國立臺灣大學社會科學院社會學系

學士班學生論文

Department of Sociology

College of Social Science

National Taiwan University

Bachelor’s Thesis

叢聚與分隔：比較主動與中介者觀點下的三角連結閉合

Clustering but also separating: comparing triadic closure from the perspective of the initiator and the intermediator

劉羽芯

Liou, Yu-Shin

指導教授：李宣緯 博士

Advisor: Hsuan-Wei, Lee Ph.D.

中華民國 112 年 4 月

April, 2023

國立臺灣大學學士班學生論文
口試委員會審定書

叢聚與分隔：
比較主動與中介者觀點下的三角連結閉合
Clustering but also Separating:
Comparing Triadic Closure From The Perspective of The Initiator and The Intermediator

本論文係劉羽芯君（B08305056）在國立臺灣大學社會學系完成之學士班學生論文，於民國 112 年 3 月 30 日承下列考試委員審查通過及口試及格，特此證明

口試委員(3 位)：

____________________________（簽名）

（指導教授）

________________ ________________

________________ ________________

系 主 任 ：________________________（簽名）

（是否須簽章依各院系規定）

誌謝

本文由中央研究院主題計畫支持（自我核心完整網絡：由接觸觀點整合社會網絡分析兩大研究傳承；編號 AS-TP-109-H02）感謝傅仰止老師，如果沒有老師的經濟支持，我能夠專心做研究、享受大學生活的時間想必會少了許多。也十分感謝明宜老師每次主持會議的辛苦，研究的過程中能有老師的溫柔提點，真的是十分幸運的事情。也感謝國賢、孟瑢老師在口試時給予的建言與鼓勵，開設了學士論文寫作的仲恩老師，生活忙碌卻依然抽空陪我們討論。感謝修課的大家，從來沒有抱怨過這個研究很無聊，願意打起精神聽我數三角形，每一個的回饋對我來說都十分珍貴。感謝我的好朋友珀瑞在社會統計上的大力協助，彥儀在社會學理論上的敦促，聰明又勤奮的廷雅，每次跟你聊完都想打開檔案再改一次，感謝紫葳以及孜頤的討論跟陪伴，還有晧庭、以及我親愛又可靠的姊姊岱宜，一直以來在技術上的支援，沒有你們，我不會有勇氣打開電腦科學的潘朵拉之盒，而依然相信希望的存在。

感謝社會所的麒文、禹文，至承，聽著大家的報告，我才漸漸知道研究應該是什麼樣子，或者原來還可以是這個樣子，祝福大家未來一切順利。為了避免一路寫到偶爾拜訪助理室的倉鼠球球，最後要感謝的是我的指導老師宣緯，老師對我總是充滿信心，很多時候比我對我自己有的還要多、太多了，常常讓我有點緊張，不過老師說我行，所以我行，明白這個道理以後漸漸的也學會了很多東西。我明白學術是一條漫長的路，而學士論文只是這個研究的一個小小中繼站，往後還有許多的路要走，雖然充滿了不確定性，但我也相信老師，相信最後我們都能得到一個好結果。最後也祝老師身體健康，鐵人三項可以拿到好成績。

2023.04.21 劉羽芯

中文摘要

區域叢集（local clustering）在社會網絡的三角連結閉合當中扮演了重要的角色。人們經常透過共同朋友去認識新朋友，與自己朋友的朋友成為朋友，這在社會網絡當中稱為三角連結閉合（triadic closure）。本文著重探討三角連結閉合的機制，解答為什麼網絡當中某些三角連結最終走向閉合，有些卻沒有。我們主張三角連結的閉合，與三角連結所鑲嵌的結構有關，是三角連結當中各個行動者的區域叢集交互作用下的結果。本文資料取自 26 個科系，919 位大學部學生，大一到大三完整的社群媒體使用紀錄，實證資料展示了班級網絡當中三角連結的形成，以及閉合的過程。本文的主要發現是，如果一位行動者尚未形成緊密的朋友圈，即處於低區域叢集的狀況底下，那他會更傾向於主動發起接觸，促成三角連結的閉合。尤其是針對那些同樣也尚未形成緊密的朋友圈，或者是受歡迎程度高的行動者。但對於已經形成緊密朋友圈的行動者，他們不傾向主動接觸朋友圈疏離的行動者，除此之外，他們的舊朋友也更不容易主動接觸新加入朋友圈的新朋友。總體來說，行動者的低區域叢集有益於行動者與朋友的朋友接觸，而兩人之間中介者的區域叢集則有抑制三角連結閉合的效果。本文主要的研究貢獻在於我們區分了區域叢集如何在不同的狀況底下促進，或者抑制三角連結的閉合。

關鍵字：區域叢集係數、社會網絡、三角連結閉合

ABSTRACT

Local clustering plays an essential role in social network triadic closure. People become friends with a friend of their friend. This is known as triadic closure in network research. However, it is still unclear why some of the triads remain unclosed in the network while others eventually close under the effect of local clustering. In this article, I argued that the effect of local clustering does not work independently. Instead, tie formation between actors is the outcome of the interaction effect between the local clustering coefficient of actors and the intermediator between them. Using college student social media usage records, I investigated the open triad formation and closure process. The main finding of this research is that if an actor has not formed a close social circle (i.e., has a low local clustering coefficient), they are more likely to initiate contact to close the triad, especially with those who also have low local clustering coefficients and popular actors, even if those popular alters were less likely to be attracted by them. This indicates that the initiator's local clustering coefficient promotes triadic closure. However, if a person has already formed a closed social circle, their old friend would be less likely to initiate contact with his new friend. This represents the intermediator's local clustering coefficient, which suppresses triadic closure. The main contribution of this research is identifying how local clustering promotes or suppresses triadic closure in different situations.

CONTENTS

誌謝 i

中文摘要 ii

ABSTRACT iii

CONTENTS iv

Chapter 1 Introduction 1

1.1 The puzzle of clustering formation and consistency 1

1.2 Identify two distinct perspectives of triadic closure 5

1.3 The rival theory of clustering: inequality of popularity 9

1.4 The empirical limitation of obtaining structural information 11

Chapter 2 Data and methods 14

2.1 Data 14

2.2 Build up ego-centric networks 15

2.3 Survival model 16

2.4 Dependent variables: define the event occurring and survival time 16

2.5 Control variables 20

2.6 Independent variables 20

2.7 Description statistics 23

Chapter 3 Results 29

3.1 Initiator perspective model 30

3.2 Intermediator perspective model 34

Chapter 4 Discussion 39

REFERENCES 42

Chapter 1 Introduction

Clustering is a common feature of social networks. It refers to the tendency for a group of nodes to have a relatively high density of connections within the group but a relatively low density of connections to nodes outside the group. Clustering exists in various social networks, including the small group in class, the clique in congress, and industrial clustering. There has been a wide range of research in network analysis, focusing on the impact of network clustering (Ahuja, 2000; Bearman & Moody, 2004; Sorenson & Audia, 2000; Tsai, 2001) and the application of clustering (Fortunato, 2010; Girvan & Newman, 2002; Schaeffer, 2007) such as graph theory, developing network detection algorithms and machine learning tasks. The existence of clustering indicated that the tie formation between a pair of nodes is not only related to the actors themself but also to the level of clustering of the structure they embedded. For example, if two of my friends already belong to a close social circle, they might be less likely to contact each other. In this article, I investigated how the tie formation between a pair of actors was related to the structure they embedded.

1.1 The puzzle of clustering formation and consistency

Transitivity and homophily are the main mechanisms behind clustering formation in current explanations. Homophily, like the old saying "birds of a feather flock together," indicates that actors tend to cluster together according to their similarities. (McPherson et al., 2001; Verbrugge, 1977). For example, people of the same gender, racial group, or age are likelier to become friends (González et al., 2007; Moody, 2001; Shakya et al., 2017;

Shrum et al., 1988; Smith et al., 2014). Homophily describes clustering as a phenomenon occurring at the dyad level and only considers actors and the attribute of their interactions.

Moreover, transitivity outlines the impact of a third party in the tie formation process. The central concept of transitivity is a triad. Triad is a set of three actors, the most basic form of a group in human society (Simmel, 1950). I decompose a network into a series of triads; if links in two dyads already exist, the third pair will also tend to form a tie (Rapoport, 1953). More intuitively, transitivity can be interpreted as "A friend of my friend is also my friend." It has been proven to exist among Twitter user friendship networks(Lou et al., 2013). In sociological explanation, having mutual friends indicates that two individuals have a higher possibility in the same social circle; from a psychological perspective, people tend to hold a positive attitude toward those with a positive relationship with their friends (Cartwright & Harary, 1956; M. Granovetter, 1992; M. S. Granovetter, 1973; Heider, 1946). This process is called triadic closure. (Figure 1-1 A) On the contrary, if the third edge remains inexistent, the triad is called the unclosed triad, as shown in Figure 1-1 B The forbidden triad might cause cognitive imbalance and be considered to affect an actor negatively. For example, having an unclosed triad in a personal friendship network significantly increases the suicide possibility of female adolescents. (Bearman & Moody, 2004)

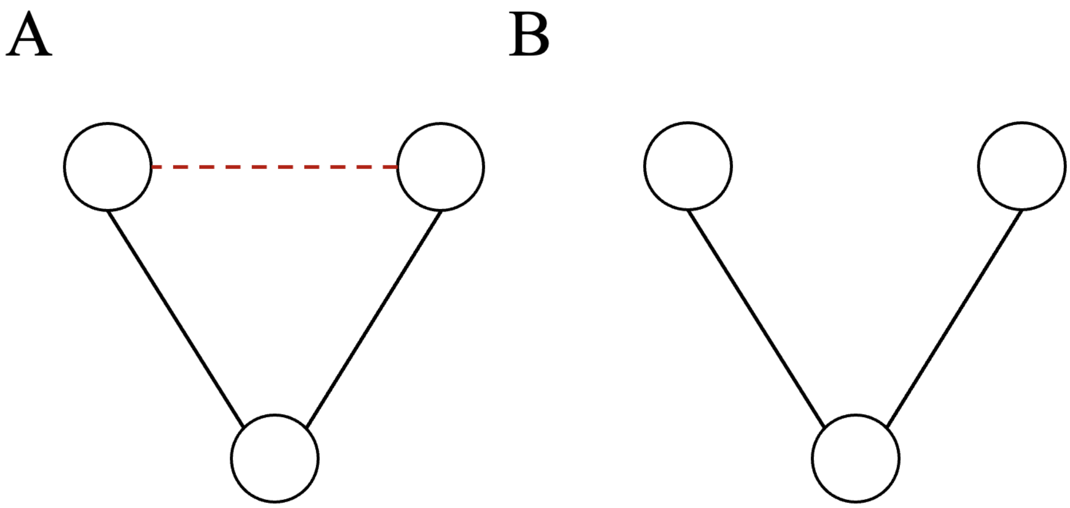

Figure 1-1 Each line represents the relationship between nodes. (A) Triadic closure: The tie formation in the third edge (red line) is called triadic closure. (B) Unclosed triad: The third edge in a triad reminds un-existed.

