



Complex delay dynamics on railway networks from universal laws to realistic modelling

Bernardo Monechi^{1,2*} , Pietro Gravino^{2,3} , Riccardo Di Clemente^{4,5}  and Vito D.P. Servedio⁶ 

*Correspondence:

mone.berna@gmail.com

¹Institute for Scientific Interchange Foundation, Turin, Italy

²SONY Computer Science Lab, Paris, France

Full list of author information is available at the end of the article

Abstract

Railways are a key infrastructure for any modern country. The reliability and resilience of this peculiar transportation system may be challenged by different shocks such as disruptions, strikes and adverse weather conditions. These events compromise the correct functioning of the system and trigger the spreading of delays into the railway network on a daily basis. Despite their importance, a general theoretical understanding of the underlying causes of these disruptions is still lacking. In this work, we analyse the Italian and German railway networks by leveraging on the train schedules and actual delay data retrieved during the year 2015. We use these data to infer simple statistical laws ruling the emergence of localized delays in different areas of the network and we model the spreading of these delays throughout the network by exploiting a framework inspired by epidemic spreading models. Our model offers a fast and easy tool for the preliminary assessment of the effectiveness of traffic handling policies, and of the railway network criticalities.

Keywords: Complex systems; Networks; Delay dynamics; Modelling; Spreading

1 Introduction

Transportation networks are a critical infrastructure with an enormous impact on local, national, and international economies [1, 2]. At the micro-level, the commuting time impacts the economy and the shape of cities [3], directly influencing our life style and choices [4]. At the macro-level the travel time regulates trade and stimulates economic activities [5]. Railroads, when correctly developed, foster interregional trades within the country, and increase income levels within the city boundaries [6]. Whether we are traveling in crowded coaches of the commuters rail during rush hours or in the long-medium high speed train trips, the reliability of the transportation system is a key factor in determining travel behavior [7]. The travel time unreliability in rail service can have substantial consequences for its users [8] and the growth of cities [9]. Delays are one of the causes that corrupt the travel time reliability of the transportation systems. To address delay propagation, traditional approaches apply either stochastic models [10] or use propagation algorithms on a timed event graph representation of a scheduled railway system [11]. Complex equations describing delay were obtained and solved by means of iterative refinement al-

gorithms to predict positive delays in urban trains [12], while traffic control models were proposed to manage the safety of timetables after perturbations occur [13].

The development of models addressing the dynamics of trains over railway networks are usually focused either on delimited aspects, e.g., single lane dynamics [14] or the issuing of schedule [15, 16], or are comprehensively including a wide variety of different microscopic ingredients that can be hardly validated by real data [17, 18]. In the last years a new perspective bloomed to provide a different understanding of delay dynamics through the lens of Complex Systems. Railway networks have largely been studied and exhibit small-world scaling properties [19], their topology have geo-spatial restriction [20] and structural redundancy [21]. On the other end, a railway delay dynamics framework within a complex systems approach is still lacking. Our aim is to develop a model capable of reproducing the actual delay structure and the emergence of congested areas, by exploiting some essential ingredients without relying on the detailed knowledge of the microscopic mechanisms underlying delay generation. A similar approach has already been successfully applied in the context of the Air Transport System of USA and Europe [22, 23], where interaction models were used to characterize and forecast the spreading of delays among flights. These modelling schemes are driven by the necessity of simple and novel frameworks to study transportation systems, allowing fast scenario simulations for schedule testing, which could be easily applied to resilience studies that are nowadays performed independently of the dynamics taking place over the network [21].

Following a similar approach, we develop a new data driven model of delay propagation among trains diffusing over a railway network. We start by inferring universal laws governing the emergence of delays over the railway network as a consequence of the occurrence of adverse conditions that are not depending on the (possible) interactions between trains. Hence, inspired by models for epidemic spreading [24], we introduce the *Delay Propagation Model* as a novel framework to asses and test how such emerging delays spawn and spread over the network. We name the first kind of delays as “exogenous delays”, while the second ones stemming from the interaction between trains as “endogenous delays”. While the occurrence of exogenous delays can be detected and analysed from our data, the endogenous delays, on the contrary, since the (supposed) direct interaction between trains cannot be inferred from the dataset we collected, will be modeled with a one parameter interaction. The effects and consistency of such endogenous delay modelling scheme will be checked *a posteriori* by means of numerical simulations. Our model is close to an SIS model of epidemic spreading [24], since trains can either get infected by delay, recover from it and then get infected again. We show that this model is capable of reproducing the empirical distribution of delays measured in the data as well as the emergence of large congested areas. Moreover, we show that the removal of the delay propagation mechanism prevents the modeled system from generating large disruptions, hence strongly suggesting the existence of this kind of interaction in the real system. Finally, we propose an application of our model in studying a scenario where the propagation of a localized delay leads to the emergence of a vast non-functioning area in the Italian Railway Network.

The paper is structured as follows: in the first section we describe the dataset and some empirical findings that inspired the assumptions used in the development of the model; in the second section we discuss the model in detail and will show its capability of reproducing empirical findings such as the distribution of delays and the size of congested areas;

we will also demonstrate the usefulness of the model as a scenario simulation tool; in the last section we will summarize our findings and discuss possible future developments.

2 Methods

2.1 Data

In order to extract relevant patterns related with the emergence of train delays, we collected and analysed data about the daily operations of the Italian and German Railway Systems during the year 2015. Such data were collected in the first case through the ViaggiaTreno website [25] and for the second one through the OpenDataCity website [26]. The information we acquired allowed for a complete reconstruction of the schedules of the trains, the structure of the Railway Network and the delays that affected the trains during their movements. For more information about the data collection please refer to Additional files 1–3. The choice of the Italian and German Systems as subjects for our studies lies in their similarities in structure and management. In fact, their networks have remarkable and comparable sizes (41,315 and 16,723 km, respectively) and densities (8.22 and 12.46 km² per km of tracks, respectively [27]). Moreover, these two countries share also a crucial characteristic: in both railway systems traffic is handled mainly by a single national company. In other countries with similar networks, like in the United Kingdom, the railway network is managed by different companies resulting in a more complex system where trains are additionally subject to the commercial policies of different operators. Figure 1 shows a set of topological analyses of the two railway networks, together with a preliminary analysis of the traffic load and delays in the systems. Following existing literature, we refer to a railway network as the network whose nodes represent stations that are connected by a link whenever there is a train connecting them with two consecutive stops in its schedule. The network is directed since it is possible that a connection between two stations is travelled just in one direction. In this paper we refer to the nodes in a railway network either as “nodes” or as “stations”. Moreover, the action performed by a train travelling from station A to station B will be referred as “travelling over the link” connecting A and B. In both networks the distribution of the degree k is peaked on $k = 4$ and for larger degrees has exponential-like decrease [19, 28] proportional to e^{-k/k_0} with $k_0 \simeq 4.5 \pm 0.1$. This finding is in agreement with other geographical networks [29]. The network assortativity (Fig. 1A) [30] is for Italy and Germany, 0.18 and 0.24 respectively. This indicates that while there is a slight preference for stations with the same degree to be connected each other, yet the various degrees are mostly mixing, i.e., smaller and larger stations are typically connected. The local clustering coefficient (Fig. 1D), defined as the fraction of pairs of neighbours of a given station that are connected over all pairs of neighbours of that station [31], can be used to infer the redundancy of the network. When a disruption occurs, a station with an high clustering coefficient can be bypassed easily. Complementarily, betweenness centrality (Fig. 1E) is a measure of the centrality of a node in a network based on the number of shortest paths that pass through it [31]. This measure highlights how strategic a given station is at a global level.

