The Length of Bridge Ties: Structural and Geographic Properties of Online Social Interactions

Yana Volkovich
Barcelona Media Foundation
yana.volkovich@barcelonamedia.org

Salvatore Scellato
University of Cambridge
ss824@cam.ac.uk

David Laniado
Barcelona Media Foundation
david.laniado@barcelonamedia.org

Cecilia Mascolo
University of Cambridge
cm542@cam.ac.uk

Andreas Kaltenbrunner
Barcelona Media Foundation
andreas.kaltenbrunner@barcelonamedia.org

Abstract

The popularity of the Web has allowed individuals to communicate and interact with each other on a global scale: people connect both to close friends and acquaintances, creating ties that can bridge otherwise separated groups of people. Recent evidence suggests that spatial distance is still affecting social links established on online platforms, with online ties preferentially connecting closer people.

In this work we study the relationships between interaction strength, spatial distance and structural position of ties between members of a large-scale online social networking platform, Tuenti. We discover that ties in highly connected social groups tend to span shorter distances than connections bridging together otherwise separated portions of the network. We also find that such bridging connections have lower social interaction levels than ties within the inner core of the network and ties connecting to its periphery. Our results suggest that spatial constraints on online social networks are intimately connected to structural network properties, with important consequences for information diffusion.

Introduction

Social science has been analyzing and discussing the relationship between spatial distance and social structure for more than 60 years. There is evidence that closer pairs of individuals are more likely to develop social bonds than distant ones (Merton 1948; Festinger, Schachter, and Back 1950). In particular, the probability that two individuals engage in social interactions quickly decays as an inverse power of their relative geographic distance (Stewart 1941). Indeed, it seems that social structure does not escape the first law of geography: *"everything is related to everything else, but near things are more related than distant things"* (Tobler 1970).

Even if spatial proximity strongly affects how people create and arrange their social ties, the tools and techniques developed to understand and analyze social network structure have largely ignored such spatial constraints. The main focus of social network analysis has been to study the structure of interpersonal ties by considering mainly the topological properties of the resulting network: individuals are represented as nodes in an abstract, dimensionless space, connected by pairwise relations (Wasserman and Faust 1994). While ignoring spatial distances might bear no influence in some geographically limited cases, this is not possible when dealing with large-scale online social systems.

The social structure of online networks

The wealth of methods and procedures of social network analysis have been used to study and analyze the social structure arising among users of online social services (Kumar, Novak, and Tomkins 2006; Mislove et al. 2007; Ahn et al. 2007). A strong difference with respect to the previous sociological studies is that these systems accumulate millions of users across all the planet, allowing researchers to easily gather large amounts of data about online interactions. Hence, space and distance might be important factors shaping these online social networks.

On the other hand, the online nature of these systems has inspired many authors to suggest that physical distance might be losing its influence: as online communication tools become faster and more reliable, allowing people to interact regardless of where they are located, the costs imposed by geographic distances might be vanishing. As a result, distance might be effectively *"dead"*, ceasing to play a role in online communication (Cairncross 2001). Nonetheless, some recent initial results demonstrate that the probability of social connection between two individuals on online social networking services still decreases with their geographic distance, although the exact relationship between these two variables is still unclear (Liben-Nowell et al. 2005; Backstrom, Sun, and Marlow 2010). Overall, the evidence is that social factors or spatial factors alone cannot capture the properties of real-world social systems over space (Scellato et al. 2011).

Structure, tie strength and distance

Another important aspect of social networks, even in their online versions, is that not all ties are equal: each bond between two individuals can be characterized by a particular level of interaction strength, denoting, for instance, whether they are close friends or just acquaintances. In online social services the amount of information about user interactions can be successfully used to quantitatively estimate how much an online connection binds two users together (Kahanda and Neville 2009; Gilbert and Karahalios 2009). The importance of tie strength is directly connected to the hypothesis that ties with different strength occupy different positions in the network structure: weak ties are more likely to connect together otherwise separated portions of a network, playing an important role in information diffusion[1] and resilience to network damage (Granovetter 1973; Onnela et al. 2007). The fact that some social ties act as bridges between otherwise separate communities, closing "structural holes" that would otherwise disconnect the social fabric, has been widely discussed in sociology. The evidence suggests that individuals in these structural positions can be more powerful or more innovative (Burt 1992). Yet, space is constraining network structure as well: in fact, social communities tend to be limited in their geographic span, denoting a potential relationship between structural position and geographic distance (Onnela et al. 2011).

