Information Systems Frontiers
https://doi.org/10.1007/s10796-024-10477-y

Securing Network Resilience: Leveraging Node Centrality for Cyberattack Mitigation and Robustness Enhancement

Essia Hamouda[1] · Mohsen ElHafsi[2] · Joon Son[1]

Accepted: 19 February 2024

Abstract

In response to the dynamic and ever-evolving landscape of network attacks and cybersecurity, this study aims to enhance network security by identifying critical nodes and optimizing resource allocation within budget constraints. We introduce a novel approach leveraging node centrality scores from four widely-recognized centrality measures. Our unique contribution lies in converting these centrality metrics into actionable insights for identifying network attack probabilities, providing an unconventional yet effective method to bolster network robustness. Additionally, we propose a closed-form expression correlating network robustness with node-centric features, including importance scores and attack probabilities. At the core of our approach lies the development of a nonlinear optimization model that integrates predictive insights into node attack likelihood. Through this framework, we successfully determine an optimal resource allocation strategy, minimizing cyberattack risks on critical nodes while maximizing network robustness. Numerical results validate our approach, offering further insights into network dynamics and improved resilience against emerging cybersecurity threats.

Keywords Centrality measure · Node detection · Performance evaluation · Optimization · Network security

1 Introduction

With the growing dependence on interconnected technologies, cybersecurity threats are becoming increasingly prevalent. Cyberattacks typically target specific nodes that we refer to as critical nodes in communication networks. These nodes are generally vulnerable components of the network and are critical for protecting the network infrastructure and performance.

Attackers typically exploit network vulnerabilities before targeting critical nodes. For instance, if an attacker successfully infiltrates essential Internet providers, they can potentially disrupt global internet services, resulting in widespread outages and significant disruptions to various online services and businesses. Similarly, infrastructure networks, such as the Internet Backbone, Internet of Things (IoT) networks, and Power Grids, are integral to our daily lives. They are enticing targets for cyberattacks due to their historical lack of robust security measures, the substantial impact that their compromise can have, and the inherent difficulty in patching entire networks due to their large scales.

Critical nodes can be the target of protective measures and defensive monitoring for beneficial purposes. An unusual traffic pattern in a network could indicate that a computer or a critical node has been hacked, and data is being transmitted to unauthorized destinations, or that a computer is being subjected to a denial-of-service attack. Therefore, to ensure normal network functions are preserved, it is essential to explore network anomalies and vulnerabilities by identifying critical nodes, proactively monitoring them, and implementing appropriate security measures to prevent or mitigate such attacks.

The critical node detection problem is fundamental to a variety of interesting applications. For example, the design of network routing applications is often based on the selection of a set of nodes that form a path that connects a source

✉ Essia Hamouda
ehamouda@csusb.edu

Mohsen ElHafsi
melhafsi@ucr.edu

Joon Son
json@csusb.edu

[1] Information and Decision Sciences Department, California State University San Bernardino, University Parkway, San Bernardino 92407, CA, USA

[2] School of Business Administration, University of California Riverside, University Avenue, Riverside 92521, CA, USA

Published online: 01 March 2024

and a destination. Hence data transmission's security is proportional to the criticality of the selected nodes. Unless the nodes are secured, the higher the number of critical nodes are in the path, the more the application is vulnerable. Moreover, compromising even one node in the path, such as the router, the hub or the switch in Fig. 1, may render the network nonfunctional. Hence detecting and analysing critical nodes is worth taking into account when designing secured network applications.

While several studies have proposed various methods for identifying critical nodes in networks (Shen et al., 2013; Lalou et al., 2018a; Arulselvan et al., 2009a), it is important to acknowledge that not all of these measures are universally effective, especially when dealing with specific types of network vulnerabilities. This is particularly evident when network fragmentation is a high priority, as illustrated in Fig. 1, where the compromise of a router, for example, results in the fragmentation of the network into two distinct components.

Furthermore, it is worth noting that some of these identification methods rely on solving intractable NP-hard optimization problems. While these approaches may provide valuable insights, they can also be computationally expensive and may not be practical in all situations (Lalou and Tahraoui, 2018; Arulselvan et al., 2009b; Walteros et al., 2019).