Social exchange theory might explain the consistency of clustering, which posits that clustering is the outcome of people's tendency to reduce potential risks in interaction. Interactions among actors are the result of conscious calculations. Actors prefer to maintain links that maximize return on investment (Blau, 2017). In a high-density network, actors would have more common neighbors, leading to increased information flow and decreased costs of supervision and sanction (Sorenson & Audia, 2000). People tend to share the same values (M. Granovetter, 1992) and hold collective norms. Hence, interacting with actors in the same clustering is less risky.

Further, generalized exchange, which increases social trust and solidarity, is more likely to occur within the clustering. Actors in the same clustering would have common neighbors. In the generalized exchange, the reward resource contributors receive is not directly from the recipient but from the third party. (Yamagishi & Cook, 1993) Common neighbors play a mediated role in the generalized exchange and foster its occurrence. Empirical studies have shown that the density of the inter-organization network is positively associated with the transfer of complex knowledge. Therefore, the high network density might increase innovation implications (Hansen, 1999; Obstfeld, 2005; Sorenson et al., 2006; Tsai, 2001). Conversely, the sparsity in the intra-firm collaborative network negatively affects organizational development.

However, the explanations provided by homophily and transitivity theory are controversial. Homophily suggests that people tend to choose friends based on similarities, but individuals may have multiple attributes such as gender, race, and interests and can find some degree of similarity with others in certain contexts. (Levine & Kurzban, 2006) While it can be advantageous to maintain a high-density network, reaching out to people beyond one's immediate circle can also be advantageous. This concept is exemplified in the theory of structural holes, which refers to how an individual embedding in a network might provide certain advantages or disadvantages. (Burt, 1985, 2004) The theory of structure hole describes the lack of tie between entities. In clustering networks, the lack of connection between clusters, or an unclosed triad, can be seen as a structural hole. A bridge position in the structural hole (Figure 1-2) increased an actor's information nexus and empowers them to control information flow. As a result, brokerage between clusters could become social capital for the actor, enabling the actor to develop diverse strategies. An empirical study of a worker network in a US electronics company founds that workers in a bridge position were more likely to learn alternative behavior or opinion. Therefore, they were more likely to come up with a good idea (Burt, 2004).

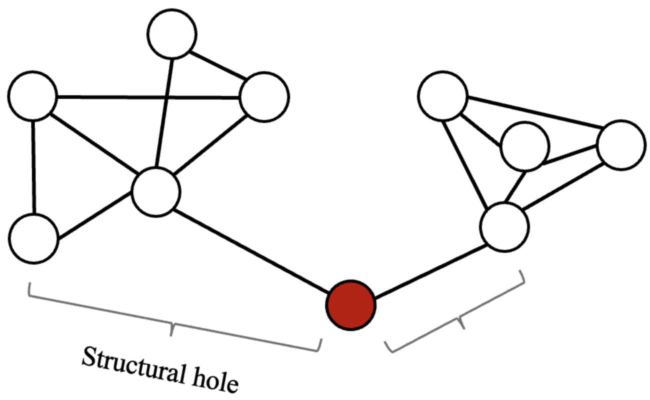

Figure 1-2 Structural hole

In this viewpoint, both inter- and intra-clustering links provide valuable information, although they differ in the type of information they provide. Complex knowledge transfers more easily within clusters, whereas links between clusters increase information heterogeneity. Therefore, if the tie formation process is based on people's conscious calculus, as the exchange theory suggests, people should also tend to contact the actor in distinct clusters. Similarly, while transitivity can explain the origin of clusters, it cannot explain their consistency. According to balance theory (Cartwright & Harary, 1956; Heider, 1946), actors near structural holes should tend to close forbidden triads to prevent cognitive imbalance. The structural hole should not continue to exist. In summary, if “a friend of a friend tends to be a friend,” why doesn’t the network evolve into a completely connected network rather than leaving unclosed triads that form structural holes?

1.2 Identify two distinct perspectives of triadic closure

To solve this puzzle, I reconsider the process of triadic closure. The typical way to quantify clustering is the local clustering coefficient. It is calculated by looking at the number of ties between the neighbors of a given node and dividing it by the number of possible connections between them (Barabási & Albert, 1999; Price, 2011; Stephen & Toubia, 2009). This measurement observes triadic closure from the perspective of the intermediary node in a triad (Figure 1-3 A). However, the theoretical explanation of triadic closure is actually from the initiator's perspective, which observes triadic closure from the position of the endpoint in the triad (Figure 1-3 B). In social exchange theory, it is believed that actors tend to contact actors with more common neighbors than themselves. Similarly, transitivity explains an actor's incentive to contact a friend of their friend.

There are two perspectives in the process of triadic closure: The initiator perspective observes the actors directly related to the triadic closure tie. On the other hand, the intermediator perspective observes triadic closure from the center position, as shown in Figure 1-3 B. It is a subtle difference in triadic closure. The former shows the triadic closure between the focal person (ego) and an alter in an ego-centric network crossing a length-2 path (degree-2 alter), while the latter is the closure between two of the ego's degree-1 alters. Instead of directly participating in the triadic closure process, the ego plays the intermediator role in the triadic closure from an intermediator's perspective. The distinction between "I become a friend with a friend of my friend" and "two of my friends become friends with each other" lies in the initiative.

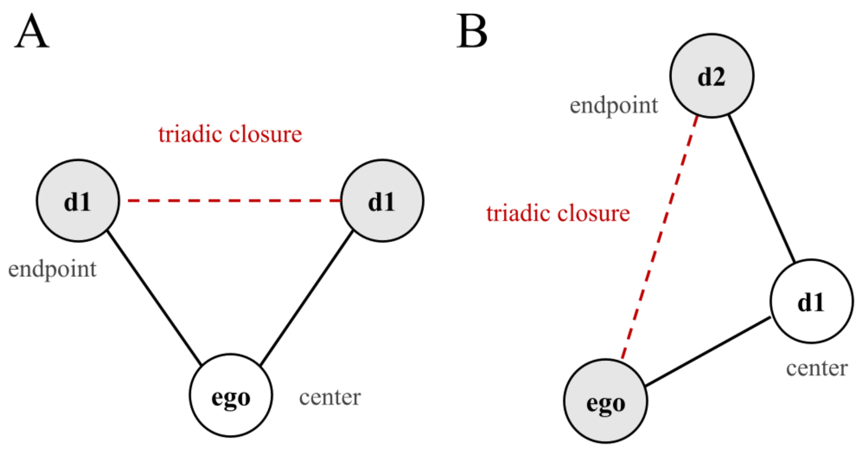

Figure 1-3 (A) The intermediator perspective triadic closure: tie formation between two of ego's degree-1 alters. The degree-1 alters ("d1") in the endpoint position in the triad, while the ego is in the center position. (B) The initiator perspective triadic closure: tie formation between ego and degree-2 alter("d2"). Ego and degree-2 alter in the endpoint position in the triad.

We need to integrate these two perspectives to gain a comprehensive understanding of the cluster formation process and consistency. Based on the clustering feature network, the local clustering coefficient of the intermediator can increase the probability of triadic closure. A high local clustering coefficient of the intermediator suggests that the intermediator belongs to a tightly-knit social circle, making it more likely for two degree-1 alters of the ego to contact each other. Conversely, a high local clustering coefficient of the endpoints decreases the likelihood of triadic closure. This means that when an actor is already part of a close social circle, they are less likely to initiate contact with someone outside of their social circle.

I argue that the relationship between the local clustering coefficient and the possibility of triadic closure changes based on the other actor's local clustering coefficient in the triad (as illustrated in Figure 4). In other words, the interaction effect between local clustering in different positions promotes triadic closure while also maintaining sparsity.

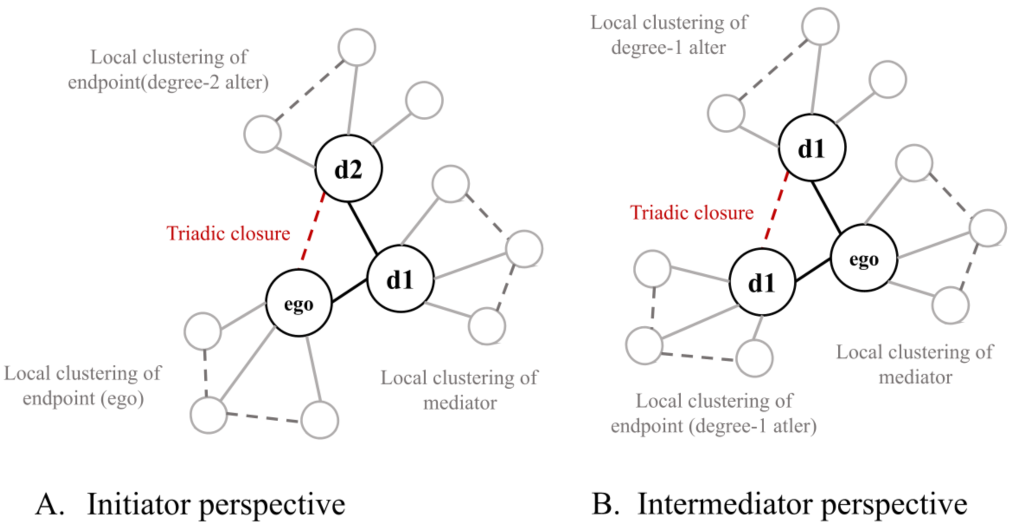

Figure 1-4 The local clustering coefficient in each position and triadic closure: “d1” refer to degree-1 alter, while “d2” refer to degree-2 alter.

