Wrocław University of Science and Technology

Marcin Kulisiewicz

Wrocław University of Science and Technology
Faculty of Computer Science and Management
Department of Computational Intelligence

Event-driven Temporal Social Networks

Dissertation supervisor: prof. dr hab. inż. Przemysław Kazienko

Auxiliary supervisor: dr inż. Radosław Michalski

A thesis submitted for the degree of
Doctor of Philosophy
Wrocław, 2019

Abstract

The Universe undergoes a constant change. This affects every subject of our world, and every method used to describe real-world systems should take its dynamics into account. One of the most powerful concepts representing a complex system is a network. In this work, I study temporal social networks that extend static networks by another degree of freedom - time. With this additional dimension, temporal networks are able to model systems' changes capturing individual interactions between network nodes. In particular, I propose a set of new entropy-based measures that are capable of quantifying temporal networks dynamics, alongside algorithms for their iterative and parallel computation. Next, by applying those to real-world cases, I demonstrate that human beings are much more selective in their communication over time. In addition, I show that communication within social communities that exist in examined real-world social networks have different entropy, which can be potentially utilized for group recognition. Furthermore, to provide a meaningful method of comparing temporal social networks of different sizes, I define the normalization method for entropy-based measures.

In literature, there is a lack of temporal network models that are able to model humans' cognitive processes. Making it another contribution to the field of temporal network modeling, I fill this gap with CogSNet - a new event-based and cognition-driven temporal network model. It provides reinforcement by discrete events and continuous forgetting mechanism that is capable of modeling human perception and cognition processes of mutual interactions. The CogSNet model significantly outperforms all other reference models commonly used in the literature that was proved on the real-world data set. The model is capable to take into account various interactions as well as the heterogeneous nature of human behavior.

Both entropy-based measures and the CogSNet model provide different but complementary information about the dynamics of social networks. This has been shown by means of experimental studies.

The following thesis contribution is meant to extend our knowledge about dynamic systems we are part of. And by doing so, to provide to the field new tools to deal with the complexity of reality.

Dedication

I dedicate this work to my beloved wife.

Contents

List of Tables

List of Figures

Notation

Symbols

- G - graph
- V - set of nodes
- n - number of nodes in the network, size of set V
- v_i - i-th node
- E - set of edges
- e_j - j-th edge
- t_{ij} - time of event between nodes v_i and v_j
- c_{ij} - number of events processed for pair of nodes v_i and v_j
- w_{ij} - weight of relation between nodes v_i and v_j
- L - trace life time
- SC - set of successions
- ev_{ijk} - an event between nodes v_i and v_j at time t_k
- e_{ij} - edge between nodes v_i and v_j
- S_N - N-th order entropy

- S_N^M - maximum N-th order entropy
- d_{ij} - distance between nodes v_i and v_j
- τ_{ij} - duration/latency between nodes v_i and v_j
- w - number of windows in static graph sequence
- h - network horizon

Acronyms

- CogSNet - Cognition-driven Social Network
- ES - event sequence

Chapter 1

Introduction

Panta Rhei - everything flows. This philosophical statement made by Heraclitus of Ephesus refers to constant changes of everything around us. Slower or faster, things transform from one form or state to another. Those changes are not independent - objects, things, people - they interact with each other in many ways. Those objects can be perceived as a large network that evolves all time. This idea is called the temporal network. In opposite to a static network approach, where objects are connected with links that do not disappear and emerge, temporal networks are trying to model some dynamics in that matter. From the beginning of the concept, temporal networks were built with the most granular particles - events. An event is an interaction of two objects in time and can be understood in many ways: face-to-face meetings, email, handshake, an electromagnetic signal, airplane flight, touch, etc.

Typically, network science utilizes static networks to model systems, usually making an assumption or simplification that connections between nodes are permanent. However, many systems are time-varying in their nature. In social networks, friendship networks are commonly presented as static ones. In fact, they rely on temporal interactions like face-to-face meetings or communication via a number of mediums. A network of airports - also usually presented as a static network - could not exist without flights between them - temporal events. The question arises here if there is some benefit of using a temporal approach instead of a static one? Authors in [71] show that in fact there is an advantage in terms of controllability of real-world networks.

Similar findings in other aspects probably wait to be discovered. However, in some cases, it may happen that the static approach would be more appropriate.

For over a decade researchers have been intensively investigating the domain of temporal networks looking for an understanding of nature of this abstract representation of reality. They came up with many solutions and applications, however, in this area, there are still much more questions than answers. First, despite many proposed methods for representations of temporal networks, still, not a single one has been pointed out as the right one as it is for static networks. This might be because modeled systems require so different approaches that it is impossible to pick one method for all of them. Next, a lot of measures are invented, however, there is still a lack of versatile measure for temporal network dynamics, especially due to the number of different representations. Researchers see the difference between human behavior in social media, in face-to-face communication, in other media, etc. [132], but those differences should be investigated even more. In the last couple of years, temporal networks are in the range of interest of neurologists which are looking for applications of some social measures method to brain activity [117], simultaneously pointing out that much more work needs to be done to find appropriate utility. In [38] Holme, in his broad survey about temporal networks, indicates that most of the work focuses mostly on adapting static network methods and measure into temporal networks but this might not be the right way. The nature of temporal networks may require fresh notions and a different approach than only adding time dimension into old methods. Likewise, difficulties appear in visualizations methods of temporal networks - they are still far from convenient. Some of them are sufficient only for small networks, others allow to explore temporal networks interactively [1], but there is a lack of good approach, especially for large networks. After over ten years of active research temporal networks are still a growing area in many disciplines (see Fig. 1.1). Even that most of the papers involve a more standard (static) approach, temporal networks are an important part of network science. The future will show if temporal networks will become a major representation of real-world objects or processes.

Yet, it is still hard to confidently declare whether using temporal networks for modeling reality is a better way than a static approach. For sure, in some cases, it is reasonable to properly incorporate a time dimension, especially in dynamic systems. Yet, to my best knowledge, there

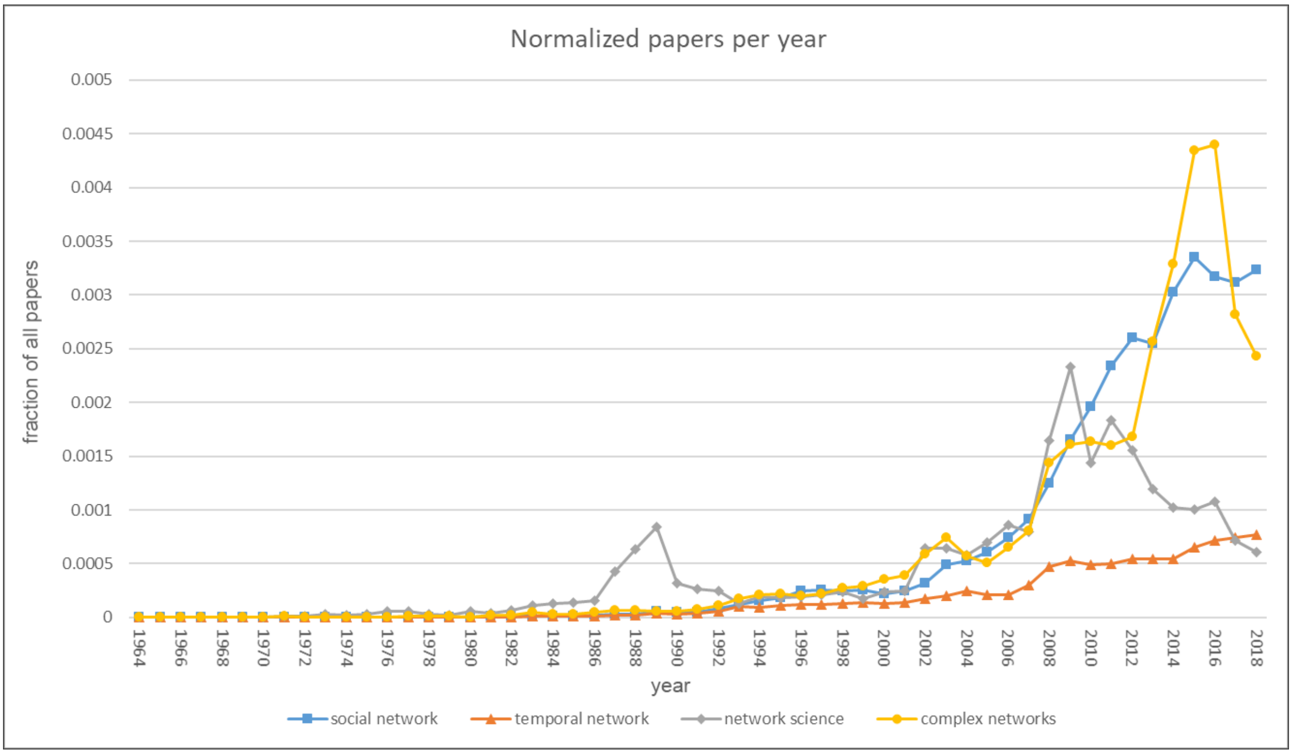


Figure 1.1: Normalized number of papers with given keyword per year, according to Scopus database. Number of documents is normalized with overall number of scientific publications. Used keywords: "social network", "temporal network", "network science" and "complex networks". Accessed on 2018-12-13

is no data that can confirm that statement. The question that might be asked here is - if temporal networks are some upgrade for static networks or just another way to model systems. The answer is not obvious and needs to take into account a broad spectrum of modeled systems. On one hand, we have some topology given by entities - people, companies, computers, etc. - and relations between them respectively, e.g., friendships, contracts, physical connections and on the other hand there are some interactions between them like messages, transactions, signals. Topological structure and temporal interactions need to conform. Latapy et al. in [67] introduces an elegant formalism that tries to deal with both structural and temporal aspects of temporal networks. Comparing temporal networks given with different representations is also still problematic since it requires transformations from one representations to another, which not always can be made without temporal information lost. Authors proposed a number of measurements - a number of nodes, neighborhoods, compactness, cliques, clustering coefficient and even more - that imitate those from static networks with an additional dimension of time. Yet, even with a number of defined measurements, researchers are not able to unambiguously state if there is some border beyond which the temporal network's approach is better than static networks. In the past literature, there is also still a lack of methods that are able to model human cognition. Another open question is what kind of temporal network structure speeds up spreading [42].

1.1 Dissertation motivation and objectives

In the literature, to my best knowledge, there is a lack of measures for event sequence, that are able to measure temporal networks dynamics. Such measures could be useful not only for a better description of temporal networks and a better understanding of its nature but also for more practical applications such as analysis of social groups in social temporal networks or detection of network behavior changes. Another aspect that is missing in the literature is a temporal network model that would allow simulation of the human cognition process. This process is dynamic and continuous, yet bases on discrete events such as meetings or messages.

My primary goal in this dissertation is to improve the state of knowledge in the area of temporal networks by introducing a new set of measures for one of the most basic forms of temporal networks - an event sequence. Proposed measures will be an important contribution to computer science as a method to describe and analyze data collected by many ICT systems. Also, proposed computation methods for measures will be designed to be efficient to handle big amount of data.

The second goal of this work is to propose a new temporal network model that is able to simulate human behavior on the cognition level. It would be another tool for data analysis, especially related to computational social science. To preserve the usefulness of the proposed model, the parameters of the model should be optimized.

Besides contribution to computer science, hopefully, my work will have a noticeable impact on other domains. Useful methods for temporal networks analysis would potentially have an impact on sociology, biology, medicine and economy.

1.2 Dissertation structure

This dissertation is organized as follows. First, the state-of-the-art in the area of temporal networks is presented: common representations, measures, processes, application, and challenges with open questions. In the next chapter, a new temporal network model with memory is introduced. The following chapter contains an introduction of a new set of entropy-based measures for event-described, along with algorithms for its iterative and parallel computations. The next chapter consists of experiments and analysis of proposed temporal network description methods to examine their complementarities. In the final chapter, the summary of the dissertation is presented, alongside ideas for future work.

Chapter 2

Temporal networks: state of the art

In this chapter state of the knowledge in the area of temporal networks is presented alongside challenges and open questions. Firstly, basic definitions and the most popular representations of temporal networks are described, followed by a review of developed measurements. The next section focuses on processes that can be modeled with temporal networks, followed by a section about temporal network applications. The chapter concludes with challenges and open questions in the area.

2.1 Graph definition

To begin with a proper presentation of temporal networks, here, some basic definitions from the network science are presented. One of the basic concepts is a network defined with a simple graph as follows:

Definition 2.1 *A Graph is a tuple $G = (V, E)$, where $V = \{v_1, ..., v_n\}, n \in \mathbb{N}_+$ is the set of vertices and $E = \{e_1, ..., e_k\}, k \in \mathbb{N}_+$ is the finite set of edges between them. An edge is a tuple $e_{ij} = (v_i, v_j, w_{ij})$, where $v_i \in V, v_j \in V, v_i \neq v_j, w \in [0, 1]$.*

This definition is valid both for directed and undirected graph. However, for undirected graph $e_{ij} = e_{ji}$, while for directed graph $e_{ij} \neq e_{ji}$. Weight w_{ij} states for how strong nodes v_i and v_j are connected. For unweighted graph weight is always equal 1. A directed, weighted graph can be also interpreted as Social Network (SN) if $V = \{v_1, ..., v_n\}$ is considered as a set of individuals and E is a set of edges representing directed social relationship between nodes v_i and v_j.

2.2 Representations

In this section, representations of temporal networks introduced in the literature, in the past, is presented. It is important to note that so far there is no established representation that is suitable for all applications. It is one of the reasons that, on one hand, temporal networks seem to be useful for many applications, but on the other hand they are not easy to use. Among many proposed representations there are few that are the most common in literature, e.g., event sequence, interval graph, snapshots [41].

2.2.1 Event sequence

The first representation is one of the basic representations of a temporal network. Event sequence (ES) or contact sequence is a list of contacts with an additional time attribute. It can be given as time-stamp (for instant events or those which duration time is negligible, e.g., text messages, emails, etc.) or with an interval (for events that takes some finite time, e.g., face-to-face meetings, phone calls, flights). While it is a representation that preserves all information about temporal network there is still a lack of convenient tools for graphical visualization. Most of empirical temporal network data [38] are given as event sequence as it is the most natural way to register temporal activity. In Fig. 2.1b an event sequence is presented on a timeline of nodes.

In this thesis set of interactions between network individuals are formally defined as Event Sequence (ES):

Definition 2.2 *An Event Sequence is a tuple* $ES = (V^e, EV)$*, where* $V^e = \{v_1^e, ..., v_{n^e}^e\}, n^e \in \mathbb{N}_+$ *is the set of social entities and* $EV = \{ev_1, ..., ev_{k^{ev}}\}, k^{ev} \in \mathbb{N}_+$ *is the finite set of events* *(contacts) between them. Each event* ev_{ijkl} *is a tuple:* $ev_{ijkl} = (v^e_i, v^e_j, t^e_k, id^e_l)$*, where* $v^e_i, v^e_j \in V^e, v^e_i \neq v^e_j$ *and* $t^e_k \in T^e$*. Here,* T^e *represents a discrete time dimension consisting of timestamps* $T^e = \{t_1, ..., t_{m^t}\}, m^t \in \mathbb{N}_+$ *in which a particular event occured or is assigned to. The set* $ID = \{id_1, ..., id_{n^{id}}\}$ *contains unique event identifiers* $id_{e^{id}} \in ID, n^{id} \in \mathbb{N}_+$*. The set of nodes* V^e *cannot possess any insolated nodes, i.e.* $\forall_i(v^e_i \in V^e \Leftrightarrow \exists_{jkl}(e_{ijkl} \in EV \vee e_{jikl} \in EV))$*.*

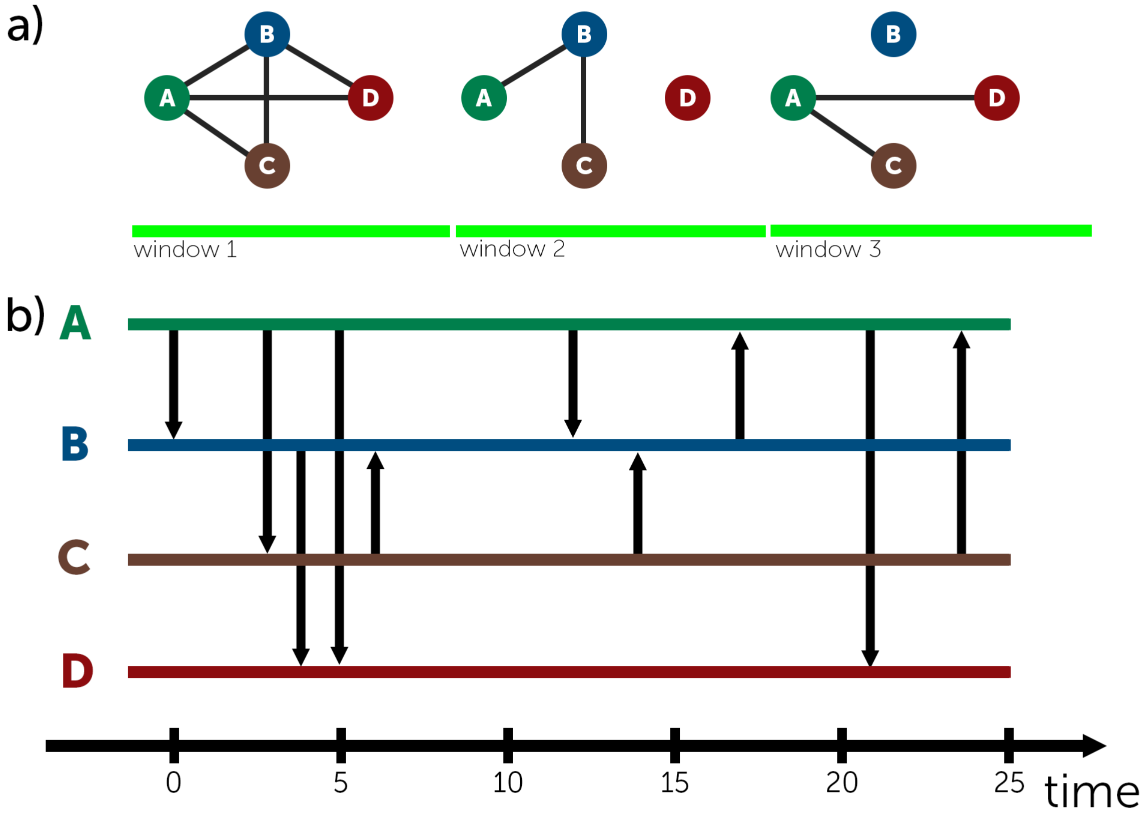


Figure 2.1: a) Event sequence mapped into a sequence of graphs - each window size is equal 9 time units. b) Event sequence presented as a timeline of nodes. In this case, contacts are considered as directed. Indirect timeline would not have arrows indicating direction.

Another way to present an event sequence is a timeline of contacts, which is less a representation and more way to visualize or understand a temporal network given with event sequence. In Fig.2.2 timeline of contacts is presented.

Both presented ways to visualize event sequence are rather tough to analyze, especially for bigger graphs.

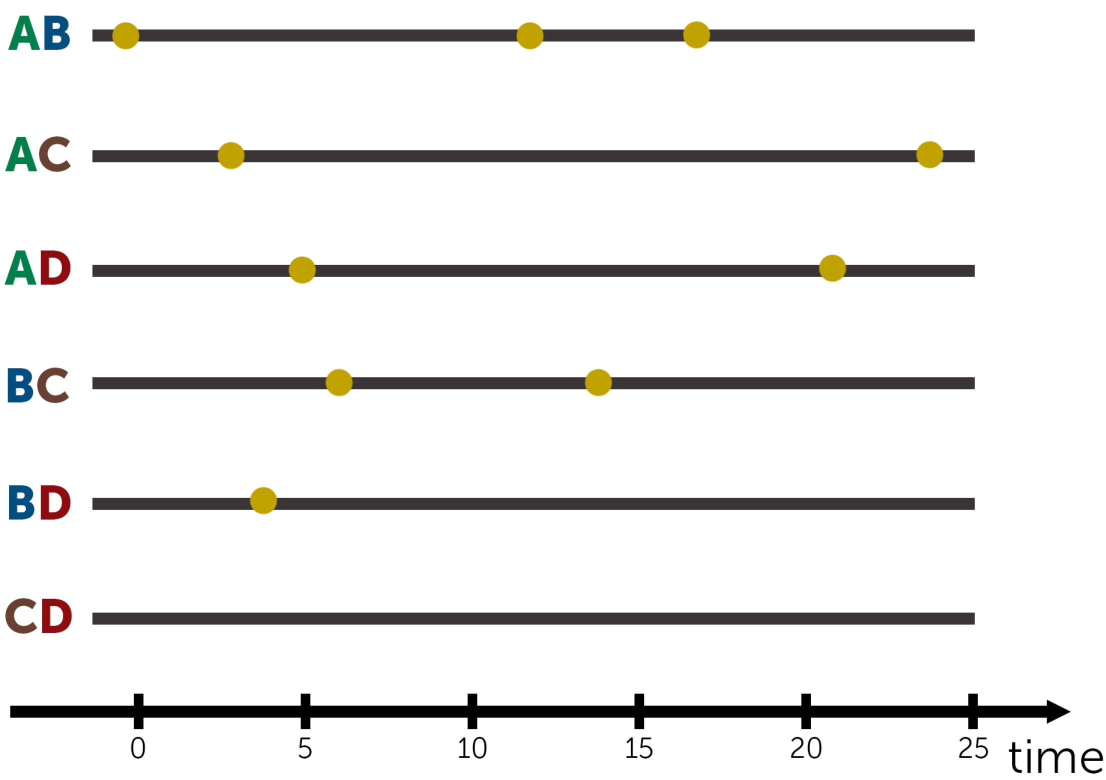


Figure 2.2: Event sequence presented as a timeline of contacts. Yellow points indicate when contact emerges. Here a version of indirected communication is presented - directed version would contain also lines for reverse communication.

2.2.2 Weighted graph

Looking for simplification of an event sequence the most intuitive approach seems to be an aggregation of a sequence into a static graph. Such an approach, however, loses all information about temporal aspects of network and frequency of communication. A kind of intermediate solution is to use a weighted graph where nodes represent participants of communication and each edge states, whether there was a contact between a pair of nodes and weight of an edge, represents frequency of communication - given as a number of contacts or a fraction relative to overall number of contacts in the network [65]. This approach may be partially good to simulate some properties of the network, however more precise prediction requires information about the order of contacts which are missing in this representation. To conclude, this approach favors the usage of a great number of methods for static networks while misses a lot of temporal information. A weighted graph is also called a Social Network [122] and is formally defined as:

Definition 2.3 *A Social Network SN on Event Sequence* $ES = (V^e, EV)$ *is a tuple* $SN = (V, E)$ *where* $V = v_1, ..., v_n, n \in \mathbb{N}_+$ *is the set of vertices and* $E = e_1, ..., e_{k^e}, k \in \mathbb{N}_+$ *is the set of edges between them. Each vertex* $v_i \in V$ *represents an individual* v_i^e *from Event Sequence and each edge* e_{ij} *corresponds to the directed social relationship from* v_i *to* v_j*, such that* $E = \{(v_i, v_j, w_{ij}) : v_i \in V, v_j \in V, v_i = v_i^e, v_j = v_j^e \text{ and } \forall_{ij}(\exists_{kl} ev_{ijkl} \in EV \Leftrightarrow e_{ij} \in E), w_{ij} \in [0,1]\}$ *Here, value* $w_{ij} = \frac{n_{ij}^e}{n_i^e}$ *denotes the importance (weight, strength) of the relationship between individuals, such that* n_{ij}^e *is the number of events* ev_{ijkl} *from* v_i^e *to* v_j^e *in ES (regardless* k, l *and* n_i^e *is the number of all events initiated by* v_i^e *(outgoing from).*

2.2.3 Static graph sequences and time-window graphs

Temporal network can be represented also with time-ordered sequence of static graphs [89, 77] or multilayer graph in which layers are representation of time periods [11, 57, 68], see Fig. 2.1a. In [18] authors proposed a useful framework to study multilayer networks including temporal networks. This representation seems to be a powerful way for temporal network analysis because it allows for the application of some algorithms from static graphs - vastly developed in the past in literature - along with visualization methods. However, there are some drawbacks of this solution and all of them come to the resolution of used periods. If the resolution is low - such representation is usually denoted as a time-window graph - each layer in the multilayer network will bring a lot of useful topological information while losing some of the information about temporal aspects. In opposite, if the resolution is high - such structure is usually called graph sequence - a lot of temporal information will be preserved, while layers may bring insignificant topological information due to a very sparse structure. In an extreme situation in each layer, there will be only one edge. Depending on an examined system it may be reasonable to use this representation while studying, e.g., the spread of influence or disease spreading it is hard to decide which of aspects is more important - temporal or topological - and graph sequence seems to be a rather poor choice. Yet another problem is how to choose boundaries of periods as well as a starting point. Consider the following example: studying communication patterns of people we can look at aggregated networks of communication during: day, week, month, etc. The different aggregation will provide as different information e.g. daily aggregated network may reveal some patterns during the week while look at the entire month some patterns about seasons or holidays [61]. Yet another idea is proposed in [25], where authors propose aggregation of network edges into some limited space. This solution is valuable for large graph processing, due to limited available memory. The proposed approach at a given point of time take recent n number of nodes of edge stream and aggregate them into a static network.

2.2.4 Reachability graphs

This representation based on a concept of the time-respecting path which states that there exist a direct edge between nodes v_i and v_j or time-ordered set of edges between them. In the reachability graph, an edge between nodes v_i and v_j means that in the temporal network a time-respecting path between those nodes exists. This concept was proposed in [82] and later independently revisited in [124]. Further research was done in [16, 41, 89, 38]. The idea of showing the capability of a network to pass information between nodes could very useful, however, it may lead to very dense (full or near full) graphs very fast [40, 37] which loses a lot of information in both topological and temporal aspects. Another drawback of this representation is that the existence of edge states for at least one time-respecting path between nodes and do not contain information about multiple paths.

2.2.5 Concurrency graphs

This representation is based on an idea of concurrent partnerships in sexual network proposed in [60]. In [35] author proposed an approach of event-sequence aggregation into a static network using concurrency of contacts: 1) in a certain period of time, 2) before and after a certain period of time and 3) in an entire network but using an exponential threshold to get rid of low weight edges. Of course, this kind of aggregation causes some loss of temporal information, however, it returns an analytically convenient static network.

2.2.6 Difference graph

Snapshots of a temporal network were used in [85] to compute a difference in links between consecutive time steps. Authors show that applying Minimum Vertex Cover can capture key properties on the structure of changes in a dynamic network [85].

2.2.7 Transmission graph

Most of the representations previously presented aggregate a temporal network into a static graph with nodes matching individuals in the contact sequence. In [100], authors propose different approach - nodes in the static network represents pairs of individuals in a contact sequence while edges represent if there is a possible transition of disease i.e. if a pair of edges share a vertex in the same time.

