# Modeling Airport Congestion Contagion by SIS Epidemic Spreading on Airline Networks

Köstler, Klemens ; Gobardhan, Rommy ; Ceria, Alberto; Wang, Huijuan

**DOI**
10.1007/978-3-030-36687-2_32

**Publication date**
2020

**Document Version**
Accepted author manuscript

**Published in**
Complex Networks and Their Applications VIII

**Citation (APA)**
Köstler, K., Gobardhan, R., Ceria, A., & Wang, H. (2020). Modeling Airport Congestion Contagion by SIS Epidemic Spreading on Airline Networks. In H. Cherifi, S. Gaito, J. F. Mendes, E. Moro, & L. M. Rocha (Eds.), *Complex Networks and Their Applications VIII : Proceedings of the 8th International Conference on Complex Networks and Their Applications, COMPLEX NETWORKS 2019* (Vol. 1, pp. 385-398). (Studies in Computational Intelligence; Vol. 881 ). Springer. https://doi.org/10.1007/978-3-030-36687-2_32

**Important note**
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

# Modeling Airport Congestion Contagion by SIS Epidemic Spreading on Airline Networks

Klemens Köstler[1], Rommy Gobardhan[2], Alberto Ceria[2], and Huijuan Wang[2]

[1] Faculty of Aerospace Engineering, Delft University of Technology, Delft, 2628CD, The Netherlands

[2] Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, 2628CD, The Netherlands

H.Wang@tudelft.nl

**Abstract.** We model airport congestion contagion as an SIS spreading process on an airport transportation network to explain airport vulnerability. The vulnerability of each airport is derived from the US Airport Network data as its congestion probability. We construct three types of airline networks to capture diverse features such as the frequency and duration of flights. The weight of each link augments its infection rate in SIS spreading process. The nodal infection probability in the meta-stable state is used as estimate the vulnerability of the corresponding airport. We illustrate that our model could reasonably capture the distribution of nodal vulnerability and rank airports in vulnerability evidently better than the random ranking, but not significantly better than using nodal network properties. Such congestion contagion model not only allows the identification of vulnerable airports but also opens the possibility to reduce global congestion via congestion reduction in few airports.

**Keywords:** Airline transportation network, epidemic spreading, network vulnerability

## 1 Introduction

Networks are ubiquitous in nature and support the flow of goods, propagation of information and transmission of epidemic disease. Therefore, networks and processes that unfold on them have been the subject of many studies in a wide range of fields such as mathematics, engineering and social sciences [1–3, 5]. The Susceptible-Infected-Susceptible (SIS) epidemic spreading process is one of the most widely studied dynamic processes on networks [4, 6–11]. An individual is either susceptible $S$ or infected $I$ at any time $t$. Each infected node infects each of its susceptible neighbors with an infection rate $\beta$. Each infected node can be recovered with a recovery rate $\delta$. Both infection and recovery processes are independent Poisson processes. The ratio of the infection rate over recovery rate $\tau{=}\beta/\delta$ is called the effective infection rate. In the SIS model on a given network, a critical epidemic threshold $\tau_c$ exists. When $\tau > \tau_c$, a non-zero fraction of nodes are infected in the meta-stable state. In contrast, the epidemic dies out if

$\tau < \tau_c$. The average fraction of infected nodes $v(t)$, also known as the prevalence, estimates the vulnerability of the network subject to an epidemic whereas the infection probability $v_i(t)$ of an individual $i$ indicates the vulnerability of node $i$ subject to an epidemic. Recent literature has focused on the understanding of how network topology influences the epidemic threshold, the prevalence [12, 13] and the nodal infection probability [14]. However, it is still unknown to what extent the SIS model can identify the vulnerability of a network or nodes subject to other challenges such as traffic congestion.

From a complex network perspective on air transport, initial research focused on analyzing the topological structure of an air traffic system [15, 16] demonstrating small-world and scale-free characteristics. Previous studies have explored network topology by analyzing subsets of a network based on geography and airlines/alliances [17, 18]. Recent research effort has been devoted to network resilience or vulnerability with respect to delay propagation [3, 19–21] and random failures [22, 23]. To date, there are few studies that have investigated traffic congestion based on network dynamics, although queuing models have been studied [24].

In this paper, we investigate the limits of modelling airport congestion contagion as an SIS process on air transportation networks in identifying airport vulnerability. The US Airport Network data [25] is considered. We construct three types of airport networks to capture diverse features such as the frequency and duration of flights. We perform the SIS spreading process with varying effective infection rate $\tau$ upon each of these three types of networks. The nodal infection probability in the meta-stable state is derived as an indicator of the vulnerability of the corresponding airport. We define and derive the actual airport vulnerability as the duration of traffic congestion over the total operation time using the US Airport Network data. An airport is congested if the number of flight movements is greater than its declared capacity. Our model is evaluated by comparing the nodal infection probabilities derived from the epidemic spreading model with the actual airport vulnerability determined from day-to-day operations.

The content of this paper is arranged as follows. Section 2 defines, derives and characterizes the actual airport vulnerability. Section 3 describes the SIS epidemic model and network construction. Section 4 evaluate our model via its capability to reproduce the airport vulnerability distribution and ranking of airports in vulnerability. Section 5 illustrates the possible generalization of the model. Section 6 summarizes our key findings and discusses possible future work.