The initiator perspective of triadic closure suggests that when the local clustering coefficient of an alter in the endpoint is low, it indicates that the alter has not yet joined a specific cluster. In this scenario, the alter may seek to connect with a degree-2 alter to gain an opportunity to join a cluster for potential benefits. As a result, degree-2 alters with high local clustering may be particularly attractive to the ego. For egos with the same level of local clustering coefficient, the higher the local clustering coefficient of the degree-2 alters, the greater the possibility of ego initiating contact.

Hypothesis 1. The lower the local clustering, the higher the possibility of ego initiating contact degree-2 alter, especially toward those degree-2 alter with high local clustering. The reverse is also expected to hold true.

Hypothesis 2. The possibility of triadic closure occurring will increase with degree-1 alter's local clustering coefficient decrease. Moreover, this effect will be amplified as ego's local clustering increases.

In other words, if a degree-1 alter has a close social circle where most of their friends know each other, then the rest of their friends who do not know each other will also tend to become friends, further raising the density within the cluster. Moreover, if degree-1 alters are already embedded in a high-density network with a high local clustering coefficient, the ego's local clustering effect will increase the possibility of triadic closure.

1.3 The rival theory of clustering: inequality of popularity

The power law degree distribution characterizes the distribution of degrees in a network, which typically includes a small number of highly connected nodes and many nodes with only a few connections. This property has been found in the diverse fields of empirical network data, such as worldwide websites, citation networks, and social commerce networks (Barabási & Albert, 1999; Price, 2011; Stephen & Toubia, 2009). In friendship networks, the power law degree distribution describes the unequal distribution of popularity, indicating that a small number of individuals have many friends while many others have relatively few friends. The power law is distinct from the "structural perspective" of transitivity, which refers to the tendency of nodes to form clusters or groups in a network. The origin of popularity inequality is often seen as a consequence of pairwise choice rather than a result of network structural forces (Feld & Elmore, 1982). Nodes may have higher popularity because of social context or other relative reasons, such as sitting in the center of the room, being assigned the role of counselor, or possessing certain traits valued by group members. Some researchers have argued that the impact of transitivity in the tie formation process needs to be emphasized more. They suggest that many of the findings related to transitivity can be attributed to the unequal distribution of popularity, which is a stronger predictor of network patterns. Popularity inequality may be a significant factor affecting the formation of network ties. Therefore, when examining network phenomena, it is crucial to take into account the significant role that popularity inequality plays in shaping network structure. (Doreian & Krackhardt, 2001; Feld & Elmore, 1982; Guo & Kraines, 2009; Hummon & Doreian, 2003). Power law distribution and clustering are two distinct aspects of network structure. The former highlights the importance of highly connected nodes, while the latter refers to the tendency of nodes to form tightly-knit groups or communities. I argue that popularity inequality and clustering are not necessarily rival factors in the tie formation process. Instead, the clustering coefficient can moderate the effect of popularity, allowing these two distinct properties to coexist in real-world networks. By considering the clustering coefficient, we may improve our understanding of the factors that shape network structure. I hypothesize that a node's popularity increases the likelihood of it being chosen by another node, but this effect is weakened by local clustering around the endpoint, which inhibits the formation of inter-clustering ties.

Hypothesis 3: A higher popularity of a degree-2 alter increases the likelihood of triadic *closure. However, the ego's local clustering coefficient weakens the popularity effect.*

In addition, I propose that the clustering coefficient of intermediator nodes can strengthen the effect of popularity and promote triadic closure.

Hypothesis 4: A higher popularity of a degree-1 alter increases the likelihood of triadic closure. Moreover, the superiority of the degree-1 alter's popularity increases with the ego's local clustering coefficient.

1.4 The empirical limitation of obtaining structural information

To fully understand the emergence and persistence of clusters, it is important to consider not only each individual triad and the network surrounding each point in the triad but also the popularity of the nodes involved. Researchers must capture the network structure across three degrees: from ego to degree-2 alter, and the degree-1 alter of degree-2 alter, which are three degrees of separation from ego. However, this can be a common empirical limitation in social network research. The traditional approach to measuring personal networks in conventional network research is through the name generator method. (Burt, 1984). It involves asking individuals about specific aspects of their friendships, such as “with whom do you talk about important matters?” and “with whom have you done something?” (Lange et al., 2004). However, the name generator is vulnerable to survey fatigue, which limits the number of questions that can be asked. As a result, it can only capture small-scale networks, typically involving less than five people, and is far from reflecting real-world networks (Marin & Hampton, 2007; Pustejovsky &

Spillane, 2009). One of the limitations of the name generator is that it only allows researchers to measure the local clustering coefficient of the focal person (ego) in a triad. The local clustering coefficient of the ego's friend (degree-1 alter) and friend of a friend (degree-2 alter) remains unknown.

Investigations conducted on small groups, such as adolescents in the same class (Değirmencioğlu et al., 1998; Frank et al., 2013) , can extend our understanding of network structure, although they are limited to friendship networks consisting of strong ties. However, traditional survey data does not capture real-time points of tie formation, which makes it impossible to determine how long it takes for triadic closure or whether local clustering increases or decreases triadic closure.

The development of social media provides a new opportunity for researchers, but it also presents new challenges. User-generated data from social media offers significant advantages as every contact between users is automatically recorded with an accurate time stamp. This extends the range of relationships that researchers can observe, no longer limiting it to strong ties. However, compared to the traditional survey method, user-generated data lacks user demographic characteristics, which are essential control variables in network research. Researchers often derive data from open-source websites like Yahoo QA, Stack Overflow, or Twitter, where most users are anonymous, and basic background information such as gender and age are missing, not to mention personality measures. One way to overcome this limitation is to obtain authorized social media usage records from a group of users and conduct surveys simultaneously. However, social media contact data contains users' multiple social circles, including neighborhoods, friends of clubs, and colleagues. Alters entering and leaving the network at different times may interfere with the triad formation process. Lastly, the range of people ego has ever contacted on social media might be in the hundreds or thousands, making it almost impossible to conduct surveys and obtain authorized social media data from all the users focal people have ever contacted.

To overcome the mentioned limitations, this article used Facebook usage records from 26 college departments in Taiwan. This dataset contains the contact records of students in the same department, ranging from their first year to their junior year (2012~2016), enabling us to observe the whole network formation process. Furthermore, our dataset also includes their offline life survey, which contains information on the students' gender, birth year, the time they spend on social media per week, and personality. By controlling for students' offline characteristics, we can identify network-level factors' characteristics and understand the triadic closure process. Our research's main contribution is overcoming the empirical limitations and illustrating how the local clustering coefficient works in the triadic closure process. Comparing triadic closure from the intermediator and initiator's perspective, we found that clustering promotes a specific group of actors linked together but limits actors' extension across the degree network.

Chapter 2 Data and methods

2.1 Data

The data utilized in this study were collected using a clustering random sampling method to select a sample. The population was all newly graduated students in Taiwan in 2016, divided into technological and non-technological universities based on the official distinction made by the Ministry of Education in Taiwan. From this pool, 47 departments comprising 3,850 students were randomly sampled from all university departments in Taiwan. Additionally, 20 departments in the finance area were randomly sampled, resulting in a total of 1,579 students. The research team obtained permission from department offices to contact graduated students through teaching assistants or class representatives and requested that they complete an online survey questionnaire and authorize Facebook usage data. The dataset consisted of two parts: the first part was an online questionnaire that investigated their background information, lifestyle, personality, and values, and the second part was Facebook contact records, which were crawled from the user account homepage of the providers. The Facebook records spanned from September 1, 2012, to June 30, 2015, corresponding to the beginning of the first year and the end of their junior year.

To ensure sufficient data for observing the class network evaluation process, only classes with response rates over 70% were included in our analysis. This resulted in a sample of 919 students from 26 classes.

2.2 Build up ego-centric networks

People interact with each other on Facebook in various ways, including liking, commenting, and tagging. However, commenting and tagging are less frequent than liking since they require more cognitive effort (Pletikosa Cvijikj & Michahelles, 2013). Comments and tags are believed to reveal a deeper meaning of the relationship between two individuals. Based on this understanding, I defined contact between two students as "having ever commented or tagged each other on Facebook." Using this definition, I created an ego-centric contact network for each student, where the student would be the ego in their ego-centric contact network. Those classmates who had ever commented or tagged ego became the ego's degree-1 alters, while those who had contacted degree-1 alters but not ego were considered degree-2 alters. Our research aimed to compare the pattern of triadic closure between the initiator perspective and the intermediator perspective. The initiator perspective indicated a degree-2 alter transfer into a degree-1 alter. I analyzed each pair of ego and degree-2 alter, and the triads consisted of ego, degree-2 alter, and one or more degree-1 alters.

On the other hand, from the intermediator's perspective, the triads consisted of an ego and two degree-1 alters. The one who contacted the ego earlier was called the "old degree-1 alter," while the one who contacted the ego relatively later was referred to as the "new degree-1 alter." As the study focused on peer networks within the same class, all alters in the ego-centric network were students from the same class as the ego.

2.3 Survival model

Survival analysis is a quantitative research approach that examines the time until a designated event of interest occurs. In this study, the Cox proportional hazards model was proposed to explore the relationship between triadic closure and the factors of interest. A critical analytical issue in conducting survival analysis is censoring. As the analysis only considered triadic closure before the end of the third-year semester, the data were right-censored. The dependent variable in survival analysis comprises two variables: a dummy variable indicating whether the event occurred and a time variable measuring the time until the event. Survival analysis originated in public health research, where the event of interest was often a healthy outcome, such as death or disease incidence. Therefore, the time variable in survival analysis is also known as "survival time." The primary objective of survival analysis is to investigate whether the independent variable, x, affects the dependent variable, y, in two dimensions: the occurrence of the event and the duration of survival time.