2.2.8 Representations summary

Representations presented above are not the only ones, however the most common to be found in the literature. In general, it is hard to judge which representations are good and which are not. It strictly depends on an application and authors of those representations show usefulness in particular cases. Holme in [38] divides temporal network representations into *lossy* and *lossless* which is another way to compare them. However, it is hard to put a line between those two groups. Another aspect that needs to be taken into account is the availability of methods to analyze a network given in a particular representation. In Fig. 2.3 an attempt to compare temporal networks representations is presented using Gartner Magic Quadrant. Please note, that particular positions of points in the quadrant are estimated and have no numerical meaning. Some of the representations are especially hard to put e.g. static networks sequence, which can have very high resolution, but then it will be very hard to analyze. Grey dashed line represents a kind of trade-off that researchers using particular representations must do in many cases - if a network shows high resolution of temporal information, it will not be easy to analyze. Some representations, like the transmission graph, are located above the line which indicates that they are able to preserve some aspects of temporal information while making it easier to analyze.

2.3 Measures of temporal networks

In the previous section, a number of representations of temporal networks are presented. Many of them utilize a static network concept, thus some measures for static networks can be used to describe their topology. However, static measures are lacking information about the temporal aspects of the network. Hence, many of them seem not to be an appropriate way to describe temporal networks just by a straightforward application. The additional degree of freedom of temporal networks - time - demands many measures to be rethought and redeveloped. Also, different representations require different measurements. Researchers made many attempts to this challenge of developing a number of measures for temporal networks. Some of those measures focus more on topology and some of them more on temporal aspects. There are also some attempts that try to capture both aspects at once.

The existence of so many temporal networks representations and measures raises the question - if temporal networks are somehow comparable among themselves? Intuitively it can be assumed that in order to compare two networks with different representations, they need to be mapped into a common form. However, it can not be done in many cases and if even so, such a mapping may cause a temporal information loss. Assuming that all temporal networks are created based on a sequence of events, probably the proper way is to compare sequences. However, there are not many measures that can be applied to event sequence and intuitively describe it.

In this section, the most popular concepts in temporal network measures are presented.

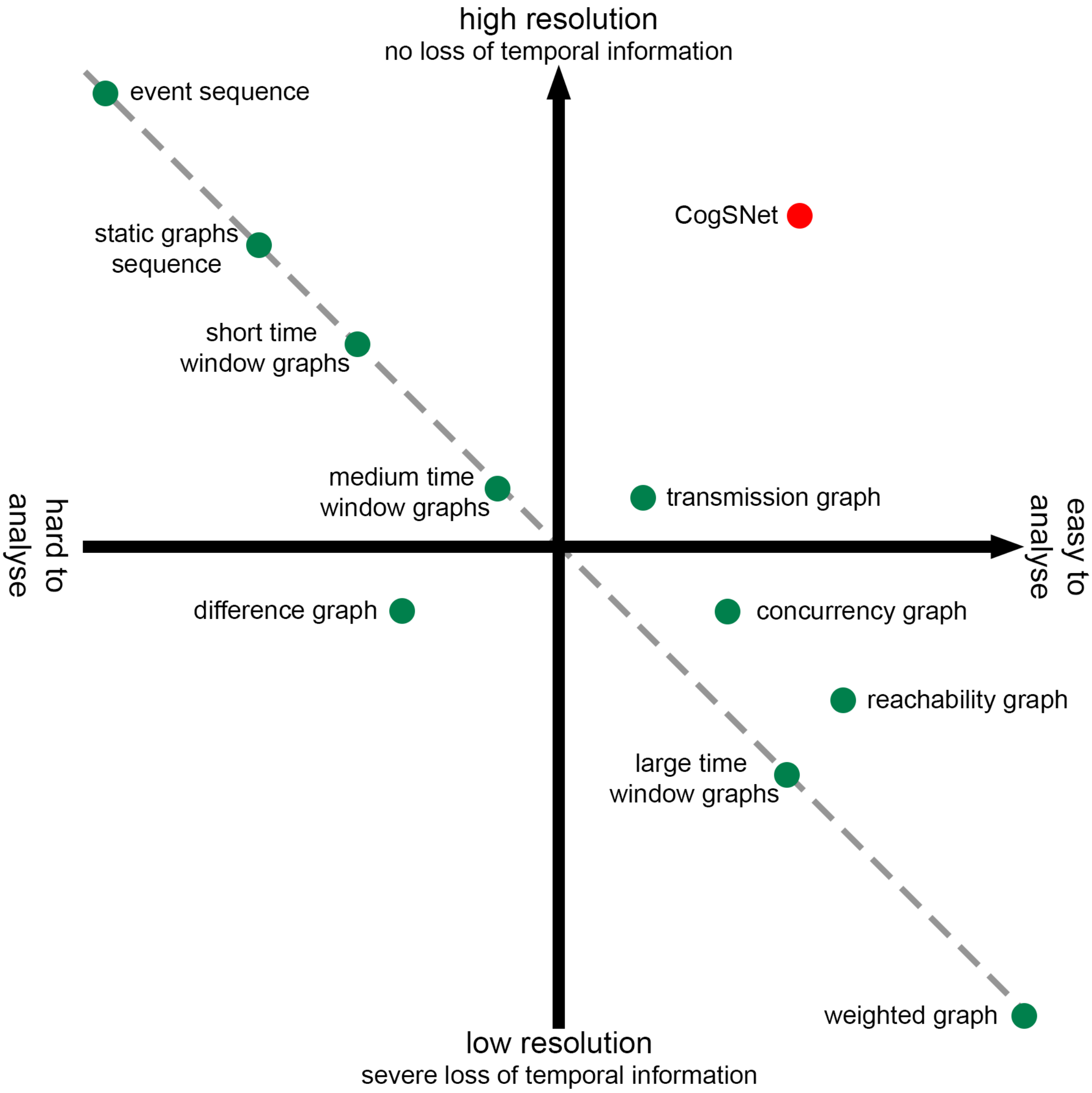


Figure 2.3: Gartner Magic Quadrant for temporal network representations. Vertical axis refers to the resolution of the network or in other words how much temporal information is preserved. The horizontal axis refers to the complexity of the analysis. It may be also understood as the availability of methods to analyze.

2.3.1 Time-respecting path

One of the basic ways to describe a relation between two nodes in the network is by measuring the distance between them. In static networks distance between nodes, v_i and v_j is a minimum number of jumps over existing edges to get from node v_i to v_j and the set of visited edges is called path. In temporal networks, there is a similar concept. Authors use different names: time-respecting path [54, 39], time-ordered path [92], journey [128, 26] or non-decreasing path [16] but the idea remains the same. In simple words, a time-respecting path usually is a sequence of temporal edges (events) connecting two nodes with non-decreasing time. If we consider a system with events with non-negligible duration time, a time-respecting path will consist of a sequence of events that each event: a) begins after the end of the previous event (except the first event) or b) overlaps with the previous event. The sense of those approaches depends on the examined system, e.g., for studying flight connections 'a)' approach fits better, while for an epidemic process on face-to-face meeting 'b)' approach seems to be the right approach.

In some cases, it is reasonable to introduce an additional requirement for a time-respecting path. In [92] authors point out that it is a good idea to introduce some limit for waiting time between consecutive events on a path, after which path will be cut off. The notion behind this comes from empirical data investigation. In air transportation data it is reasonable that people would not like to wait more than a day for connecting flight and in most cases, it would be a couple of hours. If the waiting time is longer - than the path would not be considered as a viable option. Similarly, considering phone call data, if we want to examine information spreading through the network we should take into account that people may forget to pass information after some time. The cutoff threshold will vary depending on the examined system and data - it should be different for air transportation, information spreading in email communication and epidemic spreading (even different viruses would require a specific threshold). Such diversification of events source is taken into account in a new temporal network model - CogSNet - that is described in Chapter 3.

Without regard to the specific approach of time-respecting path definition, the most obvious component of the path is its length, or in other words, the distance between two nodes.

In static networks, a path length is a number of edges between the first and last node, alternatively sum of weights on edges, which allows for a simple definition of the shortest path. In a temporal network, it would be very useful to have a similar property defined, however, due to the time aspect, it is not that obvious. In the literature, the nomenclature in this subject is not uniform and sometimes might be a little confusing. The ambiguity comes from the fact, that some authors consider path length as a time that is required to get from one node to another, while other authors consider it more like in static networks - a number of edges between two considered nodes. In this work, I would like to distinguish **distance** that measures a number of events between nodes and **duration** (or latency) that measures time. Pan and Saramaki in [92] define *temporal distance* from node v_i to v_j as the shortest time it takes to reach v_j from v_i starting at time t along temporal paths:

$$\tau_{i,j}(t) = (t' - t) + \delta t \tag{2.1}$$

where *t'* is time of the first event from v_i in the shortest path and δt is the path duration.

In [114] authors define *shortest temporal distance* $d_{ij}(h, t^{min}, t^{max})$ as a minimum number of time windows that are required to connect two nodes through available events, assuming that in a single window there can be a finite number of hops denoted as h, where distance is given for node i and j in finite time window from t^{min} to t^{max}, and h is the maximum number of hops. Authors in [7] define length of temporal path as number of hops between nodes and similar notion in [6] is called *distance*.

Whatever the approach to the definition of path length is, it is crucial to use the appropriate one to a specific application. Once defined distance or duration is a base for many further measures.

2.3.2 Centrality measures

A property in static networks that describes how a node is placed in the network in relation to other nodes is called centrality. With the defined distance between nodes in a temporal network, it is pretty straightforward to apply it to centrality measures for a temporal network, just by replacing paths with time-respecting paths and its distance or duration. Newman in [86] defines *closeness centrality* $C^C(i)$ for node v_i which promotes nodes that are closer to other nodes in the static network:

$$C_i^C = \frac{n}{\sum_j d_{i,j}} \tag{2.2}$$

where n is a number of nodes in the network and d_{ij} is the distance of node v_i to other nodes. In other words, this is an invert value of average distance of node i to other nodes. To apply this notion to temporal networks we can simply apply one of previous definitions of temporal distance defined *temporal closeness centrality* [115] for node v_i in a sequence of static graphs:

$$C_i^C(h) = \frac{1}{w * (n-1)} \sum_{j \neq i \in V} d_{ij}(h) \tag{2.3}$$

where n is a number of nodes in the network, V is a set of nodes, w is a number of static graphs in the sequence or time windows, h is network horizon - a maximum number of contacts that information can pass in single window and $d_{ij}(h)$ is the temporal distance between nodes i and j for a given horizon. This idea was also developed by Kim and Anderson in [56].

Another way to describe the centrality of a node is betweenness, which is a fraction of shortest paths that contain given node overall shortest paths in the network [122]:

$$C_i^B = \frac{\sum_{i \neq j \neq k} |s_{ijk}|}{\sum_{i \neq j \neq k} |s_{jk}|} \tag{2.4}$$

where $|s_{ijk}|$ is a number of shortest paths that contain node v_i and $|s_{jk}|$ is a number of all shortest paths in the network. When applying this notion to temporal networks it is obvious that temporal shortest path should be applied. However, before that, we need to take into account that at different points of time since the network is dynamic over time, node betweenness may vary. In [115] authors proposed a definition of temporal betweenness centrality for a node in a given point of time:

$$C_i^B(t) = \frac{1}{(n-1)(n-2)} \sum_{\substack{j \in V \\ j \neq i}} \sum_{\substack{k \in V \\ k \neq i \\ k \neq j}} \frac{|s_{ijk}(t)|}{|s_{jk}(h)|} \tag{2.5}$$

where n is a number of nodes in the network, V is a set of vertices, $|s_{jk}(h)|$ is a number of all shortest path with given horizon h and $|s_{ijk}(t)|$ is a number of temporal shortest paths from node v_j to v_k such as v_i is holding message from v_j at time t that will be passed to v_k in the future. One of the applications of betweenness is to identify important nodes in the network in terms of passing information - leaders, commanders, hubs, etc. In temporal networks it is also important how long a node holds information - the longer it holds, the more important it is. Authors of the above definition proposed also overall temporal betweenness centrality for a node in temporal network $G(t^{min}, t^{max})$, that is average betweenness over time windows in the network:

$$C_i^B = \frac{1}{w} \sum_{t=1}^{w} C_i^B((t \times w) + t^{min}) \tag{2.6}$$

where w is a number of windows in the network.

Last but not least group of centrality measures are those which measure how central are nodes in terms of random diffusion in the network. In static networks we have PageRank [122] which promotes nodes that are linked by other important nodes, eigenvector centrality [88] which finds nodes that are connected to the most influential nodes or Katz centrality [53] which extends notion of eigenvector centrality in way that a node is important if it is linked from other important nodes or if it is highly linked. In temporal networks, Holme and Sarämaki in [41] proposed an algorithm of a generalization of the eigenvector centrality:

Algorithm
1. Start with a centrality value 1 at each vertex.
2. At every contact between vertices v_i and v_j, let the C^E values after the contact at timestep t be

$$C_i^E(t+1) = \zeta C_i^E(t) + (1-\zeta) C_j^E(t)$$

and

$$C_j^E(t+1) = \zeta C_j^E(t) + (1-\zeta) C_i^E(t)$$

where ζ is a rate of how much centrality is passed from one node to another.

Larger ζ promotes recent contacts to have more influence on final centrality. This notion was further developed in [30] to solve some problems in previous approach.

2.3.3 Connectivity

One of the fundamental concepts in networks is the knowledge of whether nodes are connected with a path or not. So-called connectivity gives a minimum number of nodes that need to be removed from the network to separate remaining nodes into components. In directed static networks there are two definitions: 1) weak connectivity - which states that there are directed paths between all nodes at least in one direction and 2) strong connectivity - which states that there are directed between all nodes in both directions. We can apply this idea to temporal networks using time-respecting paths. In [90] authors proposed definitions of: 1) strong connectedness, where two nodes v_i and v_j of a time-varying graph are strongly connected if v_i is temporally connected to v_j and also v_j is temporally connected to v_i and 2) weak connectedness, where two nodes v_i and v_j of a time-varying graph are weakly connected if v_i is temporally connected to v_j and also v_j is temporally connected to v_i in the underlying undirected time-varying graph. Authors went a step further and proposed also a definition of a strongly connected component as a set of nodes of a time-varying graph G is a temporal strongly connected component of G if each node of the set is strongly connected to all the other nodes in the set.

2.3.4 Diameter and efficiency

Yet another measurement that is important in networks is a diameter, which states how far are nodes in the network from each other. Diameter in static networks is the longest shortest path and a simple transition into temporal networks would be to take a length of the longest time-respecting path or its latency - depending on the aspect of interest. Those concepts of topological diameter and temporal diameter of a temporal network was introduced in [128] and further investigated in [15]. A related concept to diameter - efficiency - is average over inverse path lengths of all paths. For temporal networks, temporal efficiency for a pair of nodes was proposed in [114]:

$$E_{ij}(t^{min}, t^{max}, h) = \frac{1}{d_{ij}(t^{min}, t^{max}, h)} \tag{2.7}$$

which is an inverse of distance between nodes in given time window (t^{min}, t^{max}) and horizon h, and global temporal efficiency:

$$E(t^{min}, t^{max}, h) = \frac{1}{n(n-1)} \sum_{ij} E_{ij}(t^{min}, t^{max}, h) \tag{2.8}$$

which is average temporal efficiency over all pairs of nodes. Authors in [114] also proposed a concept of local efficiency for a particular node and average in the entire network.

2.3.5 Inter-event times

In temporal networks based on contact events, we can take a closer look at the time between recorded events. There are a couple approaches to characterize inter-event times in order to be able to replicate a network conserving contact dynamics. In [32], in addition to established *Aggregate Pairs* approach [31, 14, 50, 28, 133], authors proposed two new approaches *Aggregate Nodes* and *Any Contact*. *Aggregate Pairs* is an approach where we analyze distribution of times between contacts in each pair of nodes separately. *Aggregate Nodes* approach focuses on the distribution of times between events for particular nodes. *Any Contact* is a holistic approach which aggregates all events and analyzes inter-event time for all nodes together. While those approaches aggregated obtained distribution of edges and nodes respectively for convenient network characterization, authors of work in [95] and [96] have shown that it is not always meaningful. Another way to characterize inter-event time was proposed in [29] is *burstiness parameter*. This parameter describes how a distribution of inter-event times $\mathrm{P}(\tau)$ - with mean σ^τ and standard deviation m^τ - differs from Poison distribution and it is given with formula:

$$B = \frac{(\sigma^\tau/m^\tau - 1)}{(\sigma^\tau/m^\tau + 1)} = \frac{(\sigma^\tau - m^\tau)}{(\sigma^\tau + m^\tau)} \tag{2.9}$$

However, authors pointed out that this definition is meaningful when both the mean and the standard deviation of $\mathrm{P}(\tau)$ exist, which is always the case for real-world finite signals [29]. The conclusion derived from available publications can be that using the smallest blocks building a temporal network is a good way to characterize the network, however, it has to be done with caution.

2.3.6 Patterns and motifs

Yet another way to examine networks is to look at persistent patterns and repeats of motifs. In [64] authors proposed an approach to find subgraphs in time windows that are more frequent than others. Such a set of windows that contain a significantly frequent (above some threshold) subgraph is called support set $S(G')$. Another way to define persistent patterns in temporal networks was proposed in [17], where the authors proposed measurement called *adjacency correlation* which results with Pearson correlation between two consecutive time windows in terms of edges appearance. They define this measure as:

$$\gamma_j = \frac{\sum_{i \in N(j)} A_{i,j}^{(x)} A_{i,j}^{(y)}}{\sqrt{(\sum_{i \in N(j)} A_{i,j}^{(x)})(\sum_{i \in N(j)} A_{i,j}^{(y)})}} \tag{2.10}$$

where $A_{i,j}$ is an adjacency matrix, x and y denote consecutive time windows and *N(j)* is a set of nodes that appears in at least one of examined two windows.

Motifs are a similar concept to persistent patterns, yet showing another aspect of the network. In static network motifs (also called graphlets) are, roughly speaking, small subgraphs (usually contains 3 or 4 nodes) that appear in a network and by counting them we can say something about network topology [48, 8, 80, 129]. In a temporal network, there are couple proposed approaches. As a simple transition from static networks authors in [12] find motifs in snapshots of a temporal network. In [93] authors proposed a similar approach which extends the previous one a little bit, by counting a given motif M if all of its edges appear in a given time period, rather in strictly defined time windows. Also, in [131] authors proposed a similar approach based not on time windows but on time difference between two edges with the same node. Another approach is proposed in [73] where authors are using Markov chains for motifs detection and Bayesian framework to optimize a number of motifs. Authors in [58] try to bring a deeper understanding of the financial market by analysis of daily patterns of the financial network. They find that *"The 'social' dynamics of financial interactions are highly stable and* *little affected by external shocks such as the occurrence of the global financial crisis"*. Temporal motifs are used in [75] to show the existence of homophily in the social network of students.

2.4 Processes and temporal networks

One of the utilities of temporal networks is the capability of modeling and simulating dynamic processes that take place on a given topology. There are several known directions where temporal networks seem to be the right choice in terms of precise modeling of a dynamic process. Modeling a spread of epidemic is one of the obvious choice since a lot of work has been done in static networks over last nearly four decades [33, 106, 112, 87, 103, 126, 107] and many authors pointed out that network dynamic should be taken into account as it has vital impact on disease process outcome [83, 49, 120, 21, 24, 121]. Quite a number of papers were written on this subject in temporal networks. Authors examined the impact of network topology on speed of epidemic outbreaks working on aspects like: burstiness [52, 81, 102, 65, 40] when it speed up or slow down disease spreading, basic reproductive number R_0 for temporal networks [102, 101] or vaccination problem [69, 36] which can help in real-world application of temporal networks. Another group of processes in networks is about information spreading and by the information we can understand the news, rumors, opinions, novelty, etc. Of course, depending on the character of spread information processes would act differently but there are several models that are capable to simulate some of the behavior. The most known models in static networks are: voter model [72], linear threshold model [55], independent cascades model [123]. Those methods were adapted by researchers for temporal networks. Cascade model for temporal networks was developed in [51], Voter model was adopted to temporal networks in [43]. Epidemic processes and information spreading are considered as similar in many ways and not without reasons, authors in [84] shows that the spread of news - temporal and dynamic process - can be modeled with SIR model known from static networks. Yet another type of process that happens in temporal networks is an evolution of the network itself. How networks grow, if the pace of growth is constant, if nodes disappear - answer for those questions may lead someone to describe such process more formally. Authors in [4] try to describe the growth of online social network - in this case, Twitter - using average node degree and try to fit a Leskovec model [70] to predict Twitter network evolution.

2.5 Applications of temporal networks

Temporal networks provide tools to examine a vast number of systems. Data collected over the years, that we can find in open datasets library, consists of information about communications, messages, physical proximity, biological processes, vehicles movement, social media activities, computer networks traffic, etc. Each of those areas can be investigated in terms of various aspects like: disease spreading, robustness to failures, traffic control, effectiveness, social phenomena. In this chapter brief review of many applications will be presented.

Human communication is one of the most researched areas. People communicate using different kinds of media starting from emails [22, 119, 78], short text messages [132, 113] which is *one-to-one* type of communication which takes place in time but has negligible duration time. This kind of data is usually used for research information spreading but also to find some patterns of communication to discover the specific structure of networks like hierarchy or groups. In such research, authors use only metadata, the content of these messages is rarely used. Even more natural human communication is simple meeting - face-to-face contact [19, 13, 44, 110, 111] - and phone calls [91, 92] are similar in the way that contact not only is between two people in time but also has some duration. Such additional information sometimes is used as a parameter of methods for the research mentioned before in this chapter. Vast popularity of social media provides enormous amount of data of *one-to-many* information broadcast on blogs [63], Twitter [46, 133], Facebook [23] etc. Such data is utilized, among others, to examined and predict information spreading in terms of how far it can reach in the network like for example in [45]. Apart of mentioned applications, there are also authors who work on epidemic spreading in human network [83, 44, 110] as well as in animal - livestock - network [49, 120]. By the analysis of contact data in a hospital, authors in [76] try to describe how infection can propagate through the network.

Temporal networks find its application also in others - maybe less intuitive - areas like biology, medicine or ecology, which shows its interdisciplinary utility. Authors in many papers find profits using a temporal network to examine protein interactions [10, 116]. Other biological interactions were examined in [27] using time-varying network data. Authors in [117] examined brain activity using temporal network methods and measures to ,, apply these measures to a resting-state fMRI dataset to illustrate their utility". Some authors in [94, 98] used temporal networks in research on relations between species considering, e.g., food webs. A broad analysis of different layers of social networks shows a strong correlation between the communication network and the strength of friendship [75].

In this section, a wide spectrum of temporal networks application is only briefly presented. However, we can observe that the temporal network could be a very universal tool for dynamic systems analysis.

Chapter 3

CogSNet - Cognition-driven Temporal Social Network

In this chapter new temporal network model with forgetting mechanism is presented. This model is a result of international cooperation with Prof. Bolesław K. Szymański from Rensselaer Polytechnic Institute (Troy, NY, US), Prof. Christian Lebiere from Carnegie Mellon University (Pittsburgh, PA, US) and Prof. Omar Lizardo from University of California (Los Angeles, CA, USA). Original results are presented in [79].

3.1 Motivation

A lot of temporal network models have been proposed in the literature so far, however many of them are very general. Some applications of networks require network models to be tuned for special purposes. The human cognition process can be considered as such a case. Current literature lacks a temporal network model that would be able to simulate the cognition process of a human being. Furthermore, proposed models are usually discrete in the sense that state of the network changes only in discrete points, such as events. In the case of sparse events, a lack of dynamics between events is a serious drawback. Here, the goal of this research is to propose a temporal network model that would be able to model the human cognition process, preserving the continuous nature of its dynamics.

3.2 General idea

One of the applications of temporal networks is to model social networks. Those are usually used to model human behavior. People are involved in many social activities, but some of them are remembered well and some fade away. Meetings, calls, messages - all of those leave a trace in human memory [5]. If the event is repeated - the trace in memory will be stronger. Furthermore, relationships between people much depend on memory. Here an assumption is made, that a larger number of interactions between people indicate a stronger relationship between them. Hence, considering human perception, memory has to be taken into account. Among many available temporal network models, it is hard to find one that considers that aspect too. Here a new temporal network model is proposed, with a mechanism that simulates human perception in terms of forgetting. Forgetting usually is a rather continuous process over time, apart from rare situations, e.g., accidents or diseases. The proposed mechanism uses a continuous function to model that process. While forgetting can be modeled with any continuous monotonic function - linear, power, logarithmic, etc., - in this dissertation two are used: power and exponential, as the most accurate based on cognitive studies [47, 99]. Proposed functions can be tuned with parameters to best fit to given data. Idea behind this model - called CogSNet (Cognition-driven Social Network) - is to take event sequence of social activities, e.g., face-to-face meetings, phone calls, text messages and to give each event some reinforcement peak μ (value from 0 to 1) that will leave trace in memory, which is also considered here as a weight of relationship. This trace will decay over time according to used forgetting function with parameter λ, which indicates the pace of forgetting. Due to the fact, that both used functions reach 0 in the infinity, we assume that the trace of a social event will disappear if its level reaches some forgetting threshold θ. For more convenient and intuitive use of the proposed model, we combine all three parameters by introducing *trace lifetime* L which is time after memory will decay from μ to a level of θ. Exact formulas for this transformation and forgetting mechanism are to be found in further sections. In Fig. 3.1 an example of one relation

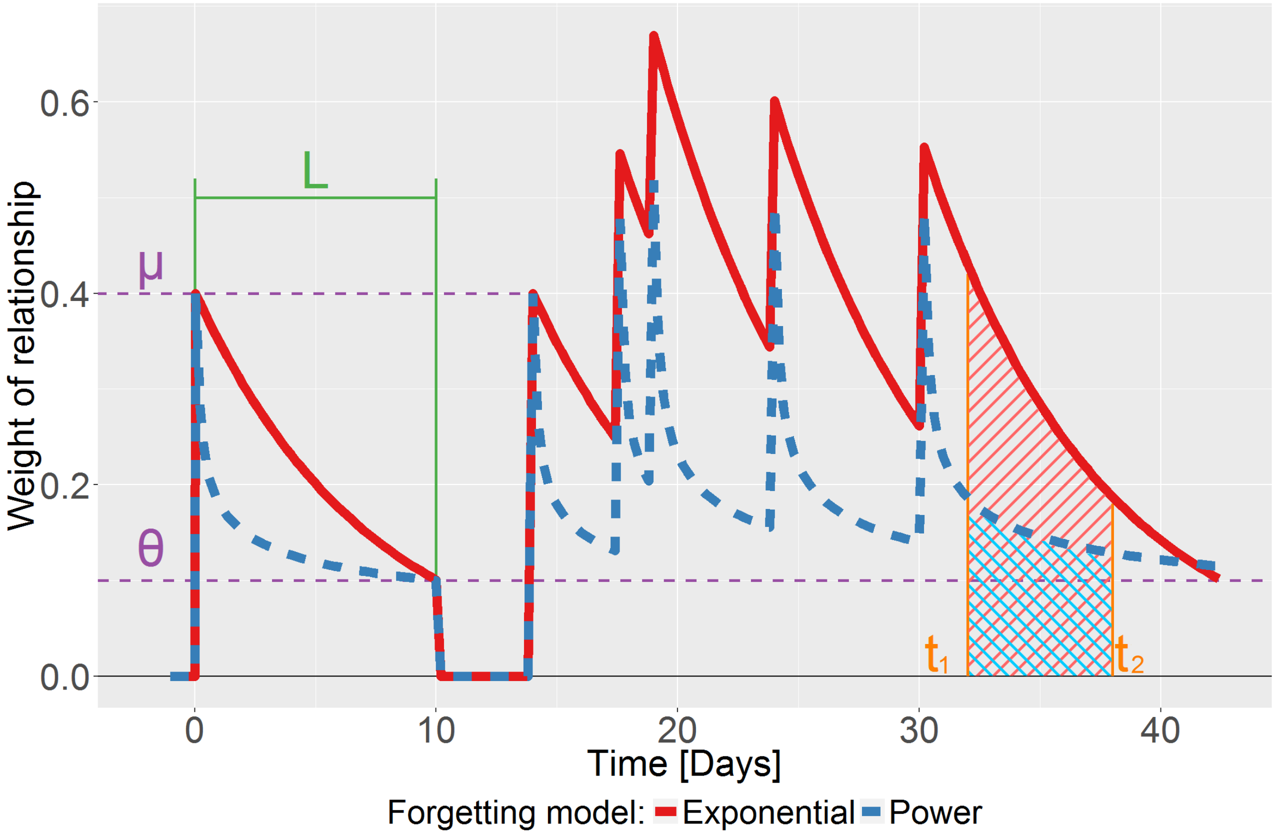


Figure 3.1: Dynamics of a relationship in CogSNet with exponential and power forgetting functions and with parameters set to $\mu = 0.4$, $\theta = 0.1$, and L=10 days.

modeled with CogSNet is presented.