## 2 Traffic vulnerability of an airport

In this section, we first introduce the US Airport Network data. Afterwards, we define and derive the airport vulnerability from the data and explore its distribution. The actual vulnerability of each airport will be later used as a benchmark to evaluate our SIS model.

### 2.1 Data

The US Airport network data was obtained from the Bureau of Transportation Statistics (BTS) which contains information regarding the flight schedule in the U.S. from 1987 until present [25]. Flight schedules are further distinguished by the computer reservation system (CRS), known as the planned schedule under optimal operation conditions, and the actual schedule. To illustrate our methods, we use the data spanning two weeks' high season period from July 1st 2019 until July 14th 2019 because flight schedule and rotations repeat. The data set contains $N = 349$ airports and $E = 645299$ flights. Each flight contains additional information such as: Tail-Number, Date, Origin, Destination, the actual and planned Departure/Arrival Times.

### 2.2 Definition and statistical properties

We define the vulnerability of an airport as its duration of traffic congestion over its total operation time, i.e. the probability of an airport being congested. An airport is congested if the actual number of movements (the number of arrival and departure flights during operation) per hour is greater than its declared capacity. The declared capacity of an airport at a given hour is approximately the number of movements scheduled/planned for that hour, at which a reasonable level of service (LOS) can be provided. Typically, an airport declares a capacity up to 85-95% of the maximum throughput capacity, which is the expected number of movements that can be achieved within one hour on a runway system without violating air traffic management rules (assuming continuous aircraft demand). From the U.S. airport network data, we derive the state of airport $i$ at any hour $t$ as congested $X_i(t) = 1$ or not $X_i(t) = 0$ based on whether the actual number of movements is larger than the scheduled number of movements at time $t$ or not. Airport $i$'s vulnerability $\phi_i = \frac{1}{m} \sum_{t=1}^{m} X_i(t)$ is thus the fraction of hours that airport $i$ is congested. All time hours within the two week period were considered (except for hours between 0 and 6 due to their low number of movements) and indexed as $[1, 2, ..., m]$, where $m = 18 \cdot 14 = 252$.

Figure 1 presents the probability density function of airport vulnerability. The maximal vulnerability is less than 0.5 and the distribution is slightly heterogeneous.

## 3 SIS Model on Airline Transportation Networks

We model the contagion of airport congestion as an SIS spreading process on an airline transportation network. Section 3.1 introduces how we construct the three types of airline transportation networks. Section 3.2 describes the individual-based mean field approximation method to derive nodal infection probabilities of the SIS model in the meta-stable state, given the underlying network, the infection rate $\beta$ and the recovery rate $\delta$.

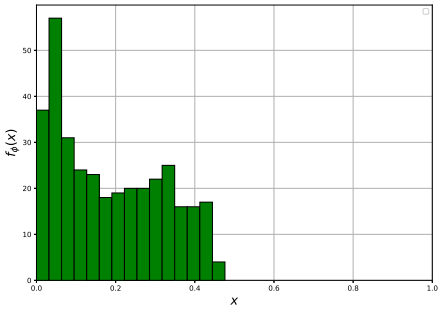



Fig. 1: The probability density function $f_\phi(x)$ of the vulnerability $\phi$ of an airport.

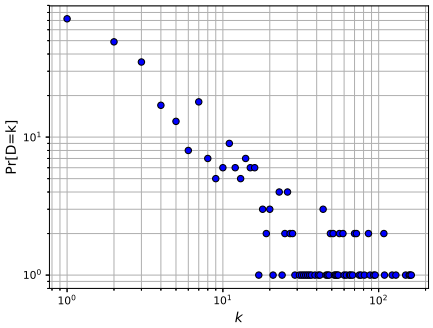



Fig. 2: Degree distribution of the airport networks.

### 3.1 Network Construction

We propose to construct three types of networks from the data to capture diverse domain information, because the SIS spreading process unfolds differently depending on the underlying networks. All three types of networks are un-directed and constructed from the Airport Network Data over the two weeks' period. Network $G_1$ is unweighted where two airports (nodes) are connected if there is at least one direct flight between them. Hence, the weight of each link is $w_{ij} = 1$. Network $G_2$ and $G_3$ are both weighted and have the same network topology as network $G_1$. We assume that the infection rate of a link is proportional to the weight of that link. In $G_2$, the weight of the link between node $i$ and $j$ equals $w^*_{ij} = F_{ij} + F_{ji}$ the total number of flights $F_{ij}$ from $i$ to $j$ plus the number of flights $F_{ji}$ from $j$ to $i$ within the given two weeks' period. This construction is motivated by the hypothesis that the more frequent the flights between two airports, the more likely that delay/congestion is transferred from one airport to the other. Moreover, the flight time between airports may also affect the congestion propagation. In order to compensate for a delayed departure, airplanes can re-optimize their velocity to remain on schedule at the destination airport. Network $G_3$ aims to capture these effects by weighting each link by the inverse of its average flight time, $w^*_{ij} = \frac{1}{E[T_{ij}]}$. A smaller average flight time may suggest that flights delayed at the departure airport likely affect the arrival time at the destination airport, while a larger average flight time gives more room to re-optimize the flight velocity.

