

Research Article

A Data-Driven Air Transportation Delay Propagation Model Using Epidemic Process Models

B. Baspinar¹ and E. Koyuncu²

¹*Controls and Avionics Research Group, Aerospace Research Center, Istanbul Technical University, Istanbul, Turkey*

²*Department of Aeronautical Engineering, Istanbul Technical University, Istanbul, Turkey*

Correspondence should be addressed to B. Baspinar; baspinarb@itu.edu.tr

Received 12 April 2016; Revised 8 July 2016; Accepted 26 July 2016

Academic Editor: Mahmut Reyhanoglu

Copyright © 2016 B. Baspinar and E. Koyuncu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In air transport network management, in addition to defining the performance behavior of the system's components, identification of their interaction dynamics is a delicate issue in both strategic and tactical decision-making process so as to decide which elements of the system are "controlled" and how. This paper introduces a novel delay propagation model utilizing epidemic spreading process, which enables the definition of novel performance indicators and interaction rates of the elements of the air transportation network. In order to understand the behavior of the delay propagation over the network at different levels, we have constructed two different data-driven epidemic models approximating the dynamics of the system: (a) flight-based epidemic model and (b) airport-based epidemic model. The flight-based epidemic model utilizing SIS epidemic model focuses on the individual flights where each flight can be in susceptible or infected states. The airport-centric epidemic model, in addition to the flight-to-flight interactions, allows us to define the collective behavior of the airports, which are modeled as metapopulations. In network model construction, we have utilized historical flight-track data of Europe and performed analysis for certain days involving certain disturbances. Through this effort, we have validated the proposed delay propagation models under disruptive events.

1. Introduction

The air transportation industry and its role in modern life are rapidly growing. It is expected that the number of commercial flights will almost double from 26 million to 48.7 million and 13.5 trillion passenger-kilometers will be flown by 2030, which is almost the triple of what is flown by airlines today (see [1]). However, the airspaces have a fixed amount of capacity, and the number of airports to be built is not sufficient to accommodate such increase in the demand. Therefore, the Air Traffic Management (ATM) system must go under an operational transformation in order to increase its efficiency to deal with these challenges.

New procedures and concepts that are being developed in SESAR and NextGen are leading to a global paradigm shift from air traffic "control" to efficient air traffic "management," which requires redesigning the ATM system and defining the performance indicators of the elements of the ATM network. In addition to defining the performance of the components of

ATM system, identification of the interactions between them is also important during tactical decision making so as to decide which elements of the system should be "controlled" or "intervened." For example, deciding how much capacity reduction on the airport should be applied under severe weather condition or which flight should be directed without disturbing entire traffic network is not a trivial issue. By considering the progressively increasing complexity in the airspace structures and increasing workloads of the operators on the ground, a decision support system should answer such question to the human operators so as not to compromise the operational efficiency even under stressful events.

In managing the modeling of the traffic network and managing the traffic flow, several researchers focused on queuing network modeling for the propagation of local delays in the air traffic network. The MITRE Corporation has developed two different National Airspace System (NAS) simulation models for simulation of delay propagation on the nationwide airport and airspace network in the United States.

The first one is the National Airspace System Performance Analysis Capability (NASPAC) [2] and the second one is the Detailed Policy Assessment Tool (DPAT), which is the successor of the NASPAC [3]. When the capacity of an airport is reduced due to external events, DPAT is able to propagate delays across the network, but it does not utilize the information regarding aircraft itineraries, which might lead to unreliable predictions. On the other hand, there are also agent-based simulation models for delay propagation, such as The Future ATM Concepts Evaluation Tool (FACET) [4]. LMINET [5] and LMINET2 [6] are national queuing network models that model the airports as $M(t)/E_k(t)/1$ queues. While LMINET does not use aircraft itineraries, LMINET2 utilizes this information. The Approximate Network Delays (AND) model is another popular model [7, 8]. The modeling approaches in AND model and LMINET2 are similar. However, calculating strategies of the local queuing delays are different. As an alternative to these strategies, epidemic spreading models can be used to simulate the delay propagation in air transportation network.

In a methodological manner, epidemic spreading and delay propagation resemble each other. In epidemic spreading, some kind of disease, such as flu, influenza, ebola, and HIV/AIDS, is propagated between living beings by the mobility of individuals. The delay spreading in air transportation is realized in a similar vein. A local disturbance occurs in a specific airport; this disturbance causes an infection, and this infection spreads along network with the help of flights. Because of this similarity, dynamical models that are used to simulate the epidemic spreading process can be used to simulate the delay spreading in air transportation.

In the epidemic model literature, the basic compartment model was first introduced by Bernoulli [9] in the 18th century. In 1927, Kermack and McKendrick created the SIR model [10]. The compartments in this model are susceptible (S), infected (I), and recovered (or immune) (R). This is a good and simple model for many infectious diseases including measles, mumps, and rubella. It is possible to immunize people against these diseases. Another model was SIS model [11, 12] that can be derived from the SIR model under the assumption that the individuals have no immunity to the disease. These two models were also represented with vital dynamics [13] that have death and birth situations. Moreover, the SIR model [14] was developed that has temporal immunity. The SIR model was also presented with carriers [14] that continue to carry the infection. In these models, individuals are able to infect others immediately when they are infected. However, some diseases have latent or exposed times before the infectious state. To simulate these kinds of diseases, E (exposed) compartment was presented and some models are constructed using this compartment as SEIS, SEIR, and SEIRS models [14]. And M compartment was also introduced in some models, such as MSIR and MSEIR models [15], to simulate the disease where an individual is born with a passive immunity from its mother. Furthermore, the network theory [16–18] was also applied to these models to simulate the situations when infection and recovery rates are heterogeneous. Moreover, metapopulation approach [18] was also represented to lump the individuals as subpopulations. In

some researches [19], air transportation was used to define the mobility between populations and spread the disease among them. However, the methodology of epidemic models was not applied to air transportation to construct a delay spreading model that is the topic of this study.

This paper focuses on building a conceptual model for the air transportation network under stress and defining parameters that are specific to the elements of the system, that is, airports, air sectors, OD pairs, and flights. In order to understand the dynamics of delay propagation in air transportation network, we have utilized epidemic spreading processes by assuming the characteristics of disease spreading and delay propagation are similar. We have introduced two different epidemic models approximating the air transportation network and enabling the capture of different aspects of the system: (a) flight-based epidemic model, which is a SIS epidemic model that concentrates on individual flights where each flight can be in susceptible or infected states, and (b) airport-based epidemic model, which is a metapopulation model, which allows us to define the collective behavior of the airports. In our data analysis, we have utilized real flight-track data of Europe and selected certain days affected by certain disturbances. The physical parameters of the network are transformed into the parameter set of the introduced epidemic models, which allow us to understand interaction rates and recovery rates of the elements of the system under similar disruptive events.

The rest of the paper is organized as follows. Section 2 explains the basics of epidemic spreading process. Section 3 presents our data-driven delay propagation based on epidemic spreading process. Finally, discussion on the results of the air transport network analysis for the sample days involving certain disturbances is given in Section 4 before concluding remarks.

