Applications of Temporal Graph Metrics to Real-World Networks

John Tang, Ilias Leontiadis, Salvatore Scellato, Vincenzo Nicosia, Cecilia Mascolo, Mirco Musolesi, and Vito Latora

Abstract Real world networks exhibit rich temporal information: friends are added and removed over time in online social networks; the seasons dictate the predator-prey relationship in food webs; and the propagation of a virus depends on the network of human contacts throughout the day. Recent studies have demonstrated that static network analysis is perhaps unsuitable in the study of real world network since static paths ignore time order, which, in turn, results in static shortest paths overestimating available links and underestimating their true corresponding lengths. Temporal extensions to centrality and efficiency metrics based on temporal shortest paths have also been proposed. Firstly, we analyse the roles of key individuals of a corporate network ranked according to temporal centrality within the context of a bankruptcy scandal; secondly, we present how such temporal metrics can be used to study the robustness of temporal networks in presence of random errors

J. Tang · I. Leontiadis · S. Scellato · C. Mascolo
Computer Laboratory, University of Cambridge 15 JJ Thomson Avenue,
Cambridge CB3 0FD, UK

V. Nicosia
Computer Laboratory, University of Cambridge 15 JJ Thomson Avenue,
Cambridge CB3 0FD, UK

Laboratorio sui Sistemi Complessi, Scuola Superiore di Catania,Via Valdisavoia 9,
95123 Catania, Italy

M. Musolesi (✉)
School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
e-mail: mirco.musolesi@acm.org

V. Latora
Laboratorio sui Sistemi Complessi, Scuola Superiore di Catania,Via Valdisavoia 9,
95123 Catania, Italy

School of Mathematical Sciences, Queen Mary, University of London, E1 4NS London, UK

Dipartimento di Fisica e Astronomia and INFN, Universitá di Catania and INFN, Via S. Sofia 64,
95123 Catania, Italy

P. Holme and J. Saramäki (eds.), *Temporal Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-642-36461-7_7,

and intelligent attacks; thirdly, we study containment schemes for mobile phone malware which can spread via short range radio, similar to biological viruses; finally, we study how the temporal network structure of human interactions can be exploited to effectively immunise human populations. Through these applications we demonstrate that temporal metrics provide a more accurate and effective analysis of real-world networks compared to their static counterparts.

1 Introduction

Temporal graph metrics [48, 49] represent a powerful tool for the analysis of real-world dynamic networks, especially with respect to the aspects for which time plays a fundamental role, such as in the case of spreading of a piece of information or a disease. Indeed, existing metrics are not able to characterise the temporal structure of dynamic networks, for example in terms of centrality of nodes over time. For these reasons, new metrics have been introduced, such as temporal centrality, in order to capture the essential characteristics of time-varying graphs. A detailed description of the metrics used in this chapter can be found in [40].

In this chapter we will discuss a series of possible applications of temporal graph metrics to the analysis of real-world time-varying networks. This chapter is structured as follows. We will cover our work in this area and, finally, we will discuss contributions in this fields by other researchers and potential future applications, in particular in the area of the modelling of epidemic spreading.

More specifically, in Sect. 2 we analyse the roles of key individuals according to temporal centrality within the context of the Enron scandal [49]. In Sect. 3 we study how such temporal metrics can used to study the robustness of temporal networks in presence of random errors and intelligent attacks [45]. Then, Sect. 4 we present a containment scheme for mobile phone malware which can spread via short range radio transmission [50, 51]. Finally, in Sect. 5 we discuss existing and potential applications to human epidemiology, outlining some research directions in these areas.

2 Corporate Networks

2.1 Overview

The Enron Energy Corporation started as a traditional gas and electrical utility supplier; however, in the late 1990s their main money making business came from trading energy on the global stock markets [18]. In December 2001, the Enron Energy Corporation filed for bankruptcy after it was uncovered that fraudulent accounting tricks were used to hide billions of dollars in debt [23]. This led to the eventual conviction of several current and former Enron executives [8, 55].

The investigation also brought to light the reliance of the company on traders to bring in profits using aggressive tactics culminating in intentional blackouts in California in Summer 2001. With both control over electricity plants and the ability to sell electricity over the energy markets, Enron traders artificially raised the price of electricity by shutting down power plants serving the State of California and profiting by selling electricity back at a premium [7].

During the investigation into the Enron accounting scandal, telephone calls, documents and emails were subpoenaed by the U.S. government and as such the email records of 151 user mailboxes were part of the public record consisting of approximately 250,000 emails sent and received during the period between May 1999 to June 2002 (1,137 days), leading up to the bankruptcy filing. None of the emails were anonymised and so they provide unique semantic information of the owner of each mailbox.

2.2 Temporal Graph Construction

In our analysis, we use the dataset prepared by Shetty and Adibi [47]. Since we do not have a complete picture of the interactions of users outside of the subpoenaed mailboxes we concentrate on email exchanges between the core 151 users only. Taking this email dataset, we process the complete temporal graph from 1999 to 2002 with undirected links, using windows of size $w = 24\,\text{h}$ and horizon $h = 1$. If an email was exchanged between two individuals in a temporal window, a link between the two nodes representing those individuals will be added to the graph representing the temporal snapshot for that time.

2.3 Semantic Value of Temporal Centrality

Figure 1 plots the static and temporal centrality rankings of employees calculated using closeness and betweenness. Examining the static centralities (left column) we note that there is little difference between the top five employees using static closeness or betweenness. Also plotting the static degree centrality of each node[1], we notice similar rankings suggesting that static analysis only favours employees who interacted with the most number of other people. Temporal closeness and temporal betweenness yield different rankings amongst the top five and the calculated Kendall-tau correlation coefficient [31] (Table 1) confirm that static-to-static metrics are strongly correlated ($\simeq 0.7$). Also note that there is low correlation (<0.4)

[1]The static degree centrality is defined as the number of edges connected to a node i, normalised by the total possible neighbour nodes $(n-1)$ [56].

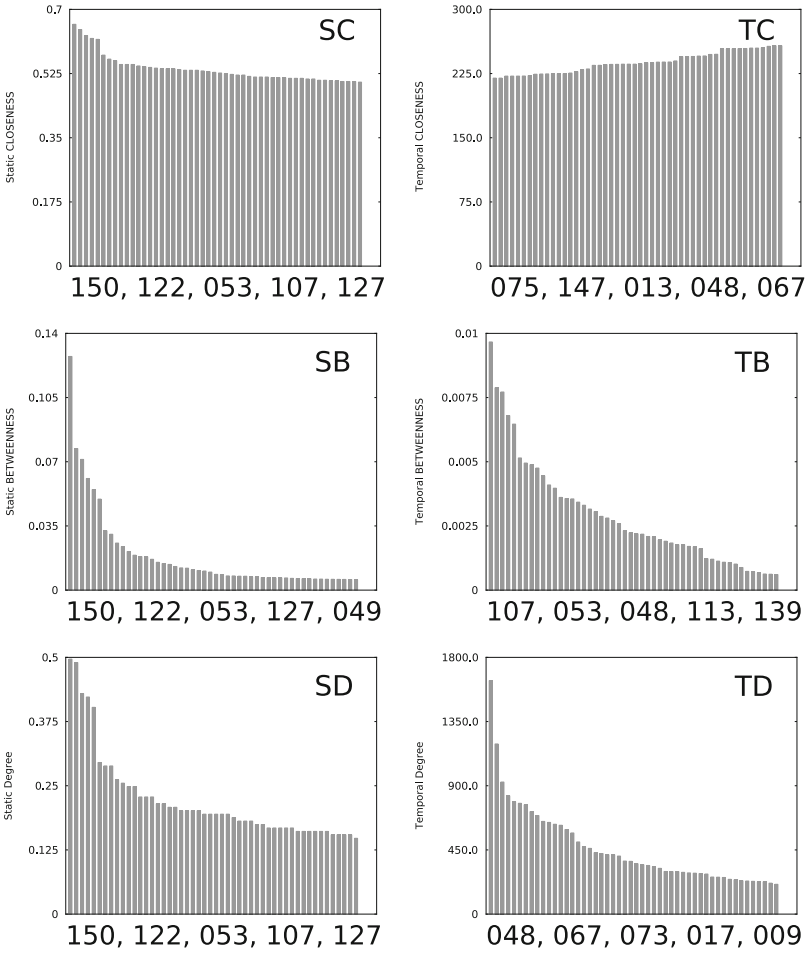

Fig. 1 Ranked distribution of top 50 statically (S) and temporally (T) central nodes. From top row: Closeness (C), Betweenness (B), and Degree (D). Top 5 node ID's listed under each plot

between temporal metrics and static degree demonstrating that temporal analysis is not dependent on the number of people an individual interacts with.

