A Visualization Approach for Simulating and Analyzing Infection Spread Dynamics Using Temporal Networks

Jean R. Ponciano[1,2], Gabriel P. Vezono[1], Claudio D. G. Linhares[3]

[1] Faculty of Computing, Federal University of Uberlândia, Uberlândia, Brazil
[2] School of Applied Mathematics, Getulio Vargas Foundation (FGV), Brazil
[3] Institute of Mathematics and Computer Sciences, University of São Paulo, São Carlos, Brazil
jean.ponciano@fgv.br, gvezono@gmail.com, claudiodgl@usp.br

Abstract.

Temporal networks have been widely used to model instances of a domain of interest and their time-evolving interaction, including modeling individuals and face-to-face contacts throughout time. In the context of infection spread, such individuals can, e.g., remain susceptible, recovered, or be infected at a particular time. Understanding the infection spread behavior (its speed and magnitude, for instance) is crucial for quick and reliable decision making. Network visualization strategies can help in this task as they allow easy identification of who infected whom and when, epidemics outbreak, and other relevant aspects. This paper presents a visualization approach for the simulation and analysis of infection spread dynamics that considers different infection probabilities and different levels of social distancing (inter-group interaction). We performed quantitative and visual experiments using three real-world social networks with distinct characteristics and from two different environments. Our findings reveal the overall influence of different levels of inter-group interaction and infection probabilities in the infection spread dynamics and also demonstrate the usefulness of our approach for enhanced local (individual- or group-level) investigations.

Categories and Subject Descriptors: E.1 [**Data Structures**]: Graphs and networks; I.5 [**Pattern Recognition**]: Miscellaneous; I.6 [**Simulation and Modeling**]: Applications; J.3 [**Life and Medical Sciences**]: Health

Keywords: Infection propagation, Visualization, Temporal networks, Dynamic networks, Community detection

1. INTRODUCTION

Information visualization strategies create graphical and interactive representations of abstract data, thus comprising an effective resource to gain insights and guide decision making [Ponciano et al. 2020]. It is useful, for example, to provide awareness by graphically showing real-world problems. In the context of health/epidemic data, an efficient visual analysis may assist experts in the study and decisions related to infection spread dynamics, such as in the recent COVID-19 pandemic [Dong et al. 2020; So et al. 2020].

In epidemics, the infection propagates through an interaction (usually face-to-face) involving an infected and a susceptible individual. Networks represent a useful and commonly used structure to model this type of data, i.e., data composed of interactions involving a domain's instances. In a network, the instances (e.g., individuals) are represented by nodes and the interactions that connect two nodes (e.g., face-to-face contacts) are represented by edges. When the information of *when* each interaction occurs is known, such network is called *temporal network* [Holme and Saramäki 2012].

The temporal network dynamism, in which there are time intervals with bursts of interactions as

well as periods of low (or none) activity, makes such network useful to simulate and study from fake news [Jr. and Frizzon 2019] to infection spreading [Linhares et al. 2019]. Another relevant property of real-world networks is the community formation, i.e., the presence of groups of individuals that interact more often among themselves than with individuals from other groups [Linhares et al. 2019]. Knowing these communities, that may be related to, e.g., household members or people in workplaces and school classes, allows a better understanding – and consequently a more effective decision making – regarding the control of the spread speed and magnitude [Ghalmane et al. 2019].

Temporal networks have been widely used to simulate infection spread dynamics through a variety of infection models [Linhares et al. 2019]. Although the employment of visualization strategies comprehends a useful tool that allows, e.g., contact tracing and the understanding of the transmission path and spread speed [Carroll et al. 2014; Linhares et al. 2019], few studies consider visualization strategies in their analysis. Most of them focus on stochastic models, geographical information, or statistical properties of the infected population [Tepper and Thiébaut 2017].

This paper presents a visualization approach for performing different simulations of infection spread dynamics that take place in real-world environments, represented in our study by temporal social networks. We simulate infection propagations by using the well-established infection models *Susceptible-Infected (SI)* [Prakash et al. 2014] and *Susceptible-Infected-Recovered (SIR)* [Barrat et al. 2008; Rocha and Blondel 2013]. Our approach considers different levels of social distancing and different infection probabilities, aspects that are perceived in real-world scenarios when the individuals avoid (or not) contacts or adopt (or not) protective measures such as the use of face masks. To the best of our knowledge, our work is the first to simultaneously consider both aspects in this type of study.

Our main contributions are: (i) a strategy for simulating different levels of social distancing that considers community detection and the number of inter-community interactions in a temporal network; (ii) a visualization approach that employs three layouts (MSV, TAM, and node-link diagram) along with a well-established node ordering method (CNO [Linhares et al. 2019]) for the analysis of infection spread dynamics; (iii) an extended version of the interactive software DyNetVis [Linhares et al. 2020] that implements our methodology and is freely available for potential users. With our approach, one can efficiently perform different infection spread simulations and identify key nodes, the transmission path (who infected whom and when), the epidemics outbreak, and other relevant behaviors. This article is an extension of the conference paper [Ponciano et al. 2021] published in the *Proceedings of the 2021 Brazilian Symposium on Databases (SBBD)*. This extended version provides new experiments to support our approach's usefulness and efficiency. It now includes analyses containing (i) a third real-world network from a domain not previously considered (InVS network), (ii) a second infection dynamics model (SIR), and (iii) a third layout (node-link diagram).

The rest of this paper is organized as follows. Section 2 presents related concepts and Section 3 discusses related work. Section 4 then describes our methodology, parameters used, the three evaluated real-world temporal networks, and the performed experiments. Section 5 discusses the results and limitations. Section 6 presents the conclusion and future work.

2. RELATED CONCEPTS

2.1 Infection models

The *Susceptible-Infected (SI)* model comprehends a simple infection dynamics where each node is associated with one of two states at a particular time, susceptible (S) or infected (I). All nodes start susceptible, except for a node chosen to be initially infected (patient zero). At each subsequent timestamp, each interaction involving an infected node and one of its susceptible neighbors (adjacent nodes) has a probability λ of infecting such neighbor (Figure 1(a)). In the SI model, an infected node does not recovery or turn susceptible again.

(a) SI (b) SIR

Fig. 1. Two infection models and considered states. (a) SI. (b) SIR. The symbol λ stands for the infection probability. Contrary to SI, individuals in SIR recover after being infected for r_t time steps.

There are also models suitable for other situations. As illustrated in Figure 1(b), one such example is the *Susceptible-Infected-Recovered (SIR)* model. It is analogous to SI with the exception that it considers that an infected node recovers after being infected for r_t time steps [Barrat et al. 2008; Linhares et al. 2019]. In the SIR model, recovered nodes do not re-infect or turn susceptible again.

2.2 Network Visualization

Several strategies were developed over the past years to enhance the visualization of temporal networks and thus improve the user perception regarding network evolution and existing patterns [Beck et al. 2016]. Figure 2 presents different layouts to analyze temporal networks and also illustrates their usefulness for analyzing infection spread dynamics. Figure 2(b) shows a node-link diagram [Battista et al. 1994], the most common strategy to highlight network topology and structure. In this layout, nodes are represented by circles and edges are straight lines. The temporal evolution is considered through animation. Without animation, the temporal information is ignored and the aggregated network (i.e., all timestamps at once) is considered. In the context of infection spread dynamics, this strategy is useful to identify node state and transmission path at particular times or at the end of the observation period.

The *Massive Sequence View (MSV)* [Linhares et al. 2019] is a layout focused on the network temporal evolution. As shown in Figure 2(c), nodes are vertically displayed and timestamps are horizontally positioned. Whenever there is an edge connecting two nodes, a vertical line involving them is drawn at the corresponding timestamp. Figure 2(d) illustrates how MSV can be employed for analyzing an infection spread. In this case, MSV shows only those edges through which the infection propagates, thus highlighting the infection path from infected to susceptible nodes. This approach, along with a transparency effect on non-infected nodes, facilitates the identification of who infected whom and when the transmission occurred. Such information is useful to recognize the importance of particular nodes and act to regulate the infection spread [Linhares et al. 2019].

The *Temporal Activity Map (TAM)* [Linhares et al. 2017] is a variant of MSV where nodes are represented by squares instead of circles (thus giving a better sense of continuity) and all edges are omitted. As illustrated in Figure 2(e), node color can be used to map labels (or other metadata) or structural properties such as the node degree. When studying infection spread dynamics, TAM is useful to highlight the node state (e.g., susceptible or infected), facilitating the identification of how groups of nodes with the same state evolve [Linhares et al. 2019] (Figure 2(f)).

Layouts for temporal network visualization can be categorized into animated (when the time dimension is mapped using consecutive frames over time) and timeline-based (when the time dimension is shown along with nodes and/or edges as a single picture) [Beck et al. 2016]. Timeline-based layouts are better than the others when the analysis depends on more than two timestamps [Beck et al. 2016], which is a key need in the type of analysis we consider in our study (e.g., tracking who infected whom and when). Note that both MSV and TAM belong to the timeline-based category.