3.3 Research question

At first, an exact research question needs to be asked. As was pointed out earlier in this chapter, memory is a significant compound of human cognition. Thus, taking it into account should provide a model that better fit real-world scenario than other known methods. The goal of this research is to answer the question: whether real-world social networks can be modeled using CogSNet more accurately than with other common temporal network models?

3.4 The CogSNet model

The human brain records events as they arrive, but only a small fraction of the incoming information is stored in long-term memory due to its limited capacity. The forgetting mechanism dictates that the chance to recall a given event decreases gradually as we move forward from the time of first exposure [125]. In some sense, it is similar to graph-streaming [25] and feed-based social media network cascades [109] scenarios where incoming events are ordered by their arrival time but only some of them are kept.

Accordingly, the CogSNet model uses a *forgetting function* f to account for the decreasing probability of keeping aging of memory traces over time. Forgetting is thus a monotonically non-increasing function of time with $f(0) = 1$ and $f(t) \geq 0$ for all $t > 0$. It is defined by two parameters: *reinforcement peak* $0 < \mu \leq 1$ and *forgetting threshold* $0 < \theta < \mu$.

Here, the Definition 2.1 of the social network is used. The model evolves in discrete steps as follows. For each pair of nodes (v_i, v_j), the system maintains two variables: t_{ij}, which represent the time of the most recent event for this pair of nodes, and c_{ij} which holds the count of events processed for this pair of nodes. Initially, both t_{ij} and c_{ij} are set to 0, as are the weights of all edges, i.e., for all pairs of nodes (v_i, v_j), $w_{ij}(0) = 0$.

When an event happens at a time t in the modeled social network, it is processed in chronological order by the model. First, the weight of the corresponding edge is updated according to the following equation:

$$w_{ij}(t) = \begin{cases} \mu_{ijc_{ij}+1}, & \text{if } w_{ij}(t_{ij})f(t - t_{ij}) < \theta, \\ \mu_{ijc_{ij}+1} + w_{ij}(t_{ij})f(t - t_{ij})(1 - \mu_{ijc_{ij}+1}), & \text{otherwise,} \end{cases} \tag{3.1}$$

where: μ_{ijk} is the value of reinforcement peak that results from the k^{th} event that impacts the edge (v_i, v_j).

Here, the value of reinforcement peak μ_{ijk} depends on the engagement and emotions invoked by the event that is either directly or indirectly related to the edge (v_i, v_j). An example of an event indirectly related to this edge could be node v_i talking about node v_j or any situation that reminds node v_i about node v_j. The values of μ can be individualized to node v_i perception of relation with node v_j at event k. The values of μ may also be dependent on event types: $\mu_{ijk} \in \{\mu_1, \mu_2, \mu_3, \mu_4, \ldots\}$, e.g., $\mu_1 = 0.5$ for emails, $\mu_2 = 0.55$ for phone calls, $\mu_3 = 0.8$ for meetings, $\mu_4 = 0.9$ for joint collaboration in projects, etc.

Finally, the processing of the current event updates both variables associated with the updated edge (v_i, v_j) as follows: $t_{ij} = t, c_{ij} = c_{ij} + 1$.

At any time t of the model evolution, the user can obtain the value of the weight of an arbitrary edge (v_i, v_j) by computing the following equation.

$$w_{ij}(t) = \begin{cases} 0, & \text{if } w_{ij}(t_{ij})f(t - t_{ij}) < \theta, \\ w_{ij}(t_{ij})f(t - t_{ij}), & \text{otherwise.} \end{cases} \tag{3.2}$$

The weight $w_{ij}(t)$ of an edge e_{ij} between two nodes at any user selected time t is computed as follows. Once all the relevant events up to time t are processed, we simply set $w_{ij}(t) = w_{ij}(t_{ij}) * f(t - t_{ij})$. If the result is less than the forgetting threshold θ, $w_{ij}(t)$ is reduced to zero and the edge is no longer considered. A threshold is needed with forgetting functions, such as power and exponential forgetting, that are positive for non-negative arguments. Otherwise, an edge would get the positive weight at creation and would always stay positive, i.e., all created memory traces would never cease to exist. The reinforcement peak μ defines the impact of an event on the weight of the edge relevant to this event. This peak is a global model parameter here. In principle, the peak can be adjusted according to the event or node type to allow for individualized event perception.

In general, the forgetting function $f(\Delta t)$ over time interval Δt can be of any type (linear, power, logarithmic, etc.), but here, informed by work in the cognitive psychology of memory [47] we evaluate only two such functions: the exponential function f^{exp}, and the power function f^{pow} defined as:

$$f^{exp}(\Delta t) = e^{-\lambda \Delta t}. \tag{3.3}$$

$$f^{pow}(\Delta t) = \max(1, \Delta t)^{-\lambda}. \tag{3.4}$$

where λ denotes the forgetting intensity; typically $\lambda \in [0, 1]$. The use of max in the power function ensures that perception of events that happened less than a time unit ago is not changed by forgetting. The time unit in which the forgetting function is expressed scales the values of the parameters. In experiments one hour is used as the time unit.

To simplify optimal parameters search as well as to provide meaningful interpretation of parameters, we aggregate all three parameters into the *trace life time* L defining the time after which an unreinforced memory trace is forgotten, i.e., too hardly recalled. In the model, L is the time over which the forgetting function reduces the edge weight from μ to θ causing the edge to be removed, cf. Fig. 3.1. For the exponential forgetting function, equation (3.3), trace life time L^{exp} is:

$$L^{exp} = \frac{1}{\lambda} \ln\left(\frac{\mu}{\theta}\right), \tag{3.5}$$

while for the power function, the formula is:

$$L^{pow} = \left(\frac{\mu}{\theta}\right)^{\frac{1}{\lambda}}. \tag{3.6}$$

3.5 Reference models

In Fig. 3.2, we compare the CogSNet model with the previous proposals for representing temporal network dynamics. The most common approach for representing social network dynamics is to use interaction sequences [41]. Under this method, each event is time-stamped and the weights are added to the edges connecting nodes involved in this event, cf. Fig. 3.2(I) and 3.2(IIa). Moreover, a given edge is active (exists) only at a given time t. This is the most granular approach as it is capable of tracking all the events occurring between nodes while preserving the temporal order of events.

In contrast, a static binary network representation, as shown in Fig. 3.2(IIb), aggregates all events by making all edges time-independent (cf. Section 2.2.2). Consequently, an edge

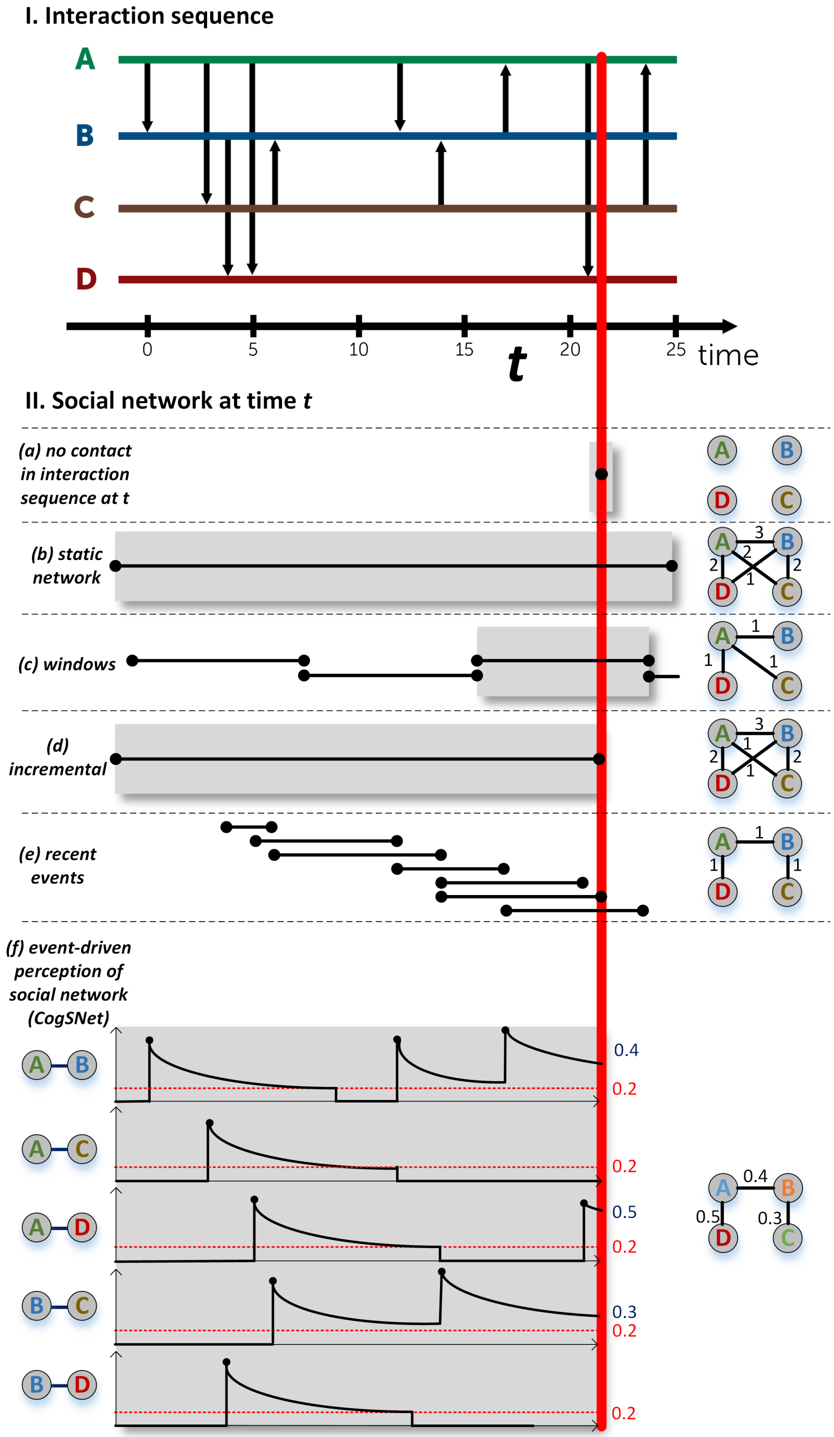


Figure 3.2: Various approaches to modeling dynamic networks and edge weighting (relation strengths) for the 4-node network at a given time t: **(I)** and **(IIa)** interaction sequence; **(IIb)** static (time-aggregated) network; **(IIc)** sliding windows; **(IId)** incremental network, all events from the beginning time t_0 to the current time t are considered; the frequency of interaction in the period constitutes the frequency-based reference FQ; **(IIe)** network based on $n = 3$ recent events, used for recency-based reference RC; **(IIf)** Cognition-based Social Network model, CogSNet, introduced here.

exists between a pair of nodes an event between these nodes occurred at least once in the whole observed period [34]. Such an edge representation throws away information on the temporal the ordering of events, making it impossible to study dynamic processes in static networks. The incremental network solution accumulates events only up to the current time t of analysis. The classical approach, used early in [59, 9, 82], views a dynamic network as a series of time-ordered sequences of static graphs, see Fig. 3.2IIc (cf. Section 2.2.3). More recently, this method was applied to modeling network and community evolution [127, 104]. The drawback of this approach is that it does not preserve the ordering of interactions within time slices. Applying a simple frequency-based aggregation creates a frequency-based, FQ, metric, cf. Fig. 3.2IId (cf. Section 2.2.2). Taking into account only a given number of the most recent events leads to the recency-based, RC, model, cf. Fig. 3.2(IIe). Both of these models are used here as baseline models. Fig. 3.2(IIf) shows an example of a dynamic social network generated using the CogSNet model. All other social network models presented in Fig. 3.2(IIa-e) can also be represented by CogSNet by setting appropriate parameter combinations to achieve, as needed, no decay, instant decay, and so forth. In this way CogSNet can be thought of as a universal generative dynamic model for temporal social networks, encompassing previous approaches as special cases.

3.6 Validation and results

In order to verify the proposed model and baseline models, real-world dataset - NetSense [113] - is used to validate empirically if CogSNet model is able to reproduce the dynamics of social relations in a dynamic social network.

3.6.1 Experimental setup

NetSense dataset consists of two parts. The first includes the time-stamps and duration/length of phone calls and text messages collected for each student participating in the study. Each student phone device recorded all connections/messages, including those to the phones of people

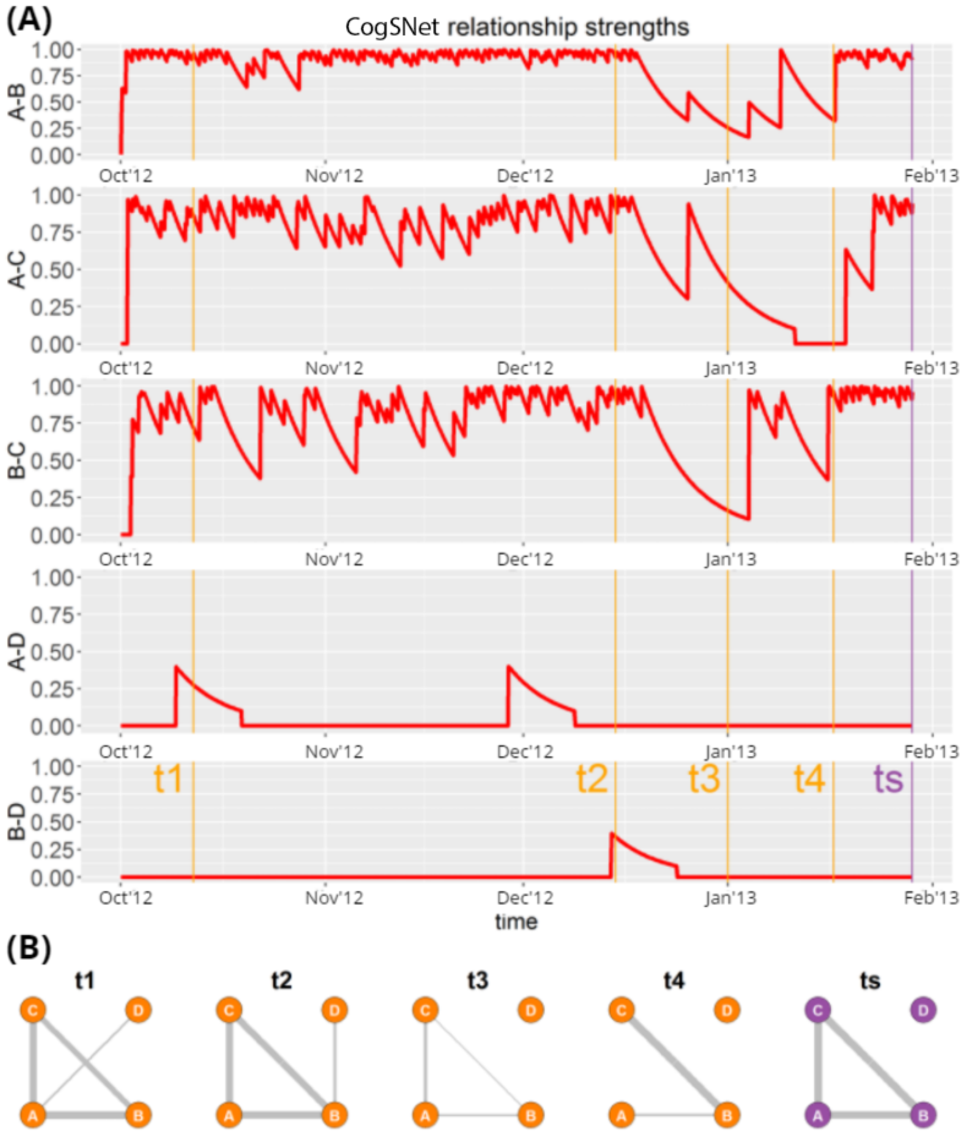


Figure 3.3: The four-node CogSNet network for the sample of real NetSense data, $\mu = 0.4$, $\theta = 0.1$, and L=10 days; nodes A, B, C, D correspond respectively to participants with ids 40997, 11360, 10841, and 1232. **(A)** Relation strengths according to the CogSNet model over 4-month period (one term). **(B)** Network snapshots at four time-stamps t_1-t_4 and at the survey time t_s.

outside of the test group, so recording was done on both sides of the communication, by sending and receiving calls/messages, if both belonged to the study. The second part of the dataset includes surveys containing peers enumerated by the participants at the end of each term in response to the following question: *"In the spaces below, please list up to 20 people (friends, family members, acquaintances, or other people) with whom you spend time communicating or interacting".* The data contains 6,290,772 human mobile phone communication events, including both calls and text messages. These are augmented by 578 surveys containing self-reports of top contacts.

This dataset is used to study the evolution of two coupled social networks of university students. The first is a behavioral network representing interactions between individuals in the form of the records of their mobile calls and text messages. The second one has perceptual edges defined by the personal nominations. These nominations are based on students' perception of the corresponding relations as one of the top twenty most interacting peers in the surveys administered to participants. These surveys cover the first four semesters of the student's college experience (beginning of Freshman year to the end of the Sophomore year). The list of nominations predicted from the CogSNet network model purely from the communication event data is compared with the list of nominations collected in a given survey. Fig. 3.3 shows an example of a dynamic social network generated from a subset of NetSense data using the CogSNet model.

3.6.2 Quality measure

To compare the performance of all the models, we use a *Jaccard* metric, (see equation 3.7), which measures the ratio of the number of nominations produced by the model that are also ground truth nominations listed in the corresponding survey divided by the number of unique nominations on both lists. The values of Jaccard metric for a single surveyed student participant v_i has been computed as follows:

$$Jaccard(v_i) = \frac{|V_i^{CogSNet} \cap V_i^{survey}|}{|V_i^{CogSNet} \cup V_i^{survey}|}, \tag{3.7}$$

where V_i^{survey} is the set of up to 20 peers enumerated in the survey by the participating student v_i; $V_i^{CogSNet}$ is the set of $|V_i^{survey}|$ neighbors of this student in the CogSNet network with the largest non-zero weights on edges to this student on the day on which the given survey was administered. Using Jaccard metric, CogSNet model for different parameters is compared to reference models.

3.6.3 Use case

Let us consider small scale example to understand idea of CogSNet and its verification. For example that is presented earlier in this work, see Fig.3.2, CogSNet model at time t represents a weighted network. A rank of nodes for each node can be delivered e.g. node A would have rank: [1.D - with value 0.5, 2. B - with value 0.4], next node B would have rank: [1. A - with value 0.4, 2. C - with value 0.3] and so on. For the use of this example it can be assumed that node A complete a survey at time t, and lists nodes: [D,B,E], where nodes D and B exists in CogSNet model, but Node E not. In such situation Jaccard measure for node A in given survey have value:

$$Jaccard(A) = \frac{|V_A^{CogSNet} \cap V_A^{survey}|}{|V_A^{CogSNet} \cup V_A^{survey}|} = \frac{\{D, B\} \cap \{D, B, E\}}{\{D, B\} \cup \{D, B, E\}} = \frac{2}{3} = 0.(6), \tag{3.8}$$

Such computation is done for all nodes in the network and then an average value is taken as a result.

3.6.4 Results and their analysis

In this section results of verification of CogSNet model are presented along with their analysis.

Fig. 3.4 shows the results of this comparison over the range of parameters corresponding to reported values for a memory life time of one day to 43 weeks. As reported in [20], the ability to recall information about social interactions starts to degrade after about one week. The experiments using NetSense dataset reveal that the performance is the highest when the forgetting of unreinforced memory traces happens within one week. The results remain satisfactory for forgetting thresholds within the range from a couple days to even two weeks. With the threshold within a two-weeks range, the model with either power or exponential forgetting is to be preferred over any baseline model. The performance of memory models using the power and exponential forgetting functions are similar to one another, but with a limited range of parameters, exponential forgetting tends to achieve a slightly higher value of Jaccard metric than the power function does [3] for some set of parameters μ and θ. We do not observe such superior performance of power forgetting here both functions have a similar peak of Jaccard metric albeit for the different lifetime values. It can be also observed that the performance of the model depends highly on parameters μ and θ, see Fig. 3.4A for $\mu = 0.3$ and $\theta = 0.2$ - Jaccard measure of the model fall even below Frequency-based baseline for the most of values of lifetime L.

When comparing the results of surveys with the states of the CogSNet network at the times of the surveys, the Jaccard metric is as high as almost 30% for the exponential function for L=2 days and about 28% for other parameters at about one week of memory lifetime. The performance of the model with power forgetting function achieves 28-29% of the Jaccard metric within the range 3-8 days for different sets of parameters μ and θ. The distant second is the recency-based RC model which delivers a much lower Jaccard metric of 17.8%.

To compare the performance of particular sets of parameters between themselves non-parametric statistical Kruskal-Wallis tests were performed. Results of pairwise tests performed for selected sets of parameters are presented in Table 3.1. The outcome of tests shows no statistical difference among the performance of the most of parameters with one exception for power forgetting function, $\mu = 0.3$ and $\theta = 0.2$.

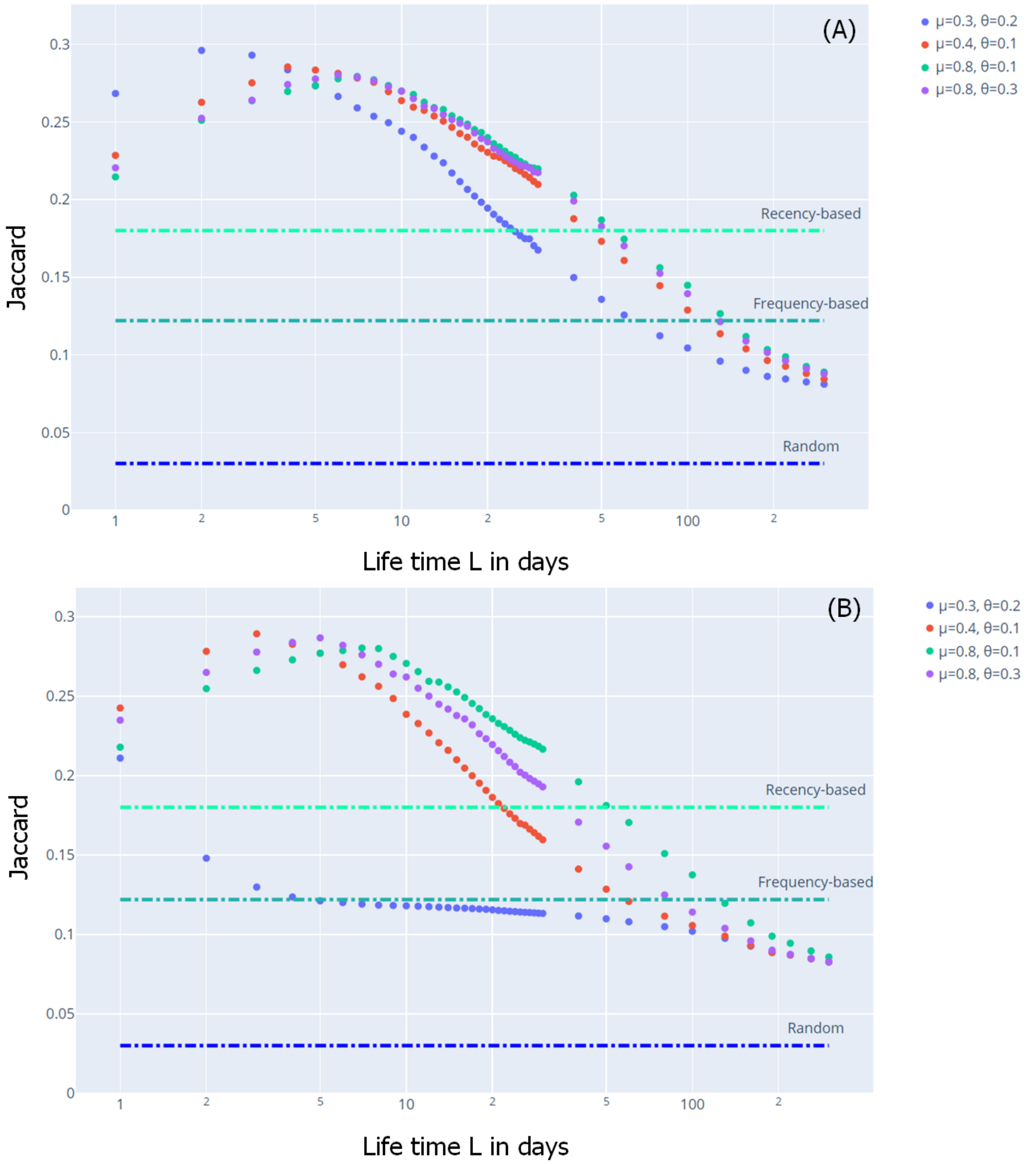


Figure 3.4: The plot of the Jaccard metric as a function of life time L. The metric measures the overlap between the two sets of peers, one identified by CogSNet and the other listed by students in the survey. These plots are compared to the results achieved by the three baseline models: recency-based with the best results obtained with the number of recent events set to 400, frequency-based, and random. The results are plotted for the CogSNet running with various parameters for **(A)** exponential and **(B)** power functions.

forgetting function	parameters	exponential				power			
		μ=0.3, θ=0.2	μ=0.4, θ=0.1	μ=0.8, θ=0.1	μ=0.8, θ=0.3	μ=0.3, θ=0.2	μ=0.4, θ=0.1	μ=0.8, θ=0.1	μ=0.8, θ=0.3
exponential	μ=0.3, θ=0.2	—	0.0567	0.0228	0.0311	0.0000	0.6462	0.0341	0.2123
	μ=0.4, θ=0.1		—	0.6395	0.7490	0.0000	0.0263	0.7844	0.4118
	μ=0.8, θ=0.1	*		—	0.8784	0.0000	0.0098	0.8711	0.2157
	μ=0.8, θ=0.3	*			—	0.0000	0.0135	0.9889	0.2678
power	μ=0.3, θ=0.2	***	***	***	***	—	0.0000	0.0000	0.0000
	μ=0.4, θ=0.1		*	***		***	—	0.0149	0.1237
	μ=0.8, θ=0.1	*				***	*	—	0.2800
	μ=0.8, θ=0.3				*	***			—

Table 3.1: Numerical and symbolic p-values of pairwise Kruskal-Wallis test comparing CogSNet model performance - computed with Jaccard measure - for selected sets of parameters and forgetting functions. The top-right section of the table shows p-value of statistical test, while bottom-left section contains symbols which are corresponding to p-values in the top-right section in following manner: '*' - p-value is less or equal 0.05, '**' - p-value is less or equal 0.01 and '***' - p-value is less or equal 0.005. Empty field state for p-value over 0.05.