Cross referencing the top two employee identifiers with their position within the organisation (Table 2) we identify a secretary (150) and managing director (122) as central nodes for both static closeness and betweenness; however, both temporal closeness and betweenness consistently selected employees in trading roles (053, 075, 107, 147). A secretary and a managing director are certainly important for information dissemination and central to many communication channels, as detected by static measures. However, instead the top trading executives are exclusively favoured by temporal analysis.

Table 1 Kendall-tau correlation coefficients between centralities

	SB	SC	SD	TB	TC	TD
SB	1.00	0.57	0.69	0.41	0.24	0.43
SC	–	1.00	0.70	0.36	0.22	0.31
SD	–	–	1.00	0.39	0.28	0.48
TB	–	–	–	1.00	0.43	0.34
TC	–	–	–	–	1.00	0.40
TD	–	–	–	–	–	1.00

Table 2 Roles of top centrality nodes

ID	Role	Notes
9	(Unknown)	
13	Legal	Senior Legal Specialist
17	Manager	
48	Executive	
53	Trader	
54	President	Former Head of Trading
67	Vice President	Enron Wholesale Services
73	Trader	
75	Director of Trading	
107	Trader	Head of Online Trading
122	Managing Director	
127	Chairman and CEO	
139	Director	
147	Trader	
150	Secretary	Assistant to Greg Whalley

To show that temporal analysis does not simply uncover nodes with the most interactions with other people, we also plot the temporal degree (TD) calculated as the total number of emails sent and received by each node i. Since there is a low correlation (<0.4) with temporal closeness and betweenness this shows that temporal analysis is not dependent on the number of emails sent and received by each individual.

2.4 Effectiveness of Central Nodes on Dynamic Processes

2.4.1 Trace-Driven Simulation Setup

To evaluate the role and the centrality of the employees identified by temporal and static analysis, we consider two dynamic processes. First, we simulate a simple information *dissemination process* over the temporal graph constructed from the Enron traces. The process is simulated as follows. We select the top N nodes from the ranking based on temporal closeness centrality. We place an identical message m into their (infinite) buffers. We refer to any node that has received a copy of this

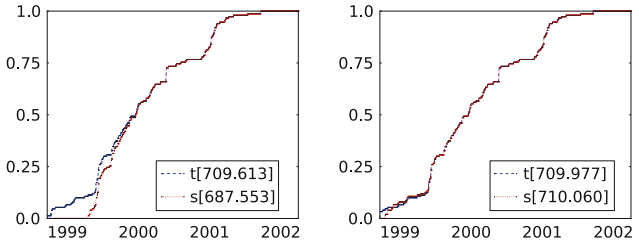

Fig. 2 Dissemination process: Dissemination ratio starting from top 2 (*left*) and top 10 (*right*) closeness source nodes. Area under curve reported in legend for temporal (t) and static (s) centrality

message as *reached*. We then replay the contact trace through time and as reached nodes make contact with an unreached node u, the message is replicated into the buffer of node u. We assume that messages are transferred instantaneously and only the first neighbour in a time window can be reached. We then repeat this for static closeness centrality and plot the dissemination ratio across time for both.

Second, to model the role of individuals as part of an information *mediation process*, we borrow concepts from the more commonly known epidemic immunisation process where the dissemination ratio of a contagion spreading throughout a static network is measured before and after certain nodes are immunised against the contagion [2]. This is analogous to measuring the spread of information (the contagion) before and after important individuals are removed from the network (such as going on holiday or being discharged) since our conjecture is that removing mediators will impact the network communication efficiency greatly.

In the trace-driven simulation, instead of a single message spreading within the organisation, we seed all employees with a different message that needs to be delivered to all other employees. This models multiple channels of communication. In order to derive a baseline performance, we start by calculating the dissemination ratio when no nodes are removed. We then remove the top N individuals identified by temporal betweenness and rerun the information spreading process. Nodes which are removed cannot receive or pass on messages. We then repeat the same process for comparison using static betweenness centrality for the ranking.

2.4.2 Evaluating Information Dissemination and Mediation

We present plots using $N = \{2, 10\}$ for information dissemination (Fig. 2) and information mediation (Fig. 3). As we can see the different pairs of traders identified by temporal analysis are better than the arbitrary nodes selected by static analysis for both disseminating information through the organisation and acting as mediators between communication channels. In the information dissemination case, although the final dissemination is the same across the long period of time, the two traders selected by temporal analysis disseminate information quicker. Only after increasing

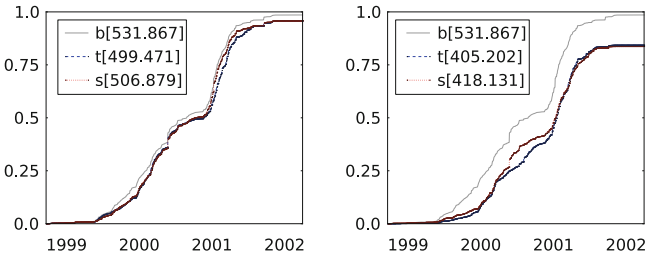

Fig. 3 Mediation process: Dissemination ratio after removing top 2 (*left*) and top 10 (*right*) betweenness nodes. Area under curve reported in legend for temporal (t), static (s) and baseline (b) where no nodes are removed

to 10 nodes the static analysis presents similar results. In the information mediation case, the final dissemination ratios for both temporal and static centrality nodes slightly decreases by removing the nodes but are comparable. However, removing the two traders gives an overall more prolonged drop in message dissemination. In the case of the removal of 10 nodes, the individuals identified by means of the temporal metrics slow the dissemination process further compared to static ones.

2.5 Summary of the Findings

This study has demonstrated the advantages of using temporal network analysis over traditional static, aggregated graphs. Although centrality rankings are quantitative measures of node importance, their physical meaning is very much qualitative. For this reason, we have shown that temporal centrality provides a more accurate identification of key people in a corporate social network where each persons role is known. Taking a second perspective, we demonstrate that temporal centrality can identify nodes which can spread and mediate information, better than static analysis. This demonstrates the importance of temporal information for applications, which are applied to real networks.

3 Network Robustness

3.1 Overview

The study of real-world communication systems by means of complex network models has provided insightful results and has vastly expanded our knowledge on how single entities create connections and how these connections are used for communication or, more generally, interaction [3]. In particular, technological networks

such as the Internet and the World Wide Web have been under scrutiny in terms of structure and dynamic behavior [22, 27]. More recently, with the widespread adoption of mobile and opportunistic networks, it has become important to develop new analytical tools to keep into account network dynamics over time [30, 32] and how this affects phenomena such as information propagation [10, 14].