The clustering and the betweenness outline the typical small world topology of transportation networks [19]. German stations show a distribution between 10 and 100 trains per day, narrower than the train distribution of the Italian stations, which instead display a broader distribution, suggesting a more heterogeneous handling of train traffic load (Fig. 1F). Finally, Fig. 1G shows the histogram of the average delays of trains aggregated by

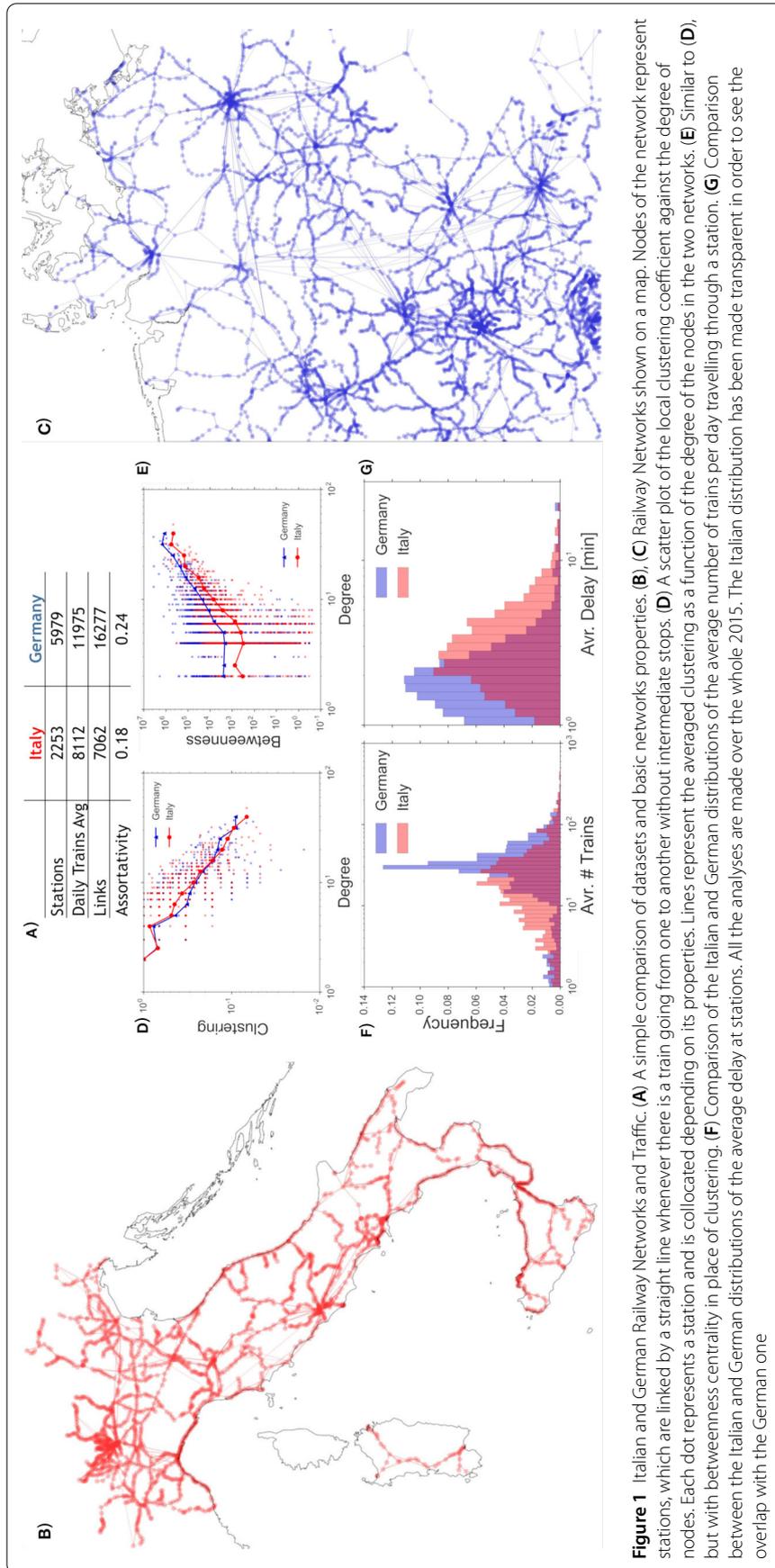
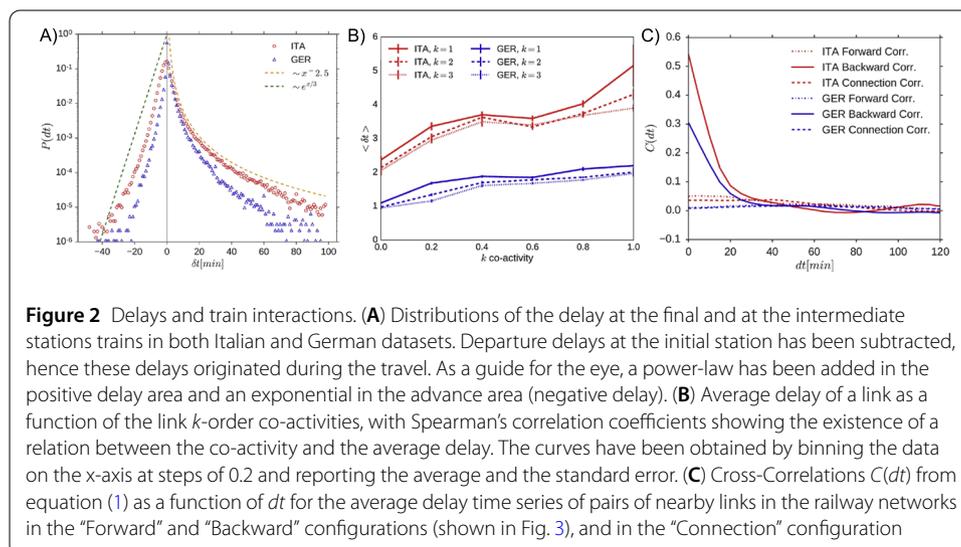


Figure 1 Italian and German Railway Networks and Traffic. **(A)** A simple comparison of datasets and basic networks properties. **(B), (C)** Railway Networks shown on a map. Nodes of the network represent stations, which are linked by a straight line whenever there is a train going from one to another without intermediate stops. **(D)** A scatter plot of the local clustering coefficient against the degree of nodes. Each dot represents a station and is collocated depending on its properties. Lines represent the averaged clustering as a function of the degree in the two networks. **(E)** Similar to **(D)**, but with betweenness centrality in place of clustering. **(F)** Comparison of the Italian and German distributions of the average number of trains per day travelling through a station. **(G)** Comparison between the Italian and German distributions of the average delay at stations. All the analyses are made over the whole 2015. The Italian distribution has been made transparent in order to see the overlap with the German one

stations for the two nations. Both distributions show a peak and heavy tail, whereas compared to the German distribution, the Italian distribution is clearly shifted towards higher delays. The topological similarities between the two networks suggest that the differences in delays and traffic load might be the result of differences in the trains dynamics.

