a location-based application may have extra attached information such as location ID [35]. Similar to location-based social networks, CDRs (call detail records) datasets are the log files of users reordered over time. These logs include the details of user communications and the attached information of location ID. As per our literature exploration, many researchers used this location ID to draw the homophily of the social networks [37, 40, 41]. A study examined existing location-based human mobility trend evaluation techniques and categorized them into mainly three classes, i.e., user, place, and trajectory-based modeling [42–44].

Homophily refers to a social grouping concept where people with common interests tend to morph into a single group [45]. In literature, homophily is broadly based on two

approaches, i.e., *induced* and *choice* homophily [46]. The combined effect of social triads is observed with homophily, and it is determined that choice homophily plays a vital role in building observed homophily [47]. Research findings also illustrated that making triads within homophilic regions is statically higher [47].

To summarize, triad creation and critical exploration in a social network help to understand social relationships that further assist in many applied areas already discussed. In literature, many research contributions have been conducted to exploit social triads for various aspects, though there is a need to further understand how location information can affect social triads and homophily.

3. Problem Formulation and Evaluation Measures

The formulation of the problem is stated as follows. Let $G = (V, E)$ be a graph representing a static social network of users and their communication links, where $V = \{v_1, v_2, \dots, v_{|V|}\}$ is a set of actors/users in a social network and $E \subset V \times V$ is a set of social links between users. $e_{ij} \in E$ shows the existence of a communication link between v_i and v_j users. Let $T = \{\Delta = (v_i, v_j, v_k) | v_i, v_j, v_k \in V\}$ be a set of triads.

Definition 1. (CT: closed triads). Let $CT = \{\Delta = (v_i, v_j, v_k) | \Delta \in T \wedge e_{ij}, e_{ik}, e_{jk} \in E\}$ be the set of closed triads.

Definition 2. (OT: open triads). Let $OT = \{\Delta = (v_i, v_j, v_k) | \Delta \in T \wedge e_{ij}, e_{ik} \in E \wedge e_{jk} \notin E\}$ be the set of open triads.

Definition 3. (HL: user home location). Let $L = \{l_1, l_2, \dots, l_{|L|}\}$ is a set of locations, where l_n denotes a distinct location. Let $HL = \{h_1, h_2, \dots, h_{|V|}\}$ be a set of user home locations, where h_n denotes a home location for user v_n . $h_n = \text{Home Location}(v_n, L)$

According to the location-based social network, every user forms a social connection at a specific location. For v_n , the function $\text{Home Location}(v_n, L)$ identifies one location from L as home location h_n based on home location algorithm stated in [48].

Definition 4. ($\Theta A, \Theta B, \Theta C$: types of triads).

For $\Delta = (v_i, v_j, v_k)$, let

$$\begin{aligned} \Theta A &= \{\Delta | \Delta \in CT \wedge h_i = h_j = h_k\}, \\ \Theta B &= \{\Delta | \Delta \in CT \wedge h_i = h_j \wedge h_i \neq h_k\}, \\ \Theta C &= \{\Delta | \Delta \in CT \wedge h_i \neq h_j \neq h_k\}. \end{aligned}$$

Definition 5. (ψ : homophily coefficient).

$$\psi = \{\psi_{xy} | \psi_{xy} = \text{Homophily}(v(h_x), v(h_y))\}, \quad (1)$$

where $\begin{cases} h_x, h_y \in HL \\ h_x \neq h_y \end{cases}$.

Let $\psi =$ be a set of homophily, where ψ_{xy} denotes homophily of graph for two sets of vertices. $v(h_n)$ denotes a set of all the vertices belonging to h_n home location. Function

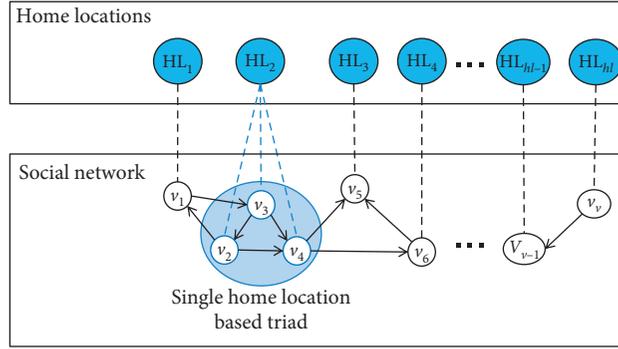


FIGURE 1: Formation of triad closure based on home locations.

