Resilient and Secure Network Design for Cyber Attack-Induced Cascading Link Failures in Critical Infrastructures

Yezekael Hayel[†‡] and Quanyan Zhu[‡]

[†]LIA/CERI, University of Avignon, Avignon, France, Email: yezekael.hayel@univ-avignon.fr

[‡]Department of Electrical and Computer Engineering, New York University, USA

E-mail: {yezekael.hayel,quanyan.zhu}@nyu.edu

***Abstract*—Designing resilient and secure critical infrastructure networks against cascading failures is essential to mitigate natural disasters and unanticipated cyber attacks. This paper develops a dynamic and distributed resilient mechanism that enables efficient link recoveries and hence reduce the impact of cascading failures. The evolving system-wide link failure is captured by a continuous-time epidemic process over a network using mean-field approximation techniques. The optimal resilient policies are designed to maximize the global connectivity of the network in the worst-case failure state, while minimizing the cost of recovery.**

I. Introduction

Cascading failures have imposed major threats to critical infrastructures due to their potential catastrophic consequences on our nation's economy and social well-being. Local link failures in networks, such as power grids, transportation and water systems, can induce global propagation of failures within the network and across interdependent networks [1]. Due to the increasing integration of the information communication technologies (ICTs), the cause of such failures is no longer limited to natural disasters, and system's own reliability issues. Instead, cyber attacks become a major concern for causing unexpected failures, and the network failure can be further exacerbated by strategic adversarial behaviors [2].

One way to mitigate such attacks is to build resilience mechanisms into networks in order to reduce the impact of the unanticipated link failures and maintain basic functions of the networks. In this work, we propose a distributed and dynamic mechanism that allows the network to self-repair its compromised links in real time by taking to account the link failure dynamics and the consequences of failure propagation over the network. The network performance is measured by the algebraic connectivity of the network, and it evolves as the link states change according to the rates of failures and recovery. The proposed mechanism provides an optimal recovery mechanism for each link that balances the cost of preparation, and the global network performance.

A. Related Work and Contributions of the Paper

In recent literature, resilient/secure network designs have been studied in [3], [4]. In [3], the author studies the network connectivity using a novel centrality measure that quantifies sensitivity of the size of the largest connected component to node removals. They show that the network resilience can be greatly improved via a few edge rewires without introducing additional edges in the network. In another work [4], the authors propose a decentralized algorithm to control the deletion and creation of links to keep the network connectivity, measured using the algebraic connectivity metric. Many works have been proposed to find graphs that can maintain global connectivity. For example, in [7], the authors study the problem of designing the topology, i.e., assigning the probabilities of reliable communication among sensors (or of link failures) to maximize the rate of convergence of average consensus, when the link communication costs are taken into account. They show that under a Bernouilli random topology, a necessary and sufficient condition is for the algebraic connectivity of the mean graph topology to be strictly positive. In our setting, the random topology comes from the failure dynamics. This approach captures the dynamics of the cascading failure as well as the network system topology. One key step in this work is to approximate the random graph by a dynamic graph using the Markov differential equation on each link. The resulting continuous-time dynamical system will be leveraged to design optimal recovery policies and build distributed resilient mechanism across the network. The contributions of this paper are summarized as follows:

(i) Optimal link recovery problem is formulated based on a dynamic cascading failure framework captured by a continuous-time epidemic process over links of a network.

(ii) Dynamic and distributed resilient network mechanism is designed to provide network resiliency to cascading failures.

(iii) Resilient mechanisms are developed for worst-case scenario of the propagation failures.

B. Organization of the Paper

The rest of the paper is organized as follows: Section II describes the system model and the problem formulation of optimal resilient mechanism. Section III presents the algorithmic solution to the problem. Section IV concludes the paper.

II. System Model

The network is modeled as a graph $\mathcal{G}$ composed of nodes and undirected links. The set of nodes $\mathcal{N}$ is finite $\{1,\dots,N\}$ and we denote by A the adjacency matrix, i.e. it is a symmetric matrix size $N\times N$ and $A_{ij}>0$ if and only if node i is directly connected to node j. We assume that for all node i, we have $A_{ii}=0$. We denote by $\mathcal{V}$ the set of edges and $|\mathcal{V}|=E$.