2.2 Train interaction on railways networks

To focus on the analysis of delays, their outbreak and evolution, we choose trains instead of stations as our reference systems. The intermediate delays for a train i travelling from station A to station B on the link e is defined as $\Delta_i t(e)$. The departure delays at the initial station have been subtracted to analyse only the delays that have been generated during the travel. Hence $\Delta_i t(e)$ can be negative whenever the train is in advance, resulting in the train waiting at the station for the correct time of departure. Figure 2A shows the delay distributions for both national systems, considering the delays at intermediate stations along the path of a train, or just at the final station. We observe similar shapes of these distributions, with the Italian one exhibiting broader tails than the German. More than 10% of the train stops are on-time. The distributions of both countries exhibit an asymmetric pattern: the right tail (labeled Delay) shows a power-law like behaviour compatible with a q -exponential distribution [32], while the left tail (labelled Advance) has an exponential steeper slope. We expect this distribution to be the result of the interplay between the occurrence of adverse conditions and the interaction between trains, influencing their dynamics. Despite the fact that the microscopic details in our data do not allow for a precise investigation of possible interactions between trains, we can highlight how the possibility of interaction might affect the delays in railway networks. Hence, we study the relation between the *first order co-activity* of a link, i.e., the probability that at a certain time t (with time steps of 30 minutes) a link which is active has at least one active neighbouring link, to the average delay of the link itself. In Fig. 2B we show this quantity as a function of the average delay for both the Italian and German cases. We notice a slight increase in average delay as the co-activity increases, confirmed also by the Spearman's coefficients, 0.43 for Italy and 0.56 for Germany. Hence, the possibility of interaction between trains in a certain part of the railway network seems to increase the delay localized in that area. Note



that we have defined the *first order co-activity* between neighbouring links, i.e. links that have at least one node in common. We can define a *k-order co-activity* considering links connecting at least two nodes that are less than $(k - 1)$ links apart following the shortest-path connecting them. We report the same measurements for $k = 2$ and $k = 3$ showing that in general, the same relation with the average delay is confirmed even though the curves are shifted towards lower values. This indicates that considering the interaction between non-neighbouring links is relevant but might include less important contributions to the delay. Hence, in the following we will limit our analysis to the first order neighbour links and assume that interactions between trains are possible only when they are in nearby links. Figure 2B supports the thesis of a propagation effect but an important feature has yet to be determined. In fact the direction of the propagation still has to be determined.

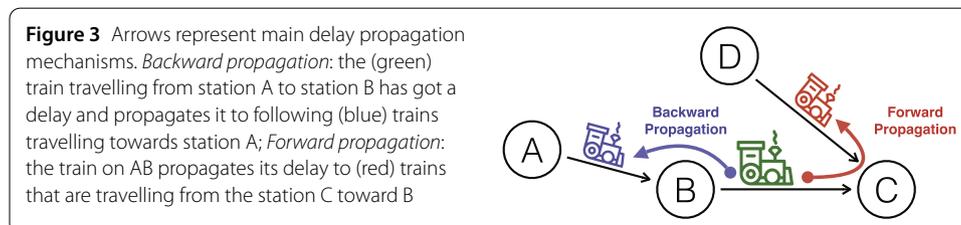
2.3 Defining possible interactions

Let us consider a train i travelling between two stations A and B (i.e., on the link $e = AB$, see Fig. 3) with some delay. We can argue that the propagation of this delay to other trains can occur only if they travel on the neighbour links of e . Due to the fact that the railway networks are directed, there are four different configurations of the links with respect to e :

- (i) links entering A , i.e. trains moving towards the last station crossed by i ;
- (ii) links entering B , i.e. trains travelling towards the same station i is currently travelling to;
- (iii) links exiting from A , i.e. trains departed from the last station crossed by i ;
- (iv) links exiting from B , i.e. trains leaving the station i is currently travelling to.

We can exclude the last two case: (iii) given that all the trains in such configuration will have no interaction with i ; (iv) is less important because it describes scheduled connections. In the latter case, schedules foresee extra-time between the two trains exactly to avoid delay propagation. We checked whether the propagation occurs in the case (i) of *backward propagation*, in the case (ii) of *forward propagation*, whose definitions are depicted in the graphic Fig. 3. To discriminate which mechanism is at play, we measured the average delay time sequence $\Delta t(e)$ of each link e of the network, defined as the average delay of all the trains that are currently travelling on e . Successively we measured the cross-correlation functions of the average delay time series of all the pairs of links, i.e., $CC_{e,e'}(dt) = \sum_t \Delta t_e(t) \Delta t_{e'}(t + dt) / \sigma_e \sigma_{e'}$ being e and e' generic neighbours links of the network, σ_e and $\sigma_{e'}$ the standard deviation of the whole time series $\Delta t_e(t)$ and $\Delta t_{e'}(t)$. Then we averaged, aggregating the pairs of neighbours links according to their configuration (forward, backward, etc). In this way, for each of the four configuration, we obtained an average cross-correlation function. In the backward propagation configuration can be defined:

$$C_B(dt) = \langle CC_{e,e'}(dt) \rangle_{(e,e') \in B} \quad (1)$$



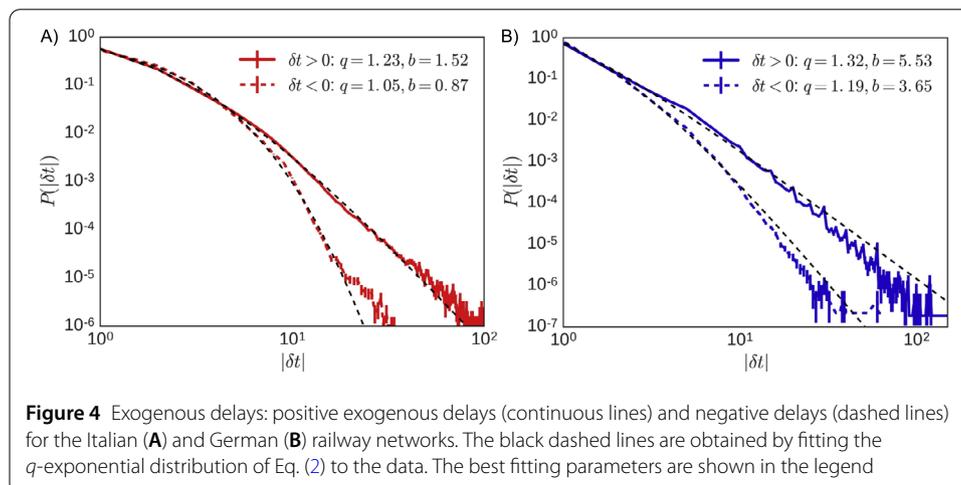
with \mathcal{B} as the ensemble of links pairs. For both networks the Backward mechanism is dominating while the Forward can be neglected (Fig. 2D). The high-speed layer of the railway network shows the similar backward mechanism, while there is no cross-correlations between the delays of high-speed vs. low-speed (see Additional file 1 Fig. S1) acting as two independent layers.

