Vulnerability of networks of interacting Markov chains

BY L. KOCAREV*, N. ZLATANOV AND D. TRAJANOV

Macedonian Academy for Sciences and Arts, Skopje, Macedonia

The concept of vulnerability is introduced for a model of random, dynamical interactions on networks. In this model, known as the influence model, the nodes are arranged in an arbitrary network, while the evolution of the status at a node is according to an internal Markov chain, but with transition probabilities that depend not only on the current status of that node but also on the statuses of the neighbouring nodes. Vulnerability is treated analytically and numerically for several networks with different topological structures, as well as for two real networks—the network of infrastructures and the EU power grid—identifying the most vulnerable nodes of these networks.

Keywords: vulnerability; VulnerabilityRank; Markov chains; networks

1. Introduction

In everyday life, we are surrounded by, and increasingly dependent on, engineered network structures. Therefore, the reliability of a network is crucial for its design and operation. The impact of component (link or node) failure on the performance of the network as a whole depends in part on network topology and in part on the flow occurring along the links. This impact is measured by network reliability. Traditionally, there are two approaches to evaluate the reliability of a stochastic network (Sanso & Soumis 1991).

— Static approach: the network is considered as a graph. In this case, the reliability of the network is related to some measure of connectivity of the network (Biechelt & Tittman 1991).
— Dynamic approach: the network is considered as a flow network that carries commodities from the origin to the destination nodes to satisfy a demand. In this case, the reliability of the network is related to the ability of the network to transmit a level of flow (Chan et al. 1997; Kishimoto 1997).

Related to the concept of reliability is a concept of vulnerability, which has been introduced in several fields including psychology, sociology, political science, economics, epidemiology, biology, environmental and geosciences, and engineering (McEntire 2005). In dictionary definitions of ‘vulnerable’, a common denominator is references to deliberate actions (threats), e.g. ‘susceptible to attack’ and ‘open

*Author for correspondence (lkocarev@ucsd.edu).

One contribution of 10 to a Theme Issue ‘Experiments in complex and excitable dynamical systems’.

This journal is © 2010 The Royal Society
to attack or assault by armed forces’ (Merriam-Webster 2006). However, there is no generally accepted definition of the concept of vulnerability even if we consider only technical, or engineering, applications. Below we will give a few examples of possible definitions of vulnerability in relation to technical systems.

Einarsson & Rausand (1998) studied industrial systems, and defined vulnerability as ‘the properties of an industrial system; its premises, facilities and production equipment, including its human resources, human organization and all its software, hardware and net-ware, that may weaken or limit its ability to endure threats and survive accidental events that originate both within and outside the system boundaries’. Berdica (2002) defines vulnerability in the road transportation system as ‘a susceptibility to incidents that can result in considerable reductions in road network serviceability’. In the field of information security, vulnerability is commonly thought of as a weakness in the security system that might be exploited to cause harm or loss. Morakis et al. (2003) define vulnerability as a ‘measure of the exploitability of a weakness’. In structural engineering, the term vulnerability is often used to capture the ‘susceptibility of a component or a system to some external action’. Thus, a structure is vulnerable if ‘any small damage produces disproportionately large consequences’ (Agarwal et al. 2003). Finally, vulnerability is also a topic in mathematics. In the branch of discrete mathematics called graph theory, vulnerability implies a lack of resistance of the graph to the deletion of vertices and edges (Barefoot et al. 1987).

In the network literature, there are different approaches to the concept of vulnerability. One trend relates the vulnerability or robustness of a network to its connectivity (Albert et al. 2000; Paul et al. 2005; Newman & Ghoshal 2008), while others relate it to the decrease in efficiency when some vertices or edges are under attack (Holme et al. 2002; Crucitti et al. 2003, 2004). In this paper we study the vulnerability of networks of interacting Markov chains. The concept of interactions on networks is not new, and has appeared in various forms in a variety of fields. The influence model (Asavathiratham 2000) differs from other previous models of interactions (such as the stochastic Ising model, cellular automata, the infinite particle system, the voter model, the interactive Markov chain) in several ways: the two most important differences are (i) each site (node) may contain an arbitrary (finite) local chain and (ii) the network may have an arbitrary (finite) graph and influence structure. The influence model (Asavathiratham 2000; Asavathiratham et al. 2001; Roy et al. 2002) is a simple (and mathematically tractable) model of random, dynamical interactions on networks. It consists of a network of nodes, each with a status that evolves over time. The evolution of the status at a node is according to an internal Markov chain, but with transition probabilities that depend not only on the current status of that node but also on the statuses of the neighbouring nodes.

