ARTICLE

Check for updates

Modularity of online social networks acts as a reliable predictor of both whole-network and ego-network characteristics over time

Yizhang Zhao[1,4✉], Wei Bai[1,4], Tianyu Qiao[2✉] & Weidong Wang[3✉]

In the digital age, social interactions have increasingly shifted online, necessitating a deeper understanding of the structure and dynamics of online social networks and their societal impacts. This study examines the stability of network modularity on relationship-based social media platforms and its predictive power for both whole-network structures and ego-network characteristics, using a nationally representative longitudinal dataset of Generation Z interactions on a popular social media platform in China. Our findings reveal that network modularity is a stable network attribute over time, which suggests that individuals tend to maintain existing contacts on relationship-based platforms and that the community structures of their online social networks are likely to persist. Notably, the initial level of network modularity significantly correlates with both whole-network and ego-network characteristics in subsequent periods, thus highlighting modularity's power to predict long-term network characteristics. These insights contribute to social network theory by deepening our understanding of how the existence of smaller communities within network structures influences interpersonal interactions in digital communication, with broader implications for how social networks evolve within the landscape of social media.

[1] Tsinghua University, Beijing, China. [2] Peking University, Beijing, China. [3] Renmin University of China, Beijing, China. [4] These authors contributed equally: Yizhang Zhao, Wei Bai ✉email: zhaoyizhang@tsinghua.edu.cn; qiaotianyu@pku.edu.cn; wwd@ruc.edu.cn

Introduction

In the era of digital technologies and social media, a significant amount of social activity has shifted to virtual platforms (Ellison and Boyd 2013; Thulin et al. 2020). As a result, individuals' online networks, formed by their online activities and interactions, have become a reflection, extension and supplement to their offline social networks (Baym 2015). This development has important implications for information spreading, the maintenance of community and the cultivation of social capital in the digital era (Bakshy et al. 2012; Ellison et al. 2007; Ellison et al. 2014). However, our understanding about the following two key questions remains limited: What are the characteristics of online networks on social media platforms, and how do they evolve over time? Answering these questions will deepen the understanding of human interactions in the digital era and enrich the relevant literature.

In general, there are two basic types of social media platforms – relationship-based and public. Relationship-based platforms, which include Facebook, WhatsApp, WeChat and LinkedIn, involve certain existing relationships, whether these are friendships, classmate relationships, family connections or professional affiliations (Ellison et al. 2007; Mangaleswaran 2017). By contrast, public platforms, such as X (formerly Twitter), Instagram, Weibo and TikTok are designed for public interaction and sharing with a broad audience. Even though people can form new connections through relationship-based platforms, most user interactions there are between people who know each other or have mutually agreed to connect. Although the broad connectivity afforded by public platforms has a significant impact on information diffusion and social mobilisation, the connections and interactions on relationship-based platforms represent people's core networks and social capital (Johnston et al. 2013; Harwit 2017). Therefore, research into the structure and dynamics of social networks on relationship-based platforms is of particular importance in the digital era.

Despite the widespread adoption of relationship-based platforms and their prevalence in people's daily lives, they have received far less research attention than public platforms (Kapoor et al. 2018), due primarily to the limited availability of data to researchers. Unlike public platforms, from which data can be more easily 'scraped' and analysed, relationship-based networks require verification from individual users, which makes it challenging to collect relevant data on a large scale. Most datasets in previous studies have featured small-scale, short-duration data collection with a low level of representativeness (Liu et al. 2018; Shen and Gong 2019; Pang 2022). In this paper, we aim to address the knowledge gap using a large-scale longitudinal dataset with a particular focus on network modularity – an important structural attribute that reveals communities or clusters in a network. In an era of widespread connectivity, it is especially relevant to examine whether network modularity tends to change or persist over time because this provides valuable insights into the structure and dynamics of online social networks.

To achieve our research goals, we collected the digital traces of a Generation Z cohort in China on WeChat – an essential tool in the daily lives of many Chinese people, with approximately 1.4 billion monthly active users and a usage rate covering 75% of China's population (Turner 2024). The respondents comprised a representative sample of the 2013 middle school entry cohort, which included all students from 221 middle school classes across China. Online data collection began in 2018, when most respondents were aged 17–18, and continued for three years until 2021. During this period, respondents experienced major life events, such as high school graduation and the COVID-19 pandemic, while we gathered data from their interactions with their former middle school classmates. This dataset enabled us to investigate how Generation Z – digital natives and the primary users of social media (PrakashYadav and Rai 2017) – engage with peers and how their online social networks evolve over time. More importantly, its properties allowed us to examine network dynamics at both the whole-network and ego-network levels.

We found that while network characteristics such as average degree, network density, and clustering coefficient may change over time, network modularity of relationship-based online social networks is significantly stable. This implies that the community divisions in relationship-based online social networks are relatively fixed. The results also showed that initial modularity is significantly correlated with the characteristics of whole networks and ego networks in subsequent periods, which suggests that network modularity can act as a reliable predictor of whole-network and ego-network characteristics over time. Our findings contribute to the existing literature by highlighting the importance of network modularity in the digital context and illuminating the structure and dynamics of online social networks over time.

Literature review

Modularity as a key network attribute. Numerous intricate systems observed in both natural phenomena and societal contexts can be effectively described using the framework of networks (Strogatz 2001; Albert and Barabási 2002; Newman 2003). Although networks may exhibit a diverse range of structures and configurations, the existence of smaller communities or modules within networks appears to be common (Flake et al. 2002; Girvan and Newman 2002; Guimerà et al. 2004). This prevalent characteristic highlights the fact that many networks are organised based on subgroups, which form the underlying organisational structure of complex networks.

Understanding the structures of communities or modules in a network holds meaningful insights, which range from revealing contact-based associations in social networks to identifying customer segments with shared preferences on e-commerce platforms. For instance, community-based analysis can help identify patterns of interaction between different disciplinary groups, revealing mechanisms of knowledge sharing, innovation, and the formation of cross-disciplinary partnerships (Locatelli et al. 2021). Similarly, identifying customer groups with similar interests improves targeted marketing strategies and personalised services on online platforms (Chattopadhyay et al. 2021). In recent decades, the size of social networks has significantly expanded with the rise in social media users, who often form groups or communities, such as circles of friends or interest-based groups, within these networks. Identifying these communities helps clarify the patterns of online interactions and information dissemination (Chouchani and Abed 2020).