At the same time, the problem of understanding whether real systems can sustain substantial damage and still maintain acceptable performance has been extensively addressed [1, 5]. Various measures of network robustness have been defined and investigated for several classes of networks, evaluating how different system can be more or less resilient against random errors or targeted attacks thanks to their underlying structural properties [15, 26].

Nonetheless, it is still unclear how to approach the study of robustness of networks by taking into account their time-varying nature: by adopting a static representation of a temporal network, important features which impact the actual performance might be missed. Thus, it becomes important to develop a robustness metric which takes into account the temporal dimension and gives insights on how a mobile network is affected by damage or change. Particularly, the fact that links are not always active means that information spreading can be delayed or even stopped and that *relative ordering in time of connection events may affect the creation of temporal paths in mobile networks*.

Our main goal is to design a novel framework for the analysis of robustness in time-varying networks. We adopt *temporal network metrics* [48] to quantify network performance and define a measure of robustness against generic network damages. At first, we study its performance on random network models to understand its properties; then we apply our method to study a real mobile network, describing how temporal robustness gives a more accurate evaluation of system resilience than static approaches.

3.2 A Framework to Evaluate Robustness of Temporal Networks

The study of robustness of complex networks has mainly focused on describing how a given performance metric of the network is affected when nodes are removed according to a certain rule. The underlying assumption is that the absence or malfunctioning of some nodes will cause the removal of their edges and, thus, some paths will become longer, increasing the distances between the remaining nodes, or completely disappear, resulting in the loss of connectivity in the whole system. In this work we will study the problem of defining and analyzing robustness in evolving networks: as a consequence, *we need to use a performance metric which includes the temporal dimension in its definition*. We choose to adopt temporal efficiency as the performance metric. We then consider random and independent failures for every node and we evaluate how the system tolerates increasing level of malfunctioning nodes.

Since a temporal graph is continuously evolving, we can evaluate how temporal efficiency changes over time by considering a value τ and evaluating $E_G(t)$ as the relative temporal efficiency of the temporal graph in the time window $[t-\tau, t]$. The effect of τ is to effectively impose an upper bound on the temporal distances, as all paths longer than τ simply do not exist. As a consequence, τ should be chosen so that any communication whose delay is longer than τ itself can be ignored.

Given a temporal graph G, we define a damage D as any structural modification on it and we define G_D as the graph resulting by the effect of the damage D on G. A damage may be the deactivation of some nodes or the removal of some edges at a particular time t_D. Because of damage D, some temporal shortest paths will be longer or will not exist any more, thus, we expect that the temporal efficiency will eventually reach a new steady value $E_{G_D} \leq E_G$. It is important to evaluate the new value of the temporal efficiency on a new temporal graph which still contains the deactivated nodes, in order to obtain a decrease in efficiency. Otherwise, we might obtain a smaller temporal graph which is more efficient than the original graph, although it has lost much of its structure. Hence, we do not consider highly dynamic systems where nodes can be constantly added or removed. Instead, we focus on evaluating the service degradation in a more controlled environment where only a number of existing nodes could fail.

We define the loss in efficiency $\Delta E(G, D)$ caused by the damage D on the temporal graph G as $\Delta E(G, D) = E_G - E_{G_D}$. Finally, we define the *temporal robustness* $R_G(D)$ of the temporal graph against the damage D as

$$R_G(D) = 1 - \frac{\Delta E(G, D)}{E_G} = \frac{E_{G_D}}{E_G} \tag{1}$$

This value is normalized between 0 and 1 and it measures the relative loss of efficiency caused by the damage: if the damage does not impact the efficiency of the graph ($E_{G_D} = E_G$) then its robustness is 1, while if the damage destroys the efficiency of the graph ($E_{G_D} = 0$) the robustness drops to 0. Temporal efficiency is a particularly suitable metric to study temporal network robustness as it denotes both longer temporal paths and the lack of paths among temporally disconnected nodes at the same time. Nonetheless, other metrics have been used to assess robustness in static systems: for instance, there could be scenarios where fast communication with small delays can be more important than global connectivity, thus other measures can be adopted. Provided that these measures can be extended to the temporal case, they can be easily integrated in our framework.

3.3 Robustness of Random Temporal Networks

In this section, we present a numerical analysis of temporal robustness for different classes of random temporal networks: an Erdős–Rényi temporal model, a Markovian temporal model and mobility-based temporal model.

3.3.1 Random Temporal Network Models

An Erdős–Rényi (ER) random graph with N nodes and parameter p is created by independently including each possible edge in the graph with probability p and it is denoted as $G(N, p)$ [19]. We generalize this model to the temporal case by creating a sequence of T ER random graphs $G(N, p)$ and we denote the resulting temporal graph as $G(N, p, T)$.

The temporal ER network model does not provide temporal correlations between consecutive graphs in the sequence. We thus consider a model where link evolution is described by a Markov process, thus enabling memory effects in network dynamics. We consider a complete graph G with N nodes. At every discrete timestep t each link may or may not be present: a temporal graph is created where the existence of each link evolves according to a two-state discrete Markov process. We denote with p the probability that a link present at time t will be removed at time $t + 1$ and with q the probability that a link which is not present at time t will be added at time $t + 1$. The steady probability of link presence then is $P_{ON} = \frac{q}{p+q}$: as a consequence, each observation of the temporal graph appears as an ER random graph with each edge present with probability P_{ON}.

We also create a random model of a temporal network by using mobility models. In this case we are introducing topological constraints: a key difference with the previous temporal models is that each node is not equally likely to connect with all the other nodes, due to the effect of spatial distance. We consider $N = 100$ nodes moving in a square area $1{,}000 \times 1{,}000$ m and we define a communication range r: at every time step, we create a graph where nodes are connected if their Euclidean distance is shorter than r. Thus, we change the probability of link presence P_{ON} by varying the communication range. Then, a temporal graph can be defined as the sequence of graphs extracted at each time step while the nodes move. We investigate two different mobility models that are implemented using the Universal Mobility Model Framework [37]: Random Waypoint Model (RWP) and Random Waypoint Group Model (RWPG). In RWP each node selects uniformly at random a location towards which it moves with speed uniformly distributed in a fixed range $[5, 40]$ mph. As the node reaches its destination, it waits for a randomly distributed time in $[0, 120]$ s and repeats the above steps until the end of the simulation.

In RWPG nodes are divided into two classes: there are M group leaders and $N - M$ group followers. Every group followers has its own leader so that the N nodes are divided into equally-sized groups. Each group leader selects a random target and moves towards it, according to the RWP mobility model. Group members do not select any target; instead, they follow their group leader according to the *pursuit force* [37] which is set to give a group span of 200 m.

3.3.2 Numerical Evaluation

We numerically evaluate temporal efficiency $E_G(t)$ over time, adopting a time window of $\tau = 100$, for a graph with $N = 100$ nodes: after an initial phase,

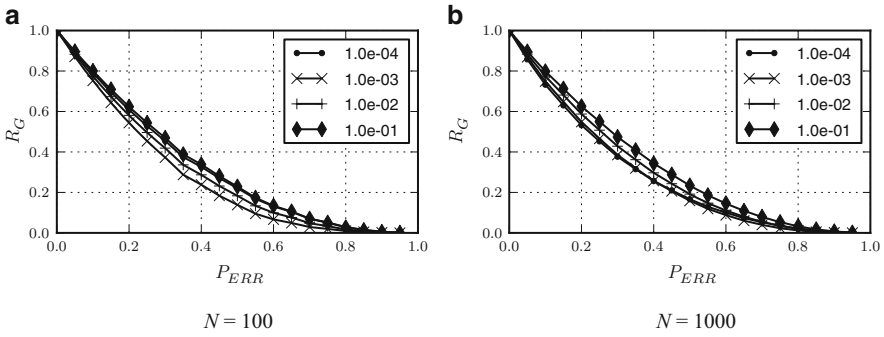

Fig. 4 Temporal robustness R_G as a function of probability of error P_{ERR} in the ER random model for different link probability p (a) 100 nodes; b) 1000 nodes). The size of the system has no impact on temporal robustness: furthermore, the system fails smoothly as the probability of error increases

the random temporal graph reaches an equilibrium state and we compute the steady value of temporal efficiency. We run each simulation for 2τ steps and we compute the average value of temporal efficiency over the last τ steps. All results have been averaged over 100 different runs. We evaluated numerically temporal robustness by deactivating each node independently with increasing probability P_{ERR}. We measure temporal efficiency before and after the failure, when the network reaches a new equilibrium state.