In this paper, we introduce a new concept called VulnerabilityRank: it takes into account the network topology, node dynamics and potential node interactions in calculating the nodes’ influence and their relative priority. The behaviour of the influence model depends strongly on the so-called influence matrix (to be defined precisely below). We define the vulnerability of the network as the stationary distribution $\pi$ of its influence matrix, which is the normalized left eigenvector of the influence matrix associated with the eigenvalue 1. For the binary influence model, it measures a steady-state probability of the node failure. The vector $\pi$ is
called the VulnerabilityRank: it shows what are the most influential (vulnerable) nodes in the network. VulnerabilityRank is treated analytically and numerically for several networks with different topological structures. We compute the VulnerabilityRank for the interdependency matrix of an infrastructure’s network and the EU power grid, showing the most influential (vulnerable) nodes. Our work differs from two of the most relevant papers of this kind, including concepts of PageRank (Brin & Page 1998) and SecureRank (Miura-Ko & Bambos 2007), in the following ways: in both (Brin & Page 1998; Miura-Ko & Bambos 2007), the nodes are static and do not change with time, while, in our work, the dynamics of each node is governed by an arbitrary (finite) Markov chain.

2. Preliminaries: influence model

The influence model is suggested by Asavathiratham (2000) as a model of random, dynamical interactions on networks. We refer the reader to Asavathiratham (2000) for a full account of the model and its properties; here, we give a brief description of the model. We define the directed graph of an $n \times n$ matrix $A$, denoted by $G(A)$, as the directed graph on nodes 1 to $n$, where a directed edge from $i$ to $j$, denoted by $(i,j)$, exists if and only if $a_{ij} \neq 0$. The edge weight is given by $a_{ij}$. Consider a graph with $n$ nodes, referred to as sites; each site has a status value that varies over time as it is ‘influenced’ by the neighbours. Assume that we are given an $n \times n$ matrix $D = [d_{ij}]$ ($d_{ij} \geq 0$). We further assume that $D$ is a stochastic matrix, that is, $\sum_j d_{ij} = 1$ for each $i$. The graph $G(D^T)$ will be called the network influence graph. An edge $(i,j)$ exists on this graph if the status of $j$ can be influenced by the status of $i$. The weight on edge $(i,j)$ can be interpreted as the amount of influence that $i$ exerts on $j$ relative to the total amount of influence that $j$ receives. The total amount of influence received by any site is equal to the sum of incoming edge weights, which is 1, because $D$ is the stochastic matrix. Let $m_i$ be the order of the local Markov chain at the site $i$ for $1 \leq i \leq n$. Let $s_i(k)$ and $p_i(k)$ be the status vector and the next-status probability mass function (PMF) vector of site $i$ at time $k$. Let $s(k) = [s_1, \ldots, s_n]^T$ and $p(k) = [p_1, \ldots, p_n]^T$ denote the state and probability vectors of length $(m_1 + \cdots + m_n)$. For each pair of $i$ and $j$, the state transition matrix $A_{ij}$ is an $m_i \times m_j$ non-negative matrix whose rows sum to 1. The influence matrix is defined as $H = D^T \otimes \{A_{ij}\}$, where $\otimes$ denotes the Kronecker product. The evolution equations of the influence model are defined as

$$p^T(k+1) = s^T(k)H$$

and

$$s^T(k+1) = \text{MultiRealize}[p^T(k+1)],$$

where the operation MultiRealize$[p(k+1)]$ treats each block of PMFs within $p^T(k+1)$ separately, independently realizing the new status of the vectors block by block.