3.7 Properties of the model

A novel Cognition-driven Social Network (CogSNet) model is introduced, one that captures the impact of human memory on the perception of accumulated events and on decisions to form, maintain, or dissolve social relations. The model explicitly represents human memory dynamics, such as the gradual decay of memory traces over time. With suitable data, it can be extended to include additional cognitive aspects, such as individual levels of sensitivity to relevant events, emotions or distractions during perception of events. Hence, the model is capable of capturing the dynamics of social interactions in natural settings from the cognitive perspective of each participant. The results reveal that the perception of the depth of interactions between people is well captured by the CogSNet model. At any given point in time, the model can compute the current strength of memory traces, including the impact of discrete events creating or reinforcing these traces. Furthermore, the reinforcement peak value can be adjusted for a given person and for an individual event, cf. also Equation (3.1). What is observed here, is a partial manifestation from human memory. However, even taking into account that the model was built based on a single data source, over 6 million telephone calls and messages among the NetSense study participants, the model provides good accuracy in predicting the salience of social contacts over all 578 surveys completed by 184 participants. This accuracy most likely could have been increased if the parameters had been individually adjusted for each participant.

In future work, an extension of the model can be made, which includes accounting for distractions during interactions, individualized strength, asymmetric of interactions of significance to participants (e.g., hierarchical relationships), and the impact of forms of interactions and of associated emotions. Hence, the CogSNet model represents an important first step towards modeling social network dynamics through the prism of human cognition. In literature, there are three approaches that are used widely to understand human cognition: cognitive bandwidth (CB), dual-process morality (DPM) and implicit association tests (IAT). In [66] those methods are compared and their common drawback is pointed out - the lack of perspective that takes a cultural background into account. The proposed approach makes such considerations possible and this challenge is a promising one of the future work directions.

The CogSNet model can be also used to measure the overall strength of the relationship over time. Modeling the forgetting mechanism with a continuous function allows computing integral in some period, which can be considered as quantification of relationship strength in a given period. In Fig.3.1, this idea is presented in period from t_1 to t_2. As can be observed, the value of the integral depends on the used forgetting function. This is also an interesting direction that could be developed in the future.

3.8 Conclusions

In this chapter, a new Cognition-driven Social Network is presented. This is an important step toward a fully cognition-aware temporal social network model that will be capable to capture more complex human behavior. The proposed model one of the application of temporal networks, here in cognition science, but it is not limited only to that. A brief idea is presented further in this dissertation.

Chapter 4

Entropy-based measures for temporal network dynamics

Through centuries people communicate with each other in many ways. Nowadays, the amount of communications channels available for us is even hard to count - starting from text messages, emails, multiple communicators, social media, phones, video calls and, of course, good old face-to-face conversation. Just by looking at interactions we may have the impression that the structure of such communication is random and we can not tell with ease if the dynamic of it is changing or not. Some recent studies show that we people do not communicate randomly [130, 108, 97, 19], however, there are still some open questions related to its dynamics. In this chapter, we introduce a new set of entropy-based measures for temporal networks. The notion of entropy in temporal networks was used before in [22] to detect organization structure in the email communication network. Yet, we propose measures that give some insights into temporal network dynamics.

4.1 Motivation

My research motivation is to develop a measure that is able to quantify a stream of events. Such a measure should be able to assign a value to a stream and be meaningful while comparing different streams. Since my research focuses on temporal networks, the proposed measure should provide a possibility of comparing temporal networks of different sizes. Further, this measure should be also useful to show the dynamic of a temporal network given by event sequence. My assumption is that dynamics could change in time, hence the proposed measure should be able to incorporate this factor too.

4.2 Research questions

The area of temporal networks has been researched intensively over a couple of last years, however, there is no established temporal network representation as to the final standard. It might be, that it is impossible to choose one. Yet, as it is mentioned in previous chapters, among a number of representations there is one of the basic - event sequence - which I use to compute entropy-based measures. Using such represented temporal networks, I try to answer questions: 1) if human communication is in fact not random and 2) if there is some trend in human communication dynamics that we can observe?

4.3 General idea

Here, I propose a new set of entropy-based measures for temporal network given as an event sequence, which would provide some insights about human communication dynamics. In [62], we proposed three new approaches to measure temporal network dynamics, all based on entropy. These are the following:

1. first-order entropy, which based on the probability of node appearance in event sequence,
2. second-order entropy, based on probability of event appearance in the temporal network or in other words probability of interaction between unique pair of nodes
3. third-order entropy, based on probability of succession, which is a pair of consecutive events.

Each approach is able to capture a different aspect of network dynamics and has a potential for different applications.

Those measures based on Shannon entropy 4.1 - well-known concept from physics and information science.

$$S = -\sum_{i \in O} p(i) \ln(p(i)) \tag{4.1}$$

where $p(i)$ is occurrence probability of state or object i, and O is the set of all possible states/objects [105].

The proposed measures were developed in collaboration with Professor Boleslaw K. Szymanski from Rensselaer Polytechnic Institute, Troy, NY, USA [62] during my research visits at RPI in 2017-18.

4.3.1 First-order entropy

At the beginning, the definition of *the first-order entropy* - also called *node entropy* - will be presented. It is based on a probability of occurrences of individual nodes $v_i, v_j \in V$, i.e. humans participating in interactions – events $ev_{ijk} = \{v_i^s, v_j^r, t_k\}$. This measure can be considered in three variants depending on focus of the interests:

1. node being a speaker/sender v_i^s,
2. node as a listener/receiver v_j^r or
3. node occurring as a speaker v_i^s or listener v_j^r.

Based on definition of Shannon entropy Eq. 4.1, the first-order entropy $S_1(ES)$ a.k.a. node entropy is defined for a event sequence ES that contains a fixed set of nodes V. The first-order entropy $S_1(ES)$ is defined as follows:

$$S_1(ES) = -\sum_{v \in V} p_1(v) \ln(p_1(v)) \tag{4.2}$$

where $p_1(v)$ is a probability of occurrence for node $v \in V$ in the appropriate role – the sender, receiver or any of these two. The choice of the role (and the entropy variant) depends on what kind of analysis we want to perform.

The node entropy measures the diversity of node popularity in the temporal network. In other words, greater entropy means that the nodes have a rather equal probability of occurrence and the small one denotes that some nodes occur significantly more frequently than the others. Entropy has the maximum value when probabilities for all nodes from V are equal. The equal probabilities emerge when all nodes occur the same number of times, e.g., only once or all twice, etc. Hence, equal probabilities are:

$$p_1(1) = p_1(2) = ... = p_1(n) = \frac{1}{n} \tag{4.3}$$

Then, the maximum possible value of entropy for a given set of nodes V is defined as:

$$S_1^M(ES) = -\sum_{v \in V} \frac{1}{n} \ln\left(\frac{1}{n}\right) = -|V| * \frac{1}{n} \ln\left(\frac{1}{n}\right) = \ln(n) \tag{4.4}$$

Use case

The idea of the first-order entropy can be explained for the small scale example depicted in Fig.2.1b. The network consists of four nodes: A, B, C, and D, hence the maximal entropy value is: $S_1^M(ES) = ln(4) \approx 0.6$. Such case happens if occurrence of all nodes is equally probable, i.e. $\forall_{v \in V} p_1(v) = 0.25$. For the case from Fig.2.1b, we have $p_1(A) \approx 6/10 = 0.6$, $p_1(B) = p_1(C) \approx 2/10 = 0.2$, $p_1(D) \approx 0/10 = 0$. These values together with probabilities corresponding to the maximum value of S_1^M are presented in Fig.4.1.

4.3.2 Second-order entropy

Next, *the second-order entropy* - also called *edge entropy* - will be defined. The second approach utilizes probabilities of occurrence of edge $e_{ij} \in E$. Again, based on definition of Shannon

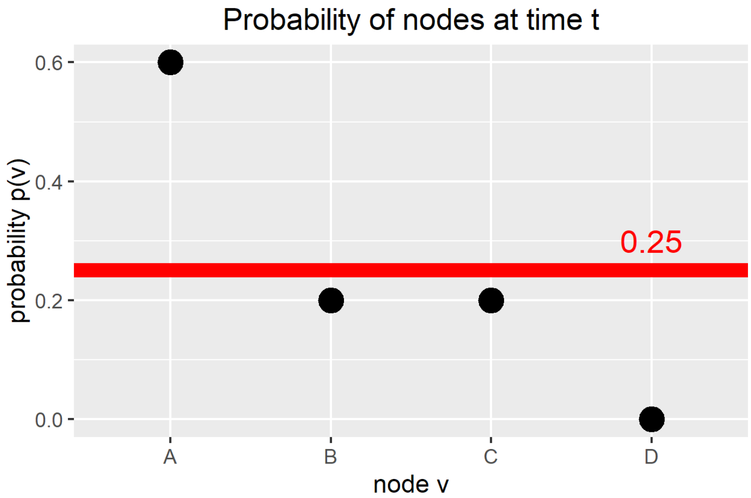


Figure 4.1: Probability of nodes appearance in the sequence by time $t = 10$. The red line denotes values of probabilities for the maximum possible entropy in the 4-node temporal network.

entropy, the second-order entropy a.k.a. edge entropy is defined as follows:

$$S_2 = -\sum_{e_{ij} \in E} p_2(e_{ij}) \ln(p_2(e_{ij})) \tag{4.5}$$

where $p_2(e_{ij})$ is a probability of edge e_{ij}, i.e. probability that events ev_{ijk} are related to edge e_{ij}. This entropy of the temporal network provides information about how uncertain (random) pairs of nodes (individuals) interact with each other. The greater edge entropy value reflects that the distribution of participating pairs is close to uniform distribution while the smaller value means that some pairs interact more frequently than the others.

Now the maximum value of edge entropy can be defined, assuming that probabilities for all possible edges are equal, i.e. all possible pairs of nodes v_i^s, v_j^r appear in the same number of events ev_{ijk}. The number of possible edges is $|\Omega(E)| = n(n-1)$. Then, it is:

$$p_2(e_{12}) = p_2(e_{13}) = ... = p(e_{n-1}) = ... = p_2(e_{n(n-1)}) = \frac{1}{n(n-1)} \tag{4.6}$$

With this probability, the maximum value of edge entropy would be defined as:

$$\begin{aligned} S_2^M &= -\sum_{e_{ij}\in\Omega(E)} \frac{1}{n(n-1)} \ln\left(\frac{1}{n(n-1)}\right) = \\ &= -n(n-1) * \frac{1}{n(n-1)} \ln\left(\frac{1}{n(n-1)}\right) = \ln(n(n-1)) \end{aligned} \tag{4.7}$$

For larger number of nodes (large n), maximum value of the second-order entropy can be approximated: $S_2^M \approx 2S_1^M$.

4.3.3 Third-order entropy

Finally, *the third-order entropy* - also called *succession entropy* - definition will be presented. This approach base on a probability of occurrence two particular node pairs (edges) in events one directly after another. We refer to such a pair of edges as succession. Event sequence ES is a list of M events ordered by time: $ES = (ev_1, ev_2, ..., ev_k, ev_{k+1}, ..., ev_M)$, and $ev_k = (s_i, r_j, t_k), ev_{k+1} = (v^s_{i'}, v^r_{j'}, t_{k+1}) \Leftrightarrow t_k \leq t_{k+1}$. For two consecutive events ev_k and ev_{k+1}, we can extract participating nodes $s_i, r_j, s_{i'}, r_{j'}$, respectively, i.e. edges $e_{ij}, e_{i'j'} \in E$. Such two edges define the single kth edge succession occurrence $sc_k = (e_{ij}, e_{i'j'})$ and the set of distinct successions (unique pairs of edges) is denoted by SC. An idea of successions is presented with simple example in Fig. 4.2. Obviously, it may happen that $e_{ij} = e_{i'j'}$. The set of all potentially possible successions is $\Omega(SC)$ with size $|\Omega(SC)|$. This size is limited by the maximum size of the edge set E for a given set of nodes V: $|\Omega(SC)| = |\Omega(E)|^2 = n^2(n-1)^2$.

Using probability of succession we can define succession entropy:

$$S_3 = -\sum_{sc\in SC} p_3(sc) \ln(p_3(sc)) \tag{4.8}$$

where $p_3(sc)$ is a probability of edge succession sc.

The value of succession entropy quantifies information about how uncertain (random) is the presence of a particular succession of edge pairs in the event sequence. Similarly to

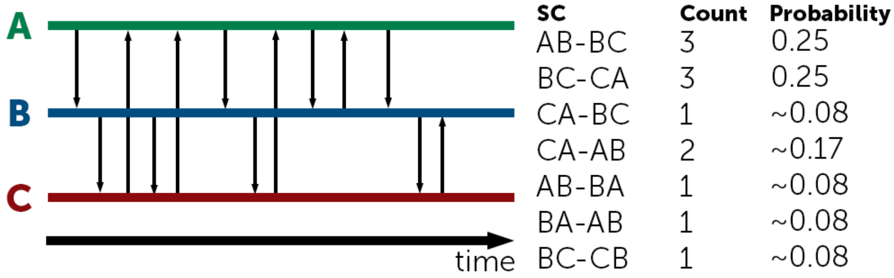


Figure 4.2: A sample network with a table of estimated probabilities of successions. Please note that successions that do not appear in the sequence (count equal 0) are not presented.

previous approaches, we can find the maximum value of succession entropy by assuming an equal distribution of succession probabilities:

$$p_3(sc_1) = p_3(sc_2) = p_3(3) = ... = p_3(|\Omega(SC)|) = \frac{1}{|\Omega(SC)|} \tag{4.9}$$

For these probabilities, the maximum value of succession entropy would be:

$$\begin{aligned} S_3^M = -\sum_{sc \in \Omega(SC)} \frac{1}{|\Omega(SC)|} \ln\left(\frac{1}{|\Omega(SC)|}\right) = -|\Omega(SC)| \frac{1}{|\Omega(SC)|} \ln\left(\frac{1}{|\Omega(SC)|}\right) = \\ = \ln(|\Omega(SC)|) = \ln\left(|\Omega(E)|^2\right) = 2\ln(n(n-1)) = 2S_2^M \end{aligned} \tag{4.10}$$

For larger quantity of nodes (large n): $S_3^M \approx 4S_1^M$

4.4 Algorithms

In order to examine the computational performance of the proposed measures, appropriate algorithms are designed for single core and parallel calculations [citation of complexity paper]. The single core algorithms are direct transition from equations 4.2, 4.5 and 4.8 while parallel algorithms are optimized in terms of performance in the distributed computing environment. Algorithms 1,2 and 3 presents the single core algorithms for each entropy-based measures.

Simple analysis of Algorithms 1, 2 and 3 shows that complexity is $O(n)$. Those algorithms compute the entropy value, the first-, second- and third-order, respectively, for the single tem-

Algorithm 1 First-order entropy algorithm

Input: ES – sequence of events ev_{ijk}
Output: *entropy* – the first-order entropy for the event sequence

```
size(ES) ← number of events in ES
V ← set of unique nodes {extract distinct nodes from ES}
entropy ← 0
for each v ∈ V do
    n ← count events ev_ijk from ES containing node v {v = v_i or v = v_j / v = v_i / v = v_j}
    {depending on the entropy type: both nodes / sender / receiver}
    entropy ← entropy + (n/size(ES)) * log(n/size(ES))
end for

return entropy
```

Algorithm 2 Second-order entropy algorithm

Input: ES – sequence of events ev_{ijk}
Output: *entropy* – the second-order entropy for the event sequence

```
size(ES) ← number of events in ES
E ← set of node pairs {extract distinct network edges e_ij - unique pairs of nodes from ES}

entropy ← 0
for each e_ij ∈ E do
    n ← count events ev_ijk from ES that contain the pair of nodes from e_ij
    entropy ← entropy + (n/size(ES)) * log(n/size(ES))
end for

return entropy
```

Algorithm 3 Third-order entropy algorithm

Input: ES – sequence of events ev_{ijk}
Output: *entropy* – the third-order entropy for the event sequence

```
S ← set of successions {2-event sequences, i.e. two consecutive events ¡ev_ijk,ev_i'j'k'¿ existing
in ES}
size(S) ← number of successions {size(S) = size(ES) − 1}
entropy ← 0
US ← set of distinct successions {pairs of network edges ¡us = e_ij, e_i'j'¿ extracted from S}
for each us ∈ US do
    n ← count occurences of us in S
    entropy ← entropy + (n/size(S)) * log(n/size(S))
end for

return entropy
```

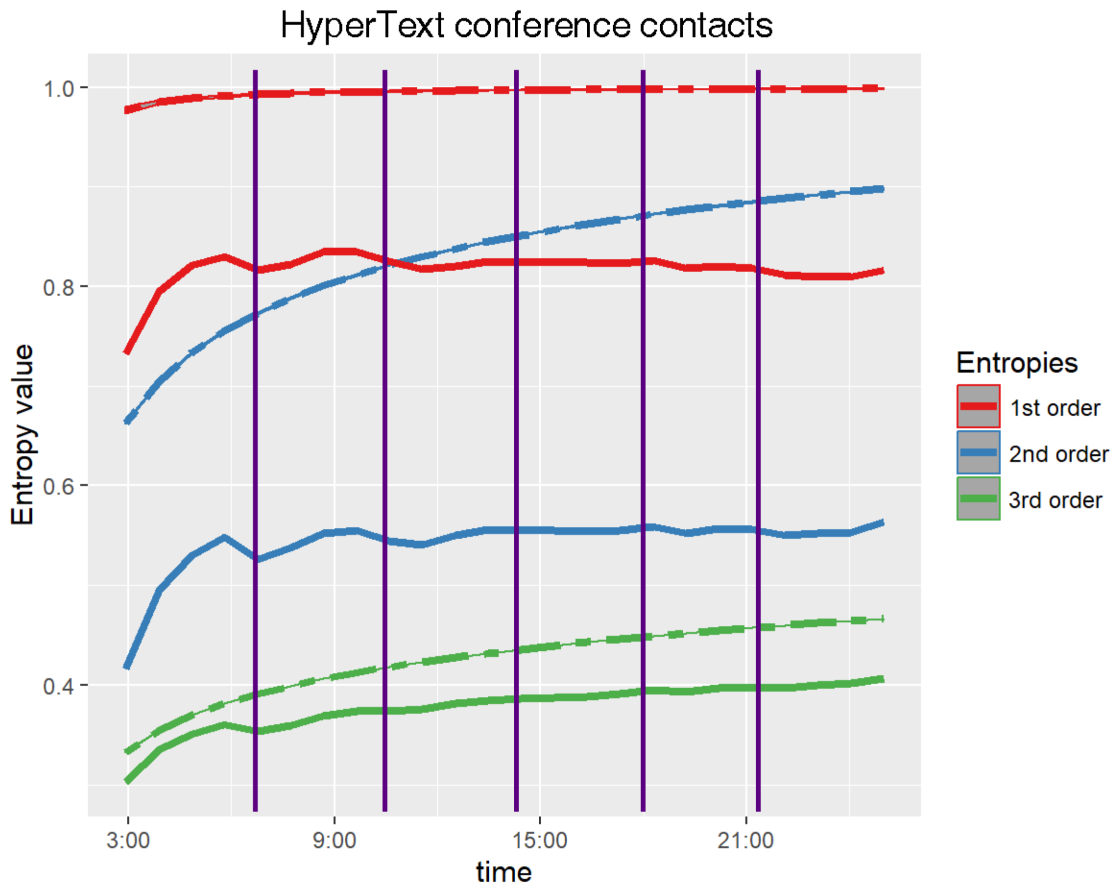


Figure 4.3: Entropy value for a real event sequence - face-to-face contacts during the HyperText conference. A solid line represents entropy of a real accumulative network. Dashed lines correspond to the artificial random network with similar properties (the same number of nodes and events); they show maximum values of entropies. Purple vertical lines are calculation points – every 1000 events, where entropies are computed in the accumulative way to measure network dynamics. All entropy values were normalized to show differences in entropies of various orders. Normalization is performed using the maximum possible entropy value at the end of the stream [62].

poral network given as an event sequence ES. However, to examine entropy changes over time in the temporal network that reflects its dynamics, the entropy values have to be separately computed for each time increment δt. Alternatively, they can be calculated for each event increment. This notion is presented in Fig.4.3.

The computation of entropy in an accumulative way requires another algorithm. The designed Algorithm 4 is generalized for all three types of entropies (node, edge and succession). The generalized algorithm has complexity $O(n)$, however joint complexity with algorithms for entropies has $O(n^2)$.

Algorithm 4 Accumulative network computing algorithm

Input: ES – sequence of events ev_{ijk}
$jump$ – number of events accumulated in every iteration
Output: $entropies$ – list of accumulative entropies of a given order

$entropyFuction \leftarrow$ Algorithm 1, Algorithm 2 or Algorithm 3
$size(ES) \leftarrow$ number of events in ES
$step \leftarrow jump$
$entropies \leftarrow []$ {the list of entropy values is initially empty}
while $step < size(ES)$ **do**
  $sequence \leftarrow ES[1 : step]$ {extract a sequence of length $step$ from ES starting from the first event}
  $entropy \leftarrow entropyFunction(sequence)$
  $entropies.push(entropy)$ {append a new entropy value to the list of computed entropies}
  $step \leftarrow step + jump$ {increase the length of the next sequence}
end while

return $entropies$

For a large event sequence, the computation may take quite a long time, so to speed up calculations, Algorithm 5 for the parallel environment was designed. The original R library for parallel computing [2] was used for parallelization purposes.

All crucial software code implemented for entropy computation is published in the GitHub repository and is publicly available at `https://github.com/MKul/tnEntropy`.

4.5 Experimental setup

This chapter contains the description of the experimental setup to examined proposed measures and the effectiveness of algorithms. Description of experiments on entropy measures will be provided, followed by experiments on algorithms' effectiveness.

4.5.1 Analysis of entropy-based measures

To examine proposed entropy-based measures, entropy values for four different dataset with data of real human interactions were computed: (1) face-to-face meetings at HyperText con-

Algorithm 5 Accumulative network parallel computing algorithm

Input: ES – sequence of events ev_{ijk}
$jump$ – number of events accumulated in every iteration
Output: $entropies$ – list of accumulative entropies of a given order

$entropyFuction \leftarrow$ Algorithm 1, Algorithm 2 or Algorithm 3
$size(ES) \leftarrow$ number of events in ES
$step \leftarrow jump$
$steps \leftarrow []$
while $step < size(ES)$ **do**
  $steps.push(step)$
  $step \leftarrow step + jump$
end while
$entropies \leftarrow []$
{for loop in parallel - distribute processing among multiple computing cores}
for each $step$ in $steps$ **do**
  $sequence \leftarrow ES[1 : step]$ {extract $step$ first events from ES}
  $entropy\_value \leftarrow entropyFunction(sequence)$
  $entropies.push(entropy\_value)$
end for

return $entropies$

ference, (2) text messages exchanged between students for 6 semesters (NetSense), (3) email communications in the manufacturing company, and (4) face-to-face interactions between patients and hospital staff members. The timeline of entropy was computed by taking a window from the beginning of network existence to some points in time - usually evenly distributed. In other words, entropy was computed cumulatively for an online stream of interaction data. To provide the baseline for real event sequences, 100 artificial event sequences were generated - for each dataset with the same numbers of nodes, events, and timestamps by randomly reselecting pairs of nodes involved in each event. In static networks, such a procedure would be called rewiring. Distributions of random event sequences are discussed later in Subsection on datasets. The average value of entropy for random event sequences is computed and compared against the values for the real network using Z-score – the distance measure that, in general, shows the number of standard deviations by which the value of entropy for real sequence is above the mean value of random streams. The negative values of Z-score mean that entropies for real data are smaller than random ones and greater the difference is more negative Z-scores are. The general concept of experiments is presented in Fig. 4.4.

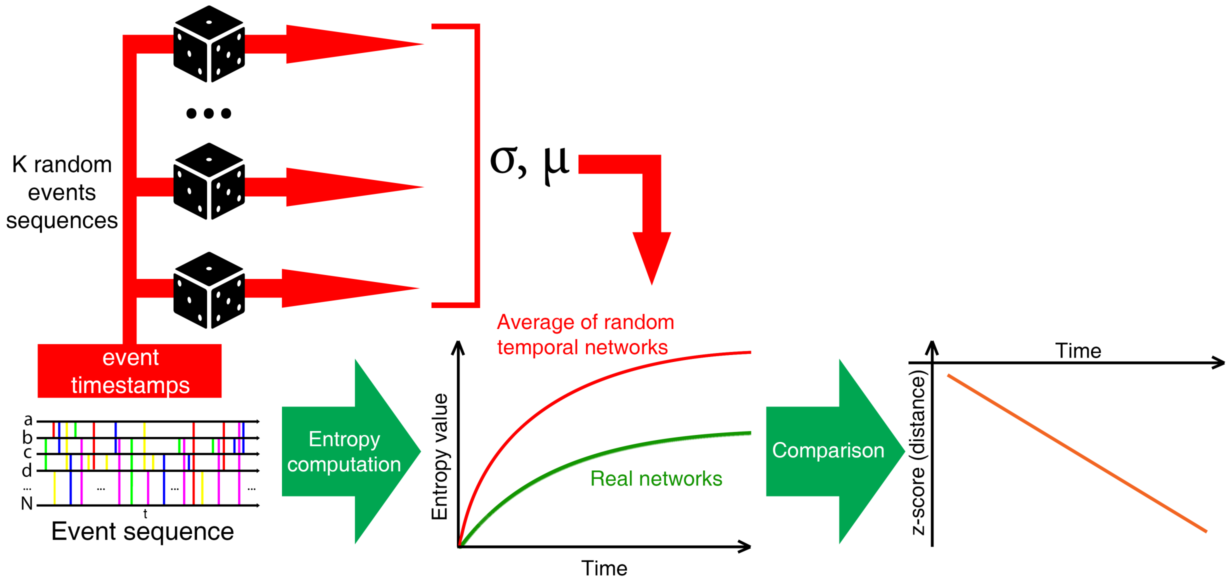


Figure 4.4: General schema of experiments. K=100 was used. From the original (real) event sequence, event time-stamps are extracted as a base for random sequence generator. Entropy value is computed for real event sequence and artificial sequences. We compare results for real data with summarized results for artificial data using Z-score.