2. Epidemic Process Model

Epidemics models are based on the population dynamics, where the individuals in a population are divided into the states or “compartments.” The simplest epidemic models include at least two compartments, which are susceptible (S) and infected (I). S represents individuals who are healthy but susceptible to becoming infected, and I represents the individuals who are infected but able to recover. Specifically, the infected individuals can transfer from I to S with some recovery rate δ , and healthy individuals can transfer into I with some infection rate β , where the dynamics of the population (or network) induce this behavior. The name of the epidemic depends on the configuration of these compartments in the model. In order to capture more features in the population, depending on their behaviors specific to their natures, some models also include additional compartments such as R (removed or recovered), E (exposed), and Q (quarantined). For instance, SIS represents the susceptible-infected-susceptible models or SIR represents the susceptible-infected-recovered models, where the individual follows the given cyclic paths indicated in their names. SIS and SIR models are the most common models, which are depicted in Figure 1, and others are typically based on these

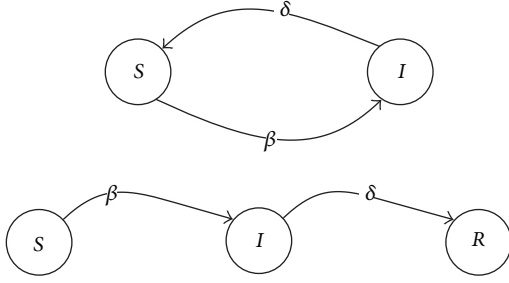


FIGURE 1: Epidemic processes with S susceptible, I infected, and R removed (or recovered) states and β infection and δ recovery rates based on SIS and SIR models.

basic models. Note that the difference between SIS and SIR models is about immunity, where the infected individuals after recovery become immune permanently to disease in SIR model, while recovered individuals go into the susceptible state S repeatedly in SIS model. Considering the delay issue in air transportation, permanent recovery is not possible. Therefore, we focused on SIS based models throughout this study.

Let $N \in \mathbb{R}$ be the total number of individuals in population. Let $N_i(t) \in \{0, 1, \dots, N\}$ be the number of infected individuals at some time t . The evaluation of states can be described through Markov process. If the population is assumed as a well-mixed population, that is, all individuals affect and are affected by all other individuals equally, then infection rates and recovery rates of all individuals in the population become β and δ for all individuals, respectively. Thus, the dynamics of SIS model can be written as [11, 12]

$$\begin{aligned} \dot{p}_i &= \beta p_i p_s - \delta p_i, \\ \dot{p}_s &= -\beta p_i p_s + \delta p_i, \end{aligned} \quad (1)$$

where $p_i = N_i/N$ and $p_s = (N - N_i)/N$ are the fraction of infected and susceptible individuals. By holding $p_s = 1 - p_i$, these equations can be translated into the following form:

$$\dot{p}_i = \beta p_i (1 - p_i) - \delta p_i. \quad (2)$$

Equation (2) is a deterministic equation, which is derived by assuming N is very large and can be analytically solved. Considering the solution of (2), for $\beta > \delta$, the disease never dies out, which is called “endemic equilibrium.” Moreover, Weiss and Dishon [20] proved that if $\beta > \delta$ and $N \rightarrow \infty$, expected time $E[\tau]$ to reach disease-free equilibrium grows exponentially with N . However, Ball [21] showed that the behavior of the population can be different while N is not large enough.

Such population models involve many simplifications that may limit building a proper model of behavioral dynamics. For instance, homogenous β infection rate and δ recovery rate assume an individual equally affects and is equally affected by others. Moreover, classification into a low number of compartments of the entire population might not capture the full behavioral dynamics. The more realistic approach could be a model taking into account all the states of all

individuals in the population separately and allowing for arbitrary interactions among them, which leads to network-based approaches.

2.1. Network Model. Network-based approach is more suitable to model the air transportation. Air transportation network system can be described as continuous-time Markov process with constant transition rates between the states on a graph. Let us consider a directed graph $G = (V, E)$, where V represents the set of nodes on G or individuals in the population that correspond to flights and E is the set of edges or connectivity between individuals. If node i is directly affected by node j , then $\beta_{ij} > 0$; otherwise $\beta_{ij} = 0$. Each node i has its own recovery rate δ_i as well. By accounting the SIS model, dynamics of epidemic process is given in a form of the following differential equation [16–18]:

$$\dot{p}_i = -\delta_i p_i + \sum_{j=1}^N \beta_{ij} p_j (1 - p_i), \quad (3)$$

where $p_i \in [0, 1]$ gives the probability of an individual i being infected in any time t .

2.2. Metapopulation Model. Considering physical network system, for example, air transportation networks, instead of the modeling of each state of N individuals in the system, M subpopulations can be modeled such that $M \ll N$. This approximation allows us to reduce the number of nodes in the model. In this situation, a node i corresponds to a subpopulation (e.g., airport) instead of individuals (e.g., flights). Such metapopulation models are defined assuming each subpopulation $i \in \{1, 2, \dots, M\}$ is well mixed and has a homogenous constant recovery rate δ_i . The similar set of differential equation, which is seen in (3), can describe the dynamics of epidemic spreading in metapopulation [18]. Let $p = (p_1, p_2, \dots, p_M)^T$ denote the state vector of the system, let $\mathfrak{D} = \text{diag}(\delta_1, \delta_2, \dots, \delta_M)$ denote the diagonal matrix of recovery rates, and let $\mathfrak{B} = [\beta_{ij}]$ denote the matrix of infection rates. Then (3) becomes

$$\dot{p} = (\mathfrak{B} - \mathfrak{D}) p - [(\mathfrak{B}p) \circ p], \quad (4)$$

where \circ denotes the Hadamard product or element-wise product implying element by element multiplication. This time, $p_i \in [0, 1]$ is the fraction of i th subpopulation that is infected.

3. Delay Propagation Models

Delay propagation over air traffic network has stochastic nature due to multiple decision makers within the system and has similarities with the mechanism of epidemic spreading. Considering the elements of the air transportation system, for example, airports, airspaces, flight routes, and so forth, an air traffic network can be modeled as a directed graph with nodes and edges similar to epidemics network model. In this work, we have constructed two different models through airport-based and flight-based approaches. The flight-based

approach, which is a SIS model, uses individual flight as the nodes in the transport network, where each flight can be in susceptible or infected states, in other words, the candidate to be delayed due to certain disturbances or already delayed. In this model, being infected is defined as to be delayed greater than 15 minutes. The second approach focuses on airports, which are considered as the subpopulations of the network system. This model, which is a metapopulation model, allows us to define the collective behavior of the airports. In order to define parameters of the network, we have utilized historical flight-track data of Europe and performed analysis for certain days involving certain disturbances.

Algorithm 1 is used to construct both airport-based or flight-based epidemic models. Estimation process for the parameters of the network, which are infection rates and recovery rates, is done at the end of the algorithm. Flight information dataset \mathfrak{F} allows us to estimate infection rates through data-driven statistical analysis. The flight information has two components: \mathfrak{F}_P flight plan information and \mathfrak{F}_A flight-track data (actual flown data), that is, $\mathfrak{F} : \{\mathfrak{F}_P, \mathfrak{F}_A\}$. Then, state vectors at time t and $t + 1$ are calculated. Then, infection rates are estimated using \mathfrak{B} , $p(t)$, and $p(t + 1)$ by the help of differential equations of epidemic spreading process (3) and (4).

