

2018 ASEE Mid-Atlantic Section Spring Conference: Washington, District of Columbia Apr 6

The importance of assessment of vulnerability for improving the robustness of a computer network

Mr. Dilnesa T Nukuro, University of the District of Columbia

Dilnesa Nukuro was born and raised in Ethiopia. He studied Electrical and Computer Engineering at University of the District of Columbia and is planning to graduate in Spring 2018. His research interests include the application of wireless sensor networks and cybersecurity. He was awarded the IEEE Region 2 Project showcase 2nd place at New Jersey in 2017.

Dr. Paul Cotae, University of the District of Columbia

Dr. Paul Cotae, Associate Professor of Electrical and Computer Engineering has more than 25 years of experience in the communication field (research and education). He received a Dipl. Ing. and a M.S. degrees in communication and electronic engineering in 1980 from the Technical University of Iassy and a Ph.D. degree in telecommunications from "Politechnica" University of Bucharest, Romania in 1991, and a Master in Applied Mathematics in 1998 from the University of Colorado at Boulder. From 1994 to 1998 he spent four years at the University of Colorado at Colorado Springs and the University of Colorado at Boulder as a Fulbright Scholar and Visiting Associate Professor doing research and teaching in ECE department and APPM department. He served also as a consultant to Navsys Corp., Colorado Springs, in 1997. From 2002 to 2008 he was with the Department of Electrical and Computer engineering at the University of Texas at San Antonio (UTSA). From 1984 to 2001, he was with the Department of Electrical Engineering, Technical University of Iassy, where he conducted research and teaching in the area of digital communications as a Full Professor at the same department. Since 2008, he has been with the University of the District of Columbia as an Associate Professor. His current research interests include multiple access, modulation and coding, mobile communications, and digital communication systems. He has authored or coauthored more than 100 papers in these areas and four books. Dr. Cotae serves as an Associate Editor for IEEE Communication Letters, EURASIP Journal on Signal Processing, Elsevier International Journal of Computers and Electrical Engineering (JCEE), and he has been on the Technical Program Committee and Session chair of the IEEE Conferences GLOBECOM (2003-2011), VTC Spring 2005, 2006 and ICC 2005-2011. He is a Senior Member of IEEE, member of ASEE, member of HKN (Eta Kappa Nu) and SIAM. He is cited in Who's Who in American Education, Who's Who in America, and in Who's Who in the World. He has been nominated two times for the best teacher award (2005 and 2006) in the College of Engineering at UTSA. He is the IEEE Vice Chair Washington Section and IEEE ComSoc chair Whashington DC Section. He is Faculty Fellow for the ONR-ASEE Summer Faculty Research Program 2009-2011.

The importance of protecting vulnerability to improve the robustness of a computer network

Dilnesa Nukuro, Nguyen Quoc, Paul Cotae

dilnesa.nukuro@udc.edu, quoc.nguyen@udc.edu, pcotae@udc.edu

University of the District of Columbia

Department of Electrical and Computer Engineering

4200 Connecticut Ave, Washington, DC, 20008, USA

Abstract

Robustness and resilience are often thought of in terms of a system's capacity to maintain functionality in the face of external perturbations. Robustness is one of the major issues for complex networks. The robustness of the network is evaluating if the network's normal functions are affected in case of an external perturbation. Improving the robustness of any network system requires analysis of its vulnerability to external perturbations¹. The outcome of the analysis would be the solution for random failures or adversarial attacks happening to an element of the network. The main focus of this paper is to study the evolution of evaluating the robustness of complex networks, specifically, the vulnerability of the clustering of the network to the failure of the network element. Our specific focus is to identify vertices whose failure will cause critically damage to the network by corrupting its clustering³. Identifying the real problem is key to discovering the solution because any significant change made to the clustering, resulting from element-wise failures, could reduce network performance. By using the mathematical algorithms, we can formulate the vulnerability analysis as an optimization problem, prove its NP-completeness and non-monotonicity, and the algorithms that we formulate will identify the vertices most important to clustering.