The influence matrix $H$ is, in general, not stochastic. However, its dominant eigenvalue is 1. Assuming for simplicity that all its eigenvalues are distinct,
the evolution of the status of the PMF approaches, the left eigenvector $\pi$ corresponding to eigenvalue 1, that is,

$$E(s^T(k)) = E(s^T(0))H^k \rightarrow \pi,$$

as $k \rightarrow \infty$, where the notation $E(\cdot)$ is used for expectation or the expected value.

In what follows, we discuss in more detail the binary influence model.

3. VulnerabilityRank

The binary influence model is the special case of an influence model for which each $A_{ij}$ is equal to the $2 \times 2$ identity matrix $I_2$, $A_{ij} = I_2$. For the binary influence model, the status of the site $i$ is represented by $s_i$, $s_i \in \{0,1\}$. The values 0 or 1 may represent any two different statuses such as ‘on’ versus ‘off’, ‘healthy’ versus ‘sick’ or ‘normal’ versus ‘failed’. Let $s = [s_1, \ldots, s_n]^T$. The binary influence model refers to the following equation:

$$p(k + 1) = Ds(k) \quad (3.1)$$

and

$$s(k + 1) = \text{Bernoulli}[p(k + 1)]. \quad (3.2)$$

Since $D$ is a stochastic matrix, it follows that $p_i(k) \leq 1$ for each $k$ and $i$. The operation $\text{Bernoulli}[p(k + 1)]$ in equation (3.2) can be thought of as flipping $n$ independent coins to realize the entries of $s(k + 1)$, where the probability of the $i$-th coin turning up heads (status 1) is $p_i(k + 1)$. If $\Gamma(D^T)$ is an ergodic (irreducible and aperiodic) graph, then the only recurrent states in a binary influence model are the all-ones and all-zeros consensus states. In this case,

$$\lim_{k \rightarrow \infty} D^k = 1\pi^T, \quad (3.3)$$

where $\pi = [\pi_1, \ldots, \pi_n]^T$ is the left eigenvector corresponding to the eigenvalue 1, which is normalized, so that $\pi^T 1 = 1$, and $1^T = [1, \ldots, 1]$ is the vector of length $n$.

For the binary influence model, we call the vector $\pi$ the VulnerabilityRank for the network $\Gamma(D^T)$. We call $\pi_j$ the vulnerability of the node $j$. The following simple observation explains why the vector $\pi$ is called the VulnerabilityRank. Assume that the network is ergodic. Assume further that, for a given $j$, $s_j(0) = 1$ if $i = j$ and $s_i(0) = 0$ if $i \neq j$. Then the probability of reaching the all-ones consensus state is $\pi_j$. Indeed, since the model is ergodic, it follows $p_i > 0$; moreover, $\sum \pi_i = 1$. From equation (3.3) we conclude that

$$\lim_{k \rightarrow \infty} E(s(k) | s(0)) = \lim_{k \rightarrow \infty} D^k s(0) = 1\pi^T s(0) = \sum_i s_i(0)\pi_i.$$

The last equation indicates that all sites have the same probability of reaching the status 1. Let $s_j(0) = 1$ and $s_i(0) = 0$ for all $i \neq j$. Then, the last equation reduces to $\pi_j$. Note that if the value 1 represents the failed status of the network, the VulnerabilityRank describes what is the vulnerability of each site $i$—it is exactly the value $\pi_i$. In the special case, if the site $j$ is such that $\pi_j = \max_{1 \leq i \leq n} \pi_i$, then $j$ is the most vulnerable site of the network.
It is easy to see that if $s_j(0) = \cdots = s_{jk}(0) = 1$ and $s_i(0) = 0$ otherwise, then
\[
\lim_{k \to \infty} E(s(k) \mid s(0)) = \pi_{ji} + \cdots + \pi_{jk}.
\]
Therefore, if $\pi_{ji} + \cdots + \pi_{jk} > 0.5$, then the probability of reaching the all-ones consensus state for the binary influence model with initial conditions $s_{jk}(0) = 1$ and $s_i(0) = 0$ otherwise is greater than 0.5.