Graph-based features like centrality (Das et al., 2018) have proven to be an effective method for critical node identification. However, in regards to communication networks, there are still some gaps. Some researchers focus exclusively on node centrality measures, particularly betweenness and eigenvector centrality, without exploring other types of centrality measures. In contrast, other researchers (Mitchell et al., 2019) emphasize the significance of edge centralities and their potential for network data analytics and event detection. However, it is essential to recognize the underlying assumptions made by the centrality measures, such as how information moves and replicates as it spreads across the network. Unfortunately, these assumptions are often overlooked.

This paper focuses on communication network and introduces a novel approach that leverages insights derived from the computation of node centrality scores using four widely-recognized centrality measures. What distinguishes our work is its unique ability to convert these centrality metrics into actionable probabilities for predicting network attack. Furthermore, we introduce a closed-form expression for measuring network robustness, a significant contribution that has reshaped our understanding of network security. We further modeled the network security problem as a nonlinear optimization problem, subject to a budget constraint. In particular, we solved a resource allocation problem aimed at minimizing the probability of cyberattacks on critical nodes while simultaneously maximizing network robustness. Contrary to existing research, we integrated the results obtained from four widely used centrality methods to gain a holistics view of the network, its probability of an attack and its robustness.

The remainder of this paper is organized as follows.

Section 2 describes existing research within this domain and explores its practical applications. Section 3 provides an overview of the network model utilized in our study, along with its underlying assumptions. Section 4 offers a concise explanation of four centrality methods, which are fundamental measures of node importance in network analysis. Section 5 outlines various network properties that are central to our analysis, shedding light on the network's key characteristics and behaviors. Section 6 introduces a network optimization model designed to provide insights into strategies that enhance network security. Section 7 discusses the integration of the four critical node centrality methods into the network, demonstrating their role in bolstering network security and offering practical implementation insights. Section 8 summarizes our research findings and outlines future research directions of this work.

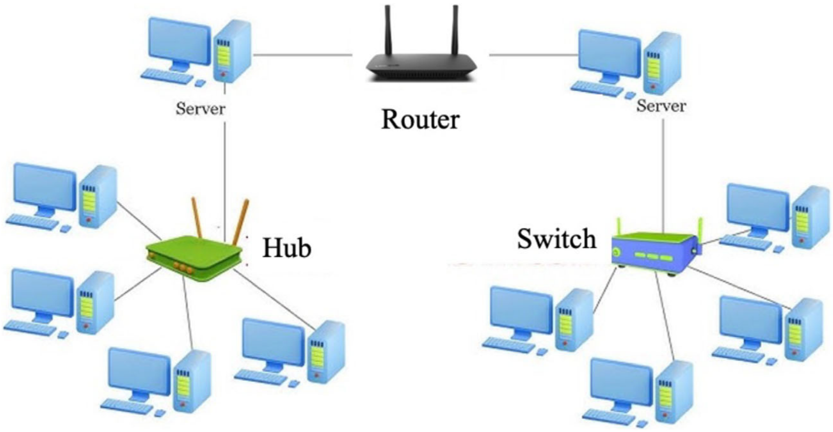


Fig. 1 Communication network where the removal of a node such as the router or the hub is sufficient to disconnect the network

2 Related Work and Applications

The identification of critical nodes has attracted significant attention in various fields, including its application in social networks for anomaly detection. A range of techniques (Li et al., 2022; Rajalakshmi et al., 2023; Helmi et al., 2021; Ganguli et al., 2020; Zaki et al., 2023; Riquelme and Vera, 2022), have been utilized to address this problem.

While our primary focus in this paper centers on communication networks, network vulnerability, and security issues, it is crucial to acknowledge the broad relevance and importance of this problem in other domains. The critical node identification problem has found applications in different fields such as transportation (Gupta et al., 2023), social networks (Mazlumi and Kermani, 2022), biology (Liu et al., 2020), public health (Alozie et al., 2022), fraud detection (Ke et al., 2022), intrusion detection (Dang et al., 2023), image processing (Zhang et al., 2022), and astronomical data analysis (Ahmed et al., 2016; Lalou et al., 2018b).