Finally, we generalize the choice of the weights in Networks $G_2$ and $G_3$ as

$$w_{ij} = \left( \frac{w^*_{ij}}{\max\limits_{k,l} w^*_{k,l}} \right)^{\alpha} . \tag{1}$$

The normalization of each initial link weight $w^*_{ij}$ by the maximum in each network ensures that the link weights remain in the range $[0, 1]$. The final choice

of the weight $w_{ij}$ is generalized as a polynomial function of $w^*_{ij}$, the normalized number of flights in $G_2$ and the normalized average flight duration in network $G_3$. The infection rate of a link with weight $w_{ij}$ is $\beta \cdot w_{ij}$. Following our hypothesis, $\alpha \geq 0$.

### 3.2 Individual-Based mean-field approximation of SIS model

We derive nodal infection probabilities via mean-field approximations instead of simulating the SIS stochastic processes for efficiency. The N-Intertwined Mean-Field Approximation (NIMFA) is one of the most precise individual-based mean-field approximations of the SIS model [5], assuming that the states of neighboring nodes are uncorrelated. For the SIS model with the infection rate $\beta$ and recovery rate $\delta$, the single governing equation for a node $i$ in NIMFA is

$$\frac{\mathrm{d}v_i(t)}{\mathrm{d}t} = -\delta v_i(t) + \beta(1 - v_i(t)) \sum_{j=1}^{N} w_{ij} v_j(t) \tag{2}$$

where $v_i(t)$ is the infection probability of node $i$ at time $t$, and each link weight $w_{ij}$ augments the infection rate of link $(i,j)$. In the meta-stable state, $\frac{\mathrm{d}V(t)}{\mathrm{d}t} = 0$, where $V(t) = [v_1(t)\ v_2(t)\ \cdots\ v_N(t)]^T$, $\lim_{t\to\infty} v_i(t) = v_{i\infty}$ and $\lim_{t\to\infty} V(t) = V_\infty$. The infection probability of each node $V_\infty$ in the meta-stable state can be derived. The trivial, i.e. all-zero, solution indicates the absorbing state where all nodes are susceptible. The non-zero solution of $V_\infty$, if exists, points to the existence of a meta-stable state with a non-zero fraction of infected nodes. Or else, the meta-stable state can be figured as 0 or not existing. The infection probability of each node in the meta-stable state depends only on the effective infection rate $\tau = \frac{\beta}{\delta}$ and the underlying (weighted) network.

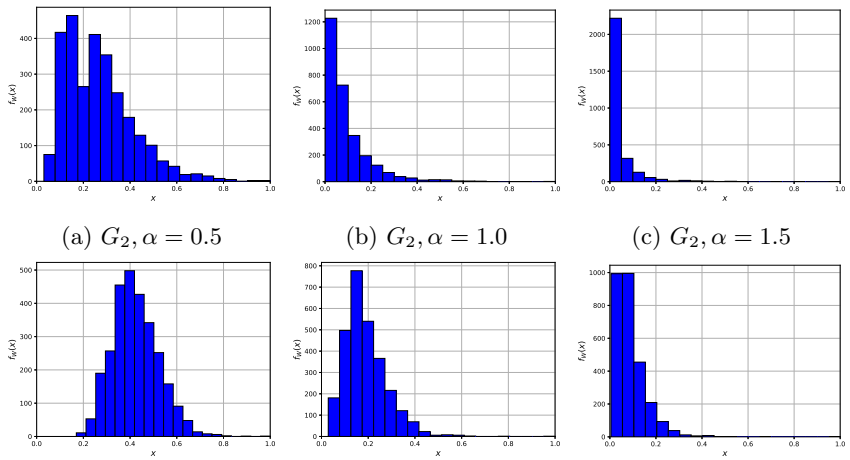
## 4 Results and Discussion

Since the network underlying the SIS spreading process influences the infection probability of the nodes, we first analyze basic properties of the three types of networks constructed, especially the effect of the scaling parameter $\alpha$ on link weight distribution, in section 4.1. Furthermore, we evaluate our SIS spreading model on an airport network ($G_1$, $G_2$ or $G_3$) in identifying airport vulnerability using nodal infection probability via the model's capability to capture the airport vulnerability distribution (section 4.2) and the ranking of airports in vulnerability (section 4.3).

### 4.1 Network Characteristics

All three types of networks have the same network topology. Their degree distribution, as shown in fig. 2 approximates a power-law degree distribution $Pr[D = k] \sim k^{-\gamma}$. As introduced in section 3.1, the links weights in $G_2$ and $G_3$ are

scaled $w_{ij} = \left( \frac{w^*_{ij}}{\max_{l,k}(w^*_{l,k})} \right)^{\alpha}$ by a scaling factor $\alpha$. An increase in $\alpha$ results in a smaller link weight. The epidemic threshold $\tau_c$ has been shown to approximate $\lambda_{max}(W))^{-1}$, where $W$ is the weighted adjacency matrix of the underlying network whose element $w_{ij}$ is the weight of link $(i, j)$ as defined in section 3.1 [10,26]. Hence, the epidemic threshold tends to be larger as the scaling factor $\alpha$ increases. As $\alpha$ increases, a similar average infection probability can be obtained by a larger effective infection rate $\tau$.



