HAL open science

Predicting Epidemic Risk from Past Temporal Contact Data

Eugenio Valdano, Chiara Poletto, Armando Giovannini, Diana Palma, Lara Savini, Vittoria Colizza

▶ **To cite this version:**

Eugenio Valdano, Chiara Poletto, Armando Giovannini, Diana Palma, Lara Savini, et al.. Predicting Epidemic Risk from Past Temporal Contact Data. PLoS Computational Biology, 2015, 11 (3), pp.e1004152. 10.1371/journal.pcbi.1004152 . hal-01224104

HAL Id: hal-01224104

https://hal.sorbonne-universite.fr/hal-01224104

Submitted on 4 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Predicting Epidemic Risk from Past Temporal Contact Data

Eugenio Valdano [1,2], Chiara Poletto [1,2], Armando Giovannini [3], Diana Palma [3] Lara Savini [3], Vittoria Colizza [1,2,4]*

1 INSERM, UMR-S 1136, Institut Pierre Louis d'Epidémiologie et de Santé Publique, F-75013 56 bd Vincent Auriol—CS 81393-75646 Paris Cedex 13, France, **2** Sorbonne Universités, UPMC Univ Paris 06, UMR-S 1136, Institut Pierre Louis d'Epidémiologie et de Santé Publique, F-75013 56 bd Vincent Auriol—CS 81393 - 75646 Paris Cedex 13, France, **3** Istituto Zooprofilattico Sperimentale Abruzzo-Molise G. Caporale Campo Boario, 64100 Teramo, Italy, **4** ISI Foundation Via Alassio 11/c, 10126 Torino, Italy

* vittoria.colizza@inserm.fr

Abstract

Understanding how epidemics spread in a system is a crucial step to prevent and control outbreaks, with broad implications on the system's functioning, health, and associated costs. This can be achieved by identifying the elements at higher risk of infection and implementing targeted surveillance and control measures. One important ingredient to consider is the pattern of disease-transmission contacts among the elements, however lack of data or delays in providing updated records may hinder its use, especially for time-varying patterns. Here we explore to what extent it is possible to use past temporal data of a system's pattern of contacts to predict the risk of infection of its elements during an emerging outbreak, in absence of updated data. We focus on two real-world temporal systems; a livestock displacements trade network among animal holdings, and a network of sexual encounters in high-end prostitution. We define the node's loyalty as a local measure of its tendency to maintain contacts with the same elements over time, and uncover important non-trivial correlations with the node's epidemic risk. We show that a risk assessment analysis incorporating this knowledge and based on past structural and temporal pattern properties provides accurate predictions for both systems. Its generalizability is tested by introducing a theoretical model for generating synthetic temporal networks. High accuracy of our predictions is recovered across different settings, while the amount of possible predictions is system-specific. The proposed method can provide crucial information for the setup of targeted intervention strategies.

OPEN ACCESS

Citation: Valdano E, Poletto C, Giovannini A, Palma D, Savini L, Colizza V (2015) Predicting Epidemic Risk from Past Temporal Contact Data. PLoS Comput Biol 11(3): e1004152. doi:10.1371/journal.pcbi.1004152

Editor: Samuel Alizon, CNRS, FRANCE

Received: July 28, 2014

Accepted: January 23, 2015

Published: March 12, 2015

Copyright: © 2015 Valdano et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was funded by: EC-Health contract no. 278433 (PREDEMICS—Prevention of Emerging Zoonotic Viruses with Pandemic Potential using Multidisciplinary Approaches—predemics.biomedtrain.eu). ANR contract no. ANR- 12-MONU-0018 (HARMSFLU—Harmonising multiple scales for data-driven computational approaches to the modeling of influenza spread—www.harmsflu.org). Programme Galilee (http://www.campusfrance.org/fr/galilee) contract no. 28144NH. Italian Ministry of Health (www.salute.gov.it) contract no. IZS AM 04/11 RC; EC-ANIHWA (Animal Health and Welfare ERA-

Author Summary

Following the emergence of a transmissible disease epidemic, interventions and resources need to be prioritized to efficiently control its spread. While the knowledge of the pattern of disease-transmission contacts among hosts would be ideal for this task, the continuously changing nature of such pattern makes its use less practical in real public health emergencies (or otherwise highly resource-demanding when possible). We show that in

Net—www.anihwa.eu) contract no. ANR-13-ANWA-0007-03 (LIVEpi). "Pierre Louis" School of Public Health of UPMC (www.ed393.upmc.fr). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

such situations critical knowledge to assess the real-time risk of infection can be extracted from past temporal contact data. An index expressing the conservation of contacts over time is proposed as an effective tool to prioritize interventions, and its efficiency is tested considering real data on livestock movements and on human sexual encounters.

Introduction

Being able to promptly identify who, in a system, is at risk of infection during an outbreak is key to the efficient control of the epidemic. The explicit pattern of potential disease-transmission contacts has been extensively used to this purpose in the framework of theoretical studies of epidemic processes, uncovering the role of the pattern's properties in the disease propagation and epidemic outcomes [1, 2, 3, 4, 5, 6, 7, 8]. These studies are generally based on the assumption that the entire pattern of contacts can be mapped out or that its main properties are known. Although such knowledge would be a critical requirement to conduct risk assessment analyses in real-time, which need to be based on the updated and accurate description of the contacts relevant to the outbreak under study [9], it can hardly be obtained in reality. Given the lack of such data, analyses generally refer to the most recent available knowledge of contact data, implicitly assuming a non-evolving pattern.

The recent availability of time-resolved data characterizing connectivity patterns in various contexts [10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22] has inevitably weakened the non-evolving assumption, bringing new challenges to the assessment of nodes' epidemic risk. Traditional centrality measures used to identify vulnerable elements or influential spreaders for epidemics circulating on static networks [1, 2, 4, 23, 24, 25, 26, 27, 28, 29, 30] are unable to provide meaningful information for their control, as these quantities strongly fluctuate in time once computed on the evolving networks [19, 31]. An element of the system may thus act as *superspreader* in a past configuration of the contact network, having the ability to potentially infect a disproportionally larger amount of secondary contacts than other elements [32], and then assume a more peripheral role in the current pattern of contact or even become isolated from the rest of the system [19]. If the rules driving the change of these patterns over time are not known, what information can be extracted from past contact data to infer the risk of infection for an epidemic unfolding on the current (unknown) pattern?

Few studies have so far tried to answer this question by exploiting temporal information to control an epidemic through targeted immunization. They are based on the extension to temporal networks [33, 34] of the so-called acquaintance immunization protocol [4] introduced in the framework of static networks that prescribes to vaccinate a random contact of a randomly chosen element of the system. In the case of contacts relevant for the spread of sexually transmitted infections, Lee et al. showed that the most efficient protocol consists in sampling elements at random and vaccinating their latest contacts [33]. The strategy is based on local information gathered from the observation and analysis of past temporal data, and it outperforms static-network protocols. Similar results are obtained for the study of face-to-face contact networks relevant for the transmission of acute respiratory infections in a confined setting, showing in addition that a finite amount of past network data is in fact needed to devise efficient immunization protocols [34].

The aim of these studies is to provide general protocols of immunization over all possible epidemiological conditions of the disease (or class of diseases) under study. For this reason, protocols are tested through numerical simulations and results are averaged over starting seeds and times to compare their performance. Previous work has however shown that epidemic

outcomes may strongly depend on the temporal and geographical initial seed of the epidemic [35], under conditions of large dynamical variability of the network and absence of stable structural backbones [19]. Our aim is therefore to focus on a specific epidemiological condition relative to a given emerging outbreak in the population, resembling a realistic situation of public health emergency. We focus on the outbreak initial phase prior to interventions when facing the difficulty that some infected elements in the population are not yet observed. The objective is to assess the risk of infection of nodes to inform targeted surveillance, quarantine and immunization programs, assuming the lack of knowledge of the explicit contact pattern on which the outbreak is unfolding. Knowledge is instead gathered from the analysis of the full topological and temporal pattern of past data (similarly to previous works [33, 34]), coupled, in addition, with epidemic spreading simulations performed on such data under the same epidemiological conditions of the outbreak under study. More specifically, we propose an egocentric view of the system and assess whether and to what extent the node's tendency of repeating already established contacts is correlated with its probability of being reached by the infection. Findings obtained on past available contact data are then used to predict the infection risk in the current unknown epidemic situation. We apply this risk assessment analysis to two large-scale empirical datasets of temporal contact networks—cattle displacements between premises in Italy [19, 36], and sexual contacts in high-end prostitution [16]—and evaluate its performance through epidemic spreading simulations. We also introduce a model to generate synthetic time-varying networks retaining the basic mechanisms observed in the empirical networks considered, in order to explain the results obtained by the proposed risk assessment strategy within a general theoretical framework.

Results and Discussion

The cattle trade network is extracted from the complete dataset reporting on time-resolved bovine displacements among animal holdings in Italy [19, 36] for the period 2006–2010, and it represents the time-varying contact pattern among the 215,264 premises composing the system. The sexual contact network represents the connectivity pattern of sexual encounters extracted from a Web-based Brazilian community where sex buyers provide time-stamped rating and comments on their experiences with escorts [16].