Keywords: Robustness, Vulnerability, Network, Algorithms, and complex network

Introduction

Most complex real-world systems attract the attention of the many engineers studying a wide range of fields such as power grid systems, computer science, and social science³. There is research done on the structural and dynamic properties of real-world networks, and most networks are found to show a power-law degree distribution and scale-free. The vulnerability of a network characterizes its inability to withstand the effect of node or link failures. The robustness is the ability of a network to remain functional after initial attack either. The robustness can be characterized by the integral size of the connected component during a whole attacking period. The percolation threshold is the critical fraction of the remaining nodes or links that lead to the collapse of the network, which is usually predicted by using a statistical physics method call percolation theory⁵. The network attack and failure have growing fast time to time. Robustness is one of the major issues for the complex network, such as the World Wide Web, transportation network, communication network, biological networks, and social information network. The robustness of the network evaluates the network normal function is affected in case of external perturbation. The Vulnerability is usually used for to indicate the lack of robustness and resilience of the complex system. To improve the robustness of the real world system, it is important to obtain key insight into structural vulnerabilities of the network representing them. The reason is

that to analyze and understand the effect of the failure of individual components on the level of clustering in the network. Using clustering is the major property of the network.

METHODOLOGY: COMPLEX NETWORK AND RELATED RESEARCH.

A complex network has several ways to represent different network methods. Our methodology is represented by a graph G with N node, M edges⁴. The purpose of the methodology is to show the steps of the complex network. Those are measuring the robustness and manipulating the robustness of a network. Our one of the methods is measuring the robustness by using metrics which is graph connectivity, the diameter relative size of largest components, and an average size of the isolated cluster. There several related has done with this methods, however, ours improving their methods⁵. The related paper has done such as graph connectivity (Dinh et al. 2012b), Laplacian matrix (Fiedles, 1973), and using graph percolation (Collaway et al, 2000). Network vulnerabilities can be quantified based on pair-wise connectivity (Dinh et al), Eigenvector (Allesina and Pascual, 2009), geodesic length (Holme et al., 2002), etc. Another possible measurement relies on the average clustering coefficient proposed by Watts and Strogatz. The Fast Adaptive Greedy Algorithm developed by Kuhnle et al is novel and promising. For a 10 node simple network, the Fast Adaptive Greedy Algorithm will identify the most important vertices whose failure will cause maximum degradation of network clustering. Further investigation will have to be done for a 100 and 500 node network.

Table for Notations	
N	Number of vertices/nodes ($N = V $)
M	Number of edges/links ($M = E $)
d_u	The degree of u
$N(u)$	The set of neighbors of u
$T(u)$	The number of triangles containing u
$C(u), C(G)$	Clustering coefficients of u and G
$\tilde{C}_v(u), \tilde{C}_v(G)$	Clustering coefficients of u and G after removing node v from
$G[S]$	The sub-graph induced by $S \subseteq V$ in G
$tr(u, v)$	The number of triangles containing both u, v .

Figure: 1 Notation.

The Algorithm and Methodology

In the following, the explanation of the algorithms will be explaining. This algorithm is used as a MATLAB code and we are going to give detail notations.³⁻⁴ The first set of notion explain the graph representing a complex network which is $G=(V, E)$. V is for the set of N nodes and E is the set of edge containing M connections. Node $u \in V$, denote by d_u and $N(u)$ the degree of u and the set of u 's neighbors, respectively. Now, we are going to over Triangle free graphs A graph and G is said to be triangle-free if no three vertices of G form a triangle of edges. This help to verify that the graph G is triangle-free or not is tractable by computing the trace of A , where A is the adjacency matrix of G . Next is about clustering measure functions. For that the flowing is given a node $u \in V$, there are d_u adjacent vertices of u in G and there are $d_u(d_u - 1)/2$ possible

edges among all u's neighbors. In this area, a couple of references are used to check the latest matrix multiplying result which is (Gall, 2014).