(a) $G_2, \alpha = 0.5$ (b) $G_2, \alpha = 1.0$ (c) $G_2, \alpha = 1.5$

(d) $G_3, \alpha = 0.5$ (e) $G_3, \alpha = 1.0$ (f) $G_3, \alpha = 1.5$

Fig. 3: Probability density function $f_W(x)$ of the weight $W$ of a random link in Network $G_2$ and $G_3$, when the link weight is scaled with a parameter $\alpha = \{1, 0.5, 1.5\}$.

The heterogeneous link weights in a network influences not only the average probability of infection but also the infection probability of each node [8, 14]. We explore further how the link weight definitions in $G_2$ and $G_3$, especially the scaling of link weights, influence the link weight distribution. Our model would be more general, if $G_2$ and $G_3$ could generate diverse link weight distributions. Figure 3 shows the probability density function of link weights in $G_2$ and $G_3$, when $\alpha = 0.5, 1$ and 1.5. When link weights are not scaled, i.e. $\alpha = 1$, network $G_2$ and $G_3$ manifest different distributions: link weights in $G_2$ are more heterogeneously distributed than those in $G_3$. A scaling parameter $\alpha < 1$ reduces the heterogeneity whereas a scaling factor $\alpha > 1$ increases the heterogeneity of link weights. The construction of network $G_2$ and $G_3$ and the link weight scaling via parameter $\alpha$ allow our model to capture diverse link weight distributions and consequently diverse nodal infection probability distributions.

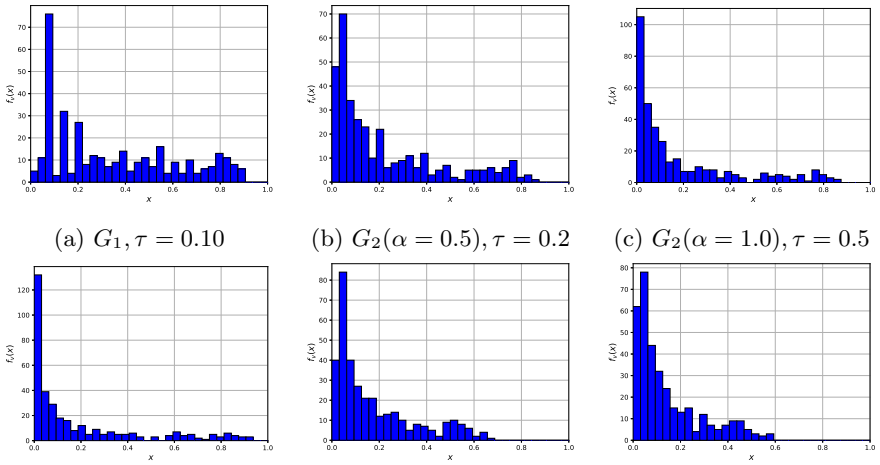
### 4.2 Model evaluation via vulnerability distribution

We consider the scaling parameter $\alpha = 0, 0.5, 1.5$ as examples. When $\alpha = 0$, all three networks are the same, i.e. $G_1 = G_2(\alpha = 0) = G_3(\alpha = 0)$. For each network type, each scaling parameter $\alpha$ and each effective infection rate $\tau \in (0, 2]$ with step size 0.1, we compute the infection probability for each node in the meta-stable state. Given the network type and $\alpha$, we identify the $\tau$ value at which the average nodal infection probability is the closest to the average airport vulnerability, denoted as $\tau_o$ (see Table 1). In other words, the average airport vulnerability is used to calibrate our model, i.e. to derive the optimal effective infection rate $\tau_o$. As the scaling factor $\alpha$ of $G_2$ or $G_3$ increases, the average link weight decreases and the epidemic threshold of the weighted network $\lambda_{max}(W))^{-1}$ increases, which requires a larger value of $\tau_o$ to reproduce the target average airport vulnerability. This has been confirmed by Table 1.

Table 1: The effective infection rate $\tau_o$ when the average nodal infection probability is the closest to the average airport vulnerability and epidemic threshold $\lambda_{max}(W))^{-1}$ for diverse network types and and link weight scaling parameter $\alpha$ values.

| | $\tau_o$ | | | $\lambda_{max}(W)^{-1}$ | | |
|---|---|---|---|---|---|---|
| $\alpha$ | G1 | G2 | G3 | G1 | G2 | G3 |
| 0.50 | 0.10 | 0.20 | 0.10 | 0.200 | 0.057 | 0.049 |
| 1.00 | 0.10 | 0.50 | 0.20 | 0.200 | 0.123 | 0.115 |
| 1.50 | 0.10 | 1.40 | 0.50 | 0.200 | 0.216 | 0.254 |

Figure 4 demonstrates the probability density function of nodal infection probability corresponding to the network type, link weight scaling parameter $\alpha$ and $\tau_o$ combinations mentioned in table 1. The nodal infection probability distribution in Network $G_2$ and $G_3$ better resemble the airport vulnerability distribution shown in Figure 1 than that in $G_1$, which is relatively homogeneous over a broad rang of infection probability. The maximal nodal infection probability obtained in Network $G_3$ better approximates the maximal airport vulnerability than that in network $G_2$. These findings support the initial hypothesis that airport congestion spreads through airports that can be reached in a short time. Still, the maximal infection probability in $G_3$ is larger than the maximal airport vulnerability. A possible explanation of this is that the SIS model does not take into account operational measures implemented by airports. Daily airport operations focuses on reducing long periods of infection because they are detrimental.