The five-years data of the livestock trade network show that stationary properties at the global level co-exist with an active non-trivial local dynamics. The probability distributions of several quantities measured on the different yearly networks are considerably stable over time, as e.g. shown by the in-degree distribution reported in Fig. 1A, where the in-degree of a farm measures the number of premises selling cattle to that farm. These features, however, result from highly fluctuating underlying patterns of contacts, never preserving more than 50% of the links from one yearly configuration to another (Fig. 1C), notwithstanding the seasonal annual pattern due to repeating cycles of livestock activities [37, 38] (see S1 Text). Similar findings are also obtained for the sexual contact network (Fig. 1B-D), where the lack of an intrinsic cycle of activity characterizing the system leads to smaller values of the overlap between different configurations ($< 10\%$). In this case we consider semi-annual configurations, an arbitrary choice that allows us to extract six network configurations in a timeframe exhibiting an approximately stationary average temporal profile of the system, after discarding an initial transient time period from the data [16]. Different time-aggregating windows are also considered (see the Materials and Methods section and S1 Text for additional details).

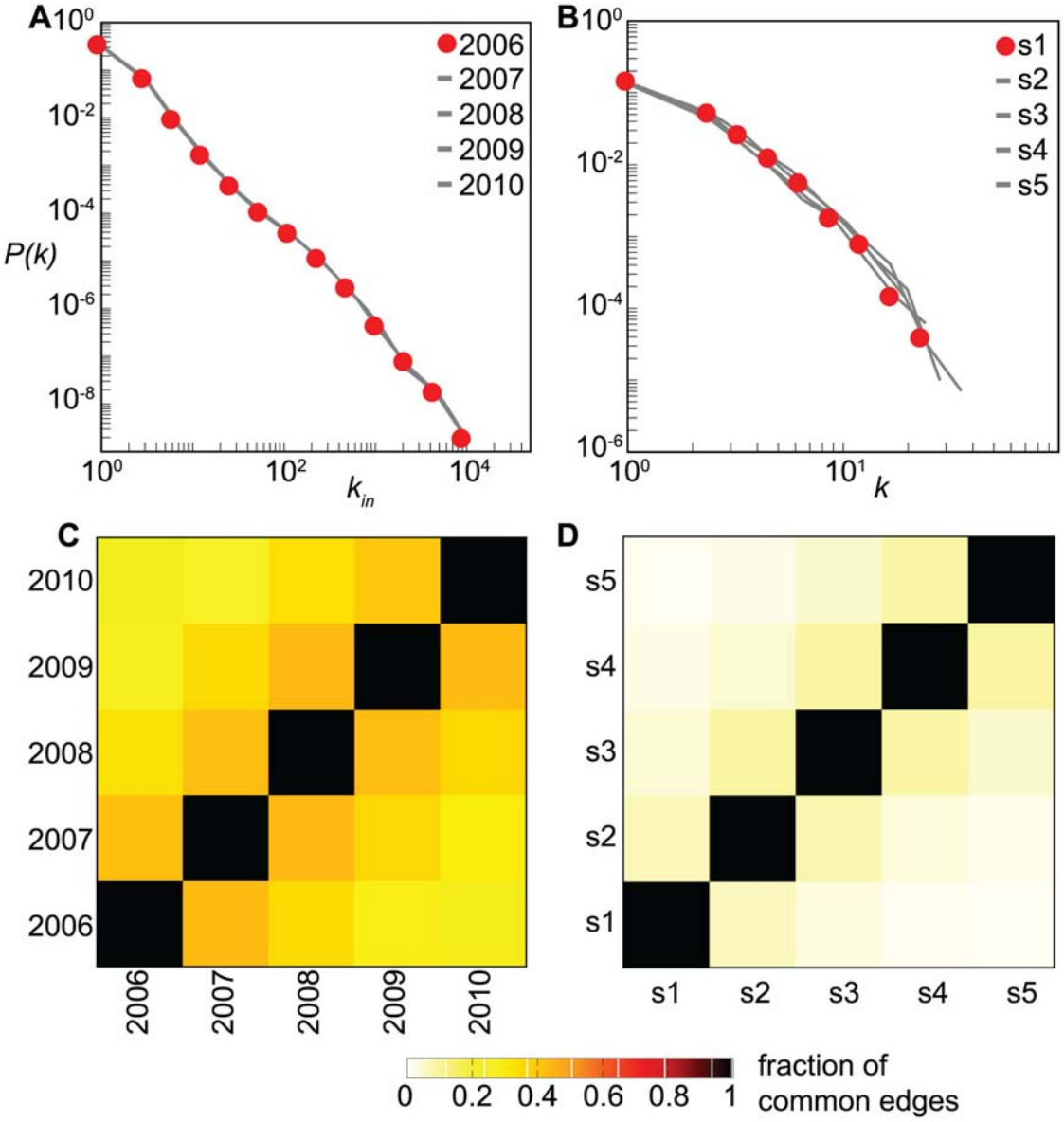

Fig 1. Structural and temporal properties of the cattle trade network and of the sexual contact network. (*A*), (*B*): premises in-degree distributions in the cattle trade network and sex customers degree distribution in the sexual contact network, respectively. Distributions for different configurations of the networks are superimposed in both cases. (*C*), (*D*): fraction of common edges contained in two configurations of the network, for the cattle trade network and the sexual contact network, respectively. In (*B*), (*D*) *s* stands for semester, the aggregation interval of each configuration.

doi:10.1371/journal.pcbi.1004152.g001

Loyalty

The observed values of the overlap of the time-resolved contact networks in terms of the number of links preserved are a measure of the degree of memory contained in the system. This is the outcome of the temporal activity of the elements of the system that reshape up to 50% or 90% of the contacts of the network (in the cattle trade case and in the sexual contact case, respectively), through nodes' appearance and disappearance, and neighborhood restructuring. By framing the problem in an egocentric perspective, we can explore the behavior of each single node of the system in terms of its tendency to remain active in the system and re-establish

connections with the same partners vs. the possibility to change partners or make no contacts. We quantitatively characterize this tendency by introducing the *loyalty* θ, a quantity that measures the fraction of preserved neighbors of a node for a pair of two consecutive network configurations in time, $c-1$ and c. If we define $\mathcal{V}_i^{c-1}$ as the set of neighbors of node i in configuration $c-1$, then $\theta_i^{c-1,c}$ is given by the Jaccard index between $\mathcal{V}_i^{c-1}$ and $\mathcal{V}_i^c$:

$$\theta_i^{c-1,c} = \frac{\left|\mathcal{V}_i^{c-1} \cap \mathcal{V}_i^c\right|}{\left|\mathcal{V}_i^{c-1} \cup \mathcal{V}_i^c\right|}. \tag{1}$$

Loyalty takes values in the interval [0, 1], with $\theta = 0$ indicating that no neighbors are retained, and $\theta = 1$ that exactly the same set of neighbors is preserved ($\mathcal{V}_i^{c-1} = \mathcal{V}_i^c$). It is defined for discrete time windows (c, c+1) and in general it depends on the aggregation interval chosen to build network configurations.

In case the network is directed, as for example the cattle trade network, θ can be equivalently computed on the set $\mathcal{V}_{in,i}^c$ of incoming contacts or on the set of neighbors of outgoing connections, $\mathcal{V}_{out,i}^c$, depending on the system-specific interpretation of the direction and on the interest in one phenomenon or the opposite. This measure originally finds its inspiration in the study of livestock trade networks, where a directed connection from holding A to holding B indicates that B purchased a livestock batch from A, which was then displaced along the link direction $A \rightarrow B$. If we compute θ on the incoming contacts of the cattle trade network, we thus quantify the propensity of each farmer to repeat business deals with the same partners when they purchase their cattle. This concept is at the basis of many loyalty or fidelity programs that propose explicit marketing efforts to incentivize the reinforcement of loyal buying behavior between a purchasing client and a selling company [39], and corresponds to a principle of exclusivity in selecting economic and social exchange partners [40, 41]. Analogously, in the case of the sexual contact network we consider the point of view of sex buyers. Formally, our methodology can be carried out with the opposite point of view, by considering out-degrees with loyalties being computed on out-neighbors. Our choice is arbitrary and inspired by the trade mechanism underlying the network evolution.

Other definitions of similarity to measure the loyal behavior of a node are also possible. In S1 Text we compare and discuss alternative choices. For the sake of clarity all symbols and variables used in the article are reported in Table 1. Finally, other mechanisms different from fidelity strategies may be at play that result in the observed behavior of a given node. In absence of additional knowledge on the behavior underlying the network evolution, we focus on the loyalty θ to explore whether it can be used as a possible indicator for infection risk, as illustrated in the following subsection.

The distributions of loyalty values, though of different shapes across the two datasets, display no considerable variation moving along consecutive pairs of configurations of each dataset (Fig. 2A-B and S1 Text), once again indicating the overall global stability of system's properties in time and confirming the results observed for the degree. A diverse range of behaviors in establishing new connections vs. repeating existing ones is observed, similarly to the stable or exploratory strategies found in human communication [42]. Two pronounced peaks are observed for $\theta = 0$ and $\theta = 1$, both dominated by low degree nodes for which few loyalty values are allowed, given the definition of Eq. (1) (see S1 Text for the dependence of θ on nodes' degree and its analytical understanding). The exact preservation of the neighborhood structure ($\theta = 1$) is more probable in the cattle trade network than in the sexual contact network ($P(\theta = 1)$ being one order of magnitude larger), in agreement with the findings of a higher system-wide memory reported in Fig. 1. Moreover, the cattle trade network exhibits the presence of high loyalty values (in the range $\theta \in [0.7, 0.9]$), differently from the sexual contact network where P

Table 1. List of variables and their description.