Homophily ($v(h_x), v(h_y)$) takes two sets of vertices, i.e., $v(h_x)$ and $v(h_y)$, and initially counts the cross-home location edges $e_{v(h_x),v(h_y)} \forall e_{v(h_y),v(h_x)}$ as p and non-cross-home location edges $e_{v(h_x),v(h_x)} \forall e_{v(h_y),v(h_y)}$ as q . Then, it finds the expected cross-home location edges as $\xi = ((p + q)/2)$. After that, the homophily coefficient is calculated using the following equation [49].

$$\psi_{xy} = 1 - \frac{\xi}{p}. \quad (2)$$

Correlation Coefficient. Correlation coefficient among $\Theta A, \Theta B, \Theta C$ types of triads and homophily is defined in

$$r(\psi, \Theta) = \frac{\sum(\psi - \bar{\psi})(\Theta - \bar{\Theta})}{\sqrt{\sum(\psi - \bar{\psi})^2} \sqrt{\sum(\Theta - \bar{\Theta})^2}} \quad (3)$$

4. Social Triads in Location-Based Social Networks

A social network is the communication graph among many users. Datasets such as telecom call logs or location-based social network data have the details of the user's interaction and a hint of location information. Each record of the datasets represents a time-stamped location-based social link between two users in communication.

4.1. Triadic Closure Property in Homophilic Environment. Triadic closure refers to the communication of three nodes. Every closed triad can be either transitive or intransitive, depending upon the type of communication occurring [50]. Each node of the triads belongs to one specific location, treated as its home location. The location of home for each user or node is identified using the *home detection algorithm* [48]. While critically examining the formation of the closed triad, we identified and hence proposed three cases of triads, listed as follows:

- (1) All users of the triad belong to the same home locations

- (2) Any two triad users belong to one home location, and the remaining user belongs to any other home location
- (3) All users of the triad belong to three different home locations

Figure 2 states an example of a social network based on a CDRs subdataset. In this figure, each hexagon shows a region of the telecommunication signal cell. A social network over the cellular signal region represents a communication graph, and each cell is considered as a home location of inside nodes. The green-colored hexagon is taken as a reference cellular signal region in the stated example, and other red-colored hexagons are considered out location cellular signal region. As described before, these three triad classes are also illustrated in Figure 2.

We named the three possible triads as Class A, Class B, and Class C for differentiation and further exploration. Our research first investigates each class, classifies it into transitive triads or intransitive triads, and then examines all possible combinations of social triads in a directed graph. Figure 3 illustrates a detail overview of all possible triads and defines them into three classes. Code underneath each triad represents the category, and the naming convention of the social triad is explained in [51]. However, we improvise the category and naming convention by adding an alphabet at the start of the code as a class name and by adding an extra digit as its variant. In the code $B210A1$, B is the class name, $210A$ is the existing naming convention, and 1 is the variation number.

4.2. Accumulative Homophily in Triadic Closure. Call detail records (CDRs) and online location-based social networks have extra associated information, i.e., location ID. In our research, we incorporated the location ID into identified homophily in a network. We utilize the existing home detection algorithm to identify the home location for each user [48]. In location-based social networks, by home location, we mean the most visited and stayed at place. The algorithm identifies one location out of all visited places as a home location. Further, we measure the correlation between the three classes of triads and homophily.

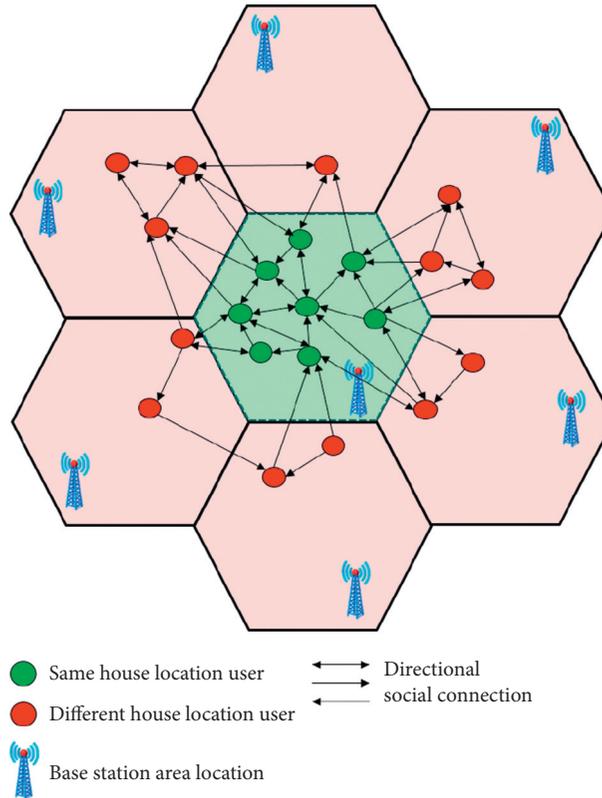


FIGURE 2: Illustration of social triads in location-based networks.

A triad is a group of three nodes, in which each node belongs to specific home locations. However, homophily is calculated based on only two groups. Initially, we calculate homophily using (2), and then we averaged them. For three home locations, e.g., h_x , h_y , and h_z , accumulative homophily is measured, as stated in

$$\text{Acc}_{(\psi_{xyz})} = \frac{\psi_{xy} + \psi_{xz} + \psi_{yz}}{3}. \quad (4)$$

5. Datasets Characteristics and Observations

5.1. Data Description. In support of research, we incorporated one large call detail record (CDR) and two online location-based datasets, i.e., Gowalla and Brightkite [20]. The CDR dataset used in this study is provided by a Chinese mobile telecommunication company. The dataset contains 702,000 subscribers along with user demographic information. The data is logged over the period of one year, which has more than half a billion social ties.