(a) $G_1, \tau = 0.10$ (b) $G_2(\alpha = 0.5), \tau = 0.2$ (c) $G_2(\alpha = 1.0), \tau = 0.5$

(d) $G_2(\alpha = 1.5), \tau = 1.4$ (e) $G_3(\alpha = 0.5), \tau = 0.1$ (f) $G_3(\alpha = 1.0), \tau = 0.2$

Fig. 4: Probability density function of the infection probability of a node in all network type, link weight scaling parameter $\alpha$ and $\tau_o$ combinations mentioned in Table 1.

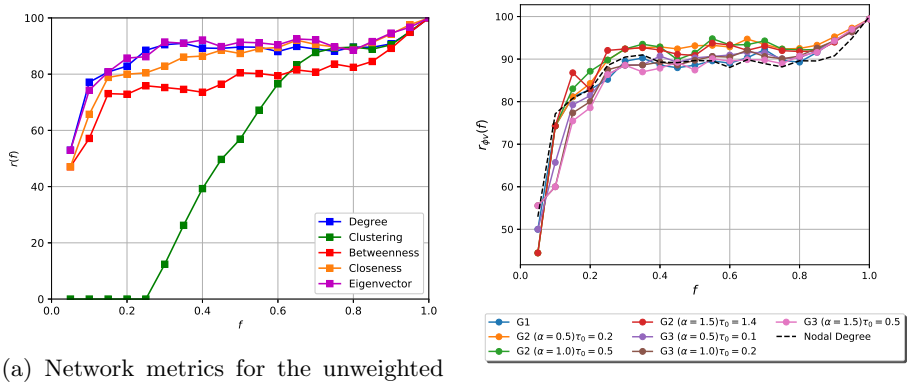
### 4.3 Model evaluation via airport ranking in vulnerability distribution

Identifying the most vulnerable airports is crucial from operation point of view. The quality of using nodal infection probability to rank airports in vulnerability can be evaluated as follows. A node with a higher infection probability is supposed to have a higher vulnerability. Nodal infection probability and airport vulnerability are used, respectively, to rank the nodes/airports. The corresponding rankings can be recorded by two vectors $R^v = [R^v_{(1)}, R^v_{(2)}, ..., R^v_{(N)}]$ and $R^\phi = [R^\phi_{(1)}, R^\phi_{(2)}, ..., R^\phi_{(N)}]$ where $R^v_{(i)}$ is the index of the node with the $i-th$ highest infection probability and $R^\phi_{(i)}$ is the index of the $i-th$ most vulnerable airport. How precise nodal infection probability could be used to rank airports in vulnerability can be quantified by the top $f$ recognition rate

$$r_{\phi v}(f) = \frac{|R^\phi_f \cap R^v_f|}{|R^\phi_f|} \tag{3}$$

where $R^\phi_f$ and $R^V_f$ are the sets of nodes ranking in the top $f$ fraction according to their vulnerability and infection probability respectively and $|R^\phi_f| = fN$ is the number of nodes in $R^\phi_f$. The recognition rate $r_{\phi v}(f)$ measures the quality of using nodal infection probability to recognize or identify the top $f$ fraction most vulnerable airports. Given the fraction $f$, a higher recognition rate suggests that the nodal infection probability could better rank the nodes in their vulnerability.

We first evaluate the recognition rate of our model that is composed of the network type, link weight scaling parameter $\alpha$ and the corresponding $\tau_o$ as given in Table 1. As shown in fig. 5b, the recognition rate $r_{\phi v}(f)$ in relation to $f$ is relatively stable as the network type, parameter $\alpha$ and $\tau_o$ vary. Moreover, our model performs better than the random strategy, that selects randomly $f$ fraction of nodes as the top $f$ fraction most vulnerable ones and that leads to the recognition rate $r(f) = f$.



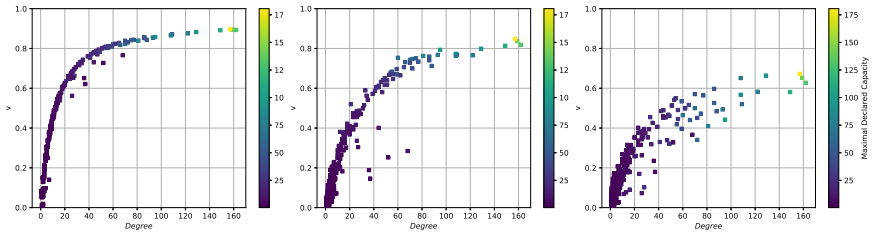

(a) Network metrics for the unweighted network.

(b) Optimal cases depicted in Table 1.

Fig. 5: Recognition rate $r(f)$ using (a) network centrality metric and (b) nodal infection probability where the model parameter $\alpha$, $\tau_o$ and network type are given in Table 1.