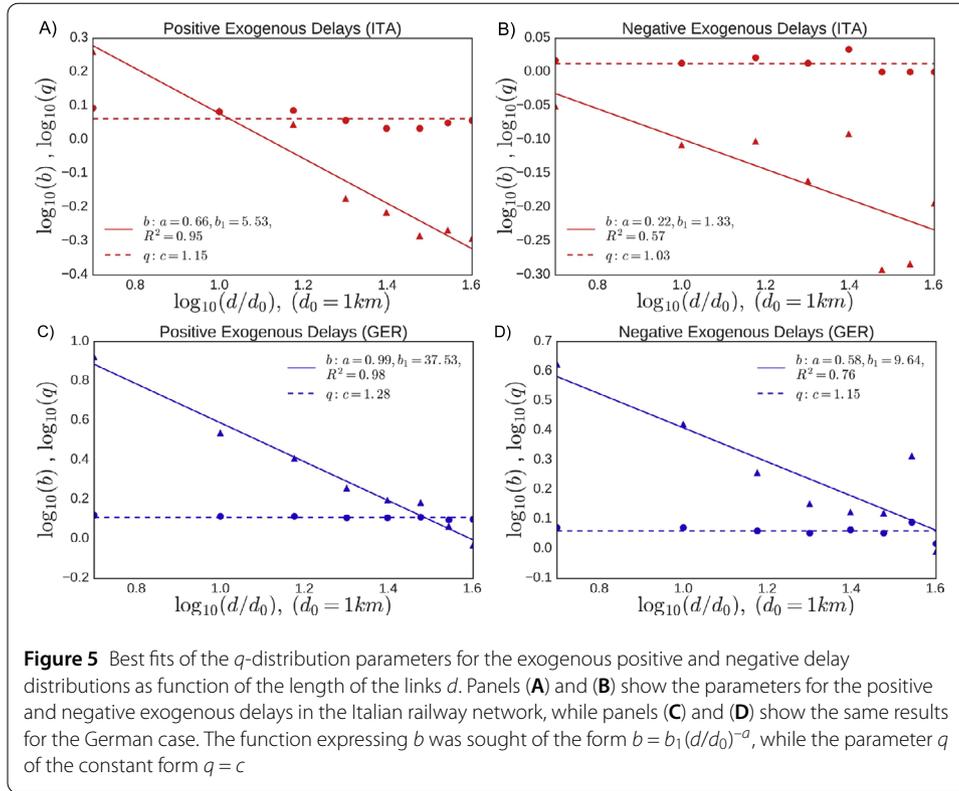
2.4 Exogenous generation of delay

We define two kinds of delays: endogenous and exogenous. By “endogenous” we mean that its origin is inside the railway system dynamics, i.e. it has been caused by another train. Conversely, by “exogenous” we mean that its cause is of a different nature: strikes, malfunctioning, bad weather or anything else which is not the result of the interaction with another train. We measure directly this types of delay in our datasets. Let us consider a train i travelling from a station A to a station B on the link e and further to a station C on the link e' . It will travel first on the link e and then on the link e' . The delay are indicate respectively as $\Delta t_i(e)$ and $\Delta t_i(e')$. If there is a increase in the delay $\Delta t_i(e') > \Delta t_i(e)$ it might be “endogenous” or “exogenous”. The exogenous delay is defined as $\delta t = \Delta t_i(e') - \Delta t_i(e)$, the variation of the train delay traveling on links whose neighbouring links were empty or hosted trains perfectly on time. It is worth noticing that δt might also be negative, for example, if the train managed to make up for lateness. Results are reported in Fig. 4, showing the distribution of positive exogenous delays as well as negative ones for the Italian and German cases. In order to model these distributions, we adopted the same approach used in [32] for departure delays. We fitted both the positive and negative parts of the distributions with q -exponential functions, where the parameter q modulates from an exponential distribution $q \rightarrow 1$ to a fat-tailed distribution for $q \in (1, 2]$ [33]:

$$e_{q,b}(\delta t) \propto (1 + b(q-1)\delta t)^{1/(1-q)} \quad \text{with } q \in [1, 2], b > 0. \quad (2)$$

It has been shown that such distribution can be obtained starting from a poissonian process $p(\delta t|\alpha) = \alpha e^{-\alpha t}$, where α is a random variable extracted from of n independent gaussian random variables X_i with $\langle X_i \rangle = 0$ and $\langle X_i^2 \rangle \neq 0$, so that $\alpha = \sum_{i=1}^n X_i^2$ [32]. In this way it can be proven that $n = 2/(q-1) - 2$, i.e. the parameter q , is related to the number of random variables composing α . The parameter b is proportional to the average value of α ,

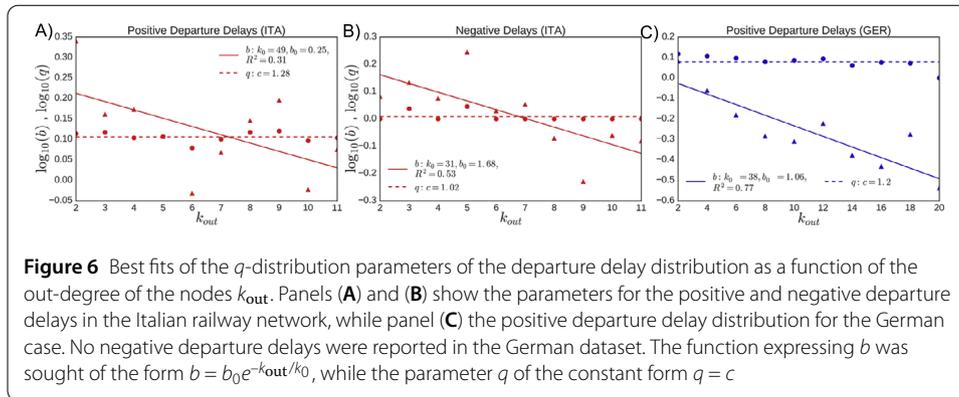




so that large values of b at fixed q result in a distribution biased toward shorter delays. The parameter q quantifies how much equation (2) deviates from being exponential, which is the case $q = 1$. This model has already been applied to the departure delays in the UK railway system, showing that the value of q where so that $4 \leq n \leq 11$ and thus estimating the number of independent occurrences contributing to the delay. For the positive exogenous delays in the Italian and German case respectively, we found $q = 1.23$ and $q = 1.32$, corresponding to $n \simeq 7$ and $n \simeq 4$. The negative part of the distribution is exponential-like for the Italian railway network and broader for the German, this outlines the delay recovery strategies in the second case. To characterize the effect of the spatial distance on the delay distribution, we subdivided the links e of the railway networks in classes according to the geodesic distance $d(e)$. Figure 5 shows the behavior of the q and b parameters of the q -exponential fit as functions of $d(e)$. The parameter q remains constant in every case, while on the other hand the parameter b decreases as $b \sim d^{-a}$. Figures S6–S9 of Additional file 1 show the different best fit for equation (2) as $d(e)$ varies. This result suggest that while the causes of the delay remain the same, the distribution of disturbances gets closer to a power-law as the length of the links increases, this outlines a relation between link length and delay. Finally, we can assume that the occurrence of positive or negative exogenous delays on links, $P(\delta t > 0|d(e))$ and $P(\delta t < 0|d(e))$, are not roughly constant with $d(e)$ and hence are not depending on the length of the links (Fig. S10 of Additional file 1).

2.5 Generation of delay at departure

Departure delays, i.e. the delay a train acquires right before leaving the first station on its route, cannot be considered in principle completely exogenous. In other words, due to the fact that different trains in our datasets can actually be the same physical train (e.g., the



same convoy travelling back and forth along the same path on the railway network – this is denoted as “rotation” –), the delay at departure might suffer from the influence of the traffic. However, railway administrators envisage suitable time buffers at the endpoints of the paths of each train so that it is reasonable to assume, at least as crude approximation, that departure delay is exogenous in character. It has already been shown that this kind of delay can be described by a q -exponential distribution in [32]. However, the dependence on the parameters of the obtained distributions with respect to the topological properties of the network has not been investigated yet. Following the same spirit of the previous paragraph for the exogenous delay on links, we divide the nodes in the network (the train stations), with respect to their out-degree. The out-degree k_{out} represents roughly the number of different railway lines starting from a certain stations and hence can be considered as a proxy for the complexity of the station itself. Once the nodes of the networks have been divided according to k_{out} , we fitted these distributions using a q -exponential following the procedure defined in [32] (see Fig. S12, Fig. S13 and Fig. S14 of Additional file 1).