2.4 Dependent variables: define the event occurring and

survival time

I proposed the Cox proportional hazards model to examine the relationship between triadic closure and the factors of interest. The first analytical problem in conducting survival analysis is censoring. Since I only analyzed triadic closure before the end of the third-year semester, our data were right-censored. Our dependent variable for survival analysis consisted of two variables: a dummy variable that recorded whether the event of interest occurred and a time variable that recorded the time until the event occurred. Survival analysis originates from public health research, where the event of interest is typically a health outcome, such as death or disease incidence. As a result, the time variable in survival analysis is also called "survival time." The main goal of survival analysis is to answer whether the independent variable, x, affects the dependent variable, y, in two dimensions: whether the event occurred and the length of survival time.

In this article, the survival time started when an open triad form ended when the alters were in the endpoint in the triad contact. From the initiator's perspective, a triad's survival time started when a degree-1 alter contacted another alter who had never contacted ego before. As the alter became a degree-2 alter, the survival time would start. The survival time stopped counting either when the degree-2 alter contacted ego or when the study period ended. Notably, all comments and tags in our data had a specific initiator and receiver. Therefore, we could divide events into ego-initiated closure and degree-2 alter-initiated triadic closure. For example, if ego commented on a degree-2 alter's post, it would be an ego-initiated closure.

I graphically illustrated the triadic closure process with sample contact records (Table 2-1) and survival data (Table 2-2) (see Figure 2-1). I considered class member A as the focal person (ego). Firstly, I observed from the initiator's perspective. B became a degree-1 alter in A's ego-centric network when B commented on A's post on 2013/02/01 at 12:01:05. Subsequently, C became a degree-2 alter since B tagged C on 2013/02/03 at 17:05:01. At the same time, the survival time started counting until 2013/02/06 at 17:01:49, when C commented on A and became a degree-1 alter. From the records, we determined that the triad A-B-C was an ego-initiated closure triad, and its survival time was three days, from 2013/02/03 at 13:12:01 to 2013/02/12 at 19:05:23.

Table 2-1 Sample contact record

Time Stamp	Contact between class members
2013/02/01 12:01:05	B commented A's post
2013/02/03 17:05:01	B tagged C
2013/02/04 20:14:10	D tagged A
2013/02/06 17:01:49	C commented A
2013/02/08 20:14:48	D commented B

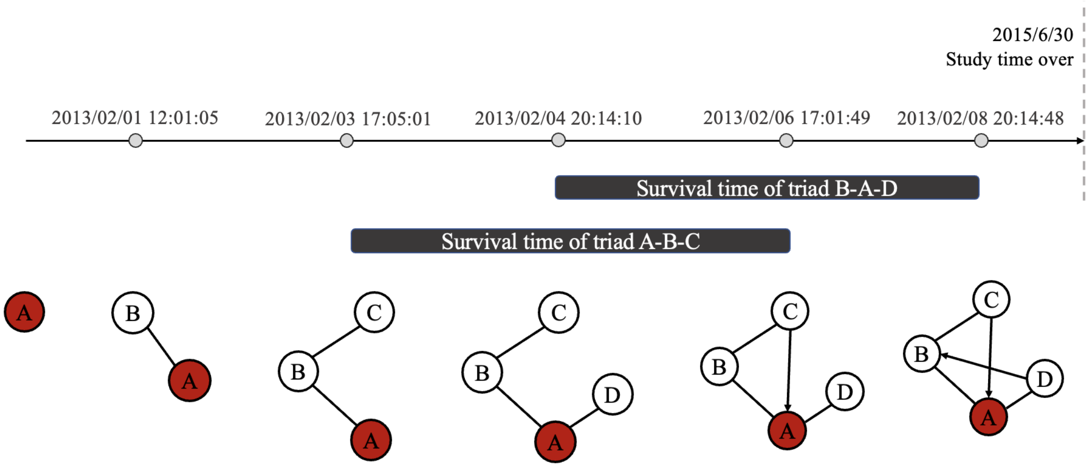

Figure 2-1 Network graph corresponding to sample contact records (consider class member A as ego)

Table 2-2 Sample survival data, according to different definitions of triadic closure event and the perspective of triadic closure

	Model	Definition of Failure	Triad	Survival Time	Failure
Initiator perspective	Model 1A	ego initiated contact degree-2 alter	A-B-C	3	0
	Model 1B	degree-2 alter initiated contact ego	A-B-C	3	1
Mediator perspective	Model 2A	old degree-1 alter initiated contact new degree-1 alter	B-A-D	4	0
	Model 2B	new degree-1 alter initiated contact old degree-1 alter	B-A-D	4	1

Secondly, I observed triadic closure from the intermediator's perspective. The survival time started from the contact between an alter and an ego with a degree-1 alter(s). Survival time stopped counting either when the new degree-1 alter contacted the old degree-1 alter or when the study time ended. The survival time started when D tagged A, who already had a degree-1 alter B, and ended when D contacted B. Since D, the one who contacted A relatively late, initiated the contact, we recorded it as a new degree-1 alter-initiated triadic closure. Overall, I built four survival models by applying different definitions of triadic closure events and perspectives of triadic closure.

2.5 Control variables

As for network-level variables, I controlled for the alter's degree and the number of mutual alters between ego and degree-2 alters. To focus on the effect of the local clustering coefficient, I also controlled for the individual-level characteristics of alters in the endpoint. The first offline control variable was extroversion. Individuals with extroverted personalities are more likely to build relationships on social media because they are more active than other users. The second offline control variable was gender. The effect of gender is similar to that of personality. Compared with males, females spend more time on social media, are more emotionally devoted, and have more robust network closures. Age was the last offline control variable. The term "relative age effect" (Mavilidi et al., 2022) individuals born earlier than their peers may have an advantage because they have more time to develop physical and intellectual abilities. Those who are older among their peers are also more popular (van Aalst & van Tubergen, 2021) . This advantage might last until adulthood.

2.6 Independent variables

Since I focused on pairwise contact, I used the average number of alters tagged during the survival time to measure the alter's popularity. The original distribution of the average number of ego being tagged was highly skewed, with skewness values of 10.854 in initiator and 7.803 in perspective triads and intermediator perspective triads, respectively. This distribution fit the description of the inequality of popularity. To prevent the deviation of popularity from interfering with the estimated results, I took the logarithm of the average number of actors being tagged as popularity. The distribution of ego's popularity is illustrated below. (Figure 2-2)

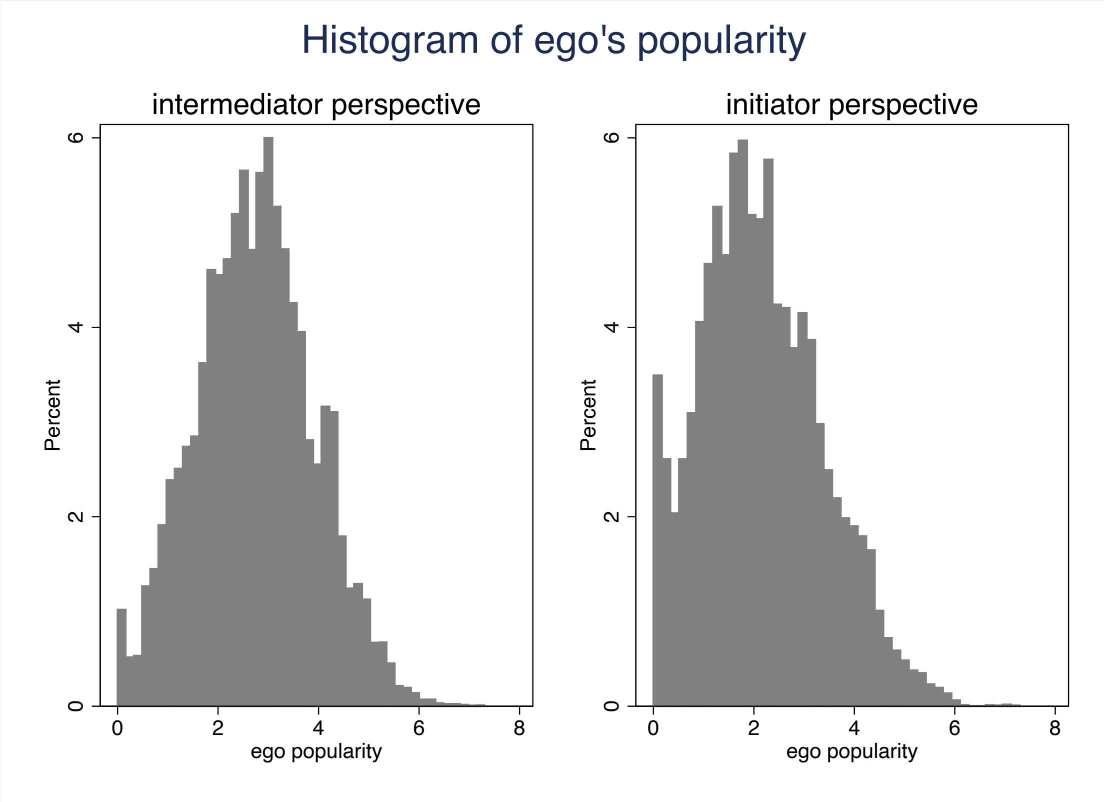

Figure 2-2 Histogram of initiator perspective (left) and intermediator perspective (right) ego's popularity

The independent variables in the intermediator perspective models (Table 2-3) were ego's local clustering coefficient, degree-2 alter's local clustering coefficient, degree-2 alter's popularity, and their corresponding interaction terms, which corresponded to Hypotheses 1 and 3. In the models from the intermediator perspective, the independent variables were the popularities of two degree-1 alters, the local clustering coefficient of each point, and their corresponding interaction terms. The interaction between the degree-1 alter and ego corresponded to Hypothesis 2, while the interaction between the degree-1 alter's popularity and ego's local clustering coefficient corresponded to Hypothesis 4.