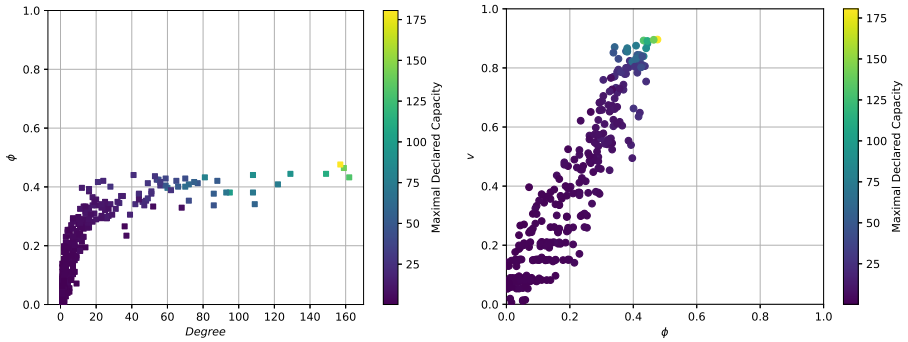
We may also use network centrality metrics that quantifies diverse nodal properties in a network [31] to identify the most vulnerable airports. Five centrality metrics in network $G_1$ are considered: degree (number of links a node possesses), clustering coefficient (the probability that two random neighbors of a node are connected), betweenness (the number of shortest paths between any pair of nodes that traverse the node), closeness (the average distance of a node to any other node in number of links) and the principal eigenvector component (the corresponding component of the node in the principal eigenvector of the adjacency matrix). A node with a high centrality metric is supposed to have a high vulnerability. The recognition rate when using each of these centrality metric is given in fig. 5a. These centrality metrics perform mostly better than the random strategy, i.e. $r(f) > f$. The degree and principal eigenvector component perform the best and similarly. Their similar performance is due to the fact that these two metrics are positively correlated in power-law networks and the $G_1$ indeed has a power-law degree distributions as shown in fig. 2. fig. 5b shows that the recognition rate using nodal infection probability can be slightly but not evidently higher than the degree (thus centrality metrics in general).

Beyond the parameter sets given in Table 1, we explore systematically the recognition rate of our model when the underlying network is either $G_2$ or $G_3$, with all possible parameter combinations where $\alpha = 0.5, 1, 1.5$ and $\tau = 0.1, 0.2, ..., 2$. The recognition rate $r(f)$ in relation to $f$ is relatively independent of the choice of the underlying network and parameters [3] and is close to the recognition rate of degree. fig. 6 shows that the nodal infection probability is positively correlated with the degree of a node, independent of the choice of the parameters. This explains why nodal infection probability achieves similar recognition rate to the degree.




(a) $G_1, \tau = 0.1$ (b) $G_2(\alpha = 0.5), \tau = 0.2$ (c) $G_3(\alpha = 1.5), \tau = 0.5$

Fig. 6: Scatter plot of the infection probability versus the degree of a node.




(a) Validation (b) $G_1$, $\alpha = 0.5$, $\tau = 0.1$

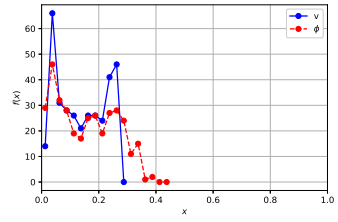
Fig. 7: Scatter plot of the vulnerability versus degree of a node in (a) and of the infection probability versus vulnerability of a node in (b).

The capacity of an airport can be represented by its maximal declared capacity, i.e. the maximal number of flights planned in an hour. As shown in fig. 7b, airports with a large capacity tends to have a large vulnerability, degree

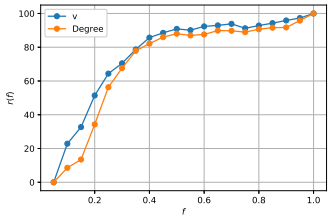
[3] except that the recognition rate drops dramatically if the infection rate is low when most nodes have an infection probability close to zero

and nodal infection probability. Furthermore, the linear relation between nodal infection probability and nodal vulnerability as shown in fig. 7b suggests that nodal infection probability may better suggest the vulnerability level of a node than the degree.

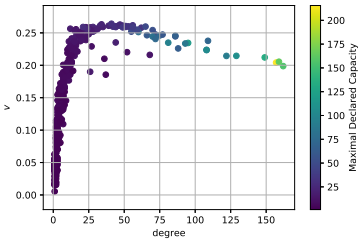
## 5 Model generalization



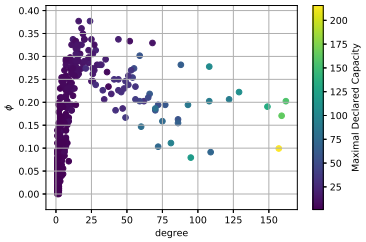

(a) $v$ and $\phi$ distributions comparison




(b) $v$ degree recognition rate




(c) $v$ degree correlation




(d) $\phi$ degree correlation

Fig. 8: Generalized SIS model with heterogeneous recovery rates to identify the generalized airport vulnerability $\phi$. The underlying network is $G_2$, $\alpha = 0.5$, $\tau = 0.02$, $\theta = 1.2$ and $c = 0.02$. (a) Probability density function of the generalized airport vulnerability and the nodal infection probability respectively;(b) Recognition rate using nodal infection probability and degree respectively; (c) Scatter plot of the infection probability of node versus its degree and (d) Scatter plot of nodal vulnerability versus its degree.