Figure 6 shows the behaviour of the parameters q and b of the q -exponential distribution as functions of k_{out} for the positive and negative departure delays in the two considered railway networks. Negative departure delays were never reported in the German dataset and hence we assume they are not present. Despite the fact that better proxies for station complexity than k_{out} might exist (weighting each link with the actual number of railway lines on it is a valuable alternative example), it is possible to see that we have again a constant parameter q indicating that the sources of delays can be assumed to be the same independently from the station, while on the other hand the parameter b decreases exponentially with k_{out} . The small value of R^2 in the case of Italy might reflect the above mentioned possibility of having a non negligible endogenous contribution to the departure delays because of train rotations.

In Germany the departure delay can be considered constant and independent on the size of the station. In Italy the $P(\delta t > 0 | k_{out})$ grows linearly with k_{out} meaning that stations with high degree are generating larger disruptions in the network.

3 Modelling a realistic rail transport system

The analyses proposed so far suggest the existence of two sources of delays: *exogenous delays* spontaneously occurring due to external adverse conditions at departure and during travel, depending on the topological properties of the Railway Network; *endogenous delays* resulting from the interaction between trains. While in the first case we were able

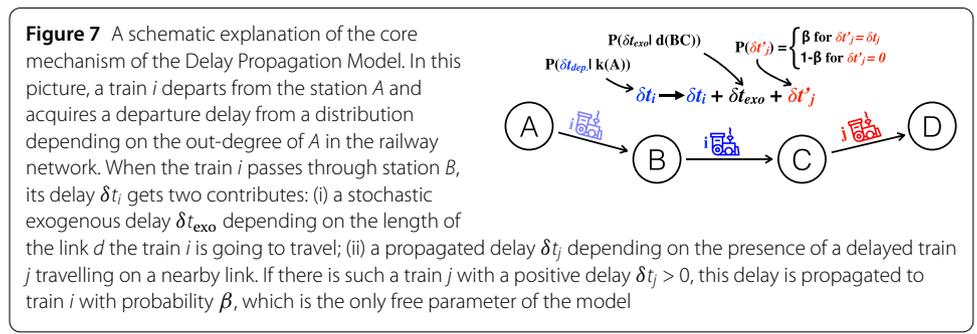
to characterize the statistical laws governing the emergence of delays, we cannot directly investigate the mechanisms of interaction between trains. Following the ideas exploited to model real world epidemics [34, 35] which have already been proven effective on air traffic delay modeling [22], we will define a propagation process for delays i.e. we will suppose that trains can spread their delays from one another according to some fixed probability and in certain conditions. Such probability will be derived by comparing the results from the simulations with the model and the empirical data.

3.1 The delay propagation model

We define the model scheme by leveraging the train-to-train propagation mechanism, with the aim of reproducing the real dynamics of delay spreading across the railroad network. The model starts reproducing the normal schedule of trains on a certain day. Train schedules are organized so that each train has its own “time window” to travel over a certain link along its path. Each train departure is at a fixed time and at a certain station and each intermediate station is going to be visited at a given time. However, our interest is in the deviance from the expected schedule. If some disruptions occurs a train might go out of the window of use for a certain link and overlap the window of another train for the same link. In this case, the second train will be forced to wait for its path to be cleared. Hence, the model adopts three different sources of delay as depicted in Fig. 7:

Departure delay This delay is assigned at the beginning of the path (originating station) of each train and is considered exogenous and unrelated with the current traffic conditions at the departing stations or in the nearby links. We assign to a train either a positive or negative delay according to the empirically found law of the corresponding probabilities $p_{dep}^+ = P(\delta t_{dep} > 0 | k_{out})$ and $p_{dep}^- = P(\delta t_{dep} < 0 | k_{out})$ respectively (and no delay with complementary probability $1 - p_{dep}^+ - p_{dep}^-$) (see Fig. S15 of Additional file 1). Once the sign of the delay has been decided, we assign a positive or negative delay value so that $|\delta t_{dep}| \sim e_{q(k_{out}), b(k_{out})}(\delta t)$, i.e. distributed according to a q -exponential distribution with the parameters q and b depending on k_{out} and on the sign of the delay itself as in Fig. 6.

Exogenous Link Delay This delay is assigned whenever a train starts travelling on a link for the first time (Fig. 7 reports an exemplification of the model). Considering a train i passing from the link AB to the link BC , we adopt the same modelling scheme as in the departure delay case assigning to the train a positive delay with probability $p_{exo}^+ = P(\delta t_{exo} > 0 | d(BC))$, a negative one with probability $p_{exo}^- = P(\delta t_{exo} < 0 | d(BC))$ and no delay with $1 - p_{exo}^+ - p_{exo}^-$, where $d(BC)$ is the geodesic length of the links BC (see Fig. S10 of Additional file 1 for the corresponding law of probability). Having



decided the sign of the delay, we fix the parameter of a q -exponential distribution q and b according to $d(BC)$ (Fig. 5) and we extract the magnitude of the delay so that $|\delta t_{\text{exo}}| \sim e_{q(d(BC)), b(d(BC))}(\delta t)$.

Delay Propagation We model the interaction between trains with a mechanism of propagation of delay from other trains to train i . The investigations performed on the datasets suggest that such interaction can occur only when nearby trains are in the *Backward propagation* configuration of Fig. 3.

Following the picture of the model in Fig. 7, when the train i starts travelling on the link BC leading to the C station is susceptible of propagation from the train j , which is travelling from C towards D . Inspired by the SIS models of epidemic spreading [24], in which agents can get infected and recover from the illness with a fixed probability, we introduce the delay propagation parameter $\beta \in [0, 1]$ describing the probability that the delay of the train j propagates to i . We assume that the propagation occurs at the time i starts travelling onto BC . When train i travels on BC from time t_1 to time t_2 , we check whether delayed trains are travelling on links in the Backward propagation configuration with respect to BC in $[t_1, t_2]$. Thus, we randomly pick one of those trains, say j , and with probability β its delay δt_j is added to the delay of i .

The model simulates the theoretic schedule applying all these delay-generating mechanisms. More in detail, we start the dynamics of each train i by adding a departure delay δt_{dep} according to the *Departure delay* mechanism. After the departure, each time the train starts traveling over a link in its route, we update its delay according to the following rule:

$$\delta t_i \rightarrow \delta t_i + \delta t_{\text{exo}} + \delta t_j, \quad (3)$$

where δt_{exo} is sample following *Exogenous Link Delay* mechanism, so that the delay will be positive with probability p_{exo}^+ and negative with p_{exo}^- and the parameters of the sampling distributions are depending on the length of the link as in Fig. 5. The term δt_j the contribution of the Delay Propagation which will different from 0 with probability β and just if there is at least one delayed train in the right configuration. This mechanism for intermediate stations reproduces the general noise associated with external causes, but does not account for long correlations such those present in case of large scale adverse conditions, e.g., bad weather, or national strikes. Finally, when the train reaches its final destination it is simply removed from the simulation. We used the described model to reproduce the delay spreading dynamics starting from a theoretic timetable and local exogenous delays distribution. Results of the simulation for both the Italian and German national railways systems are reported in the next section.