There are several efforts to reduce visual clutter and improve aesthetic aspects of network visualizations, such as the choice of nodes/edges positioning, sizes, colors, and others [Beck et al. 2016]. In node-link diagrams, a popular node positioning strategy involves the employment of a force-directed algorithm through which nodes are rearranged using spring-force to approximate those more frequently

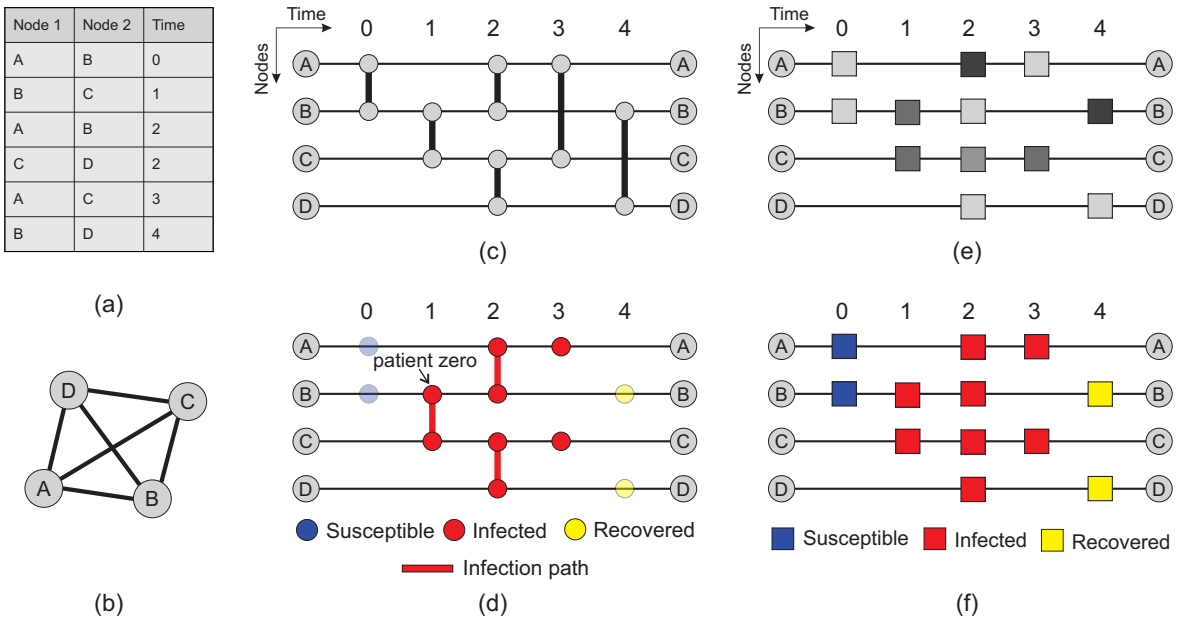

Fig. 2. Different layouts for temporal network visualization and how they can be used in the analysis of infection spread dynamics: (a) Raw data; (b) Node-link diagram; (c) MSV; (d) MSV with infection path; (e) TAM; (f) TAM highlighting node state.

connected [Brandes 2008]. For MSV, there are strategies that focus on re-positioning nodes [Linhares et al. 2019], sampling relevant edges [Ponciano et al. 2021], and changing the network temporal resolution [Ponciano et al. 2021]. The adoption of a high-quality node positioning strategy is important as changes in the node ordering can reduce edge length and consequently overlaps and visual clutter [Ponciano et al. 2022].

The *community-based node ordering (CNO)* [Linhares et al. 2019] is a visual scalable node positioning method developed to optimize the analysis of temporal networks in MSV and TAM layouts. CNO's procedure is composed of three steps: first, the method detects network communities by employing a well-established detection algorithm, such as Louvain [Blondel et al. 2008] or Infomap [Rosvall and Bergstrom 2008]; then, CNO uses node positioning methods to find a suitable layout placement for each community (second step – inter-community positioning) and for each node inside each community (third step – intra-community positioning). The produced layout has reduced (average) edge

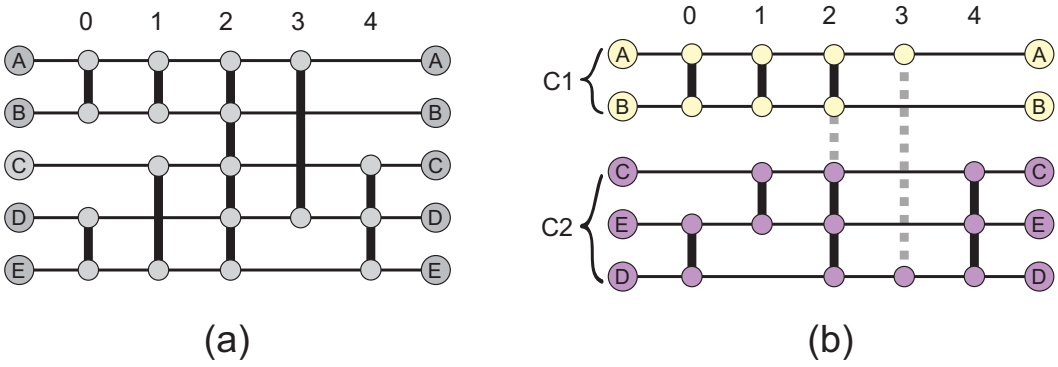

Fig. 3. Community-based node ordering (CNO) applied to MSV. (a) Original node positioning. (b) CNO. After CNO, the user may associate different colors to different communities and discard inter-community edges (dotted lines). 'C1' and 'C2' represent two different communities.

length and, consequently, less overlaps and visual clutter. As a result, hidden visual patterns emerge, for example, the presence of key nodes or groups and interesting time intervals [Linhares et al. 2019; Ponciano et al. 2021]. As illustrated in Figure 3(b), through interactive tools the user can associate different colors to different communities and also discard inter- or intra-community edges, thus reducing the amount of visual information and focusing the analysis on regions of interest.

3. RELATED WORK

Different visualization tools have been developed over the past years to analyze health/epidemic data, as for example the interactive dashboard proposed in [Dong et al. 2020] to track COVID-19 reported cases. There are three main categories of visualization applications when dealing with health data [Carroll et al. 2014]: geographic information systems (GIS), that allow the identification of particular cases, spatial trends, and disease clusters, among others; molecular epidemiology, that considers, e.g., the visualization of serotypes and genetic variants; and social temporal network analysis, commonly used to understand infection outbreaks and trace contacts and infection paths.

VoroGraph [Dunne et al. 2015] is a visual analytics tool that considers different visualization strategies based on networks and geographic information to assist epidemiologists in their analysis. Another example of temporal network visualization strategy useful in this context is the radial layout developed in [Tepper and Thiébaut 2017], through which the user can identify individuals involved in the infection spread, trace contacts, and analyze how an intervention would influence the evaluated scenarios. More recently, temporal network visualization has also been used in the study of COVID-19 transmission risk in different countries (e.g., South Korea and China), often through spatial visualizations and/or node-link diagrams [Park 2020; So et al. 2020].

Without considering infection dynamics, the timeline-based layouts MSV and TAM have been successfully used to analyze structural and temporal patterns in networks from different domains, including global and local-level patterns on Twitter [Ponciano et al. 2021; Linhares et al. 2019] and in hospital [Linhares et al. 2019] and school environments [Ponciano et al. 2020; Ponciano et al. 2021]. Their capabilities were also demonstrated by a recent user study that included feedback from 74 participants [Linhares et al. 2021]. Given these layouts' strengths and characteristics, they have recently been adapted to enhance analyses and decision making in health contexts. Applications include the simulation and analysis of infection spreads [Linhares et al. 2019] and the inspection of patients' clinical exams results [Linhares et al. 2021; Linhares et al. 2022].

In this paper, we also use MSV and TAM visualizations to study infection spread dynamics. The main reasons are their successful application in previous studies allied to the advantages of using timeline-based layouts for the type of analysis we are interested in (recall Section 2.2). Table I summarizes the differences between our study and others described throughout this section. The comparison criteria are whether the study offers partial data visualization (e.g., via animation frames or snapshots) or visualization of all data at once (via a timeline-based layout), whether it enables group-level analysis, and whether it considers different infection probabilities and interaction restriction levels. To the best of our knowledge, our work is the first to employ timeline-based visualizations to study infection propagation while simultaneously taking into account different levels of inter-group interaction (i.e., interaction restriction to simulate social distancing) and different infection probabilities (simulating the adoption of protective measures).

In [Chang et al. 2020], the authors consider social distancing as the limitation of non-household interactions while preserving household contacts unaltered. Even though household members comprehend a relevant group of nodes that should be considered in the analysis of infection spread, other groups based on sociological factors may exist (e.g., individuals in a workplace or in a school class). Groups such as these represent useful information and could be modeled as network communities [Kitchovitch and Liò 2010]. Although the analysis of the network community structure (e.g., the number and den-

Table I. Comparison between our study and others. The comparison criteria are whether the study offers partial data visualization (e.g., via animation frames or snapshots) or visualization of all data at once (via a timeline-based layout), whether it enables group-level analysis, and whether it considers different infection probabilities (Infec. prob.) and interaction restriction levels (Interac. restr.).

	Partial data vis.	All data vis.	Group analysis	Infec. prob.	Interac. restr.
[Dunne et al. 2015]	✓		✓		
[Tepper and Thiébaut 2017]		✓		✓	
[Park 2020]			✓		
[So et al. 2020]	✓				
[Linhares et al. 2019]		✓		✓	
Our study	✓	✓	✓	✓	✓

sity of inter- and intra-community contacts) is crucial for defining effective immunization strategies, few studies consider this information [Ghalmane et al. 2019]. Our approach focus on this type of group analysis.

4. VISUAL ANALYSIS

This section presents three case studies using the real-world networks *High School* [Mastrandrea et al. 2015], *Primary School* [Gemmetto et al. 2014], and InVS [Génois et al. 2015]. We consider these networks because they were proposed to analyze infection spreads. Moreover, they are publicly available and also have been extensively explored in visualization studies (e.g., [Linhares et al. 2020]). First, Section 4.1 describes the methodology we use to simulate and analyze infection spreads. Then, Section 4.2 describes the adopted parameter values, Section 4.3 discusses interaction tools provided by the system used to perform the visual analysis, and Section 4.4 presents the networks. In the sequence, Sections 4.5, 4.6 and 4.7 present the experiments and results. Figures 6, 7, 8, and 9 can be accessed in higher resolutions at `https://tinyurl.com/44k797ca`.