.

The declared capacity of an airport at a given hour can also be defined as e.g. the number of scheduled flight movements divided by 90%, assuming airports plan only 90% of their declared capacity to ensure their quality of service when the number of movements amounts to 10% more than what has been scheduled. In this case, the number of nodes with a high vulnerability around 0.4 is reduced (see fig. 8a). The nodal vulnerability increases first and decreases afterwards

with the degree, in contrast to the monotonic increase of the previously defined vulnerability with degree (see fig. 8d and fig. 1). This suggests that large (degree) airports may have a greater capability at dealing with operational delay due to better infrastructure, motivating us to generalize our model by setting the recovery rate positively dependent on its degree. The node strength (sum of the link weights incident to a node) in a weighted network corresponds to the degree in an unweighted network. Consider following heterogeneous recovery rate:

$$\delta_i' = \delta(c + (\frac{s_i}{s_{max}})^\theta) \tag{4}$$

where $s_i, s_{max}$ are respectively the strength of node $i$ and the maximum node strength in the network. With this generalization, nodal infection probability may also increase and afterwards decreases with the degree (see fig. 8c), allowing our model to better rank airport vulnerability than degree (see recognition rate in fig. 8b).

This example illustrates that the SIS model with heterogeneous recovery rates beyond heterogeneous infection rates could well address the fact that large airports may have a higher capability at dealing with operational delay by changing their operation through adjusting departure/arrival times of flights or assigning optimal taxiing routes. Such generalized SIS model could also better identify vulnerable airports than centrality metrics like the degree.

## 6 Conclusion

We model airport traffic congestion contagion as an SIS epidemic spreading process on an airline transportation network, aiming to identify airport's vulnerability (probability of being congested) using its infection probability derived from our model. This model is evaluated via its capability to reproduce the distribution of nodal vulnerability and to rank airports in vulnerability. Our model evidently outperforms the random ranking of airports in vulnerability but performs only slightly better than (similarly to) airport vulnerability ranking using airports' topological properties such as the degree. We propose three methods to construct the underlying weighted airline network to capture diverse information such as flights frequency and duration and let the infection rate of each link is proportional to its link weight. Still, such SIS model with heterogeneous infection rate leads to a higher infection probability for a node with a higher degree, which explains the similar performance between our model and nodal degree in airport vulnerability ranking.

The definition of airport vulnerability can be generalized to take into account various targeting service levels. We demonstrate that the vulnerability of an airport is then not necessarily positively correlated with the degree anymore. Correspondingly, we generalize our model to be heterogeneous in recovery rate. The generalized SIS model, where the infection probability of a node does not monotonically increases with nodal degree, performs evidently better than the degree in airport vulnerability ranking.

Beyond identifying vulnerable airports, which is of crucial operation relevance, our congestion contagion model allows further development of optimization strategies. Since the congestion in one airport may lead to the congestion of other airports, which ones should we invest in vulnerability reduction in order to minimize the global congestion? Such questions require further systematic validation and construction of the model to taken into account e.g. the heterogeneous recovery rate and other operational factors.

## Acknowledgement

This work is supported by Netherlands Organisation for Scientific Research NWO (TOP Grant no. 612.001.802).

## References

1. Kiss, I. Z., Miller, J. C., Simon, P. L.: Mathematics of epidemics on networks. Cham: Springer 598 (2017). doi:10.1007/978-3-319-50806-1
2. Zanin, M., Lillo, F.: Modelling the air transport with complex networks: A short review. Eur. Phys. J. Spec. Top. 215(1), 5-21 (2013). doi:10.1140/epjst/e2013-01711-9
3. Baspinar, B., Koyuncu, E.: A data-driven air transportation delay propagation model using epidemic process models. J. Aero. Eng. 2016 (2016). doi:10.1155/2016/4836260
4. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. and Vespignani, A.: Epidemic processes in complex networks. Rev. Mod. Phys. 87, 925 (2015).
5. Van Mieghem, P., Omic, J., Kooij, R.: Virus spread in networks. IEEE TNET. 17(1), 1-14 (2009). doi:10.1109/TNET.2008.925623
6. Li, D., Qin, P., Wang, H., Liu, C., Jiang, Y.: Epidemics on interconnected lattices. EPL. 105(6), 68004 (2014). doi:10.1209/0295-5075/105/68004
7. Qu, B., Wang, H.: SIS epidemic spreading with correlated heterogeneous infection rates. J. Phys. A 472, 13-24 (2017). doi:10.1016/j.physa.2016.12.077
8. Qu, B., Wang, H.: SIS epidemic spreading with heterogeneous infection rates. IEEE TNSE. 4(3), 177-186 (2017). doi:10.1109/TNSE.2017.2709786
9. Li, C., van de Bovenkamp, R., Van Mieghem, P.: Susceptible-infected-susceptible model: A comparison of N-intertwined and heterogeneous mean-field approximations. Phys. Rev. E 86(2), 026116 (2012).doi:10.1103/PhysRevE.86.026116
10. Van Mieghem, P.: The N-intertwined SIS epidemic network model. Computing 93(2-4), 147-169 (2011). doi:10.1007/s00607-011-0155-y
11. Li, C., Wang, H., Van Mieghem, P.: Epidemic threshold in directed networks. Phys. Rev. E 88(6), 062802. (2013). doi:10.1103/PhysRevE.88.062802
12. Yang, Z., Zhou, T.: Epidemic spreading in weighted networks: An edge-based mean-field solution. Phys. Rev. E 85(5), 056106 (2012). doi:10.1103/PhysRevE.85.056106
13. Lu, D., Yang, S., Zhang, J., Wang, H., Li, D.: Resilience of epidemics for SIS model on networks. Chaos Interdiscip. J. Nonlinear Sci. 27(8), 083105 (2017). doi:10.1063/1.4997177
14. Qu, B., Li, C., Van Mieghem, P., Wang, H.: Ranking of nodal infection probability in susceptible-infected-susceptible epidemic. Scientific reports 7(1), 9233 (2017). doi:10.1038/s41598-017-08611-9