The local clustering coefficient $C(u)$ is defined:

$$C(u) = \begin{cases} \frac{2T(u)}{d_u(d_u-1)} & d_u > 1 \\ 0 & otherwise \end{cases} \quad (1)$$

T is the number triangle containing u . It is clear that $0 \leq C(u) \leq 1$ for any $u \in V$ for any node $v \in V$, let $C_v(u)$ denote the clustering coefficient of u in $G[V \setminus \{v\}]$. Finally, define $tr(u, v)$ as the number of triangles containing both vertices u and v . In the above we have seen that how the clustering coefficients works, now we are going to see the algorithms of average clustering coefficient. Theoretically known that the average local clustering coefficient $C(G)$ of a graph G is a measure indicating how much the vertices of G tend to cluster together. This is because $0 \leq C(u) \leq 1$ for every node $u \in V$, $C(G)$ is also normalized and can only take the value in the range $[0,1]$ inclusively. For example, $C(G) = 0$ when G is a triangle-free graph and $C(G) = 1$ when G is a clique or a collection of cliques.³⁻⁴ The higher the clustering coefficient of G the more closely the graph locally resembles a clique.

$$C(u) = \begin{cases} \frac{2T(u)}{d_u(d_u-1)} & d_u > 1 \\ 0 & otherwise \end{cases} \quad C_i = \frac{|\{e_{uk}: v_j, vk \in N_i, e_{jk} \in E\}|}{k_i(k_i-1)}$$

$$2T(u) = \sum_{i \in V} \sum_{j \in V} e_{ui} e_{uj} e_{ij} \quad C = \frac{1}{n} \sum_{i=1}^n C_i \quad C(G) = \frac{1}{N} \sum_{u \in V} C(u)$$

$$C_v(G) = C(G[V \setminus \{v\}])$$

When we combine in one:

$$C(u) = 1 + \frac{nx}{1!} + \frac{n(n-1)x^2}{2!} + \dots \quad C(G) = \frac{1}{N} \sum_{u \in V} C(u)$$

Fast Adaptive Greedy

In the following, we are going to formulate the CSA problem as an integer program. Let $(e_{ij})_{i,j \in V}$ be the adjacency matrix of G .

Lemma 1. For $u \in V$, $T(u)$ can be calculated in the following way:

$$2T(u) = \sum_{i \in V} \sum_{j \in V} e_{ui} e_{uj} e_{ij} \quad (2)$$

Proof: The summand $e_{ui} e_{uj} e_{ij} = 1$ iff i, j are neighbors of u , and if edge (i, j) is in the graph; that is, vertices u, i, j form a triangle. We formulate CSA as an integer program in the following way. Let $x_i = 1$ if i is included in the set S , and $x_i = 0$ otherwise.

Integer Program 1. Min $\sum_{u \in V: d(u) > 1} \sum_{i \in V} \sum_{j \in V} \frac{e_{ui} e_{uj} x_i x_j x_u}{d_u(d_u-1)(N-K)}$
Such that $\sum_{u \in V} x_u < k$,

Notice that sum 2 compute the ALCC of the residual graph after removing S. As we see in above section, corollary 1 there always exists as a node the removal of which will not increase the ALCC; thus an optimal solution to the program is an optimal solution to CSA.

Lemma 2. In a graph G, the following statements hold:

1. $C(G) = 0$ if and only if G is a triangle free network.
2. $C(G) = 1$ if and only if G is a clique or contain only separated cliques.