4.1 Methodology

In this paper, we are interested in simulating and analyzing infection spread dynamics through temporal network visualization. From the simulation point of view, we are interested in enabling and performing infection spread simulations that take into account different levels of social distancing and infection probabilities. On the visualization side, we want to enable users to identify relevant patterns and behaviors that may help in decision-making processes. Examples include identifying key nodes and the infection outbreak (at global and group levels), and tracking the transmission path (who infected whom and when). As stated in Section 3, our visual analysis is mainly supported by the timeline-based layouts MSV and TAM.

To analyze how an infection spreads among nodes of a temporal network, we consider two key aspects. The first one concerns the number of interactions involving the existing nodes. Inspired by [Chang et al. 2020], we associate the number of interactions with social distancing, which is modeled in our context by keeping all intra-community interactions unaltered while varying the allowed number of inter-community ones. To perform such variation, β randomly chosen inter-community edges are maintained in the network, while the others are discarded. As an example, if $\beta = 0$, all inter-community edges are discarded and consequently there is a full lockdown – which is impracticable in real-world scenarios since there are many essential services [Canabarro et al. 2020]. The second aspect is related to the probability λ of a susceptible node be infected after interacting with an infected one. As an example, if $\lambda = 1$, every susceptible node will certainly become infected after interacting with an infected node.

The simulation process involves grouping the individuals (nodes) and reducing inter-group interactions according to the desired level of social distancing. Any clustering strategy based on node

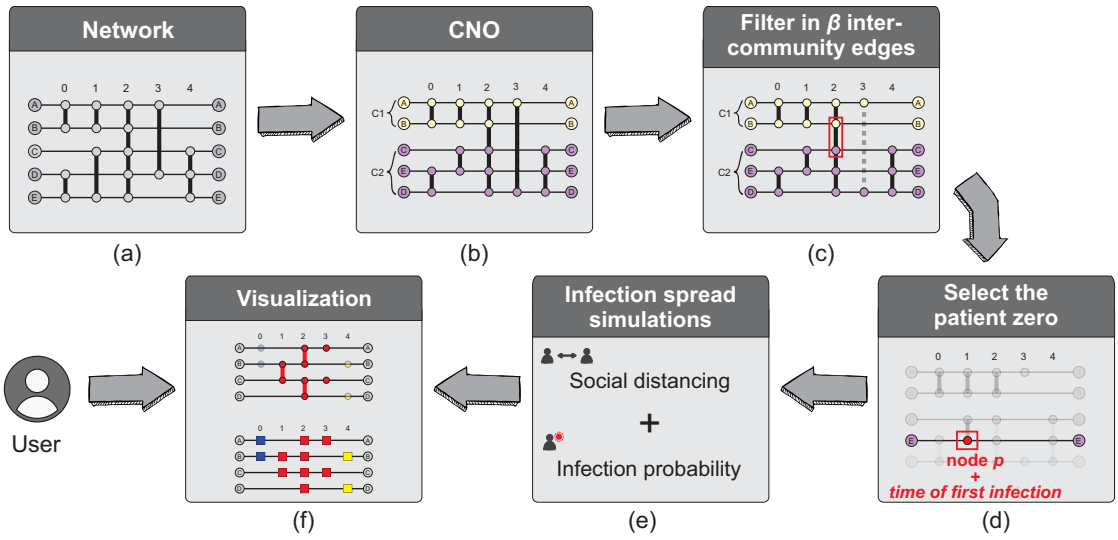

Fig. 4. Methodology step-by-step. First, CNO is executed for a given temporal network (a, b) and β inter-community edges are filtered in to simulate social distancing (c). Then, we select the patient zero p and the time of first infection (d) and use these information, along with the chosen social distancing compliance and infection probability, to simulate SI and SIR infection spreads (e). Finally, users can analyze the produced visualizations to study the propagation dynamics.

interaction can be used in this step. We rely on network community detection because communities represent valuable information in real-world networks, inside and outside the context of infection spread analysis [Fortunato and Hric 2016; Ghalmane et al. 2019]. Given that we are also interested in visualizing the network and the simulated infection spread, we use CNO, a method that detects existing communities and also efficiently places the nodes in the MSV or TAM layout (recall Section 2.2).

The step-by-step of our methodology is illustrated in Figure 4. After running CNO over a given network and filtering in β inter-community edges to simulate social distancing (Figure 4(b,c)), we select the patient zero (node p) and the time of the first infection (Figure 4(d)). The choice of p could be either arbitrary or related to a node structural property such as degree or betweenness. After that, we perform (i) a SI infection spread simulation considering p and the chosen probability λ, or (ii) a SIR infection spread simulation considering p, λ, and a chosen recovery time r_t (Figure 4(e)). Finally, users can analyze the produced visualizations to study the propagation dynamics, e.g., by tracking intra-community infection spread, identifying who infected whom and when, and perceiving other relevant behaviors (Figure 4(f)). In order to have a more precise quantitative evaluation, given p, we repeat the entire process ten times for each of the combinations (β,λ) we consider in our experiments.

4.2 Parameter definition

Although our focus is not the analysis of the COVID-19 transmission, we rely on studies that consider its behavior to define which level of social distancing (β value) should be considered in our experiments. While [Canabarro et al. 2020] consider social distancing with 75% of compliance, other studies (e.g., [Chang et al. 2020; Daghriri and Ozmen 2020]) recommend a compliance of at least 80%. We go in the same direction as these studies and also consider that the number of interactions should be reduced in 80%. We thus employ $\beta = 0.2$ (recall that β refers to the number of maintained inter-community interactions) and, as a counterpoint, $\beta = 0.8$ (compliance of 20%). In summary, the combinations (β,λ) tested in our experiments consider either $\beta = 0.8$ or $\beta = 0.2$.

The infection probabilities considered in our combinations are $\lambda = 0.05$ and $\lambda = 0.7$, values that intend to simulate different levels of protective measures during the interactions, e.g., the use (or not) of face masks. Unless stated otherwise, the patient zero (node p) is set as the most active node of the

most active CNO community. The time of the first infection is the first timestamp p appears in the network.

To be consistent with previous studies, CNO is configured as follows. The adopted community detection algorithm (CNO first step) is *Infomap* for the Primary School and *Louvain* for the High School and InVS. These algorithms are recommended when dealing with these networks because of their high performance and relation with ground-truth groups (departments / school classes) in these cases [Linhares et al. 2020]. As an example, each community detected by Infomap in the Primary School network corresponds precisely to a school class with its students and teacher. Details about the networks are given in the following section and we justify the use of community detection instead of the ground-truth groups in Section 5. To define community and node positioning (CNO second and third steps), we follow the CNO creators [Linhares et al. 2019] and also adopt *Recurrent Neighbors (RN)*, a method that places neighboring (i.e., connected) nodes close to each other in the layout [Linhares et al. 2017].

4.3 User interaction

All visual analyses were performed using DyNetVis [Linhares et al. 2020], an interactive, freely available, and open-source software that was extended for this study to consider social distancing simulation. Our extended version is available at `www.dynetvis.com` (v4.1).

DyNetVis implements all layouts from Figure 2 and several interaction tools and functionalities that enable users to adapt and coordinate the visualizations, as illustrated in Figure 5. As a result, the network data and the simulated infection spread can be analyzed through different and complementary perspectives, thus amplifying the user analytical capability. Examples of interaction available in all layouts include zoom in/out (Figure 5(a)) and the selection of groups of nodes or time intervals of interest (Figure 5(c, d)). All layouts are also coordinated such that nodes and time intervals selected in one of them are automatically selected in the others. Moreover, users can switch between layouts with a single mouse click (e.g., from TAM to MSV (Figure 5(a, b)) or from TAM to the node-link diagram (Figure 5(d, e))). Since the node-link diagram is not a timeline-based layout, a common strategy is to map the temporal information into animation frames. DyNetVis allows us to play/pause the

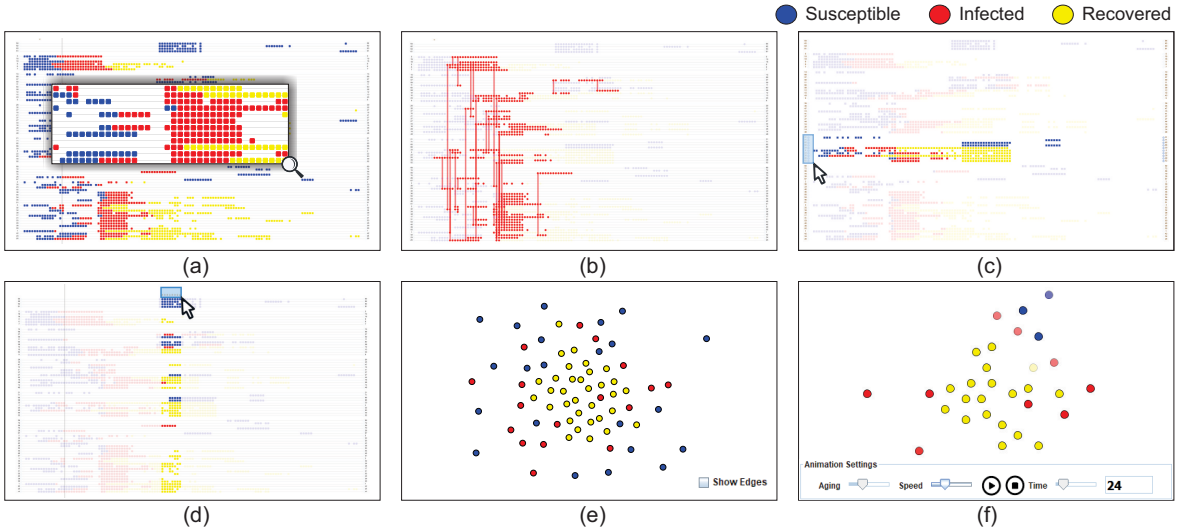

Fig. 5. Interaction. (a) Zoom in/out. (b) Moving from TAM (a) to MSV (b) with a single mouse click. (c) Selection of groups of nodes of interest. (d) Selection of time intervals of interest. (e) Moving from TAM (d) to the node-link diagram (e) with a single mouse click. Nodes and time intervals selected in a layout are automatically selected in the others. It is also possible to hide edges (e) and play/pause the animation (f) in the node-link diagram.

animation as well as to adapt the transition speed and fading factor effect (Figure 5(f)). Finally, edges can also be hidden in the node-link diagram (Figure 5(e)).