Algorithm 1 (modeling: data-driven parameter estimation).

```

input:  $\mathfrak{F}_P$  and  $\mathfrak{F}_A$ 
output:  $\mathfrak{B} = [\beta_{ij}]$ ,  $\mathfrak{D} = \text{diag}(\delta_1, \delta_2, \dots, \delta_M)$ 
Generate  $\mathfrak{B}$  through  $\mathfrak{F}_A$ ;
foreach  $i \in \{1, 2, \dots, N\}$  or  $\{1, 2, \dots, M\}$  do
    Calculate  $p_i(t) = I_{\text{dep}}^i(t) + I_{\text{arr}}^i(t)/N^i(t)$ ;
foreach  $i \in \{1, 2, \dots, N\}$  or  $\{1, 2, \dots, M\}$  do
    Calculate  $p_i(t + 1)$ ;
Estimate  $\mathfrak{D}$  (via (3) or (4));

```

Once the recovery rates are estimated for a given time period involving specific events in the network, such as severe weather effects, unusual industrial activities, terror attacks at airports, and air traffic controller strikes, we would like to show whether these rates can be used to simulate for another day under stress due to similar kinds of events through the epidemic process. Algorithm 2 is used to simulate another time period through flight plan data \mathfrak{F}_P by utilizing $\widehat{\mathfrak{D}}$ estimated recovery rates. The following subsections give the details of the algorithms.

Algorithm 2 (simulation: state prediction).

```

input:  $\mathfrak{F}_P, \widehat{\mathfrak{D}}$ 
output:  $\mathfrak{B} = [\beta_{ij}]$ ,  $p(t + 1)$ 
Generate  $\mathfrak{B}$  via  $\mathfrak{F}_P$ ;
foreach  $i \in \{1, 2, \dots, N\}$  or  $\{1, 2, \dots, M\}$  do
    Calculate  $p_i(t) = I_{\text{dep}}^i(t) + I_{\text{arr}}^i(t)/N^i(t)$ ;

```

foreach $i \in \{1, 2, \dots, N\}$ or $\{1, 2, \dots, M\}$ **do**

Estimate $p_i(t + 1)$ (via (3) or (4));

3.1. Airport-Based Metapopulation Model. In order to construct an airport-based network model, traffic flow over Europe is analyzed through the real flight data. Note that by considering the license agreement with our data provider, in our analysis, we did not span long time intervals. Therefore, it is not logical to come to a general conclusion by the comparative results that are given throughout this paper.

Through the traffic flow data, it is observed that approximately 20% of all daily movements in Europe are originated from or come to 8 busiest airports in Europe (in 2015). Moreover, in Europe, most of the airports' hourly movements are less than 4. These kinds of airports operating far from their capacity limits do not induce an additional delay in the network. By accounting this reason, while constructing the model, such minor airports were taken as a single aggregated airport. The network model focused on the European region and the flights from/to non-European airports are regarded as an incoming/outcoming flow. After this simplification, in the network model, the total number of airports in European network has been reduced to 103 airports including 102 major European airports and a single aggregated airport. A detailed study on reduction strategy and airport-centric analysis can be seen in Baspinar et al.'s paper [22], which focuses on the construction of stochastic queuing network model for European air traffic flow.

In airport-based epidemic model, 103 individuals (i.e., airports) exist, and the β_{ij} infection rates between the airports i and j are determined with their flow rates. For any airport pair m and n that has no flow between them, the infection rate is $\beta_{mn} = 0$. So, it is obvious that $\beta_{mm} = 0$ and $\beta_{nn} = 0$. Specifically, the infection rate $0 \leq \beta_{ij} < 1$ from an airport j to another i is represented through normalized directed traffic flow for the given time window. An airport has inflow from different airports; normalized inflow for an airport denotes the infection rates of this airport. For example, let airport E have 2, 5, and 3 flights from airports F , G , and H , respectively. In this situation, $\beta_{EF} = 0.2$, $\beta_{EG} = 0.5$, and $\beta_{EH} = 0.3$. Infection rates for an airport correspond to normalized inflow to this airport. Figure 2 depicts the infection rates between the 10 busiest airports in Europe (i.e., evaluated through the traffic from 13:30 to 16:30 on June 02, 2015).

Time window selection is a delicate issue in transportation networks as any interaction between the subpopulation emerges after an arbitrary delay time; in other words, the network involves time delays in their state transitions. Therefore, the flight times are critical parameters to capture infection rate transitions while solving the differential equations. Considering this issue, the distribution of flight times in Europe for a given day is analyzed and shown in Figure 3. It is observed that 79% of flights have flight times less than 2.5 hours and 89% of flights have flight times less than 3.5 hours, and the average of flight time is 116 minutes. We have chosen to use 2-hour delay spreading periods to enable us to see the complete transition on the dynamics of the network.

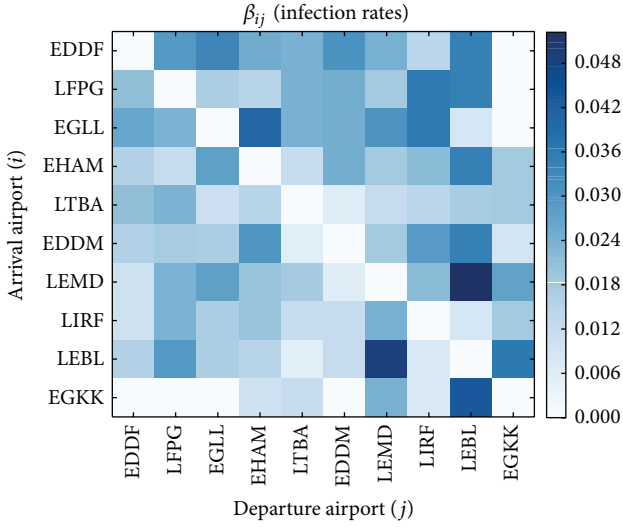


FIGURE 2: Infection rates β_{ij} between 10 busiest airports in Europe calculated through the movements between 13:30 and 16:30 on June 02, 2015.

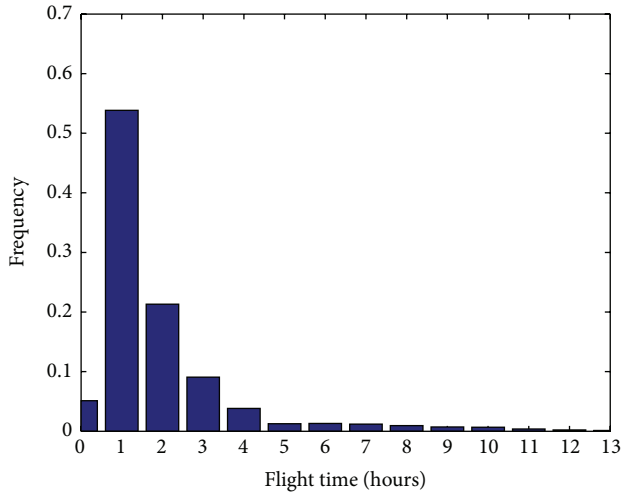
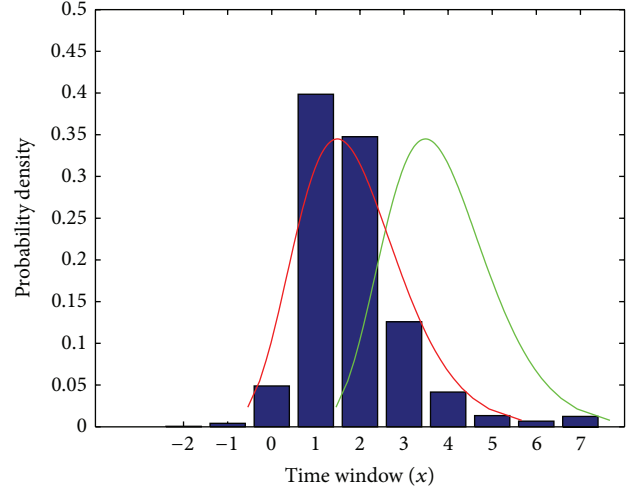


FIGURE 3: Distribution of flight times for the flights circulating in European traffic network in 2015.

