From calls to communities: a model for time-varying social networks

Guillaume Laurent1,2, Jari Saramäki2, and Márton Karsai1,a

1 Laboratoire de l'Informatique du Parallélisme, INRIA-UMR 5668, IXXI, ENS de Lyon, 69364 Lyon, France
2 Department of Computer Science, Aalto University School of Science, P.O. Box 15400, 00076 Aalto, Finland

Received 17 June 2015 / Received in final form 11 September 2015
Published online 18 November 2015 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2015

Abstract. Social interactions vary in time and appear to be driven by intrinsic mechanisms that shape the emergent structure of social networks. Large-scale empirical observations of social interaction structure have become possible only recently, and modelling their dynamics is an actual challenge. Here we propose a temporal network model which builds on the framework of activity-driven time-varying networks with memory. The model integrates key mechanisms that drive the formation of social ties – social reinforcement, focal closure and cyclic closure, which have been shown to give rise to community structure and small-world connectedness in social networks. We compare the proposed model with a real-world time-varying network of mobile phone communication, and show that they share several characteristics from heterogeneous degrees and weights to rich community structure. Further, the strong and weak ties that emerge from the model follow similar weight-topology correlations as real-world social networks, including the role of weak ties.

1 Introduction

The emergent structure and dynamics of social interactions are consequences of individual and collective decision-making processes [1]. Understanding the driving forces behind social network formation has been a great challenge that has attracted lots of attention during the last decades. Even though conventional experimental approaches such as surveys have helped to understand fundamental social mechanisms [2], it has not been possible scale such approaches up to sizes required for observing features of large-scale social structures, let alone their detailed dynamics. However, technological advances have helped to overcome these limitations through detailed digital recording of social interactions [3,4]. Following these advancements, a lot of effort has been put on mapping social structures from digital traces, and on establishing methodology for the analysis of social networks connecting millions of individuals. These efforts have led to discovering the small-world architecture [5], the heterogeneous network structures induced by preferential attachment mechanisms [6,7], the different roles of strong and weak ties [8,9] and their relationship to triangle formation [10] and the emergence of communities [11,12].

The earliest data-driven studies of social networks were mainly focused on static network structure. However, it is evident that this approach both unnecessarily discards available information – electronic interaction traces typically come with time stamps – as well as hinders detailed understanding of the mechanisms shaping network structure. In the recent years, steps have been taken to come around these limitations, with increased focus on both network dynamics (changes in link structure, such as in Refs. [13] or [14]) and temporal network structure [15] (time-stamped interaction events between individuals giving rise to e.g. temporal motifs [16,17]). In particular, studies of temporal networks of human interaction have revealed tie dynamics with memory and reinforcement processes [18,19], bursty dynamical patterns [20,21], temporal homophily in communication motifs [22] as well as different strategies of communication and tie maintenance [23].

Modelling temporal networks with processes unfolding on the time varying structure is still one of the key challenges in the field. Empirical studies employing null models with shuffled interaction sequences [24–26] have pointed out that bursty interactions together with features such as weight-topology correlations strongly influence the speed of any type of information diffusing on the temporal network [24,26–29]. However, a limiting factor of this approach is that it only allows one to remove existing temporal inhomogeneities and correlations from empirical sequences in order to understand their effects. Because of this, the null model approach does not provide ways for continuously adjusting the level of inhomogeneities or temporal correlations. This limitation can be overcome by resorting to artificial models of temporal networks [30,31] that allow fine-tuning the level of different features one by
one. One promising direction is the activity-driven framework [31] that assumes heterogeneous activity of agents. It provides a general foundation which can be extended by incorporating additional temporal and structural correlations. Studies of the activity-driven time-varying network model have demonstrated that memory effects can explain the emergence of strong and weak ties in social networks [19], while endo- and exogeneous tie formation processes may control the evolution of business alliance networks [32]. The same model has been used to study how time-varying interactions influence the dynamics of random walkers [33,34], rumour spreading [19], and epidemic processes [35].

In this paper, our aim is to consider further social mechanisms that would make models of time-varying social networks more realistic. We introduce a temporal network model with adjustable community structure and emergent weight-topological correlations by extending the activity-driven time-varying network model with three additional mechanisms. These are (i) reinforcement processes to model memory-driven interaction dynamics of individuals [19]; (ii) focal and cyclic closure motivated by the work of Kumpula et al. [12,36] to capture mechanisms responsible for the emerging community structure; and (iii) a node removal process. Using this temporal network model we demonstrate the effect of the scalable community structure and social reinforcement on information spreading through time-varying interactions. After the introduction of the model, we discuss its temporal behaviour, and match the emerging network properties with empirical observations. We discuss the present higher-order correlations and their effects on spreading processes, and finally we conclude our work and discuss possible future directions.

2 Model

Our model definition builds on the activity-driven time-varying network (ADN) model introduced by Perra et al. [31], integrating memory and reinforcement processes as well as closure mechanisms with the original model definition. Our ultimate goal is to provide a temporal network model with adjustable community structure and weight-topology correlations, in order to understand their role in shaping the emergent network structure and co-evolving dynamical processes.