4.4 Networks

The *High School* network [Mastrandrea et al. 2015] was collected in Marseille, France and comprehends face-to-face interactions among 327 students (nodes) collected during five days. There are 47,586 edges distributed in 1,211 timestamps when considering the temporal resolution as a 5-minute interval. In this resolution, each timestamp represents 5 minutes and considers all edges that occur during this interval. There are nine classes in this network. They focus on different subjects and are divided into four categories: MP (mathematics and physics), PC (physics and chemistry), PSI (engineering studies), and BIO (biology).

The *Primary School* network data [Gemmetto et al. 2014] was collected during two days in Lyon, France and refers to face-to-face interactions among 232 students and ten teachers (242 nodes). There are 55,046 edges distributed in 389 timestamps when considering the temporal resolution as a 5-minute interval per timestamp. There are five grades and each grade is divided into two classes (A and B). Each of the ten classes has an assigned teacher.

The third network, named InVS [Génois et al. 2015], is composed of face-to-face interactions collected during two weeks (ten workdays) between 92 employees of the French Institute of surveillance in public health (nodes). Since this network is not as dense as the other two (see Table II), the adopted temporal resolution was chosen such that each timestamp comprehends a 15-minute interval. There are 92 nodes, 2,737 edges and 1,097 timestamps with this resolution configuration. The institute is composed of five departments: Chronic Diseases and Traumatisms (DMCT), Health and Environment (DSE), Human Resources (SRH), Scientific Direction (DISQ), and Logistics (SFLE), and there is metadata information about which node belongs to which department. Nodes are also categorized into "residents", "linkers" or "wanderers", depending on the number of interactions they have with nodes from other departments [Génois et al. 2015]: "residents" have more internal links than external ones, i.e., they connect mainly with nodes from their own department; "linkers" have approximately 50% of external links, i.e., they act as a bridge to other departments; and "wanderers" are more linked with nodes from other departments. Most nodes are "residents". Finally, the five departments are located on two floors in the building: the ground floor contains DISQ, DMCT, and SFLE, and the first floor contains DSE and SRH.

All three networks were collected using RFID sensors from badges. An interaction (edge) is computed whenever two individuals have a frontal approximation (between 1 and 1.5 meters). Table II summarize the networks' properties. Note that the Primary School network presents a higher number of edges per timestamp on average (EPT) when compared to the High School. This behavior indicates that students from the primary school interact a lot more among themselves when compared with the students from the high school, a situation that could facilitate the infection propagation. Note also that the InVS network is not as dense as the other two. It presents the lowest EPT, even with the highest temporal resolution (TR = 15 minutes).

Table II. Face-to-face networks properties. 'EPT' refers to the average number of edges per timestamp. '#Labels' refers to the number of classes or departments, depending on the network. 'TR' refers to the adopted temporal resolution (in minutes).

Network	**#Nodes**	**#Edges**	**#Times**	**EPT**	**#Labels**	**TR**
High Sch.	327	47,586	1,211	39.29	9	5
Primary Sch.	242	55,046	389	141.5	10	5
InVS	92	2,737	1,097	2.49	5	15

Table III. Quantitative evaluation for the High School network considering the average result of ten simulations adopting SI for each combination (β,λ). 'ATI' and 'MTI' refer to the average and the median timestamp of infection, respectively.

	(β - λ)	ATI	MTI	#inf.	#non-inf.
High School	(0.8-0.7)	36.1 ± 68.5	21	327	0
	(0.8-0.05)	722.1 ± 337.3	773	260	67
	(0.2-0.7)	69.4 ± 97.3	29	327	0
	(0.2-0.05)	678 ± 351.4	689	141	186

4.5 High School case

Table III presents the average result of a quantitative evaluation that considers ten simulations adopting SI for each combination (β,λ) in the High School network. In the two cases where $\lambda = 0.7$, all nodes become infected regardless of β. This is expected, since higher infection probabilities tend to require few interactions between two individuals to transmit the infection (recall that if $\lambda = 1$, a single contact would infect a susceptible node, for instance). In these two cases, however, the spread speed varies according to β in a way that the lower its value (i.e., the higher the social distancing compliance), the slower the infection spread (perceived by the higher median time of infection – MTI). By fixing β and varying λ, one can see that a low λ value leads to more non-infected nodes and slower spreads. The combination ($\beta = 0.2$, $\lambda = 0.05$) represents the best-case scenario in this network: only 43.11% of the nodes (on average) became infected at the end of the observation period. This supports the importance of reducing the number of interactions while adopting protective measures to regulate the infection spread.

Figure 6 presents an overview of the four evaluated combinations for the High School. Through this visualization, one can easily see that almost all nodes become infected in the first day of the network when considering $\lambda = 0.7$, regardless of the β value (Figure 6 – first row), but that a lower β slows

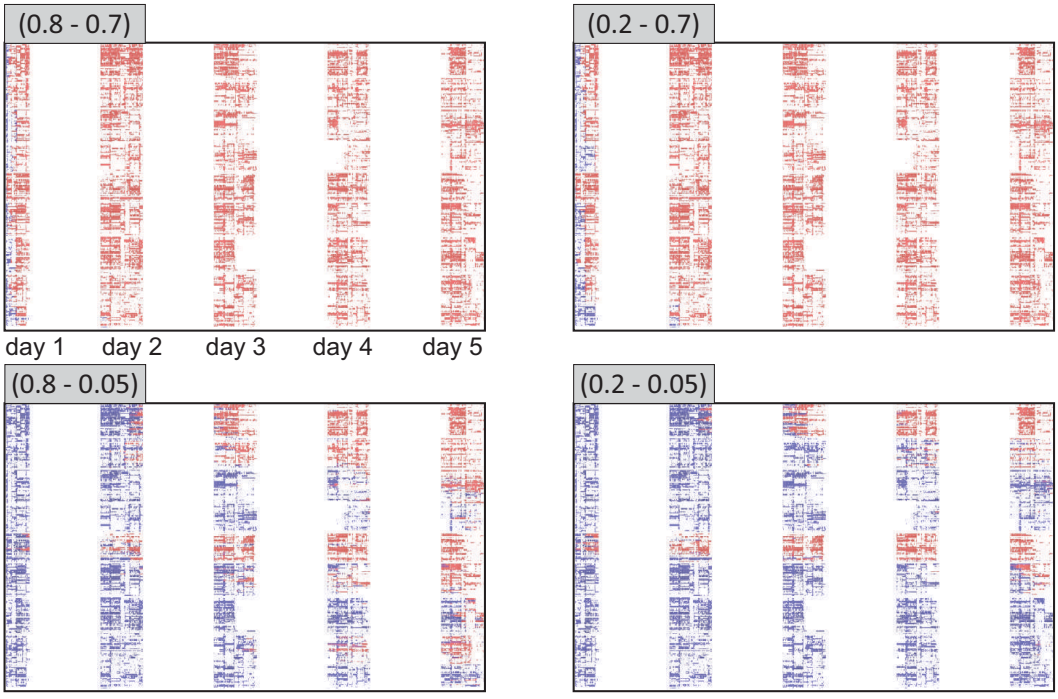

Fig. 6. TAM layouts showing how different combinations (β, λ) affect the infection spread behavior in the High School over five days. SI infection model. Almost all nodes become infected in the first day of the network when considering $\lambda = 0.7$, regardless of the β value, but a lower β slows down the spread speed ($\beta = 0.2$, $\lambda = 0.7$). Combination ($\beta = 0.2$, $\lambda = 0.05$) produces the best-case scenario where less than half of the nodes become infected.

Fig. 7. MSV infection path for the High School network ($\beta = 0.2$, $\lambda = 0.05$). SI infection model. The first three classes infected in the network are the ones focused on biology (2BIO2 – that contains the patient zero –, 2BIO1, and 2BIO3).

down the spread speed ($\beta = 0.2$, $\lambda = 0.7$). In the same way, it is easy to note that the combination ($\beta = 0.2$, $\lambda = 0.05$) produces the best-case scenario where less than half of the nodes become infected (43.11%, as mentioned earlier).

Figure 7 shows the MSV layout for the combination ($\beta = 0.2$, $\lambda = 0.05$) while highlights the transmission path and the involved classes. Through this visualization, one can see that the patient zero belongs to 2BIO2, a class of students that mainly study subjects related to biology [Mastrandrea et al. 2015]. In the High School, students interact more often with others from classes focused on the same subjects [Mastrandrea et al. 2015]. This interaction behavior highly impacts the infection spread: the first three classes infected in the network are exactly the ones focused on biology (2BIO2 – that contains the patient zero –, 2BIO1, and 2BIO3). Note also that classes PC and PC* (the ones based on physics and chemistry) had zero infected students.