Modularity is a crucial network attribute for identifying community structures, and its importance lies in several aspects. First, researchers often use network modularity to measure to what extent a network can be divided into distinct yet cohesive communities or modules (Newman and Girvan 2004; Newman 2006; Chakraborty et al. 2017). Second, it plays a significant role in detecting communities in a given network, the process of which is usually framed as a combinatorial task with the goal of optimising modularity (Porter et al. 2009; Fortunato 2010; Newman 2012; Zhang et al. 2009; Chen et al. 2014). Therefore, modularity is crucial in identifying and understanding the organisational structure within complex networks. Additionally, previous studies suggest that networks with higher modularity may exhibit greater resilience during disruptions (Shekhtman et al. 2015). However, if the overlapping parts between modules

Table 1 Differences between public and relationship-based platforms.

Difference	Public platforms (e.g. X, Instagram)	Relationship-based platforms (e.g. Facebook, WeChat)
Audience type	Broader and unrestricted audience, including both known and unknown individuals, with less emphasis on existing relationships	Audience primarily individuals with existing relationships (e.g. friends, family, colleagues)
Privacy settings	Privacy controls secondary to content visibility. Content often public by default, aimed at a wide audience	Strong emphasis on privacy controls. Content typically shared within a closed or semi-closed group
Primary user intent	Broadcasting content, engaging with a large audience, and gaining public visibility or influence	Sharing personal updates, conducting private and semi-private communication (e.g. messaging, group chats) and nurturing and maintaining existing relationships

fail under stress, the network becomes more fragile and prone to disintegration (Bagrow et al. 2015). Thus, modularity is also a key factor in assessing the resilience and robustness of networks.

Although previous research on networks has highlighted the significance of modularity, little is known about its long-term evolution, and even less is known about the modularity of online social networks. Understanding the modularity of online networks and their long-term changes is essential for network research in the digital era, given their increasing prominence.

Features of relationship-based platforms. Social media platforms can be categorised into public platforms, such as X (formally Twitter), Instagram, Weibo and TikTok, and relationship-based platforms, such as Facebook, WhatsApp, WeChat and LinkedIn. Whereas the former prioritises accessibility by a broad audience and participation in large-scale public discourse, the latter is based primarily on certain existing relationships. Despite sharing common ground in facilitating connections and communications, the two types of platforms have notable differences in terms of audience type, privacy settings and primary user intent, as summarised in Table 1.

In terms of audience type, unlike public platforms that aim for a broad and unrestricted audience, the primary audience on relationship-based platforms consists of limited and trusted groups (Hayes et al. 2016). These groups are based primarily on existing relationships, such as friendships, family ties or professional affiliations (Ellison et al. 2007; Davis et al. 2020). For example, because WeChat's connections are typically formed through personal or professional networks, it is more reflective of users' core social structures (Harwit 2017). By contrast, public platforms allow interactions with both known and unknown individuals, enabling connections to be made with both acquaintances and strangers.

Regarding privacy settings, compared to public platforms, relationship-based platforms place stronger emphasis on user control over personal information. Users can restrict the visibility of their content to selected individuals, thereby enhancing privacy. For example, Facebook allows users to switch their profile privacy from 'public' to 'private', thereby controlling who can view their posts. Similarly, WeChat users' posts on 'Moments' are visible only to approved contacts, and interactions such as 'likes' and comments can be seen only by mutual friends. Meanwhile, WeChat users can further fine-tune the visibility of their posts, setting access limits for specific time periods, such as three days, one month or six months (Huang et al. 2020; Zhang et al. 2022). This control over content visibility ensures personal exchanges within trusted networks, which may lead to an increased perception of trust in both the platform and its members.

Relationship-based platforms also differ from public platforms in terms of primary user intent. Whereas public platforms are designed for broadcasting content, sharing, public discussions and engagement in hashtags or trending posts, relationship-based platforms are typically designed to foster more intimate interactions. Relationship-based platform users' primary intent is generally to partake in private and semi-private communication, share personal updates and interact with existing acquaintances (O'Hara et al. 2014). For the majority of their users, relationship-based platforms are primarily for maintaining and reinforcing existing connections (Subrahmanyam et al. 2008).

These characteristics of relationship-based online platforms increase their likelihood of functioning as an extension of users' offline networks and being a reflection of users' core social networks (Kane et al. 2014; Harwit 2017). Existing studies on relationship-based social platforms have examined how they facilitate intimate, closed social networks as well as their impact on users' behaviours, health, communication patterns and community building (Qi and Wang 2018; Cuesta et al. 2019; Zheng et al. 2022; Zhou et al. 2024). However, little is known about the answers to the following fundamental questions: What are the structures of individuals' online social networks on relationship-based platforms? In an era of widespread connectivity, do online social networks based on existing relationships consist of smaller communities or modules, as they do in offline social networks? If so, how do the modular structures of online social networks change over time? Answering these questions is key to fully grasping the features and implications of online social networks.

Network structure and dynamics over time. Studies on offline social networks have highlighted the importance of network structures for their functions and implications (Klyver et al. 2008; Jackson et al. 2017; Muller and Peres 2019). In addition to network structure, it is also essential to understand how networks change and adapt over time (Holme and Saramäki 2012), because social networks are not static collections of connections, but dynamic systems that expand and evolve, disintegrate and disappear, and reactivate and reconstruct (Chen et al. 2022). For both network structure and network dynamics, the internal modular composition is a crucial aspect, as it not only reflects the fundamental characteristics of the network but also affects changes in the network's structure (Sinha 2014). For instance, researchers have found that modular structures may provide a foundation for social capital accumulation, such as trust between individuals, which is essential for the development of complex societies. Moreover, the potential co-evolution of community structures and cooperative behaviours within the network suggests that modularity could be a key factor in shaping network functions (Marcoux and Lusseau 2013).

Regarding online social networks, most existing studies in this area have focused on public platforms, for which data is more accessible. Community structures within these networks are less stable and more fluid, with modules forming and dissolving as users engage in different conversations, events, or trending topics

(Rossetti and Cazabet 2018). The continuous formation and breakdown of interpersonal connections leads to the constant evolution of communities, which further drives the dynamics of the overall network structure. For example, events such as political protests and natural disasters often drive spikes in online activity and the formation of new communities centred around shared goals or interests (Borge-Holthoefer et al. 2011), which may dissolve a short time after the events fade from the public consciousness.

In terms of relationship-based online social networks, due to the limited access and restricted visibility of content and interactions on these platforms, most studies have either employed a convenience sample that lacks representativeness or relied on questionnaire surveys to obtain users' subjective assessments of their usage of these platforms (Pang 2018; Agrawal 2021; Athukorala 2021; Cao et al. 2024). This has hindered the breadth and depth of the research, creating a notable gap in our understanding of the structures and dynamics of relationship-based online social networks (Weller and Kinder-Kurlanda 2015). Moreover, observing changes over time requires long-term data collection on relationship-based platforms, which becomes even more challenging. Consequently, despite the crucial importance of relationship-based online social networks in people's daily lives, our understanding of their structures and dynamics remains insufficient.