As reported in Fig. 4, the temporal ER model fails smoothly as we increase the fraction of removed nodes, without any sudden disruption for any value of P_{ERR}. This is a main difference with respect to what happens in the static case: for a static ER random graph there may exist a critical value of P_{ERR} which causes a breakdown of the network in several disconnected components [1]. This is not true for temporal robustness, as new paths can still appear after the damage as the network rearranges its connections. Time provides more redundancy and, hence, more resilience. Moreover, we also note that temporal robustness does not depend on system size: since it is normalized with respect to the value of temporal efficiency before the damage, it depends only on the relative drop in efficiency, not on absolute values.

As shown in Fig. 5a, temporal robustness is affected by probability of error P_{ERR} in the same way as in the temporal ER model: the system fails gradually as more nodes are removed. However, for intermediate values of P_{ON} robustness has lower values. At the same time, high and low values of P_{ON} provide the same robustness, even if the absolute value of temporal efficiency can be very different, thanks to the normalization of temporal robustness.

In the case of mobility-based temporal networks, reported in Fig. 5b, both RWP and RWPG exhibit a similar behavior: again, the network loses efficiency in a smooth way and temporal robustness is not affected by P_{ON} in this case as the spatial characteristics of the network are mainly affecting the resulting robustness.

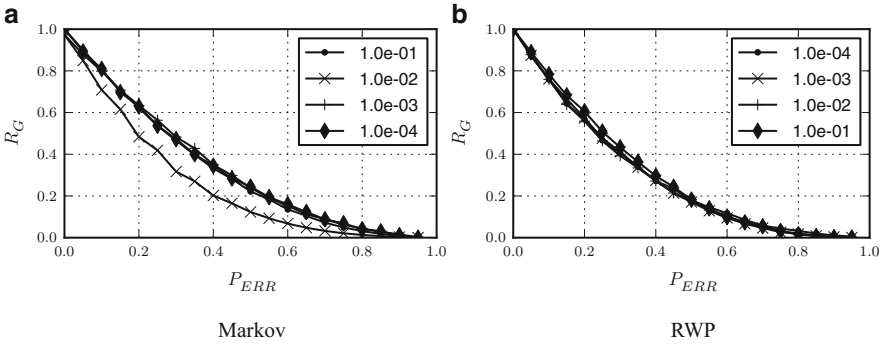

Fig. 5 Temporal robustness R_G as a function of probability of error P_{ERR} and for different values of P_{ON} for a) the Markov-based and for b) the RWP random model (RWPG does not deviate from RWP)

3.4 Case Study: Cabspotting

We have seen that temporal networks do not exhibit sudden breakdowns when nodes are being removed and that various temporal network models exhibit analogies in their resilience. We now shift our attention to real time-varying networks: our aim is to understand whether temporal robustness gives us more information than static robustness in a real case and to investigate whether random models can offer a good approximation to real networks.

3.4.1 Dataset

This case study is based on Cabspotting, a publicly available dataset of mobility traces: the Cabspotting project tracked taxi cabs in San Francisco traveling through all the Bay Area for about 2 years with the aim of gathering data about city life [41]. The vehicles were equipped with GPS sensors and every device was periodically updating its position and uploading it to a central server to be stored, along with the timestamp of the record. Thus, it is possible to reconstruct each taxi's trajectory over space. For pictorial representations of the dataset, please refer to the project website [4].

We have selected an area of about 20 km×20 km around the city of San Francisco and we have extracted 24 consecutive hours of mobility traces, corresponding to Wednesday, 21 May 2008. After this, we have generated an artificial contact trace by defining a communication range of 200 m for the vehicles, which roughly corresponds to WiFi connectivity range in similar scenarios [11]: whenever two cars are within this distance they can communicate to each other. Time granularity is in seconds, so we have a sequence of 86,400 graphs with 488 nodes and more than 350,000 contacts among them. The average contact duration is about 2 min while the average inter-contact time is more than 2.5 h.

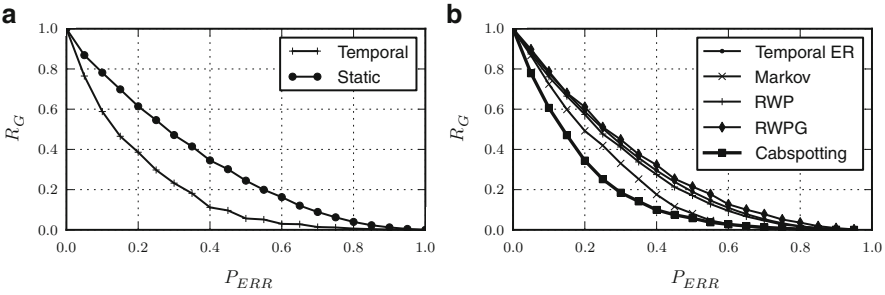

Fig. 6 Comparison between temporal robustness R_G and static robustness as a function of probability of error P_{ERR} for the Cabspotting dataset (**a**). The static approach overestimates system robustness. *Right*: Comparison between the temporal robustness R_G of the dataset and random null models with the same number of nodes N and P_{ON} of the Cabspotting temporal network (**b**)

3.4.2 Analysis

We study the reaction of the Cabspotting temporal network to random failures and compare it to our findings on random models. We adopt numerical simulation, but since the temporal dynamics of this network is not stationary, we can not compare the temporal efficiency E_G before and after a certain error, because the two temporal window will likely have already different properties. Instead, we fail nodes according to P_{ERR} at the very first time step of the temporal sequence of graphs: in this way, we can compare the average temporal efficiency over all the time for the original network and for the damaged one. We adopt a value of $\tau = 3{,}600$, which allows us to consider temporal paths up to 1 h, even if such longer paths can not contribute much to temporal efficiency.

The first comparison that we show in Fig. 6a is between static robustness and temporal robustness for the Cabspotting temporal network. In this case static robustness is computed on the static graph obtained by aggregating all the contacts in the trace and adopting static global efficiency as performance measure. Since the resulting static graph contains more than 100,000 edges it is clearly an overestimation of the communication properties of the real system, as not all these links are continuously available over time and some paths can not be used due to temporal ordering constraints. Indeed, static robustness appears much larger than the temporal counterparts: only temporal robustness is able to capture the realistic communication capabilities of the system and how they are affected by random failures.

Then, we attempt to understand if the various random temporal network models we have studied can be used to approximate the robustness properties of the real scenario. For each model, we compute the temporal robustness as a function of P_{ERR} for a network with the same number of nodes N and the same P_{ON} measured in the Cabspotting temporal network (about 0.005), using the same simulation parameters as in the real scenario. As reported in Fig. 6b, all temporal networks present the same

trend in network robustness, albeit random models have higher values of temporal robustness than the real network. Interestingly, the closest match is the Markov-based temporal model, while the mobility-based models are closer to the ER model than to the Cabspotting network, even if this is actually a mobility-based contact network. However, the assumptions used in mobility models require homogeneity of space and absolute freedom to move continuously and independently in a boundless area, while in reality taxis are usually constrained to move on streets and bridges and they often move together along the same direction or stop together in a particular place to wait for customers (i.e., airport or stations). The Markov model, instead, introduces the type of time correlations that appear to better mimic the real scenario. In fact, the most important aspect that needs to be captured is time ordering of events: in random mobility models connections do not follow particular time patterns, whereas real traces do (rush hour, working hours, human sleeping cycles). Only temporal robustness can take into account these unique characteristics.