2.1 Social mechanisms

The first mechanism to be integrated with the activity-driven model is memory in terms of frequently repeated interactions between a node and its peers who have been contacted before. This has been considered earlier by Karsai et al. [19] with reinforcement processes, where a node remembers its connected neighbours, and depending on its degree \( n \) at time \( t \), either creates a new link with probability \( p(n) = c/(n + c) \), or interacts with a neighbour with probability \( 1 - p(n) = n/(n + c) \) to reinforce an existing tie. The introduction of memory defines a non-Markovian dynamics where past interactions influence the present decision of a node. The larger a node’s egocentric network size \( n \), the smaller the chance for it to create a new link in the current iteration. This captures the realistic assumption that interactions in a social network are not random but mainly focussed at close friends whose number is strongly limited [14,37]. This decision-making mechanism naturally leads to the emergence of weak and strong ties, weight heterogeneities, and a more realistic degree distribution in the aggregated network structure [19]. Note that in this formulation \( c \) scales the strength of memory, which may capture the attitude of being a social keeper or social explorer [23]. Here, for simplicity, we set \( c = 1 \) for each node and leave the exploration of the effect of varying memory strengths for future studies [38].

The second mechanism we consider is that of tie creation with closure processes. Here, cyclic closure, the tendency to form network triangles, shapes the social structure at mesoscopic scale, and is responsible for the emergence of communities [39]. The mechanism of focal closure, on the other hand, is independent of network structure and represents the formation of ties between individuals with shared attributes or interests. It is driven by the propensity to seek cognitive balance between connected egos [8,40] as suggested by earlier theories in sociology [10,41]. An applicative definition of cyclic and focal closure in general is given by Kumpula et al. [12,36], who model cyclic closure as biased local search. To the contrary, focal closure is modelled as an unbiased global random search.

Finally, the third ingredient of the model is the process of node removal. It allows the model network to reach an equilibrium state where its overall structural characteristics become invariant of time.

2.2 Model definition

Our model takes \( N \) initially disconnected nodes that are each assigned a priori with an activity probability per unit time \( a_i = \eta x_i \). Here, \( x_i \) is the activity potential of node \( i \), drawn from a suitable distribution \( F(x_i) \) \( x_i \in [0,1] \) bounded by a minimum activity \( \epsilon \), and \( \eta \) is a time rescaling factor determining the average number of active nodes per unit time to be \( \eta \langle x \rangle N \). The activity potentials of humans are commonly observed to be heterogeneous [42–44]. We approximate this with a power-law activity potential distribution \( F(x_i) \sim x_i^{-\gamma} \) with exponent \( \gamma = 2.8 \), based on empirical observations and earlier modelling work [19]. In addition, without loss of generality, we fix the parameters \( \eta = 1 \), \( \epsilon = 10^{-3} \) and \( \Delta t = 1 \) [19].

At each iteration step, nodes are updated in random order. At one iteration step a node may become active with probability \( a_i = \eta x_i \) or be deleted with probability \( p_d \). If the node \( i \) becomes active, then, depending on its current degree \( n_i \), it can attempt to form a new link with probability \( p(n_i) = 1/(n_i + 1) \), or otherwise interact via an existing link with probability \( 1 - p(n_i) \). In the latter case it randomly selects one of its neighbours \( j \) with
probability \( p_{ij}^{W} = w_{ij}^{t}/\sum_{k \in V^{t}} w_{ik}^{t} \) weighted by the number of their previous interactions. The two nodes then interact and increase their link weight \( w_{ij}^{t} \) by \( \delta \) (reinforcement process). On the other hand, if the node decides to form a new link, it may follow different scenarios. In all cases, the new tie will initially have unit weight \( w = 1 \). If the degree of the focal node is 0, it randomly picks another node from the entire network \( j \) (focal closure) and forms a tie. Otherwise, it attempts to create a new link with the triadic closure mechanism. First, it chooses one of its neighbours \( j \) randomly with a weighted probability \( p_{ij}^{W} \). If \( j \) has no other neighbours than \( i \), node \( i \) looks for another random node to interact with (focal closure) and forms a link. Otherwise, it looks for a random neighbour \( k \) of \( j \) (\( i \neq k \)) with a weighted probability \( p_{jk}^{W} \). If \( k \) is not an already existing neighbour of \( i \) (\( k \notin V^{t}_{i} \)), the two nodes interact with probability \( p_{\Delta} \) and close the triad by forming a link. Otherwise, with probability \( 1 - p_{\Delta} \), node \( i \) follows the focal closure mechanism and instead forms a link with a randomly selected node (other than \( j \) and \( k \)). Finally, if \( k \) is already a neighbour of \( i \), that is, \( k \in V^{t}_{i} \), the two nodes interact and increase the weight of their existing link by \( \delta \) (reinforcement process). Note that during a single step, a node can participate in temporal interactions by creating a link (if it is active), or by receiving link(s) from other active nodes. At the end of each iteration step, all nodes finish their active interactions but remember their weighted egocentric network, i.e., their already connected neighbours \( j \in V^{t}_{i} \) and the weight \( w_{ij}^{t} \) of interactions with each of them. If a node is deleted all its connections are removed and a new node is added to the network in the next iteration step in order to keep the network size constant. For a pseudocode version of the algorithm, see Appendix.