Data and methods
Data. In this study, we constructed a database of social media records from a nationally representative sample of a Generation Z cohort. In 2013, we randomly selected 10,279 Grade 7 students (mostly aged 12–13) from 221 classes across 112 schools in 26 Chinese counties using a probability proportional to size sampling method. The class served as the smallest sampling unit, and all students within the selected classes were invited to participate in the baseline survey, with follow-up surveys conducted in 2014, 2015, 2017, and 2019. Over the years, due to the reallocation of students into different classes and the recombination of students when transitioning from middle school to high school, widespread connections among the respondents were formed beyond original class and school boundaries, but more than 99% of the connections remained within the same county. Therefore, we focused on the within-county interactions and used the county as the unit for analysis.

In 2018, when most participants were 17–18 years old, we requested their consent to access their WeChat posts and interactions with their former middle school classmates. WeChat was chosen because it is China's predominant social media platform, with over one billion monthly active users. A total of 8636 participants provided their consent. From May 2018 to April 2021, we collected all the respondents' publicly shared posts and documented their interactions with their middle school classmates. The dataset aggregates 1,048,999 public posts and 1,030,710 interactions within their peer network. As we had data from only 8636 out of 10,279 participants, we implemented a weighting procedure to ensure that the distributions of key parameters aligned with those of the overall sample. Our findings showed that the weighted results were similar to those of the unweighted results.

Methods
Social network analysis. In this study, we focused on the changes in network modularity as well as several other network characteristics over time to examine the dynamics of network structure. We introduce the indicators below, and the formulations of each indicator are detailed in the Supplementary Note.

To reflect the connectivity of the network, we calculated the average degree, the average degree of neighbours, and the network density in each county. The average degree refers to the average number of connections or edges each individual (node) has. The average degree of neighbours measures the average number of connections of an individual's neighbours, showing whether the nodes are connected to other nodes that are themselves well connected. The network density measurement quantifies how many edges are present in a network relative to the maximum number of edges that could possibly exist, which captures the global connectivity of the entire network.

To measure the aggregation of the network, we used the average clustering coefficient as the indicator. The clustering coefficient is a network metric that measures the extent to which nodes in a network tend to cluster together. The average clustering coefficient is the mean of the clustering coefficients of all the nodes in the network, which provides an overall measure of the extent of clustering in the network.

Finally, to indicate the overall accessibility of nodes in a network, we calculated the average closeness centrality. The closeness centrality of a node measures how close a node is to all other nodes in the network, based on the shortest paths between them. The average closeness centrality takes the average value of the closeness centrality across all nodes in the network, reflecting how efficiently nodes can reach other parts of the network.

Network modularity measures the extent to which a network can be divided into non-overlapping communities, where nodes within each community have more connections than with nodes outside the community. We used a greedy modularity maximisation algorithm to calculate the network modularity, which identifies communities in a network by iteratively improving the division of nodes into groups (Newman 2004). In practice, if the value is greater than 0.3, it suggests a significant community structure in the network.

Dynamics of ego networks. To depict the dynamics of ego networks over time, we set an examination window of three months. We not only examined the dynamics of the whole-network structure in this manner over a three-year period but also calculated two indicators to depict the mobility of ego networks. One of the indicators is the proportion of nodes whose neighbours in the current time window remain completely unchanged compared with the previous time window, which is defined as follows:

$$P_m^{\mu+1} = \frac{\sum_{i=1}^{N_m} \lambda(s_i^\mu, s_i^{\mu+1})}{N_m} \tag{1}$$

where s_i^μ is the set of neighbours that individual node i interacts with in Period μ; $\lambda(s_i^\mu, s_i^{\mu+1})$ is the Kronecker delta, which takes the value of 1 if $s_i^\mu = s_i^{\mu+1}$ (i.e. node i interacts with the same set of neighbours in Period $\mu + 1$ as in Period μ) and 0 otherwise.

The other indicator is the average number of new contacts that an individual node has in the current time window compared with the previous period, as shown in the following equation:

$$\eta_m^{\mu+1} = \frac{\sum_{i=1}^{N_m} \left| s_i^{\mu+1} - s_i^\mu \right|}{N_m} \tag{2}$$

where $|s_i^{\mu+1} - s_i^\mu|$ denotes the number of contacts that node i interacts with only in Period $\mu + 1$ but not in Period μ.

Results
The stability of network modularity. We first examined the evolution of online social networks in 26 counties across China every three months over a three-year period from 2018 to 2021. We then calculated the coefficient of variation (CV) in each
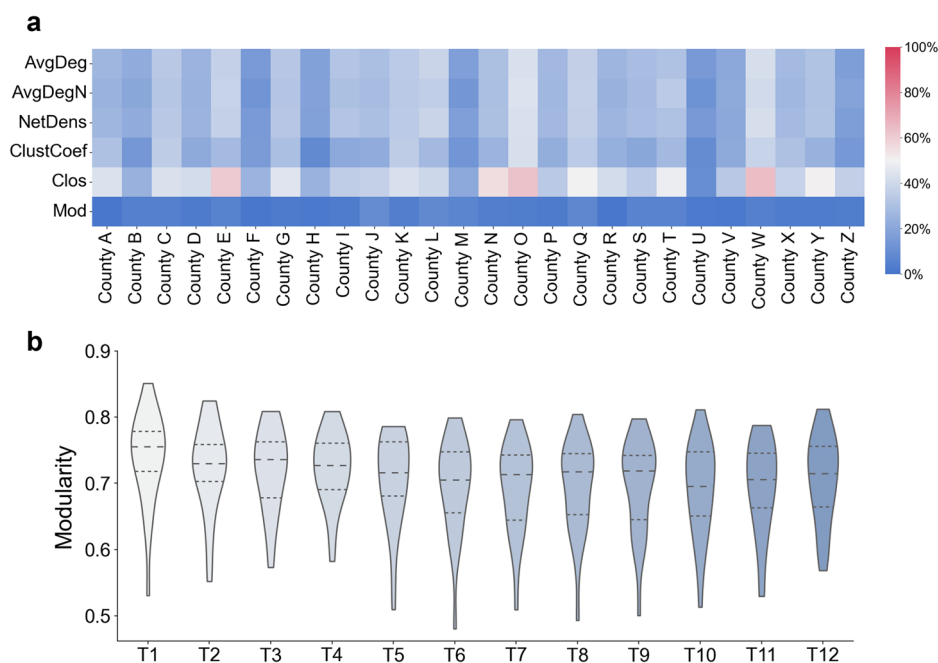
Fig. 1 The stability of network modularity across different counties and the distributions over time. a The coefficients of variation (CV) of different network attributes across 26 counties. The abbreviated terms in the figure are AvgDeg (average degree), AvgDegN (average degree of neighbours), NetDens (network density), ClustCoef (clustering coefficient), Clos (closeness centrality), and Mod (modularity). **b** The distributions of network modularity over the 12 periods, with the mean values and upper and lower quartiles plotted. Tn represents the nth period. $N = 8636$.

cluster, which measures the level of relative variability of network attributes. The CV is defined as the ratio of the standard deviation to the mean, with a value below 10% typically indicating a low level of variability. As shown in Fig. 1a, among all network characteristics examined in this study, network modularity – the measure of the strength of community divisions within a network – is the most stable feature over time. While the CV values of other network characteristics vary from 8% to 64%, the CV values of network modularity consistently fall below 10% across all counties. The significant stability of network modularity suggests that once community structures are established in a social network, they are highly likely to persist over the long term.