Brightkite and Gowalla are openly available location-based social network datasets [20, 52]. Both datasets are gathered using the online social networking website. Websites maintain user check-in data by fetching mobile GPS location data. These services create an environment that enables people to build a social connection with nearby people. The Brightkite dataset contains 58,228 nodes and 214,078 edges, and Gowalla contains 196,591 nodes and 950,327 edges. In the data cleaning phase, we removed missing or wrong data types and empty

rows. In the CDRs dataset, each record is represented as in the following column format.

Caller ID	Call type	Callee ID	Time
Duration	Call type	LAC ID	CELL ID

5.2. Observations. Call duration is one of the key attributes of the calling dataset. While mining the CDR dataset and investigating the social networks, we observed some interesting call duration facts. Figure 4 shows the relation of call duration and number of calls. We found two big spikes in the number of calls according to the call duration. We have found that the maximum number of calls has call duration in the range of either 10 to 30 seconds or 1 min to 2 min. This observation infers that people mostly prefer to have a short duration communication to convey their message. One research shows that direct calls are a kind of strong communication and are considered the baseline for the strong ties [53].

CDR logs contain another important item, i.e., the location ID attribute, which identifies the area from which the call was made. Initially, we applied the home location algorithm and inferred the home location based on the call logs, and then we segregated all users according to the location ID. Figure 5 shows the distribution of users based on location ID.

We carefully monitored the communication behavior of the people within each location.

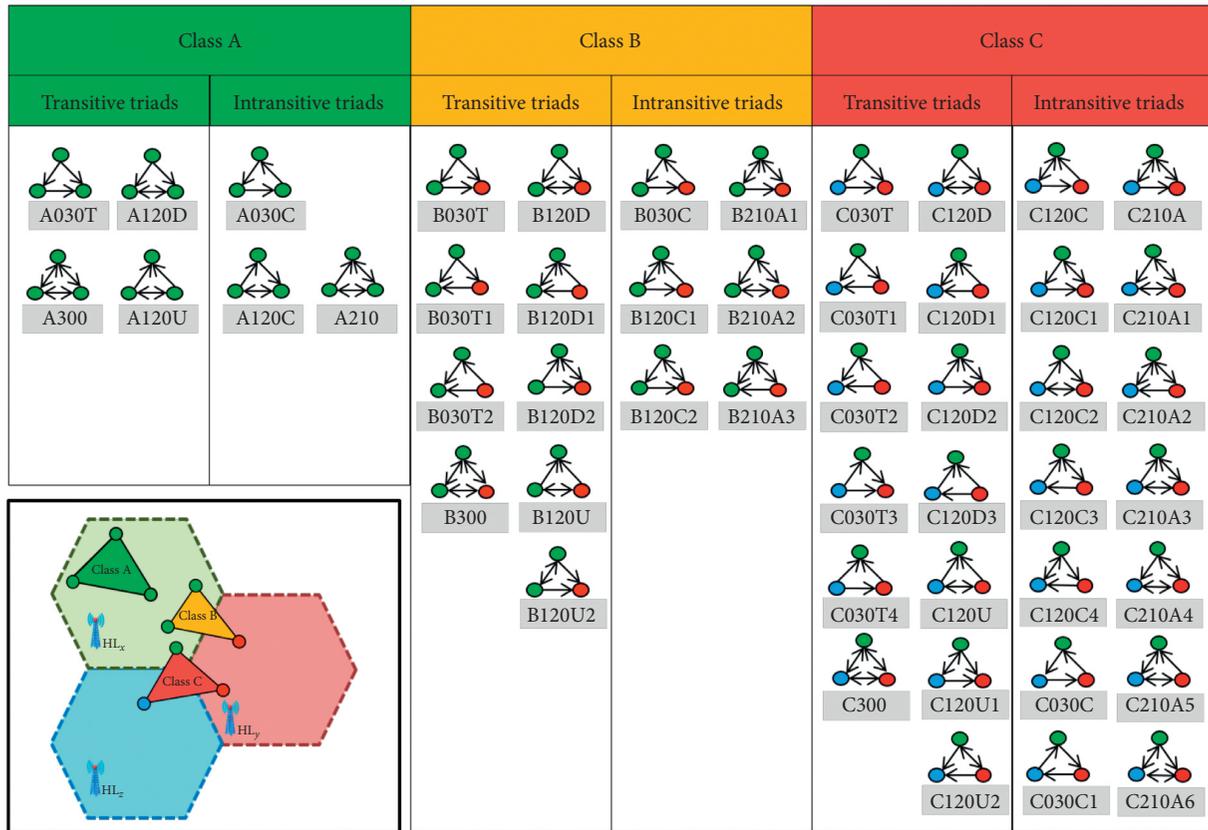


FIGURE 3: A fine-grained classification of social triads in location-based homophilic networks with all variations.