Table 2-3 Definition of variable

	Description			
dependent var.				
failure	Model1A	Model1B	Model2A	Model2B
	ego initiated contact degree-2 alter	degree-2 alter initiated contact ego	old degree-1 alter initiated contact new degree-1 alter	new degree-1 alter initiated contact old degree-1 alter
survival time	days it takes for triadic closure (failure)			
control var.				
age	integer, all of students were either 22 or 23			
degree	number of alter contacted with new degree-1 alter in survival time			
extroversion	range from 1 to 4, 4 was high extroversion			
gender	1=female, 0=male			
mutual neighbor	number of degree-1 linked between ego and degree-2 alter in initiator perspective triad			
independent var.				
lcc	alter's local clustering coefficient			
popularity	average number of alter be tagged during the survival time (the logarithm transformation)			

2.7 Description statistics

Table 2-4 illustrates the characteristics of the initiator perspective triads, which were grouped based on their closure status: no closure until study time ended, ego-initiated closure, and degree-2 alter-initiated closure. In the initiator perspective, the study observed 21,900 sets of initiator perspective triads, of which 34.6 percent closed during the study period with an average survival time of 582.571 days. The pattern of the local clustering coefficient was consistent with hypothesis 1, as ego-initiated closure triads had the lowest ego's local clustering coefficient, while degree-2 alter-initiated closure triads had the lowest degree-2 alter's. This suggested that the lower the alter's local clustering coefficient, the higher the possibility that the alter would initiate contact. Additionally, both ego-initiated and degree-2 alter-initiated closures triads had higher popularity than no-closure triads.

In Table 2-5, the study recorded 36,563 intermediator perspective triads during the study period, with an average survival time of 564.555 days and a closure (failure) rate of 15.4%, which was lower than that of initiator perspective triads. The popularity of degree-1 alters was higher in triads that eventually closed compared to those that did not close until the research period ended, which was consistent with the idea of power law. However, the average of ego's local clustering coefficient in the no closure triads was higher than the other two triad groups, contradicting hypothesis 2.

Table 2-4 Initiator triad description statistics, mean, and their standard error

	Traid			
	no closure	degree-2 alter initiated closure	ego initiated closure	Total
dependent var.				
Failure (%)				34.6
survival time	761.321	242.357	246.71	582.571
	(270.515)	(235.205)	(236.717)	(357.151)
control var.				
degree-2 alter 's degree	5.754	6.831	4.697	5.768
	(6.206)	(6.908)	(5.608)	(6.273)
degree-2 alter's age	22.419	6.108	5.561	5.495
	(0.595)	(0.58)	(0.589)	(0.591)
degree-2 alter's level of extroversion	2.54	2.641	2.534	2.557
	(0.73)	(0.761)	(0.732)	(0.737)
degree-2 alter's gender (1=female)	0.539	0.584	0.521	0.544
	(0.498)	(0.493)	(0.5)	(0.498)
ego's degree	6.085	4.862	7.165	6.049
	(6.306)	(5.566)	(7.195)	(6.378)
ego's age	22.407	22.412	22.42	22.41
	(0.576)	(0.571)	(0.56)	(0.572)
ego's level of extroversion	2.535	2.528	2.646	2.553

	(0.731)	(0.729)	(0.773)	(0.739)
ego's gender (1=female)	0.538	0.52	0.581	0.542
	(0.499)	(0.5)	(0.494)	(0.498)
mutual neighbor	5.311	6.108	5.561	5.495
	(5.054)	(5.349)	(5.112)	(5.126)
independent var.				
degree-2 alter 's lcc	0.647	0.532	0.656	0.628
	(0.241)	(0.224)	(0.233)	(0.241)
degree-2 alter popularity	1.804	2.93	2.876	2.185
	(0.986)	(1.319)	(1.223)	(1.212)
ego's lcc	0.645	0.65	0.535	0.627
	(0.245)	(0.246)	(0.224)	(0.245)
ego popularity	1.793	2.836	2.877	2.161
	(0.994)	(1.24)	(1.347)	(1.216)
number of observation	3,679	3,895	14,326	21,900

Table 2-5 Intermediator perspective triad description statistics, mean, and their standard error

	Traid			
	no closure	d1 new initiated closure	old d1 initiated closure	Total
dependent var.				
failure(%)				0.154
survival time	625.061	223.409	241.619	564.555
	(286.518)	(236.213)	(247.481)	(313.926)
control var.				
new degree-1 alter's age	22.453	22.405	22.437	22.448
	(0.586)	(0.578)	(0.6)	(0.587)
new degree-1 alter's degree	14.934	14.856	12.155	14.704
	(9.927)	(9.663)	(8.592)	(9.835)
new degree-1 alter's level of extroversion	2.592	2.633	2.538	2.591
	(0.768)	(0.718)	(0.7)	(0.759)
new degree-1 alter's gender(female=1)	0.587	0.606	0.534	0.584
	(0.492)	(0.489)	(0.499)	(0.493)
old degree-1 alter's age	22.419	22.474	22.487	22.429
	(0.587)	(0.571)	(0.553)	(0.584)
old degree-1 alter's degree	15.571	11.373	14.144	15.147

	(10.37)	(8.512)	(9.515)	(10.24)
old degree-1 alter level of extroversion	2.589	2.513	2.69	2.592
	(0.74)	(0.775)	(0.823)	(0.75)
old degree-1 alter’s gender (female=1)	0.577	0.6	0.618	0.582
	(0.494)	(0.49)	(0.486)	(0.493)
independent var.				
new degree-1 alter's lcc	0.632	0.523	0.629	0.624
	(0.228)	(0.223)	(0.242)	(0.23)
old degree-1 alter's lcc	0.629	0.614	0.525	0.62
	(0.225)	(0.237)	(0.223)	(0.228)
ego's lcc	0.576	0.519	0.519	0.567
	(0.204)	(0.216)	(0.217)	(0.207)
new degree-1 popularity	2.266	3.125	3.177	2.403
	(1.025)	(1.374)	(1.32)	(1.127)
old degree-1 alter popularity	2.366	3.149	3.108	2.484
	(1.146)	(1.253)	(1.373)	(1.205)
number of observation	30,921	2,686	2,956	36,563

Chapter 3 Results

Firstly, I used the Kaplan-Meier estimator to estimate the probability of triad survival (i.e., not closing) over time. As shown in Figure 3-1, the survival curves of triadic closure exhibit similar patterns, with a steep decline in the early stages and a gradual flattening over time. This suggests a high risk of triadic closure in the first few months, which gradually decreases over time. This indicates that the ego is more likely to initiate contact with a degree-2 alter in the first few months after the alter becomes a degree-2 alter. Similarly, the two degree-1 alters of the ego are also more likely to contact each other in the first few months after the new degree-1 alter first contacts the ego.

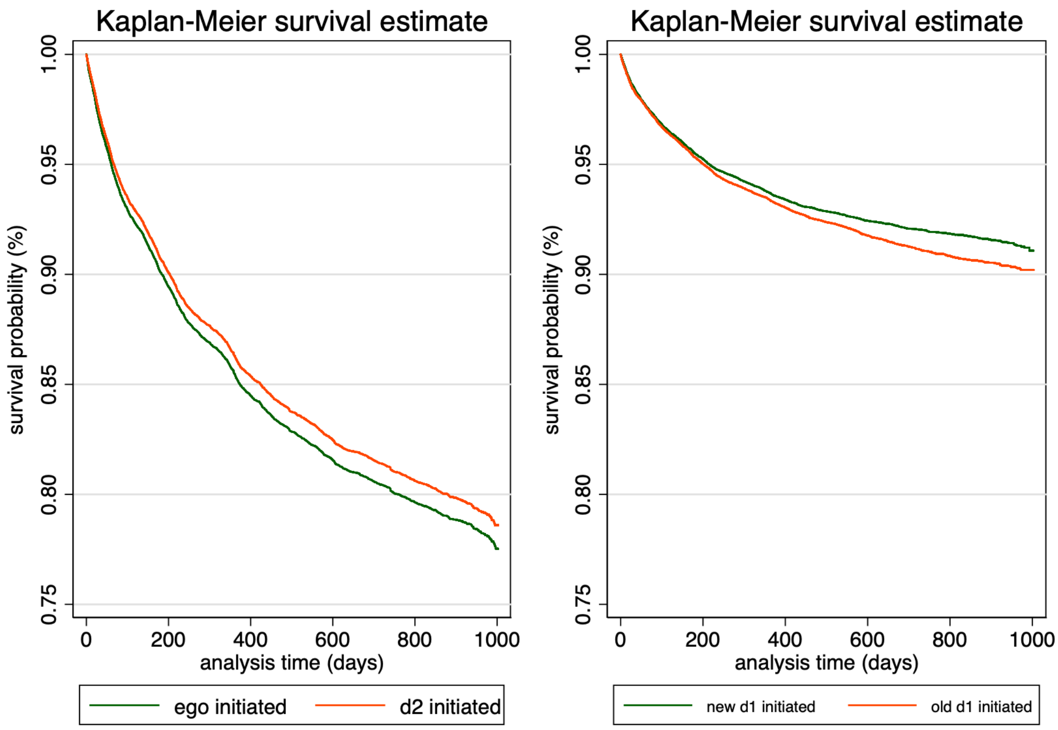

Figure 3-1 Kaplan-Meier estimate survival curve for initiator perspective triadic closure (left) and intermediator perspective triadic closure (right)

3.1 Initiator perspective model

The result of initiated perspective survival analysis is displayed in Table 3-1 (Model 1A and Model 1B) The parameter estimate of interest in initiator perspective models was the ego's lcc # degree-2 alter's lcc and ego's lcc # degree-2 alter popularity. Both interaction terms significantly impacted the hazard ratio of triadic closure. Figure 3-2 graphically shows the interaction term ego's lcc # degree-2 alter's lcc, which addresses hypothesis 1. I could observe the main effect of lcc in ego-initiated triadic closure. The higher the ego's lcc, the lower the relative hazard ratio. Degree-2 alter initiated model illustrated the same tendency. Corresponding to hypothesis 1, the alter with lower lcc would have a higher possibility of initiating contact to closure. However, the higher the degree-2 alters lcc, the lower the relative hazard ratio. The higher the degree-2 alter's lcc, the lower the possibility of ego-initiated contact. The effect of degree-1 alters lcc would increase as ego's lcc decrease. Intuitively, if ego had not formed its social circle, it would be more willing to initiate contact across the clustering, especially toward that degree-2 alter who also had not formed a social circle. Low degree-2 alter' lcc could appeal ego initiated contact, but this superiority would decrease the higher the ego's lcc. The effect ego's lcc contributed to the closure event depended on degree-2 alter's lcc, but this phenomenon could only be observed in ego initiated model. In the degree-2 alter initiated model, the change of ego' lcc did not affect the relationship between the relative hazard ratio and degree-2 alter's lcc. Moreover, comparing the scope of the two graphs, I could find that ego's lcc affects the relative hazard ratio more dramatically than degree-2 alter's lcc.