Notation	Description
c	index for network configurations
θ or $\theta_i^{c-1,c}$	loyalty of node i between configurations c–1, c
$\mathcal{V}_{in,i}^c$, $\mathcal{V}_{out,i}^c$	set of in(out)-neighbors of i in config c
L, D	loyalty classes (loyal, disloyal)
ε	loyalty threshold
s	epidemic seed
τ	duration of the outbreak early stage
$\mathcal{I}_s^c$	set of infected nodes for outbreak starting from s in config c
$\pi_D^{c-1,c}(s)$, $\pi_L^{c-1,c}(s)$	infection potentials for class D (L) computed for seed s between configs c–1, c
k	degree (in-degree for the cattle trade network)
$T_{DD}^c(k)$, $T_{DL}^c(k)$, $T_{LD}^c(k)$, $T_{LL}^c(k)$	transition probability from one loyalty class to another
ρ_i^c	epidemic risk for node i in config c
$\mathcal{I}_{s,h}^c$, $\mathcal{I}_{s,l}^c$	set of infected nodes with high(low) epidemic risk
P_h, P_l	probability of a high(low) risk node to be infected
$v = P_h/P_l$	risk ratio between P_h, P_l, measure of accuracy
$\omega_s^{c-1,c}$	predictive power (fraction of infected nodes for which it is possible to compute the epidemic risk)
b, d	node probability of becoming active or inactive
p_α	node probability of keeping an in-neighbor
α	number of kept in-neighbors
β_{in}	number of new in-neighbors
β_{out}	number of new out-neighbors
γ, δ	exponents of the distributions of β_{in}, β_{out}

doi:10.1371/journal.pcbi.1004152.t001

(θ) is always equal to zero in that range except for one pair of consecutive configurations giving a positive probability for $\theta = 0.8$. Farmers in the cattle trade network thus display a more loyal behavior in purchasing cattle batches from other farmers with respect to how sex buyers establish their sexual encounters in the analyzed sexual contact dataset.

For the sake of simplification, we divide the set of nodes composing each system into the subset of *loyal nodes* having θ greater than a given threshold ϵ, and the subset of *disloyal nodes* if instead $\theta < \epsilon$. We call hereafter these classes as *loyalty statuses* L and D, respectively, and we will later discuss the role of the chosen value for ϵ.

Epidemic simulations and risk of infection

Both networks under study represent substrates offering potential opportunities for a pathogen to diffuse in the corresponding populations. Sexually transmitted infections spread among the population of individuals through sexual contacts [43, 44], whereas livestock infectious diseases (e.g. Foot-and-mouth disease [45], Bluetongue virus [46], or BVD [47]) can be transmitted from farm to farm mediated by the movements of infected animals (and vectors, where relevant), potentially leading to a rapid propagation of the disease on large geographical scales.

As a model for disease-transmission on the network of contacts we consider a discrete-time Susceptible-Infectious compartmental approach [48]. No additional details characterizing the course of infection are considered here (e.g. recovery dynamics), as we focus on a simplified theoretical picture of the main mechanisms of pathogen diffusion and their interplay with the

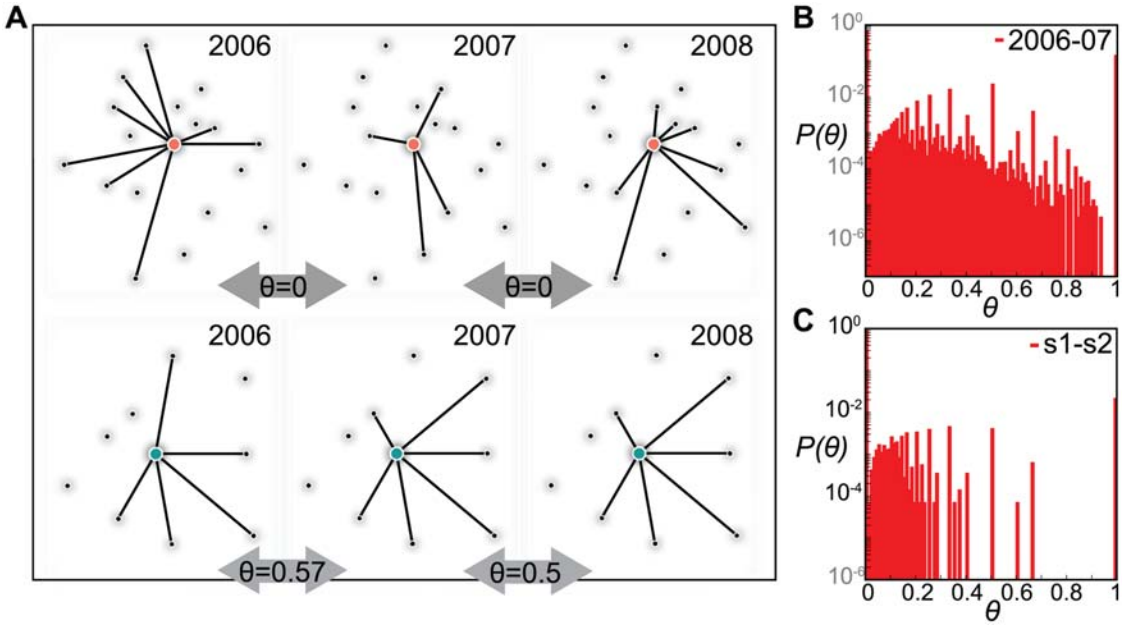

Fig 2. Loyalty. (*A*) Visualization of the neighborhood of two different farms in the cattle trade network (orange node, characterized by low loyalty, and green node, characterized by high loyalty) and corresponding loyalty values computed on three consecutive configurations (2006, 2007, 2008). (*B*), (*C*): Loyalty distributions in the cattle trade network and in the sexual contact network, respectively. Histograms refer to the first pair of consecutive configurations for visualization purposes, all other distributions being reported in S1 Text and showing stability across time.

doi:10.1371/journal.pcbi.1004152.g002

network topology and time-variation, for the prediction of the risk of infection. The aim is to provide a general and conceptually simple framework, leaving to future studies the investigation of more detailed and realistic disease natural histories.

At each time step, an infectious node can transmit the disease along its outgoing links to its neighboring susceptible nodes that become infected and can then propagate the disease further in the network. Here, we consider a deterministic process for which the contagion occurs with probability equal to 1, as long as there exist a link connecting the infectious node to a susceptible one. Although a crude assumption, this allows us to simplify the computational aspects while focusing on the risk prediction. The corresponding stochastic cases exploring lower probabilities of transmission per link are reported in S1 Text.

We focus on the early phase of the spreading simulations, defined as the set of nodes infected up to simulation time step $\tau = 6$. This choice allows us to study invasion stage only, while the epidemic is no more trivially confined to the microscopic level. Additional choices for τ have been investigated showing that they do not alter our findings (see S1 Text). Network configurations are kept constant during outbreaks, assuming diseases spread faster than network evolution, at least during their invasion stage. Examples of incidence curves obtained by the simulations are reported in S1 Text.

Livestock disease spread is often modeled by assuming that premises are the single discrete units of the spreading processes and neglecting the possible impact of within-farm dynamics [49]. This is generally considered in the study of highly contagious and rapid infections, and corresponds to regarding a farm as being infected as soon as it receives the infection from neighboring farms following the transport of contagious animals. Under this assumption, both case studies can be analyzed in terms of networks of contacts for disease transmission. In addition, for sake of simplicity, we do not take into account the natural definition of link weights

on cattle network, representing the size of the moved batches. In S1 Text we generalize our methodology to the weighted case, including a weighted definition of loyalty, reaching results similar to the unweighted case.

We consider an emerging epidemic unfolding on a network configuration c and starting from a single node (seed s), where the rest of the population of nodes is assumed to be initially susceptible. The details on the simulations are reported in the Material and Methods section. We define $\mathcal{I}_s^c$ the set of nodes infected during the early stage invasion. In order to explore how the network topology evolution alters the spread of the disease, we consider an outbreak unfolding on the previous configuration of the system, c–1, and characterized by the same epidemiological conditions (same epidemic parameters and same initial seed s). By comparing the set of infected nodes $\mathcal{I}_s^{c-1}$ obtained in configuration c–1 to $\mathcal{I}_s^c$, we can assess changes in the two sets and how these depend on the nodes' loyalty. We define a node's infection potential $\pi_L^{c-1,c}(s)$ ($\pi_D^{c-1,c}(s)$) measuring the probability that a node will be infected in configuration c by an epidemic starting from seed s, given that it was infected in configuration c–1 under the same epidemiological conditions and provided that its loyalty status is L (D):

$$\pi_L^{c-1,c}(s) \stackrel{def}{=} \text{Prob}\left[i \in \mathcal{I}_s^c \mid i \in \mathcal{I}_s^{c-1} \text{ and } i \in \{L\}\right],$$

$$\pi_D^{c-1,c}(s) \stackrel{def}{=} \text{Prob}\left[i \in \mathcal{I}_s^c \mid i \in \mathcal{I}_s^{c-1} \text{ and } i \in \{D\}\right],$$

where i is a node of the system. π_L and π_D thus quantify the effect of the temporal stability of the network at the local level (loyalty of a node) on the stability of a macroscopic process unfolding on the network (infection). They depend on the seed chosen for the start of the epidemic, on the pair (c–1, c) of network configurations considered along its evolution, and also on the threshold value ϵ assumed for the definition of the loyalty status of the nodes.