Figure 1b presents the distributions of network modularity over the 12 periods. First, it shows that the network modularity values for almost all counties exceed 0.5, and the average values are approximately 0.7 over the 12 periods, which indicates the prevalence and persistence of a pronounced community structure within counties. In addition, compared to the other network attributes (see Supplementary Appendix Fig. S1 for the distributions of other network attributes), whose average values fluctuate with the frequency of WeChat posts, network modularity does not exhibit such fluctuation. For example, the sixth period covered summer holidays, during which the respondents updated more posts and interacted more frequently with their WeChat contacts, which led to changes in all network attributes except for modularity. This suggests that the modularity of online networks is not affected by the level of activity of platform users.

Relationship between initial network modularity and other network attributes in subsequent periods. Given that network modularity exhibits a high level of stability, we examined whether the initial value of network modularity can be used to predict other network attributes. Figure 2 illustrates the correlation between the network modularity in the first period (May 2018 to July 2018) and other network attributes in subsequent periods. It shows that except for the clustering coefficient, network

modularity has a stable correlation with all other network attributes. Specifically, as shown in Fig. 2b, the correlation coefficients between initial network modularity and average degree, the average degree of neighbours, network density, and closeness centrality during the seventh period (November 2019 to January 2020) are -0.49, -0.45, -0.64, and -0.44, respectively (p-value < 0.05). These correlation coefficients are similar to those of the twelfth period (February 2021 to April 2021), which are -0.42, -0.41, -0.59, and -0.40 (p-value < 0.05), respectively, as indicated in Fig. 2c.

The consistent negative correlations suggest that networks with more pronounced community divisions tend to have relatively poorer connectivity and cohesiveness. More importantly, despite fluctuations in the values of most network attributes, their relationships with network modularity remain stable over time and are not affected by the number of posts and the frequency of interactions at different times. Given the above findings and considering that modularity is a highly stable network attribute, we may use the modularity of a network at any given time to predict various network attributes in the future, thereby eliminating the need for prolonged observation.

Dynamics of ego networks. Having examined the relationship between the modularity of online social networks and the long-term structural characteristics of the whole network, we further investigated whether modularity affects the dynamics of ego networks over time. To this end, we constructed two indicators to evaluate ego-network turnover.

One is the proportion of individuals whose contacts remain completely unchanged compared with the previous time window. For example, a value of 50% indicates that half the individuals in the analytical sample interact exclusively with the same contacts as in the previous three months, without losing or adding any new contacts. Using a window of three months at a time, we obtained 11 values for each county over the three-year period. Figure 3a presents the mean values of this indicator and the 95%
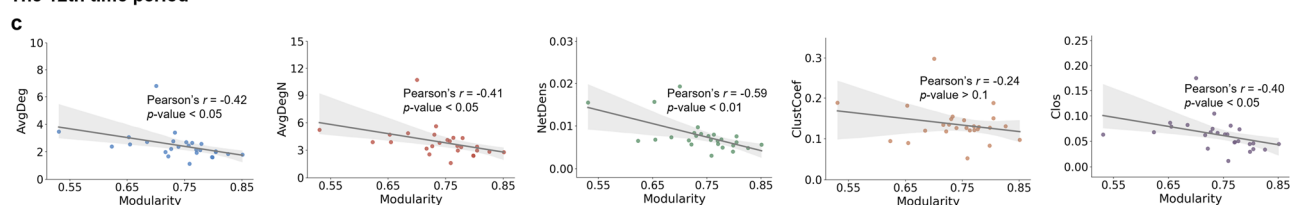
The 1st time period

a
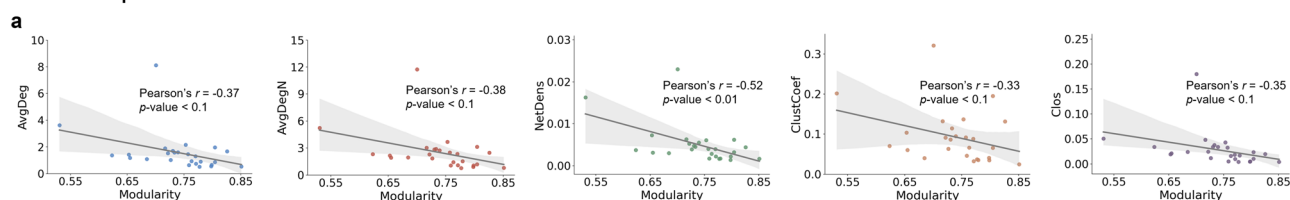

The 7th time period

**b**
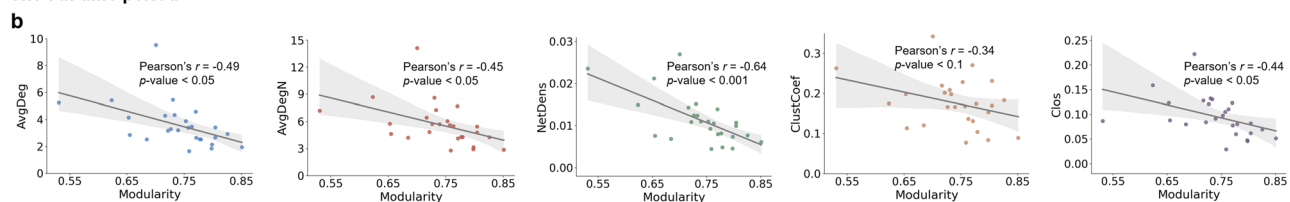
The 12th time period

c

Fig. 2 The relationship between network modularity at T1 and other network attributes at T1, T7, and T12. The correlation between the network modularity at T1 (May 2018 to July 2018) and other network attributes across 26 counties at T1 (**a**), T7 (**b** November 2019 to January 2020) and T12 (**c** February 2021 to April 2021). The lines are linear fits based on scatterplots, with the 95% confidence interval shaded. $N = 8636$.