15. Barat, A., Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A.: The archiecture of complex network weights. Proceedings of the National Academy of Sciences of the United States of America 101(11), 3747-3752 (2014). doi:10.1073/pnas.0400087101
16. Guimerà, R., Mossa, S., Turtschi, A., Amaral, L.: The worldwide air transportation network - anomalous centrality, community structure, and cities' global roles. Proceedings of the National Academy of Sciences of the United States of America 102(22), 7794–7799 (2005). doi:10.1073/pnas.0407994102
17. Reggiani, A., Signoretti, S., Nijkamp, P., Cento, A.: Network measures in civil air transport - a case study of Lufthansa. Lecture Notes in Economics and Mathematical Systems 613, 257–282 (2009). doi:10.1007/978-3-540-68409-1-14.
18. Han, D. D., Qian, J. H., Liu, J. G.: Network topology and correlation features affiliated with European airline companies. Phys. A 388(1), 71–81 (2009). doi:10.1016/j.physa.2008.09.021
19. Fleurquin, P., Ramasco, J. J., Eguiluz, V. M.: Systemic delay propagation in the US airport network. Scientific reports 3, 1159 (2013). doi:10.1038/srep01159
20. Ciruelos, C., Arranz, A., Etxebarria, I., Peces, S.: Modelling Delay Propagation Trees for Scheduled Flights. USA/Europe Air Traffic Management Research and Development Seminar 11, (2015).
21. Baspinar, B., Koyuncu, E.: A Data-Driven Air Transportation Delay Propagation Model Using Epidemic Process Models. International Journal of Aerospace Engineering (2016). doi:10.1155/2016/4836260
22. Chi, L. P., Cai, X.: Structural changes caused by error and attack tolerance in US airport network. Int. J. Mod. Phys. B, 18, 2394-2400 (2004). doi:10.1142/S0217979204025427
23. Wilkinson, S.M., Dunn, S., Ma, S.: The vulnerability of the European air traffic network to spatial hazards. Natural hazards 60(3), 1027-1036 (2012). doi:10.1007/s11069-011-9885-6
24. Lacasa, L., Cea, M., Zanin, M.: Jamming transition in air transportation networks. J. Phys. A 388(18), 3948-3954 (2009). doi:10.1016/j.physa.2009.06.005
25. United States Bureau of Transportation Statistics. http://www.transtats.bts.gov
26. Wang, H., Li, Q., D'Agostino, G., Havlin, S., Stanley, H. E. and Van Mieghem, P.: Effect of the interconnected network structure on the epidemic threshold. Phys. Rev. E. 88, 022801 (2013).
27. Gang, Y., Tao, Z., Jie, W., Zhong-Qian, F., & Bing-Hong, W.: Epidemic spread in weighted scale-free networks. Chinese Physics Letters, 22(2), 510 (2005). doi:10.1088/0256-307x/22/2/068
28. Odoni, A., De Neufville, R.: Airport systems: Planning, design, and management. McGraw-Hill Professional (2003).
29. Dunn, S., Wilkinson, S. M.: Increasing the resilience of air traffic networks using a network graph theory approach. J. Trans. Res. E 90, 39-50 (2016). doi:10.1016/j.tre.2015.09.011
30. Ciruelos, C., Arranz, A., Etxebarria, I., Peces, S., Campanelli, B., Fleurquin, P., Ramasco, J.J.: Modelling delay propagation trees for scheduled flights. In: Proceedings of the 11th USA/EUROPE Air Traffic Management R&D Seminar, pp. 23-26. Lisbon, Portugal (2015).
31. Li, C., Li, Q., Van Mieghem, P, Stanley, H. E. and Wang, H.: Correlation between centrality metrics and their application to the opinion model. The European Physical Journal B, 88:65 (2015).
32. Fleurquin, P., Ramasco, J.J., Eguiluz, V.M.: Systemic delay propagation in the US airport network. Scientific reports 3, 1159 (2013). doi:10.1038/srep01159