As an example, we consider a binary influence model with a fully connected network of six nodes and the influence matrix $D$ given by
\[
D = \begin{bmatrix}
1/60 & 7/15 & 7/15 & 1/60 & 1/60 & 1/60 \\
1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\
19/60 & 19/60 & 1/60 & 1/60 & 19/60 & 1/60 \\
1/60 & 1/60 & 1/60 & 1/60 & 7/15 & 7/15 \\
160/60 & 1/60 & 1/60 & 7/15 & 1/60 & 7/15 \\
1/60 & 1/60 & 1/60 & 11/12 & 1/60 & 1/60
\end{bmatrix}.
\]

It is easy to compute the eigenvector $\pi$,
\[
\pi^T = [0.03721 \ 0.05396 \ 0.04151 \ 0.3751 \ 0.206 \ 0.2862].
\]

Therefore, site 4 is the most vulnerable site and $\pi_4 + \pi_6 > 0.5$.

For the case of heterogeneous Markov chains, the VulnerabilityRank is defined as the left eigenvector $\pi$ corresponding to eigenvalue 1 of the influence matrix $H$. We will deal with the heterogeneous influence model in a forthcoming paper, while here a simple illustrative example is presented. Assume that we consider a model of two interdependent nodes, for which one node represents a power plant (node number 2) and the other node represents a customer (node number 1). Define the matrix $D$ as
\[
D = \begin{bmatrix}
0.5 & 0.5 \\
1 & 0
\end{bmatrix},
\]
meaning node 1 receives influence 0.5 from itself and from node 2, while node 2 is completely influenced by node 1. Assume further that node 1 can have only two statuses, ‘demand’ and ‘normal’ (not demand), and node 2 has three possible statuses ‘normal’, ‘alert’ or ‘failed’.

The influence matrix for heterogeneous Markov chains is defined as $H = D^T \otimes \{A_{ij}\}$, where $\otimes$ denotes the Kronecker product. Assume that the matrices $A_{11}$, $A_{12}$ and $A_{21}$ are given by
\[
A_{11} = \begin{bmatrix}
\text{demand} & \text{normal} \\
\text{demand} & 0.5 & 0.5 \\
\text{normal} & 0.5 & 0.5
\end{bmatrix},
\]
\[
A_{12} = \begin{bmatrix}
\text{normal} & \text{alert} & \text{failed} \\
\text{demand} & 0 & 0.5 & 0.5 \\
\text{normal} & 0.5 & 0.5 & 0
\end{bmatrix},
\]
and
\[
A_{21} = \begin{bmatrix}
\text{normal} & \text{normal} \\
\text{alert} & 0.5 & 0.5 \\
\text{failed} & 0 & 1
\end{bmatrix}.
\]
We see (from $A_{12}$) that when node 1 is in the ‘demand’ state then node 2 can evolve to the ‘alert’ or ‘failed’ status with equal probability, but it cannot go to the ‘normal’ status. When node 1 has the ‘normal’ status, node 2 can be in the ‘normal’ or ‘alert’ status, while it cannot be in the ‘failed’ state. Similarly, from $A_{21}$, it follows that, when node 2 is in the ‘normal’ or ‘alert’ status, then node 1 can change the state to ‘demand’ or ‘normal’ with equal probabilities. When node 2 is in the ‘failed’ state, node 1 can evolve only to the ‘demand’ status. For $k \to \infty$, PMF approaches the left eigenvector $p$ corresponding to eigenvalue 1, which, for the considered example, is equal to

$$\begin{bmatrix} 0.4444 & 0.5556 & 0.2778 & 0.5 & 0.2222 \end{bmatrix}.$$ 

Therefore, for this example, the most vulnerable node is node 1 (customer) in the ‘normal’ status.