This means that $p(t + 1)$ is the fraction of subpopulations after 2 hours from $p(t)$ and can be approximated through the updated β_{ij} infection rates and evaluated δ_i recovery rates by solving differential equations (3) or (4).

Once we have $p(t)$ probability vector of fractions of subpopulations at time t , we can evaluate $p(t + 1)$ through the transition equation (4) with \mathfrak{B} and \mathfrak{D} matrices. As the \mathfrak{B} parameter set is derived through the data, which are basically normalized traffic flow between the airports, the only unknown parameter set is \mathfrak{D} recovery rates specific to airports. For a specific airport i , $p(t)$ can be defined as

$$p_i(t) = \frac{I_{\text{dep}}^i(t) + I_{\text{arr}}^i(t)}{N^i(t)}, \quad (5)$$



■ Data
 — $g_1(x) = g(x | 10.25, 0.37)$
 — $g_2(x) = g_1(x + 2)$

FIGURE 4: Distribution of arrival time windows according to estimated time of arrivals, which can be approximated by $g_2(x) \sim f_{\text{ETA}}$ gamma distribution.

where $I_{\text{dep}}^i(t)$, $I_{\text{arr}}^i(t)$, and $N^i(t)$ denote the number of delayed flights at departure and arrival stages and the total number of flights in the airport i at time t , respectively.

The following subsection explains flight-based epidemic model and the recovery rate estimation, which are set upon airports.

3.2. Flight-Based Epidemic Model. In the flight-based epidemic model, this time, flights are represented as the individuals in a population. In this approach, the infection rates $[\beta_{ij}]$ are the composition of its two components, which are $[\beta_{ij}^1]$ associated with f_{ETA} probabilistic distribution on the time of arrivals and $[\beta_{ij}^2]$ associated with ground-waiting times between consecutive flights. Then, $[\beta_{ij}]$ is given as their linear combinations as follows:

$$[\beta_{ij}] = w_1 [\beta_{ij}^1] + w_2 [\beta_{ij}^2], \quad (6)$$

where $w_1, w_2 \in [0, 1]$ and $w_1 + w_2 = 1$. In order to provide general approximation, we have chosen to use $w_1, w_2 = 0.5$ in our simulations.

The model is based on the idea that the delayed flights (i.e., infected) would affect the flights at the departure or arrival phase in the destination airports. Considering the arrival slot, which is typically $\Delta t = 15$ min, arrival delay profile of each airport was determined by comparing the estimated time of arrivals and the actual time of arrivals of the flights. A typical arrival delay profile of an airport (the name of the airport is intentionally masked) is shown in Figure 4. As seen in the figure, the distribution can be represented by $g_2(x) \sim f_{\text{ETA}}$ gamma distribution. Note that an infection rate of a flight to another depends on this distribution; in

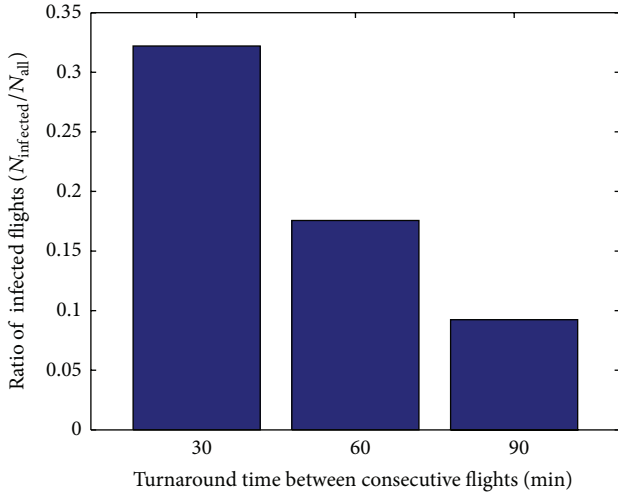


FIGURE 5: Ratio of being infected according to ground-waiting time.

other words, a delayed (infected) flight affects the flights mostly on their actual arrival slots and this effect naturally diminishes as time goes to infinity. The gamma function on delay profile for the airport constructs $[\beta_{ij}^1]$ infection relations. The infection rates β_{ij}^1 subject to sequential time slots are set to $g_2(t+k)$, where we have assumed that $k \in \{0, 1, \dots, 5\}$ by accounting the general behavior of the system observed in our analyses (seen in Figure 4). Once we have constructed the first infection relation matrix $[\beta_{ij}^1]$, we have normalized it, which is $\sum_{i=1}^N \beta_{ij}^1 = 1$. For example, suppose that flight₁ will arrive to airport D in time window t , and flight₅, flight₆, flight₇, and flight₈ are at the departure or arrival phase in the destination airport D at t , $t+1$, $t+1$, and $t+2$, respectively. Suppose that there is no other flight at airport D from t to $t+5$. Then, the infection rates due to flight₁ are estimated through $g_2(t)$ function. By considering flight₁ and flight₅, whose movements are in the same time window, through $g_2(t)$ at $t=0$ (see Figure 4), the infection rate β_{51}^1 is chosen as equal to 0.1. Other infection rates due to flight₁ are chosen as $\beta_{61}^1 = 0.3$, $\beta_{71}^1 = 0.3$, and $\beta_{81}^1 = 0.32$ utilizing $g_2(t)$ at $t=1$ and $t=2$. Then, all infection rates due to flight₁ are normalized, and real infection rates are calculated as $\beta_{51}^1 = 0.098$, $\beta_{61}^1 = 0.294$, $\beta_{71}^1 = 0.294$, and $\beta_{81}^1 = 0.314$.

In order to define the second infection rate matrix $[\beta_{ij}^2]$, time for aircraft ground-waiting τ_G , which is the time between the consecutive flights, is examined. If a flight has an infection (i.e., delayed at least +15 minutes) and the time between terminated flight and the next flight for the same aircraft is smaller than τ_{TAT} nominal turnaround time of the airport, we expect that the consecutive flight will also be infected. In the model estimation phase, the aircraft ground-waiting times τ_G between the actual arrival time and estimated departure time for the next leg are obtained through the historical flight information $\mathfrak{F} : \{\mathfrak{F}_p, \mathfrak{F}_A\}$, and the relation between the rates of being delayed according to their ground-waiting is shown in Figure 5 for a typical airport. This analysis allows us to define the impact of ground-waiting

times on being delayed. It is evident that the rate of being infected will be smaller while the ground-waiting times is larger as both the inherited delay and the airport induced delay will be easily absorbed. The rate of being infected can be naturally defined as a fraction of N_{infected} number of infected flights over N_{all} number of all flights within the given time interval. We have observed through our data-driven analysis if the ground-waiting time is greater than 115 minutes, the infection rate goes to zero. Using these rates, $[\beta_{ij}^2]$ matrix is generated. For example, suppose that an aircraft has two different flights and flight₁ is predecessor of flight₈. In this situation, flight₁ can infect flight₈, and the infection rate β_{81}^2 can be defined through ground-waiting time. Supposing the ground-waiting time is slightly less than 30 min, then β_{81}^2 would be estimated as 0.32 as shown in Figure 5.