3.5 Summary of the Findings

These two results provide evidence that *temporal robustness is a more accurate measure to be used on mobile networks instead of standard static approaches*. Therefore, when testing protocols and applications to be deployed in mobile networks, a temporal study is more meaningful and should not be substituted by a static approximation.

4 Mobile Malware

4.1 Overview

Smartphones are not only ubiquitous, but also an essential part of life for many people who carry such devices through their daily routine. It comes at no surprise then that recent studies have shown that the mobility of such devices mimic that of their owners' schedule [17, 54]. This fact constitutes an opportunity for devising efficient protocols and applications, but it also represents an increasing security risk: as with biological viruses that can spread from person to person, mobile phone viruses can also leverage the same social contact patterns to propagate via short-range wireless radio such as Bluetooth and WiFi. For example, when security researchers downloaded *Cabir* [53]—the first proof-of-concept piece of mobile malware—for analysis, they soon discovered the full risk potential of the mobile worm as it broke loose, replicating from the test device to external mobile phones. This event prompted the need for specially radio shielded rooms to securely test such malicious code [28].

Unlike desktop computers mobile malware can spread through both short-range radio (i.e., Bluetooth and WiFi) and long-range communication (SMS, MMS and email) [33]. Long-range malicious traffic can potentially be contained by the network operator by scanning every message against a database of known malware or spam [52], however, short-range propagation might fall under the radar of centralised service providers: effective schemes to defend against short-range mobile malware spreading are necessary. In addition, while a global patching of the devices through cellular connectivity is the natural solution and is in theory possible, in practice, due to associated costs and resource consumption, this is not ideal. For example, there are potential constraints with respect to the cellular network capacity and server bandwidth (with respect to the latter, similar issues have also investigated for software updates distribution in the Internet, see for example [25]).

4.2 Temporal Centrality Metrics for Malware

Being highly correlated with human contacts, understanding how such malware propagates requires an accurate analysis of the underlying time-varying network of contacts amongst individuals. State-of-the-art solutions on mobile malware containment have ignored two important temporal properties: firstly, the time order, frequency and duration of contacts; and secondly, the time of day a malicious message starts to spread and the delay of a patch [57, 58]. Instead, we argue that the temporal dimension is of key importance in devising effective solutions to this problem.

With this in mind, the focus of this study is to investigate the effectiveness of two containment strategies based on targetting key nodes, taking into account these temporal characteristics. We firstly investigate a traditional strategy, inspired by studies on error and attack tolerance of networks [1], exploiting static and a time-aware enhanced version of *betweenness centrality* which provide the best measure of how nodes that mediate or bridge the most communication flows. According to this strategy the nodes that act as mediators are patched to *block* the path of a malicious message. However, due to temporal clustering and alternative temporal paths, in most cases, such strategies merely *slow* the malware and does not *stop* it. In other words, a scheme based solely on immunisation of key nodes is not sufficient, instead *quick spreading* of the patch is necessary for most networks. We propose a solution based on local *spreading* of patches through Bluetooth, i.e., exploiting the same mechanism used by the malware itself. The key issue in this approach is to select the right nodes as starting points of the patching process, using *temporal closeness centrality* which ranks nodes by the speed at which they can disseminate a message to all other nodes in the network. We show that this strategy can reduce the cellular network resource consumption and associated costs, achieving at the same time a complete containment of the malware in a limited amount of time.

Table 3 Experimental datasets

	CAMBRIDGE	INFOCOM	MIT
N	18	78	100
Duration (days)	10	5	280
Contacts (avg. per day)	1,927	25,796	231
Scanning rate	30 s	2 min	5 min

4.3 Evaluation

4.3.1 Simulation Setup

To evaluate the time-aware mobile malware containment schemes, three traces of real mobile device contacts carried by humans are used: Bluetooth traces of researchers at the University of Cambridge, Computer Laboratory, as part of an emotion sensing experiment [42]; Bluetooth traces of participants at the 2006 INFOCOM conference [46]; and campus Bluetooth traces of students and staff at MIT [17]. We shall refer to these as CAMBRIDGE, INFOCOM, MIT, respectively. Table 3 describes the characteristics of each set of traces. All three datasets were constructed from mobile device co-location where participants were given Bluetooth enabled mobile devices to carry around. When two devices come into communication range of the Bluetooth radio, the device logs the colocation with the other device. For the CAMBRIDGE dataset, all 10 days are used as part of the evaluation. For the INFOCOM dataset, since devices were not handed out to participants until late afternoon during the first day, only the last 4 days are used. For the MIT dataset, we show results for the first 2 weeks of the Fall semester representing a typical fortnight of activity.

The top N_p devices are chosen according to the calculated temporal betweenness or temporal closeness centrality ranking from the temporal graph $\mathscr{G}_t^w(t_p, t_{max})$, where w is set to the finest window granularity, corresponding to the scanning rate of the devices in each dataset (for example, 30 s windows for CAMBRIDGE); and h is set to 1, since higher values of h lead to similar performance of the containment schemes. The N_m nodes that are initially infected with malicious messages are chosen uniformly randomly. The results are obtained by averaging over 100 runs for each N_p. The static centralities from the static aggregated graph over the time interval $[t_p, t_{max}]$ are also calculated for comparison.

Our evaluation is based on the following assumptions: firstly, when a node receives a patch message, it is immunised for the rest of the simulation (i.e., we assume that the malware does not mutate over time); secondly, there is always a successful file transfer between devices (errors in transmission can be taken into consideration in the assessment of the contention scheme without changing significantly the results of our work, assuming random transmission failures); thirdly, an attacker chooses nodes at random; and finally, we have no knowledge

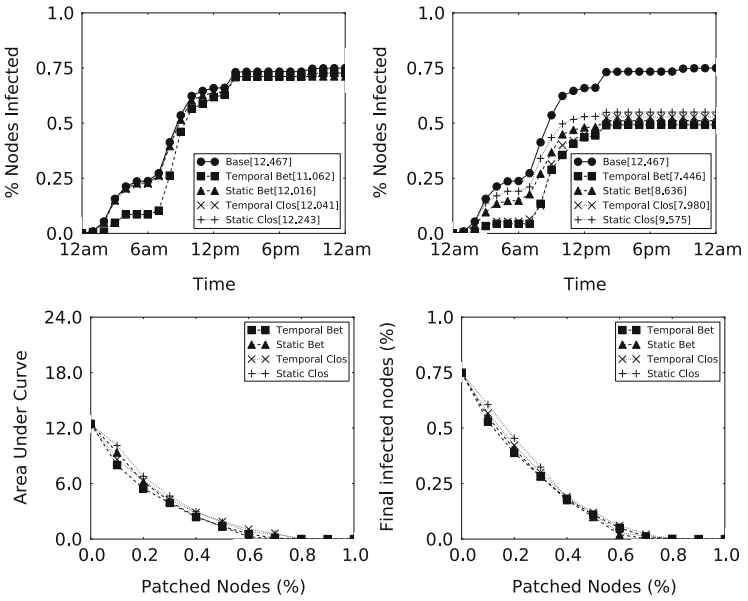

Fig. 7 INFOCOM day 4: Immunising 1 node (*top left*) and 10 source nodes (*top right*). Area under curves shown in the legend. Area (*bottom left*) and final % infected nodes (*bottom right*), as we increase the % nodes immunised (x-axis)

of which devices are compromised (otherwise the best scheme is to patch those devices immediately).