In the realm of network vulnerability assessment, much of the existing research has focused on centrality measurements, including degree, betweenness, closeness centralities, and average shortest path length (Freeman, 1978; Devkota et al., 2018; Kivimäki et al., 2016). These metrics have been instrumental in understanding network structures. For instance, recent work (Kim, 2020) showcased the practicality of centrality measures in identifying central nodes involved in malware distribution, offering insights into distinguishing critical (*i.e.*, malicious) nodes from non-critical (*i.e.*, benign) nodes.

Research efforts (Zheng et al., 2017; Wang et al., 2018) have employed graph measurements to identify critical nodes. These studies primarily relied on metrics such as average similarity and global clustering coefficients. They found that critical nodes tend to exhibit higher similarity and tighter clustering, providing valuable insights into network vulnerability. It is important to highlight that these studies did not comprehensively analyze the wide range of structural differences that may exist between critical and non-critical nodes.

In telecommunication networks, the identification of critical nodes holds paramount importance, serving both defensive and offensive purposes (Proselkov et al., 2021; Alozie et al., 2021). These nodes play a dual role, either as essential components to preserve the functionality of a communication network or as prime targets for disruption in adversarial contexts. Consider the case of terrorist and insurgent networks (Rains, 2022; Ballinger, 2023), where the primary objective is to serve the communication channels by strategically removing critical nodes to disable terrorist networks (Arulselvan, 2009).

In the context of wireless communication networks, work in Commander et al. (2007) formulated the Wireless Network Jamming Problem as a critical node identification challenge. In this scenario, the goal is to pinpoint critical nodes that, when jammed, effectively neutralize an adversary's wireless communication network.

In sensor networks (Imran et al., 2013; Shukla, 2023), beyond its role in anomaly detection, critical node identification plays a crucial role in optimizing energy utilization and extending the operational lifespan of these nodes (Mitton et al., 2009). Critical nodes, often traversed by many shortest paths, can experience faster energy depletion, potentially leading to network fragmentation (Gouvy et al., 2012).

In decentralized systems such as peer-to-peer and adhoc networks (Jain and Reddy, 2013; Hamouda et al., 2011), a major weakness is network disconnectivity. These systems typically have a weak topology that can be easily fragmented by targeting critical nodes, which maintain the network's entire connectivity. Hence, the identification of these critical nodes is essential for designing robust and secure applications (Xing et al., 2023).

The works in Dinh and Thai (2011); Dinh et al. (2010); Veremyev et al. (2015) have underscored the significance of identifying critical nodes in network vulnerability assessment. They converted network vulnerability assessment into the problem of identifying critical nodes, measuring network vulnerability by the minimum number of nodes whose removal disrupts the network pairwise connectivity. Building on this foundation, research in Shen et al. (2013, 2012b, a) expanded this concept, assessing network vulnerability across various graph types, including unit-disk graphs, power-law graphs, and dynamic graphs.

Research in Yan et al. (2011); Invernizzi et al. (2014) delved into the study of malware distribution networks using centrality metrics and empirical approaches (Kim et al., 2018). Others (Lozano et al., 2017; Faramondi et al., 2017, 2016; Lalou et al., 2018b) have framed the critical node identification problem as an optimization problem. This entails the identification of a set of nodes whose removal would significantly degrade network connectivity based on predefined metrics. While these problems can be computationally demanding, researchers have explored various approaches, including dynamic programming and integer linear programming (Ugurlu, 2022; Megzari et al., 2023; Laha et al., 2020), to seek exact solutions. Moreover, there have been efforts to devise approximation solutions with performance guarantees, employing heuristic algorithms and polynomial-time approximation algorithms (Aringhieri et al., 2016; Ventresca and Aleman, 2014; Berger et al., 2014).

Critical nodes in complex systems need to be identified for protection or removal. Removal of critical nodes decreases or minimizes a system's ability to diffuse entities such as information, goods, or diseases (Yang and An, 2020). Previous research suggested some vulnerability metrics, but there remains a lack of understanding how a metric changes (*i.e.*, upper bound and lower bound) and how it is related to the structure of a complex system (Johnson and Hogan, 2013).

