Social Networks 54 (2018) 168–178

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Embedding time in positions: Temporal measures of centrality for social network analysis

Lucia Falzon[a,b,*], Eric Quintane[c], John Dunn[a], Garry Robins[b]

[a] Defence Science and Technology Group, Department of Defence, Australia
[b] School of Psychological Sciences, The University of Melbourne, Australia
[c] School of Management, The University of Los Andes, Colombia

ARTICLE INFO

Article history:

Keywords:
Time ordered social interactions
Temporal centrality measures
Network algebra

ABSTRACT

Digital data enable researchers to obtain fine-grained temporal information about social interactions. However, positional measures used in social network analysis (e.g., degree centrality, reachability, betweenness) are not well suited to these time-stamped interaction data because they ignore sequence and time of interactions. While new temporal measures have been developed, they consider time and sequence separately. Building on formal algebra, we propose three temporal equivalents to positional network measures that incorporate time and sequence. We demonstrate how these temporal equivalents can be applied to an empirical context and compare the results with their static counterparts. We show that, compared to their temporal counterparts, static measures applied to interaction networks obscure meaningful differences in the way in which individuals accumulate alters over time, conceal potential disconnections in the network by overestimating reachability, and bias the distribution of betweenness centrality, which can affect the identification of key individuals in the network.

Introduction

Social network ties can represent different types of relational data. Traditionally, network ties have represented social relations (e.g., friendship, advice) captured using sociometric questionnaires (Marsden, 1990). Recently, the availability of electronic trace data (e.g., email, phone communications, Twitter...) has made time-stamped social interactions more common as sources of network data (Lazer et al., 2009). Social interactions differ from social relations in one key dimension: the specificity of their temporal dimension (Borgatti et al., 2009; De Nooy, 2015). This means that there is a specific beginning, end, and duration of an interaction, while this is not necessarily the case for a social relation (Borgatti et al., 2013).

Researchers typically aggregate these social interactions over a period of time to create ties that resemble social relations (Kitts, 2014). When aggregating social interactions in order to constitute ties in a network, a key issue is the temporal granularity of the aggregation. If researchers want to conserve as much of the temporal information of the interaction data as possible, they have to use a fine-grained temporal window of aggregation (e.g., an hour or a day, depending on the context) (Moody et al., 2005). However, in most cases, ties composed of social interactions aggregated over short intervals of time will not equate to a relation, and a network of these ties does not represent social structure, which is typically defined as being stable over long time frames (Laumann and Pappi, 1976). Consequently, the established interpretations of measures of position of actors in social structure (e.g., about popularity for indegree centrality or control over information for betweenness) are unlikely to be meaningful when applied to networks composed of social interactions aggregated over short time frames. By contrast, aggregating social interactions over longer time frames (e.g., several months) makes them more similar to social relations, but it reduces the temporal information of social interactions. This is, at best a missed opportunity to take time into account, but it may be problematic because it ignores potential temporal dependences of interactions (Butts, 2008).

To solve this aggregation issue, a series of temporal network measures that embed time in positions has been proposed (see review by Holme and Saramäki, 2012). These measures typically distinguish between two aspects of time: the actual time of the interaction and the sequence in which social interactions unfold (Broccatelli et al., 2016; Kovanen et al., 2011). The time of the interaction has been used to provide information about actor activity over specific time periods (e.g., number of interactions on specific

* Corresponding author at: DST Group Edinburgh, PO Box 1500, Edinburgh SA 5111, Australia.
E-mail address: Lucia.falzon@dst.defence.gov.au (L. Falzon).

https://doi.org/10.1016/j.socnet.2018.02.002
0378-8733/

days of the week or hours of the day) (Batagelj and Praprotnik, 2016; Perra et al., 2012), but also about the pace of relation formation or dissolution (Moody et al., 2005); for example, the time needed for an interaction to be reciprocated. The second aspect of time considers the sequence of interactions that permits a path across nodes to exist, irrespective of the actual time (Kim and Anderson, 2012; Moody, 2002; Pan and Saramäki, 2011). For example, the process of transitive closure requires that ties form following a specific sequence - from A to B, then from B to C and then from A to C - in order to be interpreted as such.

These two temporal aspects of social interactions (time and sequence) are typically used separately in existing temporal network measures, which may be problematic because it is not clear when interactions are part of the same sequence. For example, an email from B to A sent 30 minutes after an email from A to B can reasonably be considered as part of an AB-BA sequence and reflect reciprocation in an information exchange context. However, if the same email from B to A has been sent one week after the email from A to B (or in a more extreme case, six months after) it is not clear that it should still be considered as part of the same sequence rather than part of a new sequence. Combining sequence and time together enables us to specify - based on the context - a maximum amount of time to consider two social interactions as part of the same sequence.