The goal in this paper is to study the impact on the connectivity of a cascading dynamic failures of links. Each link $l\in\mathcal{V}_{\mathcal{G}}$ can be in one of two states 0 or 1 at each time

978-1-4799-8428-2/15/$31.00 ©2015 IEEE

t and we denote by $s_l(t)$ the state of link l at time t. If link l is broken (inactive) at time t then its state is 1, otherwise his state is 0. In the last state, we say that the link is active. We observe on figure 2 an example of such dynamic process where at time t_1 one link failed, at time t_2 other links have failed du to the propagation of the failure and at time t_3 one failed link is repaired.

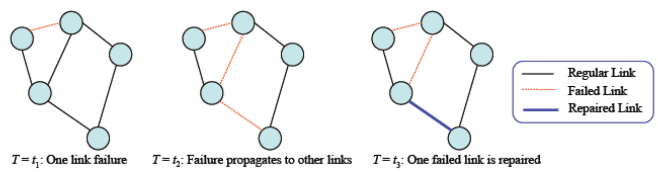

Fig. 1. Example of cascading failures and repairing.

A. Algebraic connectivity

The level of connectivity of a given graph $\mathcal{G}$ can be measured trough several metrics like vertex connective, edge connectivity or algebraic connectivity. We consider this last one as it has many applications in different domains [8] and, moreover, it has interesting topological properties like all trees have the same vertex and edge connectivity (one), whereas the algebraic connectivity of a star is higher than that of a path. The computation of the algebraic connectivity of a graph can be very difficult as the objective function is complex, being an eigenvalue of a matrix. Denoting by $d_i = \sum_{j=1}^{N} a_{ij}$ the degree of any node $i \in \mathcal{N}$, the degree matrix D is the diagonal matrix with entries d_{ii}.

Definition 1: The (weighted) Laplacian matrix $L(\mathcal{G})$ of graph $\mathcal{G}$ is the $n \times n$ matrix L defined by: $L = D - A$.

We have the following definition of the algebraic connectivity of a graph $\mathcal{G}$ [9].

Definition 2: The algebraic connectivity $a(\mathcal{G})$ of a graph $\mathcal{G}$ is the second-smallest eigenvalue of $L(\mathcal{G})$.

The magnitude of this value reflects how well connected the overall graph is and it can be used to efficiently partition the graph. The most important property of the algebraic connectivity is that this metric is greater than 0 if and only if the graph $\mathcal{G}$ is connected. The *Fielder vector*, denoted by $v(\mathcal{G})$, is the normalized eigenvector associated to the algebraic connectivity of graph $\mathcal{G}$. The level of connectivity of node i is measured trough the value $v_i(\mathcal{G})$.

B. Stochastic propagation failures and dynamic graph

We consider that the security issues may induce link failures propagation into the network. Then, our setting is dynamic and particularly the graph $\mathcal{G}$ is evolving over time following an epidemic process. Indeed, each edge $l \in \mathcal{V}$ can be in one of the two state 0 (inactive) or 1 (active) at time t. The state $l(t) = 0$ means that edge l is infected, i.e. broken/inactive, and that it is no more used to communicate between the two nodes it was linked to. Then we denote by $\mathcal{G}(t) = (\mathcal{N}, \mathcal{V}(t))$ the graph at time t, composed of the constant vertex set $\mathcal{N}$ and the set of active edges $\mathcal{V}(t)$ defined by: $\mathcal{V}(t) = \{l \in \mathcal{V} | l(t) = 1\}$.

The graph evolution is induced by the dynamic state evolution of each link $l \in \mathcal{V}$. We assume a cascade or contamination effect of links failures such that, when a link breaks, each of its neighbor links will break after an exponential time with rate δ. The duration of the link l failure follows an random exponential time with rate μ_l. In a way, μ_l is the recovery rate of the link l failure. This model is the Susceptible-Infected-Susceptible (SIS) epidemic process model [10]. We denote by $\mathcal{G}^\infty := \lim_{t\to\infty} \mathcal{G}(t)$ the graph obtained in stationary regime of the failures epidemic process. This epidemic process is well known in epidemiology but in our context the main difference is that the infection is over the links; which affects the global connectivity of the network.