4.3.2 Non-effectiveness of Betweenness Based Patching

Starting from the results of the analysis of the effects time of day has on message spreading, we now evaluate the *best case* scenario for the containment scheme based on patching nodes (without spreading the patch) and we show that this is highly inefficient since it requires a very large number of nodes to be patched via the cellular network to be effective.

Using Day 4 of the INFOCOM trace for this example, a piece of malware is started at the beginning of the day (t_m=12 a.m.) and the device(s) are patched at the same time (t_p=12 a.m.). This is the best case scenario for two reasons: first, the temporal graph in the morning is characterised by low temporal efficiency since there are very few contacts, therefore, the malware spreads slowly; secondly, devices that are immunised immediately have the best chance of blocking malware spreading routes.

Figure 7 shows the ratio of compromised devices across time when the top 1 (top left panel) and top 10 (top right panel) devices are patched after being selected using betweenness and closeness. As we can see, temporal betweenness initially perform better than static betweenness and both temporal and static closeness (quantified

by the difference in the area under each curve, shown in the legend). However, by 7 a.m. we observe a steep rise in the number of compromised devices and by the end of the day, all curves converge to the same point. We also note that *in both cases it is not possible to totally contain the malware, suggesting that more devices need to be patched*. Taking a broader view, Fig. 7 shows the area under the curve (bottom left) and final ratio of nodes infected (bottom right) as we increase the number of patched devices. Clearly, even when the malware is started at the slowest time of day for communication, we still need to patch 80 % of the devices before we can completely stop the malware from spreading; this can be considered an impractically high number of devices to patch. Similar high percentages are also required in the MIT trace with a minimum of 45 % patched nodes.

4.3.3 Effectiveness of Closeness Based Patching (Worst Case Scenario)

Since the betweenness based containment scheme is not effective, we now evaluate the closeness based *spreading* scheme with the aim of disseminating a patch message throughout the network more quickly than a malicious message. For brevity, we do not present results on spreading based on temporal betweenness centrality since it is intuitive that this metric is not designed to quantify the speed of the patching dissemination process and, for this reason, it leads to poorer results. We start our analysis by examining a *worst case* scenario using the CAMBRIDGE dataset: a researcher receives a malicious message on their device in the early hours of Friday morning (t_m = Fri 12 a.m.) and the malicious program replicates itself to any devices it meets during the day. A patch message is started a day later to try and patch all the compromised devices (t_p = Sat 12 a.m.). This can be considered as the a worst case since there are more interactions and hence more opportunity for malware to spread during the day and the patch is delayed until a day later.

Figure 8 shows the spreading rate for the malicious message versus the best (left) and worst device (right) to start the patching message. These results were obtained by running simulations considering every single device as a starting point of the patching process, and then ranking them based on three *performance metrics*:

- The area under the curve (AUC), which captures the behaviour of the infection over time with respect to the number of infected devices[2];
- The peak number of compromised devices (I_{max});
- The time in days necessary to achieve total malware containment (τ).

Since the AUC captures both the I_{max} and τ, the best and worst initial devices that were patched were selected using the AUC. Comparing all three measures, the case related to the selection of the worst device (right panel) is characterised by double AUC (2.62 vs. 1.07); a higher peak in compromised devices I_{max} (68 % vs. 60 %)

[2]The AUC is commonly used in epidemiology and medical trials [21].

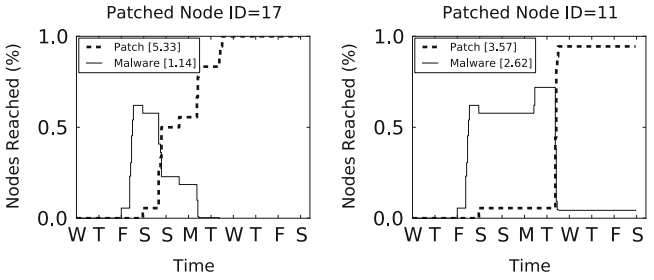

Fig. 8 CAMBRIDGE [t_m=Fri 12 a.m., t_p=Sat 12 a.m.] delivery rate (y-axis) starting a mobile worm from single node with best case patching node (*left*) and worse case patching node (*right*) shown. Area under the curve shown in the legend

and by the fact that it is not possible to fully contain the malware in a finite time τ (∞ vs. 3.3 days).

4.3.4 Sensitivity to Malware Start Time

Thus far we have only considered a single malware start time. We now take a broader view and examine the effects of a malicious message starting at different times. For each dataset the AUC, I_{max} and τ are exhaustively calculated for different malware start times at hourly intervals and increasing patch delays starting from zero (i.e., patch messages start at the same time as malicious messages) to up to 2 days. As a baseline, a naive method of randomly selecting patching nodes is also calculated, averaged over 100 runs. Figure 9 shows for each dataset the performance metrics as a function of the malware start time t_m, averaged over all patch delays. In particular, we note that the AUC and the maximum number of infected nodes I_{max} tend to follow the temporal efficiency (strictly related to human circadian rhythms); however, the total time of containment (τ) remains stable across all start times. These results demonstrate that this time-aware containment scheme is an effective method of quickly containing malware, irrespective of when the malware started. Now analysing the AUC and I_{max}, the temporal centrality curve is consistently lower than static and naive methods. Furthermore, static centrality performs worse than the naive method at some points of time; more specifically:

- For the CAMBRIDGE dataset, during the weekend a static method has a higher peak number of compromised devices (I_{max}) than the naive method, which shows that a static method is not effective at slowing down the malware from spreading.
- For the INFOCOM dataset, again I_{max} is higher than the naive method, during days 2 and 4. In addition, the AUC curve for a static method peaks with temporal efficiency during days 2, 4 and 5: this means that the malware is not contained effectively in these scenarios. Also, the total containment time (τ) is greater than that of the naive method during days 3, 4 and 5. This shows that temporal

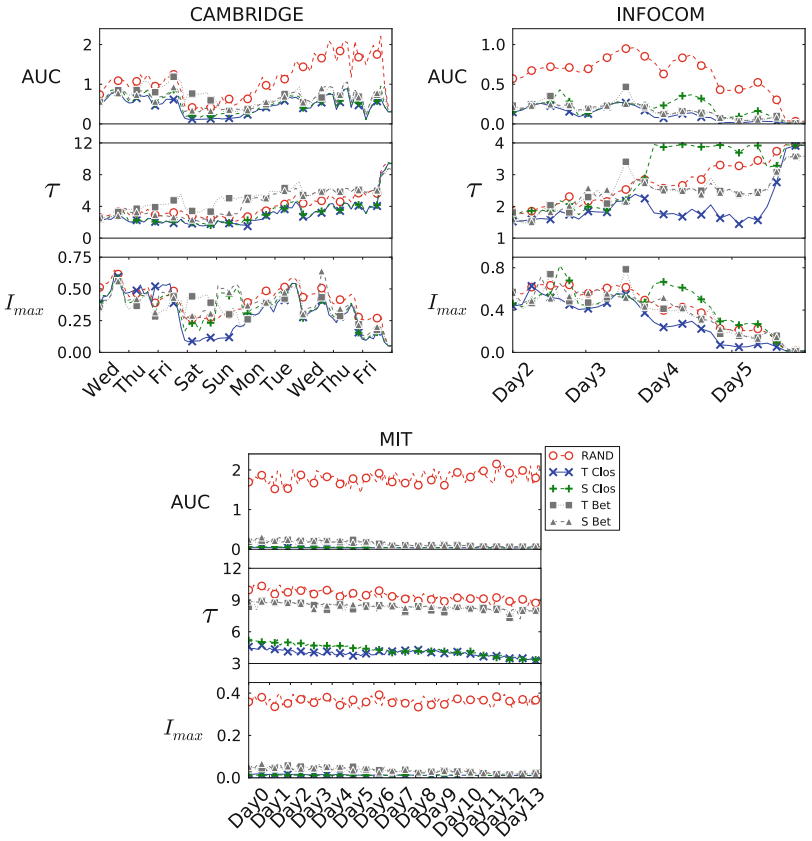

Fig. 9 Performance of temporal, static and naive node selection, across different malware start times (x-axis), averaged over all patch delays

centrality is more consistently effective for identifying the best nodes to start the patching process, compared to both static and naive methods.