The plausible interaction between tie strength, social structure and spatial distance is therefore evident: the strength of a tie is related to the position of that link within the network, while the spatial distance between two individuals affects how likely they are to be connected. Thus, these three properties represent three different facets of a single system which combines spatial and social factors and binds together individuals, affecting complex processes such as the spreading of information over social links (Rogers 1995; Newman, Barabasi, and Watts 2006) or the ability to navigate the social networks to route a message to a particular individual (Kleinberg 2000). Nonetheless, the research community still lacks a broad understanding of the interplay between the structure of a social network, the strength of its ties and the space that embeds it.

Our work

Given the importance of spatial distance on online social interaction, and the important structural properties arising from the interplay between tie strength and network structure, the main research question we address in this work is: *what is the relationship between the structural properties of online social ties and the spatial distance they span?* This question is purposefully generic, as we aim to study a series of properties of the spatial social network arising from a large-scale online social networking service, Tuenti, with more than 10 millions active users in Spain. We have access to the social ties among Tuenti members, to their online interactions and to their home locations.

Our findings support the claim that social ties arising among Tuenti members are heavily constrained by space, with *individuals at closer distance more likely to establish social connections*. The social network appears divided into an inner, well-connected core and a periphery of less connected users: *while the social links in the core tend to span shorter geographic distances, outer ties are much longer*. Finally, we observe how *interaction levels are higher inside the core and connecting to the periphery, but much lower on bridging ties*, with the net effect that social interaction between friends appears independent of spatial distance.

These results highlight how the spatial properties of an online social network are influencing its structure, which then impacts important processes taking place on the network itself, such as information diffusion. In particular, the presence of spatially-limited and well-connected cores of individuals suggest how information might be trapped inside socially and spatially confined areas, hampering diffusion to the fringe of the network. This conclusion has significant consequences for systems and applications built on top of online social platforms.

[1] A recent study confirmed this finding on the Facebook social network `http://goo.gl/giDq5`.

Dataset

In this section we describe Tuenti, the online social service under analysis, and we present some basic properties of the dataset we study, introducing the notation we will use throughout our work.

Tuenti

Tuenti[2] is an invitation-only social networking service founded in 2006 in Spain. Thanks to its widespread popularity in this country, Tuenti is often referred to as the "Spanish Facebook". As other popular social networking platforms, it allows users to set up their profile, connect with friends and share links and media items. Tuenti users can interact with each other by writing messages on each other's walls.

The dataset under analysis in this work is a full anonymized snapshot of Tuenti friendship connections as of November 2010. It includes about 9.8 million registered users (25% of Spanish population), more than 580 million friendship links and 500 million message exchanges over a period of 3 months. Since Tuenti members must choose a location of residence from a list of Spanish cities to be able to join the service[3], we are able to assign a spatial home location to each user. Tuenti was originally popular in the city of Madrid, then further gained popularity in Seville, Valencia, Malaga and Gran Canaria, progressively gaining enough traction in the entire Spanish country: in fact, the service has become pervasively used in many cities.

Notation

The main goal of this work is to investigate the interplay between social network structure, spatial distance and tie strength. We note that Tuenti members can take part in a series of online interactions with each other, ranging from explicitly declaring a social connection (a friendship) to exchanging direct wall messages. Tuenti only allows users that are friends with each other to exchange wall messages.

We model the social network among Tuenti users as a directed weighted graph $G = (V, E)$: the set of nodes $V = \{u_1, u_2, \ldots, u_n\}$ is composed of n users and the set of edges E is composed of pairs of users that are present in each other's friend lists. We define Γ_i as the set of users connected to user u_i in graph G, so that $deg_i = |\Gamma_i|$ is the number of friends of u_i. We denote as $w_{i,j}$ the number of messages user u_i posted on the wall of user u_j. When user u_i has never left a message on user u_j's wall we set $w_{i,j} = 0$.

[2] `http://www.tuenti.com`

[3] This requirement has changed after our dataset was collected.

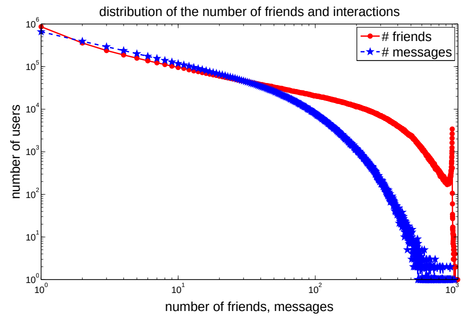


Figure 1: Complementary cumulative distribution of the number of friends and messages sent by Tuenti users.