The actual infection rate set $[\beta_{ij}]$ for flight-based epidemic model then can be given as $[\beta_{ij}] = w_1[\beta_{ij}^1] + w_2[\beta_{ij}^2]$. Note that the recovery rates for both models are yet unknown. The following subsection explains how to obtain recovery rates through data-driven analysis.

3.3. Recovery Rate Estimation for Epidemic Models. Although the recovery rate is time and day dependent, it provides a good indicator for the behavior of the system under disturbances and allows us to characterise “weakness” of the network. This enables us to identify which elements of the system should be removed/intervened under stress. In our airport-based metapopulation model, recovery rate emerges as a performance indicator for the airports, whereas it is about OD (origin/destination) pairs in our flight-based epidemic model. In order to evaluate the real values of recovery rates, the set of differential equations (3) or (4) is solved through Euler Method approximation by holding estimated $p(t)$, $p(t+1)$, and \mathfrak{B} .

In Euler Method, for the approximate solution, let the first order differential equation be $\dot{y}(t) = f(y(t), t)$ and let step size be h , where $t_n = t_0 + nh$, and let one step of Euler Method from t_n to $t_{n+1} = t_n + h$ be $y_{n+1} = y_n + hf(y_n, t_n)$, where the value of $y_n = y(t_n)$ is an approximation of the solution to the ODE at time t_n . By knowing y_n and y_{n+1} , $f(y_n, t_n)$ can be solved; hence, recovery rate matrix \mathfrak{D} can be evaluated by solving (3) or (4).

In order to understand the behavior of the system under the stress and the dynamics of delay spreading process, certain disturbing events are selected from the flight data for the analysis. The following tactical ATC headline is taken from the Network Operation Portal (NOP) of EUROCONTROL for June 2, 2015, mentioning the potential problems and delay causes within the European network.

Tactical Update at 12:35 UTC for 02/06/2015 at 11:25

Airports

EGLL (Heathrow) arrivals

Regulated due to Wx strong winds.
Moderate to high delays.

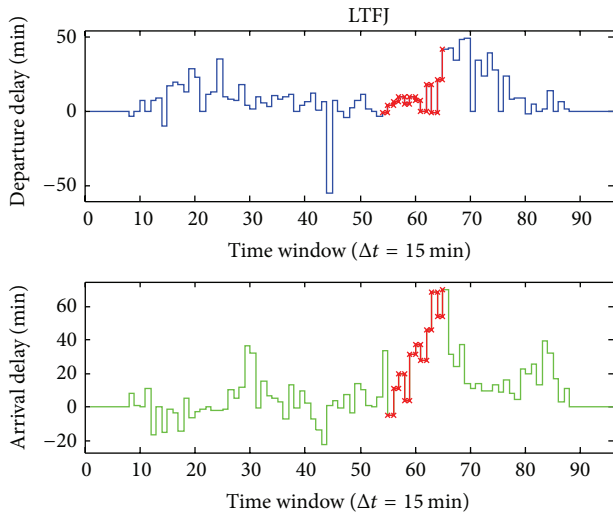


FIGURE 6: Actual delay profile of LTFJ on June 02, 2015.

LTFJ (Istanbul/S. Gokcen) arrivals

Regulated due to aerodrome capacity.
Moderate to high delays expected after
15:20 UTC.

We have selected all flights from 13:30 to 16:30 on June 2, 2015, to estimate the recovery rates over the airports and flight routes. In airport-based epidemic model, $p(t)$ is extracted through the flights from 13:30 to 14:30, and $p(t+1)$ is extracted through the flights from 15:30 to 16:30 considering 2-hour nominal spreading time. Infection rates are derived from the normalized flow rates between the airports. By utilizing Euler Method, real recovery rates for the interested period are found and given in Figure 7. It must be noted that the intention is not to provide a performance comparison between the airports; this is the reason why we have chosen a small time interval in our analysis so as not to publish true performance indicators for the airports. One can provide actual indicators for the airports by doing long term calculations.

As shown in Figure 6, some recovery rates are smaller than 0, where the smallest one is for LTFJ. This is the special case of the epidemic process models. The mean delay profile of LTFJ evaluated through the track data is shown in Figure 6 for a given short time interval (i.e., red line of the plot). Note that the ATC headline of NOP was declaring to expect high delays after 15:30 in LTFJ due to aerodrome capacity. As the capacity problem of LTFJ emerges during the analysis time, $p_{LTFJ}(t+1)$ would be significantly higher than $p_{LTFJ}(t)$. Because of this dramatic increase, the recovery rate of LTFJ is smaller than 0, which means that the airport itself induces delay. However, for example, the recovery rate of EGLL is slightly greater than 0 (i.e., $\delta_{EGLL} = 0.054$) as seen in Figure 7. It is shown in Figure 8 that EGLL has huge delays as well; however, delay trends of the airport remain almost the same within the analysis time. Hence, it can be said that EGLL has “zero recovery” but does not generate additional delay for the given time interval.

Similarly, for flight-based epidemic model, we have evaluated $p(t)$ and $p(t+1)$ for the flights from 13:30 to 16:30 on June 2, 2015, to estimate the recovery rates. The $p(t)$ probability of being infected is constructed upon departure, and $p(t+1)$ are constructed upon departure if the flight is terminated within the time interval. As we have already evaluated β_{ij} infection rates, which were driven through the data, by solving (4), δ_i recovery rates for each individual are found. Similar to airport-based epidemic model, negative recovery rates are seen as the specific case of epidemic spreading models. For a given analysis time interval, for example, 482 of 8581 flights exhibit negative $\delta_i p_i(t)$, and 80 of them arrive to EGLL suffering from the strong wind at that time. One can say that negative $\delta_i p_i(t)$ for a single flight indicates the problems at en route airspaces or the destination airport.

4. Delay Propagation Simulations through the Epidemic Models

In this section, we have discussed the results of the models and compared them with the actual delay profile of the European network. We have estimated the delay propagation trend of Europe for the disturbance on August 25, 2015, through the evaluated recovery rates. Moreover, recovery characteristics of European airports are then given by analyzing two weeks of traffic data to understand the sensitivity of the nominal recovery characteristics to the instantaneous recovery characteristics under disturbances.

Note that being infected in the epidemiological approach is an internal state of the individuals in the population; however, environmental effects or specific events can be simulated. As these kinds of events directly affect \mathfrak{B} and \mathfrak{D} , they are event dependent. Therefore, estimated event specific \mathfrak{B} and \mathfrak{D} parameters can be used to simulate the days under the effect of similar events.

The recovery rates for the airport-based epidemic model were given in Figure 7, which are evaluated through the traffic data from 13:30 to 14:30 on June 02, 2015.

The infection rates ($p_i(t)$) for 10 busiest airports in Europe, which are evaluated through the airport-based epidemic model, are shown in Figure 9(a) and the real infection rates are given in Figure 9(b). By comparing the results of the model with the real trends, it is seen that the model provides a reliable approximation on the delay spreading. The noticeable difference is seen over the time window transitions due to the constant recovery rates, which are defined for 2-hour period for each airport. In reality, this rate, of course, slightly changes over time but its differentiation remains bounded for each airport. For example, the fraction of infection at LEBL has a continuous decrease in the approximate model, whereas it has an increasing trend changing into the sharp fall in the real situation; however, fractions converge to each other at the end. The aim to construct this model is to predict the fraction of infection rate at time $t+1$ with proper accuracy, and it must be noted that selecting 2-hour time window is a requirement providing enough time to understand the dynamics of delay spreading over a network.