3 Model network analysis

In addition to the activity-driven model parameters whose values are fixed, our model has three intensive parameters, \( p_{\Delta}, p_{\Delta}, \) and \( \delta \). By varying these parameters, one can simulate a rich variety of time-varying networks with several emergent structural properties and correlations. In the following, we explore how the properties of the emerging network structure depend on time and on the intensive parameters, and whether those properties match with those of an empirical temporal network of mobile phone communications. The dataset we use here contains 633,986,311 time-stamped mobile call communication records of 6,243,322 customers of a mobile operator (market share \( \approx 20\% \)) in a European country. Customers are represented as network nodes, connected via 16,783,865 weighted mutual links with weights defined as the number of calls between customers who mutually called each other at least once during the observation period. The presented network properties of the mobile phone call (MPC) network were calculated from a static, aggregated representation of the social network structure obtained by integrating the temporal interactions over 6 months.

In the following, model networks were generated via large-scale numerical simulations with \( N = 10\,000 \) nodes, and results were averaged over 100 independent realisations (if not noted otherwise). For each realisation, we measured the network parameters by considering links that are actually present in the network, i.e. we disregarded links of removed nodes.

3.1 Temporal features

The introduced network model is inherently temporal and simulates time-varying interactions between individuals. To explore its overall temporal behaviour, we measured two general network properties as a function of time. The first property is the average degree defined as \( \langle k \rangle (t) = N^{-1} \sum_{i} k_{i}(t) \), where the sum runs over all nodes \( i \), and \( k_{i}(t) \) denotes the current degree of node \( i \), i.e. the total number of connections that node \( i \) has established since the time it was added to the network up to time \( t \) (more formally \( k_{i}(t) = |V^{t}_{i}| \)). Second, we measure the average local clustering coefficient \( C(t) \), defined as the fraction of the real and possible numbers of triangles around a node given its number of current links, averaged over the whole network [45]. \( C \) quantifies the density of triangles in a network; in social networks, the existence of communities typically gives rise to high triangle densities. Its definition takes into account links that were established by present nodes since the time of their introduction to the network. Links of deleted nodes are removed from the structure, and so they are not considered in any of these definitions.

As follows from the model definition, the process starts from a set of disconnected agents, and hence all measured properties are trivially zero at time \( t = 0 \). However, as time goes by and ties are formed via temporal interactions, \( \langle k \rangle (t) \) starts increasing until the network reaches a stationary state with constant average degree. The time it takes to reach this equilibrium state strongly depends on the choice of the node deletion probability \( p_{\Delta} \) as shown in Figure 1a. If \( p_{\Delta} \) is too small, nodes remain in the system for a long time and even nodes with small activity levels have time to evolve their egocentric network resulting in a slow relaxation to the stationary state. On the other hand, as \( p_{\Delta} \) is increased due to the finite life-time of nodes, less active nodes are removed before their egocentric structures are fully evolved. Because of this, the network reaches equilibrium faster. This explains the decrease of stationary average degree with increasing \( p_{\Delta} \). The average degree can be tuned to approximate the empirical stationary average degree of the MPC network (dashed line in Fig. 1a), measured in the aggregated empirical network in the end of the observation period. Note that the model and real temporal networks reach their stationary states in rather different ways. While the empirical data records activities in a network which is already in its stationary state [46], the model’s evolution starts from an empty network and passes through a transitory period before the stationary state is reached. Consequently, only the stationary properties of the model and the empirical
network can be meaningfully compared, but not their evolution during the initial observation period.

Measurement of the time dependence of the average clustering coefficient (see Fig. 1b) yields a fairly similar scenario. However, here, after the initial increase of $C(t)$ it reaches a maximum, followed by a decrease and relaxation to a stationary value. This is because triangles tend to emerge right in the beginning of the process, followed by the evolution of strong ties. After nodes have created their first links and closed triangles between them, interactions begin to frequently take place on existing links that are reinforced in the process. Thus the local search of active nodes is biased towards strong ties that are part of already formed triangles. Later, nodes attempt less frequently to create new triangles through the weak links that emerge throughout the dynamics by local closure. Again, there is saturation due to node removal influenced by the rarely active but surviving nodes, who keep introducing random links in the network. By choosing appropriate parameter values, the clustering coefficient can again be tuned to be comparable to its corresponding stationary empirical value (dashed line in Fig. 1b).