4 Results

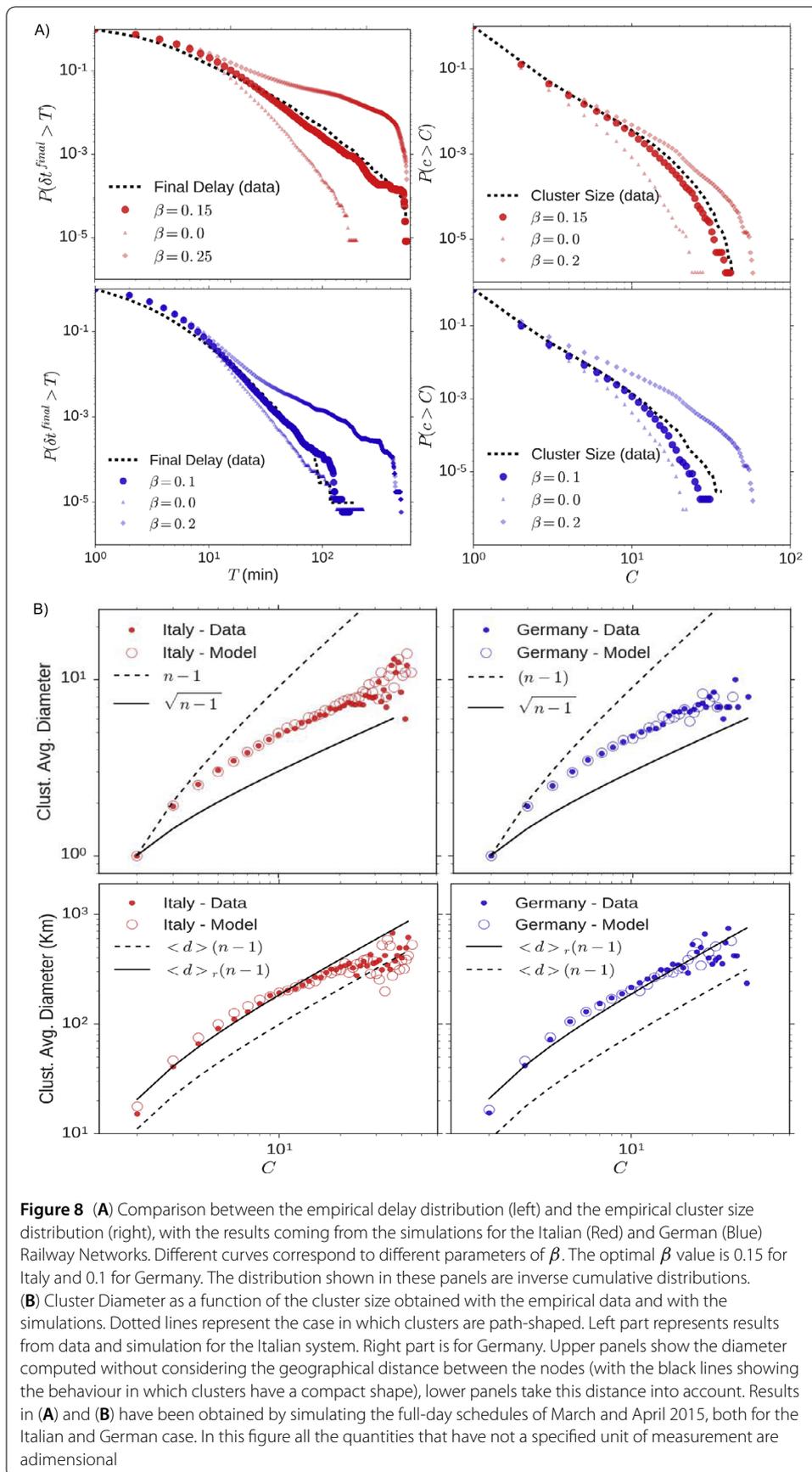
The model is based on the exogenous sources of delay that represent the spontaneous emergence of disruptions due to the aggregation of a finite number of external causes (trains malfunctions, accidents, bad weather, etc.). These sources are modelled according to a universal probabilistic law, whose parameters are inferred from the data at our disposal. The propagation parameter β modulates the mutual interaction between trains. For simplicity, we chose this parameter to be uniform all over the network, which is a strong assumption since the propagation might depend on the considered part of the railway system.

4.1 Reproducing the emergence of delays and congestion

In Additional files 1–3 we show that it is possible to estimate the best value for β by trying to reproduce the average delay of each station. These values have been estimated as $\beta = 0.15$ for Italy and $\beta = 0.1$ for Germany. In the top panels of Fig. 8A, we show that with these choices for β the system is capable of reproducing the empirical distributions of arrival delays. However, in order to disentangle the effect of the β parameter and the two other sources of exogenous delay, we also report the same result for the case in which $\beta = 0$ and β is larger than the optimal value. In the latter, we clearly see that there is an increase in the number of extremely large delays. On the other hand, turning off the interaction by setting $\beta = 0$ completely removes delays larger than 2 hours. Hence according to our model, the presence of exogenous delays by themselves would not be able to predict the presence of large disruptions. Since β represents the probability of delay propagation between two trains, these estimations give a quantitative account of the difference between the two countries. These results suggest that the Italian trains propagate each other delay more often than German trains. The microscopic reasons of this difference could be connected to different properties of the railway networks, to the peculiar geographical structure of the territory, to different delay handling policies, etc., and it is out of the scope of the present paper. As it happens in other transportation systems, we expect that disruptions occur clustered in certain areas of the network [22]. For this purpose we provide a definition of “cluster of congested stations” and how to discriminate whether a station is “congested”, i.e. when its functioning is inefficient because of the hoarding of delays. For each station we can define a threshold between the “functioning” and the “congested” status based on the value of incoming train delays. We calculated such threshold for each station as the average value of the delays of all the trains arriving at the considered station in the dataset. Considering a period of two months (March and April 2015), we examined each station every 5 minutes during the day, checking whether the average delay of all the trains moving towards that station during that lapse of time was above the average. We consider the station as “congested”, while if the average delay is below that threshold we consider it as “functioning”.

We identify, at each time step, the “clusters of congested stations” as the connected components of the railway network left once we removed all the functioning stations from the network. (i.e. the stations whose congestion might be correlated with one another because of the propagation of delay).

In order to check whether our model is able to reproduce the emergence of these areas, we focused on two measures: (i) the size of the clusters in terms of number of stations and (ii) the relation between the size and the diameter of the clusters. This latter measure is capable of giving insights about the topology of the congested clusters. Bottom panels of Fig. 8A show real and simulated cumulative distributions of cluster sizes for both countries. We find a strong accordance between model and reality when β is set to the optimal value, pointing out how the model is also able to reproduce this aspect of delays dynamics. Similarly to the arrival delay distribution, here β affects the tail of the distribution so that $\beta = 0$ (i.e. no propagation of endogenous delays) implies the absence of large congested areas. We study the diameter of the cluster defined as the distance between the farthest couple of nodes in the cluster. We compute the diameter both in the case where just the topological distance between nodes is considered and the case where each link is weighted with the geodesic distance between the stations it connects. To panels of Fig. 8B