For every edge $e_{i,j} = (u_i, u_j) \in E$ between user u_i and user u_j we introduce an associated *balanced interaction weight:*

$$\bar{w}_{i,j} = \min(w_{i,j}, w_{j,i}) + (1 - \delta_{w_{i,j},w_{j,i}})/2,$$

where $\delta_{x,y}$ is Kronecker delta (returns 1 if arguments are equal and 0 otherwise). We use the minimum of the interaction weights to emphasize reciprocated interactions, since non-reciprocated interactions may indicate spam. Therefore, for the non-reciprocated interactions we only add 1/2 no matter the difference in the numbers of messages exchanged. The distribution of the balanced interaction weights is depicted in Figure 2(b).

We define $d_{i,j}$ as the geographic distance between the cities of residence of user i and user j, computed as the great-circle distance over the planet: we define $d_{i,j} = 0$ if they report the same city of residence. In Table 1 we report the main properties of the dataset under analysis.

N	9,769,102	d_{eff}	5.8
K	587,415,363	d_{max}	9
N_{GC}/N	0.97	$\langle d_{path} \rangle$	5.2
$\langle deg \rangle$	126	$\langle D \rangle$	531.2
$\langle C \rangle$	0.200	$\langle l \rangle$	79.9

Table 1: Properties of the social network among Tuenti members: number of nodes N and edges K, size of the giant connected component GC, average node degree $\langle deg \rangle$, average clustering coefficient $\langle C \rangle$, 90-percentile effective network diameter d_{eff}, maximal distance d_{max} between two nodes in the network, average path-length between nodes $\langle d_{path} \rangle$, average geographic distance between nodes $\langle D \rangle$ [km], average link length $\langle l \rangle$ [km].

Basic properties

The social network arising among Tuenti members includes about 10 million members and more than 500 million edges: thus, the average number of friends a user has is about 126. However, the complete distribution of the number of friends is heterogeneous, as depicted in Figure 1. The vast majority of users have less than 10 friends, while about 1% of users have more than 200 friends: thus, the presence of popular users heavily affects the overall distribution of social ties among Tuenti members. The peak noticeable at 1,000 friends is due to a limit imposed by Tuenti on the number of connections made by each users. Though, apparently there are few users that manage to evade this limit.

There is a dominating giant connected component which contains about 97% of all the members, thus leaving aside about 300,000 users. The social network exhibits a relatively high average clustering coefficient of 0.2: that is, users form richly connected local clusters of nodes. At the same time, the network exhibits short path lengths between users: on average two nodes are divided by 5.2 social links and 90% of pairs of users are within 5.8 hops. These two properties, high local clustering and low average path length, confirm the small-world nature of the Tuenti social network (Watts and Strogatz 1998), as found in many other online social systems (Kumar, Novak, and Tomkins 2006; Mislove et al. 2007; Leskovec and Horvitz 2008).

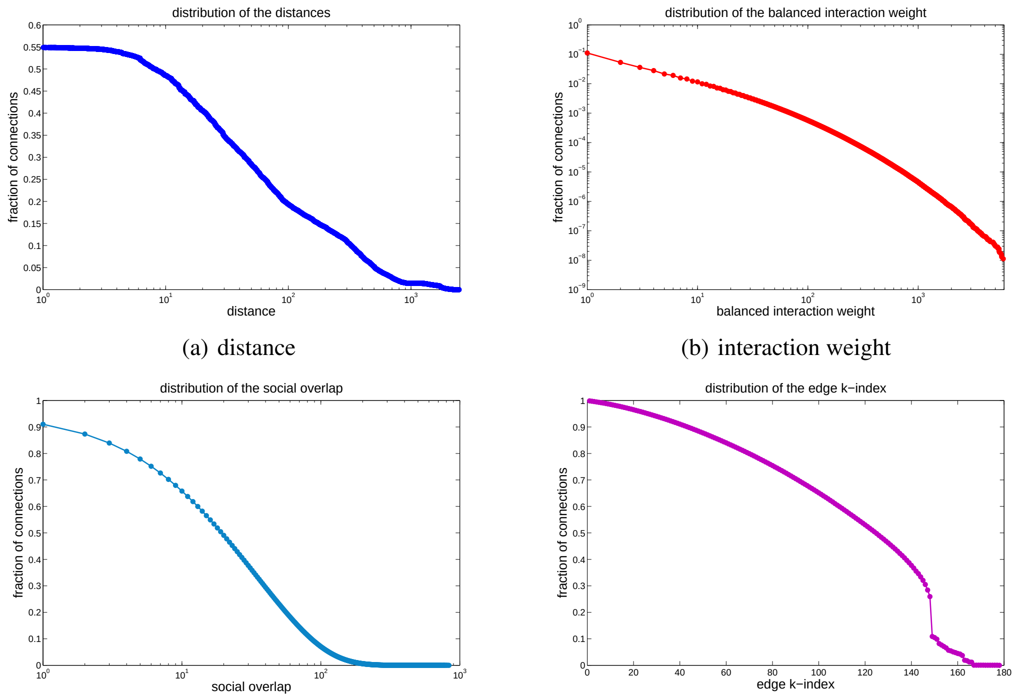
From a spatial point of view, we note that the average geographic distance between users $\langle D \rangle$ is about one order of magnitude larger than the average geographic distance between friends $\langle l \rangle$: this indicates that *spatially closer users are much more likely to engage in a social connection (e.g. become friends) than users separated by larger distances*. This is confirmed by the entire distribution of spatial distances between friends: in detail, there are about 50% of social links between users at a distance of 10 km or less (see Figure 2(a)).