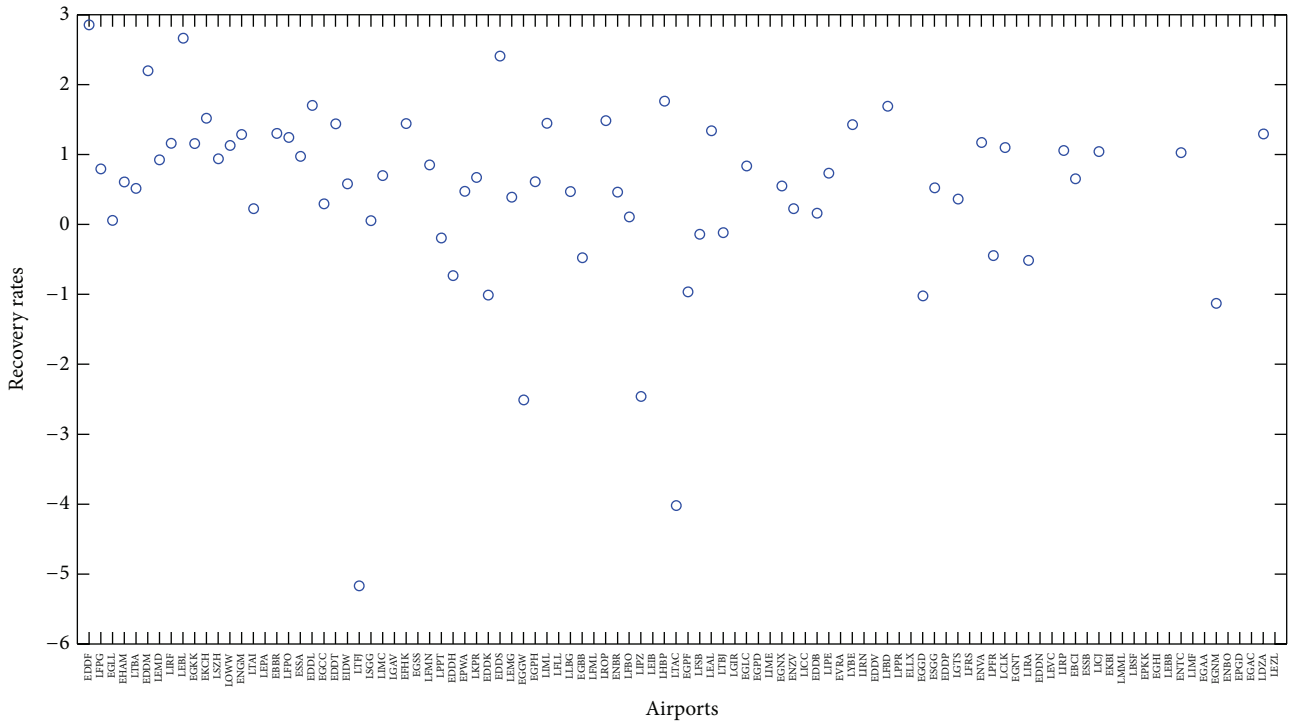


FIGURE 7: Recovery rates for airport-based metapopulation model (from 13:30 to 16:30 on June 02, 2015).

4.1. Delay Spreading Estimation with Recovery Rates. By holding the idea that similar disturbances affect the system in the same fashion, delay propagation is simulated for a day through the estimated recovery rates from the previous analysis. Specifically, the recovery rates are calculated through Algorithm 1 for a day that has similar disturbances. By utilizing these recovery rates, delay propagation is simulated through Algorithm 2 for current time.

Tactical Update for 25/08/2015 at 21:23 UTC

Airports

LTFJ (Istanbul/S. Gokcen) arrivals

Regulated due to aerodrome capacity.
Moderate to very high delays.

As mentioned before, the aim to construct model is to predict the fraction of infection rate at time $t + 1$ and selecting 2-hour time window is a requirement providing enough time to spread the delay. Flight times of flights in Europe are 2 hours in average. And a flight must arrive at the destination to spread the delay. Because of this reason, the real situation and simulation results can be matched at the end of 2-hour time period. So, simulation time is chosen as 2 hours. However, it is possible to predict the spreading for the next 4 hours, 6 hours, or a future time period. In this time, infection and recovery rates must be updated every 2 hours.

In this case, we have utilized the airport-based epidemic model to simulate the delay spreading on August 25, 2015, from 13:30 to 16:30. The recovery rates extracted from June 02, 2015, are used by assuming the similar disturbance is in effect.

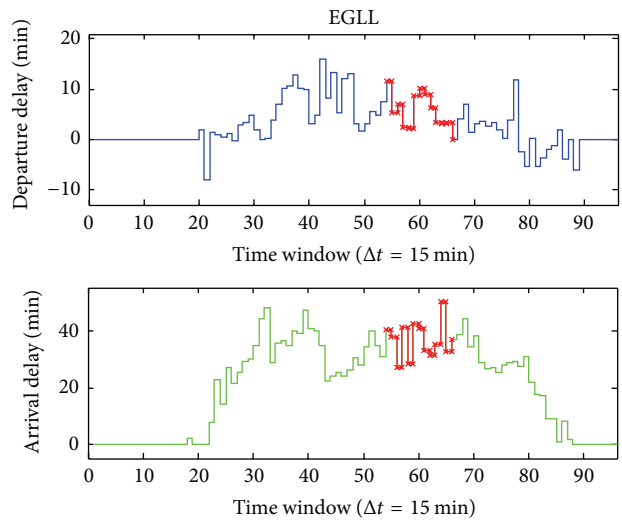


FIGURE 8: Actual delay profile of EGLL on June 02, 2015.

In the following ATC headline from EUROCONTROL's NOP, it was noted that LTFJ has a capacity shortage problem on August 25, 2015, and this problem was also noted for June 02, 2015.

By evaluating infection rates based on the traffic flow on August 25, 2015, infection fractions for 10 busiest airports in Europe are given in Figure 10(a) and the real infection ratios are also given in Figure 10(b). It is seen that the airport-based epidemic model provides accurate approximation in the prediction of infection ratios of most of the airports.