Consequently, in this paper we propose measures of individual positions in networks composed of aggregated social interactions that account for both the time and sequence of time-stamped social interaction data. We use a formal algebraic framework developed in Kontoleon et al. (2013) that allows us to explicitly consider the time and sequence of dyadic interactions, and to incorporate a decay variable, which we shall refer to as δ (delta), that relates to the maximum amount of time that can elapse between two interactions to consider them as related in a sequence. We propose three temporal equivalents to static network measures: temporal degree, temporal reachability and temporal betweenness. We demonstrate the empirical value of these temporal measures in a dataset of email interactions (70,584) among all the employees (103) of a medium size digital agency in Europe. We apply our temporal measures to the email interaction dataset and compare the properties of their distribution to their static equivalents. We show that the temporal measures differ from and complement the static measures by providing a finer grained understanding of the role of individuals in enabling and controlling information spread within the company.

Review of temporal measures

Holme and Saramäki (2012) argue that measures developed for static networks need revising when the temporal nature of network edges is explicitly analyzed. This is particularly the case when the order of interactions has an impact on the analysis question, for example measures that involve directed network paths in the analysis of contact sequences, such as reachability over time-respecting paths, i.e., ones in which each interaction in the path sequence is also ordered in time. We review measures presented in the extant literature that have been applied to time-stamped social interaction data. We include measures developed in the Computer Science and Physics literatures, yet we select only those measures that can be applied to understand actor roles in networks of human interactions. We identify two broad categories of network analysis research that give explicit consideration to the temporal nature of interactions. The first deals with the sequence of social interactions and the second is concerned with the time of social interactions. We deal with each category in turn.

Measures based on the sequence of social interactions

The sequence of social interactions as they play out in time through a network is crucial to understanding how information is transmitted and the roles that actors have in sending, receiving or propagating information, for two reasons. First, the sequence of interactions defines paths through which information, resources or diseases can be transferred. For example, Moody's (2002) study of disease transmission introduces an approach that takes into account the start and end of a sequence of social relations to track disease flow through a network. Second, the existence of a social interaction at a certain point in time can change the possibility of existence of subsequent interactions (Butts, 2008). For example, Gibson's (2005) analysis of turn taking in conversations highlights the notion of "sequential constraints" (p. 1563), which may prevent individuals from responding to a comment that has not yet occurred.

Taking into account the sequence of social interactions implies a definition of temporal path typically conceptualized using a notion of temporal geodesics (i.e., the shortest time-ordered path between two nodes). Moody (2002) observes that centrality measures appropriate for networks of time-sequenced interactions are based on the number and length of shortest time-ordered paths. Holme and Saramäki (2012) distinguish between two different types of temporal geodesics: the fewest number of interactions and the shortest path duration. They describe measures of temporal closeness centrality, based on the shortest path duration between each pair of nodes; and two versions of betweenness centrality, one based on the fewest number of interactions and the other on the shortest path duration. Praprotnik and Batagelj (2016) present an algebraic approach that makes it possible to track temporal paths through the network, and to calculate the path duration and earliest arrival times, which are then used to compute various forms of betweenness and other centrality measures. The authors demonstrate the measures on synthetic example networks and they consider an application to the analysis of journeys over transport networks.

Spiro et al. (2013) propose a dynamic measure of brokerage that takes into account the sequence in which connections are made within triads of individuals. They consider three concepts of brokerage: transfer, matchmaking and coordination. Each of these brokerage processes is specified differently, emphasizing the need to consider the sequence of events. The activities underlying the brokerage processes in question are collaborative and therefore undirected, but the sequence in which they occur induces a temporal direction enabling the identification of different brokerage roles.

Keating (2012) introduces two new centrality measures for time-ordered networks. Both are based on the notion of weakly connected temporal components (Nicosia et al., 2013), and the observation that since temporal components have a start time and end time, node importance can be conceptualized in terms of whether nodes were active "in either kick-starting or concluding processes" (Keating, 2012). Entrance centrality is thus defined as the fraction of time a given node participates at the beginning of a weakly connected temporal component; exit centrality is the analogous measure taken at the end.

Other authors (Grindrod and Higham, 2013; Kim and Anderson, 2012; Nicosia et al., 2012, 2013; Tang et al., 2010) consider a time-ordered series of network snapshots at various time-points that span the whole observation interval. From these they are able to construct temporal paths and walks that are time respecting, that is, those that do not violate the time ordering of connections. Although interactions that occur between time-points are treated as having taken place concurrently, that is, sequence is ignored at this finer level. Bajardi et al. (2011) provide an empirical illustration through