3 Network Model Description

We model a communication network as an undirected graph denoted by $G = (V, E)$, as depicted in Fig. 2 (Ryan, 2015). In this representation, V corresponds to the set of nodes, encompassing various components such as routers, switches, and similar elements, with cardinality $|V| = n$, and E denotes the set of edges or links ($E \subset V \times V$), with cardinality $|E| = m$. For each node $i \in V$ ($i \in \{1, \dots 32\}$ is the node identity in Fig. 2), we define $N(i) = j \in V | (i, j) \in E$ as the neighborhood set of node i. Let A denote the graph's adjacency matrix of size $n \times n$, (*i.e.* $A_{ij} = 1$ if there is a link between nodes i and j, otherwise $A_{ij} = 0$).

In communication networks, it is essential to acknowledge the variability in node importance. Specific nodes, such as routers, occupy central roles, and their removal could lead to the loss of network connectivity and functionality. Conversely, centrally positioned nodes, such as servers, often store significant data. In contrast, peripheral or end nodes generally play a lesser role in network operations.

Our primary objective in this study is to identify and categorize nodes within the network (or graph), distinguishing between those classified as important (or critical) and those of lesser significance. The definition of node importance varies based on the application (*e.g.*, Yin et al., 2019; Deng, 2019; Yong et al., 2022; Shi et al., 2016). In our case, we formalize node importance according to the following definition.

Definition 1 In the context of communication networks, we classify a node as important if it meets any of the following criteria: (1) It generates, receives, or transmits a significant volume of network traffic. (2) It functions as a repository for a substantial amount of sensitive data. (3) it plays an integral role in critical communication paths, such as the shortest path.

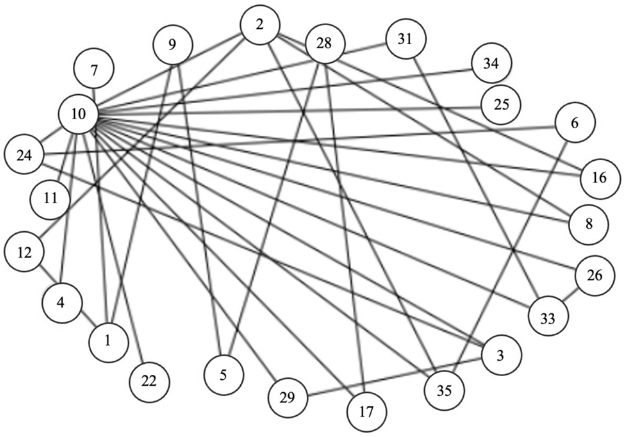


Fig. 2 Network depicted as a graph $G = (24, 32)$

The concept of centrality plays a pivotal role in graph analysis, as highlighted in references such as (Zverovich, 2021; Majeed and Rauf, 2020). It serves as a fundamental tool for identifying critical nodes within a network. Various centrality methods have been developed, each shedding light on a unique aspect of node importance, as discussed in Section 4. In our study, we delve into four distinct centrality methods to gain deeper insights into the network. Each of these methods quantifies the importance of a node i using a measure denoted as $C(i)$, generally referred to as its *centrality score*. Consequently, we refine our Definition 1 of node importance as follows:

Definition 2 A node i is deemed important if its centrality measure $C(i)$ surpasses a predefined threshold $\beta > 0$.

4 Centrality Measures

Below, we offer a concise overview of four commonly employed centrality methods and their interpretation within the context of communication networks.

4.1 Degree Centrality

Node degree centrality (Golbeck, 2013; Rodrigues, 2019; Das et al., 2018) is a metric that quantifies a node's direct connections within a network. Mathematically, the degree centrality score of node i can be computed as: $C_d(i) = \frac{\sum_{j=1}^{n} A_{ij}}{n-1}, i, j \in V$.