(a) *VulnerabilityRank for different network topologies*

In this subsection, we address the following problem: given a graph, how do we define the corresponding network influence graph. Let $G$ be a finite simple undirected connected graph with $n$ nodes, so that its adjacency matrix $A = (a_{ij})$ is a symmetric (0,1) matrix with zeros on its diagonal. We suggest two approaches to relate the network influence graph to an arbitrary graph $G$.

For the first approach, called the node–degree influence model, we define the influence matrix $D = (d_{ij})$, such that $d_{ij} = a_{ij}/\sum_i a_{ij}$. The second model, called the betweenness–centrality influence model, is defined as follows. Recall that the edge betweenness centrality is defined as

$$c_{ij} = \sum_{s \neq i, s \neq t} \frac{\sigma_{st}(e_{ij})}{\sigma_{st}},$$

where $e_{ij}$ is the edge between nodes $i$ and $j$, $\sigma_{st}(e_{ij})$ is the number of shortest paths from node $s$ to node $t$ that edge $e_{ij}$ lies on, and $\sigma_{st}$ is the total number of shortest paths from node $s$ to node $t$. The shortest path is the minimum distance between two nodes. The distance between two nodes is the sum of edge weights on that path. For this model, the matrix $D = (d_{ij})$ is defined as $d_{ij} = c_{ij}/\sum_i c_{ij}$. Edges that occur on many shortest paths have higher betweenness than those that do not. Therefore, in the betweenness–centrality influence model, edges that have higher betweenness have larger influence.

Figures 1–3 show the *VulnerabilityRank* for several networks with different topologies using the node–degree influence model. The most vulnerable site is in the scale-free network: its vulnerability is 10 times larger than the maximum vulnerabilities of small-world and Erdős–Rényi (ER) graphs. Similar figures are obtained for the betweenness–centrality influence model. The average maximum values for the VulnerabilityRank for the three graphs, scale-free, small-world and ER, when the node–degree influence model is used, are 0.0413, 0.0048 and 0.0067, respectively. The same values for the betweenness–centrality influence model are 0.1099, 0.0171 and 0.0079. We have also computed the average percentage of the nodes that need to be in the status off at time $k = 0$ so that, when time goes to infinity ($k \to \infty$), the probability of the whole network (all sites) to be in the status 1 (off) is greater than or equal to 0.5. These numbers for the node–degree
Vulnerability of networks

Figure 1. The VulnerabilityRank for the node–degree binary influence model on the small-world graph.

Figure 2. The VulnerabilityRank for the node–degree binary influence model on the scale-free graph.

Figure 3. The VulnerabilityRank for the node–degree binary influence model on the ER graph.
influence model and the scale-free, small-world and ER graphs are 21.24, 40.71 and 41.09, respectively. For the betweenness–centrality influence model, we have 9.73 for scale-free networks, 24.27 for small-world networks and 40.04 for ER graphs. For both models (the node–degree and betweenness–centrality influence models), the scale-free graph is the most vulnerable graph, while the random ER graph is the most robust graph.

(b) VulnerabilityRank for reducible graphs

Irreducibility is a desirable property because it is precisely the feature that guarantees that a Markov chain possesses a unique (and positive) stationary distribution vector \( \pi \). When \( \Gamma(D^T) \) is an ergodic graph, then computation of the VulnerabilityRank for the graph \( G \) is easy. However, when \( \Gamma(D^T) \) is reducible, further adjustment is necessary in order to ensure irreducibility. We first compute the following quantity

\[
\pi_i = \frac{1}{n} \sum_{k=1}^{n} \lim_{k \to \infty} D^k s_i(0),
\]

where \( s_i(0) = [s_{i1}(0), \ldots, s_{in}(0)]^T \), such that \( s_{ii} = 1 \), otherwise \( s_{ij} = 0 \). The VulnerabilityRank for this graph is \( \pi = [\pi_1, \ldots, \pi_n]^T \). Next, following Brin & Page (1998), we make every state directly reachable from every other state by adding a perturbation matrix to \( D \) so that

\[
D^* = \alpha D + (1 - \alpha) \frac{11^T}{n}.
\]