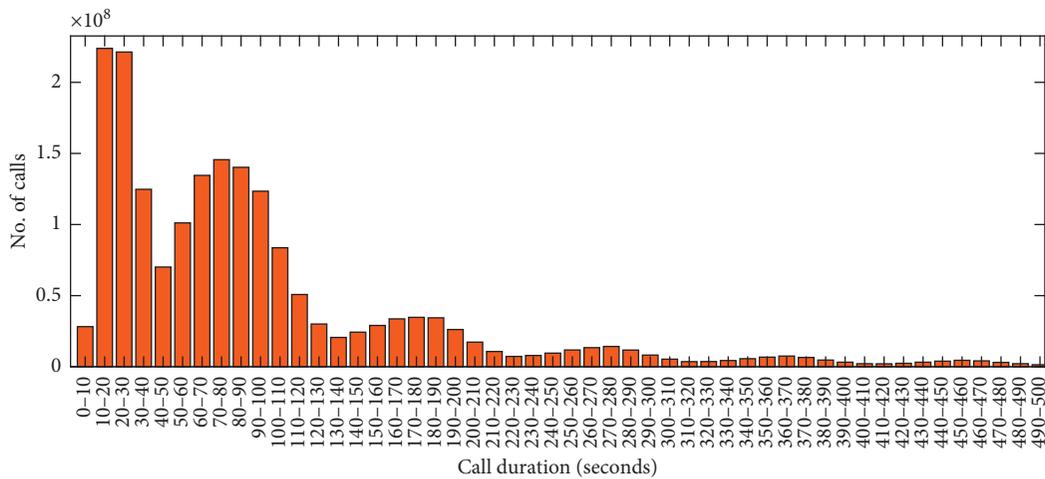


FIGURE 4: Segregation of number of calls in comparison to call duration.

During fact extraction, we found a high ratio of calls between people at the same location in comparison to that of different localities. Figure 6 is a preview of communications taking place for different locations or within the same location. Location-based cross-communication infers homophily

which is based on location, which is the key motivation aspect for this study. Figure 6 shows that the interaction taking place between people from the same location is more than that between people from different locations, which further indicates the existence of location-based homophily. This

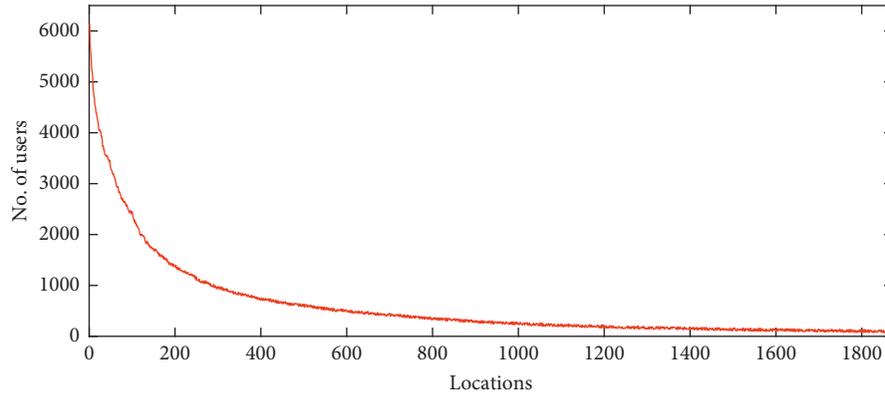


FIGURE 5: Location-based user density compiled through home location detection algorithm.

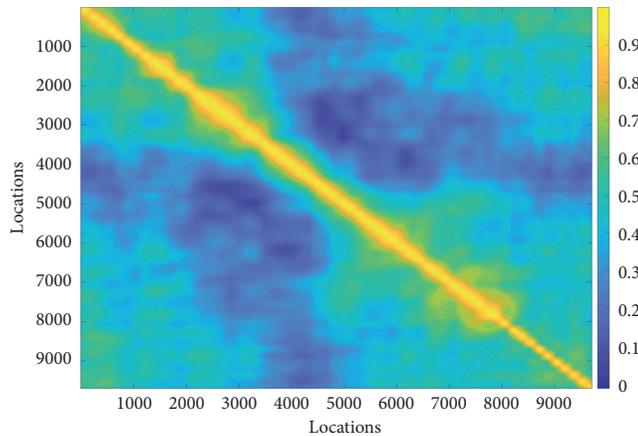


FIGURE 6: Visualization of homophily for intralocation-based user communication.

further adheres to the fact that there is a strong connection between location-based homophily and triadic social closure.

6. Results and Discussion

Our research evaluation results classify the empirical social triads into three groups based on the strong correlation between homophilic networks and social triads. We found a positive correlation between the homophily and a specific class of triads. Our findings indicate that people having the same home location are more likely to form a triad.

In this study, we incorporated two location-based large datasets and one close source CDR dataset. Figure 7 illustrates nine correlation comparisons, three for CDR, Brightkite, and Gowalla datasets. Results show the correlation between homophily and classes of triads. The y -axis shows the percentage of homophily, and the x -axis refers to the number of triads in percentage. Results shown in Figure 7 reveal that the accumulative homophily between the groups has a positive correlation with Class A triads. Simultaneously, Class A refers to a group of users triad having a common home location.