4.6 Primary School case

Table IV presents the average result of a quantitative evaluation that considers ten simulations, also adopting SI, for each combination (β,λ) in the Primary School network. By considering $\beta = 0.8$ (social distancing of 20%) and varying the infection probability λ, one can perceive, by looking at the median time of infection (MTI), that the spread speed has been slowed in more than two times when adopting $\lambda = 0.05$. For $\beta = 0.2$ (social distancing of 80%) and $\lambda = 0.05$, the transmission was more than four times slower when compared with $\lambda = 0.7$, regardless of the adopted β. As occurred with the High School, the combination that resulted in a lower spread speed and a higher number of non-infected nodes at the end of the observation period (10) was ($\beta = 0.2$, $\lambda = 0.05$), showing the importance of having few inter-group interaction along with low infection probability.

In the Primary School, the combination with MTI = 109 (of 389 timestamps) resulted in the slowest propagation, before one-third of the observation period. Recall that the slowest propagation in the

Table IV. Quantitative evaluation for the Primary School network considering the average result of ten executions for each combination (β,λ). 'ATI' and 'MTI' refer to the average and the median timestamp of infection, respectively.

	(β - λ)	ATI	MTI	#inf.	#non-inf.
Primary School	(0.8-0.7)	38.4 ± 47.0	24	241	1
	(0.8-0.05)	81.1 ± 73.6	60	240	2
	(0.2-0.7)	39.6 ± 47.3	25	241	1
	(0.2-0.05)	162.8 ± 118.5	109	232	10

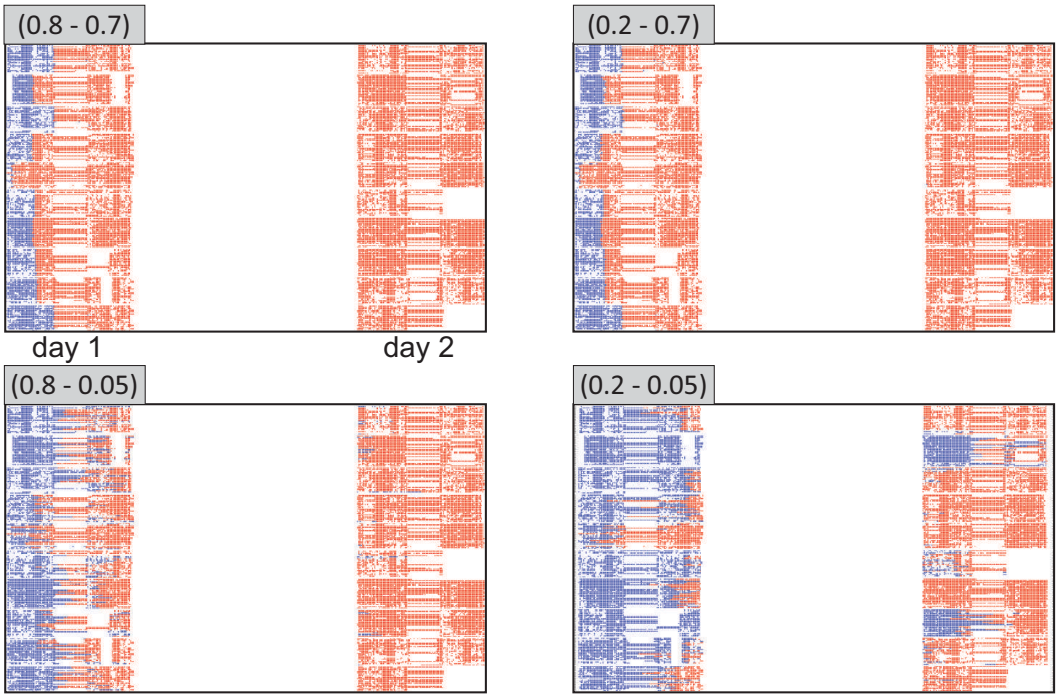

Fig. 8. TAM layouts showing how different combinations (β, λ) affect the infection spread behavior in the Primary School over two days. SI infection model. First row: for $\lambda = 0.7$, varying β did not produce a significant difference in the infection spread behavior. Second row: for $\lambda = 0.05$, different β values produced different infection behaviors.

High School infected the nodes with MTI = 773 (of 1,211), therefore after half of the observation period. This difference in the spread speed can be justified by the different number of edges per timestamp (on average) that each network contains. Since the Primary School contains almost four times more edges than the High School (Table II), its nodes tend to become infected earlier.

Figure 8 shows four TAM layouts, one for each parameter combination and the SI model. For $\lambda = 0.7$, varying β did not produce a significant difference in the infection spread behavior (see Figure 8 – first row), which is in agreement with the results from Table IV. In contrast, when assuming $\lambda = 0.05$, different β values produced different infection behaviors (Figure 8 – second row). When considering $\beta = 0.2$ and $\lambda = 0.05$, for instance, several students entered in the second day of school being susceptible to the infection, most of them belonging to two particular classes (see blue nodes in the second day of the network for ($\beta = 0.2$, $\lambda = 0.05$)).

Figure 9 illustrates the infection path for the combination ($\beta = 0.2$, $\lambda = 0.05$) using the MSV layout. With this representation, one can easily identify the patient zero (central node of class 3B) and see the evolution of the infection spread in relation to the classes. As the patient zero is from class 3B, and considering that students from the same class interact more often among themselves than with others [Gemmetto et al. 2014], the infection propagates rapidly inside class 3B. A few moments later, a student from class 3A becomes infected. Having both 3A and 3B infected one after another is not surprising, since same grade students also interact more often among themselves than with others [Gemmetto et al. 2014]. As illustrated in Figure 9, most of the inter-community interactions that allowed the infection to reach other classes happened during the first day's lunch break. During this interval, students from different classes and grades are free to interact with each other; as shown in the figure, classes 1A, 1B, 4A, and 5B ended up infected during lunch. After the first day of school, only the students from classes 2A and 2B remained susceptible. These students were infected near the second day's lunch break.

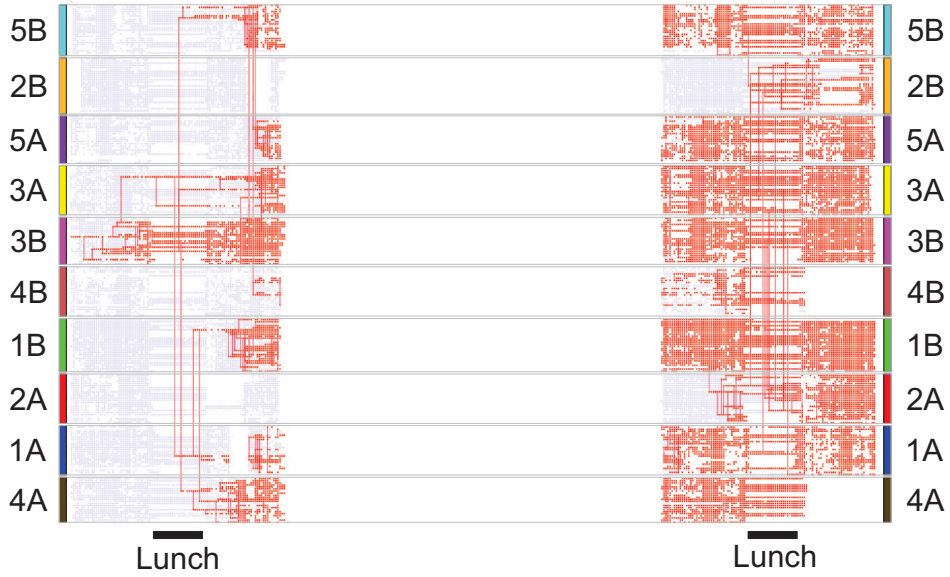

Fig. 9. MSV infection path for the Primary School network ($\beta = 0.2$, $\lambda = 0.05$). SI infection model. The patient zero is the central node of class 3B. The infection propagates rapidly inside this class and, a few moments later, students from class 3A also become infected. Having both 3A and 3B infected one after another is not surprising, since same grade students interact more often among themselves than with others.

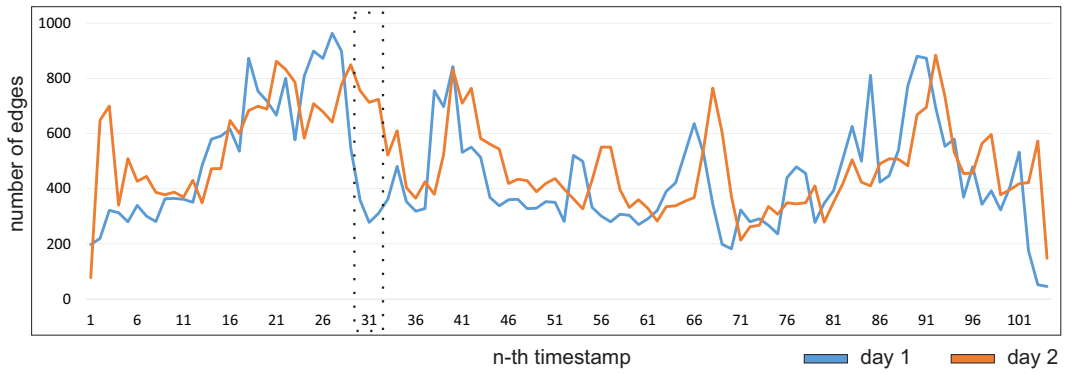

Fig. 10. Comparison of both days interaction behavior, in terms of edge count per timestamp, after filtering $\beta = 0.2$ inter-community edges. Although the chart shows both days having similar ups and downs in the number of edges per timestamp, they have different interaction behaviors (KS distance $KS_d = 0.2596$ with $p\text{-}value = 0.0017$). Two major discrepancies are found in the beginning of each day (approximately between the first and the fourth timestamps) and at the 31st timestamp of each day (see dotted region).