By exploring all seeds and computing the infection potentials for different couples of years, we obtain sharply peaked probability distributions of π_L and π_D around values that are well separated along the π axis. Results are qualitatively similar in both cases under study, with peaks reached for $\pi_L/\pi_D \simeq 2.5$ in the cattle trade network and $\pi_L/\pi_D \simeq 3$ in the sexual trade network (Fig. 3A-B). An observed infection in c–1, based on the knowledge of the epidemiological conditions and no information on the network evolution, is an indicator of an infection risk for the same epidemic in c more than twice larger for loyal farms with respect to disloyal farms. Analogously, loyal sex buyers have a threefold increase in their infection potential with respect to individuals having a larger turnover of partners. Remarkably, small values of loyalty threshold ϵ are able to correctly characterize the loyal behavior of nodes with status L. Results shown in Fig. 3A-B are obtained for $\epsilon = 0.1$. Findings are however robust against changes in the choice of the threshold value, as this is induced by the peculiar bimodal shape of the probability distribution curves for the loyalty (see S1 Text). This means that intermediate values of the local stability of the nodes (i.e. $\theta > \epsilon$) imply that a possible risk of being infected is strongly stable, regardless of the dynamics of the network evolution. Valid for all possible seeds and epidemiological conditions, this result indicates that the loyalty of a node can be used as an indicator for the node's risk of infection, which has important implication for the spreading predictability in case an outbreak emerges.

These results are obtained for temporally evolving networks where no further change induced by the epidemic is assumed to occur. Focusing on the initial stage of the outbreak, we disregard the effect of interventions (e.g. social distancing, quarantine of infectious nodes, movements bans) or of adaptive behavior following awareness [37, 50, 51, 52, 53, 54]. Such assumption relies on the study's focus on the initial stage of the epidemic that may be characterized by a silent spreading phase with propagation occurring before the alert or outbreak

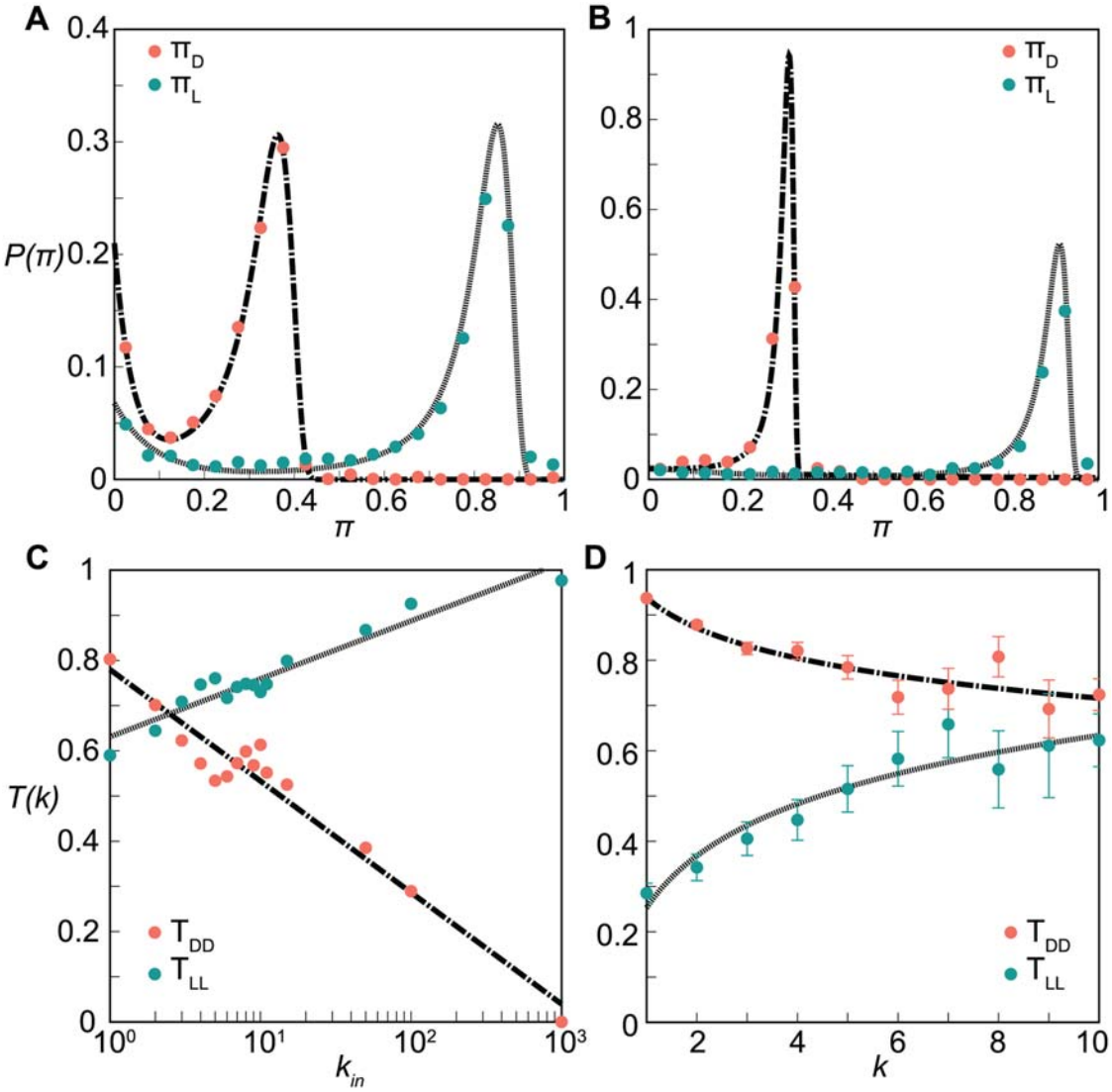

Fig 3. Infection potentials and loyalty transitions. (*A*), (*B*): Probability distributions of the infections potentials for loyal (π_L, green) and disloyal nodes (π_D, orange), for the cattle trade network and the sexual contact network, respectively. Loyalty is set with a threshold $\varepsilon = 0.1$. Dashed lines show the fit with a Landau+exponential model (see Material and Methods). (*C*), (*D*): Loyalty transition probabilities between loyal statuses ($T_{LL}(k)$, green) and disloyal statuses ($T_{DD}(k)$, orange) as functions of the degree k of the node, for the cattle trade network and the sexual contact network, respectively. Dashed lines represent the logarithmic models: $T_{DD}(k) = 0.78–0.11\log k$, and $T_{LL}(k) = 0.63+0.06\log k$ for the cattle trade network; $T_{DD}(k) = 0.94–0.10\log k$, and $T_{LL}(k) = 0.25+0.17\log k$ for the sexual contact network. Transition probabilities are computed as frequencies in the datasets under study. The error bars here represent one binomial standard deviation from these frequencies. In (*C*) the error bars are smaller than the size of the points. A single pair of configurations is considered here as example; the behavior observed is the same for all the pair of configurations.

doi:10.1371/journal.pcbi.1004152.g003

detection takes place; or, following an alert, by a contingent delay in the implementation of intervention measures.

Risk assessment analysis

The observed relationship between loyalty and infection potential can be used to define a strategy for the risk assessment analysis of an epidemic unfolding on an unknown networked system at present time, for which we have however information on its past configurations. This

may become very useful in practice even in the case of complete datasets, as for example with emerging outbreaks of livestock infectious diseases. Data on livestock movements are routinely collected following European regulations [55], however they may not be readily available in a real-time fashion upon an emergency, and a certain delay may thus be expected. Following an alert for an emerging livestock disease epidemic, knowledge of past network configurations may instead be promptly used in order to characterize the loyalty of farmers, simulate the spread of the disease on past configurations and thus provide the expected risk of infection for the farms under the ongoing outbreak. The general scheme of the strategy for the risk assessment analysis is composed of the following steps, assuming that the past network configurations $\{c-n, \ldots, c-1, c\}$ are known and that the epidemic unfolds on the unknown configuration $c+1$:

1. identify the seed s of the ongoing epidemic;
2. characterize the loyalty of the nodes from past configurations by computing $\theta_i^{c-1,c}$ from Eq. (1);
3. predict the loyalty of the nodes for the following unknown configuration $c+1$: $\theta_i^{c,c+1}$;
4. simulate the spread of the epidemic on the past configuration c under the same epidemiological conditions of the ongoing outbreak and identify the infected nodes $\mathcal{I}_s^c$;
5. compute the node epidemic risk for nodes in statuses L and D.

This strategy enables the assessment of the present infection risk (i.e. on configuration $c+1$) for all nodes hit by the simulated epidemic spreading on past configuration c ($\mathcal{I}_s^c$), not knowing their present pattern of contacts. It is based on configurations from $c-n$ to c as they are all used to build the probability distributions needed to train our approach. In the cases under study such distributions are quite stable over time so that a small set of configurations ($\{c-2, c-1, c\}$) was shown to be enough.