Not all the social connections imply online interactions between Tuenti users: about 80% of social links exhibit no messages exchanged. The distribution of the number of messages sent for all the social connections with at least one exchanged message is depicted in Figure 1. The distribution shows a heavy-tail behavior, with the majority of connections having only few interactions and few connections having many thousands messages. Hence, there is huge heterogeneity across online social interactions: in the next sections our aim will be to understand how the strength of online social ties is related to their structural position and to their spatial properties.

Structural position of social ties

As we have discussed earlier, not all the edges in a social network are equal, as they carry different importance; at the same time, edges occupy different structural positions within the network itself. Our aim now is to introduce measures that help us to understand the structural position of a social tie. Overall, a given connection between two nodes in a network can be characterized from a local point of view, thus observing only the properties of the two endpoints and their relative neighborhood, and from a global perspective, assessing the position that a particular connection has with respect to the entire network.

For example, a single link connecting together two otherwise disconnected components of a network would be of lower importance from a local point of view, as the two endpoints have no other connection in common: on the contrary, it would exhibit an extremely important global role, as it keeps the entire network connected, allowing the flow of information. Thus, we introduce two metrics that capture these different structural properties of a social tie: the *social*



(a) distance

(b) interaction weight

(c) social overlap

(d) k-index

Figure 2: Complementary cumulative distribution function of spatial length, balanced interaction weight, social overlap and k-index for the social ties in the Tuenti social network.

overlap and the k-*index*.

Local position

The structure of social networks tends to reflect that individuals usually belong to social groups: that is, friends usually know each other, creating clusters of individuals that are mutually connected to one another. In particular, when two connected users have several friends in common that is an indication that their social link is situated inside a particularly well connected social community. Hence we define the *social overlap* of an edge $e_{i,j}$ as $o_{i,j} = |\Gamma_i \cap \Gamma_j|$.

The distribution of social overlap values for all the edges in the social network under analysis is depicted in Figure 2(c): the median value is about 20, while only 10% of edges have a social overlap higher than 100. We also note that about 5% of edges have a social overlap equal to zero (not depicted), thus acting as perfect local bridges.

Global position

Understanding the position of a single edge with respect to the entire social network can be a daunting and elusive proposition: standard approaches that focus on the importance of the edge in the information flow supported by the network can be computationally unfeasible on large-scale social networks. Thus, the approach we adopt in our work is to rely on a node property, the k-index, and then derive a measure of the global position of a social link by considering the properties of its endpoints. The k-index has been

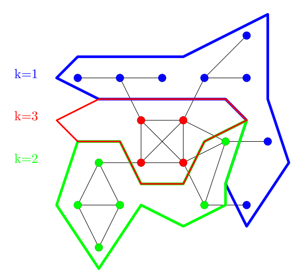


Figure 3: An example of the k-core decomposition.

found to be an indicator of influential nodes within a social network (Kitsak et al. 2010).

The k-core of a network is the maximal subgraph in which each node is connected to at least k other nodes of the subgraph (Seidman 1983). The k-index of a node is v if it belongs to the v-core but not to the $(v+1)$-core. In Figure 3 we present an example for the k-core decomposition of a small network.

The k-index of a node reveals whether it lies in a central core position with respect to the entire network, whether it lies on the periphery of the network itself, or whether it is located on a smaller core in between. The k-index thus reveals the global position of a node within the network.

In this work we are interested in social connections, therefore we define the k-*index* k_{ij} of an edge as the minimum of the k-indexes of its two endpoints. Using such definition of

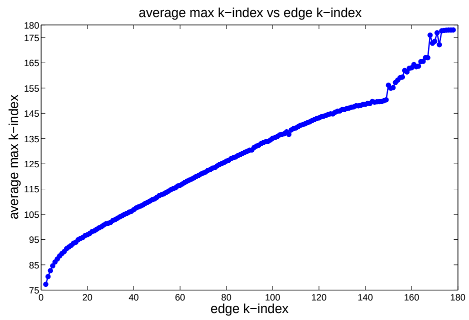


Figure 4: Average value of the maximum k-index of the two endpoints of an edge as a function of edge k-index.

the edge k-index we distinguish if the edge connects nodes inside a core or whether it links to a node in the periphery. The distribution of k-index values for all the edges in the social network is shown in Figure 2(d): we notice a large number of edges having k-index values close to 150, which denotes how edges tend to connect nodes inside the core of the network.