We initially measured the number of triads for all the three classes of the datasets and observed that the minimum

quantity for a triads can be individually calculated from each category. A sum of 2,200 triads was found for Class C. For the understanding of results and normalization, we randomly selected 2,000 triads for the three classes. Results show that higher homophily corresponds to a higher number of social triads from Class A. However, the impact of homophily related to Class B and Class C is comparatively unspecific. A consistency of positive correlation was observed in all the three datasets between homophily percentage and triads of Class A.

The regression coefficient r of the correlation was examined using (3). From the comparisons between all datasets and Class A, we found the highest value for the regression coefficient of r . Besides high regression coefficient values r and consistency, our research also discovers all results' closeness, especially for the CDR dataset.

In the analysis, we found the maximum observations of homophily within the range of 25% to 80%, and the cross-relation between Class A and homophily highlights the maximum observation of triads in the range of 5% to 70%. All the three datasets produce symmetric and positive regression trend results. The regression coefficient $r = 0.61$, $r = 0.65$, and $r = 0.55$ is measured for CDR, Brightkite, and Gowalla dataset, respectively. The r value denotes the

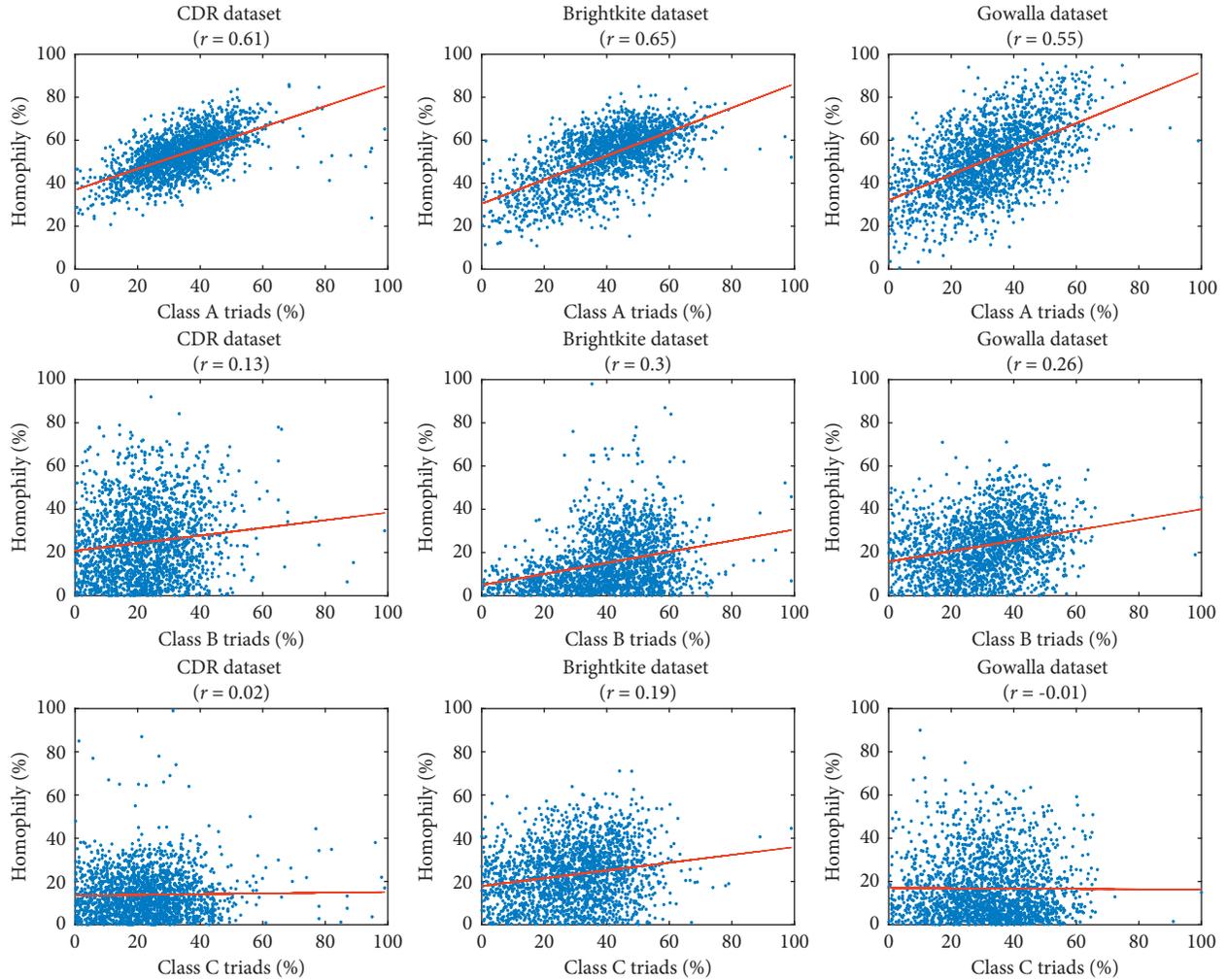


FIGURE 7: Correlation between the three classes of triadic closure and homophily using location-based datasets.