It is easy to show that if the respective spectra of \( D \) and \( D^* \) are \( \sigma(D) = \{1, \mu_2, \ldots, \mu_n\} \) and \( \sigma(D^*) = \{1, \lambda_2, \ldots, \lambda_n\} \), then \( \lambda_k = \alpha \mu_k, \quad k = 2, \ldots, n \). It should be noted here that the matrix \( D^* \) in the context of the Web’s hyperlink structure is generally called ‘the Google matrix’ and its stationary distribution is the real PageRank vector. Let \( \pi^* \) be the left eigenvector corresponding to the eigenvalue 1 of the matrix \( D^* \). Numerically, we found that \( \pi^* \approx \pi \) for values of \( \alpha \) close to 1.

We now present an example. Let

\[
D = \begin{bmatrix}
0 & 1/2 & 1/2 & 0 & 0 & 0 \\
1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\
1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\
0 & 0 & 0 & 0 & 1/2 & 1/2 \\
0 & 0 & 0 & 1/2 & 0 & 1/2 \\
0 & 0 & 0 & 1 & 0 & 0
\end{bmatrix}.
\]

This matrix is stochastic, but it is reducible, so it cannot have a unique positive stationary distribution. To force irreducibility, choose \( \alpha = 0.9 \); thus, one obtains the matrix given with equation (3.4), for which we have already found that node 4 is the most influential node.

(c) VulnerabilityRank for the SIR model

In this subsection, we study the VulnerabilityRank for a stochastic susceptible–infected–removed (SIR) model. Let \( G \) be a finite simple undirected connected graph with \( n \) nodes, so that its adjacency matrix \( A = (a_{ij}) \) is a symmetric (0,1) matrix with zeros on its diagonal. Let \( [p_i^S(k) \ p_i^I(k) \ p_i^R(k)] \) be the

Phil. Trans. R. Soc. A (2010)
three-dimensional probability vector of node $i$ at time $k$. $p^S_i(k)$, $p^I_i(k)$ and $p^R_i(k)$ are the probabilities that node $i$ at time $k$ is in the state susceptible, infected and removed, respectively. Further, let $[s^S_i(k) \ s^I_j(k) \ s^R_j(k)]$ be the three-dimensional status vector of node $i$ at time $k$. The status vector can have only one element equal to 1; the other elements are equal to 0. For $s^S_i(k) = 1$, the node $i$ has the status susceptible; if $s^I_j(k) = 1$, then node $i$ has the status infected; for $s^R_j(k) = 1$, node $i$ has the status removed. We consider the following version of the SIR model, for which the node equations are

$$p^S_i(k + 1) = 1 - p^I_i(k + 1) - p^R_i(k + 1), \quad (3.6)$$

$$p^I_i(k + 1) = s^S_i(k) \left[ 1 - \prod_{j=1}^{N} (1 - \beta a_{ij} s^I_j(k)) \right] \quad (3.7)$$

and

$$p^R_i(k + 1) = s^I_i(k) + s^R(k), \quad (3.8)$$

where $\beta$ is the probability that the infection attempt is successful, and $A = (a_{ij})$ is the adjacency matrix of the graph. Each node that is infected at time $k$ attempts to infect each of its neighbours; each infection attempt is successful with probability $\beta$ independent of other infection attempts. Each infected node at time $k$ is removed at time $k + 1$.

We consider the epidemic spreading on graphs starting with one node initially infected, and all other nodes are susceptible. In this section, we address the following problem: given a graph, what is the most important (influential, vulnerable) node. Let $N(v)$ be the average number of infected nodes when time goes to infinity and only node $v$ is initially infected. Intuitively, we say that node $v$ is the most influential node in the network when $N(v) > N(u)$ for all nodes $u$ in the network (the average number of infected nodes when time goes to infinity is the largest or the time needed for the nodes to be infected is the largest). For the SIR model, we compute the VulnerabilityRank for the node–degree influence model. The VulnerabilityRank gives the answer to our problem; indeed, the most vulnerable node computed with the VulnerabilityRank is also the most important (influential) node for the SIR dynamics on the graph $G$. As an example, we show in figure 4 the average number of infected nodes for the small-world graph versus time for two different values of $\beta$ when the most influential, the least influential and random nodes are initially infected.