The above picture is fully supported by measurements capturing the fraction of events which create new links $n_{cr}(t) = \# e_{cr}/(\# e_{cr}(t) + \# e_{rf}(t))$, or reinforce already existing ones $n_{rf}(t) = \# e_{rf}/(\# e_{cr}(t) + \# e_{rf}(t))$. Here $\# e_{cr}$ ($\# e_{rf}$) denotes the number of such events at time $t$. As seen in Figure 1c these measures take values of 1 and 0 (respectively) at time $t = 0$, as all events initially create new links. However they rapidly approach a constant value as it is visible in Figure 1d. On the other hand, as $p_d$ is increased, the number of events reinforcing already existing links is increased, while events responsible for the creation of new links become less frequent. This is in accordance with our interpretation above. The measures can be tuned to reflect empirical stationary network values here as well (black lines).

In the following, if not noted otherwise, we set $p_d = 0.5 \times 10^{-5}$ and run the simulations for $t = 150000$ time steps to ensure the model networks to always reach their stationary state with average properties invariant in time.

### 3.2 Static features

Next we concentrate on the effect of different mechanisms on the emerging network structure in the stationary state. In Figure 2, we visually demonstrate the role of the cyclic closure (a–c) and link weight reinforcement (d–f) mechanisms in the emerging model networks. In Figures 2a–2c, we have kept the reinforcement increment constant, $\delta = 1$, while varying the cyclic closure probability $p_\Delta$ between 0.5 and 0.995. The node deletion probability $p_d$ has been chosen to yield networks with suitable link density for visualisation (for exact parameters see the figure caption). When $p_\Delta$ is small, the emerging network structure is densely connected and appears more like a random structure since link creation is driven by focal closure. Nevertheless, because of the reinforcement process, weight heterogeneities emerge already in this case. More interestingly, by increasing $p_\Delta$, communities are seen to emerge, with weight-topology correlations that are in line with the Granovetterian picture [8,9], where strong ties connect nodes inside communities, while weak ties emerge between them (darker link colour = stronger link weight).

However, focal closure alone is not sufficient for the emergence of community structure as shown in Figures 2d–2f. Here, the triangle formation probability has been kept constant, $p_\Delta = 0.995$, while the reinforcement parameter $\delta$ has been varied between 0 and 1.5. Even though a triangle would almost always be closed by the local search if a suitable node were found, without link weight reinforcement (Fig. 2d), the local search will be dispersed and often result in focal closure because of hitting a neighbour with no other neighbours. However, by increasing $\delta$ the local search will get more biased towards emerging strong ties and lead to the evolution of tight community structure. Note that since $\delta$ controls the strength of the local search mechanism, it also scales the size of the communities.

To quantitatively explore the emerging network structure as the function of the intensive parameters, we have measured the behaviour of general network properties. Figures 3a and 3b show how the average degree of the model networks depends on $p_\Delta$ and $\delta$ (Figs. 3a and 3b, respectively) for different deletion probabilities $p_d$. The average degree of the empirical MPC network is displayed for comparison (dashed horizontal line). In each case, the average degree decreases when the parameters controlling the probability of cyclic closure $p_\Delta$ and the amount of link reinforcement $\delta$ are increased, indicating that these parameters together with $p_d$ strongly control $\langle k \rangle$. However, while the dependence of $\langle k \rangle$ on $p_\Delta$ is concave, $\langle k \rangle(\delta)$ is a convex function for any $p_d$. Moving beyond averages,
Fig. 2. Demonstration of the emerging structure in the time-varying network model. Panels (a)–(c) depict simulated networks with fixed $\delta = 1$ and varying $p_{\Delta} = 0.5, 0.9, \text{ and } 0.995$ ($p_d = 4 \times 10^{-5}, 2 \times 10^{-5}, \text{ and } 1.04 \times 10^{-5}$) respectively. Panels (d)–(f) shows networks with fixed $p_{\Delta} = 0.995$ with varying $\delta = 0, 0.5, \text{ and } 1.5$ ($p_d = 3.5 \times 10^{-5}, 1.7 \times 10^{-5}, \text{ and } 8 \times 10^{-6}$) respectively. Each panel depicts the actual structure of a network with $N = 500$, in its stationary state. Links are coloured according to their weight (darker link colour = stronger link weight).

Figures 3c and 3d display the behaviour of the degree distributions $P(k)$ for various $p_{\Delta}$ and $\delta$, indicating that there are strong degree heterogeneities. While the average degrees strongly depend on these parameters, the overall distribution shapes do not change much. The distribution of the empirical MPC is again displayed for reference, showing a very similar shape than those produced by the model.

The strength of social ties, measured here as the number of dyadic interactions, is commonly observed to be distributed heterogeneously in social networks. This is also the case with our empirical MPC network (black circles in Figs. 3g and 3h). Tie strength heterogeneity is an emergent property of our model networks where it comes from the preference of nodes to reinforce existing links. This mechanism is independent of the search strategies, as evident in Figure 3g, where the model weight distributions are invariant to the selection of the cyclic closure probability $p_{\Delta}$ ($\delta$ and $p_d$ are kept constant in these cases), and also in Figure 3e, where the average of these distributions increases only moderately with increasing $p_{\Delta}$. The emerging weight heterogeneities are the consequence of the intrinsic memory process, which inclines nodes to reinforce existing links rather than create new ones. Therefore, the weight heterogeneity naturally depends on the reinforcement parameter $\delta$ (see Fig. 3h). The reinforcement parameter in turn controls the average social tie strength (see Fig. 3f) by scaling the tail of $P(w)$ (Fig. 3h) whose functional shape is not affected; this has the appearance of a power law with a constant exponent.