In communication networks, This metric is frequently used to identify highly connected nodes, offering insights into information flow patterns and potential vulnerabilities of critical nodes. Nodes with a high degree centrality score are extensively interconnected, enabling rapid information dissemination throughout the network, often making them information hubs. However, such nodes may also be susceptible to attacks or malware, potentially serving as weak nodes. While their removal may not critically impact the overall network performance, it can disrupt communication and information flow.

Despite its simplicity and efficiency, degree centrality does not consider the network's global structure, hence its complexity scales at $O(m)$ (Das et al., 2018). Therefore, a node with fewer highly important neighbors might possess greater importance than a node with numerous less important neighbors.

4.2 Closeness Centrality

Node closeness centrality (Yen et al., 2013; Fernandes et al., 2023; Das et al., 2018; Rodrigues, 2019) is a measure of how close a particular node is to all other nodes in a network. It is defined as the inverse of the sum of the shortest path distances between a given node and all other nodes in the network. Mathematically, the closeness centrality score of node i can be calculated as: $C_c(i) = (n-1)/\sum_i^j d_{ij}$ where d_{ij} is the shortest path distance between nodes i and j.

In communication networks, closeness centrality proves invaluable for identifying potential malware vectors. Elevated closeness centrality scores suggest nodes with swift and efficient information transmission, including the propagation of malware to other network nodes. Nodes exhibiting significantly lower closeness centrality scores than expected, given the network's overall structure, may indicate issues such as communication link failures or node isolation caused by malware.

Closeness centrality offers a distinct advantage as it relies on a comprehensive view of the network, rendering it highly responsive to network changes. Nevertheless, this advantage comes at a computational cost, of the order of $O(mn)$ (Das et al., 2018), particularly in the context of large networks. Furthermore, closeness centrality's dependence on node reachability introduces constraints, making it unsuitable for networks with disconnected components and networks characterized by small diameters (Sariyüce et al., 2013).

4.3 Betweenness Centrality

Node betweenness centrality (Rodrigues, 2019) serves as a metric to gauge the degree to which a specific node lies on the shortest paths connecting other nodes within a network. Mathematically, the betweenness centrality score of node i is computed by summing the fraction of all-pairs shortest paths that pass through node i: $C_b(i) = \frac{2}{(n-1)(n-2)} \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$. Here, σ_{st} is the total number of shortest paths from node s to node t and $\sigma_{st}(i)$ is the number of those paths that include node i.

Nodes with high betweenness centrality score play a pivotal role in facilitating efficient information flow across the network. Serving as vital connectors between different network segments, they appear on numerous shortest paths connecting various nodes.

In communication networks and anomaly detection, nodes with high betweenness centrality scores are often targeted by attackers or malware due to their control over network flow, making them susceptible to exploitation for disruptive purposes. Their removal can result in network fragmentation, leading to congestion and diminished performance (Powell and Hopkins, 2015).

Similar to closeness centrality, betweenness centrality offers the advantage of relying on a comprehensive view of the network, as it computes the count of the shortest paths passing through each node for every pair of nodes in the network. This comes at a high computational cost of $O(m^3)$ (Das et al., 2018).

4.4 Eigenvector Centrality

Node eigenvector centrality (Rodrigues, 2019) is a measure of a node's importance, taking into account not only its number of connections but also the importance of its neighboring nodes. Nodes with high eigenvector centrality are linked to other nodes that are themselves important within the network. Mathematically, eigenvector centrality score of node i can be computed as: $C_e(i) = \frac{1}{\lambda} \sum_{j \in N(i)} (A_{ij} \times x(j))$. Here, A_{ij} denotes the element of the adjacency matrix A that corresponds to the connection between nodes i and j, and $x(j)$ represents the centrality score of node $j \in N(i)$. This relationship can be represented in matrix form as $AX = \lambda X$, where X is a vector of eigenvector centrality scores.

In the context of communication networks and the study of malware diffusion, eigenvector centrality proves invaluable in pinpointing nodes with a high potential for spreading malware throughout the network. For instance, a node that maintains connections with numerous highly central nodes, such as network hubs, is likely to have a high eigenvector centrality score, signifying its capability in the propagation of malware.