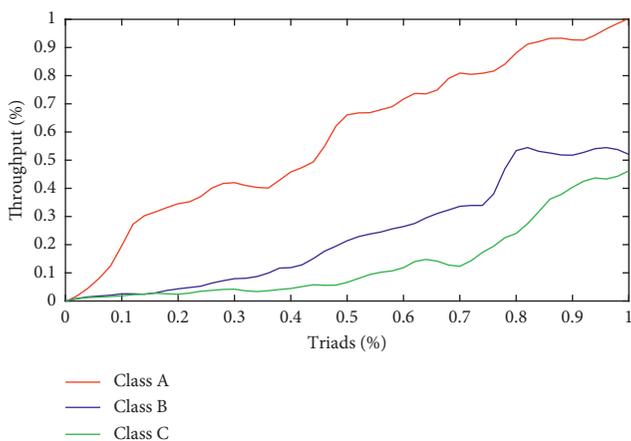


FIGURE 8: The relative throughput of triadic closure for the three classes.

existence of cause and effect relationship between the triadic closure and homophily, especially between Class A and homophily.

In the second phase of evaluation, we measured the accumulative throughput for Class A, B, and C in all the datasets. Figure 8 shows the overall throughput for the three datasets; the y -axis shows throughput percentage and the x -axis shows the number of triads in percentage. The throughput (T) is measured using (5). We used a relative throughput measure to cross-relate the results. The lowest and the highest values of the throughput were taken as reference values, and then accordingly the rest of the graph was plotted.

$$T = \frac{\text{calls made by triads}}{\text{total calls made by triads}} \times 100. \quad (5)$$

In this study, we observed that Class A triads consume the maximum amount of bandwidth. We encountered a significant rise in the throughput for Class A after 40%, which shows that people with a higher number of triads of the same home location also exchange a higher number of calls as shown in Figure 8. However, we came across the least throughput for Class B and Class C within the range of 1% to 50%. The lower values of throughput indicate the least communication among the triad users.

The throughput of Class C is comparatively less than that of Class B because all the three users of Class C were in different home locations. However, Class B, having any two users from a common home location, explains the slight increase in its throughput. This study highlights the higher throughput of Class A as compared to the rest of the classes. The results indicate that triads formed between people from the same home location have more communication rates than triads formed at different home locations.

7. Conclusion and Future Work

Triadic closure and homophily coefficient are the two mutually exclusive merits required to understand the behavior of the social network. In this study, we found the cause and effect relationship between the homophily and triad closure for the social networks formed based on the location. We have closely observed social triads' formation in a homophilic social network and found interesting relationships between them. Our study used Chinese telecommunication-based anonymized call detail records (CDRs) and two openly available location-based social network datasets, Brightkite and Gowalla. This research identifies three sets of social triad classes in a homophilic network and expresses the correlation between social triads and homophily. Examination findings opened a novel direction of measuring homophily based on multiple types of social triads. Based on the communication directions, we further organized homophilic triads into a transitive and intransitive group. In the last part of the research, we also examined the effect of a specific triadic class on a network's throughput. We will investigate the reasons for the formation of transitive and intransitive classes in homophilic networks in the future.

Data Availability

The data used can be found at <http://snap.stanford.edu/data/index.html#locnet>.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this work.

Acknowledgments

This work was supported by King Saud University, Saudi Arabia, through research supporting project number RSP-2021/184. Nauman Ali Khan acknowledges the support of the Chinese Government and Chinese Scholarship Council (CSC) for his Ph.D. studies at the University of Science and Technology, China. This research work was partially supported by Key Program of National Natural Science Foundation of China (Grant number 61631018).