In experiments, four datasets with real-world temporal networks are used. In Table 4.1 basic statistics about datasets are presented and short descriptions of used datasets are presented below:

Real event sequences

- **NetSense - text messages**. The dataset contains phone and text communication among students at University of Notre Dame. The dataset was created to map peers' social network and contains data from 3 years (6 semesters) starting from September 6, 2011. [113]
- **Hospital ward dynamic contact network**. This dataset contains the temporal network of contacts between patients, patients and health-care workers (HCWs) and among HCWs in a hospital ward in Lyon, France, from Monday, December 6th, 2010 to Friday, December 10th, 2010. The study included 46 HCWs and 29 patients [118]. Contacts were collected using proximity sensors which do not provide information about the direction

of the contact. However, for our experiments, we consider it as directed communication for easier comparison with other datasets.

- **Hypertext 2009 dynamic contact network**. The dataset was collected during the ACM Hypertext 2009 conference, where the SocioPatterns project deployed with the Live Social Semantics application. Conference attendees volunteered to wear radio badges that monitored their face-to-face proximity. The dataset published here represents the dynamical network of face-to-face proximity of 110 conference attendees over about 2.5 days [44]. The collecting method does not provide direction of contacts but for easier comparison with other datasets, we consider contacts as directed.

- **Manufacturing emails**. This is the internal email communication between employees of a mid-sized manufacturing company. The network is directed and nodes represent employees while events correspond to individual emails [78].

Table 4.1: Datasets in numbers

	# of nodes	# of events	# of unique edges	# of unique successions
HyperText conference – meetings	113	20,818	2,498	12,060
Netsense – text messages	212	28,520	896	10,710
Manufacturing company – emails	167	82,927	5,784	56,367
Hospital ward dynamic contact network	75	32,424	1,139	13,162

Equivalent artificial event sequences

Along with real-world networks, artificial networks are generated to provide a baseline for proposed measures. They are generated based on real networks as follows: number of nodes, number of events as well as time-stamps were preserved, but nodes involved in the events were drawn out from the uniform distribution. The entire process is described with the following procedure:

Procedure 1

1. Take the real event sequence ES and extract distinct nodes from event's senders and receivers – create set of nodes V.
2. Take the next event from the real event sequence, starting from the first one and keep its timestamp t_k.
3. Randomly select the sender $v_i^s \in V$ (according to selected distribution).
4. Randomly select the receiver $v_j^r \in V$ (according to selected distribution).
5. If the sender and receiver are the same, repeat step 4.
6. Create event $ev_{ijk} = (v_i^s, v_j^r, t_k)$.
7. If it is the last event in the real sequence ES – stop, otherwise go to step 2.

For each real temporal network, we generate 100 artificial networks. One may ask if the uniform distribution is the correct choice. Along with uniform distribution, normal and exponential distribution was tested but differences were not statistically significant in terms of entropy.

4.5.2 Analysis of algorithms' effectiveness

In this section, experiments on the efficiency of entropy computation are described. Three scenarios were considered: 1) single network computation on a single computational core, 2) accumulative network estimation on a single core and 3) parallel computation for the accumulative network on multiple cores. All experiments were performed in an isolated environment. The most important specification aspects of the environment are as follows:

- OS: Linux, kernel 2.6.32-696.30.1.el6.x86_64

- Processing unit: Intel Xeon E5-2670 v3 2.3 GHz, Haswell
- Memory: 512 GB
- Programming environment: R ver. 3.2.3

The performance of entropy computation is examined using artificial communication streams (temporal networks) generated from the uniform distribution. In other words, all nodes had equal probability to appear in the events. The streams contained up to 10,000,000 events and 10,000 distinct nodes. For each stream size, 20 various event streams were independently generated. The averaged values over these 20 streams together with the appropriate standard deviation are shown in charts. Some studies were carried out for smaller streams only. For each stream (network), three types of entropy were computed.

To examine the effectiveness of single network computation, single core computation is performed. Proposed implementation in R shows high efficiency for single core computation and - to my best knowledge - development of implementation for multiple cores is impossible in an efficient way.

Accumulative network entropy is the way to examine network entropy dynamic evolution. Here, entropy values are computed for the increasing (accumulative) size of the stream starting from 100 events up to a given size with a 100-event step. In other words, for a stream of size 500 events, five entropies are calculated: for 100, 200, 300, 400, and 500 events. The processing time captures all these 5 calculations. The increasing size of streams linearly increases the number of streams but their average length is additionally longer and simultaneously raises the number of distinct objects (nodes for the first-order or edges for the second-order entropies). It means that computational time grows with the stream size much faster than linearly.

Computation for the accumulative network on multiple cores has a similar structure to the previous, iterative approach. The only difference is parallel computation for subsequences of the event sequence. In most cases, it is impossible to have as many cores as the number of subsequences to evaluate. Thus, the consecutive subsequences are assigned to the cores after the previous computation is terminated. The overall idea of parallel computing of entropies for the accumulative network is presented in Fig. 4.5.

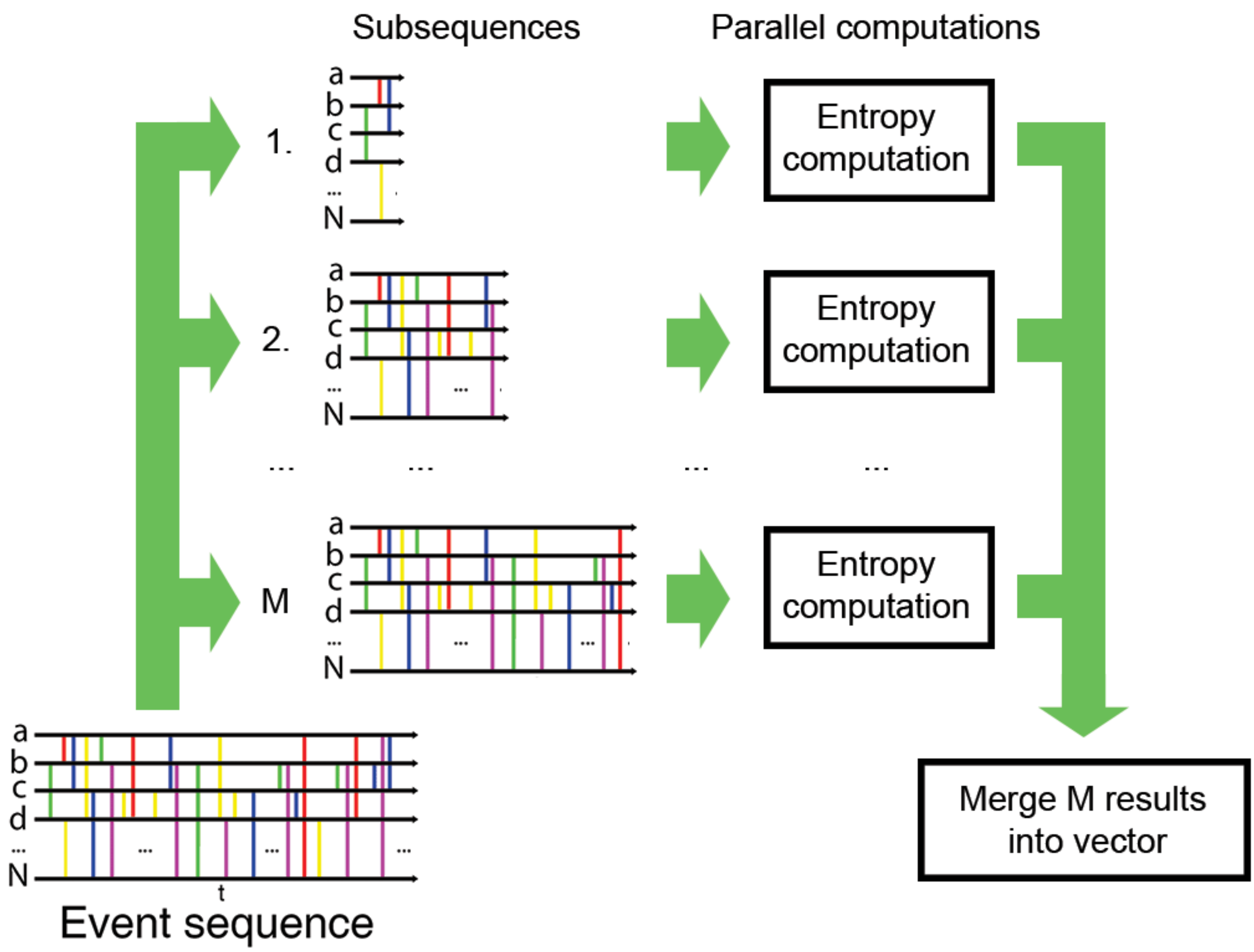


Figure 4.5: Schema of parallel computation of entropies for the accumulative network.

4.6 Results

4.6.1 Analysis of entropy measures

Here results of experiments are presented. For more clear presentation only some of the results are presented but similar results are obtained for all used datasets. The first observation about the nature of entropy is that while cumulative computation maximum entropy is non decreasing since it depends on a non-decreasing number of nodes in the network (disappearing of nodes is not considered here). The second observation is that for real event sequence all entropies seem to converge to some value, see solid lines in Fig. 4.6A and B, and similar behavior for random event sequence. By looking at normalized value of entropy (Fig. 4.6A) we can observe that random sequences seem to converge to the maximum value of entropy - nearly to 1 - which is visible for first-order entropy. The shape of second- and third-order entropy suggests the same. The value of entropy for real sequences seems to converge to some value specific for a given sequence. That we can observe in all datasets divided into parts in Fig. 4.7. Here, the third observation emerges - values to which entropy tent to converge decrease over time. In Fig. 4.7B consecutive semesters have this value decreasing while it can not be observed for random sequences. A similar observation can be made for other datasets. The next observation is that there are some exceptions from previous observation - some periods in datasets have do not fall into this specific decreasing sequence. Those exceptions, however, can be explained by deeper observation of datasets e.g. in Fig. 4.7D we can notice that the first period - 2010-12-06 - has the smallest entropy value while according to third observation it should have the biggest value. This dataset - face-to-face contacts of hospital staff and patients were divided into singular days and this particular day - 6th of December is Saint Nicolas Day. Here an assumption can be made that celebrating such day can change the typical behavior of people and thus, change expected entropy value. Similarly, in Fig. 4.7C there is an unexpected value of entropy in 2010 July, which can be explained by vacation time when employees do not replay on emails as usual. The last observation is that in a particular period, a difference between the value of entropy for real sequence and for artificial sequence seems to be bigger over time. To examine that aspect Z-score distance measure is used. In Fig. 4.6C we can observe that the value of Z-score is decreasing over time (absolute value is increasing), thus there is a bigger difference between entropy value for real sequence and artificial one over time.

4.6.2 Effectiveness results

In Fig.4.8, the performance of the single network computation on a single core is presented. It can be observed that the computational time is proportional to the number of events in the

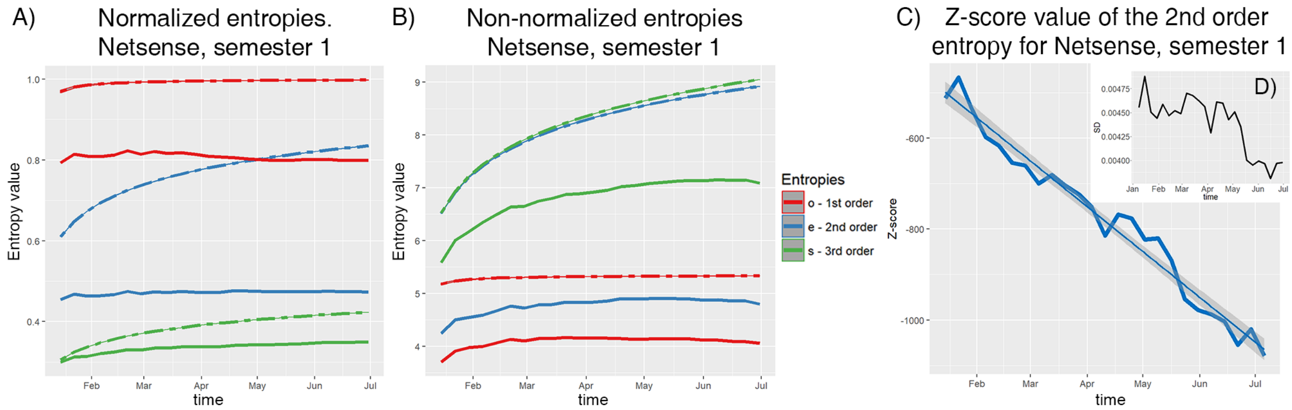


Figure 4.6: The NetSense dataset, the 1st semester. A) Values of normalized entropies. Solid lines refer to the original event sequence and dashed ones present the average value for the baseline – random sequences. B) Values of non-normalized entropies. C) Z-score for non-normalized second-order entropy with the computed trend and marked standard deviation (gray area). Inset (D) Standard Deviations (SD's) of entropies of the randomized networks; each SD is used as a denominator in the computation of the respective network Z-score.

stream. The increasing number of nodes in network increase non linearly with computational time, see the inset in Fig.4.8.

Next, another experiment to examine whether the proposed implementation is capable to process large networks is performed. In Fig. 4.9 a single network computation performance for the event sequences with 10,000 nodes and from 100,000 to 10,000,000 events in total is presented. As it can be observed, the computational time for the second- and third-order entropy for the largest stream is about 160 seconds. The first-order entropy is computed very fast with no so fast increase of time for the rise of stream size: from 0.066 sec to 0.416 sec for the largest stream. It can be also observed that the time grow is non-linear. A similar relation was seen in experiments on smaller networks and it can be explained by additional hard disk operations during computations.

Entropy computational times for the accumulative network on single core reveals the almost exponential relation to the stream size, see Fig.4.10. It means that the calculation of accumulative streams requires much more resources than the simple entropy evaluation. The number of events seems to be the most important factor in terms of computational time. Nevertheless, the number of nodes also impacts on computation time but much less, see the inset in Fig.4.10. The relation is similar to the one for the single network computation, compare to the inset in Fig.4.8.

Parallel computation of accumulative networks to examine how efficient is the parallel approach is performed. The results of parallel processing on 8 cores are presented in Fig.4.11. Only average values without standard deviation are shown for clarity. The parallel computation did not change the overall tendency of almost exponential growth. However, it reduced the computational time of the largest examined network (3,000 events, 100 nodes) by 40% in comparison to the single-core computation, see Fig.4.10. In the case of smaller networks, the gain is smaller or calculations take even more time due to an additional cost for parallelization itself. With the growth of the stream size, however, the parallel approach reveals its superiority much more.

To examine the influence of a number of cores on computational time an experiment on an artificial stream with 3,000 events and 100 nodes is performed, independently on different numbers of computational cores: from 2 to 12. The results are depicted in Fig. 4.12. It can be observed, that the benefit of parallel computation is visible for the second- and third-order entropies, while the first-order entropy provides no significant gain for more computational cores.

4.7 Discussion

Obtained results show a couple of interesting things about examined temporal networks but also ask some questions. Decreasing entropy in nature is something opposite to universe tendency, an entropy of particles rises over time due to the theory of thermodynamics. Yet, the presented results show that it is not the case in human communication. An explanation of such behavior may be that people getting know each other over time - or rather during communication - and become more selective with their interlocutors as time flows. In other words, people seem to limit their contacts to some group of people (probably colleagues, friends), which raises the probability of contact with them and in consequence - decrease the entropy of network. Such a hypothesis makes sense for presented datasets. Students in a university (NetSense dataset) in the first year have not many friends - they are in a new environment, new town, etc. - while later they make relations with other students and are also bound in project groups, classes and so on. Participants of the conference at the first two days may talk to many people, while on the last day - probably after some social event - they prefer to talk to a fewer number of people. Similar explanations can be made for other datasets. Here, a question can be raised: is there any terminal value to which entropy will converge? This is an interesting aspect and can be further investigated.

Entropy value depends also on temporal network size - the number of nodes, number of events - thus it is possible to compare temporal networks sourcing from the same environment e.g. same system, the same group of people, etc. Otherwise, it is not meaningful. However, a normalized value of entropy help to overcome this issue. In Fig. 4.13 values of normalized entropy are presented for different datasets. Those values - especially of second-order entropy (see Fig. 4.13B) - could be explained by nature of datasets, e.g., NetSense dataset is a group of students which are assigned to some classes which effectively limit their contacts, while manufacture employees can correspond to each other without artificial boundaries. Those are just hypotheses and a deeper investigation is required to confirm those.

Presented results show that entropy has some dynamics - it is changing in consecutive periods. However, the dynamic itself is not quantified here yet. Entropy here is a measure that shows some property of a temporal network. By analogy it can be understood as velocity in uniformly accelerated motion - it gives information at some point in time, but does not describe how it changes. Acceleration is what is required to quantify dynamics. In temporal networks measure of dynamics could be derivative of entropy. In Fig.4.14 this idea is presented using small temporal network sample presented earlier in this work (see. Fig.2.1).

4.8 Conclusions

In this chapter research on the new set of entropy-based measures is presented. Some interesting conclusions can be derived. First, people do not communicate randomly - a known fact now confirmed in temporal networks environment. The difference between human and random communication is even more clear over time - longer contact sequence is observed, the bigger difference is. Second, even more interesting is that the entropy of human communication decrease over time. It is an interesting observation that reflects one aspect of human nature. Next, some ideas about measuring the dynamics of a temporal network were presented, however much deeper investigation is needed. Finally, the effectiveness of iterative and parallel algorithms for entropy computation was examined. Parallel computation can be used to increase the effectiveness of accumulative entropy computation. Computations of second- and third-order entropies benefit the most on parallel computation.

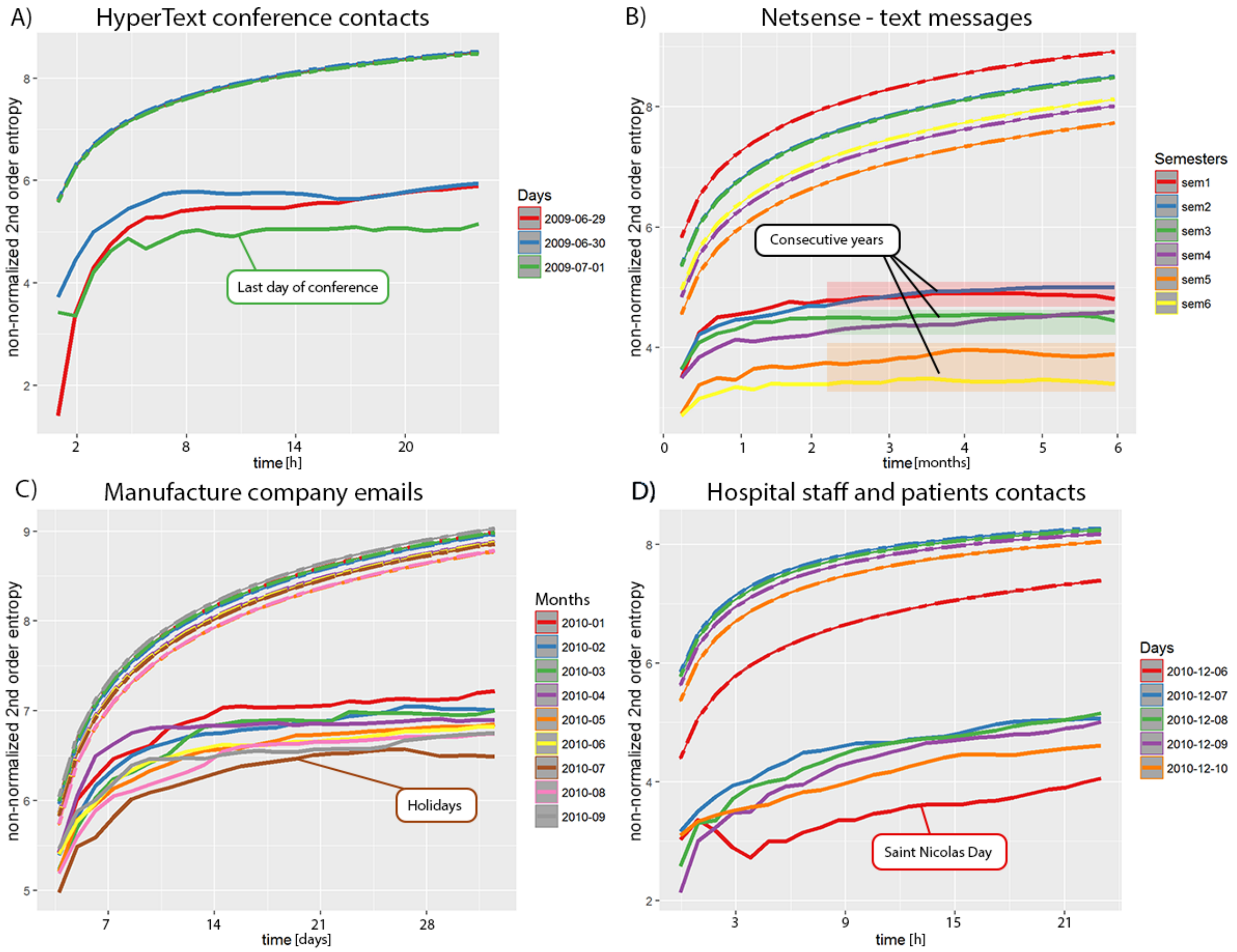


Figure 4.7: Value of non-normalized second-order entropy for all examined datasets. Solid lines refer to the real event sequences; upper dashed lines – to average values of random sequences. Each dataset is divided into parts for more convenient analysis. Parts were selected empirically. Different level of entropies for random sequences (especially for NetSense and hospital) comes from either a smaller or greater number of interacting nodes in a given period. A) In consecutive days of conference entropy of communication decreases which is especially clear for the last day of the conference. B) Students tend to be more selective in their communication in later semesters than at the beginning of studies. C) Manufacturing company employees communicate with similar dynamics over time but a decreasing tendency of the entropy can be still observed with the exception of one month probably related to the holiday period. D) Hospital staff and patients contacts show decreasing entropy over consecutive days with the exception of 6th of December, usually celebrated as Saint Nicolas Day, which may influence contact dynamics.

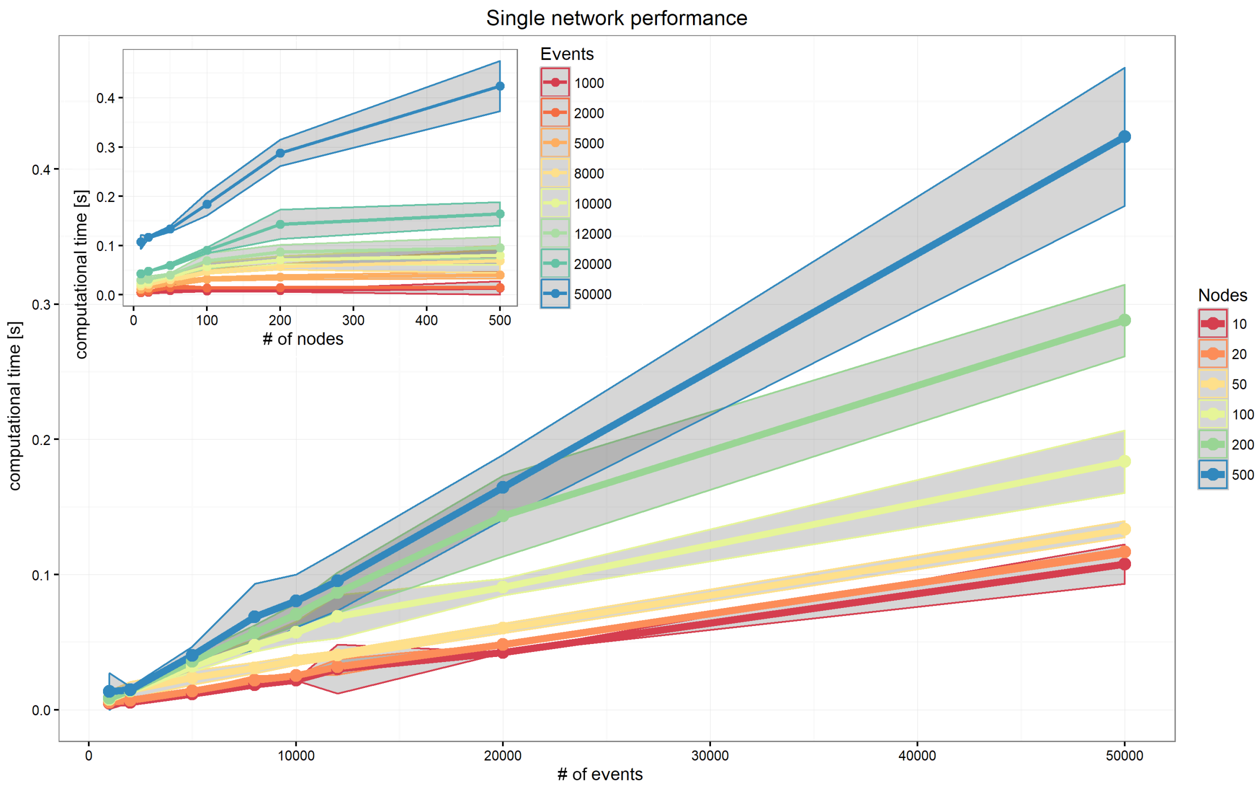


Figure 4.8: Average performance of the second-order entropy computation for the single network performed on a single computational unit. The x-axis represents the number of events in the stream and Y-axis is a computational time in seconds. Colors correspond to the number of nodes in the network. In the inset: Y-axis is the same but X-axis represents the number of nodes while colors state for the number of events. Grey areas denote standard deviations over 20 streams.