C. Optimal resilient mechanism

Our goal is to control the cascading link failures in order to guarantee an average algebraic connectivity in the network when the failures epidemic is in his stationary regime. In fact, the stationary regime of the propagation failures epidemic is the worst graph in terms of connectivity. The control parameters are the repairing rates vector $\mu := (\mu_1, \ldots, \mu_E)$. We denote by $F(\mu)$ the cost function which is assumed to be separable, i.e. $F(\mu) = \sum_{l=1}^{E} F_l(\mu_l)$. Each edge-cost function $F_l(.)$ is monotone and continuously increasing in each μ_l. The optimization problem can be written as:

$$\max_{\mu}(a(\mathcal{G}^\infty(\mu)) - F(\mu)).$$

The stationary graph $\mathcal{G}^\infty$ may have two different features depending on the parameters and the epidemic threshold denoted by λ_c. There is no simple expression of this threshold for general graph structure [11]. In [12], the authors use mean-field approximation models to obtain an epidemic threshold in closed form characterized using a vector that makes the largest eigenvalue of a modified adjacency matrix equal to unity. Here we are interested in cases where repairing rates are not high enough so that failures will not cascade, since the network will not experience cascades if repairing rates are beyond a certain threshold, leading to trivial solutions. Hence, we look for solutions that is consistent with this assumption.

III. Resilient solution algorithm

The link state of the network $\mathcal{G}(t)$ can be represented as a vector $S(t) = (S_1(t), \ldots, S_E(t))$ where $S_l(t)$ is a random variable that denotes the state of edge l at time t. This process evolves as a continuous-time Markov process. We denote for each link l, the probability $s_l(t) := \mathbb{P}(S_l(t) = 1)$. This dynamics involve too many states and in order to compute the probability for any link to be broken at time t, requires the insights in the eigen-structure of the adjacency matrix of the links [13]. Then, computations based on the adjacency matrix are convenient only for small graphs. For complex graphs with a large number of nodes or edges, we use the N-Intertwined Mean Field Approximation (NIMFA) model proposed in [14].

Given the link adjacency matrix B (dual of the adjacency matrix A, i.e. $B_{ll'} = 1$ if and only if edges l and l' have a common node), and assuming that the number of links initially is large, we consider the following dynamics of the failure probability of a link l at time t:

$$\dot{s}_l(t) = -\mu_l s_l(t) + \delta(1 - s_l(t)) \sum_{l' \in \mathcal{V}} B_{ll'} s_{l'}(t). \qquad (1)$$

We can write in the following matrix form with $S(t) = (S_1(t), \dots, S_E(t))^T$:

$$\frac{dS(t)}{dt} = [\delta \mathrm{diag}(1 - S_l(t))B - \mathrm{diag}(\mu_l)]\, S(t),$$

with for any vector $x = (x_1, \dots, x_E)$, $\mathrm{diag}(x_l)$ is the diagonal matrix with elements $x_1, \dots, x_E$. In order to find the stationary regime of the epidemics, we have to solve the following non-linear system that comes from the stationarity of equations (1):

$$\forall l, \quad s_l^\infty = 1 - \frac{\mu_l}{\mu_l + \delta \sum_{l' \in \mathcal{V}} B_{ll'} s_{l'}^\infty}. \qquad (2)$$

We can use the following approximation of the stationary infection probability: $s_l^\infty \simeq 1 - \frac{\mu_l}{\mu_l + \delta k_l}$, $\forall l$, where k_l is the number of neighbors of link l (its degree in the link-adjacency matrix). This approximation is in fact an upper bound on the infection probability, then it gives a worst-case analysis. We obtain the following result expressed in this proposition.

Proposition 1: Given a worst-case failures propagation scenario, the optimal repairing rates $\mu^* = (\mu_1^*, \dots, \mu_E^*)$ are

$$\forall l = \mathcal{L}(i,j), \quad \mu_l^* = |v_i^\infty(\mu^*) - v_j^\infty(\mu^*)| . \sqrt{\frac{\delta k_l}{F_l'(\mu_l^*)}} - \delta k_l,$$

where v_i^∞ is the Fielder value of node i in the stationary regime of the epidemic.

In order to obtain the optimal repairing rate on each edge, we have to solve a fixed point system denoted by $\mu = H(\mu)$. The consistency test of our solution is based on the mean-field approximation model proposed in [12].

Definition 3: The solution $\mu^* = (\mu_1^*, \dots, \mu_E^*)$ is consistent with our problem if the largest eigenvalue of the symmetric matrix R is equal to one, with $R = \mathrm{diag}\left(\sqrt{\frac{\delta}{\mu_l^*}}\right) \cdot A \cdot \mathrm{diag}\left(\sqrt{\frac{\delta}{\mu_l^*}}\right)$.