It is worth noting that eigenvector centrality takes into account the global structure of the network, contributing to its computational complexity, which scales at $O(n^2)$ (Das et al., 2018). Additionally, it may not be suitable for graphs that lack strong connectivity or contain loops.

In summary, the aforementioned centrality methods are instrumental in the identification of critical nodes (*i.e.*, nodes with high centrality scores). They provide network administrators with the means to proactively address potential issues, thus safeguarding the overall performance of the network.

5 Network Properties

For the remainder of this paper, as we delve into our analysis, computations, and numerical findings, we will consider a network comprising 24 nodes and 32 edges, as depicted in Fig. 2. In this section, our focus shifts towards harnessing node centrality as a tool to glean insights and extract essential network properties. These insights will provide valuable guidance on how to enhance network security.

Table 1 Likelihood (p_i) of compromising a critical node (i) given its centrality measure $(C_x(i), x \in \mathcal{M})$ for the various centrality methods

Degree Centrality			Closeness Centrality			Betweeness Centrality			Eigenvector Centrality		
Node i	$C_d(i)$	p_i	Node i	$C_c(i)$	p_i	Node i	$C_b(i)$	p_i	Node i	$C_e(i)$	p_i
10	0.696	0.444	10	0.639	0.193	10	0.787	0.4712	10	0.641	0.320
2	0.217	0.139	2	0.500	0.151	2	0.248	0.1484	2	0.288	0.144
1	0.130	0.083	35	0.451	0.136	12	0.184	0.1101	3	0.224	0.112
3	0.130	0.083	17	0.442	0.133	17	0.158	0.0947	35	0.221	0.110
24	0.130	0.083	8	0.434	0.131	1	0.142	0.0852	33	0.219	0.110
33	0.130	0.083	16	0.434	0.131	28	0.101	0.0604	24	0.207	0.104
35	0.130	0.083	3	0.418	0.126	35	0.050	0.0300	8	0.201	0.100
$R_x(\%)$		30.96			33.91			30.25			33.01

5.1 Critical Nodes Identification

We will compute the importance (centrality) score for each node in the graph depicted in Fig. 2, using the four centrality measures introduced in Section 4. We will use Definition 2 to evaluate node importance. For the purpose of our analysis, we will focus on the set of critical nodes which score is higher than a given threshold $\beta_x \geq 0, x \in \mathcal{M} = \{d, c, b, e\}$.[1] Table 1 displays seven nodes that satisfy our criteria when $\beta_d = 0.10, \beta_c = 0.40, \beta_b = 0.05, \beta_e = 0.20$. Note that β_x can be assigned a constant value β across all methods. However, this approach results in a variable number of critical nodes for each method. To facilitate our analysis and presentation, we tailor β_x to each method to ensure a consistent number of critical nodes. Let $S_x, x \in \mathcal{M}$, be the set of these critical nodes in the network where $|S_x| = k$.

The data presented in Table 1 consistently highlight that node 10 emerges as the most critical node in the network, closely followed by node 2, across the four critical measure methods. Moreover, nodes 10, 2, and 35 consistently rank among the top 7 critical nodes. It is important to highlight that, while we've presented a subset of the network's critical nodes, our computations unequivocally identify node 7 as the least critical node across all methods ($C_d(7) = 0.043, C_c(7) = 0.239, C_b(7) = 0, C_e(7) = 0.003$). The network topology, as depicted in Fig. 2, aligns with these findings. Node 7 occupies a peripheral position in the network, distant from the majority of the nodes and linked to only one node (node 1), contributing to its low importance score. In contrast, node 10 holds a central position, boasting numerous connections that bring it into close proximity with a larger number of nodes, thereby substantiating its high importance score.

5.2 Likelihood of a Critical Node Attack

In practice, a node's susceptibility to a cyberattack is influenced by a multitude of factors, including its position within the network, the sensitivity of the data it stores, and its vulnerability to specific attack vectors. Moreover, since nodes in a network are interconnected, the likelihood of an attack is intricately tied to the unique characteristics of all nodes, particularly those in close proximity.

In our analysis, we make the assumption that these influencing factors and node attributes are encapsulated within the node centrality measures, $C_x(i), x \in \mathcal{M}$. Leveraging these centrality measures, particularly those associated with critical nodes, significantly enhances our ability to evaluate the network's vulnerability to potential cyberattacks.