The proportion of individuals who maintain the same set of contacts as in the previous period

a

b

The average number of new contacts per person

c

d
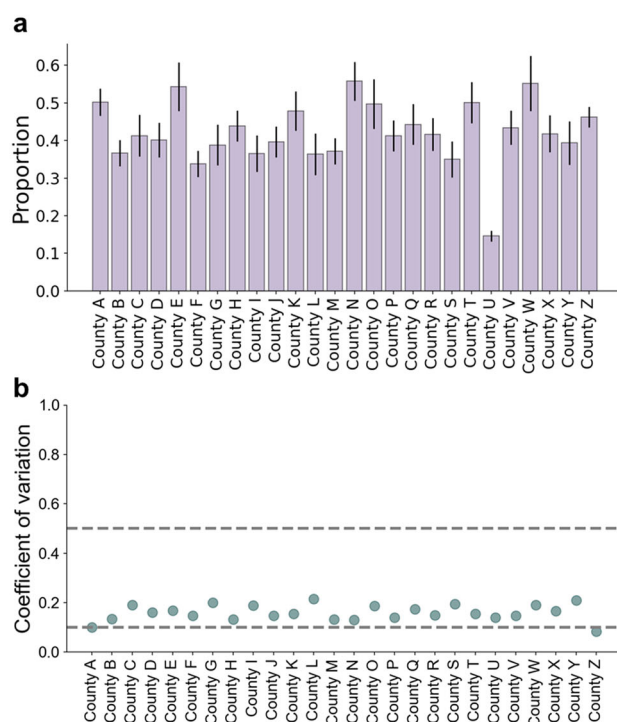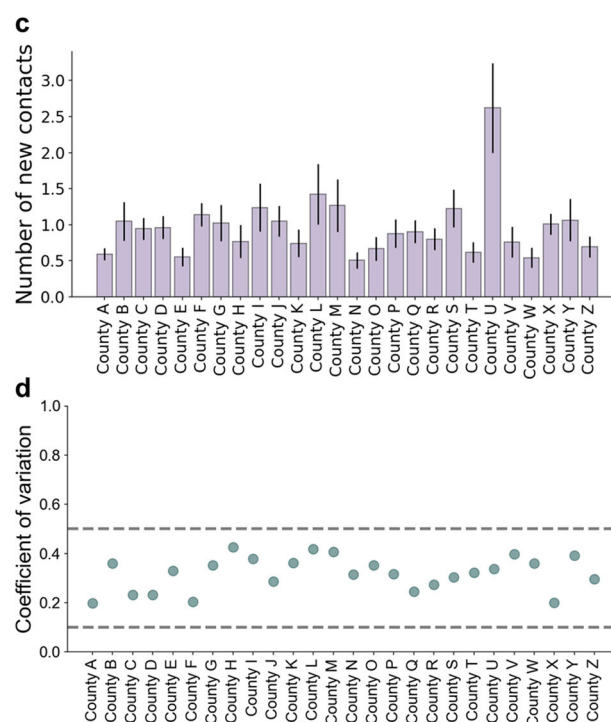
Fig. 3 The ego-network turnover across 26 counties. a The average proportion of individuals who maintain the same set of contacts as in the previous period. **b** The coefficients of variation for the proportion of individuals whose contacts remain unchanged compared with the previous period. **c** The average number of new contacts per person compared with the previous period. **d** The coefficients of variation for the average number of new contacts per person compared with the previous period. **a**, **c** Error bars indicate 95% confidence intervals. $N = 8636$.

confidence intervals for each county. The analysis shows that the average proportion varies across different counties, with the maximum value at 55% and the minimum value at 15%. It suggests that in County U, only 15% of the respondents maintained consistent interactions with the same group of contacts from one period to the next, but that this figure more than tripled in County N.

We also calculated the average number of new contacts that an individual has in the current period compared with the previous one, as shown in Fig. 3b. The results reveal that in County U, a respondent develops approximately 2.5 new connections in a succeeding period on average, while in County N, the figure is around 0.5. These findings align with those related to the first indicator, that ego-network turnover varies across different counties. In addition, as shown in Fig. 3c, the CV values are around 16% and 32% for the two indicators, respectively, which suggests a relatively high variability over time in most counties. Additional analyses showed that the weighted results are similar to the unweighted ones (see Supplementary Appendix Fig. S2 for details).
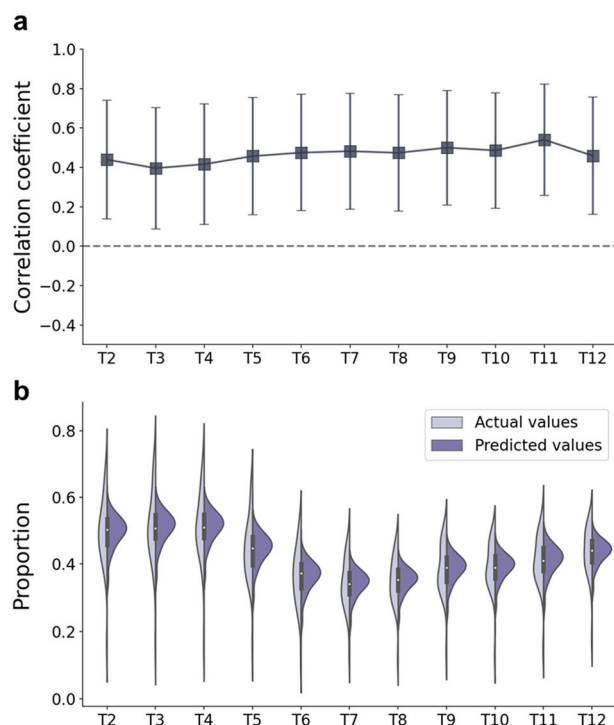
Relationship between initial network modularity and ego-network characteristics over time. We examined whether whole-network modularity predicts ego-network characteristics in the long term. Figure 4 shows the relationship between the modularity of the whole network in the first period and the two indicators of ego-network turnover in subsequent periods.

Figure 4a displays the correlation coefficients between network modularity and the proportion of individuals who maintain the

same set of contacts as in the previous period, which shows a significant positive and stable correlation over time. Figure 4c presents the correlation coefficients between network modularity and the average number of new contacts per person in the succeeding period, which shows a significant negative correlation in most cases. Figure 4b, d illustrate how the two indicators of ego-network turnover would be distributed using initial network modularity as a predictor, compared with the actual data distribution in different periods. Despite the distributions of predicted values being more concentrated than those of actual values, the predictions demonstrate a high degree of accuracy and consistency across different periods, which aligns with the results shown in Fig. 4a, c. The results collectively indicate that in networks with a more pronounced community structure, individuals are less likely to experience ego-network turnover, which means they are more likely to maintain existing relationships and less likely to develop new connections. We conducted additional analysis by applying the weighting parameter and obtained similar results (see Supplementary Appendix Fig. S3 for details).