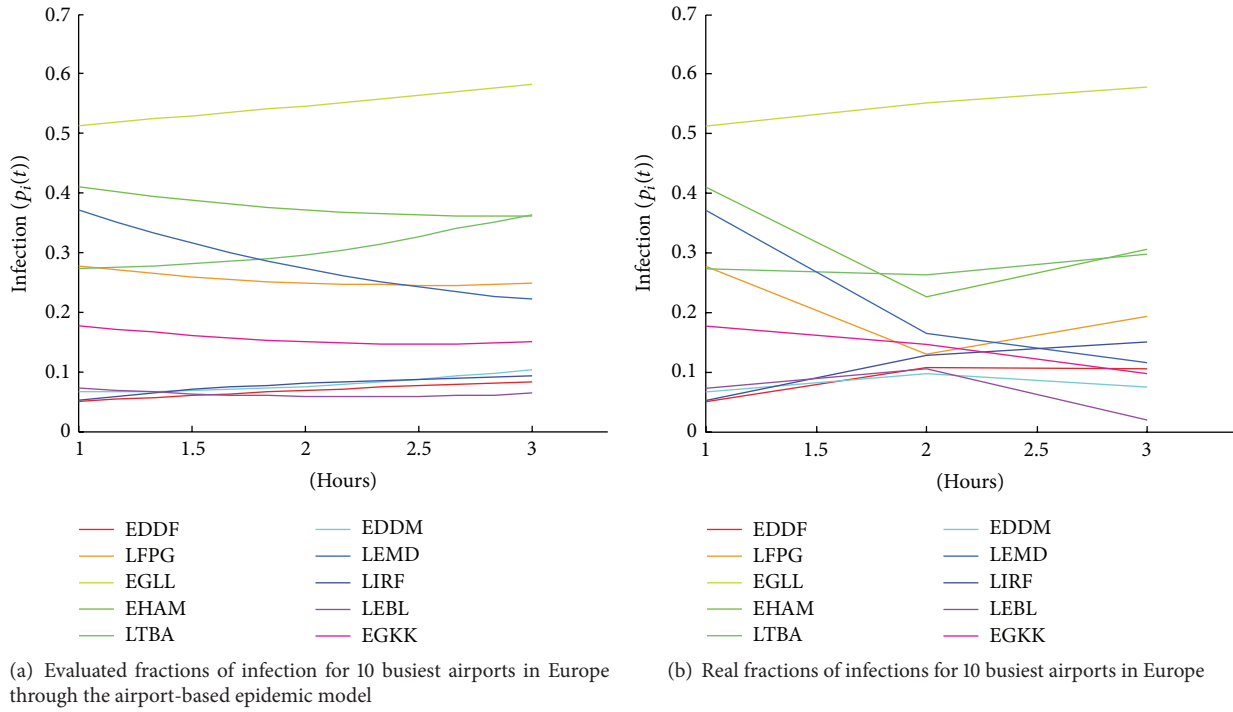


FIGURE 9: Comparison between the results of the airport-based epidemic model and real data on June 02, 2015.

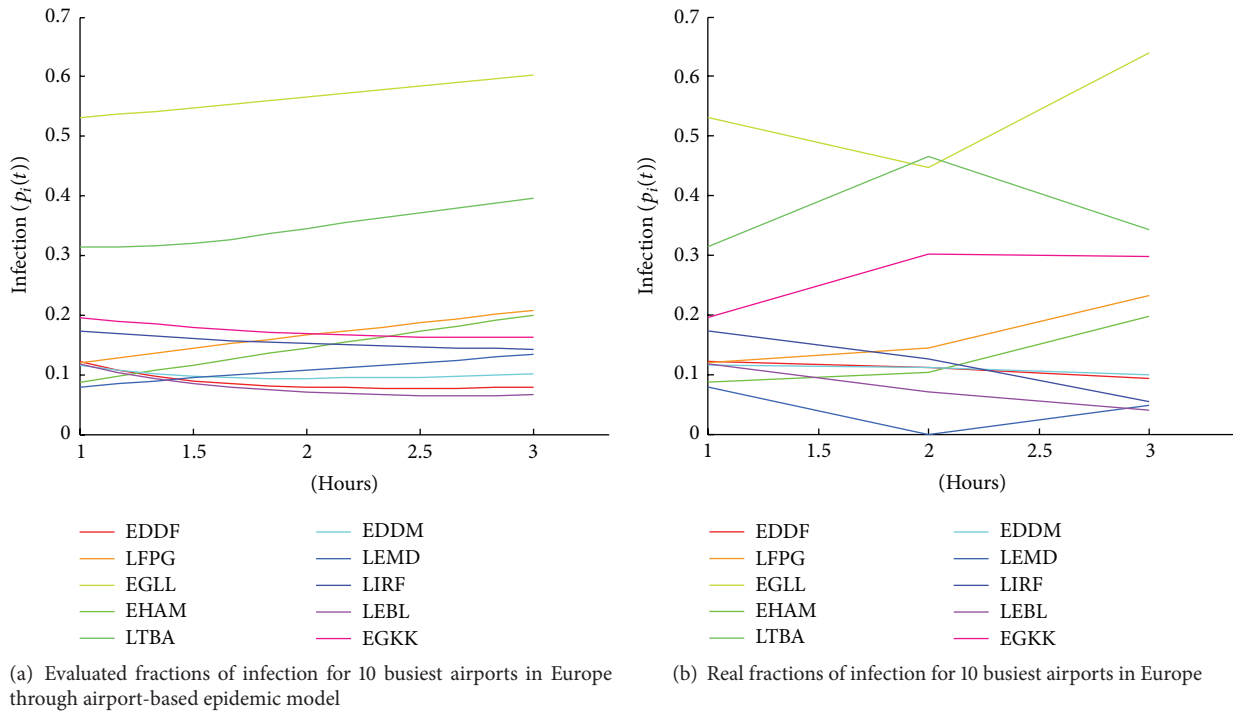


FIGURE 10: Comparison between the results of the airport-based metapopulation model and real data on August 25, 2015.

4.2. Analysis Recovery Characteristics of European Airports. To discuss typical recovery characteristics of European airports, nominal recovery rates between 13:30 and 16:30 during 2-week period (March 01, 2016, to March 14, 2016) are derived and given in Figure 11.

As seen in Figure 11, recovery rates of 10 busiest airports in Europe remain the same by comparing the instantaneous recovery characteristics under disturbances. Two of these airports have recovery rate smaller than 1, and it is observed that these airports most of the time work around their capacity

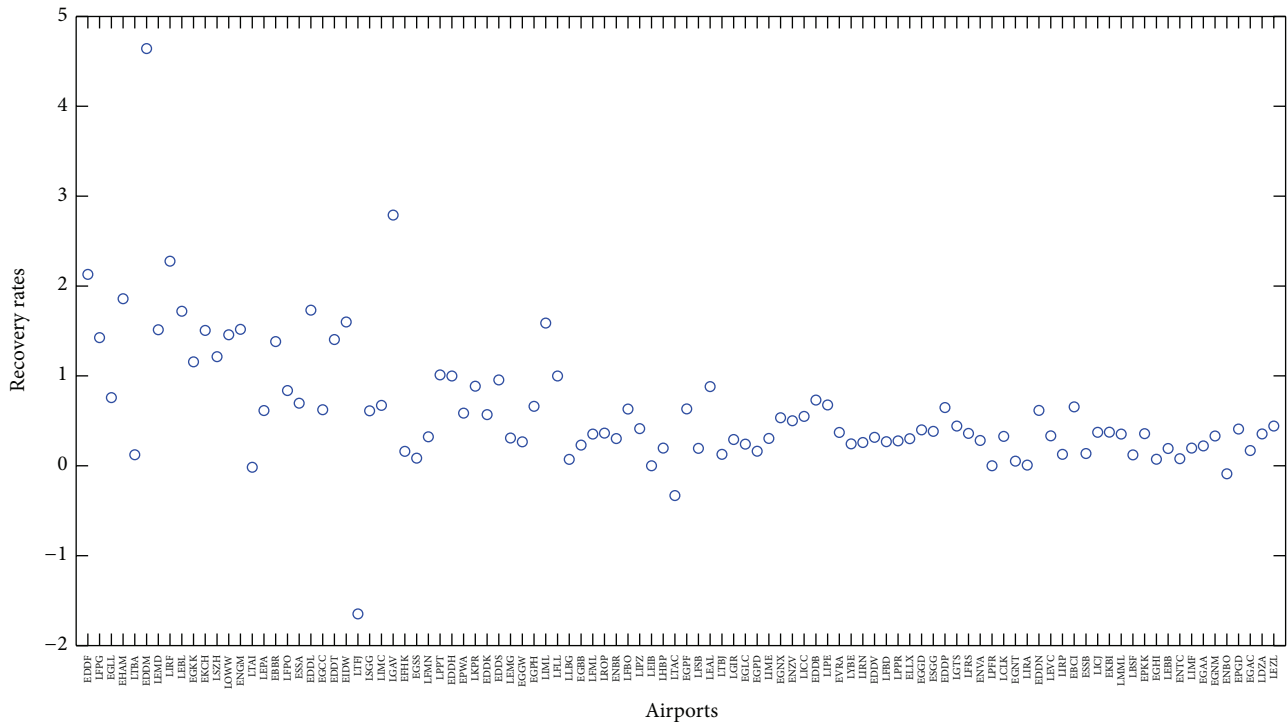


FIGURE 11: Evaluated nominal recovery rates of European airports through the airport-based epidemic model (analysis has been performed for 2 weeks: March 01–14, 2016).

limits. Furthermore, it is noted that the airports within the 10 busiest utilizing Collaborative Decision Making (CDM) in their operations, which is a procedure enabling information sharing between the airports, have higher recovery rates. A detailed discussion is avoided so as not to provide performance comparison between the airports through the conducted analysis as we have committed to the data provider.