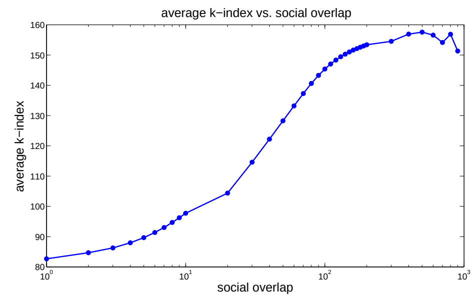
In Figure 4 we study how the k-index of an edge, which is the minimum k-index of the two endpoints, is related to the maximum k-index of such endpoints. Even edges with low k-index are connected to nodes with high k-index. This suggests that edges tend to be between nodes in the core or between nodes in the periphery and nodes in the core, while nodes in the periphery seldom connect to each other. Hence, most of the nodes in the social network can be assigned to a core section, where users have a large number of links between them, and to a periphery, with users with only few friendship connections going mainly to the core.

Structural patterns

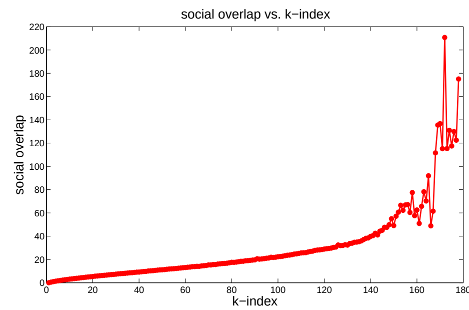
In the following analysis we will analyze the pairwise relationship between different measures of social links. In each case, as we increase one variable $\mathcal{X}$ we compute the average value of the other variable $\mathcal{Y}$, aggregating together all links that have the same value of the first one. We do not show standard deviation values as they are negligible. Hence, in order to fully grasp the relationship between the two variables, we present each comparison in both directions.

The interplay between the local properties of a social tie and its global position reveals what type of structure the entire network exhibits. The definition of social overlap and k-index allow network scenarios where links may have high k-index and low overlap, or the other way round. For instance, a network whose nodes are arranged and connected as the vertexes of a high-dimensional cube would have all edges with high k-index and zero overlap. However, social networks tend to have densely connected local communities. Thus, we would expect that edges within a large social community have both high social overlap and high k-index.

In fact, as seen in Figure 5(a), as the social overlap of an edge increases, its average k-index quickly grows as well. At the same time, as depicted in Figure 5(b), when the k-index increases, the average social overlap grows slowly, reaching



(a) Average k-index vs social overlap



(b) Average social overlap vs k-index

Figure 5: Relationship between social overlap and k-index of social links.

extremely high values only for k-index values larger than 150. These relationships confirm our findings on the structure of the network: there are inner cores where users are tightly connected to each other, while other parts of the network include more isolated users that tend to not belong to any community.

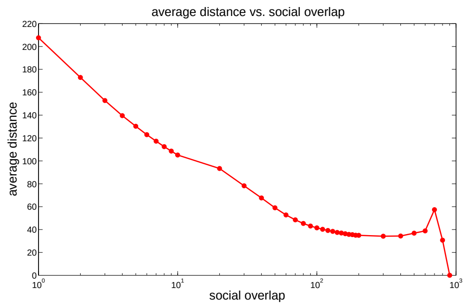
Tie strength and socio-spatial properties

The social network among Tuenti members exhibits a clear structure with well-connected inner cores and isolated nodes in the fringe. Our analysis turns now its focus on the relationship between spatial length of social links and their structural position, investigating how space influences, maybe in different ways, social ties in the core and in the periphery.

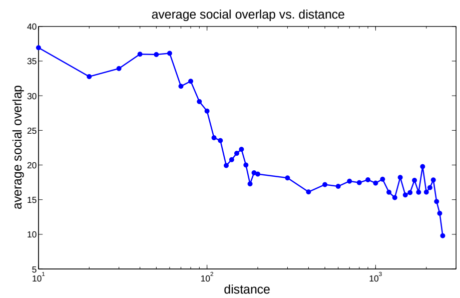
Structural position and spatial length

We first note that the geographic distance between two connected users decreases as they share more and more friends, as shown in Figure 6(a). At the same time, again in Figure 6(b), when looking at social links of increasing length, social connections which span less than 60-80 km exhibit higher values of social overlap, whereas the social overlap of longer links quickly tumbles down. This indicates how social links can be divided in *short-range* and *long-range*, with the separation distance being between 50 and 100 km, suggesting a division in intra-city and inter-city social bonds.