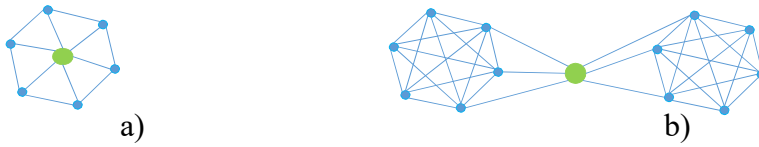


Figure: 2 Nonmonotonicity of ALCC

- (i) Suppose there exists a triangle u, v, w in G . Then $C(u) > 0$, so $C(G) > 0$. For the converse, if $C(G) > 0$, there exists $u \in V$ such that $C(u) > 0$. By definition of $C(u)$, there exists a triangle u, v, w containing u .
- (ii) Suppose $C(G) = 1$. Then, for each $u \in V$, the subgraph induced by $\{u\} \cup N(u)$ is a clique, from which G is a clique or only separated cliques. The converse follows directly from the definition of $C(G)$.

Lemma 3. For any $u \in V$ $2T = \sum_{v \in N(u)} |N(u) \cap N(v)|$ (3)

Proof: For each neighbor v of u , the number of triangles that contain both u and v is $|N(u) \cap N(v)|$. Since each triangle containing u contains exactly two neighbors of u , it follows that the summation $\sum_{v \in N(u)} |N(u) \cap N(v)|$ counts twice the number of triangles containing u .

Lemma 4. For any node

$$\frac{1}{N-1} \sum_{v \in V \setminus \{u\}} C_v(u) \leq C(u). \quad (4)$$

Proof: To prove this Lemma, we will show the following statements regarding the degree of u :

$$\frac{1}{N-1} \sum_{v \in V \setminus \{u\}} C_v(u) \leq C(u). \text{ when } d_u \leq 2. \quad (5)$$

$$\frac{1}{N-1} \sum_{v \in V \setminus \{u\}} C_v(u) \leq C(u). \text{ when } d_u > 2 \quad (6)$$

Eq. (3) is equivalent to $\sum_{v \in V \setminus \{u\}} C_v(u) \leq (N-1) C(u)$.

To find $\sum_{v \in V \setminus \{u\}} C_v(u) \leq (N-1) C(u)$. we use the fact that removing a non-neighbor node of u will not affect the local clustering coefficient $C(u)$, i.e., $\tilde{C}_v(u) = C(u)$ for $v \in V \setminus (N(u) \cup \{u\})$.

There are $(N - d(u) - 1)$ non-neighbors vertices of u in G . Thus the second term of (6) follows. To evaluate the first term of Eq. (5).

Lemma 5. Let $N_2(u) = \{v \in N(u) : d(v) = 2\}$, $N_{>2}(u) = \{v \in N(u) : d(v) > 2\}$. For each $u \in V$, $\Delta C_u(g)$ can be computed in the following way:

$$N_u = \frac{2T(u)}{Nd(u)(d(u)-1)} + \sum_{v \in N_2(u)} \frac{4T(v)(1-N) + 2tr(u,v)Nd(v) - 2T(v)d(v)}{N(N-1)d(v)(d(v)-1)(d(v)-2)} + \sum_{v \in N_{>2}(u)} \frac{T(v)}{N} \quad (7)$$

Proof: Denote the contribution of $v \in G$ to the average clustering coefficient as c_v before the removal of u and \hat{c}_v after.

$\Delta \tilde{C}_u$ can be written as $\sum_{v \in G} c_v - \hat{c}_v$. If $v \in N(u) \cup \{u\}$, then $c_v = \hat{c}_v$. If $v = u$, then

$$C_v - \hat{C}_v = \frac{2T(u)}{Nd(u)(d(u) - 1)}$$

Let $v \in N_{>2}(u)$. Then before removal of u , v is in $T(v)$ triangles. After removal, v is in $T(v) - tr(u, v)$ triangles. Hence

$$C_v = \frac{2T(u)}{Nd(u)(d(u) - 1)}$$

And

$$C_v = \frac{2T(u)}{(N-1)(d(v)-1)(d(v)-2)}$$

hence

$$C_v - \hat{c}_v = \frac{4T(v)(1-N) + 2tr(u,v)Nd(v) - 2T(v)d(v)}{(N-1)(d(v)-1)(d(v)-2)} \quad (8)$$