To make the above strategy operational, we still need to determine how we can exploit past data to predict the evolution of the loyalty of a node in future configurations (step 3) and use this information to compute nodes epidemic risk (point 5). As with all other variables characterizing the system, indeed, also θ may fluctuate from a pair of configurations ($c-1$, c) to another, as nodes may alter their loyal behavior over time, increasing or decreasing the memory of the system across time. Without any additional knowledge or prior assumption on the dynamics driving the system, we measure from available past data the probabilities of (dis)loyal nodes staying (dis)loyal across consecutive configurations, or conversely, of changing their loyalty status. This property can be quantified in terms of probabilities of transition across loyalty statuses. We thus define $T_{LL}^c(k)$ as the probability that a node with degree k being loyal between configurations $c-1$ and c will stay loyal one step after (c, $c+1$). It is important to note the explicit dependence on the degree k of the node (here defined at time c), which may increase or decrease following neighborhood reshaping (it may also assume the value $k = 0$ if the node becomes inactive in configuration c). Analogously, $T_{DD}^c(k)$ is the probability of remaining disloyal. The other two possible transition probabilities are easily obtained as $T_{LD} = 1-T_{LL}$ and $T_{DL} = 1-T_{DD}$.

Fig. 3C-D show the transition probabilities of maintaining the same loyalty status calculated on the two empirical networks for $\epsilon = 0.1$. Stability in time and non-trivial dependences on the degree of the node are found for both networks. In the cattle trade network, loyal farmers tend to remain loyal with a rather high probability ($T_{LL} > 0.6$ for all k_{in} values). In addition, this probability markedly increases with the degree, reaching $T_{LL} \simeq 1$ for the largest values of k_{in}.

Interestingly, the probability that a disloyal farmer stays disloyal the following year dramatically decreases with the degree, reaching 0 in the limit of large degree. Among the farmers who purchase cattle batches from a large number of different premises, loyal ones have an increased chance to establish business deals with the same partners the following year, whereas previously disloyal ones will more likely turn to being loyal.

A similar qualitative dependence on the degree is also found in the sexual contact network, however in this case the probability of remaining disloyal is always very high ($T_{DD} > 0.7$) even for high degrees. T_{LL} shows a relatively more pronounced dependence on k, ranging from 0.3 (low degree nodes) to 0.6 (high degree nodes). Differently from the farmers behavior, sex buyers display a large tendency to keep a high rate of partners turnover across time. Moreover, the largest probability of preserving sexual partners is obtained when the number of partners is rather large.

Remarkably, in both networks, transition probabilities are found to be stable across time and are well described by logarithmic functions (with parameters depending on the system and on ϵ) that can be used to predict the loyalty of nodes in configuration c+1 from past data (Fig. 3C-D). With this information, it is then possible to compute the epidemic risk of a node i in configuration c+1, having degree $k = k_i^c$ in configuration c and known loyalty status $\{L, D\}$ between configurations c−1 and c as follows:

$$\begin{cases} \text{if loyalty class} = D: \quad \rho_i^{c+1} = \pi_D^{c,c+1}(s)T_{DD}(k) + \pi_L^{c,c+1}(s)T_{DL}(k)\,; \\ \text{if loyalty class} = L: \quad \rho_i^{c+1} = \pi_D^{c,c+1}(s)T_{LD}(k) + \pi_L^{c,c+1}(s)T_{LL}(k)\,. \end{cases} \tag{2}$$

It is important to note that in our framework the epidemic risk is a node property, and not a global characteristic of a specific disease.

Validation

To validate our strategy of risk assessment, we test our predictions based on past data for the risk of being infected in configuration c+1 on the results of an epidemic simulation explicitly performed on the supposedly unknown configuration c+1. We consider the set of nodes $\mathcal{I}_s^c$ for which we are able to provide risk predictions and divide it into two subsets, according to their predicted risk of infection ρ_i^{c+1}. We indicate with $\mathcal{I}_{s,h}^c$ the top 25% highest ranking nodes, and with $\mathcal{I}_{s,l}^c$ all the remaining others. We then compute the fraction P_h of nodes in the subset $\mathcal{I}_{s,h}^c$, i.e. predicted at high risk, that belong to the set of infected nodes $\mathcal{I}_s^{c+1}$ in the simulated epidemic aimed at validation. Analogously, P_l measures the fraction of nodes in $\mathcal{I}_{s,l}^c$ that are reached by the infection in the simulation on c+1. In other words, P_h (P_l) represents the probability for a node having a high (low) risk of infection to indeed get infected. The accuracy of the risk assessment analysis can thus be measured in terms of the relative risk ratio $\nu = P_h/P_l$, where values $\nu \leq 1$ indicate negative or no correlation between our risk predictions and the observed infections, whereas values $\nu > 1$ indicate that the prediction is informative. For both networks we find a significant correlation, signaled by the distributions of the relative risk ratio ν peaking around values $\nu > 1$ (Fig. 4A-B). The peak positions ($\nu \simeq 1.4$ and $\nu \simeq 1.7$ for cattle and sex, respectively) are remarkably close to the benchmark values represented by the distributions computed on the training sets (red lines in Fig. 4A-B). In addition, the comparison with the distributions from a null model obtained by reshuffling the infection statuses of nodes (dotted curves peaking around $\nu = 1$ in Fig. 4A-B) further confirms the accuracy of the approach. Findings are robust against changes of the value used to define $\mathcal{I}_{s,h}^c$ or against alternative definitions of this quantity (see S1 Text).

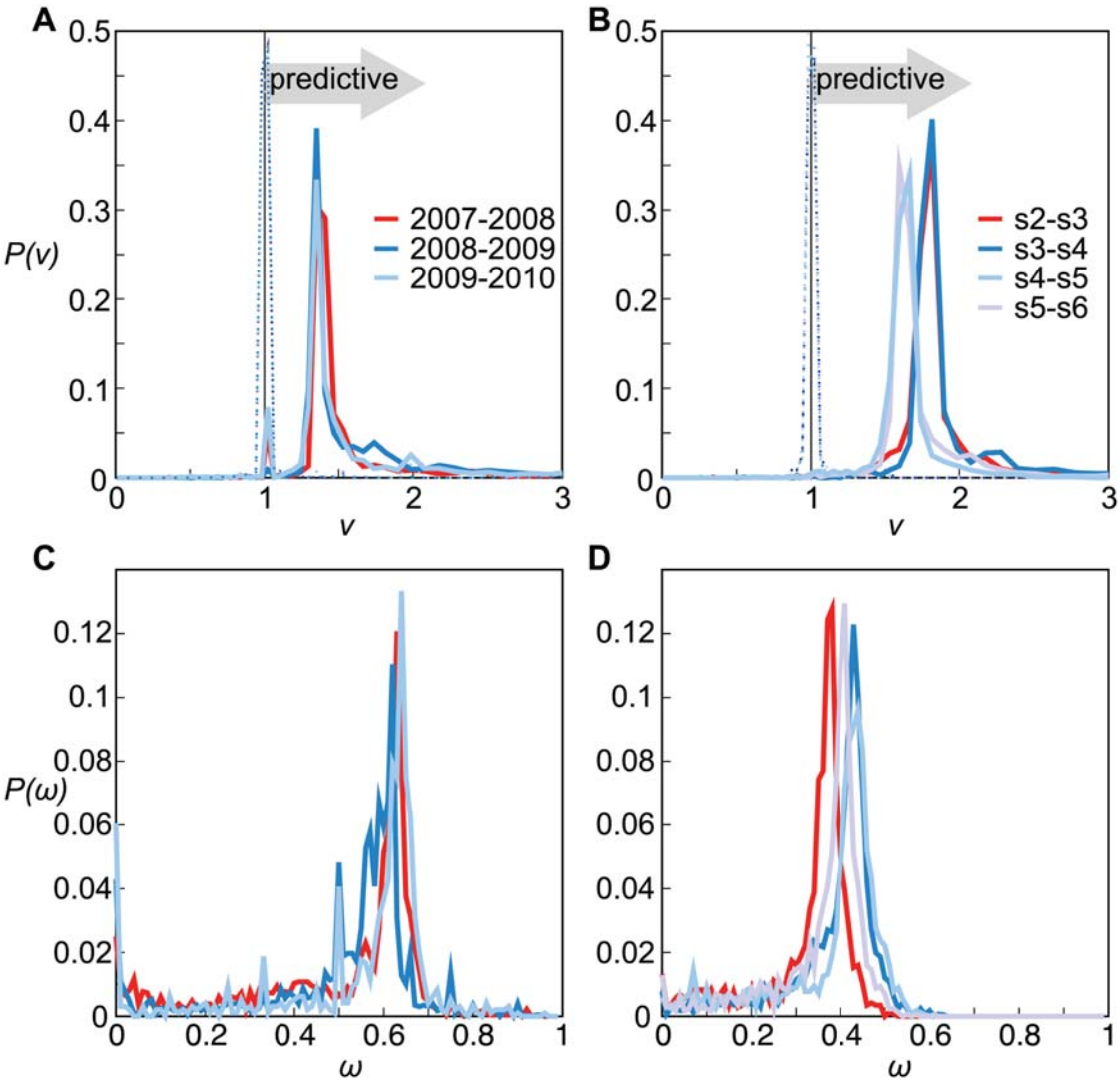

Fig 4. Validation of the risk assessment analysis. (*A*), (*B*): Probability distributions of the risk ratio *v* for the cattle trade network and the sexual contact network, respectively. Red lines are computed on training sets (2007–08 for cattle and s2-s3 for sexual contacts). The dashed lines peaking around 1 represent a null model based on reshuffling the infection statuses, i.e. randomly permuting the attribute *"actually being infected"* among the nodes for which risk assessment is performed. (*C*), (*D*): Probability distributions of the predictive power ω for the cattle trade network and the sexual contact network, respectively.