From a global point of view, the average spatial length of social links decreases as their k-index increases, as depicted in Figure 7(a): thus, social links inside the core tend to be shorter than the ones reaching the periphery of the social



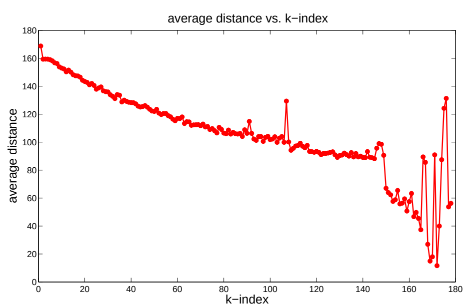
(a) Average social overlap vs link length (km)



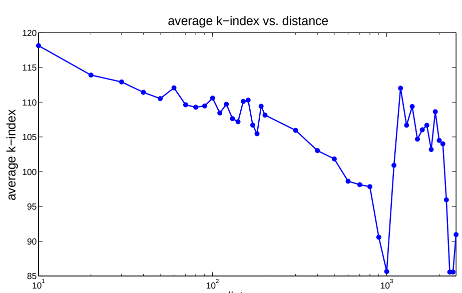
(b) Average link length (km) vs social overlap

Figure 6: Relationship between link length and social overlap.

network. Even when swapping the two variables, when the spatial length of social links increases their k-index drops accordingly, as seen in Figure 7(b).



(a) Average k-index vs link length (km)



(b) Average link length (km) vs k-index

Figure 7: Relationship between link length and k-index.

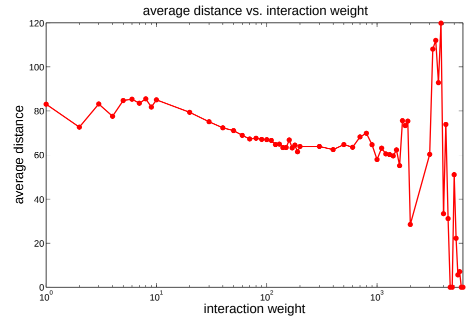


Figure 8: Average link length as a function of interaction weight.

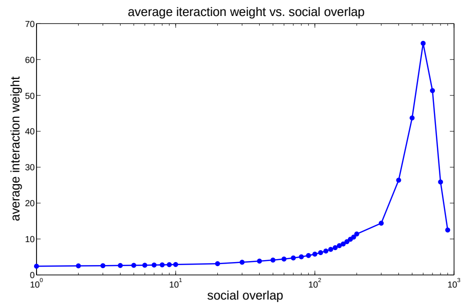
This analysis indicates how the division of the social network into a core and a periphery does not take place independently of spatial distance. Instead, social ties belonging to the core of the network, connecting together popular users with several friendship connections, tend to arise at shorter spatial distances than social ties established by less popular members in the periphery. Hence, isolated members tend to seek connections even at longer distances: this would signal that either they have no potential connection available at short distance, as they may be located in a scarcely populated area, or they are more willing to connect to individuals far away. Hence, there is evidence of *a bridging behavior of spatially longer social links, connecting together diverse portions of the network, while shorter links are tightly integrated inside social groups*.

The impact of tie strength

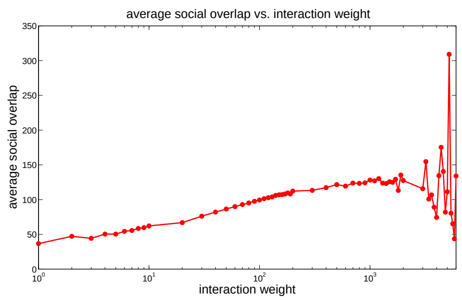
An important facet of social relationships that we capture in our Tuenti dataset is the strength of the interaction between two members: the balanced interaction weight we assign to each social link captures how likely is a social tie to be used to spread information. The importance of tie strength with respect to the structural and spatial properties lies on the fact that not all the links are equally likely to be used: to fully understand how the social network carries and spread information, we study whether tie strength is correlated with the structural position and the spatial length of a social link.

Surprisingly, our first observation is that the amount of interaction is fairly uncorrelated to spatial distance: in fact, as the interaction weight of a social link increases its spatial length remains fairly constant, as shown in Figure 8. Thus, even though the likelihood that two individuals are connected is heavily dependent on distance, when considering how much friends interact geographic distance is not a limiting factor.