References

- [1] X. Luo, C. Jiang, W. Wang, Y. Xu, J.-H. Wang, and W. Zhao, "User behavior prediction in social networks using weighted extreme learning machine with distribution optimization," *Future Generation Computer Systems*, vol. 93, pp. 1023–1035, 2019.
- [2] S. Nazir, S. Khan, H. U. Khan et al., "A comprehensive analysis of healthcare big data management, analytics and scientific programming," *IEEE Access*, vol. 8, pp. 95714–95733, 2020.
- [3] F. Masood, A. Almogren, A. Abbas et al., "Spammer detection and fake user identification on social networks," *IEEE Access*, vol. 7, pp. 68140–68152, 2019.
- [4] B. Amina and T. Azim, "Scanceplens: a framework for automatic lexicon generation and sentiment analysis of micro blogging data on China pakistan economic corridor," *IEEE Access*, vol. 7, pp. 133876–133887, 2019.
- [5] V. Gupta and R. Hewett, "Real-time tweet analytics using hybrid hashtags on twitter big data streams," *Information*, vol. 11, no. 7, p. 341, 2020.
- [6] S. Pouyanfar, S. Sadiq, Y. Yan et al., "A survey on deep learning: algorithms, techniques, and applications," *ACM Computing Surveys (CSUR)*, vol. 51, no. 5, pp. 1–36, 2018.
- [7] Q. Gong, Y. Chen, X. He et al., "DeepScan: exploiting deep learning for malicious account detection in location-based social networks," *IEEE Communications Magazine*, vol. 56, no. 11, pp. 21–27, 2018.
- [8] A. Zrnc and D. Lavbič, "The role of social connections in plagiarism detection," in *Proceedings of the International Workshop on Learning Technology for Education in Cloud*, pp. 54–63, Springer, Maribor, Slovenia, August 2015.
- [9] S. Ali, N. Islam, A. Rauf, I. Din, M. Guizani, and J. Rodrigues, "Privacy and security issues in online social networks," *Future Internet*, vol. 10, no. 12, p. 114, 2018.
- [10] S. Aral and D. Walker, "Identifying influential and susceptible members of social networks," *Science*, vol. 337, no. 6092, pp. 337–341, 2012.
- [11] L. Becchetti, P. Boldi, C. Castillo, and A. Gionis, "Efficient algorithms for large-scale local triangle counting," *ACM Transactions on Knowledge Discovery from Data*, vol. 4, no. 3, pp. 1–28, 2010.
- [12] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, vol. 8, Cambridge University Press, Cambridge, UK, 1994.
- [13] Z. Ali, M. A. Shah, A. Almogren, I. Ud Din, C. Maple, and H. A. Khattak, "Named data networking for efficient iot-based disaster management in a smart campus," *Sustainability*, vol. 12, no. 8, p. 3088, 2020.
- [14] M. E. Newman, "Properties of highly clustered networks," *Physical Review E*, vol. 68, no. 2, Article ID 026121, 2003.
- [15] K. Warren, B. Campbell, S. Cranmer et al., "Building the community: endogenous network formation, homophily and prosocial sorting among therapeutic community residents," *Drug and Alcohol Dependence*, vol. 207, Article ID 107773, 2020.
- [16] N. Muyinda, J. M. Baetens, B. De Baets, and S. Rao, "Using intransitive triads to determine final species richness of competition networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 540, Article ID 123249, 2020.
- [17] D. V. Foster, J. G. Foster, P. Grassberger, and M. Paczuski, "Clustering drives assortativity and community structure in ensembles of networks," *Physical Review E*, vol. 84, no. 6, Article ID 066117, 2011.
- [18] H. Huang, Y. Dong, J. Tang, H. Yang, N. V. Chawla, and X. Fu, "Will triadic closure strengthen ties in social networks?" *ACM Transactions on Knowledge Discovery from Data*, vol. 12, no. 3, pp. 1–25, 2018.
- [19] S. Khan, S. Nazir, and H. Khan, "Smart object detection and home appliances control system in smart cities," *Computers, Materials and Continua*, vol. 67, pp. 895–915, 01 2021.