doi:10.1371/journal.pcbi.1004152.g004

One other important aspect to characterize is the predictive power of our risk assessment analysis. Our predictions indeed are limited to the set $\mathcal{I}_s^c$ of nodes that are reached in the simulation performed on past data, proxy for the future outbreak. If a node is not infected by the simulation unfolding on configuration *c* or it is not active at that given time, our strategy is unable to provide a risk assessment for that node in the future. We can then quantify the predictive power ω as the fraction of infected nodes for which we could provide the epidemic risk, i.e. $\omega_s^{c,c+1} = \mid \mathcal{I}_s^{c+1} \cap \mathcal{I}_s^c \mid / \mid \mathcal{I}_s^{c+1} \mid$. High values of ω indicate that few infections are missed by the risk assessment analysis. Fig. 4C-D display the distributions $P(\omega)$ obtained for the two case studies, showing that a higher predictive power is obtained in the cattle trade network (peak at $\omega \simeq 60\%$) with respect to the sexual contact network (peak at $\omega \simeq 40\%$). Our methodology can

potentially be applied to a wide range of networks, other than the ones presented here, as shown with the example of human face-to-face proximity networks relevant for the spread of respiratory diseases reported in S1 Text.

We also tested whether our risk measure represents a significant improvement in prediction accuracy with respect to simpler and more immediate centrality measures (namely, the degree). Through a multivariate logistic regression, in S1 Text we show that our definition of node risk is predictor of infection even after adjusting for node degree.

Memory driven dynamical model

The results of the risk assessment analysis obtained from the application of our strategy to the two empirical networks show qualitatively similar results, indicating that the approach is general enough to provide valuable information on the risk of infection in different settings. The observed differences in the predictive power of the approach are expected to be induced by the different temporal behavior of the two systems, resulting in a different amount of memory in preserving links (Fig. 1) and different loyalty of nodes and their time-variations (Fig. 2 and 3C-D).

In order to systematically explore the role of these temporal features on the accuracy and predictive power of our approach, we introduce a generic model for the generation of synthetic temporal networks. The model is based on a set of parameters that can be tuned to reproduce the empirically observed features of the two networks, i.e.: *(i)* the topological heterogeneity of each configuration of the network described by a stable probability distribution (Fig. 1A-B); *(ii)* a vital dynamics to allow for the appearance and disappearance of nodes; *(iii)* a tunable amount of memory characterizing the time evolution of the network contacts (Fig. 1C-D). These specific properties differentiate our approach from the previously introduced models that display instantaneous homogeneous properties for network configurations [56, 57, 58, 59], reproduce bursty inter-event time distributions but without the explicit introduction of memory [33, 60, 61] or of its control [58].

Based on an iterative network generation approach (see Materials and Methods), we can build an arbitrarily large number of configurations of networks with 10^4 nodes. They are characterized by stable in-degree and out-degree heterogeneous distribution across time (Fig. 5A where high memory and low memory regimes are displayed) and by profiles for the probability distribution of the loyalty as in the empirical networks (Fig. 5B). The number of nodes with zero loyalty can be computed analytically (see Materials and Methods) and it is confirmed by numerical findings (see S1 Text). A high memory regime corresponds to having nodes in the system that display a highly loyal behavior (e.g., $\theta > 0.7$), whereas values in the range $\theta \in [0.7, 1)$ are almost absent in a low memory regime, in agreement with the findings of Fig. 2.

Applying the introduced risk assessment analysis to the synthetically generated temporal network, we recover a significant accuracy for both memory regimes (Fig. 5C). Different degrees of memory are however responsible for the fraction of the system for which a risk assessment can be made. In networks characterized by higher memory, the distribution of the predictive power ω has a well defined peak, whereas for lower memory it is roughly uniform in the range $\omega \in [0, 0.4]$ (Fig. 5D). Such a regime implies that not enough structure is maintained in the system to control more than 40% of the future infections. Our risk assessment analysis allows therefore accurate predictions across varying memory regimes characterizing the temporal networks, but the degree of memory impacts the amount of predictions that can be made. The model also shows that the analysis is not affected by the choice of the aggregating time window used to define the network configurations [61, 62, 63], as long as the heterogeneous topological features at the system level and the heterogeneous memory at the node level

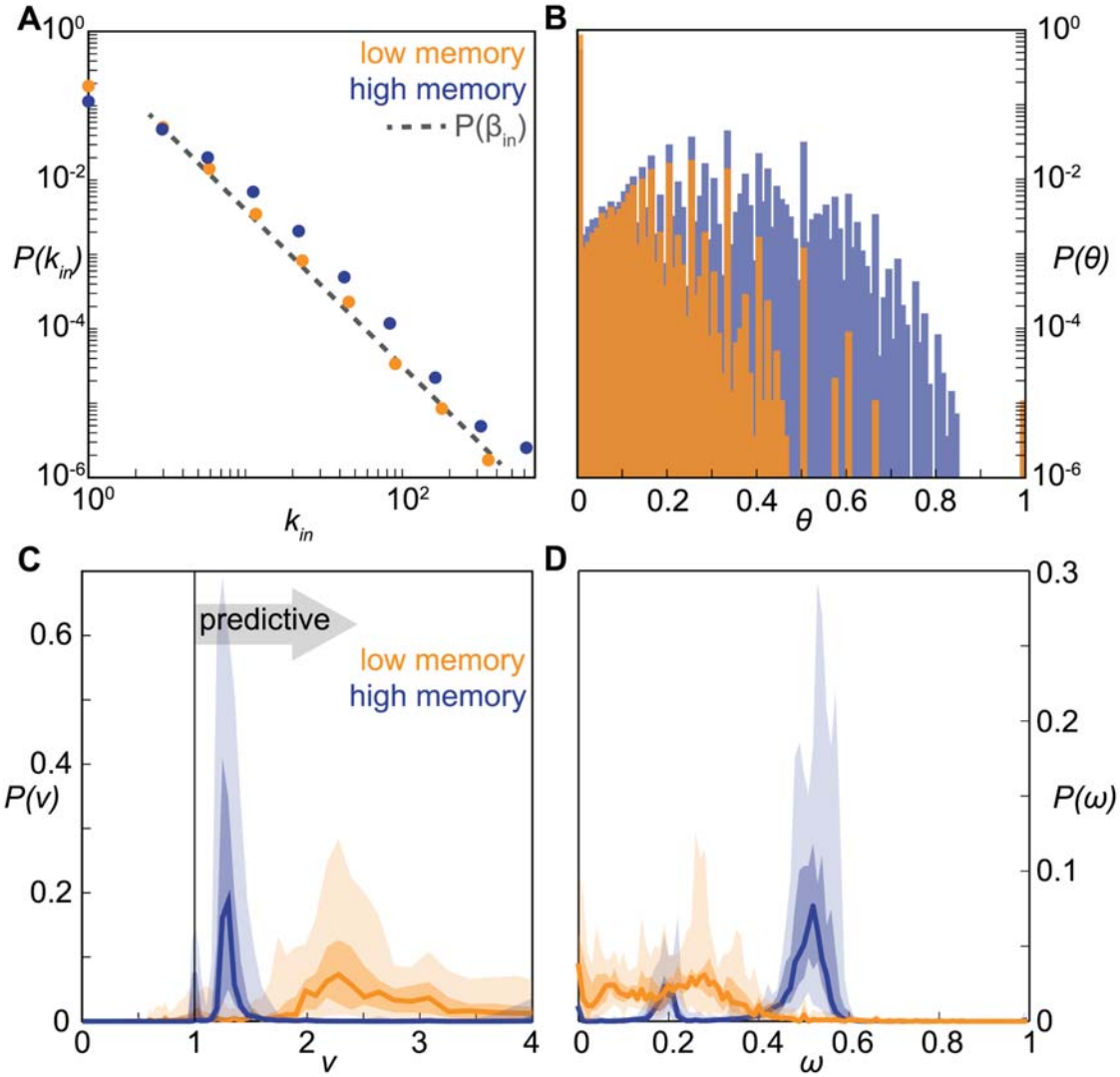

Fig 5. Memory driven dynamical model: model properties and validation of the risk assessment analysis. (*A*): Probability distributions of the node in-degree, in the low ($p_\alpha = 0.3$) and high memory ($p_\alpha = 0.7$) regimes. The slope of the distributions does not depend on p_α, and it is forced by the exponent γ of the β_{in} distribution (dashed line). (*B*): Probability distributions of the loyalty, in the low and high memory regimes. Distributions are color-coded as in panel (*a*). (*C*): Probability distributions of the risk ratio v, in the low and high memory regimes. Lines represent the median values obtained from 50 realizations of the model; darker and lighter shaded areas represent the 50% and 95% confidence intervals. (*D*): Probability distributions of the predictive power ω, in the low and high memory regimes. Medians and confidence intervals are presented as in panel (*C*). Distributions are color-coded as in panel (*A*).

doi:10.1371/journal.pcbi.1004152.g005

are kept across aggregation, as observed for the empirical networks under study (see [19] and S1 Text).