$$\sum_{v \in N_{>2}(u)}$$

Let $v \in N_2(u)$. Before removal of u , v is in $T(v)$ triangles. After removal, v is in 0 triangles, hence the result follows. One important feature of FAGA is that the produced residual ALCC values will form a non-increasing sequence. Algorithm 2 Fast Adaptive Greedy Algorithm (FAGA - fast greedy) Number the vertices

from 1 to N such that $u < v$ implies $d(u) \leq d(v)$. $S \leftarrow \emptyset$; for each $u \in V$ do $T(u) \leftarrow 0$; end for for each $(u, v) \in E$ do $tr(u, v) \leftarrow 0$; end for for $u \leftarrow n$ to 1 do for each $v \in N(u)$ with $v < u$ do for each $w \in A(u) \cap A(v)$ do Increase $tr(u, v)$, $tr(v, w)$ and $tr(u, w)$ by one; Increase $T(u)$, $T(v)$ and $T(w)$ by one; $d(u) \leftarrow A(v)$; end for end for end for for $i \leftarrow 1$ to k do for each $u \in V \setminus S$ do $N_d(u)(d(u)-1)$ for each $v \in N(v) \setminus S$ do if $d(v) > 2$ then end if if $d(v) = 2$ then $\Delta \tilde{C}_u \leftarrow \Delta \tilde{C}_u + T(v)/N$; end if end for end for $u_{max} \leftarrow \arg \max_{u \in V \setminus S} \{\Delta \tilde{C}_u\}$; Remove u_{max} from G , add u_{max} to S , and decrease N by one for each $(v, w) \in E$ and $v, w \in N(u_{max}) \setminus S$ do Decrease $T(v)$ and $T(w)$ by one; end for end for return S .

The graph and Complex Clustering Network.

The graphs shows that some connection in a small group. By setting $maxNode$ to be greatest of the node we want to consider, say 10, we can look at the first 10 nodes and their

connection but disregard all connections involving nodes which fall outside of the first 10.²⁻⁴ As we described, the MATLAB has a function which only requires the input of an adjacency matrix and automatically positions the nodes on the graph with coordinated that minimize crossing of the lines and clearly show the feature of the graph. Adjacency Matrix is a square matrix representing a simple graph of n vertices. Each element a_{ij} is 1 when there is an edge from vertex i to vertex j , and zero when there is no edge. An adjacency matrix can be used to represent a complex network consisting of clusters. Given a simple network of 10 nodes (vertices) in figure 1, the corresponding adjacency matrix is computed below:

```

0 1 1 1 1 0 0 0 0 0
1 0 0 0 0 1 1 1 0 1
1 0 0 0 0 1 1 1 1 1
1 0 0 0 0 1 1 1 1 0
1 0 0 0 0 0 0 0 0 1
0 1 1 1 0 0 1 1 1 0
0 1 1 1 0 1 0 1 1 1
0 1 1 1 0 1 1 0 0 0
0 0 1 1 0 1 1 0 0 0
0 1 1 0 1 0 1 0 0 0

```

Figure: 3 Adjacent Matrix of Simple node network.

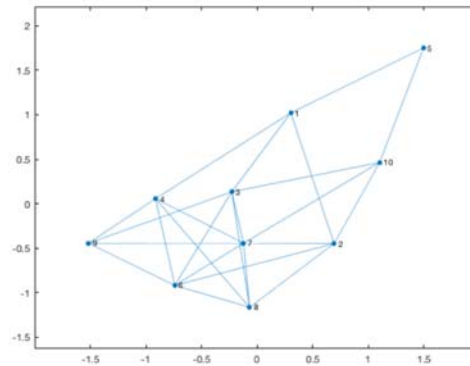


Figure 4: Simple 10 node network

Conclusions

This paper proposes robustness and resilience in term of a system's capacity to maintain functionality in the face of external perturbations. Because robustness is a major issue for the complex network. The issues of the robustness analyzed and the outcome of the analysis would be the solutions for random vulnerability to external perturbations. Main focused of the evolution was the vulnerability of the clustering of the network to the failure of the network element. Clustering vulnerability is an important aspect in assessing the robustness of complex networks, as the level of clustering has significance for a variety of applications, including a silent role in the propagation of information in a social network.⁵⁻⁶