To streamline our analysis and make meaningful comparisons, we have standardized each centrality score which we denote as p_i to ensure that the sum of centralities across all critical nodes adds up to one ($\sum_i^k p_i = 1$). This standardization process allows us to provide the following interpretation of the p_i values.

Assumption 1 The normalized centrality measure p_i for node i denotes its vulnerability to cyberattacks, with higher values indicating a greater likelihood of being targeted. Mathematically, p_i is defined as:

$$p_i = C_x(i) / \sum_{j \in S_x} C_x(j), x \in \mathcal{M}$$

and can be interpreted as the probability of attack of node i.

The above interpretation introduces a novel concept that correlates a node's centrality measure with its susceptibility to attacks. Table 1 offers insights into the probability of a critical node being targeted in an attack using all four centrality methods.

Note that in Definition 2, the threshold β can be set as the probability cutoff.

[1] d: degree, c: closeness, b: betweenness, e: eigenvector centrality methods

5.3 Network Robustness

The concept of network robustness varies based on the application and the objective (*e.g.*, Liu et al., 2017; Ghayoori and Leon-Garcia, 2013; Si et al., 2022; Cai et al., 2021; Davis et al., 2021; Lou et al., 2023; Hamouda, 2024). Within the context of network security, the robustness of a network is intricately connected to the security of its individual nodes. In this work, to quantify a network robustness, we introduce the notation R_x, $x \in \mathcal{M}$ and provide the following definition:

Definition 3 The network robustness, R_x, quantifies the probability that the identified k critical nodes in the network remain secure. It is defined as a function of the node-level probability of compromise, p_i, associated with each node $i \in S_x$ under a particular centrality method $x \in \mathcal{M}$:

$$R_x = \prod_{i \in S_x} (1 - p_i). \tag{1}$$

Note that the higher the probabilities, p_i's, the lower the robustness R_x. Thus, the more the network is vulnerable. In essence, R_x represents the network's capacity to endure potential attacks by accounting for the combined resilience of its critical nodes. As depicted in Table 1, our numerical findings illustrate variations in network robustness depending on the applied centrality method. Closeness centrality exhibits the highest network robustness at 33.91%, while betweenness centrality demonstrates the lowest at 30.25%, indicating a relatively lower level of network robustness for the latter.

The computation of R_x for centrality method $x \in \mathcal{M}$ is outlined in Algorithm 1.

Algorithm 1 Calculate Network Robustness R_x.

```
Require: Graph $G = (V, E)$, Centrality Method $x \in \mathcal{M}$
$S \Leftarrow \emptyset$
$C \Leftarrow compute\_centrality(V)$
while $|S| \leq |V|$ do
    $i = \arg\max\{C\}, \forall i \in V \setminus S$
    $S \leftarrow S \cup \{i\}$
end while
$C \leftarrow$ Sort vector $C$ in descending order
$S_x \leftarrow$ Choose the top $k$ central nodes in $S$.
$p = C / \sum_{i \in S_x} C(i)$
$R_x = \prod_{i \in S_x} (1 - p_i)$
return $R_x$
```

The computational complexity of Algorithm 1 reduces to the complexity of the most computationally demanding method. In this case, it is the betweenness centrality which complexity is of the order of $O(m^3)$ (see Section 4), where m is the number of edges in the network.

6 Modeling Network Security

In this section, our focus centers on augmenting network security through the reinforcement of measures implemented at critical nodes. This entails investing in security enhancements for individual nodes, denoted as s_i, $i \in S_{x \in \mathcal{M}}$. Striking a balance between the cost of these security measures and the potential risk of a security breach is of utmost significance when devising strategies to fortify critical nodes.

Notably, there exists an inverse relationship between the likelihood of compromising a critical node and the level of security investment. Leveraging this relationship, we formulate critical node security as a nonlinear optimization problem. Our objective is to determine the optimal values of investment or resource s_i for $i \in S_x$, that minimizes the probability of compromising critical nodes while adhering to a specified budget constraint.