Note that since the considered mechanisms are not fully independent, they influence multiple properties, leading to non-unique parameter sets corresponding to networks with similar features.

3.3 Higher-order correlations

Other than the observed degree and weight heterogeneities in the model networks, certain parameter ranges result in the emergence of rich community structure. In real social networks, communities can be characterised by higher-order correlations as they are built of closed triangles with unevenly distributed links of various strength in the structure. Stronger ties that are maintained by frequent interactions tend to connect nodes inside communities and shape the structure locally, while weak links with infrequent interactions are situated between communities and keep the social structure globally connected. To see if the modelled temporal networks display these characteristics, we have performed three sets of measures and compared our findings with empirical results.
Fig. 3. General network measures as a function of the intensive parameters. (a) and (b) ((c) and (f)) show the dependence of the average degree \( \langle k \rangle \) (average link weight \( \langle w \rangle \)) on \( p_\Delta \) and \( \delta \), respectively (for the former, \( \delta = 1 \) and the latter, \( p_\Delta = 1 \)). Various values of \( p_\Delta \) have been used (see legend in (b) (and (f))). The dashed black line depicts the average degree \( \langle k \rangle \) of the empirical MPC network. (c) and (d) depict the degree distributions \( P(k) \) of model networks with varying \( p_\Delta \) and \( \delta \); while (e) and (f) show the corresponding weight distributions \( P(w) \). In panels (c) and (e) ((d) and (f)), we kept \( \delta = 1 \) \( (p_\Delta = 1) \) and \( p_\Delta = 5 \times 10^{-5} \). Black circles denote the corresponding MPC distributions.

As we have seen earlier, both the values of \( p_\Delta \) and \( \delta \) influence the community structure. However, we expect that \( p_\Delta \) plays a dominant role here as it controls the triangular closing mechanism in the network. If \( p_\Delta = 0 \), links are created randomly and the clustering coefficient is very small as seen in Figure 4a, where a constant \( \delta = 1 \) were set for each measurement. By increasing \( p_\Delta \), cyclic closure becomes more dominant, reflected in an increasing \( C \) with very high values as \( p_\Delta \) goes to its maximum. The clustering coefficient of the MPC network is depicted as the horizontal dashed line. There are several sets of parameter values for which the model yields networks with similar clustering coefficient values, e.g. \( p_\Delta \approx 0.5 \), \( \delta \approx 1 \), and \( p_\Delta \approx 5 \times 10^{-5} \) (see the crossing points of the black dashed and green solid lines in Fig. 4a). Not surprisingly, \( \delta \) plays a weaker role in the emergence of triangles. The reinforcement mechanism introduces a bias in the local search for new ties, which leads to tighter communities and more closed triangles as we increase \( \delta \), reflected by the slightly increasing \( C \) in Figure 4b.

Second, we have measured weight-topology correlations to check if the model networks are constructed
according to the Granovetterian weak-tie structure. In his seminal paper [8] Granovetter suggests that the fraction of common friends of connected individuals is positively correlated with their tie strength (i.e. their link weight here). The fraction of common friends can be measured by the link overlap [9] defined as $O_{ij} = n_{ij}/((k_i - 1) + (k_j - 1) - n_{ij})$ where $n_{ij}$ is the number of common neighbours of nodes $i$ and $j$, and $k_i$ and $k_j$ are their degrees. To quantify weight-topology correlations, we measure the average link overlap ($O$) as a function of cumulative tie strength $P_{cum}(w)$ that measures the fraction of links with tie strength smaller than $w$ [9]. In the MPC network, this function reflects positive correlations between overlap and tie strength (black circles in Figs. 4e and 4f) in accordance with earlier observations [9]. If the model networks have their triangle-formation mechanism or weight reinforcement turned off ($p_{\Delta} = 0$ or $\delta = 0$), no such correlation is found. However, for any positive values of $p_{\Delta}$ and $\delta$, a positive weight-topology correlation emerges in accordance with the empirical observation and the hypothesis of Granovetter. For larger $p_{\Delta}$ (with constant $\delta = 1$), the function is shifted as more triangles evolve, which is increasing the average overlap (see Fig. 4c), while by increasing $\delta$ (and constant $p_{\Delta} = 0.5$) (see Fig. 4d) the correlation becomes stronger as ties with larger strength appear in the networks.