In addition, the predictive power of modularity remains consistent over time. The initial level of modularity effectively predicts ego-network turnover in three years' time. In networks with a more pronounced initial community structure, network members are more likely to maintain long-term, stable relationships with one another within these communities three years later. This reveals an important connection between whole-network structure and ego-network dynamics. Furthermore, given the stability of network modularity over time, we may also replace the initial modularity of a network with modularity at any given time to predict ego-network dynamics in the long term.

The proportion of individuals who maintain the same set of contacts as in the previous period

The average number of new contacts per person
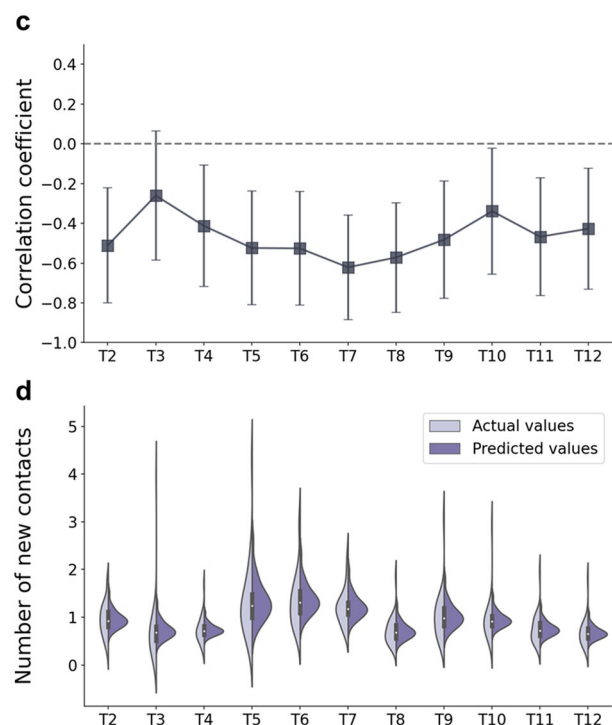
Fig. 4 The relationship between initial network modularity and ego-network characteristics in subsequent periods. a The correlations between the initial network modularity and the proportion of individuals who maintain the same set of contacts as in the previous period. **b** Comparison of actual values and predicted values that use the initial network modularity as a predictor for the proportion of individuals who maintain the same set of contacts as in the previous period. **c** The correlations between the initial network modularity and the average number of new contacts per person compared with the previous period. **d** Comparison of actual values and predicted values that use the initial network modularity as a predictor for the average number of new contacts per person compared with the previous period. **a**, **c** Error bars indicate 95% confidence intervals. $N = 8636$.

Discussion

In the digital era, people increasingly rely on social media platforms for connection and communication, and online social networks have become an essential aspect of modern social interaction (Heidemann et al. 2012). In particular, relationship-based social networks increasingly serve as a crucial element of personal communication by offering new avenues for maintaining social connections in the digital age. However, despite the growing recognition of the importance of online social networks, little is known about their structures and dynamics over time (Weller and Kinder-Kurlanda 2015).

To facilitate a better understanding of online social networks, this study employed a nationally representative sample of a Generation Z cohort taken from 221 middle school classes across 26 counties in China. Based on respondents' online interactions on WeChat as they approached or reached adulthood, we constructed 26 networks to examine the characteristics and dynamics of online networks over time. We focused particularly on one important network attribute – network modularity – examined its stability over time, and explored its predictive power for the whole-network structure and ego-network dynamics over time.

We found that, first, among various attributes of the relationship-based online networks, modularity remains consistently stable in the long term. In contrast, attributes such as network density and clustering coefficient are significantly affected by the number of updated posts and follow-up interactions within a given period, which led to fluctuations not observed in modularity. More importantly, network modularity has a relatively stable correlation with whole-network structures and ego-network dynamics in the long term, which reveals the predictive power of this indicator on long-term network characteristics.

In particular, the predictive power of network modularity on ego-network turnover reminds us of the shaping effect of whole-network structure on personal connections. Members of networks with pronounced community structures have long-term and stable relationships but fewer new connections beyond their initial circles. Such network structures are beneficial for small-group cohesion, but may impose obstacles to larger-scale collaboration. They may also exert a long-term impact with important implications for the cultivation of social capital in the digital age (Burt et al. 2022).

As for why modularity tends to remain stable over time, the underlying mechanisms still require further exploration. Research based on offline networks has found that the frameworks for individuals to interact in, provided by formal organisational structures, along with homophilic interactions between individuals, are factors that promote the formation of communities within networks (McPherson et al. 2001; Monge et al. 2008). Furthermore, social capital plays a key role in maintaining these community structures. By providing emotional, informational and material support to community members and fostering collective efficacy, social capital strengthens reciprocity and enhances cohesion within the community, thereby increasing the stability of community structures (Wellman and Wortley 1990; Sampson et al. 1997; Temkin and Rohe 1998). Future research could explore whether these mechanisms apply to online networks and compare these with those in offline networks.

This study has several limitations. First, our analysis is limited to WeChat as a single platform, and it is unclear whether the conclusions are specific to WeChat or applicable to other relationship-based platforms as well. Future research could examine other relationship-based platforms in different digital and cultural contexts to explore whether similar patterns exist and to provide a more comprehensive understanding of the role of network modularity. Second, considering the characteristics of relationship-based platforms, people are more likely to interact within their intimate social circles on these platforms, which may lead to a persistence of community division in such networks. It is not yet known to what extent these findings may apply to public platforms. Previous studies have suggested that Twitter users have highly dynamic personal networks, with a large percentage of weak ties and high turnover (Arnaboldi et al. 2013). Future studies may explore the dynamics of modularity on public social media platforms and compare these with relationship-based ones. Third, the research is based on social networks formed between middle school students and their classmates. Whether the findings could be extended to other networks formed at different life stages is worthy of further exploration.

Despite these limitations, this paper contributes to the existing literature in several ways. First, by presenting the structural characteristics of relationship-based online networks based on a national representative sample, it reveals the structural characteristics and dynamics of relationship-based online networks of the Generation Z cohort in China. The Generation Z cohort is the first generation of internet natives who rely on social media platforms to maintain their social networks. This pattern may also apply to succeeding generations, even though they are not covered in the current dataset. Therefore, the findings of this study may shed light on the evolution of online social networks for generations to come.