If a solution is not consistent, then we consider the projection of the solution on the set μ^c that verify the critical threshold. We observe that the optimal repairing rate μ_i^* of edge l depends only on local knowledge for each node i. Specifically, the only information needed is the fielder value of all its neighbor nodes, which can be obtained by estimation or information exchange. Therefore, we are able to develop a fully distributed algorithm that computes the repairing rates in order to maximize the algebraic connectivity in the stationary regime of the failures propagation, given the cost function.

IV. Conclusion

This paper deals with a model of cascading link failures induced by cyber attacks, and we develop mechanisms to provide network resilience to this type of failures. The cascading failure model has been developed using a continuous-time epidemic process over the links of a network represented as a graph. Moreover, we have investigated the optimization of the global connectivity of the network, measured using the algebraic connectivity of the graph, under this epidemic failure dynamics. Finally, we have been able to build a distributed resilient mechanism that solves the optimization problem.

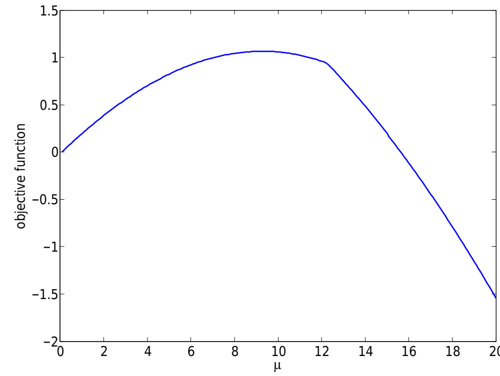

Fig. 2. Example of the objective function depending on μ in the homogeneous case, i.e. $\mu_l = \mu$ for all edge l and quadratic cost function $F(\mu) = 0.01\mu^2$, $N = 10$, $E = 15$ and $\delta = 1.5$.

References

[1] I. Dobson, B. A. Carreras, V. E. Lynch, and D. E. Newman. *Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization,* Chaos: An Interdisciplinary Journal of Nonlinear Science ,17, no. 2 (2007): 026103.

[2] P. Shakarian, H. Lei, and R. Lindelauf, *Power grid defense against malicious cascading failure,* In Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems (AAMAS '14), 2014.

[3] P. Chen, A. Hero, *Assessing and safeguarding network resilience to nodal attacks* in IEEE Communications Magazine, vol. 52, no. 11, 2014.

[4] M.M. Zavlanos, H.G. Tanner, A. Jadbabaie, and G.J. Pappas, *Hybrid control for connectivity preserving flocking*, IEEE Transactions on Automatic Control, vol. 54, no. 12, pp. 28692875, 2009.

[5] P. Di Lorenzo, S. Barbarossa, *Distributed Estimation and Control of Algebraic Connectivity over Random Graphs*, to appear in IEEE Transactions on Signal Processing, 2014.

[6] A. Ghosh, S. Boyd, *Growing Well-connected Graphs* in proceedings of CDC, 2006.

[7] S. Kar, J. Moura, *Sensor Networks With Random Links: Topology Design for Distributed Consensus*, in IEEE transactions on Signal Processing, vol. 57, no. 6, 2008.

[8] S. Hoory, N. Linial, and A. Wigderson. *Expander graphs and their applications*. Bulletin (New Series) of the American Mathematical Society, 43(4):439561, October 2006.

[9] M. Fiedler *Algebraic connectivity of graphs*. Czechoslovak Mathematical Journal, 23(98):298 305, 1973.

[10] D. Daley, and J. Gani *Epidemic modelling: an introduction*. Cambridge University Press, Cambridge, 1999.

[11] S. Ferreira, C. Castellano, and R. Pastor-Satorras *Epidemic thresholds of the Susceptible-Infected-Susceptible model on networks: A comparison of numerical and theoretical results*, in Physical Review E, vol. 86, 2012.

[12] P. Van Mieghem, J. Omic, *In-homogeneous Virus Spread in Networks*, in ArXiv 1306.2588, 2014.

[13] G. Duff and D. Naylor, *Differentiation Equations of Applied Mathematics*. New York: Wiley, 1966.

[14] P. Van Mieghem, J. Omic, and R. E. Kooij. *Virus spread in networks* IEEE/ACM Transactions on Networking, 17(1):114, February 2009.