Finally, we have checked how the clustering coefficient $C$ changes by removing a fraction of links $f_0$ in increasing order of overlap. The resulting function $C(f_0)$ can be divided into two regimes. Trivially the removal of links with $O = 0$, which connect communities, does not decrease the number of triangles in the network but only decreases the degrees of nodes by removing links not participating in triangles. Consequently, by the removal of these links the clustering coefficient must increase. On the other hand, the removal of links with $O > 0$ that connect nodes inside communities reduces the number of triangles, resulting a decrease in $C$. This behaviour is observed in the empirical system (black circles in Figs. 4e and 4f) and also for the model networks. In Figure 4e as $p_{\Delta} \rightarrow 0$ no triangles evolve in the network, and thus most of the links have $O = 0$, and the first regime is extended. In the other extreme, as $p_{\Delta} \rightarrow 1$, most of the links have non-zero overlap and the second regime dominates. The best match with the empirical curve appears when $p_{\Delta} \approx 0.5$ (red curve in Fig. 4e) and $\sim 40\%$ of links evolve with $O = 0$. Note that the second regime of this function is very sensitive to the fine-grained structure of the actual communities, which causes the difference between the model and empirical curves. In addition, by keeping $p_{\Delta} = 0.5$, we have checked the $\delta$ dependence of this function (see Fig. 4f). As the tie reinforcement $\delta$ increases, links form tighter communities and the clustering coefficient increases in this case.

**4 Effects on spreading processes**

We conclude our modelling study by demonstrating its capacity in simulating dynamical processes evolving on the temporal network structure with scalable integrated social mechanisms. For this reason, we have selected the simplest possible spreading process, the susceptible-infected (SI) model [47], where each node can be in one of two mutually exclusive states: susceptible (S) or infected (I). Initially, every node is in the state $S$, except for a random seed with state $I$. The infection passes from any node in the $I$ state to any node in the $S$ state via a temporal interaction, but independent of the direction of the actual temporal link [24]. In order to assure that the network is in the stationary state we initiated the spreading process after the temporal network of size $N = 10 000$ has evolved for $t = 50 000$ iterations. After this time we measured the $i(t) = I(t)/N$ fraction of infected nodes until $t = 150 000$, where the process has reached maximal penetration (see Figs. 5a and 5b). Note that maximal penetration is not converging to 1 as there are always new susceptible nodes introduced in the network, moreover the temporal structure is not necessarily connected. For simplicity, we set the probability of transmission per interaction event to $\lambda = 1$ (note that then the speed of transmission between two nodes is only limited by the frequency of their interactions). Further, we consider interactions as undirected and therefore they may transmit the interaction both ways.

The simulation results depicted in Figures 5a and 5b show that the spreading process slows down when either the cyclic closure probability $p_{\Delta}$ or the weight reinforcement parameter $\delta$ is increased. While in the first case...
this is mainly due to the emerging tight communities, in the second case weight-topology correlations also play a role. The same conclusions can be drawn from measuring the speed of spreading, quantified as the time $\tau_{50\%}$ when the process has reached 50% penetration. Tight communities evolve as $p_{\Delta} \to 1$, which constrains the spreading process and explains the convex shape of the curve in Figure 5c. On the other hand, weight-topology correlations appear even for small values of $\delta$ (as we have shown in Fig. 4c). These immediately slow down the speed of spreading and cause the concave shape of the function in Figure 5d. Thus, one can conclude that even though more pronounced community structure has considerable effects on the spreading process, already weak weight-topology correlations can strongly affect the speed of spreading.

Please note that for empirical temporal networks, such as the mobile call network used in this study, it is known that spreading processes are dramatically affected by the presence of temporal inhomogeneities such as burstiness that have not been incorporated to the model. Therefore, unlike for the network properties, we do not compare SI model results with the MPC.

5 Conclusions

The temporal network representation takes into account the time-varying nature of interactions between entities instead of considering only the static network structure of their connections [15]. So far, this approach has mainly focussed on analyzing empirical data, and there is still a lack of appropriate network models. In this paper we introduce a model of temporal social networks with emerging heterogeneous structure, communities, and higher-order correlations. As a starting point, we take the recently introduced modelling framework of activity-driven networks with heterogeneously distributed activities of individuals [31] and memory processes [19]. We extend this model with three mechanisms that have been earlier used for generating static weighted networks [12], such as social reinforcement, cyclic and focal closure. These mechanisms arguably drive the tie formation of individuals and are responsible for the emergence of communities and the global connectedness of the network. In addition, we introduce a node removal process allowing the model network to reach a stationary state, where its overall structural characteristics become invariant of time. By adjusting the model parameters, it is possible to recover several characteristics of a real-world temporal network of mobile phone calls. These include heterogeneous degrees and weights, rich community structure, and weight-topology correlations and weak tie properties as suggested by Granovetter [8].