Second, this study overcomes the limitations of previous offline surveys, which faced obstacles in tracking and monitoring the long-term, large-scale evolution of whole networks, thereby advancing our understanding of the long-term dynamics of network modularity. In addition to identifying community structures and revealing patterns of interactions (Newman and Girvan 2004; Newman 2006; Porter et al. 2009; Chakraborty et al. 2017; Locatelli et al. 2021), this paper highlights the stability of modularity over time, as well as the capacity of network modularity to predict the long-term conditions of other metrics at both the whole-network and ego-network levels. Although the network in this study is based on classmate relationships, which have their own unique characteristics, the presence of smaller communities or groups is common across many types of networks. Our study encourages future research to focus more on this network attribute and its role in examining long-term network evolution.

Third, the paper facilitates greater understanding of how online network structure may affect interpersonal interactions, thereby highlighting some of the mechanisms underlying the spread of information and ideas in the new context of social media. Particularly in an era of broad connectivity, smaller communities still persist in individuals' social networks. This suggests that the spread of information and ideas may still be strongly influenced by such network structures, which may amplify echo chambers and exacerbate idea polarisation. Therefore, in addition to the effect of the algorithmic recommendation system, policymakers or social media designers may also need to consider the impact of network structures. Overall, the findings of this study not only deepen our understanding of network structures on relationship-based platforms, but also remind us of the widespread impact of such structures in the context of the digital world.

Data availability

The data supporting the findings of this study are available upon request from the corresponding author [W.W.], subject to compliance with the laws of the People's Republic of China and approval from the Office of Research at Renmin University of China. The data are not publicly available due to their sensitive nature and the potential for identification, which could compromise research participant privacy.

References

Agrawal SR (2021) Adoption of WhatsApp for strengthening internal CRM through social network analysis. J Relat Mark 20(4):261–281

Albert R, Barabási AL (2002) Statistical mechanics of complex networks. Rev Mod Phys 74(1):47

Arnaboldi V, Conti M, Passarella A, Dunbar R (2013) Dynamics of personal social relationships in online social networks: a study on twitter. In: Proceedings of the first ACM conference on online social networks. ACM, pp. 15–26

Athukorala AWV (2021) Usage of WeChat by International Students and Its Usefulness: Study Based in Wuhan University China. Int J Sci Res Sci Technol 8(6):66–75

Bagrow JP, Lehmann S, Ahn YY (2015) Robustness and modular structure in networks. Netw Sci 3(4):509–525

Bakshy E, Rosenn I, Marlow C, Adamic L (2012) The role of social networks in information diffusion. In: Proceedings of the 21st international conference on World Wide Web. ACM, pp. 519–528

Baym NK (2015) Personal connections in the digital age. John Wiley & Sons

Borge-Holthoefer J, Rivero A, García I, Cauhé E, Ferrer A, Ferrer D, Moreno Y (2011) Structural and dynamical patterns on online social networks: the Spanish May 15th movement as a case study. PLoS One 6(8):e23883

Burt RS, Opper S, Holm HJ (2022) Cooperation beyond the network. Organ Sci 33(2):495–517

Cao C, Meng Q, Zhang H (2024) A longitudinal examination of WeChat usage intensity, behavioral engagement, and cross-cultural adjustment among international students in China. High Educ 87(3):661–683

Chakraborty T, Dalmia A, Mukherjee A, Ganguly N (2017) Metrics for community analysis: a survey. ACM Comput Surv 50(4):1–37

Chattopadhyay S, Basu T, Das AK, Ghosh K, Murthy LC (2021) Towards effective discovery of natural communities in complex networks and implications in e-commerce. Electron Commer Res 21(4):917–954

Chen H, Mehra A, Tasselli S, Borgatti SP (2022) Network dynamics and organizations: a review and research agenda. J Manag 48(6):1602–1660

Chen M, Kuzmin K, Szymanski BK (2014) Community detection via maximization of modularity and its variants. IEEE Trans Comput Soc Syst 1(1):46–65

Chouchani N, Abed M (2020) Online social network analysis: detection of communities of interest. J Intell Inf Syst 54(1):5–21

Cuesta M, Eklund M, Rydin I, Witt AK (2019) Using Facebook as a co-learning community in higher education. In: Social media and education. Routledge Press, pp. 55–72

Davis J, Wolff HG, Forret ML, Sullivan SE (2020) Networking via LinkedIn: an examination of usage and career benefits. J Vocat Behav 118:103396

Ellison NB, Boyd DM (2013) Sociality through social network sites. In: Dutton WH (ed), The Oxford handbook of internet studies. Oxford University Press, pp. 151–178

Ellison NB, Steinfield C, Lampe C (2007) The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. J Comput-Mediated Commun 12(4):1143–1168

Ellison NB, Vitak J, Gray R, Lampe C (2014) Cultivating social resources on social network sites: Facebook relationship maintenance behaviors and their role in social capital processes. J Comput Mediated Commun 19(4):855–870

Flake GW, Lawrence S, Giles CL, Coetzee FM (2002) Self-organization and identification of web communities. Computer 35(3):66–70

Fortunato S (2010) Community detection in graphs. Phys Rep 486(3-5):75–174

Girvan M, Newman ME (2002) Community structure in social and biological networks. Proc Natl Acad Sci USA 99(12):7821–7826

Guimerà R, Sales-Pardo M, Amaral LAN (2004) Modularity from fluctuations in random graphs and complex networks. Phys Rev E 70(2):025101

Harwit E (2017) WeChat: Social and political development of China's dominant messaging app. Chin J Commun 10(3):312–327

Hayes RA, Carr CT, Wohn DY (2016) It's the audience: differences in social support across social media. Soc Media Soc 2(4):2056305116678894

Heidemann J, Klier M, Probst F (2012) Online social networks: a survey of a global phenomenon. Comput Netw 56(18):3866–3878

Holme P, Saramäki J (2012) Temporal networks. Phys Rep 519(3):97–125

Huang X, Vitak J, Tausczik Y (2020) "You don't have to know my past": how WeChat Moments users manage their evolving self-presentation. In: Proceedings of the 2020 CHI conference on human factors in computing system. ACM, pp. 1–13

Jackson MO, Rogers BW, Zenou Y (2017) The economic consequences of social-network structure. J Econ Lit 55(1):49–95

Johnston K, Tanner M, Lalla N, Kawalski D (2013) Social capital: the benefit of Facebook 'friends. Behav Inf Technol 32(1):24–36

Kane GC, Alavi M, Labianca G, Borgatti SP (2014) What's different about social media networks? A framework and research agenda. MIS Q 38(1):275–304

Kapoor KK, Tamilmani K, Rana NP, Patil P, Dwivedi YK, Nerur S (2018) Advances in social media research: past, present and future. Inf Syst Front 20:531–558