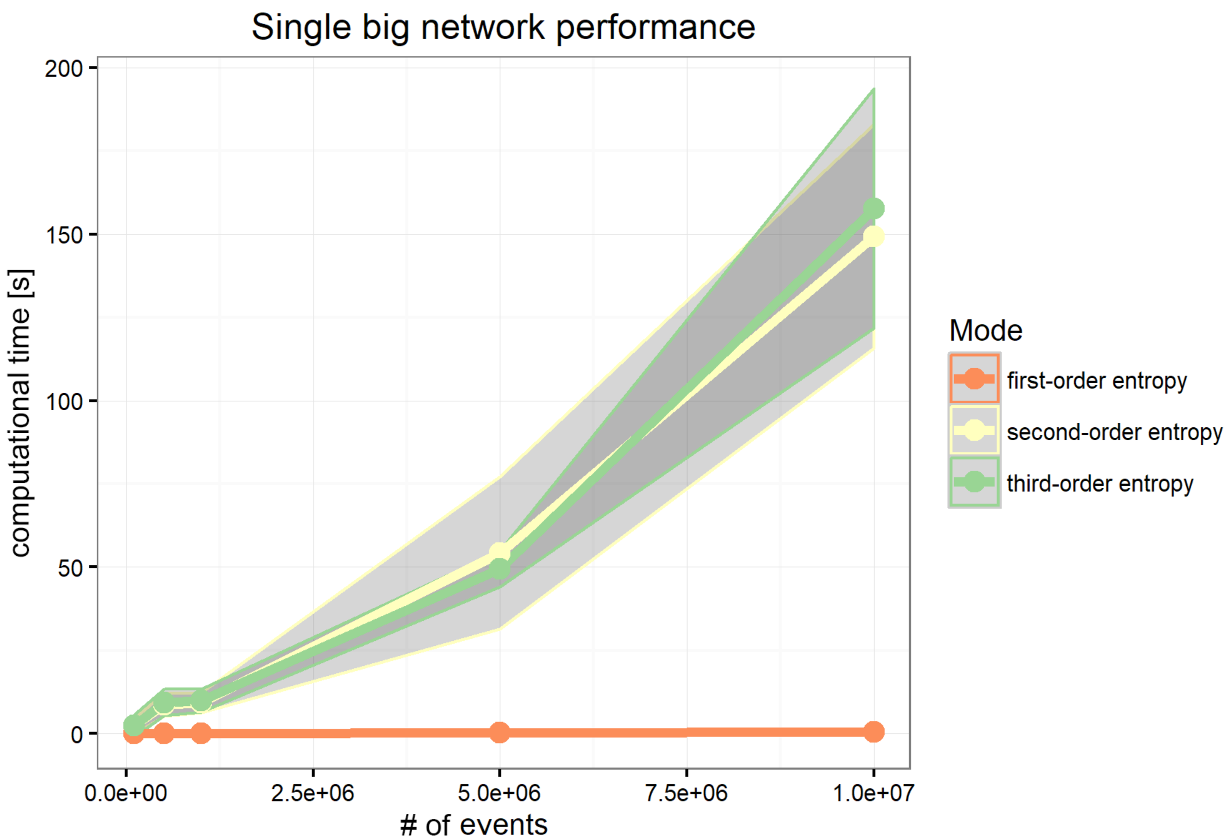


Figure 4.9: Average performance of entropy computation for large temporal networks in relation to the number of events. All examined networks have 10,000 nodes. Colors represent the type of the computed entropy: the first-order, the second-order and the third-order entropy. Grey areas depict standard deviations for 20 generated streams.

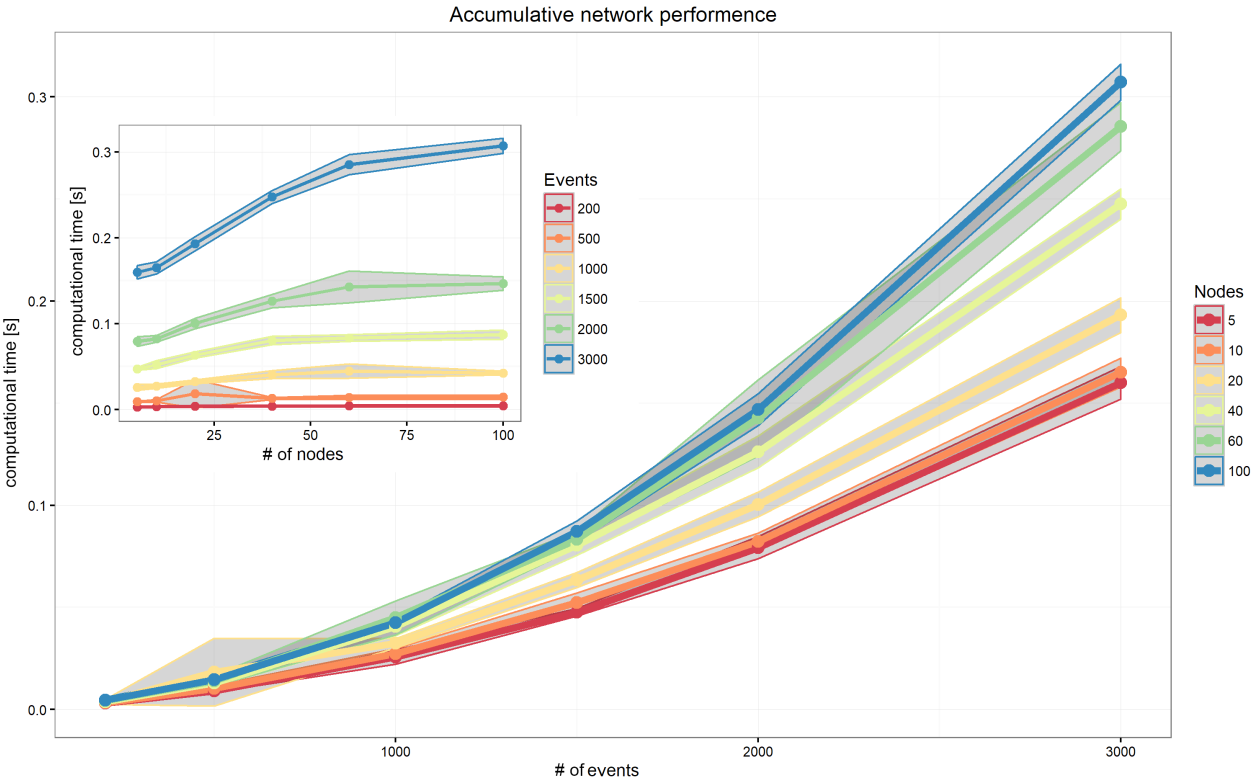


Figure 4.10: Average performance of the second-order entropy for the accumulative network performed on a single computational unit. The x-axis represents the number of events in the network while y-axis states for the processing time in seconds and colors correspond to the number of nodes in the network. In the inset, x-axis states for the number of nodes, while colors represent the number of events in the network. Grey space denotes standard deviation.

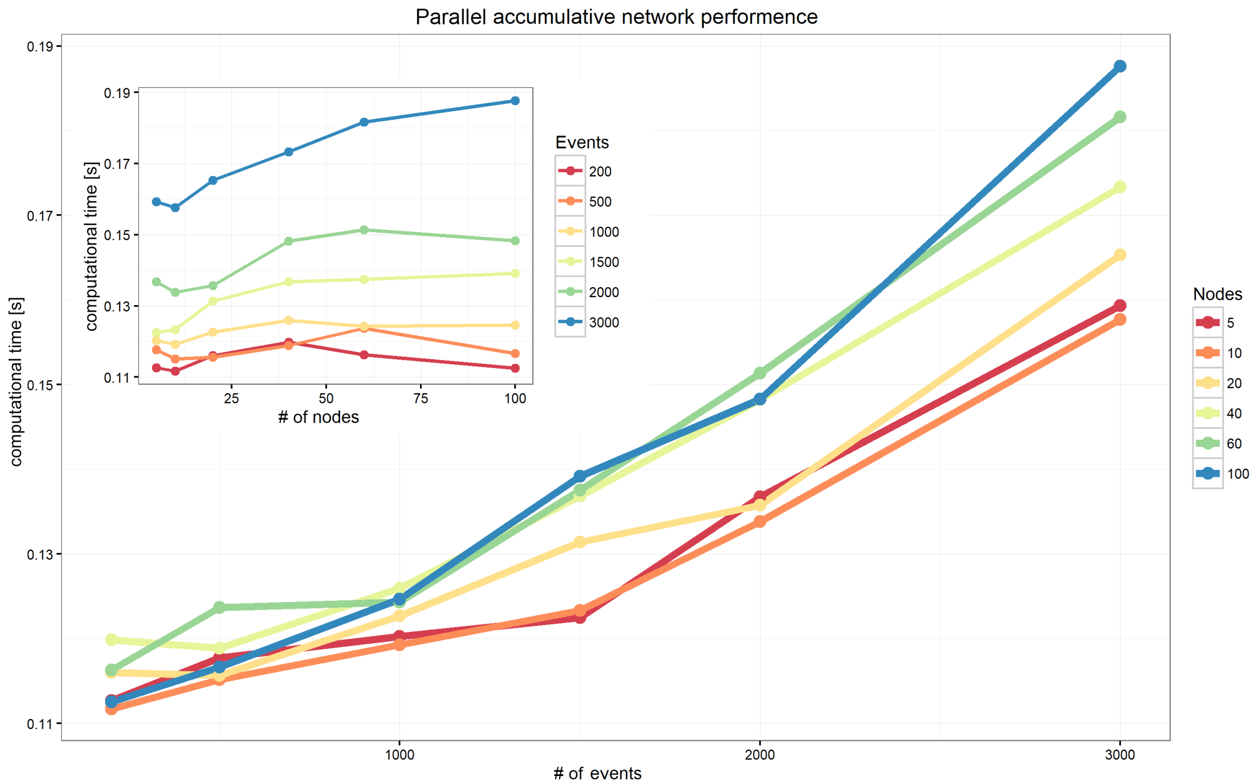


Figure 4.11: Performance of the second-order entropy computations for the accumulative network performed on 8 computational cores. Colors represent the number of nodes in the stream. In the inset, x-axis states for the number of nodes while colors distinguish the number of events in the stream.

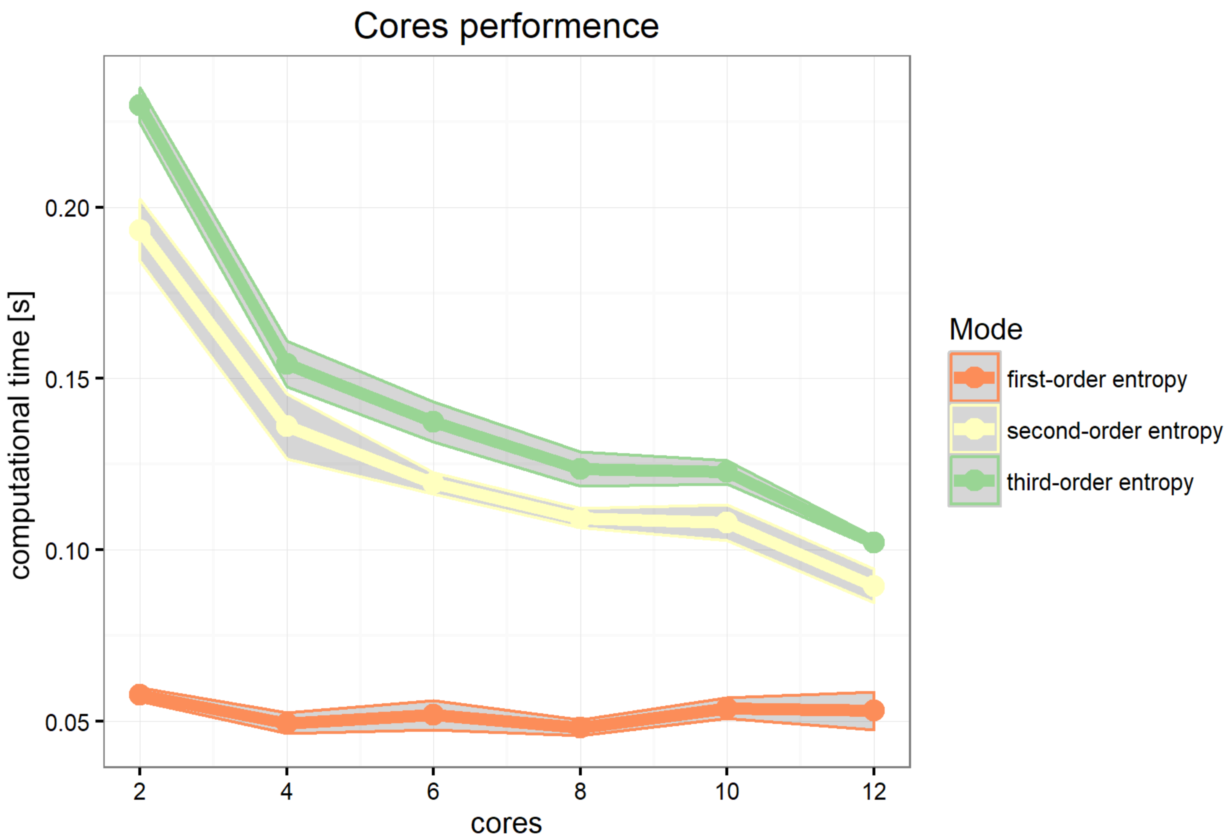


Figure 4.12: Performance of entropy computations on multiple computational units for single streams with 100 nodes and 3,000 events. The x-axis represents the number of used computational units (cores), while the y-axis corresponds to the average time in seconds. The type of the computed entropy, i.e. the first-, the second-, and the third-order are distinguished by different colors.

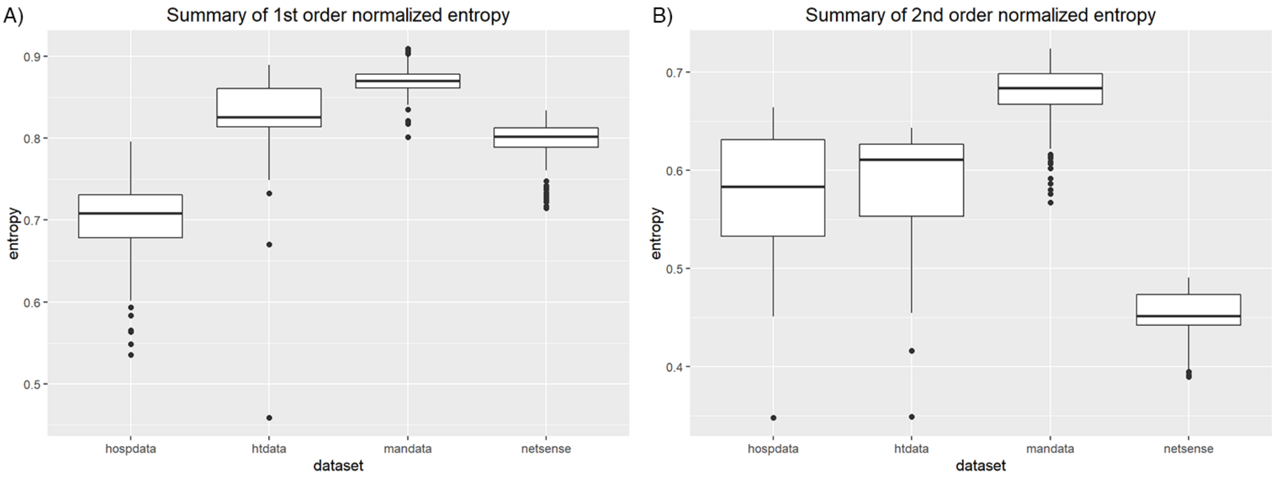


Figure 4.13: Comparison of average value of normalized entropy obtained incrementally (measured in 50 equally distributed points in dataset period): A) the first-order, B) the second-order. X-axis labels: *hospdata* - dataset of hospital face-to-face contacts, *htdata* - conference face-to-face contacts, *mandata* - email communication in manufacturing company, *netsense* - student text communication. Boxes shows median (black horizontal line) with the 1st and the 3rd quartile. Points refers to outliers and vertical lines to range of main observations.

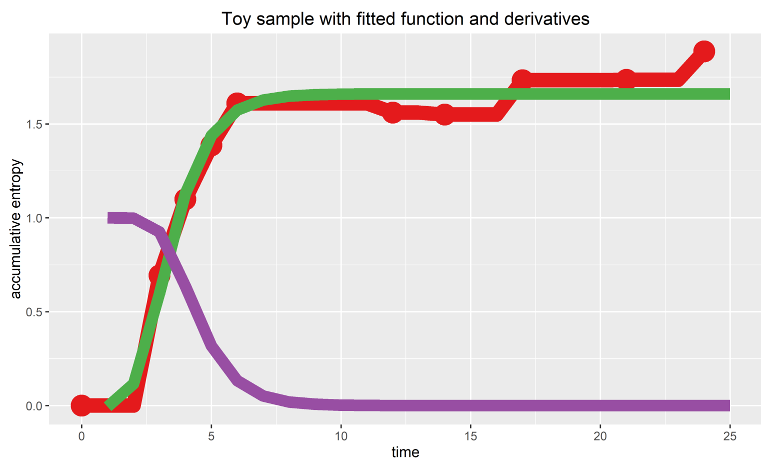


Figure 4.14: Accumulative second-order entropy value of sample temporal network with fitted logistic function and first derivative.

Chapter 5

Interdependency and applications

In this chapter, an approach to compare CogSNet and entropy-based measures is presented. Both methods proposed in the previous chapters describe an event-sequence in terms of some features. Here, I try to address the problem of whether both methods are interdependent from each other. Whatsmore, despite the fact that this work does not focus on the utilization of those methods, a couple of potential applications are presented to provide some idea in which directions those methods can be developed.

5.1 CogSNet aggregation

5.1.1 Average strength values in periods: definite integrals

As mentioned in Section 3.7, it is possible to compute an integral of CogSNet for each pair of nodes. A definite integral in a given period can be computed as an average strength of the relationship between given nodes. Please note, that the begin of the period can be given in three scenarios: 1) the period can start exactly at event time, 2) the period can start before any event or at/after a time the relationship strength drops below threshold θ and or 3) the period starts after the event and the weight still remains above the threshold. The first case corresponds to period $[t_0,t_1]$, $[t_2,t_3]$, $[t_3,t_4]$, and the following. The second case occurs for period

[t'_0,t_0] and [t_1,t_2], while the third scenario happens for [$t"_0$,t_1], see Fig. 5.1. Based on Eq.3.2, the formula for CogSNet integral can be expressed in the following way using exponential forgetting function:

$$\int w_{ij}(t)dt = \int w_{ij}(t_{ij})e^{-\lambda(t-t_{ij})} = -\frac{w_{ij}(t_{ij})e^{-\lambda(t-t_{ij})}}{\lambda} + c \tag{5.1}$$

and power forgetting function:

$$\int w_{ij}(t)dt = \int w_{ij}(t_{ij})(t-t_{ij})^{-\lambda} = w_{ij}(t_{ij})\frac{(t-t_{ij})^{(-\lambda+1)}}{-\lambda+1} + c \tag{5.2}$$

where t is a point in time, t_{ij} is a time of the last event between nodes v_i and v_j and $w_{ij}(t_{ij})$ is a weight after the last event between nodes v_i and v_j, c is the constant of integration. The above equation can be applied to calculate relation weight for a given period $[t_1, t_2]$. If the forgetting keeps the relation weight above the threshold (is not cut off) over this entire period, i.e. the course of the weight is a raw exponential function, we can use the definite integral. Note that additionally, no event may happen in this period, see the orange hatched field in Fig.3.1. In such case, the average weight (strength) of relationship, i.e. $w_{ij}(t_1, t_2)$ between node v_i and v_j in period $[t_1, t_2]$ can be easily computed as follows (using exponential forgetting function and power forgetting function, respectively):

$$w_{ij}(t_1,t_2) = \int_{t_1}^{t_2} w_{ij}(t)dt = -\frac{w_{ij}(t_{ij})e^{-\lambda(t_2-t_{ij})}}{\lambda} + \frac{w_{ij}(t_{ij})e^{-\lambda(t_1-t_{ij})}}{\lambda} = \frac{w_{ij}(t_{ij})}{\lambda}(e^{-\lambda(t_1-t_{ij})} - e^{-\lambda(t_2-t_{ij})}) \tag{5.3}$$

$$\begin{aligned} w_{ij}(t_1,t_2) = \int_{t_1}^{t_2} w_{ij}(t)dt =& w_{ij}(t_{ij})\frac{(t_2-t_{ij})^{(-\lambda+1)}}{-\lambda+1} - w_{ij}(t_{ij})\frac{(t_1-t_{ij})^{(-\lambda+1)}}{-\lambda+1} = \\ & w_{ij}(t_{ij})\frac{(t_2-t_{ij})^{(-\lambda+1)} - (t_1-t_{ij})^{(-\lambda+1)}}{-\lambda+1} \end{aligned} \tag{5.4}$$

In other cases, the definite integral of CogSNet is computed sequentially within time chunks. Each time chunk is created upon either the event or falling down weight reaching threshold θ, which may happen inside the considered period. For period [t_0,t_7] or [t''_0,t_7] in

Fig.5.1, we have seven chunks, however, there are as many as eight chunks for [t'_0,t_7] - an additional zero-valued chunk [t'_0,t_0] precedes all others. The value of definite integral of the period from t_0 to t_n equals the weighted average of weights for the smaller, component periods:

$$w_{ij}(t_0,t_n) = \frac{w_{ij}(t_0,t_1)*(t_1-t_0)+w_{ij}(t_1,t_2)*(t_2-t_1)+...+w_{ij}(t_{n-1}),t_n)*(t_n-t_{n-1})}{t_n-t_0} \tag{5.5}$$

In such case, each smaller period is computed as follows:

$$w_{ij}(t_x,t_y) = \begin{cases} 0, & \text{if there is no event } ev_{ijt_x} \text{ at } t_x \\ & \text{ and no relationship } (w_{ij}(t_x) == 0) \ , \\ \frac{w_{ij}(t_{ij})}{\lambda}(e^{-\lambda(t_x-t_{ij})}-e^{-\lambda(t_y-t_{ij})}), & \text{otherwise,} \end{cases} \tag{5.6}$$

Please note, that the exponential forgetting function is used in the above Equation 5.6, however. The power forgetting function can be used in a comparable manner.

The integral is computed for each pair of nodes using Algorithm 6. The exponential forgetting function is used there, but power forgetting function can be applied analogically.

Yet another way to look at the weight of the relationship is to consider it as a random variable. For a random variable, we can compute an average value which is, in such case, its expected value.

5.1.2 Temporal node degree based on CogSNet

A node degree - a common property of the node - can be defined for a temporal network in two ways, due to its nature: 1) for a single point of time and 2) for a period.

In the first case, it can be achieved by computing a sum of weights of relationships with other nodes at a given point of time. Then, the weight of a relationship between nodes v_i and v_j is defined in Equation 3.2. Based on this definition temporal degree of node v_i in time t can

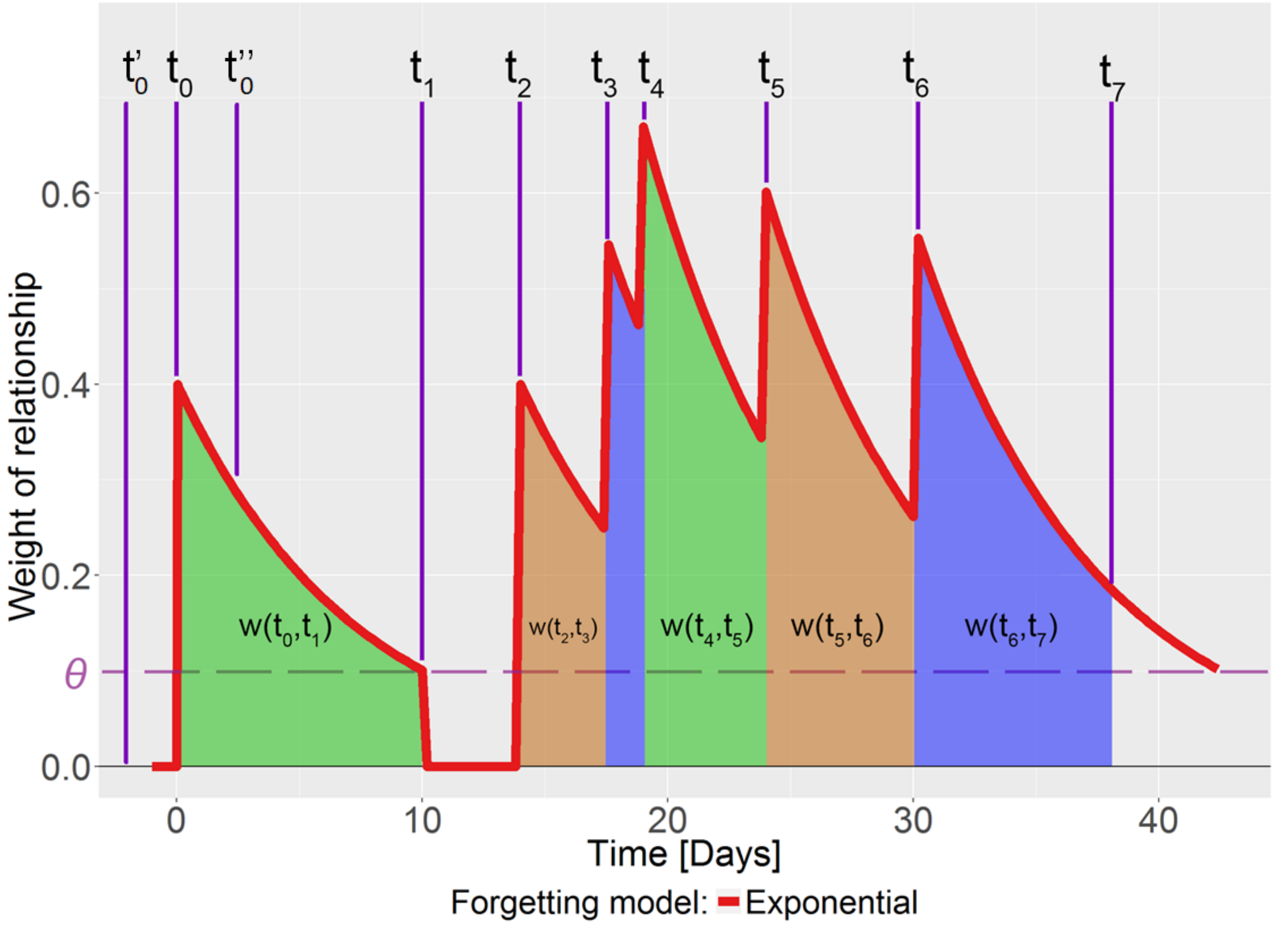


Figure 5.1: Calculation of definite integral of CogSNet for a period from t_0 to t_7. Note, that the begin of the period t_0 can be set also at t'_0 and t''_0. Each period begins with an event, except the period $[t_1,t_2]$ where t_1 is time when the weight of relationship falls below threshold level θ.

Algorithm 6 CogSNet integral algorithm

Input: ES – sequence of events ev_{ijk}
(t_{from}, t_{to}) – the begin and end of the selected period
(x, y) – nodes

Output: $integral = 0$ – definite integral of CogSNet for the exponential function in the period for nodes x, y. Replace integral in line 12, 24 and 36 for other forgetting functions, e.g. power, accordingly.