The main advantages of this model are that (a) it is able to mimic the time-varying nature of directed interactions characterising several real systems; (b) it helps in understanding the importance of different mechanisms in shaping the emerging network structure; and (c) it allows us to test their effect in a scalable fashion on evolving dynamical processes. This last point is of special importance, as e.g. spreading processes have been shown to have significantly different critical behaviour when transmitted by time-varying interactions. In addition, by varying the model parameters one can control the average degree, strength of ties, size of communities, clustering, and inter-connectedness of the network. These features can control e.g. the speed of spreading as we have demonstrated via a simple study on SI processes, but play crucial effects in case of other processes like random walks, the diffusion of information, epidemics, or social contagion processes, making the model an ideal testbed for simulating such processes.

This model can be extended in several possible ways. So far it integrates mechanisms, which shape the structure of the emerging temporal network, while still neglecting temporal inhomogeneities such as burstiness of links. Thus one important challenge would be to consider the bursty nature of human interactions or even higher-order temporal correlations. This would provide effective ways for studying the interplay between structural and temporal correlations, which may cause acceleration [48] or slowing down [24,26] of information diffusion in real temporal network.

Appendix: Algorithmic model definition

Here we describe the algorithmic implementation of the temporal network model introduced in the main text. The main function Temporal network model ($G_{t=0}$, $T$, $p_{\Delta}$, $p_{r}$, $\delta$) takes as input a weighted network with empty link set $G_{t=0} = (V, \emptyset, \{a_i\}, \{w_{ij}^0\})$ with size $N$, pre-assigned node activity probabilities per unit time $\{a_i\}$, and initial link weights $\{w_{ij}^{t=0} = 0\}$. Further parameters are $T$: the number of iterations; $p_{\Delta}$: probability of triadic closure; $p_{r}$: probability of node deletion; and $\delta$: link reinforcement increment. In the pseudo code, $n_i = |V_i|$ denotes the current degree of node $i$; rand() is a pseudo-random number generator, which returns a rational number between $[0,1]$; and $p_{ij}^t = w_{ij}^t / \sum_{k \in V_i} w_{ik}^t$ is the probability to select a random neighbour $j$ from the current neighbour set $V_i^t$ of node $i$ weighted by the number of interactions between them performed since the link was created up to time $t$ ($\sum_{j \in V_i} p_{ij}^t = 1$). Temporal events between nodes $i$ and $j$ performed at the current iteration step $t$ are denoted by $[t, i, j]$.

The main function calls two subroutines called Cyclic closure ($i$, $G_i$) and Focal closure ($i$, $G_i$, $X$) with input parameters $i$ denoting the currently active node; $G_i$ the current weighted network structure; and $X$ exception set of nodes to connect in the current call.

J.S. acknowledges support by the Academy of Finland, project No. 260427 (“Temporal Networks of Human Interactions”). We thank A.-L. Barabási for providing the MPC data.
Algorithm 1 Temporal network model

Input: $G_{t=0}$, $T$, $p_d$, $p_c$, $\delta$

for $t = 1$ to $T$ do
   for $N$ randomly selected node $i \in V$ do
      if $\text{rand}() \leq p_d$ then
         remove all links of node $i$ ▶ delete and re-insert $i$
      end if
      if $\text{rand}() \leq a_i$ then
         ▶ node $i$ is active
      end if
      if $\text{rand}() \leq p(n_i)$ then
         ▶ link creation
      end if
      if $n_i = 0$ then
         ▶ node $i$ has no neighbour
         $\text{Focal closure}(i, \{i\}, G_t)$
      else
         $\text{Cyclic closure}(i, G_t)$
      end if
      if $\text{rand}(\cdot) \leq \delta$ then
         $\text{Cyclic closure}(i, G_t)$
      else
         ▶ link reinforcement
         Select a $j$ neighbour of $i$ with probability $p_{ij}^t$
         Event: $[t, i, j]$ $w_{ij}^t = \delta$
      end if
   end for
end for

Algorithm 2 Cyclic closure: as input $i$ is the current active node; and $G_t$ is the current weighted network

Input: $i, G_t$

Select a $j$ neighbour of $i$ with probability $p_{ij}^t$
if $n_i = 1$ then
   ▶ node $j$ has one neighbour
   $\text{Focal closure}(j, \{i, j\} \cup V^t, G_t)$
else
   Select a $k(\neq i)$ neighbour of $j$ with probability $p_{jk}^t$
   if $(i, k) \notin E_t$ then
      ▶ triangle $(i, j, k)$ is not closed
      Event: $[t, i, k]$ $w_{ik}^t = 1$
   else
      $\text{Focal closure}(j, \{i, j, k\} \cup V^t, G_t)$
   end if
end if
Event: $[t, i, k]$ $w_{ik}^t = \delta$
end if

Algorithm 3 Focal closure: as input $i$ is the current active node; $X$ is the exception set of nodes to connect; and $G_t$ is the current weighted network