When considering the impact of social overlap on interaction, we discover that the interaction weight remains fairly constant for social overlap values up until 100: after this threshold the amount of interaction exponentially grows, as seen in Figure 9(a). Recall that about 90% of social links have a social overlap smaller than 100: this suggests that the extremely high levels of interaction mainly take place between users with several shared friends, which are likely

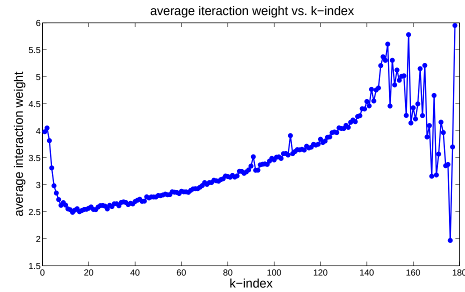


(a) Average social overlap vs interaction weight

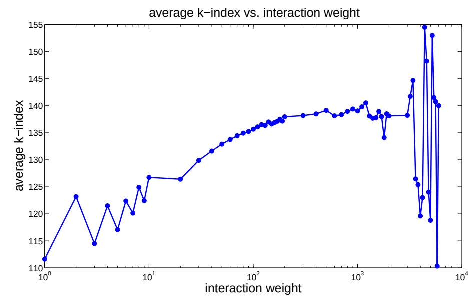


(b) Average interaction weight vs social overlap

Figure 9: Relationship between social overlap and interaction weight of social links.



(a) Average k-index vs interaction weight



(b) Average interaction weight vs k-index

Figure 10: Relationship between k-index and interaction weight of social links.

to be in the network core. In fact, as the interaction weight increases the social overlap grows only slowly, as seen in Figure 9(b).

In Figure 10 this outcome is further confirmed as we investigate the relationship between interaction weight and k-index of social links. Ties in the inner cores have the highest levels of interaction. However, we also note something strikingly more surprising: interaction weights are almost equally high for social ties with low k-index, while social links with intermediate k-index values reach a minimum of interaction weight. In agreement with these results, when we consider social ties of increasing weight we find that the k-index grows but suddenly reaches a plateau, as both links with low and high k-index have the highest levels of interaction. These relationships suggest that interaction levels tend to be higher when links are completely inside the network core, corresponding to high k-indexes, or connecting to the periphery, thus having low k-indexes. Social ties with intermediate k-index, likely to bridge together different portions of the network, experience the lowest interaction levels.

Discussion

The social network arising among Tuenti members has a well-defined structure: users have heterogeneous properties, with few popular nodes belonging to an inner core of tightly connected individuals and nodes with less connections located on the periphery of the network. Yet, nodes belonging to the core are still connected to the fringe of the network, since an overwhelming majority of users belong to the same connected component.

This strong structural division results in ties located inside the core and ties that stretch out to connect users in the periphery. The former ties are strongly embedded inside social communities, connecting together users that tend to share a large fraction of common friends. The latter connections act then as bridges, maintaining the entire network connected and allowing the flow of information. These properties are aligned to what has been often found in other online social services (Kumar, Novak, and Tomkins 2006; Mislove et al. 2007).

From a spatial perspective, the interplay of geographic distance and structural network properties is evident: social connections between users inside the core tend to have shorter geographic spans than connections stretching outside the core. Geographic closeness not only increases the likelihood of connections, but also increases the likelihood that users belong to the same, tightly connected group of individuals. Instead, social ties outside the core tend to be much longer than the other links: the length of these bridge ties is thus creating not only network shortcuts, but also spatial shortcuts. The role of these spatially long bridges is crucial to spread information over the network and, at the same time, over space.

Surprisingly, the amount of interactions appears independent of spatial distance: interaction levels appear higher inside well-connected cores and on links connecting to the fringe of the network. The effect of social overlap, instead, seems to be much weaker on online interaction, albeit it does offer a certain correlation. Overall, our findings suggest that online interaction remains largely untouched by the influence of distance, although spatial constraints still limit which social links are established.

In summary, even though online social networks connect individuals through short chains of social bonds, as epitomized by Milgram's experiment (Travers and Milgram 1969), the combined influence of geographic space and tie strength when people create these chains is extremely strong, as seen also in many offline and online networks (Dodds, Muhamad, and Watts 2003; Liben-Nowell et al. 2005). Our work provides useful and promising results to further unravel the close relationship between where users are located and whom they interact with.