Figure 10 shows a comparison between both days interaction behavior, in terms of edge count per timestamp, after filtering $\beta = 0.2$ inter-community edges. Although we see that day 2 activity presents some delay and differences in the number of edges compared with day 1, the two days have similar ups and downs. Two major discrepancies are found in the beginning of each day (approximately between the first and the fourth timestamps) and in the 31st timestamp of each day (see dotted region). In order to test whether both days have different behaviors, we employed the *two-sample Kolmogorov-Smirnov (KS) statistic*, a measure commonly used to obtain the distance between two empirical cumulative distribution functions [Hassani and Silva 2015; Ahmed et al. 2013; Ponciano et al. 2021]. The KS distance KS_d varies from 0 (identical distributions) to 1 (completely different distributions). In our case, its value represents the distance between the distributions of edge counts

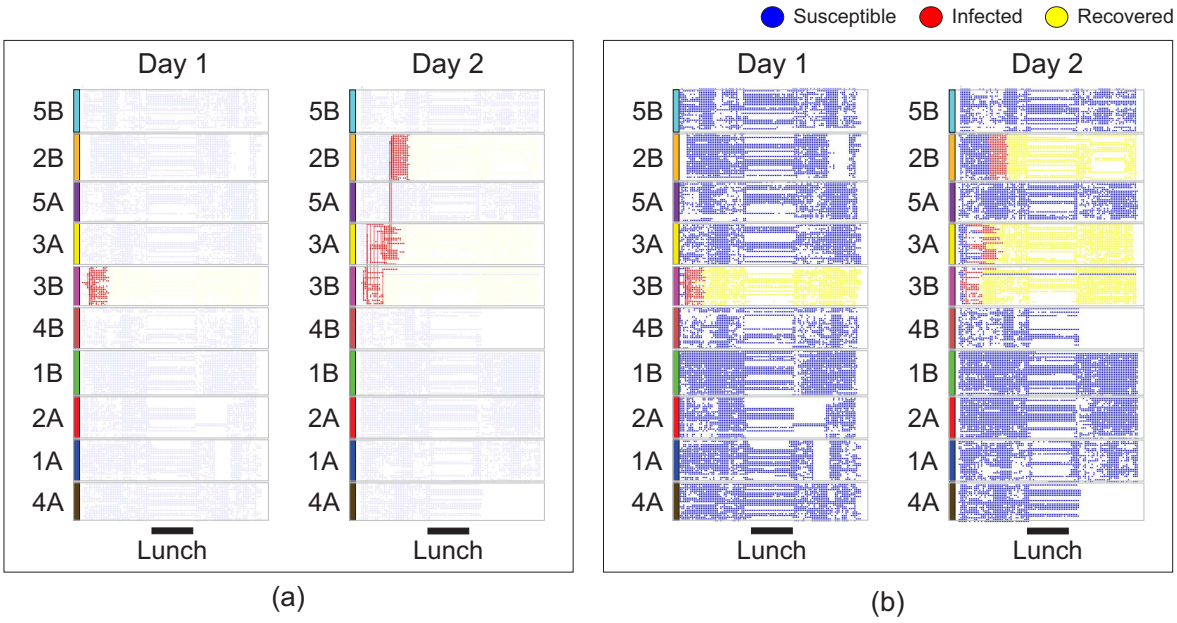

Fig. 11. Two simulations of infection propagation, one per day of the Primary School network. SIR model with ($\beta = 0.2$, $\lambda = 0.7$) and recovery time $r_t = 10$ timestamps. (a) MSV infection path. (b) TAM. There are few interactions at the beginning of day 1, which restricted the infection propagation. In contrast, students from three classes became infected in day 2: 3B, 3A, and 2B.

for day 1 and day 2 after filtering $\beta = 0.2$ inter-community edges. After obtaining $KS_d = 0.2596$ with $p\text{-}value = 0.0017$, we concluded that both days do have different distributions and, therefore, different interaction behaviors.

To help understand how different interaction behaviors affect infection propagation, Figure 11 shows the layouts produced by two simulations using SIR. More specifically, we performed two SIR simulations with the same patient zero p used in the analysis with SI, one starting in the first timestamp of day 1 where p appears and considering only this day, and another starting in the first timestamp of day 2 where p appears and considering only this day. Both simulations use the combination ($\beta = 0.2$, $\lambda = 0.7$) and $r_t = 10$ timestamps, a value empirically chosen to guarantee recovery before lunch break. With few interactions at the beginning of day 1, the infection was restricted to class 3B (recall that p belongs to this class) and all the others remained susceptible. In contrast, the number of interactions at the beginning of day 2 allowed the infection to reach other classes. Besides class 3B, 3A and 2B were also infected. As mentioned before, same grade students tend to interact a lot with each other. This behavior may explain why the infection reached 3A just a few timestamps after its spread on 3B.

Using a node-link diagram with hidden edges, Figure 12 shows snapshots of the node state (susceptible, infected, or recovered) for those classes infected in the simulations depicted in Figure 11. Timestamps $t = 30$ and $t = 316$ were chosen for analysis because they correspond to the 31st timestamp of each day (recall Figure 10). First, we can see how fast class 2B became infected: it took only four timestamps to change from "all nodes susceptible" to "all nodes infected" (Figure 12(b)). We also see that class 3B is the only one containing students that were never infected. The susceptible 3B student depicted at $t = 30$ (first day simulation, Figure 12(a)) and one of those susceptible depicted at $t = 316$ (second day simulation, Figure 12(c)) missed school these days. The other student at $t = 316$ joined school but remained susceptible because he/she had no interactions during the infection propagation — this node can be seen in Figure 11(b, Day 2). Note that the analysis of the missing students would not be possible with TAM or MSV as these layouts only show active nodes. The presented node-link diagram, on the other hand, considers all nodes from the three studied school classes.

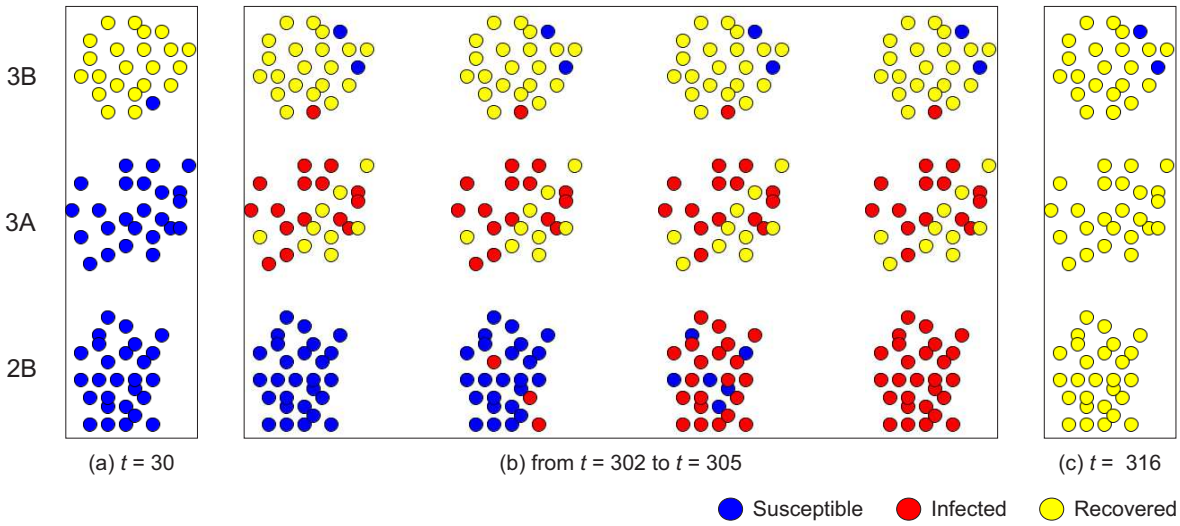

Fig. 12. Node-link diagram with hidden edges showing the node state for students and teachers of classes 2B, 3A, and 3B for particular timestamps during the simulations depicted in Figure 11. It took only four timestamps (from $t = 302$ to $t = 305$) to change the status of the class 2B nodes from "all nodes susceptible" to "all nodes infected". We also see that class 3B is the only one containing students that were never infected.