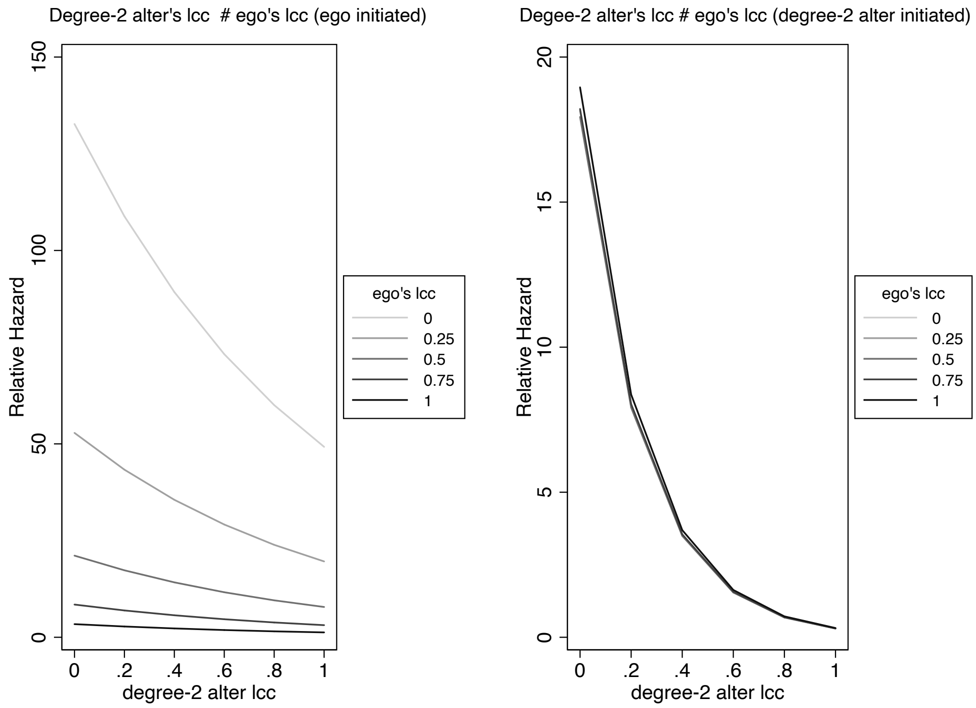

Figure 3-2 Interaction between ego's lcc and degree-2 alter's lcc

Table 3-1 Initiator perspective cox proportional hazards model, the estimates of relative hazard ratios and their standard errors (significance level: * 0.1, **0.05, ***0.001)

	Model 1A	Model 1B
	Coef./std.errors	Coef./std.errors
control var.		
independent var.		
ego's lcc	0.011***	0.123***
	(0.003)	(0.023)
degree-2 alter lcc	0.371***	0.017***
	(0.061)	(0.003)
ego lcc # degree-2 alter lcc	2.819***	1.698*
	(0.870)	(0.522)
ego popularity	2.039***	2.335***
	(0.037)	(0.043)
degree-2 alter popularity	2.646***	1.711***
	(0.095)	(0.067)
ego's lcc # degree-2 alter popularity	1.162***	1.539***
	(0.066)	(0.082)
No. of Obs.	21,900	21,900

Secondly, I examined hypothesis 3, which explored the relationship between degree-2 alter's popularity and ego's lcc. In the ego and degree-2 alter initiated model, I found that the higher the degree-2 alter's popularity, the higher the relative hazard ratio. This indicated that degree-2 alters with high popularity were more attractive and likely to initiate contact with the ego. Moreover, the interaction term between ego's lcc and degree-2 alter's popularity was significantly positive in both Model 1A and Model 1B.

The interaction effect is presented in Figure 3-3 In the ego-initiated model, I observed that the attraction of popular degree-2 alters became stronger as the ego's lcc decreased. This suggested that those who had not yet formed their social circle were more likely to be attracted by popular degree-2 alters. However, those degree-2 alters with the highest level of popularity were more likely to initiate contact with an ego who had a higher level of lcc. More specifically, degree-2 alters initially preferred an ego with lower lcc, but as their popularity grew, the ego's lcc gradually became a superiority that attracted degree-2 alters. Overall, the relationship between degree-2 alter's popularity and triadic closure was positive, increasing the possibility of ego contacting degree-2 alter, and degree-2 alter contacting ego, as expected in hypothesis 3. While ego's lcc did not reverse this relationship, it did weaken or enhance the effect of popularity.

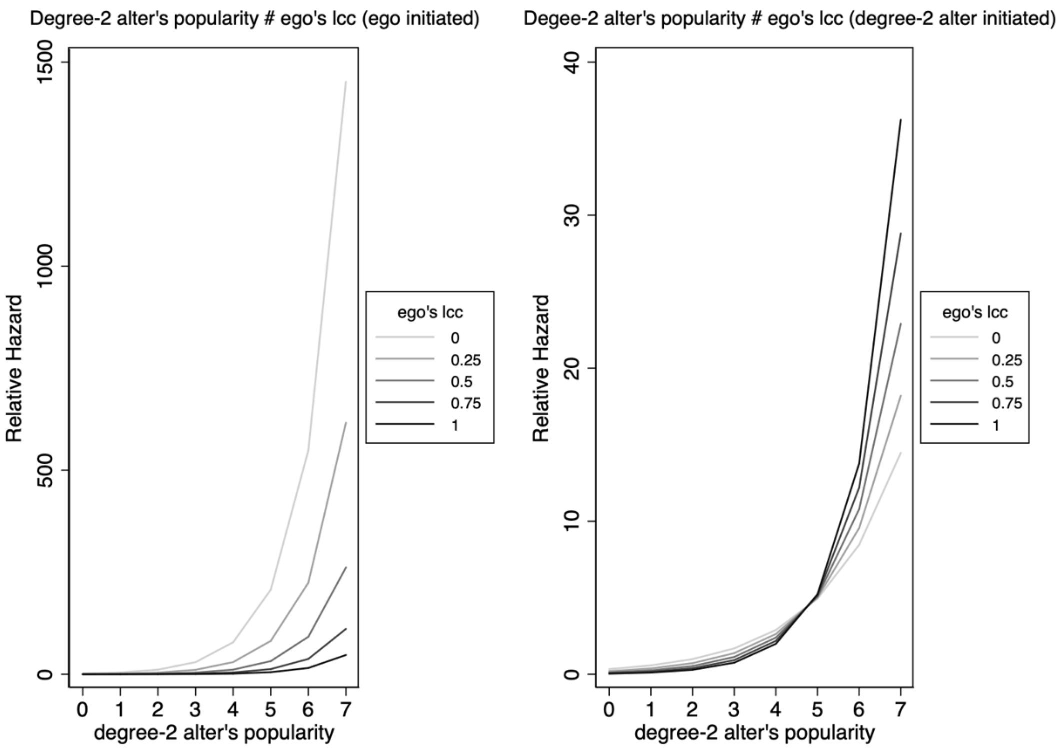

Figure 3-3 Interaction between ego's lcc and degree-2 alter's popularity

3.2 Intermediator perspective model

Table 3-2 presents the results of the intermediator perspective model (Model 2A and Model 2B). Consistent with the findings in Figure 3-2, I observed a significant main effect of degree-1 alter's local clustering coefficient (lcc), indicating that lower lcc values of degree-1 alters were associated with a higher likelihood of initiating contact, regardless of whether they were new or old degree-1 alters. Furthermore, a lower lcc of another degree-1 alter was also associated with a higher likelihood of initiating contact, which aligns with the results of the initiator perspective models. To summarize, if the endpoint alters in the triad had not established their social circle, they were more likely to initiate contact with another endpoint in the triad, particularly if the other alter had also not formed their own social circle.

Table 3-2 Intermediator perspective cox proportional hazards model, the estimates of relative hazard ratios and their standard errors (significance level: * 0.1, **0.05, ***0.001)

	Model 2A	Model 2B
	Coef./std.errors	Coef./std.errors
control var.		
Independent var.		
ego's lcc	0.357***	0.271***
	(0.125)	(0.096)
new degree-1 alter lcc	0.144***	0.297***
	(0.036)	(0.067)
old degree-1 alter's lcc	0.184***	0.082***
	(0.042)	(0.020)
ego's lcc # old degree-1 alter's lcc	5.986***	0.010***
	(2.707)	(0.005)
ego's lcc # new degree-1 alter's lcc	0.003***	4.362***
	(0.001)	(2.074)
new degree-1 alter popularity	1.189***	2.846***
	(0.059)	(0.152)
old degree-1 alter popularity	4.003***	1.761***
	(0.212)	(0.092)
ego's lcc # new degree-1 alter popularity	3.580***	1.017
	(0.330)	(0.102)
ego's lcc # old degree-1 alter popularity	0.360***	1.061
	(0.035)	(0.101)
No. of Obs.	36563	36563

Further, I investigated how ego's lcc mediates the effect of endpoint alter's lcc, as hypothesized in Hypothesis 2. To test Hypothesis 2, I examined the interaction terms. In the new degree-1 initiated model, I found a significant and positive interaction between the ego's lcc and the old degree-1 alter's lcc, whereas in the old degree-1 alter-initiated model, the interaction was significant and negative. Similarly, in the new degree-1 initiated model, I found a significant and negative interaction between ego's lcc and the new degree-1 alter's lcc, while in the old degree-1 alter-initiated model, it was significant and positive. Figure 3-4 vividly represents these interaction effects.