4. Applications

(a) Network of infrastructures

Infrastructures are vital for the operation of our society. A mathematical model that might be applied for the assessment of the compound risk of failure of interdependent infrastructure networks is provided by Sivonen (2004), in the frames of which the following groups of inter-operating items have been considered: technical infrastructure (energy supply, communications, information systems), basic services and supplies (food supply, transport logistics, mass media, healthcare, financial services) and threats (threats to data systems, illegal

\[ \text{Phil. Trans. R. Soc. A (2010)} \]
immigration, threats to food and health, threats to the environment, economic threats, crime and terrorism, disasters, international tension, war and war-like situations). These items can be broken down to a more detailed level. It is important to understand the difference between the threats and the rest of the items. The threats come from outside the system and the model will first allocate the risks to threats, because all the items depend on them, but they do not depend on anything. For each of the items, the frequency, duration and effects of failures are studied, as well as the dependency of a failure of one item on the failures of other items. The interdependencies between different infrastructures and basic services and threats are presented as a grid of different symbols. The colour symbol BLACK means that a failure in the item in the vertical column is one of the primary causes of a failure in the item in the horizontal row (e.g. relative value of 1). RED means a secondary cause (relative value of 0.1); YELLOW a rare cause (relative value 0.01); and GREEN a possible cause (relative value 0.001). (WHITE: no dependency.) If there are 200 items, there will be 40 000 possible positions in the interdependency grid.

The mean time between failures in each infrastructure or service and occurrence of a threat are classified as BLACK (less than a year), RED (1–10 years), YELLOW (10–100 years) and GREEN (more than 100 years). The durations

Figure 4. Average number of infected nodes versus time for three different initially infected nodes: (a) $\beta = 0.5$; (b) $\beta = 1$. Open circles, most vulnerable; open diamonds, least vulnerable; plus symbols, random node.

Phil. Trans. R. Soc. A (2010)
of different failures are classified as ≤0.5 day, ≤1 day, ≤0.5 week, ≤1 week, and >1 week. The direct effect of a 1 day long failure in each infrastructure or service is classified as BLACK (more than 1000 units), RED (100–1000 units), YELLOW (10–100 units) and GREEN (1–10 units). The unit can be freely chosen. It can be financial loss, loss of one human life or an abstract disadvantage measurement unit.

The results of the calculation are, for each item, as follows: probability of at least one failure in a year (to give an indication of vulnerability), combined effect of failure for 1 day and combined risk for 1 year. The items are sorted by decreasing combined risk. The compound risk connected to the infrastructures or services is the stochastically expected value of the effect of failures. This is equal to the quantity effect × probability × duration added over the different duration categories and times of occurrences of failure per year. The compound risk quantified by the method explained above is considered as the influence between the nodes, and so it gives the elements of the influence matrix \( (d_{ij}) \) (see Sivonen 2004). An example of a network influence graph, \( \Gamma(D^T) \), obtained in this way using real data from the Web page of the National Emergency Supply Agency of Finland (www.nesia.fl) is shown in figure 5.