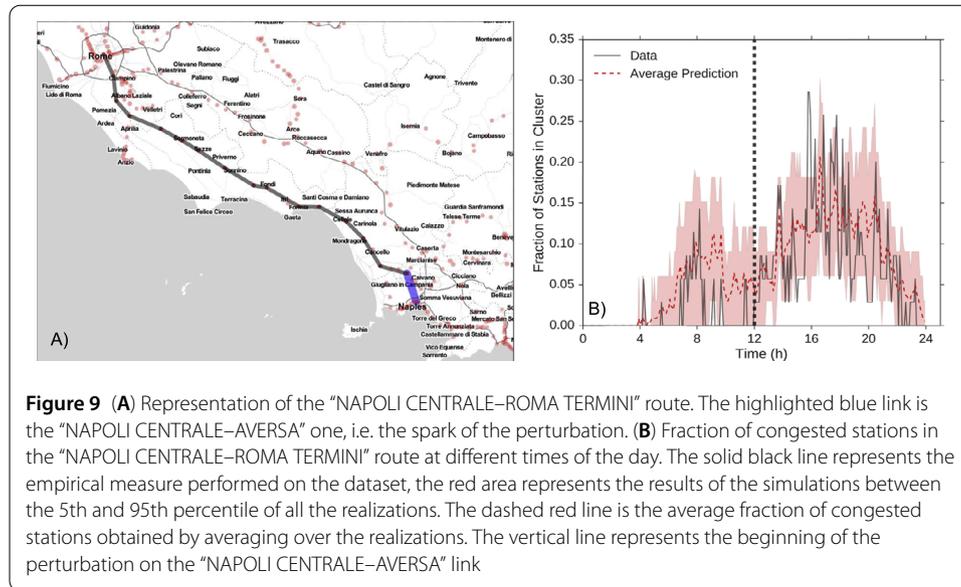


shows the dependence of the cluster diameter computed with both distances on the cluster size. In the cases where the diameter is computed by using the geodesic distance, the dotted line represents the diameter of a cluster of size n , assuming that all the links of the clusters are as long as the average link length of the network. Such pattern corresponds to the case where all the clusters are randomly sampled from the whole network without any constraint. From the lower panels of Fig. 8B we can see that while clusters do have a path-like structure, they cannot be considered randomly distributed over the networks. Instead, they seem to be deployed in areas where the links are larger than the average links length of the corresponding network, resulting in the same dependence of the dotted lines but shifted upwards. The fact that the geographical diameter of large clusters grows up to hundreds of kilometers indicates that disturbances can propagate between far away parts of the network.

The topological measure in the top panels of Fig. 8B exhibits strong deviations from the path-like behaviour especially when large clusters are involved. The mismatch between the two ways of characterizing clusters could be due to the fact that the geographical distance may hide, by stretching them, non path-like structures that are actually present in the delay propagation patterns. The appearance of these clusters, and their features, are typical outcomes of a non-trivial complex dynamics arising from train interactions. For this reason it is a remarkable result that, as can be neatly be seen in Fig. 8, the model proposed, despite its simplicity, captures this behaviour correctly. The relations between cluster size and cluster diameter is show very good agreements with the empirical measures for both nations and, perhaps more importantly, for both clusters diameter definition: the topological and the spatial one. This clear adherence with reality of the results from a model with only one parameter trained on a so small training set (only one week) is a strong proof of the fact that the model hypothesis are very likely to be correct. Our model is capable of reproducing some global patterns of the delay dynamics in railway systems. However, some discrepancies at a more microscopic level can be observed. In particular our model fails to correctly describe the behaviour of stations with low traffic. For these stations, the hypothesis of a constant in time and homogeneous coupling parameter β may not be well justified. We refer to Sect. 5 of Additional file 1 for a thorough discussion of this point.

4.2 Scenario simulation: prediction of the effects of a strong localized disturbance

According to our model, the emergence of large clusters of congested stations is due to the propagation of delay between trains. The source of this delay is however exogenous, in the sense that comes from some adverse conditions which are external with respect to the interaction of the trains. In order to investigate how the emergence of a real large cluster is linked to exogenous effects, we study the case of the cluster emerged in the eastern part of the Italian Railway Network on the 28th of February 2015. As can be seen in the Additional files 1–3 video or on the on-line visualization,^a a large congestion starts to emerge in the early afternoon around 12am and propagates to a large part of the network until night. We empirically identified the interested part of the network as the shortest-path connecting the station of “NAPOLI CENTRALE” to the station of “ROMA TERMINI”, indicated in Fig. 9A. The shortest-path has been computed weighting each link with the geographical distance between the stations it connects. Figure 9B shows the fraction of stations in this path that are congested in different times of the day, clearly indicating that such fraction



starts growing until a peak is reached in the afternoon. We have been able to identify the beginning of this adverse occurrence as a disruption on the “NAPOLI CENTRALE–AVERSA” link (highlighted in Fig. 9A) in the first part of the afternoon from 11 am to 14 pm, resulting in a large delay acquired by the trains travelling on that links. We argued that this disruption was the spark that lightened the emergence of the congested cluster.

In order to check this hypothesis, we ran a simulation in which a large delay of 100 minutes is assigned with probability 1 to each of the trains crossing the “NAPOLI CENTRALE–AVERSA” link in that period of time. Due to the non-deterministic nature of the model, we performed the simulation of 200 different realizations in order to have a set of scenarios to be compared with the real data. This comparison is shown in Fig. 9B. We can see that with this simple modification and not considering other possible correlations between the occurrences of external adverse conditions in nearby links, we are able to reproduce a pattern which is qualitatively similar to the one observed in the data, clearly pointing out that the “NAPOLI CENTRALE–AVERSA” link played a major role in the congestion of the network. Moreover, our scenarios indicates the probability of having a minor congestion in the line also in the early morning, probably due to usual minor disruption occurring on other links. The proposed scenario simulation could be easily extended to less localized adverse occurrences, which might comprehend spatially correlated disruptions due to natural events or strikes. Moreover, the same approach could be used to study more dramatic effect of node, link removal or the positive effects due to the introduction of more resilient and optimized schedules. In this cases, it is sufficient to use the same structure of the model and modify the input schedule and/or the Railway Network itself.

5 Conclusions

Railways and the railway transport systems have been historically of utmost importance for the development of modern countries. Nowadays their importance is on the rise again due to their relevance in the reduction of CO₂ emissions and their competitiveness in short and middle range movements, so that huge investment have been made in the European Union to improve their efficiency. However, the emergence of large disruptions and

the intrinsic inefficiencies seem to be endemic in this kind of transportation. The understanding of the universal features governing the dynamics of the system from a theoretical point of view could give an important contribution to the solution of such problems. The development of a universal model explaining the emergence of delays in Railway Network would allow for a quantification of the causes behind the occurrence of large disruptions and for a more direct link between the effects that localized interventions might have on the global system. In this work, we developed a novel model of delay emergence and propagation between trains moving on Railway Networks, partly based on empirical laws inferred from the data governing the accidental emergence of delays from the influence of adverse occurrences. Remarkably, the statistical models used for the description of this “exogenous delays” showed how these adverse occurrences are the result of the same finite number of causes (e.g., bad weather conditions and malfunctions) independently from the topology, as opposed to the scale of these disruptions, which are largely influenced by the part of the network the trains are travelling on. We find, in fact, that both the complexity of a station as measured by the number of different routes originating from it and the length of a route, are connected to the magnitude of the microscopic delays composing the overall delay by simple statistical laws with two parameters, whose value is the same all over the network and can be inferred and fixed by data. This kind of universality is somehow surprising and opens the possibility of easily simulating the behaviour of *any* railway system once the complete schedules and the network structure are known. These emergent delays can then spread from one train to another by means of a delay propagation probability β , which is in fact the only free parameter of our model. Despite the high level of abstraction used in our approach, our model is capable of reproducing the empirical patterns found in data when the delay-spreading probability β is fine tuned to an optimal value. The model reproduces correctly the final delay distribution and the distribution of the sizes of congested areas. Moreover, the model grasps the topological properties of such congested areas, which appears to be spreading in some cases for many kilometres through the network. According to our model, the emergence of large disruptions is the result of the interplay between the occurrence of localized exogenous delays and the propagation of such delays between trains. In other words, turning off the delay propagation mechanism prevents the system from generating extremely large delays and congested areas, pointing out that interactions are the driving force behind the emergence of major spontaneous adverse occurrences. The modelling scheme seems quite promising so far, but it still suffers from some major assumptions that might limit its predictive power. In fact, the interactions between trains could occur in longer ranges and not just between neighboring links and the propagation probability could not be uniform all over the Railway Network but instead depending on specific operational conditions. Moreover, the modelization of exogenous delays could be refined by taking into account more static topological feature of the Railway Networks or by introducing dynamical variations due to different traffic conditions. Another interesting possibility would be to increase the number of Railway Systems under study (e.g., the French SNFC sharing similar characteristics with respect to the Italian and German systems whose data is publicly available [36]), in order to understand the reasons behind the quantitative differences observed, e.g. larger delays in the Italian case. In the final part of the paper, we show how the model can be applied to study the capability of functioning of the system after a localized large disruption occurring in a single node of the network. This approach can be easily extended to more