1: G – compute a weighted static graph at t_{from} using CogSNet
2: $ES_{x,y}$ – extract communication between selected nodes x,y in period $[t_{from}, t_{to}]$
3: $weighted\_integral = 0$ – define the list to store partial integrals
4: **for** each consecutive $event\_pair \in ES_{x,y}$ (their time is t_1 and t_2) **do**
5:   **if** it is the first pair of events **then**
6:     w_{xy} – extract weight from G
7:     **if** the weight w_{xy} reaches threshold θ at t' after t_{from} and before t_1, $t_{from} < t' < t_1$ **then**
8:       $dt = t' - t_{from}$
9:     **else**
10:       $dt = t_1 - t_{from}$
11:     **end if**
12:     $weighted\_integral = weighted\_integral + \frac{w_{xy}}{\lambda}(1 - e^{-\lambda * dt)}) * dt$
13:     **if** $t' < t_1$ **then**
14:       $w_{xy} = \mu$
15:     **else**
16:       $w_{xy} = \mu + w_{xy} * e^{-\lambda * dt} * (1 - \mu)$
17:     **end if**
18:   **end if**
19:   **if** the weight w_{xy} reaches threshold θ at t_2' before t_2, $t_2' < t_2$ **then**
20:     $dt = t_2' - t_1$
21:   **else**
22:     $dt = t_2 - t_1$
23:   **end if**
24:   $weighted\_integral = weighted\_integral + \frac{w_{xy}}{\lambda}(1 - e^{-\lambda * dt)}) * dt$
25:   **if** $t_2' < t_2$ **then**
26:     $w_{xy} = \mu$
27:   **else**
28:     $w_{xy} = \mu + w_{xy} * e^{-\lambda * dt} * (1 - \mu)$
29:   **end if**
30:   **if** it is the last pair of events **then**
31:     **if** the weight w_{xy} reaches threshold θ at t_{to}' before t_{to}, $t_{to}' < t_{to}$ **then**
32:       $dt = t_{to}' - t_2$
33:     **else**
34:       $dt = t_{to} - t_2$
35:     **end if**
36:     $weighted\_integral = weighted\_integral + \frac{w_{xy}}{\lambda}(1 - e^{-\lambda * dt)}) * dt$
37:   **end if**
38: **end for**
39: $integral = weighted\_integral/(t_{to} - t_{from})$
40: **return** $integral$

be defined as follows:

$$d^T(v_i, t) = \sum_{v_j \in V, j \neq i} w_{ij}(t) \tag{5.7}$$

Every edge that connects node v_i with another node has weight w_{ij} greater than threshold θ or is equal 0 if there is no connection.

In the second case, node temporal degree can be defined in a time period using presented in the previous section CogSNet integral. Based on relation strength between nodes v_i and v_j in a period, defined in Equation 5.5, temporal degree of node v_i in a given period (t_1, t_2) can be defined as follows:

$$d^T(v_i, t_1, t_2) = \sum_{v_j \in V, j \neq i} w_{ij}(t_1, t_2) \tag{5.8}$$

Where $w_{ij}(t_1, t_2)$ is an integral on the relation weight in the period (t_1, t_2) computed using either forgetting or power forgetting function. Both approaches can be used to describe some property of a temporal network. On the one hand, a degree in a point of time can be used to describe node its importance in the network at a specific point of time. On the other hand, a degree in a period can describe the average importance of the node in the given period. Due to limited resources, no experimental studies were carried out on temporal degree. It is envisaged as an interesting direction for future work.

5.1.3 CogSNet parameters vs. second-order entropy measure

To compare entropy-based measure and CogSNet, two sample artificial event sequences (6 nodes, 20 events) are manually created:

1. with maximum second-order entropy and
2. with low second-order entropy

For the first one, it is achieved by keeping the same quantity of all unique events. The second one is done in the opposite way - the most of events occur once or twice while just a few events have multiple occurrences. Next, CogSNet integrals 5.3 are also computed for both event

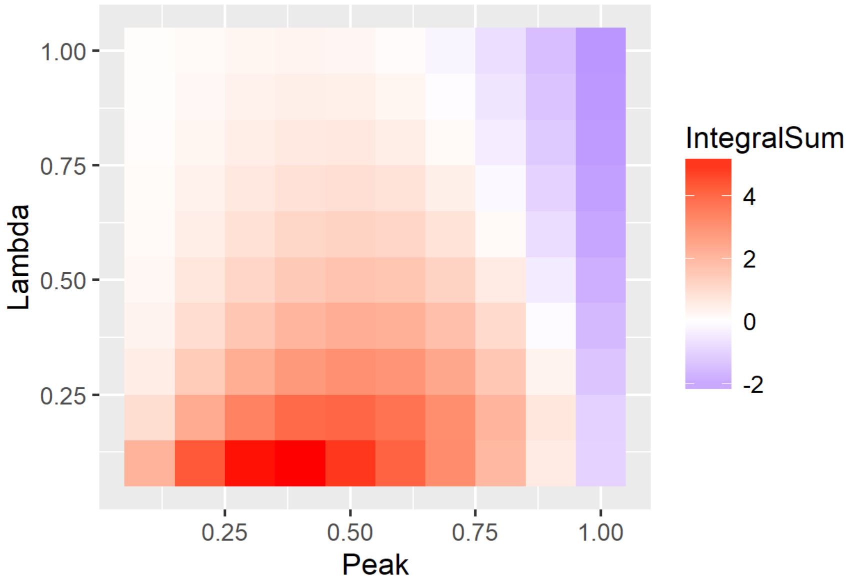


Figure 5.2: Heatmat of differences between CogSNet integral sum for low-entropy sequence and high-entropy sequence.

sequences (for the whole period) and then the sum of integrals of all node pairs is calculated. Afterward, the difference between integral sums is evaluated. Since the shape of the forgetting function depends on a couple of parameters, integrals are independently computed for a set of parameters. In Fig.5.2, the difference between CogSNet integrals for low and high entropy and various parameters is presented. It can be observed that the sum of integrals depends heavily on CogSNet parameters λ and μ. Moreover, for different parameters, greater integral can be achieved both for low and high entropy network which not directly shows that there is no correlation between average relation weight (CogSNet integral) and entropy value (second-order entropy).

5.1.4 Information equivalence between CogSNet and entropies

Experimental setup

To bring a better understanding of the presented models, yet another experiment is performed. In a small-scale sample, starting from random event sequence, events are rewired in two scenarios: 1) constant entropy and 2) constant CogSNet integral. If there is no possible rewire that cannot preserve the desired fixed value, the one is applied which results in the smallest change.

Results and discussion

Results of the described experiment are presented in Fig. 5.3. The constant entropy can be observed in the bottom plot in Fig 5.3A, while the corresponding values of CogSNet integrals are shown in the top plot. Changes in integral values have no clear explanations - they happen only for few iterations (rewirings). Moreover, in this case, the integrals only decrease, but in general, they can increase as well. In Fig 5.3B, the results of the second scenario are presented. The value of CogSNet integral remains constant throughout all iterations except one. Nevertheless, entropy value in every iteration - decreases almost linearly. It rises only in the last iteration of the simulation.

These experiments reveal that it is possible to get different values of one measure while keeping the other unchanged. This is a sign that entropy and CogSNet do not have to be closely related to each other. Along with the experiment presented in Section 5.1.3, it is clear that both entropy-based measures and CogSNet describe two distinct features of the temporal network. A deeper understanding of relationships between them two as well as the nature of information about event sequence they carry appears to be yet another direction for further investigations.

5.2 Entropies of social communities

5.2.1 Social communities

Many social groups can be distinguished in social networks in terms of, e.g., age, gender, religion, occupation, political option. These features (labels) may be given directly in the available data, i.e. we can have each node labeled with appropriate features. In many cases it is expected, that different social groups will communicate in different manners. Young people in schools may interact very dynamically, while workers in a company maybe not that much.

In the following experiment, different groups in the temporal social network are examined

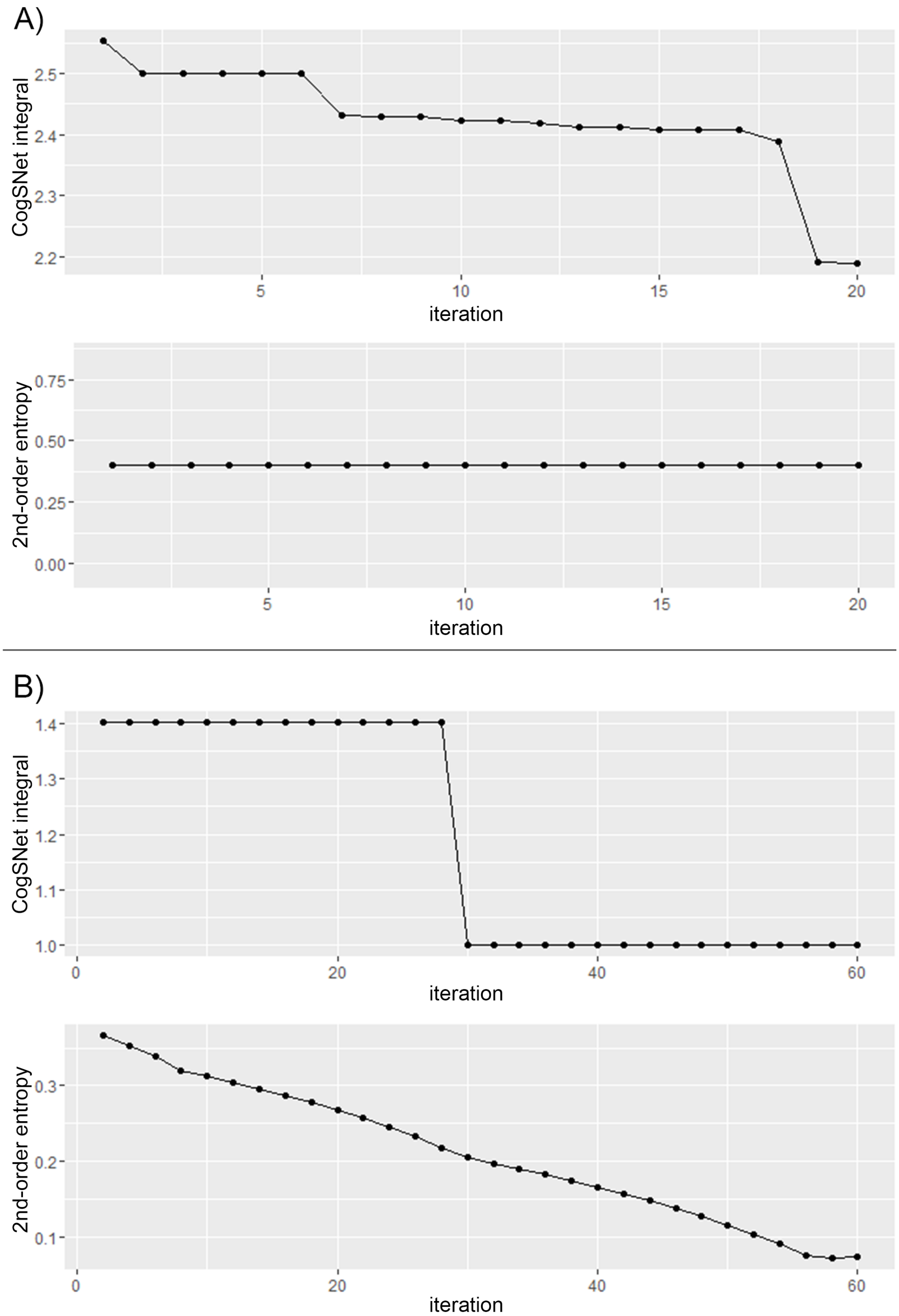


Figure 5.3: Entropy and CogSNet integral over iterations of rewirering focused on: A) constant entropy B) constant CogSNet integral.

in terms of the entropy of their communication to test whether the communication in those communities differs from each other. The dataset used is Netsense - students' network presented in the previous sections. This dataset contains also some additional features about students, namely gender, age, income, political preference, etc. For the purpose of the experiment, communication in the selected groups is extracted in the following way:

1. homogeneous communication: only communication inside the group is considered, i.e. all events with both nodes belonging to the given group and
2. heterogeneous communication: communication with other groups, i.e. all events that only one node belongs to the given group.

For such extracted communication, entropy-based measures are computed. Social communities distinguished by the following features are considered: gender and parents income.

5.2.2 Results and discussion

Some results of this experiment are presented in Fig. 5.4. In general, higher entropy value may indicate that events are more evenly distributed between nodes, which in terms of human communication can be understood that people have no special preferences with whom they want to interact. Lower entropy, whereas, may indicate that people have a limited group of friends, with whom they interact more frequently. Differences between groups are clearly visible and may help to understand human behavior, e.g., in Fig. 5.4A the group of students with a high income has low entropy comparing to other groups, while the group with low income in the opposite. It may lead to a similar conclusion that is made in [74], that people with low income have a tendency to communicate much more diverse than people with high income. A possible explanation of such observation is that poorer people have the motivation to communicate with others to increase the chance of profitable contacts that may improve their economic situation. Another interesting observation is homogeneous and heterogeneous communication in gender groups. People with the same sex tend to have a more limited group of interlocutors, while interacting with opposite sex communication may be more equally distributed. An opposite observation can be made in income groups - communication inside income group is less selective than in communication outside the group.

Such a difference in entropy between social groups in the temporal network can be exploited in clustering or in group recognition.

5.3 Conclusions

In this chapter, some basic comparisons entropy-based measures and CogSNet model are made. Both approaches provide information about the event-sequence but they are not necessarily closely related. Simulations show that it is possible to obtain different values of one measure while keeping the other one fixed. The potential application of entropy-based measures to an analysis of social communities is also appointed.

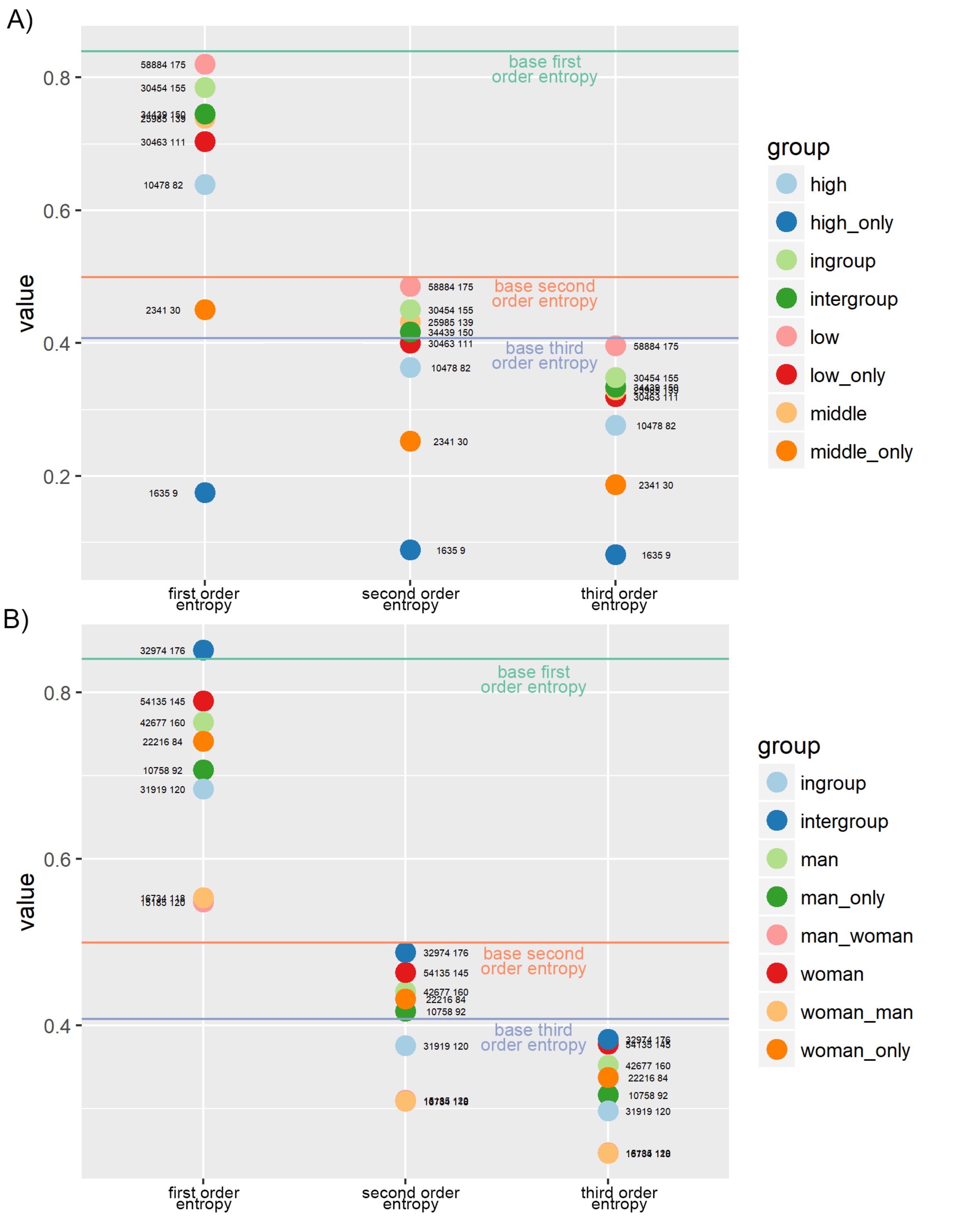


Figure 5.4: Entropy of different groups in Netsense dataset. A) groups based on income of students parents: *high/middle/low* - refers to events where one of interlocutors have high/middle/low income, *high_only/middle_only/low_only* - events only between students with high/middle/low income, *ingroup* - communication between students in the same income group, *intergroups* - events between students from different groups. B) groups based on gender: *man/woman* - refers to events where one of interlocutors is man/woman, *man_only/woman_only* - events only between men/women, $woman_man/man_woman$ - events from woman to man and man to woman respectively, assuming direct communication, *ingroup* - communication between students with the same gender, *intergroups* - events between students with opposite gender.

Chapter 6

Summary

This dissertation is focused on temporal networks based on events - highly granular pieces of information. High resolution of information requires more sophisticated methods in comparison to static networks, in order to preserve such temporal information, from one hand, and to be convenient to use form the other hand. In this work, I proposed a new temporal network model that satisfies this requirement, which is the first major contribution. CogSNet is a network model that is able to capture and simulate the process of human cognition. Conducted experiments show, that CogSNet can simulate real network communication and provide information about relations between network nodes better than baseline methods. The success of the proposed method lies in an applied continuous forgetting mechanism that is an important part of human cognition. The continuous nature of CogSNet allows applying analytic methods to compute an average weight of the relationship between individuals in the network with the integral of weight over some period of time.

The second important contribution of this work is a new set of entropy-based measures for temporal network dynamics. Proposed measures applied to real-world networks reveals an interesting characteristic of human communication - the entropy of communication is decreasing over time. That implies that people tend to be more selective in their communication over time, which can be explained that while people getting to know each other their discover their preferences about with whom they want to communicate. For proposed measures, algorithms for iterative and parallel computation are also designed, providing an efficient computational method for application in large networks. Interestingly, proposed measures show some potential in group recognition in temporal networks by showing, that different social groups in real-word temporal networks have different entropy.

In this work, I also make an approach to compare proposed methods in terms of a numerical description of a temporal network. CogSNet integral and entropy-based measures are able to provide some information about event-sequence, however their cover different aspects.

6.1 Contributions

Following contribution of my work can be enumerated:

- Design of a new temporal social network model that simulates human cognition.
- Examination of proposed model properties
- Validation of proposed model on real-world temporal network
- Definition of average strength of relationship in temporal networks
- Definition of entropy for temporal network given as event sequence
- Design of maximum and normalized entropy-based measure for temporal networks
- Algorithms design for iterative and parallel computation of entropy for single network and accumulative network
- Comparison and analysis of proposed approaches

6.2 Future Work

Many directions covered in this work showed the potential for future development. First, entropy-based measures showed an ability to describe some aspects of temporal network dynamics. However, it would be worthwhile to find applications of those measures in network dynamics prediction. Since one of the findings of this work is that entropy is decreasing over time in temporal social networks, those networks have the highest chance of benefit from the proposed measures in prediction attempts. Next, some potential is recognized of entropy-based measures in solving problems like detection of social communities from dynamic data about human activities. The hypothesis is that entropy is able to distinguish different groups in the event sequence since the groups may have a different dynamic profile of interactions (different entropy levels), e.g. within hospital staff members and separately among patients. Another direction that would be worth to develop is application of CogSNet to large temporal social networks to find optimal parameters of the model. It would be very useful to assign some parts of the parameter space to a particular communication method, e.g., phone calls or emails. Previously mentioned approaches of temporal network description - entropy based measurements and the CogSNet integral - are briefly analyzed in terms of mutual dependency. Deeper studies may bring a better understanding of those two approaches in real-world temporal networks.

6.3 Dissemination

6.3.1 Research projects

The presented work was supported by the following projects I participated in:

- *Models, Methods and Algorithms of Computational Network Science*, research project funded by National Science Centre (OPUS), led by Prof. Przemysław Kazienko, 2017-2020
- *RENOIR - Reverse EngiNeering of sOcial Information pRocessing*, H2020-MSCA-RISE-2015, no. 691152, main contractor, 2016-2019
- Diffusion of Information in Temporal Social Networks, research project funded by National Science Centre (SONATA), led by Dr. Radosaw Michalski, 2016-2020

As well by already terminated projects:

- ENGINE - European research centre of Network intelliGence for INnovation Enhancement, FP7-REGPOT-2012-2013-1, Coordination and Support Action, no. 316097, 2013-2016
- Machine learning methods for complex networks, research project funded by National Science Centre (OPUS), 2014-2017
- TRANSFoRm - Translational Research and Patient Safety in Europe, ICT-2011.11.3: Supplements to Strengthen Cooperation in ICT R D in an Enlarged European Union, FP7 247787, 2011-2015
- Mapping and analysis of the collaboration and knowledge exchange network of key experts involved in a critical business process in KGHM Polska Mied S.A., research project funded by KGHM Polska Miedź S.A. (in collaboration with lome.pl), 2015
- Preparation and delivery of algorithms and mathematical models for xDSL services prequalification, TP S.A./Orange (Polish Telecom), 2012-2013

6.3.2 Collaborators and advisors

In my work I collaborated with following researchers:

- Boleslaw K. Szymanski - computer scientist and mathematician, Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY, USA
- Christian Lebiere - cognitivist, Carnegie Mellon University, Pittsburgh, PA, USA
- Omar Lizardo - sociologist, University of Notre Dame, Notre Dame, IN, USA; University of California Los Angeles, CA, USA
- Jean Karl. Soler - medical doctor, Mediterranean Institute of Primary Care 19, Triq ir-Rand, Attard, Malta

- Derek Corrigan - computer scientists, Department of General Practice, HRB Centre for Primary Care Research, Beaux Lane House, Lower Mercer Street, Dublin, Ireland
- Tomasz Kajdanowicz - computer scientist, Faculty of Computer Science and Management, Department of Computational Intelligence, Wrocław University of Science and Technology, Wrocław, Poland
- Roxana Danger - computer scientists, Imperial College London, London, UK
- Brendan Delaney - medical doctor, Wolfson Chair of General Practice, Kings College London, Capital House, Guys Hospital, London, England
- Maciej Piasecki - computer scientist, linguist, Faculty of Computer Science and Management, Department of Computational Intelligence, Wrocław University of Science and Technology, Wrocław, Poland
- Łukasz Augustyniak - computer scientist, Faculty of Computer Science and Management, Department of Computational Intelligence, Wrocław University of Science and Technology, Wrocław, Poland
- Włodzimierz Tuligłowicz - computer scientist, Faculty of Computer Science and Management, Wrocław University of Science and Technology, Wrocław, Poland
- Adrian Popiel - computer scientist, Faculty of Computer Science and Management, Department of Computational Intelligence, Wrocław University of Science and Technology, Wrocław, Poland
- Andrzej Misiaszek - computer scientist, Faculty of Computer Science and Management, Wrocław University of Science and Technology, Wrocław, Poland

6.3.3 Research visits

The studies presented in this thesis was partially carried out during the following research visits at US universities:

- 1.05.2016 - 31.05.2016 - research visit at Rensselaer Polytechnic Institute, Troy, NY, USA
- 4.01.2017 - 4.04.2017 - research visit at Rensselaer Polytechnic Institute, Troy, NY, USA
- 22.10.2017 - 22.12.2017 - research visit at Rensselaer Polytechnic Institute, Troy, NY, USA
- 20.01.2018 - 19.02.2018 - research visit at University of California Davis, CA, USA

6.3.4 Scientific papers

I am a co-author of the following scientific papers:

- **Marcin Kulisiewicz**, Przemysław Kazienko, Bolesław K. Szymański, Radosław Michalski *Entropy Measures of Human Communication Dynamics.* Scientific Reports, volume 8, Article number: 15697 (2018).**IF=4.609**
- Jean Karl. Soler, Derek Corrigan, Przemysław Kazienko, Tomasz Kajdanowicz, Roxana Danger, **Marcin Kulisiewicz**, Brendan Delaney *Evidence-based rules from family practice to inform family practice; the learning healthcare system case study on urinary tract infections.* BMC Family Practice. 2015, vol. 16, nr 5, p. 1-6. **IF=2.16**
- Radosław Michalski, Bolesław K. Szymański, Przemysław Kazienko, Christian Lebiere, Omar Lizardo, **Marcin Kulisiewicz** (2018). *Social Networks through the Prism of Cognition.* arXiv preprint arXiv:1806.04658.
- **Marcin Kulisiewicz**, Przemysław Kazienko, Radosław Michalski *Parallel Entropy Computation for Interaction Stream Dynamics* - in reviews in the JCR-listed journal,
- Łukasz M. Augustyniak, Tomasz Kajdanowicz, Przemysław Kazienko, **Marcin Kulisiewicz**, Włodzimierz J. Tuligłowicz, *An approach to sentiment analysis of movie reviews: lexicon based vs. classification.* Hybrid artificial intelligence systems : 9th international conference, HAIS 2014, Salamanca, Spain, June 11-13, 2014 : proceedings / Marios Polycarpou

[i in.] (eds.). Cham [i in.] : Springer, cop. 2014. s. 168-178. (Lecture Notes in Computer Science. Lecture Notes in Artificial Intelligence, ISSN 0302-9743; vol. 8480)

- Andrzej Misiaszek, Przemysław Kazienko, **Marcin Kulisiewicz**, Łukasz M. Augustyniak, Włodzimierz J. Tuligłowicz, Adrian Popiel, Tomasz Kajdanowicz, *Belief propagation method for word sentiment in WordNet 3.0.* Intelligent Information and Database Systems : 6th Asian Conference, ACIIDS 2014, Bangkok, Thailand, April 7-9, 2014 : proceedings. Pt. 2 / Ngoc Thanh Nguyen [i in.] (Eds.). Cham [i in.] : Springer, cop. 2014. s. 263-272. (Lecture Notes in Computer Science. Lecture Notes in Artificial Intelligence, ISSN 0302-9743; vol. 8398)

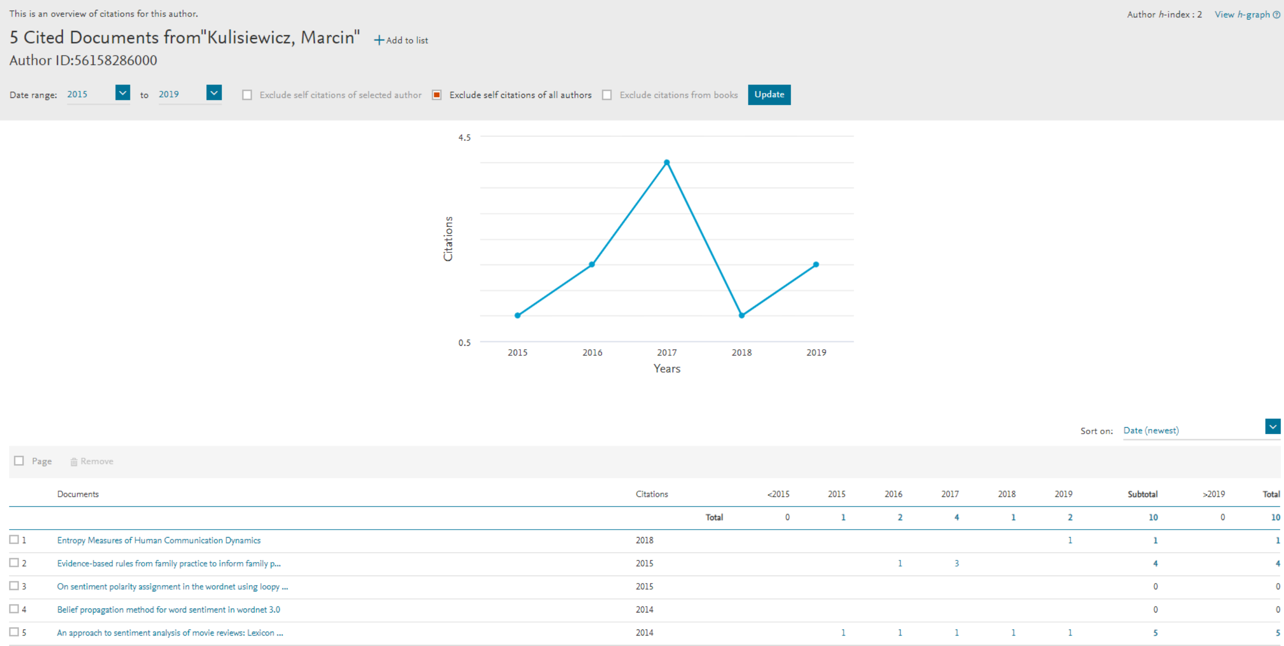
- **Marcin Kulisiewicz**, Tomasz Kajdanowicz, Przemysław Kazienko, Maciej Piasecki *On sentiment polarity assignment in the Wordnet using Loopy Belief Propagation.* Hybrid artificial intelligent systems : 10th International Conference, HAIS 2015, Bilbao, Spain, June 22-24, 2015 : proceedings / Enrique Onieva [i in.] (eds.). Cham [i in.] : Springer, cop. 2015. s. 451-462. (Lecture Notes in Computer Science. Lecture Notes in Artificial Intelligence, ISSN 0302-9743; vol. 9121)

In Fig. 6.1 citations without self-citation of my publications are presented.