Figure 6 shows the VulnerabilityRank for the infrastructure network influence graph shown in figure 5. The most vulnerable sites are the sites representing the following threats (out of 17 threats grouped in four groups: causes for severe disturbances, economic threats, environment and health treats, and political security threats): 1, weather phenomenon; 2, threats to data systems; 3, crime and terrorism; 4, strike; and 5, international logistics crisis. Threats 1 and 4 belong to the same group—‘causes for severe disturbances’; threats 2 and 5 to the group ‘economic threats’; and 3 to ‘political security threats’. Since the influence graph in this example is reducible, we compute the VulnerabilityRank using the approach described in the previous section with \( \alpha = 0.9 \).
Our next example is the EU power grid. The Union for the Co-ordination of Transmission of Electricity (UCTE) is the association of transmission system operators in continental Europe; different data can be found at the UCTE Web page (www.ucte.org). For this paper, we consider the physical energy flows for the month of January 2007. The data can be organized as a directed weighted graph with $n = 26$ nodes. Each node represents a country (or region) in the EU as follows: 1, Austria; 2, Bosnia; 3, Belgium; 4, Bulgaria; 5, Switzerland; 6, Czech Republic; 7, Germany; 8, Spain; 9, France; 10, Greece; 11, Croatia; 12, Hungary; 13, Italy; 14, Luxembourg; 15, Monte Negro; 16, Macedonia; 17, Netherland; 18, Poland; 19, Portugal; 20, Romania; 21, Serbia; 22, Slovenia; 23, Slovakia; 24, Denmark West; 25, Ukraine West; and 26, others (Albania, Belarus, Denmark East, Great Britain, Morocco, Republic of Moldavia, Norway, Sweden, Republic of Turkey and Ukraine). The weight $a_{i,j}$ represents a value (in GWh) that country $i$ exports to country $j$. For example, $a_{1,5} = 906$, $a_{1,7} = 227$, $a_{1,12} = 104$, $a_{1,13} = 121$ and $a_{1,22} = 72$ means that Austria exported, during January 2007, 906 GWh energy to Switzerland, 227 GWh to Germany and so on.

We compute the VulnerabilityRank for the influence model defined on this graph using both approaches: node–degree and betweenness–centrality. The results are shown in figures 7 and 8. For the node–degree influence model, the most vulnerable node is 9 (corresponding to France), while for the betweenness–centrality influence model, the most influential node is 7 (corresponding to Germany).

Our last example is a network representing the power grid of the EU, which contains a total of 12 741 nodes of the following six types connected by 17 798 edges:

1. gas node (a junction, intersection or end-node) – 2212 nodes;
2. electricity substation – 5096 nodes;
3. natural gas power plant – 998 nodes;
4. generic power plant – 4383 nodes;
5. DC line connected substation – 31 nodes; and
6. liquefied natural gas (LNG) terminal – 21 nodes.
The data are collected in the EU MANMADE project. The graph contains directed and undirected edges; the weights of the edges are normalized so that $D$ is the stochastic matrix. Directed edges can be grouped into two classes: ‘gas node $\rightarrow$ gas power plant $\rightarrow$ substation’ and ‘generic power plant $\rightarrow$ substation’. Since the influence graph in this example is reducible, we compute the VulnerabilityRank using the approach described in the previous section with $\alpha = 0.85$. The nodes are grouped as follows: gas nodes are grouped from node index 1 to node index $\approx 2200$; from node index $\approx 2200$ to $\approx 7300$ the graph consists of power substations; and, finally, gas power plants and generic power plants are labelled with node indexes $>7300$. Figure 9 depicts the VulnerabilityRank for the EU power grid network. The boundaries 2200 and 7300 can be seen from the figure: the VulnerabilityRank has different values for the three subnetworks. The most vulnerable nodes are power substations.
Figure 9. The VulnerabilityRank for the EU power grid.

5. Conclusions

We have suggested a method for calculating the VulnerabilityRank for networks of interacting Markov chains. The method is readily applicable for huge matrices and heterogeneous Markov chains. The method can be applied to any network, including most of the infrastructure networks, such as the power grid, gas network and transportation network, as well as to the network of infrastructures. It can also be extended to biological and social networks provided that each node can be described with a Markov chain. We stress that our concept of the VulnerabilityRank is different from both the recently introduced concepts of PageRank and SecureRank—it reduces to these concepts only when we consider the node–degree binary influence model. More general treatment of the VulnerabilityRank for the heterogeneous influence model will be a subject of interest of our future work.

We thank Hannu Sivonen for sending us the data describing the interdependencies of critical infrastructures of Finland. We thank Daniel Smilkov for computing the VulnerabilityRank for the EU power grid. This work is partially supported by the EU Commission (project MANMADE).

References


Merriam-Webster. 2006 Merriam-Webster online dictionary. See http://www.m-w.com/