complex case studies of distributed disruptions, the occurrence of strikes, the removal of trains or parts of the network, also including the possibility to introduce delay management strategies to increase the resilience of the system. The model is also useful in order to have a fast test of changes in the overall schedule of trains so to have a more precise assessment of their global effects. Finally, despite its simplicity the model is open an increase of complexity that might lead to a better adherence to the empirical findings such as long-range interactions between trains and a more detailed model of exogenous delays including correlations between nearby link and the dependence on traffic conditions.

Additional material

Additional file 1: Supplementary information: Complex delay dynamics on railway networks from universal laws to realistic modelling (PDF 5.2 MB)
Additional file 2: Supplementary interactive simulation files (ZIP 78.5 MB)
Additional file 3: Supplementary data (ZIP 480 kB)

Acknowledgements

The authors acknowledge Fabio Lamanna for the initial discussion about the datasets to be used for the work.

Funding

This work has been supported by the KREYON Project, funded by the John Templeton Foundation under contract n. 51663; VDPS acknowledges financial support from the Austrian Research Promotion Agency FFG under grant #857136. RDC as Newton International Fellow of the Royal Society acknowledges support from The Royal Society, The British Academy and the Academy of Medical Sciences (Newton International Fellowship, NF170505).

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article (and its additional files).

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare no competing interests.

Consent for publication

Not applicable.

Authors' contributions

BM and RDC conceived and designed the analysis; BM analysed the datasets; BM and VDPS designed the model; BM performed the simulations; BM, PG, RDC and VDPS wrote the manuscript. All authors read and approved the final manuscript.

Author details

¹Institute for Scientific Interchange Foundation, Turin, Italy. ²SONY Computer Science Lab, Paris, France. ³"Sapienza" University of Rome, Roma, Italy. ⁴The Bartlett Centre for Advanced Spatial Analysis, University College London, London, United Kingdom. ⁵Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, USA. ⁶Complexity Science Hub Vienna, Vienna, Austria.

Endnote

^a The visualization can be found at www.riccardodiclemente.com/trainsimulation.html

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 22 June 2017 Accepted: 30 August 2018 Published online: 17 September 2018

References

1. Banister D, Watson S, Wood C (1997) *Environment and planning B: planning and design* 24(1) p 125
2. Limao N, Venables AJ (2001) *World Bank Econ Rev* 15(3):451
3. Newman P, Kenworthy J (1999) *Sustainability and cities: overcoming automobile dependence*. Island Press, Washington
4. Train KE (2009) *Discrete choice methods with simulation*. Cambridge University Press, Cambridge

5. Weiss, DJ Nelson, A Gibson, HS Temperley, W Peedell, S Lieber, A Hancher, M Poyart, E Belchior, S Fullman, N (2018) *Nature* 553(7688):1476–4687
6. Donaldson D (2018) *Am Econ Rev* 108(4–5):899
7. Carrion C, Levinson D (2012) *Transp Res, Part A, Policy Pract* 46(4):720
8. Noland RB, Polak JW (2002) *Transp Rev* 22(1):39
9. Banister D (2008) *Transp Policy* 15(2):73
10. Meester LE, Muns S (2007) *Transp Res, Part B, Methodol* 41(2):218
11. Goverde RM (2010) *Transp Res, Part C, Emerg Technol* 18(3):269
12. Higgins A, Kozan E (1998) *Transp Sci* 32(4):346
13. D'Ariano A, Pranzo M, Hansen IA (2007) *IEEE Trans Intell Transp Syst* 8(2):208
14. Li K, Gao Z, Ning B (2005) *J Comput Phys* 209(1):179
15. Cacchiani V, Caprara A, Toth P (2010) *Transp Res, Part B, Methodol* 44(2):215
16. Şahin İ (1999) *Transp Res, Part B, Methodol* 33(7):511
17. Giua A, Seatzu C (2008) *IEEE Trans Autom Sci Eng* 5(3):431
18. Middelkoop D, Bouwman M (2001) In: *Proceedings of the 33rd conference on Winter simulation. IEEE Comput. Soc., Los Alamitos*, pp 1042–1047
19. Sen P, Dasgupta S, Chatterjee A, Sreeram P, Mukherjee G, Manna S (2003) *Phys Rev E, Stat Nonlinear Soft Matter Phys* 67(3):036106
20. Erath A, Löchl M, Axhausen KW (2009) *Netw Spat Econ* 9(3):379
21. Bhatia U, Kumar D, Kodra E, Ganguly AR (2015) *PLoS ONE* 10(11):e0141890
22. Fleurquin P, Ramasco JJ, Eguíluz VM (2013) *Scientific reports* 3
23. Campanelli B, Fleurquin P, Eguíluz V, Ramasco J, Arranz A, Etxebarria I, Ciruelos C (2014) In: Schaefer D (ed) *Proceedings of the Fourth SESAR Innovation Days, Madrid*
24. Pastor-Satorras R, Vespignani A (2001) *Phys Rev Lett* 86(14):3200
25. Viaggiatreno. <http://www.viaggiatreno.it/>. Accessed: 2018-05-29
26. OpenDataCity. <http://www.opendatacity.de/>. Accessed: 2017-05-10
27. The CIA World Factbook. <https://www.cia.gov/library/publications/the-world-factbook/>. Accessed: 2018-05-29
28. Kurant M, Thiran P (2006) *Phys Rev E, Stat Nonlinear Soft Matter Phys* 74(3):036114
29. Barthélemy M (2011) *Phys Rep* 499(1):1
30. Newman ME (2002) *Phys Rev Lett* 89(20):208701
31. Newman ME (2003) *SIAM Rev* 45(2):167
32. Briggs K, Beck C (2007) *Phys A, Stat Mech Appl* 378(2):498
33. Picoli S Jr, Mendes R, Malacarne L, Santos R (2009) *Braz J Phys* 39(2A):468
34. Gomes MF, y Piontti AP, Rossi L, Chao D, Longini I, Halloran ME, Vespignani A (2014) *PLOS Currents Outbreaks*
35. Zhang Q, Sun K, Chinazzi M, Pastore-Piontti A, Dean NE, Rojas DP, Merler S, Mistry D, Poletti P, Rossi L et al (2016) *bioRxiv* p 066456
36. SNCF Open Data. <https://ressources.data.sncf.com/explore/dataset/api-sncf/>. Accessed: 2018-05-29

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