Conclusions

We introduce a simple measure to characterize the amount of memory in the time evolution of a networked system. The measure is local and it is empirically motivated from two case studies relevant for disease transmission. By focusing on the degree of loyalty that each node has in establishing connections with the same partners as time evolves, we are able to connect an egocentric view of the system (the node's strategy in establishing its neighborhood over time) to

the system's larger-scale properties characterizing the early propagation of an emerging epidemic.

We uncover a non-trivial correlation between the loyalty of a node and its risk of being infected if an epidemic occurs, given fixed epidemiological conditions, and use this to inform a risk assessment analysis applicable to different settings with no information on the network evolution dynamics. A theoretical model generating synthetic time-varying networks allows us to frame the analysis in a more general perspective and disentangle the role of different features. The accuracy of the proposed risk assessment analysis is stable across variations of the temporal correlations of the system, whereas its predictive power depends on the degree of memory kept in the time evolution. The introduced strategy can be used to inform preventive actions in preparation to an epidemic and for targeted control responses during an outbreak emergency, only relying on past network data.

Methods

Datasets

The cattle trade network is obtained from the database of the Italian national bovine registry recording all cattle displacements due to trade transactions. We consider animal movements during a 5 years time period, from 2006 to 2010, involving 215,264 premises and 2,973,710 directed links. Nodes may be active or inactive depending whether farms sell/buy cattle in a given timeframe. The cattle network is available as S1 Dataset. From the dataset we have removed slaughterhouses ($\sim$ 1% of the nodes) as they are not relevant for transmission.

The sexual contact network is extracted from an online Brazilian forum where male sex buyers rate and comment on their sexual encounters with female sex sellers [16]. Time-stamped posts are used as proxies for sexual intercourse and multiple entries are considered separately, following previous works [16, 31]. A total of 13,855 individuals establishing 34,509 distinct sexual contacts are considered in the study, after discarding the initial transient of the community growth [16]. Nodes may be active or inactive depending whether individuals use or not the service, and join or quit the community. Six-months aggregating snapshots are chosen. A different aggregating time window of three months has been tested, obtaining similar results (see S1 Text).

Risk of infection

The distributions of the risk potentials π_L and π_D reported in Fig. 3 are modeled with a sum of Landau distribution and an exponential suppression. This family of functions depending on four parameters (see S1 Text for the specific functional form) was chosen as it well reproduces the distribution profiles of the risk potentials, and it was used to compute the nodes' epidemic risk. A goodness of fit was not performed, as this choice was automatically validated in the validation analysis performed on the whole prediction approach.

Memory driven model

The basic iterative network generation approach allows to build configuration c+1 from configuration c through the following steps:

- vital dynamics: nodes that are inactive in configuration c become active in c+1 with probability b, while active nodes become inactive with probability d;
- memory: active nodes maintain same in-neighbors each with probability p_α; then they form β_{in} new in-stubs, where β_{in} is extracted from a power-law distribution: $P(\beta_{in}) \sim \beta_{in}^{-\gamma}$;

- out-degree heterogeneity: each node is assigned β_{out} out-stubs, where β_{out} is drawn from another power-law distribution: $P(\beta_{out}) \sim \beta_{out}^{-\delta}$. Then each of the in-stubs is randomly matched to an out-stub.

The total set of nodes is considered to be fixed in time, and nodes may be active (i.e. establishing connections) or inactive (i.e. isolated) in a given configuration. All five parameters b, d, γ, p_α, δ are assumed constant in time and throughout the network. The amount of memory in the system is tuned by the interplay of the two parameters p_α and d. Starting from an arbitrarily chosen initial configuration $c = 0$, simulations show that the system rapidly evolves towards a dynamical equilibrium, and successive configurations can be obtained after discarding an initial transient of time. The parameters values used in the paper are: $N = 10^4$; $b = 0.7$; $d = 0.2$; $\gamma = 2.25$; $\delta = 2.75$; $p_\alpha = 0.3, 0.7$. The influence of such parameters on the network properties is examined in S1 Text.

If we denote with α the number of neighbors that a given node keeps across two consecutive configurations (c–1, c), we can express the loyalty simply as:

$$\theta_i^{c-1,c} = \frac{\alpha^{c-1,c}}{\left(k_i^c + \beta_{in}^c\right)} \tag{3}$$

where the superscript c for α, β_{in} indicate the values used to build configuration c. The number of nodes with $\theta = 0$ as a function of the degree can be computed analytically: $P(\theta_{c,c+1} = 0) = d + (1-d)(1-p_\alpha)^{k_c}$. Similarly, it is possible to compute the probability $f_{c,\,c+1}$ that a link present in configuration c is also present in configuration c+1. In the S1 Text we show that $f_{c,\,c+1} \simeq (1-d)p_\alpha$ and confirm this result by numerical simulations.

Supporting Information

S1 Dataset. Cattle trade network dataset. We provide the cattle trade network as yearly edge lists, from 2006 to 2010. The dataset consists in five CSV files (one for each year) compressed in a ZIP archive.
(ZIP)

S1 Text. Additional analyses. We provide a description of the seasonal pattern of cattle network (Section 1), a more in-depth characterization of loyalty (Section 2), a comparison between loyalty and other similarity measures (Section 3), the specific modeling function for the infection potential (Section 4), the robustness of the risk assessment procedure to variations in parameters and assumptions (Section 5), further analyses of the memory driven model in terms of analytical results (Section 6) and additional properties (Section 7), an extension of our methodology to take into account transmissibility lower than 1 (Section 8), and links weights (Section 9).
(PDF)

Author Contributions

Conceived and designed the experiments: EV CP VC. Performed the experiments: EV. Analyzed the data: EV CP VC. Contributed reagents/materials/analysis tools: EV CP AG DP LS VC. Wrote the paper: EV CP AG DP LS VC.

References

1. Pastor-Satorras R, Vespignani A (2001) Epidemic spreading in scale-free networks. Phys Rev Lett 86:3200–3203. doi: 10.1103/PhysRevLett.86.3200 PMID: 11290142
2. Lloyd A, May R (2001) Epidemiology: how viruses spread among computers and people. Science 292:1316–1317. doi: 10.1126/science.1061076 PMID: 11360990
3. Newman MEJ (2002) Spread of epidemic disease on networks. Phys Rev E 66, 016128. doi: 10.1103/PhysRevE.66.016128
4. Cohen R, Havlin S, ben-Avraham D (2003) Efficient immunization strategies for computer networks and populations. Phys Rev Lett 91:247901. doi: 10.1103/PhysRevLett.91.247901 PMID: 14683159
5. Keeling MJ, Eames KTD (2005) Networks and epidemic models. J R Soc Interface 2: 295–307. doi: 10.1098/rsif.2005.0051 PMID: 16849187
6. Riley S (2007) Large-scale spatial-transmission models of infectious disease. Science 316: 1298–1301. doi: 10.1126/science.1134695 PMID: 17540894
7. Colizza V, Barrat A, Barthélémy M, Vespignani A (2006) The role of the airline transportation network in the prediction and predictability of global epidemics. Proc Natl Acad Sci USA 103:2015–2020. doi: 10.1073/pnas.0510525103 PMID: 16461461
8. Brockmann D, Helbing D (2013) The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. Science 342, 1337. doi: 10.1126/science.1245200 PMID: 24337289
9. Riley S et al (2003) Transmission Dynamics of the Etiological Agent of SARS in Hong Kong: Impact of Public Health Interventions. Science 300: 1961–1966. doi: 10.1126/science.1086478 PMID: 12766206
10. Eckmann JP, Moses E, Sergi D (2004) Entropy of dialogues creates coherent structures in e-mail traffic. Proc Natl Acad Sci USA 101:14333–14337. doi: 10.1073/pnas.0405728101 PMID: 15448210
11. Onnela JP, Saramäki J, Hyvönen J, Szabó G, Lazer D, et al. (2007) Structure and tie strengths in mobile communication networks. Proc Natl Acad Sci USA 104(18):7332–7336. doi: 10.1073/pnas.0610245104 PMID: 17456605
12. Rybski D, Buldyrev SV, Havlin S, Liljeros F, Makse HA (2009) Scaling laws of human interaction activity. Proc Natl Acad Sci USA 106(31):12640–12645. doi: 10.1073/pnas.0902667106 PMID: 19617555
13. Gautreau A, Barrat A, Barthélémy M (2009) Microdynamics in stationary complex networks. Proc Natl Acad Sci (USA) 106:8847–8852. doi: 10.1073/pnas.0811113106
14. Vernon MC, Keeling MJ (2009) Representing the UK's cattle herd as static and dynamic networks. Proc R Soc B 276:469–476 doi: 10.1098/rspb.2008.1009 PMID: 18854300
15. Cattuto C, Van den Broek W, Barrat A, Colizza V, Pinton JF, et al. (2010) Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks. PLOS ONE 5(7):e11596. doi: 10.1371/journal.pone.0011596 PMID: 20657651
16. Rocha LEC, Liljeros F, and Holme P (2010) Information dynamics shape the sexual networks of Internet-mediated prostitution. Proc Natl Acad Sci USA 107:5706–5711. doi: 10.1073/pnas.0914080107 PMID: 20231480
17. Salathé M, Kazandjieva M, Lee JW, Levis P, Feldman MW, et al. (2010) A high-resolution human contact network for infectious disease transmission. Proc Natl Acad Sci USA 107:22020–22025. doi: 10.1073/pnas.1009094108 PMID: 21149721
18. Tang J, Scellato S, Musolesi M, Mascolo C, Latora V (2010) Small-world behavior in time-varying graphs. Phys Rev E 81:055101. doi: 10.1103/PhysRevE.81.055101
19. Bajardi P, Barrat A, Natale F, Savini L, Colizza V (2011) Dynamical patterns of cattle trade movements. PLOS ONE, 6(5):e19869 doi: 10.1371/journal.pone.0019869 PMID: 21625633
20. Miritello G, Moro E, Lara R (2011) Dynamical strength of social ties in information spreading. Phys Rev E 83:045102. doi: 10.1103/PhysRevE.83.045102
21. Karsai M, Kaski K, Barabási AL, Kertész J (2012) Universal features of correlated bursty behaviour. Sci Rep 2, 397. doi: 10.1038/srep00397 PMID: 22563526
22. Holme P, Saramaki J (2012) Temporal networks. Phys Rep 519:97. doi: 10.1016/j.physrep.2012.03.001
23. Albert R, Jeong H, Barabási AL (2000) Error and attack tolerance of complex networks. Nature 406:378–482. doi: 10.1038/35019019 PMID: 10935628
24. Cohen R, Erez K, ben-Avraham D, Havlin S (2001) Breakdown of the Internet under intentional attack. Phys Rev Lett 86:3682–3685. doi: 10.1103/PhysRevLett.86.3682 PMID: 11328053
25. Freeman LC (1979) Centrality in social networks: Conceptual clarification. Social Networks 1:215–239. doi: 10.1016/0378-8733(78)90021-7