4.7 InVS case

Contrary to the school networks, where every node (student or teacher) interacts mainly with nodes from its own class, InVS presents three different node interaction behaviors (recall "residents", "linkers", and "wanderers" discussed in Section 4.4). Given that the node-link diagram is useful to analyze the network topology and structure (Section 2.2), Figure 13 shows this layout with node positioning defined by a force-directed layout and colored according to (a) node department and (b) node interaction behavior. As we can see when comparing Figures 13(a,b), "residents" that belong to the same

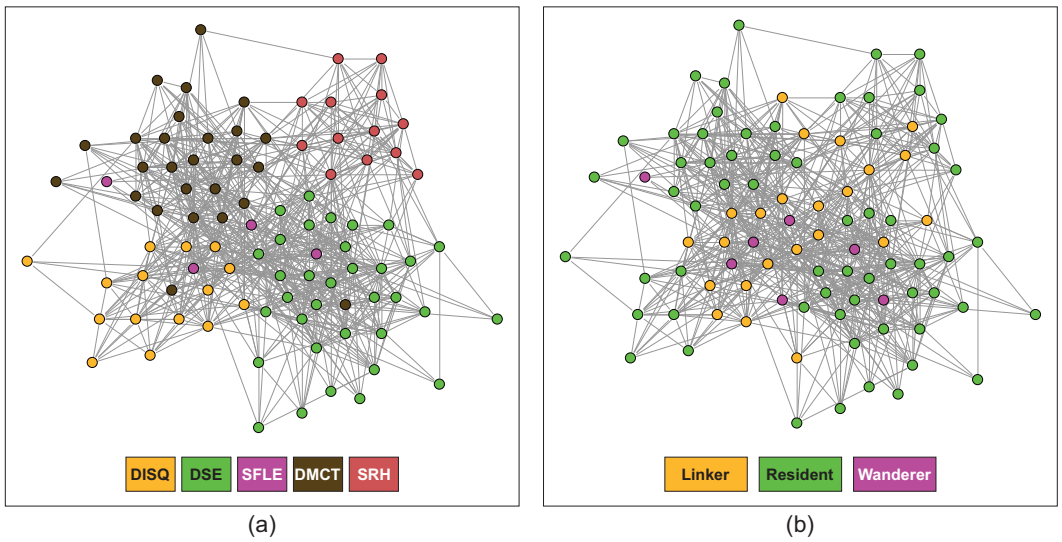

Fig. 13. Node-link diagram with node positioning defined by a force-directed layout and colored according to (a) node department and (b) node interaction behavior. "residents" that belong to the same department are positioned close to each other, while "wanderers" and "linkers" are far away or at their department periphery. Note also that all nodes from the SFLE department are wanderers.

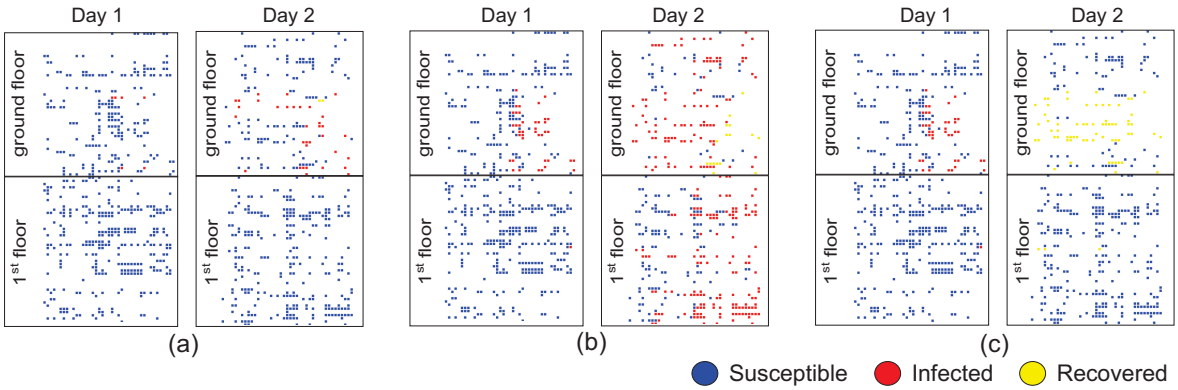

Fig. 14. TAM layouts showing SIR infection simulations. (a) "Resident" patient zero. (b) "Linker" patient zero. (c) Same simulation as in (b) but with a different recovery time r_t. All simulations use ($\beta = 0.2$, $\lambda = 0.7$). In (a) and (b), $r_t = 100$; in (c), $r_t = 50$. Nodes positioned according to the floor their departments belong to. There are less infected nodes when the patient zero is a "resident"; note also that the infection did not reach departments of the first floor in this case.

department are positioned close to each other, while "wanderers" and "linkers" are far away or at their department periphery. Note also that all nodes from the SFLE department are wanderers.

To study whether (and how) different node interaction behaviors would affect an infection spread in the network, we simulated SIR infection dynamics using ($\beta = 0.2$, $\lambda = 0.7$) but varying the patient zero p. Figure 14 shows the infection spread during the first two days of the network through a TAM layout with nodes positioned according to the floor their departments belong to. Note that this layout contains several empty spaces within each day, a consequence of the network sparsity (for comparison with the other networks, recall Figures 6 and 8). In the first simulated scenario (Figure 14(a)), p is the node of type "resident" with the highest degree in the aggregated network (considering the ten days of the network). Analogously, in the second scenario (Figure 14(b)), p is the "linker" with the highest degree. For both scenarios, $r_t = 100$ timestamps (arbitrary value). As we can see, there are less infected nodes when p is a "resident" (compare Figures 14(a,b)). We may think of "residents" as nodes that respect social distancing (they have less contact with nodes from other groups), while "linkers" do not (they interact a lot with at least two groups of nodes). Moreover, contrary to what happened when p was a "linker" (Figure 14(b)), the infection did not reach departments of the first floor when adopting the chosen "resident" p (Figure 14(a)). Not surprisingly, vaccination strategies targeting "linkers" successfully prevent infection outbreaks [Génois et al. 2015]. Our visual analysis supports this affirmation by illustrating the increase in the number of nodes infected when p is a "linker".

SIR simulations are strongly affected by the infection probability λ and by the recovery time r_t. To illustrate the impact of a different r_t value, Figure 14(c) shows a simulation analogous to the one performed in Figure 14(b), i.e., the same patient zero and ($\beta = 0.2$, $\lambda = 0.7$), but now with $r_t = 50$ timestamps. In this case, all nodes infected in day 1 recovered before the beginning of day 2, greatly restricting the number of infected nodes.

5. DISCUSSION

We presented three case studies that simulated infection propagation in real-word temporal social networks. From the visualization perspective, the representations are very important to gain knowledge and generate insights on the data, with the possibility of finding unexpected patterns, trends, and anomalies that would be difficult to find otherwise.

The visual analyses using the studied school networks allowed us to see the infection initially propagating across related classes, such as those from the same grade in Primary School or classes focused on

the same subject in the High School. Moreover, the visualization allowed the identification of classes whose students were infected later than others, and classes whose students were never infected. Using a third network, we also analyzed how different interaction behaviors lead to different infection dynamics. By understanding the infection path, experts can identify who infected whom and when, thus optimizing decision making. In school environments, the school manager could act based on her/his findings. She/He could, e.g., change a class schedule to avoid interactions during lunch and, as a consequence, prevent new infections. Likewise, companies and institutes with employees presenting different interaction behaviors could (dis)encourage specific behaviors during epidemic scenarios.

Our findings are inline with studies that show the importance of reducing the infection probability and interactions involving different groups. In our case, we used λ to consider different infection probabilities while β stands for the level of allowed inter-group interaction (an indication of social distancing compliance). Our findings, however, should only serve as a basis for future researches as several aspects of infection propagation were not considered in our study. Examples include: (i) individuals or communities may adopt different precautions to reduce their susceptibility [Kitchovitch and Liò 2010]; (ii) different age-groups (or based on other criteria) may be differently affected [Canabarro et al. 2020]; (iii) the need for analyses considering more environments (e.g., networks from domains such as hospitals or shopping malls). As expected, the lower the number of interactions and infection probability, the less infected nodes at the end of the observation period. The main challenge is to find the parameter combination that best represents existing real-world scenarios and expected ones.

For all three analyzed networks, the community detection methods returned communities highly related to the school classes and institute departments. We have considered community detection instead of the direct use of class/department information because there are several real-world networks in which the community structure is not so clear. We thus aimed to demonstrate the suitability of our approach for the analysis of a variety of real-world networks, regardless of previous knowledge about existing groups of individuals. Nevertheless, the number of communities, their sizes, the quality of the detection, and the community structure strength may impact infection propagation and should be further analyzed [Ghalmane et al. 2019; Leão et al. 2019].

For convenience, our evaluation considered only SI and SIR, two simple infection models, to analyze the presented real-world networks. There are more complex models that consider other variables, such as death of individuals and latent/exposed states. When studying a particular disease, one must first identify which infection model is more suitable given the disease's characteristics.

Even though inter-community edges are less relevant than intra-ones [Linhares et al. 2019], they may still contain relevant information that is lost with the edge filtering. A crucial decision in this context is, therefore, how to choose which edges will be removed. In our experiments, we have adopted a random removal of inter-community edges, however, more sophisticated algorithms could have been employed (e.g., SEVis [Ponciano et al. 2021] or strategies based on network centrality values).

Timeline-based layouts (e.g., MSV and TAM) plot all nodes, edges (if the case), and times as a single picture (recall Section 2.2). Although this characteristic makes them suitable for analyzing temporal networks where all network data is known in advance, they cannot be used to analyze real-time infection spread because there is no guarantee all network data will fit in primary memory. To handle these cases, animated layouts such as animated node-link diagrams or hybrid strategies (timeline plus animation) are preferred. Dealing with the user's mental map preservation and aspects such as the transition speed and fading factor effect [Linhares et al. 2021] would then become crucial.

6. CONCLUSION

Network visualization strategies include the user in the network data exploration and validation through graphical and interactive representations. In the context of infection spreads, an effective visual analysis provides means for quick and reliable identification of key individuals, epidemics out-

break, transmission path, and contact tracing, among other relevant behaviors and patterns.

This paper presented a visualization approach for the simulation and analysis of infection spread dynamics. Our approach considers two key aspects: different levels of social distancing (i.e., whether the individuals avoid interacting with others), defined in our case by the number of inter-community (inter-group) interactions, and different infection probabilities (which could vary in real-world scenarios according to the use of protective measures such as face masks, for example). We demonstrated the usefulness of our approach through case studies that considered real-world interactions modeled by three temporal social networks. Our quantitative and visual analyses showed the overall influence of different levels of inter-group interaction and infection probabilities and how local investigations lead to optimized understanding and reliable decision making.