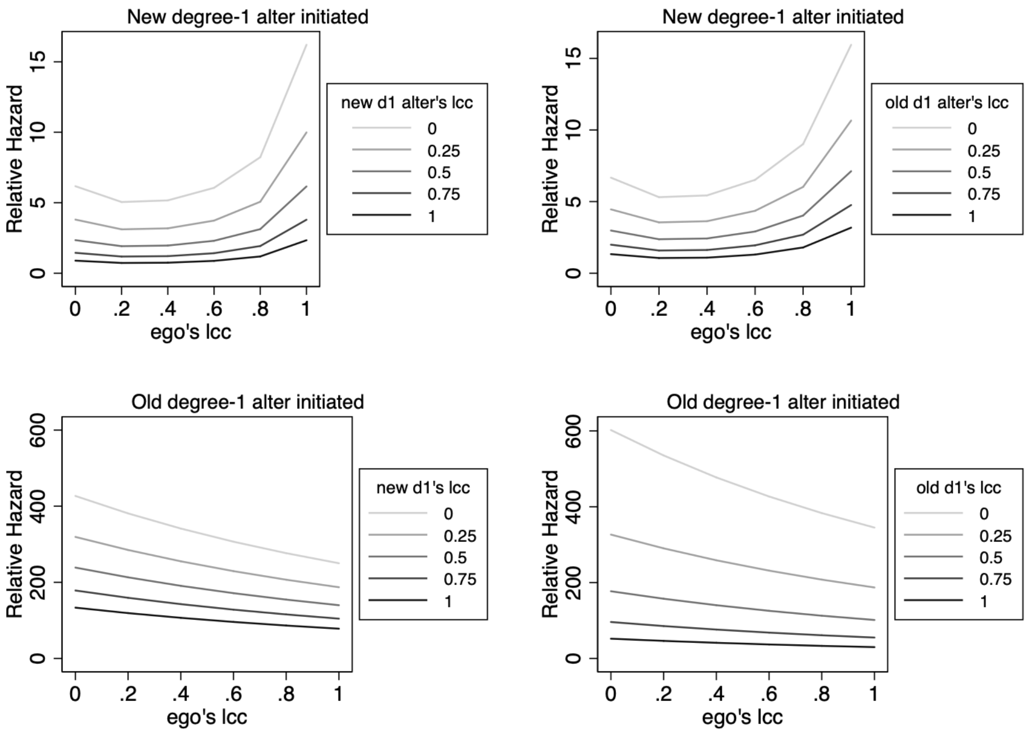

Figure 3-4 Interaction between degree-1 alter’s lcc and ego’s lcc

I first examined the role of the newcomer in the ego-centric network. In the new degree-1 alter-initiated model, I observed a U-shaped relationship between the relative hazard ratio and ego's lcc. At the start, ego's lcc decreased the relative hazard ratio. However, as the ego's lcc increased, the direction of the relationship gradually changed and became positive. The level of the new degree-1 alter's lcc weakened the U-shaped tendency rather than inverting it, making the predictability of the relative hazard ratio smoother.

Next, I focused on the effect of the old degree-1 alter. Contrary to our expectation, ego's lcc reduced the likelihood of the old degree-1 alter initiating contact. The effect of ego's lcc prevented the old degree-1 alter from contacting the new degree-1 alter, and this effect became more robust as the lcc of the new or old degree-1 alter decreased. This partially supported Hypothesis 2; the lcc of the degree-1 alter did indeed suppress the effect of ego's lcc, but ego's lcc only promoted the new degree-1 alter to take the initiative at the beginning and then decreased the possibility of the new degree-1 alter contacting the old degree-1 alter. As for the old degree-1 alter, the effect of ego's lcc was to suppress its likelihood of contacting the newcomer.

Finally, the interaction term ego's lcc # old degree-1 alter's popularity, and ego's lcc # new degree-1 alter's lcc gave us a clue about how popularity moderated the relationship between lcc and triadic closure. Figure 3-5 shows the effect of interaction terms. Firstly, it was notable that the effect of the new degree-1 alter's popularity was more substantial in the old degree-1 alter-initiated triadic closured.

The highest relative hazard ratio reached 5000, about five times more than the maxim in new degree-1 alters initiated circumstance. Consistent with our findings in Model 1A and Model 1B, a popular degree-1 alter tended to be more willing to initiate contact with another degree-1 alters, and other degree-1 alter were also more willing to contact with them. For old degree-1 alter, the effect of popularity would fade as ego's lcc increases. For new degree-1 alter, ego's lcc also weakened the tendency of old degree-1 alter contact popular new degree-1 alter, but it would improve the tendency of popular new degree-1 alterts initiated contact. In conclusion, the result partly agreed with hypothesis 3. Ego's lcc indeed decreased the effect of popularity, but for a newcomer in the ego-centric network, ego's lcc promoted their likehood of initiating contact.

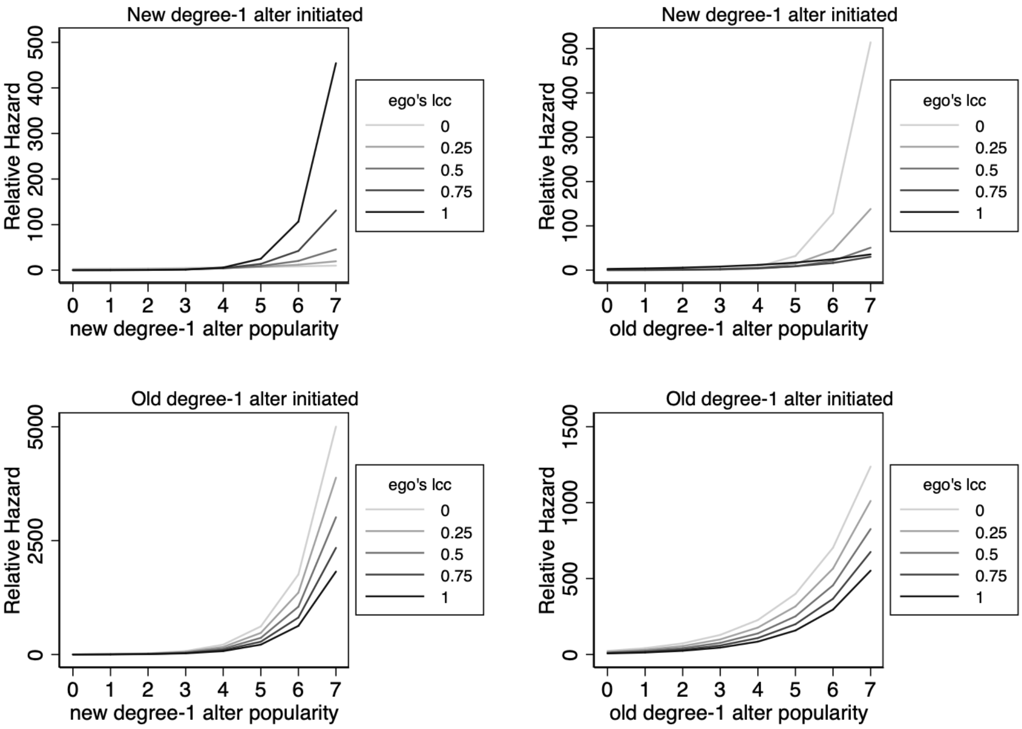

Figure 3-5　　Interaction between degree-1 alter's popularity and ego's lcc

Chapter 4 Discussion

Through the analysis of contact records on Facebook, I was able to observe the comprehensive process of peer network evolution, shedding light on the puzzle of clustering formation and consistency mechanisms. The factors promoting triadic closure and intermediator variables were found to enhance the effect, as shown in Figure 4-1

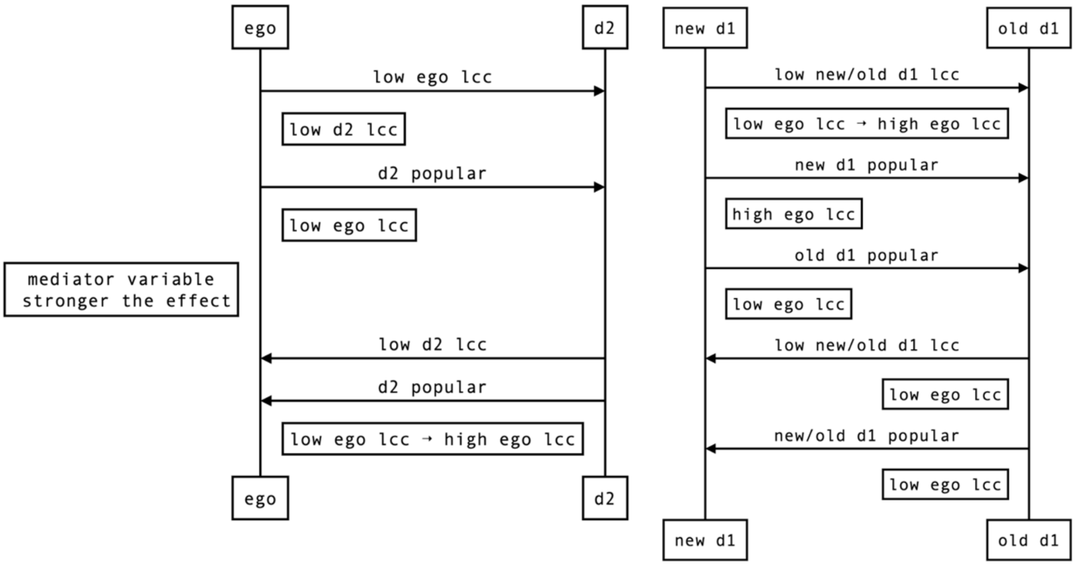

Figure 4-1 Factors increase triadic closure and intermediator variables which more potent the effect of variables (text in the text box).