26. Friedkin NE (1991) Theoretical foundations for centrality measures. Am J Sociology 96:1478–1504. doi: 10.1086/229694

27. Holme P (2004) Efficient local strategies for vaccination and network attack. Europhys Lett 68:908–914. doi: 10.1209/epl/i2004-10286-2

28. Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, et al. (2010) Identification of inuential spreaders in complex networks. Nature Phys 6:888–893. doi: 10.1038/nphys1746

29. Salathé M, Jones JH (2010) Dynamics and control of diseases in networks with community structure. PLOS Comput Biol 6:e1000736. doi: 10.1371/journal.pcbi.1000736 PMID: 20386735

30. Natale F, Savini L, Giovannini A, Calistri P, Candeloro L, et al. (2011) Evaluation of risk and vulnerability using a Disease Flow Centrality measure in dynamic cattle trade networks. Prev Vet Med 98(2–3):111–8. doi: 10.1016/j.prevetmed.2010.11.013 PMID: 21159393

31. Rocha LEC, Liljeros F, Holme P (2011) Simulated Epidemics in an Empirical Spatiotemporal Network of 50,185 Sexual Contacts. PLOS Computational Biology 7(3) (2011) doi: 10.1371/journal.pcbi.1001109 PMID: 21445228

32. Galvani AP, May RM (2005) Dimensions of superpsreading. Nature 438, 293–294. doi: 10.1038/438293a PMID: 16292292

33. Lee S, Rocha LEC, Liljeros F, Holme P (2012) Exploiting temporal networks structures of human interaction to effectively immunize populations. PLOS ONE 7(5):e36439. doi: 10.1371/journal.pone.0036439 PMID: 22586472

34. Starnini M, Machens A, Cattuto C, Barrat A, Pastor-Satorras R (2013) Immunization strategies for epidemic processes in time-varying contact networks. J Theor Biol 337: 89–100. doi: 10.1016/j.jtbi.2013.07.004 PMID: 23871715

35. Bajardi P, Barrat A, Savini L, Colizza V (2012) Optimizing surveillance for livestock disease spreading through animal movements. J Roy Soc Int (June, 2012). doi: 10.1098/rsif.2012.0289

36. Natale F, Giovannini A, Savini L, Palma D, Possenti L, et al. (2009) Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread. Prev Vet Med 92:341–350. doi: 10.1016/j.prevetmed.2009.08.026 PMID: 19775765

37. Robinson SE, Everett MG, Christley RM (2007) Recent network evolution increases the potential for large epidemics in the British cattle population. J Roy Soc Int 4: 669–674 doi: 10.1098/rsif.2007.0214

38. Kao RR, Danon L, Green DM, Kiss IZ (2006) Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain. Proc Roy Soc Lond B Biol Sci 273: 1999–2007 doi: 10.1098/rspb.2006.3505

39. Sharp B, Sharp A (1997) Loyalty Programs and Their on Repeat-Purchase Loyalty Patterns. International Journal of Research in Marketing 14(5):473–86. doi: 10.1016/S0167-8116(97)00022-0

40. Podolny JM (1994) Market uncertainty and the social character of economic exchange. Administrative Science Quarterly 39:458–483. doi: 10.2307/2393299

41. Sorenson O, Waguespack DM (2006) Structure and Exchange: Self-Confirming Dynamics in Hollywood. Administrative Science Quarterly 51:560–589.

42. Miritello G, Lara R, Cebrian M, Moro E (2013) Limited communication capacity unveils strategies for human interaction. Sci Rep 3, 1950. doi: 10.1038/srep01950 PMID: 23739519

43. Morris M, Kretzschmar M (1997) Concurrent partnerships and the spread of HIV. AIDS 11:641–648. doi: 10.1097/00002030-199705000-00012 PMID: 9108946

44. Anderson RM, May RM, Boily MC, Garnett GP, Rowley JT (1991) The spread of HIV-1 in Africa: sexual contact patterns and the predicted demographic impact of AIDS. Nature 352:581–589. doi: 10.1038/352581a0 PMID: 1865922

45. Keeling MJ, Woolhouse MEJ, Shaw DJ, Matthews L, Chase-Topping M, et al. (2001) Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape. Science 294:813–817 doi: 10.1126/science.1065973 PMID: 11679661

46. Saegerman C, Berkvens D, Mellor PS (2008) Bluetongue Epidemiology in the European Union. Emerging Inf Dis 14:539–544 doi: 10.3201/eid1404.071441

47. Tinsley M, Lewis FI, Brülisauer F (2012) Network modeling of BVD transmission Veterinary Research 43:11

48. Anderson RM, May RM (1992) Infectious Diseases of Humans Dynamics and Control. Oxford University Press, Oxford.

49. Keeling MJ (2005) Models of foot-and-mouth disease. Proc Roy Soc B 272:1195–1202. doi: 10.1098/rspb.2004.3046

50. Gross T, Dommar D'Lima CJ, Blasius B (2006) Epidemic dynamics on adaptive networks. Phys Rev Lett 96:208701. doi: 10.1103/PhysRevLett.96.208701 PMID: 16803215

51. Funk S, Salathé M, Jensen VAA (2010) Modeling the inuence of human behavior on the spread of infectious diseases: a review. J R Soc Interface 7:1247–1256. doi: 10.1098/rsif.2010.0142 PMID: 20504800

52. Shaw LB, Schwartz IB (2010) Enhanced vaccine control of epidemics in adaptive networks. Phys Rev E 81:046120. doi: 10.1103/PhysRevE.81.046120

53. Meloni S, Perra N, Arenas A, Gomez S, Moreno Y, et al. (2011) Modeling human mobility responses to the large-scale spreading of infectious diseases. Sci Rep 1:62. doi: 10.1038/srep00062 PMID: 22355581

54. Bajardi P, Poletto C, Ramasco JJ, Tizzoni M, Colizza V, et al. (2011) Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. PLOS ONE 6(1), e16591. doi: 10.1371/journal.pone.0016591 PMID: 21304943

55. European Parliament and European council (2000) 204: 1–10. Regulation (EC) No. 1760/2000 of 17 July 2000 establishing a system for the identification and registration of bovine animals and regarding labeling of beef and beef products and repealing Council Regulation (EC) No. 820/97 European Council. Off. J. Eur.Communities L.

56. Stehlé J, Barrat A, Bianconi G (2010) Dynamical and bursty interactions in social networks. Phys. Rev. E 81:035101. doi: 10.1103/PhysRevE.81.035101

57. Perra N, Gonçalves B, Pastor-Satorras R, Vespignani A (2012) Activity driven modeling of time varying networks. Sci Rep 2:469. doi: 10.1038/srep00469 PMID: 22741058

58. Starnini M, Baronchelli A, Pastor-Satorras R (2013) Modeling Human Dynamics of Face-to-Face Interaction Networks. Phys Rev Lett 110:168701 doi: 10.1103/PhysRevLett.110.168701 PMID: 23679648

59. Kársai M, Perra N, Vespignani A (2014) Time varying networks and the weakness of strong ties. Sci Rep 4:4001. doi: 10.1038/srep04001 PMID: 24510159

60. Rocha LEC, Blondel VD (2013) Bursts of Vertex Activation and Epidemics in Evolving Networks. PLoS Comput Biol 9(3): e1002974. doi: 10.1371/journal.pcbi.1002974 PMID: 23555211

61. Holme P (2013) Epidemiologically optimal static networks from temporal network data. PLOS Comp Bio. doi: 10.1371/journal.pcbi.1003142

62. Hoffmann T, Porter MA, Lambiotte R (2012) Generalized master equations for non-Poisson dynamics on networks. Phys Rev E 86:046102. doi: 10.1103/PhysRevE.86.046102

63. Ribeiro B, Perra N, Baronchelli A (2013) Quantifying the effect of temporal resolution on time-varying networks. Sci Rep 3:3006. doi: 10.1038/srep03006 PMID: 24141695