Acknowledgment

This research was funded by the Army Research Office (ARO) - Department of Defense (DOD)–Award No. W911NF-15-1-0481.

Reference

1. Milanovic, Jovica V., and Wentao Zhu. “Modelling of Interconnected Critical Infrastructure Systems Using Complex Network Theory.” IEEE Transactions on Smart Grid, 2017, pp. 1–1.,doi:10.1109/tsg.2017.2665646.
2. Buldyrev, Sergey V., et al. “Catastrophic Cascade of Failures in Interdependent Networks.” Nature, vol. 464, no. 7291, 2010, pp. 1025–1028., doi:10.1038/nature08932.

3. Huang, Xuqing, et al. "The Robustness of Interdependent Clustered Networks." *EPL (Europhysics Letters)*, vol. 101, no. 1, 2013, p. 18002., doi:10.1209/0295-5075/101/18002.
4. Albert, "Error and Attack Tolerance of Complex Networks." *The Structure and Dynamics of Networks*, 2011, doi:10.1515/9781400841356.503.
5. Wang, Xiao Fan, and Guanrong Chen. "Synchronization in Scale-Free Dynamical Networks: Robustness and Fragility." *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 1, 2002, pp. 54–62., doi:10.1109/81.974874.
6. Kuhnle, Alan. "Vulnerability of clustering under Node failure in Complex Network." SpringerLink. Springer Vienna, 01 Mar. 2017. Web. 15 Mar. 2018.

Bio

Dilnesa Nukuro

Dilnesa Nukuro was born and raised in Ethiopia. He studied Electrical and Computer Engineering at University of the District of Columbia and is planning to graduate in Spring 2018. His interests include the application of wireless sensor networks and cybersecurity. He was awarded the IEEE Region 2 Project showcase 2nd place at New Jersey in 2017.

Dr. Paul Cotae

Dr. Paul Cotae, Professor of Electrical and Computer Engineering is the Director of the PhD Program at SEAS and Director of the SEAS Research Center. His research is in Digital Communication, Information theory, Statistics and Applied Mathematics and Cybersecurity: Anomaly detection, Detection of Low Rate Denial of Service Attacks, Intrusion Detection, Information Visualization. He published more than 140 conference and journal papers, many of them at IEEE level, authored 2 books and coauthored 3 books in the area of digital communications systems. During the AY 2014-2015 he spent his sabbatical at the Center for High Assurance Computer Systems Code 5540, Naval Research Laboratory, Washington DC, 20375. Since 2009 he has been selected every summer as ONR Senior Research Fellows for the ASEE Summer Faculty Research Program at NRL. His research is sponsored by NSF, ONR, AFOSR and USAF. He received in last five years more than \$1M for his research from DOD as a sole PI for the following grants: –Army Research Office (ARO) –Award No. W911NF-15-1-0481: "Performance Data-Driven Methods and Tools for Computer Network Defense through Network Science" and Office of Naval Research - Award no. W911NF-11-1-0144 "Information-Driven Blind Doppler Shift Estimation and Compensations Methods for Underwater Wireless Sensor Networks".

Quoc A Nguyen

Electrical Engineering major at the University of the District of Columbia graduating in 2019. His interest includes Cybersecurity, Solid States Physics using Faraday's Balance and Vibrating Sample Magnetometer, MATLAB, and Artificial Intelligence.