- [20] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1082–1090, San Diego, CA, USA, August 2011.
- [21] J. Tang, Y. Chang, and H. Liu, "Mining social media with social theories," *ACM SIGKDD Explorations Newsletter*, vol. 15, no. 2, pp. 20–29, 2014.
- [22] F. Amato, A. Castiglione, A. De Santo et al., "Recognizing human behaviours in online social networks," *Computers & Security*, vol. 74, pp. 355–370, 2018.
- [23] M. M. Tadesse, H. Lin, B. Xu, and L. Yang, "Personality predictions based on user behavior on the facebook social media platform," *IEEE Access*, vol. 6, pp. 61959–61969, 2018.
- [24] M. S. Granovetter, "The strength of weak ties," *American Journal of Sociology*, vol. 78, no. 6, pp. 1360–1380, 1973.
- [25] S. Peng, Y. Zhou, L. Cao, S. Yu, J. Niu, and W. Jia, "Influence analysis in social networks: a survey," *Journal of Network and Computer Applications*, vol. 106, pp. 17–32, 2018.
- [26] I. U. Din, A. Almogren, M. Guizani, and M. Zuair, "A decade of internet of things: analysis in the light of healthcare applications," *IEEE Access*, vol. 7, pp. 89967–89979, 2019.
- [27] D. Krackhardt and M. Kilduff, "Whether close or far: social distance effects on perceived balance in friendship networks," *Journal of Personality and Social Psychology*, vol. 76, no. 5, pp. 770–782, 1999.
- [28] C. McMillan and D. Felmlee, "Beyond dyads and triads: a comparison of tetrads in twenty social networks," *Social Psychology Quarterly*, vol. 83, no. 4, pp. 383–404, Article ID 0190272520944151, 2020.
- [29] D. T. Robinson and J. W. Balkwell, "Density, transitivity, and diffuse status in task-oriented groups," *Social Psychology Quarterly*, vol. 58, no. 4, pp. 241–254, 1995.
- [30] D. Kretschmer, L. Leszczensky, and S. Pink, "Selection and influence processes in academic achievement-more pronounced for girls?" *Social Networks*, vol. 52, pp. 251–260, 2018.
- [31] G. Kossinets and D. J. Watts, "Empirical analysis of an evolving social network," *Science*, vol. 311, no. 5757, pp. 88–90, 2006.
- [32] H. Schäfer, "Relationality and heterogeneity: transitive methodology in practice theory and actor-network theory," in *Methodological Reflections on Practice Oriented Theories*, pp. 35–46, Springer, Berlin, Germany, 2017.
- [33] X. Han, S. Cao, Z. Shen et al., "Emergence of communities and diversity in social networks," *Proceedings of the National Academy of Sciences*, vol. 114, no. 11, pp. 2887–2891, 2017.
- [34] K. D. Doekhie, M. M. H. Strating, M. Buljac-Samardzic, and J. Paauwe, "Trust in older persons: a quantitative analysis of alignment in triads of older persons, informal carers and home care nurses," *Health & Social Care in the Community*, vol. 27, no. 6, pp. 1490–1506, 2019.
- [35] N. A. Khan, S. Zhang, W. Zhou, A. Almogren, I. U. Din, and M. Asif, "Inferring ties in social iot using location-based networks and identification of hidden suspicious ties," *Scientific Programming*, vol. 2020, Article ID 6667610, 16 pages, 2020.
- [36] Z. Hu, A. Lin, and P. Willett, "Identification of research communities in cited and uncited publications using a co-authorship network," *Scientometrics*, vol. 118, no. 1, pp. 1–19, 2019.
- [37] J. Luo, A. P. Sinha, and H. Zhao, "Location-sensitive friend recommendations in online social networks," in *Proceedings of the 2020 Pacific Asia Conference on Information Systems*, p. 155, Dubai, UAE, June 2020.
- [38] N. Bibi, M. Sikandar, I. Ud Din, A. Almogren, and S. Ali, "IoT-based automated detection and classification of leukemia using deep learning," *Journal of Healthcare Engineering*, vol. 2020, Article ID 6648574, 12 pages, 2020.
- [39] B. Liao, Y. Ali, S. Nazir, L. He, and H. U. Khan, "Security analysis of iot devices by using mobile computing: a systematic literature review," *IEEE Access*, vol. 8, pp. 120 331–120 350, 2020.
- [40] K. M. Kumar and B. Srinivasan, "Point-of-interest based classification of similar users by using support vector machine and status homophily," *International Journal of Machine Learning and Computing*, vol. 9, no. 5, pp. 615–620, 2019.
- [41] S. Guha and S. B. Wicker, "Do birds of a feather watch each other? homophily and social surveillance in location based social networks," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pp. 1010–1020, Vancouver, BC, Canada, March 2015.
- [42] E. Toch, B. Lerner, E. Ben-Zion, and I. Ben-Gal, "Analyzing large-scale human mobility data: a survey of machine learning methods and applications," *Knowledge and Information Systems*, vol. 58, no. 3, pp. 501–523, 2019.
- [43] G. Jadoon, I. Ud Din, A. Almogren, and H. Almajed, "Smart and agile manufacturing framework, a case study for automotive industry," *Energies*, vol. 13, no. 21, p. 5766, 2020.
- [44] M. A. Khan, S. Israr, A. S. Almogren, I. U. Din, A. Almogren, and J. J. Rodrigues, "Using augmented reality and deep learning to enhance taxila museum experience," *Journal of Real-Time Image Processing*, pp. 1–12, 2020.
- [45] M. Yohsuke, J. Hang-Hyun, T. János, K. János, and K. Kimmo, "Structural transition in social networks: the role of homophily," *Scientific Reports (Nature Publisher Group)*, vol. 9, no. 1, 2019.
- [46] D. Cepić and Ž. Tonković, "How social ties transcend class boundaries? Network variability as tool for exploring occupational homophily," *Social Networks*, vol. 62, pp. 33–42, 2020.
- [47] A. Asikainen, G. Iñiguez, J. Ureña-Carrión, K. Kaski, and M. Kivelä, "Cumulative effects of triadic closure and homophily in social networks," *Science Advances*, vol. 6, no. 19, Article ID eaax7310, 2020.
- [48] Y. Gu, Y. Yao, W. Liu, and J. Song, "We know where you are: home location identification in location-based social networks," in *Proceedings of the 2016 25th International Conference on Computer Communication and Networks (ICCCN)*, pp. 1–9, IEEE, Waikoloa, HI, USA, August 2016.
- [49] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: homophily in social networks," *Annual Review of Sociology*, vol. 27, no. 1, pp. 415–444, 2001.
- [50] D. Doran, H. Alhazmi, and S. S. Gokhale, "Triads, transitivity, and social effects in user interactions on facebook," in *Proceedings of the 2013 Fifth International Conference on Computational Aspects of Social Networks*, pp. 68–73, IEEE, Fargo, ND, USA, August 2013.
- [51] P. W. Holland and S. Leinhardt, "Local structure in social networks," *Sociological Methodology*, vol. 7, pp. 1–45, 1976.
- [52] J. Leskovec and A. Krevl, "SNAP datasets: stanford large network dataset collection," 2014, <http://snap.stanford.edu/data>.
- [53] J.-P. Onnela, J. Saramäki, J. Hyvönen et al., "Structure and tie strengths in mobile communication networks," *Proceedings of the National Academy of Sciences*, vol. 104, no. 18, pp. 7332–7336, 2007.