6.3.5 International conferences

I actively attended the following international conferences:

- HAIS 2015, 10th International Conference, Bilbao, Spain, oral presentation, publication in proceedings
- Conference on Complex Systems 2015, Phoenix, AZ, USA, oral presentation
- NetSci-X 2016, Wroclaw, Poland. Oral presentation
- Poster at NetSci 2018, Paris, France. Kulisiewicz M., Kazienko P., Michalski R., Szymański B.K.: Phenomena of Entropy in Temporal Networks



	Documents	Citations		<2015	2015	2016	2017	2018	2019	Subtotal	>2019	Total
			Total	0	1	2	4	1	2	10	0	10
1	Entropy Measures of Human Communication Dynamics	2018							1	1		1
2	Evidence-based rules from family practice to inform family p...	2015				1	3			4		4
3	On sentiment polarity assignment in the wordnet using loopy ...	2015								0		0
4	Belief propagation method for word sentiment in wordnet 3.0	2014								0		0
5	An approach to sentiment analysis of movie reviews: Lexicon ...	2014			1	1	1	1	1	5		5

Figure 6.1: Citation overview of author publications excluding self citation of all authors. Source: Scopus.com

- Poster at ICCSS 2016 - 2016 International Conference on Computational Social Science. Evanston, IL, USA (2016). Michalski R., Kazienko P., Kulisiewicz M.: Core Nodes as the Influencers in Temporal Social Networks.

Acknowledgements

I would like to express my very great appreciation to my supervisors Prof. Przemysław Kazienko and Dr. Radosław Michalski for their patient guidance, enthusiastic support, great ideas, inspiration, useful critiques of this research work, protection and friendship. I am particularly grateful for the assistance given by the whole team of Data Science Group for valuable discussions and remarks, and great atmosphere in the office. I would like to offer my special thanks to Dr. Tomasz Kajdanowicz who inspired me to begin PhD studies. I would like to thank my family, especially my beloved wife for support, patient, encouragement and love that gave me the strength to overcome difficulties during my work. I would also like to extend my thanks to my friends for support and opportunities to take a fresh look at my work.

The RENOIR project - Reverse EngiNeering of sOcial Information pRocessing - leading to this publication has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 691152. The work was partially supported by European Union, the ENGINE grant, agreement no 316097 (FP7). This work was supported by the National Science Centre, Poland, projects no. 2015/17/D/ST6/04046 and 2016/21/B/ST6/01463. The calculations were carried out in Wroclaw Centre for Networking and Supercomputing (http://www.wcss.wroc.pl), grant No 177.

Bibliography

[1] J.-w. Ahn, M. Taieb-Maimon, A. Sopan, C. Plaisant, and B. Shneiderman. Temporal visualization of social network dynamics: Prototypes for nation of neighbors. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, pages 309–316. Springer, 2011.

[2] R. Analytics and S. Weston. doparallel: Foreach parallel adaptor for the parallel package. *R package version*, 1(8), 2014.

[3] J. R. Anderson, D. Bothell, C. Lebiere, and M. Matessa. An integrated theory of list memory. *Journal of Memory and Language*, 38(4):341–380, 1998.

[4] D. Antonakaki, S. Ioannidis, and P. Fragopoulou. Utilizing the average node degree to assess the temporal growth rate of twitter. *Social Network Analysis and Mining*, 8(1):12, 2018.

[5] A. D. Baddeley, M. Kopelman, and B. A. Wilson. *The essential handbook of memory disorders for clinicians*. John Wiley & Sons, 2004.

[6] V. Batagelj and S. Praprotnik. An algebraic approach to temporal network analysis based on temporal quantities. *Social Network Analysis and Mining*, 6(1):28, May 2016.

[7] S. Bayhan, E. Hyytiä, J. Kangasharju, and J. Ott. Analysis of hop limit in opportunistic networks by static and time-aggregated graphs. In *Communications (ICC), 2015 IEEE International Conference on*, pages 3287–3292. IEEE, 2015.

[8] A. R. Benson, D. F. Gleich, and J. Leskovec. Higher-order organization of complex networks. *Science*, 353(6295):163–166, 2016.

[9] T. Y. Berger-Wolf and J. Saia. A framework for analysis of dynamic social networks. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 523–528. ACM, 2006.

[10] B. Blagoev, S.-E. Ong, I. Kratchmarova, and M. Mann. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nature biotechnology*, 22(9):1139, 2004.

[11] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin. The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1–122, 2014.

[12] D. Braha and Y. Bar-Yam. Time-dependent complex networks: Dynamic centrality, dynamic motifs, and cycles of social interactions. In *Adaptive Networks*, pages 39–50. Springer, 2009.

[13] C. Cattuto, W. Van den Broeck, A. Barrat, V. Colizza, J.-F. Pinton, and A. Vespignani. Dynamics of person-to-person interactions from distributed rfid sensor networks. *PloS one*, 5(7):e11596, 2010.

[14] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of human mobility on opportunistic forwarding algorithms. *IEEE Transactions on Mobile Computing*, 6(6):606–620, 2007.

[15] A. Chaintreau, A. Mtibaa, L. Massoulie, and C. Diot. The diameter of opportunistic mobile networks. In *Proceedings of the 2007 ACM CoNEXT conference*, page 12. ACM, 2007.

[16] E. Cheng, J. W. Grossman, and M. J. Lipman. Time-stamped graphs and their associated influence digraphs. *Discrete Applied Mathematics*, 128(2-3):317–335, 2003.

[17] A. Clauset and N. Eagle. Persistence and periodicity in a dynamic proximity network. *arXiv preprint arXiv:1211.7343*, 2012.

[18] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas. Mathematical formulation of multilayer networks. *Physical Review X*, 3(4):041022, 2013.

[19] N. Eagle and A. Pentland. Reality mining: sensing complex social systems. *Personal and ubiquitous computing*, 10(4):255–268, 2006.

[20] N. Eagle, A. S. Pentland, and D. Lazer. Inferring friendship network structure by using mobile phone data. *Proceedings of the national academy of sciences*, 106(36):15274–15278, 2009.

[21] K. T. Eames and M. J. Keeling. Monogamous networks and the spread of sexually transmitted diseases. *Mathematical biosciences*, 189(2):115–130, 2004.

[22] J.-P. Eckmann, E. Moses, and D. Sergi. Entropy of dialogues creates coherent structures in e-mail traffic. *Proceedings of the National Academy of Sciences*, 101(40):14333–14337, 2004.

[23] N. B. Ellison, C. Steinfield, and C. Lampe. The benefits of facebook friends: social capital and college students use of online social network sites. *Journal of Computer-Mediated Communication*, 12(4):1143–1168, 2007.

[24] N. Fefferman and K. Ng. How disease models in static networks can fail to approximate disease in dynamic networks. *Physical Review E*, 76(3):031919, 2007.

[25] J. Feigenbaum, S. Kannan, A. McGregor, S. Suri, and J. Zhang. On graph problems in a semi-streaming model. *Theoretical Computer Science*, 348(2-3):207–216, 2005.

[26] A. Ferreira. On models and algorithms for dynamic communication networks: The case for evolving graphs. In *In Proc. ALGOTEL*. sn, 2002.

[27] A. Fujita, J. R. Sato, H. M. Garay-Malpartida, P. A. Morettin, M. C. Sogayar, and C. E. Ferreira. Time-varying modeling of gene expression regulatory networks using the wavelet dynamic vector autoregressive method. *Bioinformatics*, 23(13):1623–1630, 2007.

[28] W. Gao, Q. Li, B. Zhao, and G. Cao. Multicasting in delay tolerant networks: a social network perspective. In *Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing*, pages 299–308. ACM, 2009.

[29] K.-I. Goh and A.-L. Barabási. Burstiness and memory in complex systems. *EPL (Europhysics Letters)*, 81(4):48002, 2008.

[30] P. Grindrod, M. C. Parsons, D. J. Higham, and E. Estrada. Communicability across evolving networks. *Physical Review E*, 83(4):046120, 2011.

[31] R. Groenevelt, P. Nain, and G. Koole. The message delay in mobile ad hoc networks. *Performance Evaluation*, 62(1-4):210–228, 2005.

[32] E. Hernández-Orallo, J. C. Cano, C. T. Calafate, and P. Manzoni. New approaches for characterizing inter-contact times in opportunistic networks. *Ad Hoc Networks*, 52:160–172, 2016.

[33] H. W. Hethcote and P. Van den Driessche. An sis epidemic model with variable population size and a delay. *Journal of mathematical biology*, 34(2):177–194, 1995.

[34] P. Holme. Network dynamics of ongoing social relationships. *EPL (Europhysics Letters)*, 64(3):427, 2003.

[35] P. Holme. Epidemiologically optimal static networks from temporal network data. *PLoS computational biology*, 9(7):e1003142, 2013.

[36] P. Holme. Analyzing temporal networks in social media. *Proceedings of the IEEE*, 102(12):1922–1933, 2014.

[37] P. Holme. Information content of contact-pattern representations and predictability of epidemic outbreaks. *Scientific reports*, 5:14462, 2015.

[38] P. Holme. Modern temporal network theory: a colloquium. *The European Physical Journal B*, 88(9):234, 2015.

[39] P. Holme, C. R. Edling, and F. Liljeros. Structure and time evolution of an internet dating community. *Social Networks*, 26(2):155–174, 2004.

[40] P. Holme and F. Liljeros. Birth and death of links control disease spreading in empirical contact networks. *Scientific reports*, 4:4999, 2014.

[41] P. Holme and J. Saramäki. Temporal networks. *Physics reports*, 519(3):97–125, 2012.

[42] P. Holme and J. Saramäki. *Temporal networks*. Springer, 2013.

[43] K. Hoppe and G. Rodgers. Mutual selection in time-varying networks. *Physical Review E*, 88(4):042804, 2013.

[44] L. Isella, J. Stehlé, A. Barrat, C. Cattuto, J.-F. Pinton, and W. Van den Broeck. What's in a crowd? analysis of face-to-face behavioral networks. *Journal of theoretical biology*, 271(1):166–180, 2011.

[45] J. Jankowski, P. Kazienko, B. K. Szymanski, P. Bródka, R. Michalski, and T. Kajdanowicz. Sequential seeding in complex networks: Trading speed for coverage. *Sci. Rep*, 7:891, 2017.

[46] A. Java, X. Song, T. Finin, and B. Tseng. Why we twitter: understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, pages 56–65. ACM, 2007.

[47] P. Jenkins, G. Earle-Richardson, D. T. Slingerland, and J. May. Time dependent memory decay. *Am J Ind Med*, 41(2):98–101, 2002.

[48] K. Juszczyszyn, P. Kazienko, and K. Musiał. Local topology of social network based on motif analysis. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, pages 97–105. Springer, 2008.

[49] R. R. Kao, D. M. Green, J. Johnson, and I. Z. Kiss. Disease dynamics over very different time-scales: foot-and-mouth disease and scrapie on the network of livestock movements in the uk. *Journal of the Royal Society Interface*, 4(16):907–916, 2007.

[50] T. Karagiannis, J.-Y. Le Boudec, and M. Vojnovic. Power law and exponential decay of intercontact times between mobile devices. *IEEE Transactions on Mobile Computing*, 9(10):1377–1390, 2010.

[51] F. Karimi and P. Holme. Threshold model of cascades in empirical temporal networks. *Physica A: Statistical Mechanics and its Applications*, 392(16):3476–3483, 2013.

[52] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki. Small but slow world: How network topology and burstiness slow down spreading. *Physical Review E*, 83(2):025102, 2011.

[53] L. Katz. A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43, Mar 1953.

[54] D. Kempe, J. Kleinberg, and A. Kumar. Connectivity and inference problems for temporal networks. *Journal of Computer and System Sciences*, 64(4):820–842, 2002.

[55] D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.

[56] H. Kim and R. Anderson. Temporal node centrality in complex networks. *Physical Review E*, 85(2):026107, 2012.

[57] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter. Multilayer networks. *Journal of complex networks*, 2(3):203–271, 2014.

[58] T. Kobayashi and T. Takaguchi. Social dynamics of financial networks. *EPJ Data Science*, 7(1):15, 2018.

[59] G. Kossinets and D. J. Watts. Empirical analysis of an evolving social network. *science*, 311(5757):88–90, 2006.

[60] M. Kretzschmar and M. Morris. Measures of concurrency in networks and the spread of infectious disease. *Mathematical biosciences*, 133(2):165–195, 1996.

[61] G. Krings, M. Karsai, S. Bernhardsson, V. D. Blondel, and J. Saramäki. Effects of time window size and placement on the structure of an aggregated communication network. *EPJ Data Science*, 1(1):4, 2012.

[62] M. Kulisiewicz, P. Kazienko, B. K. Szymański, and R. Michalski. Entropy measures of human communication dynamics. *arXiv preprint arXiv:1801.04528*, 2018.

[63] R. KUMAR, J. NOVAK, P. RAGHAVAN, and A. TOMKINS. On the bursty evolution of blogospace. international world wide web conference. In *Proceedings of the 12th international conference on World Wide Web. Hungria*, 2003.

[64] M. Lahiri and T. Y. Berger-Wolf. Structure prediction in temporal networks using frequent subgraphs. In *Computational Intelligence and Data Mining, 2007. CIDM 2007. IEEE Symposium on*, pages 35–42. IEEE, 2007.

[65] R. Lambiotte, L. Tabourier, and J.-C. Delvenne. Burstiness and spreading on temporal networks. *The European Physical Journal B*, 86(7):320, 2013.

[66] M. Lamont, L. Adler, B. Y. Park, and X. Xiang. Bridging cultural sociology and cognitive psychology in three contemporary research programmes. *Nature Human Behaviour*, 1(12):866, 2017.

[67] M. Latapy, T. Viard, and C. Magnien. Stream graphs and link streams for the modeling of interactions over time. *Social Network Analysis and Mining*, 8(1):61, 2018.

[68] K.-M. Lee, B. Min, and K.-I. Goh. Towards real-world complexity: an introduction to multiplex networks. *The European Physical Journal B*, 88(2):48, 2015.

[69] S. Lee, L. E. Rocha, F. Liljeros, and P. Holme. Exploiting temporal network structures of human interaction to effectively immunize populations. *PloS one*, 7(5):e36439, 2012.

[70] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1):2, 2007.

[71] A. Li, S. P. Cornelius, Y.-Y. Liu, L. Wang, and A.-L. Barabási. The fundamental advantages of temporal networks. *Science*, 358(6366):1042–1046, 2017.

[72] T. M. Liggett et al. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, volume 324. Springer Science & Business Media, 1999.

[73] K. Liu, W. K. Cheung, and J. Liu. Detecting stochastic temporal network motifs for human communication patterns analysis. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 533–540. ACM, 2013.

[74] L. Lotero, R. G. Hurtado, L. M. Floría, and J. Gómez-Gardeñes. Rich do not rise early: spatio-temporal patterns in the mobility networks of different socio-economic classes. *Royal Society open science*, 3(10), 2016.

[75] A. Manivannan, W. Q. Yow, R. Bouffanais, and A. Barrat. Are the different layers of a social network conveying the same information? *arXiv preprint arXiv:1801.02874*, 2018.

[76] L. Martinet, C. Crespelle, E. Fleury, P.-Y. Boëlle, and D. Guillemot. The link stream of contacts in a whole hospital. *arXiv preprint arXiv:1805.05752*, 2018.

[77] O. Michail. An introduction to temporal graphs: An algorithmic perspective. *Internet Mathematics*, 12(4):239–280, 2016.

[78] R. Michalski, S. Palus, and P. Kazienko. Matching organizational structure and social network extracted from email communication. In *International Conference on Business Information Systems*, pages 197–206. Springer, 2011.

[79] R. Michalski, B. K. Szymański, P. Kazienko, C. Lebiere, O. Lizardo, and M. Kulisiewicz. Social networks through the prism of cognition. *arXiv preprint arXiv:1806.04658*, 2018.

[80] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[81] B. Min, K.-I. Goh, and A. Vazquez. Spreading dynamics following bursty human activity patterns. *Physical Review E*, 83(3):036102, 2011.

[82] J. Moody. The importance of relationship timing for diffusion. *Social Forces*, 81(1):25–56, 2002.

[83] M. Morris and M. Kretzschmar. Concurrent partnerships and transmission dynamics in networks. *Social Networks*, 17(3-4):299–318, 1995.

[84] E. Mussumeci and F. C. Coelho. Reconstructing news spread networks and studying its dynamics. *Social Network Analysis and Mining*, 8(1):6, 2018.

[85] V. Neiger, C. Crespelle, and E. Fleury. On the structure of changes in dynamic contact networks. In *SITIS*, pages 731–738, 2012.

[86] M. Newman. Networks: an Introduction. *Oxford*, 2010.

[87] M. E. Newman. Spread of epidemic disease on networks. *Physical review E*, 66(1):016128, 2002.

[88] M. E. Newman. The mathematics of networks. *The new palgrave encyclopedia of economics*, 2(2008):1–12, 2008.

[89] V. Nicosia, J. Tang, C. Mascolo, M. Musolesi, G. Russo, and V. Latora. Graph metrics for temporal networks. In *Temporal networks*, pages 15–40. Springer, 2013.

[90] V. Nicosia, J. Tang, M. Musolesi, G. Russo, C. Mascolo, and V. Latora. Components in time-varying graphs. *Chaos: An interdisciplinary journal of nonlinear science*, 22(2):023101, 2012.

[91] J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási. Structure and tie strengths in mobile communication networks. *Proceedings of the national academy of sciences*, 104(18):7332–7336, 2007.

[92] R. K. Pan and J. Saramäki. Path lengths, correlations, and centrality in temporal networks. *Physical Review E*, 84(1):016105, 2011.

[93] A. Paranjape, A. R. Benson, and J. Leskovec. Motifs in temporal networks. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 601–610. ACM, 2017.

[94] M. Pascual, J. A. Dunne, et al. *Ecological networks: linking structure to dynamics in food webs*. Oxford University Press, 2006.

[95] A. Passarella and M. Conti. Characterising aggregate inter-contact times in heterogeneous opportunistic networks. In *International Conference on Research in Networking*, pages 301–313. Springer, 2011.

[96] A. Passarella, M. Conti, C. Boldrini, and R. I. Dunbar. Modelling inter-contact times in social pervasive networks. In *Proceedings of the 14th ACM international conference on Modeling, analysis and simulation of wireless and mobile systems*, pages 333–340. ACM, 2011.

[97] H. Pham, C. Shahabi, and Y. Liu. Ebm: an entropy-based model to infer social strength from spatiotemporal data. In *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*, pages 265–276. ACM, 2013.

[98] T. Poisot, D. B. Stouffer, and D. Gravel. Beyond species: why ecological interaction networks vary through space and time. *Oikos*, 124(3):243–251, 2015.

[99] D. Reitter and C. Lebiere. Social cognition: memory decay and adaptive information filtering for robust information maintenance. In *Proceedings of the Twenty-sixth AAAI Conference on Artificial Intelligence*, pages 242–248, 2012.

[100] C. S. Riolo, J. S. Koopman, and S. E. Chick. Methods and measures for the description of epidemiologic contact networks. *Journal of Urban Health*, 78(3):446–457, 2001.

[101] L. E. Rocha and V. D. Blondel. Bursts of vertex activation and epidemics in evolving networks. *PLoS computational biology*, 9(3):e1002974, 2013.

[102] L. E. Rocha, F. Liljeros, and P. Holme. Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLoS computational biology*, 7(3):e1001109, 2011.

[103] S. Ruan and W. Wang. Dynamical behavior of an epidemic model with a nonlinear incidence rate. *Journal of Differential Equations*, 188(1):135–163, 2003.

[104] V. Sekara, A. Stopczynski, and S. Lehmann. Fundamental structures of dynamic social networks. *Proceedings of the national academy of sciences*, 113(36):9977–9982, 2016.

[105] C. E. Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.

[106] B. Shulgin, L. Stone, and Z. Agur. Pulse vaccination strategy in the sir epidemic model. *Bulletin of mathematical biology*, 60(6):1123–1148, 1998.

[107] T. Smieszek, L. Fiebig, and R. W. Scholz. Models of epidemics: when contact repetition and clustering should be included. *Theoretical biology and medical modelling*, 6(1):11, 2009.

[108] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási. Limits of predictability in human mobility. *Science*, 327(5968):1018–1021, 2010.

[109] S. Sreenivasan, K. S. Chan, A. Swami, G. Korniss, and B. K. Szymanski. Information cascades in feed-based networks of users with limited attention. *IEEE Transactions on Network Science and Engineering*, 4(2):120–128, 2017.

[110] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, V. Colizza, L. Isella, C. Régis, J.-F. Pinton, N. Khanafer, W. Van den Broeck, et al. Simulation of an seir infectious disease model on the dynamic contact network of conference attendees. *BMC medicine*, 9(1):87, 2011.

[111] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.-F. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina, et al. High-resolution measurements of face-to-face contact patterns in a primary school. *PloS one*, 6(8):e23176, 2011.

[112] L. Stone, B. Shulgin, and Z. Agur. Theoretical examination of the pulse vaccination policy in the sir epidemic model. *Mathematical and computer modelling*, 31(4-5):207–215, 2000.

[113] A. Striegel, S. Liu, L. Meng, C. Poellabauer, D. Hachen, and O. Lizardo. Lessons learned from the netsense smartphone study. In *ACM SIGCOMM Computer Communication Review*, volume 43, pages 51–56. ACM, 2013.

[114] J. Tang, M. Musolesi, C. Mascolo, and V. Latora. Temporal distance metrics for social network analysis. In *Proceedings of the 2Nd ACM Workshop on Online Social Networks*, WOSN '09, pages 31–36, New York, NY, USA, 2009. ACM.

[115] J. Tang, M. Musolesi, C. Mascolo, V. Latora, and V. Nicosia. Analysing information flows and key mediators through temporal centrality metrics. In *Proceedings of the 3rd Workshop on Social Network Systems*, SNS '10, pages 3:1–3:6, New York, NY, USA, 2010. ACM.

[116] I. W. Taylor, R. Linding, D. Warde-Farley, Y. Liu, C. Pesquita, D. Faria, S. Bull, T. Pawson, Q. Morris, and J. L. Wrana. Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nature biotechnology*, 27(2):199, 2009.

[117] W. H. Thompson, P. Brantefors, and P. Fransson. From static to temporal network theory: Applications to functional brain connectivity. *Network Neuroscience*, 1(2):69–99, 2017.

[118] P. Vanhems, A. Barrat, C. Cattuto, J.-F. Pinton, N. Khanafer, C. Régis, B.-a. Kim, B. Comte, and N. Voirin. Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PloS one*, 8(9):e73970, 2013.

[119] A. Vazquez, B. Racz, A. Lukacs, and A.-L. Barabasi. Impact of non-poissonian activity patterns on spreading processes. *Physical review letters*, 98(15):158702, 2007.

[120] M. C. Vernon and M. J. Keeling. Representing the uk's cattle herd as static and dynamic networks. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1656):469–476, 2009.

[121] E. Volz and L. A. Meyers. Susceptible–infected–recovered epidemics in dynamic contact networks. *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1628):2925–2934, 2007.

[122] S. Wasserman and K. Faust. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.

[123] D. J. Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771, 2002.

[124] J. Whitbeck, M. Dias de Amorim, V. Conan, and J.-L. Guillaume. Temporal reachability graphs. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, Mobicom '12, pages 377–388, New York, NY, USA, 2012. ACM.

[125] J. T. Wixted and E. B. Ebbesen. On the form of forgetting. *Psychological science*, 2(6):409–415, 1991.

[126] D. Xiao and S. Ruan. Global analysis of an epidemic model with nonmonotone incidence rate. *Mathematical biosciences*, 208(2):419–429, 2007.

[127] J. Xie, M. Chen, and B. K. Szymanski. Labelrankt: Incremental community detection in dynamic networks via label propagation. In *Proceedings of the Workshop on Dynamic Networks Management and Mining*, pages 25–32. ACM, 2013.

[128] B. B. Xuan, A. Ferreira, and A. Jarry. Computing shortest, fastest, and foremost journeys in dynamic networks. *International Journal of Foundations of Computer Science*, 14(02):267–285, 2003.

[129] Ö. N. Yaveroğlu, N. Malod-Dognin, D. Davis, Z. Levnajic, V. Janjic, R. Karapandza, A. Stojmirovic, and N. Pržulj. Revealing the hidden language of complex networks. *Scientific reports*, 4:4547, 2014.

[130] K. Zhao, M. Karsai, and G. Bianconi. Entropy of dynamical social networks. *PloS one*, 6(12):e28116, 2011.

[131] Q. Zhao, Y. Tian, Q. He, N. Oliver, R. Jin, and W.-C. Lee. Communication motifs: a tool to characterize social communications. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, pages 1645–1648. ACM, 2010.

[132] Z. Zhi-Dan, X. Hu, S. Ming-Sheng, and Z. Tao. Empirical analysis on the human dynamics of a large-scale short message communication system. *Chinese Physics Letters*, 28(6):068901, 2011.

[133] H. Zhu, L. Fu, G. Xue, Y. Zhu, M. Li, and L. M. Ni. Recognizing exponential inter-contact time in vanets. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–5. IEEE, 2010.