The comprehensive process of peer network evolution shed light on the puzzle of clustering formation and consistency mechanisms. Low local clustering was indicative of not belonging to a particular social circle. In most situations, individuals who did not belong to a social group initiated contact with a friend of a friend, especially those who had also not formed their own social group. Additionally, popularity was a factor that attracted those who had not formed their own social circle, as they were more likely to initiate contact with popular degree-2 alters. However, those with the highest level of popularity preferred individuals who were already part of their social circle.

Contrary to our initial assumption, ego's local clustering coefficient also inhibited new degree-1 alters from contacting old degree-1 alters. Initially, we hypothesized that ego's local clustering would enhance the effect of degree-1 alters' local clustering, indicating that having a relatively close social circle would increase the likelihood of a new friend becoming friends with an old friend. However, ego's local clustering only increased the likelihood of new degree-1 alters taking the initiative when it was over approximately 0.4. Otherwise, the local clustering coefficient reduced the likelihood of degree-1 alters taking the initiative. This suggests that people are less likely to initiate contact with a new friend in a relatively close social circle. It may be the reason why a close social circle encourages new friends to take the initiative. The popularity of new friends could attract old friend-initiated contact, but the superiority of popularity would become weaker in a close social circle.

In conclusion, local clustering is a double-edged sword. On the one hand, it can promote an alter to take the initiative and make them more likely to attract other alters. On the other hand, the benefits of low local clustering are limited. Alters with low local clustering coefficients are more likely to connect with others who also have low local clustering and are less likely to attract popular alters. I speculate that having a low local clustering coefficient may allow an alter's network to extend more efficiently, but it also makes it less likely to access the center position in clustering or the power nodes in a network. As traditional theory suggests, low local clustering puts one at risk of cognitive imbalance (Bearman & Moody, 2004; Heider, 1946). Additionally, it provides an opportunity to link with alters who have crossed structural holes, leading to diverse sources of information (Ahuja, 2000; Burt, 2004).

Furthermore, it is worth noting that local clustering not only prevents an alter from being initiated but also prevents other alters from taking the initiative toward it. Traditional theories related to clustering explain the tendency of clustering due to people's limited energy for maintaining ties and their preference for interacting with people within their clustering. However, these explanations only address why local clustering prevents alter contact with others, but not why local clustering also makes an alter less attractive to other alters. Future studies could investigate the mechanisms underlying this phenomenon.

REFERENCES

Ahuja, G. (2000). Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly*, *45*(3), 425–455. https://doi.org/10.2307/2667105

Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*(5439), 509–512. https://doi.org/10.1126/science.286.5439.509

Bearman, P. S., & Moody, J. (2004). Suicide and Friendships Among American Adolescents. *American Journal of Public Health*, *94*(1), 89–95. https://doi.org/10.2105/ajph.94.1.89

Blau, P. M. (2017). *Exchange and Power in Social Life* (1st ed.). Routledge. https://doi.org/10.4324/9780203792643

Burt, R. S. (1984). Network items and the general social survey. *Social Networks*, *6*(4), 293–339. https://doi.org/10.1016/0378-8733(84)90007-8

Burt, R. S. (1985). Toward a Structural Theory of Action: Network Models of Social Structure, Perception and Action. *American Journal of Sociology*, *90*(6), 1336–1338. https://doi.org/10.1086/228215

Burt, R. S. (2004). Structural Holes and Good Ideas. *American Journal of Sociology*, *110*(2), 349–399. https://doi.org/10.1086/421787

Cartwright, D., & Harary, F. (1956). Structural balance: A generalization of Heider's theory. *Psychological Review*, *63*(5), 277. https://doi.org/10.1037/h0046049

Değirmencioğlu, S. M., Urberg, K. A., Tolson, J. M., & Richard, P. (1998). Adolescent Friendship Networks: Continuity and Change Over the School Year. *Merrill-Palmer Quarterl*, *44*(3), 313–337.

Doreian, P., & Krackhardt, D. (2001). Pre-transitive balance mechanisms for signed networks. *The Journal of Mathematical Sociology*, *25*(1), 43–67. https://doi.org/10.1080/0022250x.2001.9990244

Feld, S. L., & Elmore, R. (1982). Patterns of Sociometric Choices: Transitivity Reconsidered. *Social Psychology Quarterly*, *45*(2), 77–85. https://doi.org/10.2307/3033928

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3–5), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Frank, K. A., Muller, C., & Mueller, A. S. (2013). The Embeddedness of Adolescent Friendship Nominations: The Formation of Social Capital in Emergent Network Structures. *American Journal of Sociology*, *119*(1), 216–253. https://doi.org/10.1086/672081

Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, *99*(12), 7821–7826. https://doi.org/10.1073/pnas.122653799

González, M. C., Herrmann, H. J., Kertész, J., & Vicsek, T. (2007). Community structure and ethnic preferences in school friendship networks. *Physica A: Statistical Mechanics and Its Applications*, *379*(1), 307–316. https://doi.org/10.1016/j.physa.2007.01.002

Granovetter, M. (1992). Problems of explanation in economic sociology. *Networks and Organizations: Structure, Form, and Action*, 25–56.

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, *78*(6), 1360–1380. JSTOR. https://doi.org/10.1086/225469

Guo, W., & Kraines, S. B. (2009). A Random Network Generator with Finely Tunable Clustering Coefficient for Small-World Social Networks. *2009 International Conference on Computational Aspects of Social Networks*, 10–17.

https://doi.org/10.1109/cason.2009.13

Hansen, M. T. (1999). The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits. *Administrative Science Quarterly*, *44*(1), 82–111. https://doi.org/10.2307/2667032

Heider, F. (1946). Attitudes and Cognitive Organization. *The Journal of Psychology*, *21*(1), 107–112. https://doi.org/10.1080/00223980.1946.9917275

Hummon, N. P., & Doreian, P. (2003). Some dynamics of social balance processes: Bringing Heider back into balance theory. *Social Networks*, *25*(1), 17–49. https://doi.org/10.1016/s0378-8733(02)00019-9

Lange, D. D., Agneessens, F., & Waege, H. (2004). Asking Social Network Questions: A Quality Assessment of Different Measures. *Metodoloski Zvezki*, *1*(2), 351–378.

Levine, S. S., & Kurzban, R. (2006). Explaining clustering in social networks: Towards an evolutionary theory of cascading benefits. *Managerial and Decision Economics*, *27*(2–3), 173–187. https://doi.org/10.1002/mde.1291

Lou, T., Tang, J., Hopcroft, J., Fang, Z., & Ding, X. (2013). Learning to predict reciprocity and triadic closure in social networks. *ACM Transactions on Knowledge Discovery from Data*, *7*(2), 1–25. https://doi.org/10.1145/2499907.2499908

Marin, A., & Hampton, K. N. (2007). Simplifying the Personal Network Name Generator: Alternatives to Traditional Multiple and Single Name Generators. *Field Methods*, *19*(2), 163–193. https://doi.org/10.1177/1525822x06298588

Moody, J. (2001). Race, School Integration, and Friendship Segregation in America. *American Journal of Sociology*, *107*(3), 679–716. https://doi.org/10.1086/338954

Obstfeld, D. (2005). Social Networks, the *Tertius Iungens* Orientation, and Involvement in Innovation. *Administrative Science Quarterly*, *50*(1), 100–130. https://doi.org/10.2189/asqu.2005.50.1.100

Pletikosa Cvijikj, I., & Michahelles, F. (2013). Online engagement factors on Facebook brand pages. *Social Network Analysis and Mining*, *3*, 843–861. https://doi.org/10.1007/s13278-013-0098-8

Price, D. J. de S. (2011). Networks of scientific papers. In M. Newman, A.-L. Barabási, & D. J. Watts, *The Structure and Dynamics of Networks* (pp. 149–154). Princeton University Press. https://doi.org/10.1515/9781400841356.149

Pustejovsky, J. E., & Spillane, J. P. (2009). Question-order effects in social network name generators. *Social Networks*, *31*(4), 221–229. https://doi.org/10.1016/j.socnet.2009.06.001

Rapoport, A. (1953). Spread of information through a population with socio-structural bias: I. Assumption of transitivity. *The Bulletin of Mathematical Biophysics*, *15*(4), 523–533. https://doi.org/10.1007/bf02476440

Schaeffer, S. E. (2007). Graph clustering. *Computer Science Review*, *1*(1), 27–64. https://doi.org/10.1016/j.cosrev.2007.05.001

Shakya, H. B., Christakis, N. A., & Fowler, J. H. (2017). An exploratory comparison of name generator content: Data from rural India. *Social Networks*, *48*, 157–168. https://doi.org/10.1016/j.socnet.2016.08.008

Shrum, W., Cheek, N. H., & Hunter, S. MacD. (1988). Friendship in School: Gender and Racial Homophily. *Sociology of Education*, *61*(4), 227. https://doi.org/10.2307/2112441

Simmel, G. (1950). *The sociology of georg simmel* (Vol. 92892). Simon and Schuster.

Smith, J. A., McPherson, M., & Smith-Lovin, L. (2014). Social Distance in the United States: Sex, Race, Religion, Age, and Education Homophily among Confidants, 1985 to 2004. *American Sociological Review*, *79*(3), 432–456. https://doi.org/10.1177/0003122414531776

Sorenson, O., & Audia, P. G. (2000). The Social Structure of Entrepreneurial Activity: Geographic Concentration of Footwear Production in the United States, 1940–1989. *American Journal of Sociology*, *106*(2), 424–462. https://doi.org/10.1086/316962

Sorenson, O., Rivkin, J. W., & Fleming, L. (2006). Complexity, networks and knowledge flow. *Research Policy*, *35*(7), 994–1017. https://doi.org/10.1016/j.respol.2006.05.002

Stephen, A. T., & Toubia, O. (2009). Explaining the power-law degree distribution in a social commerce network. *Social Networks*, *31*(4), 262–270. https://doi.org/ 10.1016/j.socnet.2009.07.002

Tsai, W. (2001). Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of Management Journal*, *44*(5), 996–1004. https://doi.org/10.2307/3069443

Yamagishi, T., & Cook, K. S. (1993). Generalized exchange and social dilemmas. *Social Psychology Quarterly*, 235–248. https://doi.org/10.2307/2786661