- Finally, for the MIT dataset, the naive method performs extremely poorly (with high values of UAC, I_{max} and τ across all malware start times), compared to either a static or temporal method.

4.3.5 Summary of the Findings

This study demonstrates that a temporal analysis of mobile device interactions is better suited to real networks where the topology changes over time. As we have seen, a traditional strategy of patching high betweenness nodes is not effective when temporal topological information is taking into account for both the information dissemination process and also centrality calculations. Instead, we propose a

strategy that can select the best devices to spread the patch in a *competitive* fashion, using the same opportunistic encounters with other devices utilised by the malware itself.

5 Epidemics and Immunisation

A promising application area of the proposed metrics is indeed epidemics modelling in human networks [29, 36]. Currently, the vast majority of the existing models assume an underlying static network [35, 38]. By considering a static network, the *temporal order of appearance* of the links (i.e., the sequence of the contact opportunities) is somehow neglected. This fact might have significant implications on the actual realism of the mathematical model, especially in the case of small populations. Another specific aspect that might have a strong influence on the resulting model is the type of mixing patterns [39] that are present in the population.

The problem of modelling the spreading of infection in a time-varying graph and the definition of vaccination strategies given the information related to the network and epidemic dynamics (and the correlation between the two) are open and challenging research areas at the same time. Temporal centrality metrics can be used for example to prioritise the vaccination of individuals involved in an immunisation program. Moreover, temporal metrics can be used in general to study the evolution of a disease over time by providing quantitative measures of the time scale of its spreading considering the sequence of infected individuals (or geographic areas) over time.

In the recent years, some works have been focussed on the interplay between changing topology and the epidemic process taking place over the network. For example, in [44] Saramäki and Kaski present a model for studying the spreading of an infectious influenza on a dynamic small-world network, by analysing the effect of a dynamic re-wiring process on a Susceptible-Infective-Recovered-style epidemic model, deriving the equations for the epidemic threshold and spreading dynamics. In particular, the authors show how the epidemic saturation time scale varies with the size of the network and the initial conditions. In [34] the authors present the results of different vaccination strategies by simulating the dynamics of sexual disease spreading in empirical contact sequences of individuals. More specifically, the authors analyse the largest outbreak, the average outbreak sizes, and the relative advantages of the different strategies as a function of the infectivity and the duration of the infective state.

In general, social encounter networks, a typical class of time-varying networks, are attracting an increasing attention in the epidemiology community [16]. Researchers have been employing RFID and sensor techniques for extracting contact traces in order to accurately reconstruct patterns of interactions among individuals, also with respect to the duration of the contacts (see for example [6, 10, 43]). These models might also be scaled up to study disease outbreaks in cities [20].

Epidemic models have also been applied to diffusion of ideas, behaviour or lifestyle choices in social networks [9], for example in order to study the spreading of obesity [12] or smoking [13]. Another interesting and open area is the characterisation of the spreading of an epidemic and the simultaneous dissemination of information about it that might modify the behaviour of individuals (i.e., the dynamics of the underlying time-varying networks). A work in this direction is [24].

We believe that the application of the proposed metrics to these fields is indeed very promising and might contribute to increase the accuracy, and therefore, the realism of existing models and to develop new ones allowing researchers to extract valuable information and insights from them. In particular, the centrality metrics presented in this chapter and the accompanying one [40] in this book can be used to identify the key spreaders in order to define effective containment and immunisation strategies. For example, vaccination can be based on priorities assigned to the temporal centralities of the individuals in case of human diseases, whereas, as far as computer viruses are concerned, *white worms* used for patching the systems could be distributed starting from the nodes with the highest temporal closeness centrality.

6 Final Remarks

In this chapter we presented some key applications of temporal metrics to various domains such as centrality analysis in a social network, robustness and epidemiology of computer viruses and diseases. We have shown that temporal graph metrics are able to provide information about the structure of the time-varying networks and the dynamics of processes happening over them that is not possible to extract through the classic static metrics and graph representations. We hope that the case studies presented in this chapter and the discussions of open problems in the field might be considered as a starting point and a source of inspiration for future applications in these and other fields.

Acknowledgements This work was funded in part through EPSRC Project MOLTEN (EP/I017321/1) and the EU LASAGNE Project, Contract No.318132 (STREP).

References

1. Albert, R., Jeong, H., Barabási, A.-L.: Error and attack tolerance of complex networks. Nature **406**(6794), 378–382 (2000)
2. Barrat, A., Barthélemy, M., Vespignani, A.: Dynamical Processes on Complex Networks. Cambridge University Press, Cambridge (2008)
3. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U.: Complex networks: structure and dynamics. Phys. Rep. **424**(4–5), 175–308 (2006)
4. Cabspotting Project. http://cabspotting.org/ (2009). Accessed 27 Feb 2013
5. Callaway, D.S., Newman, M.E.J., Strogatz, S.H., Watts, D.J.: Network robustness and fragility: percolation on random graphs. Phys. Rev. Lett. **85**(25), 5468–5471 (2000)

6. Cattuto, C., Van den Broeck, W., Barrat, A., Colizza, V., Pinton, J.-F., Vespignani, A.: Dynamics of person-to-person interactions from distributed RFID sensor networks. PLoS ONE **5**(7), e11596 (2010)
7. CBS News. Enron traders caught on tape. http://www.cbsnews.com/stories/2004/06/01/eveningnews/main620626.shtml (2004). Accessed 27 February 2013
8. CBS News. Former Enron trader pleads guilty. http://www.cbsnews.com/stories/2004/06/16/national/main623569.shtml (2004). Accessed 27 February 2013
9. Centola, D.: The spread of behavior in an online social network experiment. Science **329**(5996), 1194–1197 (2010)
10. Chaintreau, A., Hui, P., Crowcroft, J., Diot, C., Gass, R., Scott, J.: Impact of human mobility on opportunistic forwarding algorithms. IEEE Trans. Mobile Comput. **6**(6), 606–620 (2007)
11. Chaintreau, A., Le Boudec, J.Y., Ristanovic, N.: The age of gossip: spatial mean field regime. In: Proceedings of SIGMETRICS '09, pp. 109–120. ACM, New York (2009)
12. Christakis, N.A., Fowler, J.H.: The spread of obesity in a large social network over 32 years. N. Engl. J. Med. **357**(4), 370–379 (2007)
13. Christakis, N.A., Fowler, J.H.: The collective dynamics of smoking in a large social network. N. Engl. J. Med. **358**(21), 2249–2258 (2008)
14. Clauset, A., Eagle, N.: Persistence and periodicity in a dynamic proximity network. In: Proceedings of DIMACS Workshop on Computational Methods for Dynamic Interaction Networks, Rutgers University, Piscataway, 24–25 September 2007
15. Crucitti, P., Latora, V., Marchiori, M., Rapisarda, A.: Error and attack tolerance of complex networks. Physica A **340**, 388–394 (2004)
16. Danon, L., House, T.A., Read, J.M., Keeling, M.J.: Social encounter networks: collective properties and disease transmission. J. R. Soc. Interface **9**(76), 11 (2012)
17. Eagle, N., Pentland, A.: Reality mining: sensing complex social systems. Pers. Ubiquit. Comput. **10**(4), 255–268 (2006)
18. Elkind, P., McLean, B.: The Smartest Guys in the Room: The Amazing Rise and Scandalous Fall of Enron. New York, Penguin (2004)
19. Erdős, P., Rényi, A.: On the evolution of random graphs. Publ. Math. Inst. Hungar. Acad. Sci. **5**, 17–61 (1960)
20. Eubank, S., Guclu, H., Anil Kumar, V.S., Marathe, M.V., Srinivasan, A., Toroczkai, Z., Wang, N.: Modelling disease outbreaks in realistic urban social networks. Nature **429**(6988), 180–184 (2004)
21. Evans, C.H. Jr., Ildstad, S.T.: Small Clinical Trials: Issues and Challenges. National Academy of Sciences Press, Washington, DC (2001)
22. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the Internet topology. In: Proceedings of SIGCOMM '99, pp. 251–262. ACM, New York (1999)
23. Federal Energy Regulatory Commission. Addressing the 2000–2001 Western Energy Crisis. http://www.ferc.gov/industries/electric/indus-act/wec.asp (2010). Accessed 27 Feb 2013
24. Funka, S., Gilada, E., Watkinsb, C., Jansena, V.A.A.: The spread of awareness and its impact on epidemic outbreaks. Proc. Natl. Acad. Sci. U.S.A. **106**, 6872–6877 (2009)
25. Gkantsidis, C., Karagiannis, T., Vojnovic, M.: Planet scale software updates. In: Proceedings of SIGCOMM '06. ACM, New York (2006)
26. Holme, P., Kim, B.J., Yoon, C.N., Han, S.K.: Attack vulnerability of complex networks. Phys. Rev. E **65**(5), 056109 (2002)
27. Huberman, B.A., Adamic, L.A.: Growth dynamics of the world-wide web. Nature **401**(6749), 131 (1999)
28. Hypponen, M.: F-secure weblog: the grand opening! http://www.f-secure.com/weblog/archives/00000568.html (2005). Accessed 27 Feb 2013
29. Keeling, M.J., Rohani, P.: Modeling Infectious Diseases in Human and Animals. Princeton University Press, Princeton (2007)
30. Kempe, D., Kleinberg, J., Kumar, A.: Connectivity and inference problems for temporal networks. J. Comput. Syst. Sci **64**(4), 820–842 (2002)
31. Kendall, M.G.: A new measure of rank correlation. Biometrika **30**(1–2), 81–93 (1938)