Conclusions

In this paper we have studied the interplay between spatial distance, interaction strength and structural properties in social ties arising among members of an online social platform. We have analyzed a large-scale dataset collected from Tuenti. We found how the social network appears divided in a tightly connected core and a periphery with less connected users, with some edges acting as bridges to keep the entire network connected. These bridges span longer geographic distances, creating shortcuts both over the network and over space. To our surprise, the interaction levels tend to be more correlated with the structural properties of a social tie rather than to its spatial length.

Our findings shed for the first time light on how spatial constraints influence network structure: since the structural properties of a network are of crucial importance in influencing processes taking place on online social networks, this work demonstrates that space is equally fundamental.

Acknowledgements

Yana Volkovich acknowledges support from the Torres Quevedo Program from the Spanish Ministry of Science and Innovation, co-funded by the European Social Fund.

References

Ahn, Y. Y.; Han, S.; Kwak, H.; Moon, S.; and Jeong, H. 2007. Analysis of topological characteristics of huge online social networking services. In *Proceedings of WWW 2007*.

Backstrom, L.; Sun, E.; and Marlow, C. 2010. Find me if you can: improving geographical prediction with social and spatial proximity. In *Proceedings of WWW 2010*.

Burt, R. S. 1992. *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.

Cairncross, F. 2001. *The Death of Distance: How the Communications Revolution Is Changing our Lives*. Harvard Business School Press.

Dodds, P. S.; Muhamad, R.; and Watts, D. J. 2003. An experimental study of search in global social networks. *Science* 301(5634):827–829.

Festinger, L.; Schachter, S.; and Back, K. 1950. *Social Pressures in Informal Groups*, volume 57. Harper.

Gilbert, E., and Karahalios, K. 2009. Predicting tie strength with social media. In *Proceedings of CHI 2009*. New York, NY, USA: ACM.

Granovetter, M. S. 1973. The strength of weak ties. *The American Journal of Sociology* 78(6):1360–1380.

Kahanda, I., and Neville, J. 2009. Using transactional information to predict link strength in online social networks. In *Proceedings of ICWSM 2009*.

Kitsak, M.; Gallos, L. K.; Havlin, S.; Liljeros, F.; Muchnik, L.; Stanley, H. E.; and Makse, H. A. 2010. Identification of influential spreaders in complex networks. *Nature Physics* 6(11):888–893.

Kleinberg, J. M. 2000. Navigation in a small world. *Nature* 406(6798):845.

Kumar, R.; Novak, J.; and Tomkins, A. 2006. Structure and evolution of online social networks. In *Proceedings of KDD 2006*.

Leskovec, J., and Horvitz, E. 2008. Planetary-scale views on a large instant-messaging network. In *Proceedings of WWW 2008*.

Liben-Nowell, D.; Novak, J.; Kumar, R.; Raghavan, P.; and Tomkins, A. 2005. Geographic routing in social networks. *PNAS* 102(33):11623–11628.

Merton, R. 1948. The social psychology of housing. *Current Trends in Social Psychology* 163 – 217.

Mislove, A.; Marcon, M.; Gummadi, K. P.; Druschel, P.; and Bhattacharjee, B. 2007. Measurement and Analysis of Online Social Networks. In *Proceedings of IMC 2007*.

Newman, M.; Barabasi, A.-L.; and Watts, D. J. 2006. *The Structure and Dynamics of Networks*. Princeton Studies in Complexity. Princeton University Press, 1 edition.

Onnela, J.-P.; Saramäki, J.; Hyvönen, J.; Szabó, G.; Lazer, D.; Kaski, K.; Kertész, J.; and Barabási, A.-L. 2007. Structure and tie strengths in mobile communication networks. *PNAS* 104(18):7332–7336.

Onnela, J.-P.; Arbesman, S.; González, M. C.; Barabási, A.-L.; and Christakis, N. A. 2011. Geographic constraints on social network groups. *PLoS ONE* 6(4):e16939.

Rogers, E. 1995. *Diffusion of innovations*. New York,: Free Press.

Scellato, S.; Noulas, A.; Lambiotte, R.; and Mascolo, C. 2011. Socio-Spatial Properties of Online Location-Based Social Networks. In *Proceedings of ICWSM 2011*.

Seidman, S. B. 1983. Network structure and minimum degree. *Social Networks* 5(3):269 – 287.

Stewart, J. Q. 1941. An inverse distance variation for certain social influences. *Science* 93(2404):89–90.

Tobler, W. R. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* 46:234–240.

Travers, J., and Milgram, S. 1969. An experimental study of the small world problem. *Sociometry* 32(4):425–443.

Wasserman, S., and Faust, K. 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press.

Watts, D. J., and Strogatz, S. H. 1998. Collective dynamics of 'small-world' networks. *Nature* 393(6684):440–442.