As future work, we intend to analyze the impact of the community size in the infection spread. Furthermore, we plan a follow-up study that considers other infection models and a more rigorous parameter analysis, as well as simulations using networks from other environments. Finally, our plans also include automatically suggest potential initial patients for analysis and conduct a user evaluation with health experts (e.g., epidemiologists) to validate usefulness and usability.

Acknowledgements

This research was supported by Coordenação de Aperfeicoamento de Pessoal de Nível Superior, São Paulo Research Foundation (FAPESP - Grants number 2020/10049-0, 2020/07200-9, 2016/17078-0), and the School of Applied Mathematics at Getulio Vargas Foundation (FGV). The authors also thank SocioPatterns for making available the network data sets used in this paper.

REFERENCES

Ahmed, N. K., Neville, J., and Kompella, R. Network sampling: From static to streaming graphs. *ACM Trans. Knowl. Discov. Data* 8 (2): 7:1–7:56, June, 2013.

Barrat, A., Barthélemy, M., and Vespignani, A. *Dynamical Processes on Complex Networks*. Cambridge University Press, 2008.

Battista, G. D., Eades, P., Tamassia, R., and Tollis, I. G. Algorithms for drawing graphs: an annotated bibliography. *Computational Geometry* 4 (5): 235 – 282, 1994.

Beck, F., Burch, M., Diehl, S., and Weiskopf, D. A taxonomy and survey of dynamic graph visualization. *Computer Graphics Forum* 36 (1): 133–159, 2016.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* vol. 2008, pp. P10008, 2008.

Brandes, U. pp. 1–6. In M.-Y. Kao (Ed.), *Force-Directed Graph Drawing*. Springer US, Boston, MA, pp. 1–6, 2008.

Canabarro, A., Tenorio, E., Martins, R., Martins, L., Brito, S., and Chaves, R. Data-driven study of the covid-19 pandemic via age-structured modelling and prediction of the health system failure in brazil amid diverse intervention strategies. *medRxiv*, 2020.

Carroll, L. N., Au, A. P., Detwiler, L. T., Fu, T.-c., Painter, I. S., and Abernethy, N. F. Visualization and analytics tools for infectious disease epidemiology: a systematic review. *Journal of biomedical informatics* vol. 51, pp. 287–298, 2014.

Chang, S., Harding, N., Zachreson, C., Cliff, O. M., and Prokopenko, M. Modelling transmission and control of the covid-19 pandemic in australia. *ArXiv* vol. abs/2003.10218, 2020.

Daghriri, T. and Ozmen, O. Quantifying the effects of social distancing on the spread of covid-19. *Journal of Vaccines & Vaccination*, 2020.

Dong, E., Du, H., and Gardner, L. An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases* 20 (5): 533–534, 2020.

Dunne, C., Muller, M., Perra, N., and Martino, M. Vorograph: Visualization tools for epidemic analysis. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '15. Association for Computing Machinery, New York, NY, USA, pp. 255–258, 2015.

Fortunato, S. and Hric, D. Community detection in networks: A user guide. *Physics Reports* 659 (Supplement C): 1 – 44, 2016.

Gemmetto, V., Barrat, A., and Cattuto, C. Mitigation of infectious disease at school: targeted class closure vs school closure. *BMC infectious diseases* 14 (1): 695, Dec., 2014.

Génois, M., Vestergaard, C. L., Fournet, J., Panisson, A., Bonmarin, I., and Barrat, A. Data on face-to-face contacts in an office building suggest a low-cost vaccination strategy based on community linkers. *Network Science* 3 (3): 326–347, 2015.

Ghalmane, Z., El Hassouni, M., and Cherifi, H. Immunization of networks with non-overlapping community structure. *Social Network Analysis and Mining* 9 (1): 45, 2019.

Hassani, H. and Silva, E. A kolmogorov-smirnov based test for comparing the predictive accuracy of two sets of forecasts. *Econometrics* vol. 3, pp. 590–609, 08, 2015.

Holme, P. and Saramäki, J. Temporal networks. *Physics Reports* 519 (3): 97–125, October, 2012.

Jr., L. G. and Frizzon, G. Fake news and brazilian politics – temporal investigation based on semantic annotations and graph analysis. In *Anais do XXXIV SBBD*. SBC, Porto Alegre, RS, Brasil, pp. 169–174, 2019.

Kitchovitch, S. and Liò, P. Risk perception and disease spread on social networks. *Procedia Computer Science* 1 (1): 2345 – 2354, 2010. ICCS 2010.

Leão, J., Laender, A., and de Melo, P. A multi-strategy approach to overcoming bias in community detection evaluation. In *Anais do XXXIV SBBD*. SBC, Porto Alegre, RS, Brasil, pp. 13–24, 2019.

Linhares, C., Ponciano, J., Pereira, F., Rocha, L., Paiva, J., and Travençolo, B. Visual analysis for evaluation of community detection algorithms. *Multimedia Tools and Applications* 79 (25): 17645–17667, 2020.

Linhares, C. D. G., Lima, D. M., Bones, C. C., Rebelo, M. F. S., Gutierrez, M. A., Traina, C., and Traina, A. J. M. I-covidvis – a visual analytics tool for interoperable healthcare databases using graphs. In *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*. pp. 125–130, 2021.

Linhares, C. D. G., Lima, D. M., Ponciano, J. R., Olivatto, M. M., Gutierrez, M. A., Poco, J., Traina, C., and Traina, A. J. M. ClinicalPath: a visualization tool to improve the evaluation of electronic health records in clinical decision-making. *IEEE Transactions on Visualization and Computer Graphics*, 2022.

Linhares, C. D. G., Ponciano, J. R., Paiva, J. G. S., Rocha, L. E. C., and Travençolo, B. A. N. DyNetVis - an interactive software to visualize structure and epidemics on temporal networks. In *2020 IEEE/ACM ASONAM*. pp. 933–936, 2020.

Linhares, C. D. G., Ponciano, J. R., Paiva, J. G. S., Travençolo, B. A. N., and Rocha, L. E. C. pp. 83–105. In , *Visualisation of Structure and Processes on Temporal Networks*. Springer International Publishing, Cham, pp. 83–105, 2019.

Linhares, C. D. G., Ponciano, J. R., Paiva, J. G. S., Travençolo, B. A. N., and Rocha, L. E. C. A comparative analysis for visualizing the temporal evolution of contact networks: a user study. *Journal of Visualization* 24 (5): 1011–1031, Oct, 2021.

Linhares, C. D. G., Ponciano, J. R., Pereira, F. S. F., Rocha, L. E. C., Paiva, J. G. S., and Travençolo, B. A. A scalable node ordering strategy based on community structure for enhanced temporal network visualization. *Computers & Graphics* vol. 84, pp. 185 – 198, 2019.

Linhares, C. D. G., Travençolo, B. A. N., Paiva, J. G. S., and Rocha, L. E. C. DyNetVis: a system for visualization of dynamic networks. In *Proceedings of the Symposium on Applied Computing*. SAC '17. ACM, Marrakech, Morocco, pp. 187–194, 2017.

Mastrandrea, R., Fournet, J., and Barrat, A. Contact patterns in a high school: A comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PLOS ONE* 10 (9): 1–26, 09, 2015.

Park, J. Y. Spatial visualization of cluster-specific covid-19 transmission network in south korea during the early epidemic phase. *medRxiv*, 2020.

Ponciano, J., Vezono, G., and Linhares, C. Simulating and visualizing infection spread dynamics with temporal networks. In *Anais do XXXVI Simpósio Brasileiro de Bancos de Dados*. SBC, Porto Alegre, RS, Brasil, pp. 37–48, 2021.

Ponciano, J. R., Linhares, C. D. G., Faria, E. R., and Travençolo, B. A. N. An online and nonuniform timeslicing method for network visualisation. *Computers & Graphics* vol. 97, pp. 170–182, 2021.

Ponciano, J. R., Linhares, C. D. G., Melo, S. L., Lima, L. V., and Travençolo, B. A. N. Visual analysis of contact patterns in school environments. *Informatics in Education* 19 (3): 455–472, 2020.

Ponciano, J. R., Linhares, C. D. G., Rocha, L. E. C., Faria, E. R., and Travençolo, B. A. N. Combining clutter reduction methods for temporal network visualization. In *Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing*. SAC '22. Association for Computing Machinery, New York, NY, USA, pp. 1748–1755, 2022.

Ponciano, J. R., Linhares, C. D. G., Rocha, L. E. C., Faria, E. R., and Travençolo, B. A. N. A streaming edge sampling method for network visualization. *Knowledge and Information Systems* vol. 63, pp. 1717–1743, Apr, 2021.

Prakash, B. A., Vreeken, J., and Faloutsos, C. Efficiently spotting the starting points of an epidemic in a large graph. *Knowledge and information systems* 38 (1): 35–59, 2014.

Rocha, L. E. C. and Blondel, V. D. Bursts of vertex activation and epidemics in evolving networks. *PLOS Computational Biology* vol. 9, pp. 1–9, 2013.

Rosvall, M. and Bergstrom, C. T. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences* 105 (4): 1118–1123, 2008.

So, M., Tiwari, A., Chu, A., Tsang, J., and Chan, J. Visualizing covid-19 pandemic risk through network connectedness. *International Journal of Infectious Diseases* vol. 96, pp. 558 – 561, 2020.

Tepper, J. G. and Thiébaut, D. Data visualization of agent-based simulation of an infectious spread. In *INFOCOMP 2017*, 2017.