5. Conclusion

In this study, the methodology of epidemic spreading was examined, and this method was applied to delay propagation in air transportation system. Two different epidemic models were constructed, which are airport-based metapopulation model and flight-based epidemic model utilizing the SIS epidemic model. The initial parameters estimation was performed through the historical flight record data to evaluate true recovery rates. The relation between the estimated recovery rate and the actual dynamics of the air transportation network has been analyzed to understand their behaviors under disruption. In the implementation phase, we have validated the accuracy of the models in capturing delay spreading behavior for certain events. The recovery characteristics of European airports were defined to understand the sensitivity of the nominal recovery characteristics to the disruptive events. Note that providing detailed discussion was avoided so as not to provide a performance comparison between the airports through our analysis as we have committed to the data provider throughout the paper.

The presented model can also provide active flow management by integrated control mode heuristics. As infection

rates in the airport-based model are extracted from plan data, when a flight is cancelled, this will directly affect the infection rates. In addition to cancellation, ground delays also change the infection rates. This idea is also true for the flight-based model; it can also be applied when active flow management measures are being taken to avoid congestion. In this model, each flight has a specific infection rate that relies on its two components: arrival time window and ground-waiting time. When a flight (j) is cancelled, related infection rates (β_{ij}^1 and β_{ij}^2) can be set to zero. These parameters can also be controlled through constructed heuristics in case ground delay is applied. Implementation of this practice also enables active flow management. ATM-induced delay related to configuration management can also be simulated by these models. A set of heuristics can be used to control recovery and infection rates for dynamic capacity-demand balancing (DCB). The recovery parameters on arrival airport can be adapted to simulate different kinds of operation modes. In the flight-based epidemic model, as the infection and recovery rates are flight specific, the flow rate management over OD pairs or flight cancellation simulations can be performed.

A potential future work will be to control or to bound the infection rates for improving the recovery rates of the elements in the systems. Moreover, sensitivity analysis in the estimation of the recovery rate will be conducted to understand the impacts of different disruptive events.

Competing Interests

The authors declare that they have no competing interests.

References

- [1] Airbus, "Global market forecast flying by numbers 2015–2034," Tech. Rep. 14029463, Airbus, 2015.
- [2] I. Frolow and J. H. Sinnott, "National airspace system demand and capacity modeling," *Proceedings of the IEEE*, vol. 77, no. 11, pp. 1618–1624, 1989.
- [3] F. Wieland, "Limits to growth: results from the detailed policy assessment tool," in *Proceedings of the 1997 16th AIAA/IEEE Digital Avionics Systems (DASC '97)*, vol. 2, pp. 9.2-1–9.2-8, IEEE, Irvine, Calif, USA, October 1997.
- [4] K. D. Bilimoria, B. Sridhar, G. Chatterji, K. Sheth, and S. Grabbe, "Facet: future atm concepts evaluation tool," *Air Traffic Control Quarterly*, vol. 9, 2001.
- [5] D. Long, D. Lee, J. Johnson, E. Gaier, and P. Kostiuk, "Modeling air traffic management technologies with a queuing network model of the national airspace system," Tech. Rep., National Aeronautics and Space Administration, Langley Research Center, 1999.
- [6] D. Long and S. Hasan, "Improved prediction of flight delays using the LMINET2 system-wide simulation model," in *Proceedings of the 9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO '09)*, Hilton Head, SC, USA, September 2009.
- [7] N. Pyrgiotis, *A stochastic and dynamic model of delay propagation within an airport network for policy analysis [Ph.D. thesis]*, Massachusetts Institute of Technology, 2012.
- [8] N. Pyrgiotis, K. M. Malone, and A. Odoni, "Modelling delay propagation within an airport network," *Transportation Research Part C: Emerging Technologies*, vol. 27, pp. 60–75, 2013.
- [9] K. Dietz and J. A. Heesterbeek, "Daniel Bernoulli's epidemiological model revisited," *Mathematical Biosciences*, vol. 180, pp. 1–21, 2002.
- [10] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 115, pp. 700–721, 1927.
- [11] Y. Zhou and H. Liu, "Stability of periodic solutions for an SIS model with pulse vaccination," *Mathematical and Computer Modelling*, vol. 38, no. 3-4, pp. 299–308, 2003.
- [12] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Analysis and control of epidemics: a survey of spreading processes on complex networks," *IEEE Control Systems*, vol. 36, no. 1, pp. 26–46, 2016.
- [13] H. W. Hethcote, "Three basic epidemiological models," in *Applied Mathematical Ecology*, S. A. Levin, T. G. Hallam, and L. J. Gross, Eds., vol. 18 of *Biomathematics*, pp. 119–144, Springer, Berlin, Germany, 1989.
- [14] V. Capasso and V. Capasso, *Mathematical Structures of Epidemic Systems*, vol. 88, Springer, Berlin, Germany, 1993.
- [15] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Review*, vol. 42, no. 4, pp. 599–653, 2000.
- [16] C. Li, R. van de Bovenkamp, and P. Van Mieghem, "Susceptible-infected susceptible model: a comparison of N -intertwined and heterogeneous mean-field approximations," *Physical Review E*, vol. 86, no. 2, Article ID 026116, 2012.
- [17] P. Van Mieghem and J. Omic, "In-homogeneous virus spread in networks," <http://arxiv.org/abs/1306.2588>.
- [18] A. Lajmanovich and J. A. Yorke, "A deterministic model for gonorrhoea in a nonhomogeneous population," *Mathematical Biosciences*, vol. 28, pp. 221–236, 1976.
- [19] Y. Ben-Zion, Y. Cohen, and N. M. Shnerb, "Modeling epidemics dynamics on heterogeneous networks," *Journal of Theoretical Biology*, vol. 264, no. 2, pp. 197–204, 2010.
- [20] G. H. Weiss and M. Dishon, "On the asymptotic behavior of the stochastic and deterministic models of an epidemic," *Mathematical Biosciences*, vol. 11, pp. 261–265, 1971.
- [21] F. Ball, "Stochastic and deterministic models for SIS epidemics among a population partitioned into households," *Mathematical Biosciences*, vol. 156, no. 1-2, pp. 41–67, 1999.
- [22] B. Baspinar, N. K. Ure, E. Koyuncu, and G. Inalhan, "Analysis of delay characteristics of European air traffic through a data-driven airport-centric queuing network model," in *Proceedings of the 14th IFAC Symposium on Control in Transportation Systems*, Istanbul, Turkey, May 2016.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