32. Kostakos, V.: Temporal graphs. Physica A **388**(6), 1007–1023 (2009)
33. Leavitt, N.: Mobile phones: the next frontier for hackers? Computer **38**(4), 20–23 (2005)
34. Lee, S., Rocha, L.E.C., Liljeros, F., Holme, P.: Exploiting temporal network structures of human interaction to effectively immunize populations. PLoS ONE **7**(5), e36439 (2012)
35. Liljeros, F., Edling, C.R., Amaral, L.A.N.: Sexual networks: implications for the transmission of sexually transmitted infections. Microb. Infect. **5**(2), 189–196 (2003)
36. May, R.M.: Network structure and the biology of populations. Trends Ecol. Evol. **21**(7), 394–399 (2006)
37. Medina, A., Gursun, G., Basu, P., Matta, I.: On the universal generation of mobility models. In: Proceedings of IEEE/ACM MASCOTS '10, Miami Beach, FL, August 2010
38. Newman, M.E.J.: Networks: An Introduction. Oxford University Press, Oxford (2010)
39. Newman, M.E.J.: Mixing patterns in networks. Phys. Rev. E **67**, 026126 (2003)
40. Nicosia, V., Tang, J., Mascolo, C., Musolesi, M., Russo, G., Latora, V.: Graph metrics for temporal networks. In: Saramäki, J., Holme, P. (eds.) Temporal Networks. Springer, Berlin (2013)
41. Piorkowski, M., Sarafijanovic-Djukic, N., Grossglauser, M.: CRAWDAD data set epfl/mobility (v. 2009-02-24). Downloaded from http://crawdad.cs.dartmouth.edu/epfl/mobility (2009). Accessed 27 Feb 2013
42. Rachuri, K.K., Musolesi, M., Mascolo, C., Rentfrow, P.J., Longworth, C., Aucinas, A.: EmotionSense: a mobile phones based adaptive platform for experimental social psychology research. In: Proceedings of UbiComp '10. ACM, New York (2010)
43. Salathé, M., Kazandjieva, M., Lee, J.W., Levis, P., Feldman, M.W., Jones, J.H.: A high-resolution human contact network for infectious disease transmission. In: Proceedings of the National Academy of Sciences of the USA, Washington, vol. 107, pp. 22020–22025, December 2010
44. Saramäki, J., Kaski, K.: Modelling development of epidemics with dynamic small-world networks. J. Theor. Biol. **234**(3), 413–421 (2005)
45. Scellato, S., Leontiadis, I., Mascolo, C., Basu, P., Zafer, M.: Evaluating temporal robustness of mobile networks. IEEE Trans. Mobile Comput. **12**(1), 105–117 (2013). http://doi.ieeecomputersociety.org/10.1109/TMC.2011.248
46. Scott, J., Gass, R., Crowcroft, J., Hui, P., Diot, C., Chaintreau, A.: CRAWDAD data set cambridge/haggle (v. 2009-05-29). Downloaded from http://crawdad.cs.dartmouth.edu/cambridge/haggle (2009). Accessed 27 Feb 2013
47. Shetty, J., Adibi, J.: Discovering important nodes through graph entropy the case of Enron email database. In: Proceedings of the 3rd International Workshop on Link Discovery, pp. 74–81. ACM, Chicago (2005)
48. Tang, J., Musolesi, M., Mascolo, C., Latora, V.: Temporal distance metrics for social network analysis. In: Proceedings of WOSN '09, Barcelona, August 2009
49. Tang, J., Musolesi, M., Mascolo, C., Latora, V., Nicosia, V.: Analysing information flows and key mediators through temporal centrality metrics. In: Proceedings of ACM SNS '10, Paris, April 2010
50. Tang, J., Mascolo, C., Musolesi, M., Latora, V.: Exploiting temporal complex network metrics in mobile malware containment. In: Proceedings of the 12th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WOWMOM'11), Lucca, June 2011
51. Tang, J., Kim, H., Mascolo, C., Musolesi, M.: STOP: socio-temporal opportunistic patching of short range mobile malware. In: Proceedings of the 13th IEEE Symposium on a World of Wireless, Mobile and Multimedia Networks (WOWMOM'12), San Francisco, June 2012
52. Van Ruitenbeek, E., Courtney, T., Sanders, W.H., Stevens, F.: Quantifying the effectiveness of mobile phone virus response mechanisms. In: Proceedings of DNS '07, 2007
53. Virus description: Bluetooth-worm:symos/cabir. http://www.f-secure.com/v-descs/cabir.shtml (2004). Accessed 27 Feb 2013
54. Wang, P., Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.-L.: Understanding the spreading patterns of mobile phone viruses. Science **324**(5930), 1071–1076 (2009)

55. Washington Post. Enron fraud trial ends in 5 convictions. http://www.washingtonpost.com/wp-dyn/articles/A23034-2004Nov3.html (2004). Accessed 27 Feb 2013
56. Wasserman, S., Faust, K.: Social Networks Analysis. Cambridge University Press, Cambridge (1994)
57. Zhu, Z., Cao, G., Zhu, S., Ranjan, S., Nucci, A.: A social network based patching scheme for worm containment in cellular networks. In: Proceedings of INFOCOM '09, IEEE, Rio de Janeiro, April 2009
58. Zyba, G., Voelker, G.M., Liljenstam, M., Mehes, A., Johansson, P.: Defending mobile phones from proximity malware. In: Proceedings of INFOCOM '09, IEEE, Rio de Janeiro, April 2009