Input: $i$, $X$, $G_t$

Select a random node $j \in V \setminus X$
Event: $[t, i, j]$ $w_{ij}^t = 1$

References

1. C. Castellano, S. Fortunato, V. Loreto, Rev. Mod. Phys. 81, 591 (2009)
2. S. Wasserman, K. Faust, in Social Network Analysis: Methods and Applications (Structural Analysis in the Social Sciences), 1st edn. (Cambridge University Press, 1994), Vol. 8
3. D. Lazer et al., Science 323, 721 (2009)
4. A. Vespignani, Science 325, 425 (2009)
5. D.J. Watts, Proc. Natl. Acad. Sci. 99, 5766 (2002)
6. D.J. de Solla Price, Science 149, 510 (1965)
7. R. Albert, A.-L. Barabási, Rev. Mod. Phys. 74, 47 (2001)
8. M.S. Granovetter, Am. J. Sociol. 78, 1360 (1973)
9. J.P. Onnela et al., Proc. Natl. Acad. Sci. 104, 7332 (2007)
10. A. Rapoport, Bull. Math. Biophys. 19, 257 (1957)
11. S. Fortunato, Phys. Rep. 486, 75 (2010)
12. J.M. Kumpula et al., Phys. Rev. Lett. 99, 228701 (2007)
13. G. Palla, A.-L. Barabási, T. Vicsek, Nature 446, 664 (2007)
14. J. Saramäki et al., Proc. Natl. Acad. Sci. 111, 942 (2014)
15. P. Holme, J. Saramäki, Phys. Rep. 519, 97 (2012)
16. L. Kovanen et al., in Temporal Networks, Understanding Complex Systems, edited by P. Holme, J. Saramäki (Springer, Berlin, Heidelberg, 2013), pp. 119–133
17. L. Kovanen, M. Karsai, K. Kaski, J. Kertész, J. Saramäki, J. Stat. Mech. 2011, P11005 (2011)
18. K. Zhao, J. Stehlé, G. Bianconi, A. Barrat, Phys. Rev. E 83, 056109 (2011)
19. M. Karsai, N. Perra, A. Vespignani, Sci. Rep. 4, 4001 (2014)
20. A.-L. Barabási, Nature 435, 207 (2005)
21. A.-L. Barabási, Burst: The Hidden Pattern Behind Everything We Do, 1st edn. (Dutton Adult, 2010)
22. L. Kovanen, K. Kaski, J. Kertész, J. Saramäki, Proc. Natl. Acad. Sci. 110, 18070 (2013)
23. G. Miritello, R. Lara, M. Cebrian, E. Moro, Sci. Rep. 3, 1950 (2013)
24. M. Karsai et al., Phys. Rev. E 83, 025102 (2011)
25. M. Kivelä et al., J. Stat. Mech. 2012, P03005 (2012)
26. G. Miritello, E. Moro, R. Lara, Phys. Rev. E 83, 045102 (2011)
27. R. Lambiotte, L. Tasson, J.-C. Delvenne, Eur. Phys. J. B 86, 320 (2013)
28. T. Takaguchi, T. Nakamura, P. Holme, PLoS One 8, e68989 (2013)
29. B. Min, K.I. Goh, A. Vazquez, Phys. Rev. E 83, 036102 (2011)
30. L. Rocha, A. Decuyper, V. Blondel, Dynamics On and Of Complex Networks, in Modeling and Simulation in Science, Engineering and Technology, edited by A. Mukherjee et al. (Springer, New York, 2013), Vol. 2, pp. 301–314
31. N. Perra, B. Gonçalves, R. Pastor-Satorras, A. Vespignani, Sci. Rep. 2, 469 (2012)
32. M.V. Tomassetti et al., Sci. Rep. 4, 5679 (2014)
33. N. Perra et al., Phys. Rev. Lett. 109, 238701 (2012)
34. M. Starnini, R.P. Satorras, Phys. Rev. E 87, 062807 (2013)
35. S. Liu, N. Perra, M. Karsai, A. Vespignani, Phys. Rev. Lett. 112, 118702 (2014)
36. J.M. Kumpula et al., Comput. Phys. Commun. 280, 517 (2009)
37. R.I.M. Dunbar, J. Hum. Evol. 22, 469 (1992)
38. E. Ubaldi et al., to be published (2015)
39. G. Bianconi, R.K. Darst, J. Iacovacci, S. Fortunato, Phys. Rev. E 90, 042806 (2014)
40. F. Heider, J. Phys. 21, 107 (1946)
41. K.H. Wolff, The Sociology of Georg Simmel (Nabu Press, 2011)
42. A. Chmiel, K. Kowalska, J.A. Holyst, Phys. Rev. E 80, 066122 (2009)
43. J.-P. Onnela et al., New. J. Phys. 9, 179 (2007)
44. C. Cattuto et al., PLoS One 5, e11596 (2010)
45. M. Newman, Networks: An Introduction, 1st edn. (Oxford University Press, USA, 2010)
46. G. Krings, M. Karsai, S. Bernardsson, V.D. Blondel, J. Saramäki, EPJ Data Science 1, 4 (2012)
47. A. Barrat, M. Barthélemy, A. Vespignani, Dynamical Processes on Complex Networks, 1st edn. (Cambridge University Press, 2008)
48. Rocha et al., PLOS Comput. Biol. 7, e1001109 (2011)