Klyver K, Hindle K, Meyer D (2008) Influence of social network structure on entrepreneurship participation—a study of 20 national cultures. Int Entrep Manag J 4:331–347

Liu L, Qu B, Chen B, Hanjalic A, Wang H (2018) Modelling of information diffusion on social networks with applications to WeChat. Phys A 496:318–329

Locatelli B, Vallet A, Tassin J, Gautier D, Chamaret A, Sist P (2021) Collective and individual interdisciplinarity in a sustainability research group: a social network analysis. Sustain Sci 16:37–52

Mangaleswaran R (2017) Fulfilling social relationship needs – a descriptive study on users of Whatsapp. Int J Interdiscip Res Res Arts Humanit 2:175–181

McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: homophily in social networks. Annu Rev Socio 27(1):415–444

Monge P, Heiss BM, Margolin DB (2008) Communication network evolution in organizational communities. Commun Theory 18(4):449–477

Marcoux M, Lusseau D (2013) Network modularity promotes cooperation. J Theor Biol 324:103–108

Muller E, Peres R (2019) The effect of social networks structure on innovation performance: a review and directions for research. Int J Res Mark 36(1):3–19

Newman ME (2003) The structure and function of complex networks. SIAM Rev 45(2):167–256

Newman ME (2004) Fast algorithm for detecting community structure in networks. Phys Rev E 69(6):066133

Newman ME, Girvan M (2004) Finding and evaluating community structure in networks. Phys Rev E 69(2):026113

Newman ME (2006) Modularity and community structure in networks. Proc Natl Acad Sci USA 103(23):8577–8582

Newman ME (2012) Communities, modules and large-scale structure in networks. Nat Phys 8(1):25–31

O'Hara KP, Massimi M, Harper R, Rubens S, Morris J (2014) Everyday dwelling with WhatsApp. In: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing. ACM, pp. 1131–1143

Pang H (2018) WeChat use is significantly correlated with college students' quality of friendships but not with perceived well-being. Heliyon 4(11):e00967

Pang H (2022) Connecting mobile social media with psychosocial well-being: Understanding relationship between WeChat involvement, network characteristics, online capital and life satisfaction. Soc Netw 68:256–263

Porter MA, Onnela JP, Mucha PJ (2009) Communities in networks. Not Am Math Soc 56(9):1082–1097

PrakashYadav G, Rai J (2017) The Generation Z and their social media usage: a review and a research outline. Glob J Enterp Inf Syst 9(2):110–116

Qi GY, Wang Y (2018) Investigating the building of a WeChat-based community of practice for language teachers' professional development. Innov Lang Learn Teach 12(1):72–88

Rossetti G, Cazabet R (2018) Community discovery in dynamic networks: a survey. ACM Comput Surv (CSUR) 51(2):1–37

Sampson RJ, Raudenbush SW, Earls F (1997) Neighborhoods and violent crime: a multilevel study of collective efficacy. Science 277(5328):918–924

Shekhtman LM, Shai S, Havlin S (2015) Resilience of networks formed of interdependent modular networks. N J Phys 17(12):123007

Shen C, Gong H (2019) Personal ties, group ties and latent ties: connecting network size to diversity and trust in the mobile social network WeChat. Asian J Commun 29(1):18–34

Sinha S (2014) The importance of community: how modular organization of social networks affects their collective dynamics. Stud Microecon 2(1):49–61

Strogatz SH (2001) Exploring complex networks. Nature 410(6825):268–276

Subrahmanyam K, Reich SM, Waechter N, Espinoza G (2008) Online and offline social networks: use of social networking sites by emerging adults. J Appl Dev Psychol 29(6):420–433

Temkin K, Rohe WM (1998) Social capital and neighborhood stability: an empirical investigation. Hous Policy Debate 9(1):61–88

Thulin E, Vilhelmson B, Schwanen T (2020) Absent friends? Smartphones, mediated presence, and the recoupling of online social contact in everyday life. Ann Am Assoc Geogr 110(1):166–183

Turner A (2024) How many people use WeChat? User Statistics & Trends, https://www.bankmycell.com/blog/number-of-wechat-users/

Wellman B, Wortley S (1990) Different strokes from different folks: community ties and social support. Am J Socio 96(3):558–588

Weller K, Kinder-Kurlanda K (2015) Uncovering the challenges in collection, sharing and documentation: the hidden data of social media research? Proc Int AAAI Conf Web Soc Media 9(4):28–37

Zhang XS, Wang RS, Wang Y, Wang J, Qiu Y, Wang L, Chen L (2009) Modularity optimization in community detection of complex networks. Europhys Lett 87(3):38002

Zhang Y, Ran X, Luo C, Gao Y, Zhao Y, Shuai Q (2022) Only visible for three days": Mining microblogs to understand reasons for using the Time Limit setting on WeChat Moments. Comput Hum Behav 134:107316

Zheng C, Song X, Li J, Chen Y, Dong T, Yang S (2022) Why do we "like" on WeChat moments: the effects of personality traits and content characteristics. Front Psychol 13:772547

Zhou Y, Abuduxukuer KS, Wang C et al. (2024) WeChat usage and preservation of cognitive functions in middle-aged and older Chinese adults: indications from a nationally representative survey, 2018–2020. BMC Public Health 24(1):1783

Acknowledgements

Author contributions

Y.Z., W.B., T.Q. and W.W jointly conceptualised the study. W.W. collected data and constructed the dataset for analysis. Y.Z., W.B. and T.Q. conducted the data analysis. Y.Z. and W.B. wrote the first draft of the manuscript. All authors reviewed and edited the final draft and approved the final version for publication.

Competing interests

The authors declare no competing interests.

Ethics approval

Ethical approval was provided by the Ethics Committee of the National Survey Research Center at Renmin University of China on April 11 2018 (RN2018041101), confirming that the study complies with ethical review requirements for scientific research and authorising its conduct in accordance with the approved research protocol. The study was conducted in accordance with relevant guidelines of the Declaration of Helsinki.

Informed consent

Informed consent was obtained from all participants. Between April 15 and 30, 2018, with the approval of the Ethics Committee of the National Survey Research Center at Renmin University of China, participants were invited in writing to grant the research team permission to access their online posts and interactions with their middle school classmates on WeChat for further academic research. The invitation message included the following statement, 'If you agree to participate, please provide your WeChat ID, which we will consider as your informed consent.' Participants were informed of the purpose of the research, potential risks, data protection measures, and their right to withdraw at any time. All survey participants were aged 16 or above at the time of invitation, and their individual consent was sufficient for participation in this study in accordance with the current legal framework in China.

Additional information

Correspondence and requests for materials should be addressed to Yizhang Zhao, Tianyu Qiao or Weidong Wang.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.