Let $\mathbf{s} = (s_1, \ldots, s_k)$ denote the allocated resources to nodes $i \in \{1, \ldots k\}$. We use the notation $R_x(\mathbf{s})$ to indicate that the robustness R_x is computed under resource allocation $(\mathbf{s})$. Furthermore, we express p_i as a function of s_i and denote it as $p_i(s_i)$. In this paper, we assume that $p_i(s_i) = p_i(0) - \alpha s_i$, where α is a predefined positive constant, and $p_i(0)$ is the probability of node i being compromised when $s_i = 0$, *i.e.* the probability of a node attack before any security investment. α can be interpreted as the sensitivity of the probability of attack to the security investment level. It is worth noting that alternative functions, including quadratic functions, can be considered to establish the relationship between p_i and s_i.

In the following analysis, we explore two distinct scenarios.

6.1 Scenario 1: Enhancing Security of all Critical Nodes

Below, we present an optimization model designed to achieve efficient resource allocation within the defined budget constraint, denoted as B (as per Eq. 4). In simpler terms, our goal is to allocate B units of resources (*i.e.* dollars) to enhance the security of critical nodes, thereby minimizing their vulnerability to cyberattacks and maximizing network robustness.

$$\max R_x(\mathbf{s}) = \prod_{i \in S_x} (1 - p_i(s_i)). \tag{2}$$

$$\text{s.t.} \quad 0 \leq p_i(s_i) = p_i(0) - \alpha s_i \leq 1 \tag{3}$$

$$\sum_{i \in S_x} s_i \leq B \tag{4}$$

$$s_i \geq 0, \forall i \in S_x. \tag{5}$$

The formulation outlined above constitutes a nonlinear optimization problem where constraint Eq. 3, defines the probability of a cyberattack occuring at node i. Note that in constraint Eq. 3 the constant α is the rate at which an attack probability decreases with each dollar of resource invested to secure a node. It can also be thought of as a scaling parameter for the budget level B. Feasible values of α are bounded above by $min\{p_i(0)/B\}$ and below by 0. The choice of α may stem from theoretical insights or expert judgment, derived from mathematical models capturing system dynamics. Alternatively, it can be estimated using empirical data, practical observations, and experiments or simulations to observe the impact of investment changes on attack probability. In summary, estimating α is an iterative process, balancing theoretical expectations with real-world observations.

Proposition 2 *The function $R_x(\mathbf{s})$ is convex.*

Proof The proof of convexity is straightforward. We begin by applying the logarithm (log) of the product terms of $R_x(\mathbf{s})$, resulting in $log(\prod_{i\in S_x}(1-p_i(s_i)))$. This simplifies to: $\sum_{i\in S_x} log(1-p_i(s_i))$. Recognizing that $(1-p_i(s_i))$ is convex in s_i, the log preserves convexity, and since the sum of convex functions is convex, we conclude that the function $R_x(\mathbf{s})$ is convex. □

Proposition 3 *The optimal allocation $s_i^*, i \in S_x$ is obtained using Algorithm 2:*

Algorithm 2 Optimal Resource Allocation.

```
Require: Probabilities $p$, budget $B$, set of critical nodes $S_x$
Ensure: $S_x \neq \emptyset$
Rank the probabilities $p_i(0)$ in descending order, and indexed in ascending order.
Let $s_i^*$ denote the optimal allocation to node $i$.
Let $s_1^* \leftarrow min\{\frac{p_1(0)}{\alpha}, B\}$
for all $i = 2, \ldots, |S_x|$ do
    $s_i^* \leftarrow min\{\frac{p_i(0)}{\alpha}, B - \sum_{j=1}^{i-1} s_j^*\}$
    $i \leftarrow i + 1$
end for
return $\mathbf{s}^* = (s_1^*, \ldots, s_k^*)$
```

Proof The Lagrangian can be written as:

$$\mathcal{L}(s,\lambda,\mu) = R_x(\mathbf{s}) + \sum_{j=1}^{k}\lambda_j(p_j(0)-\alpha s_j - 1) + \mu(\sum_{j=1}^{k} s_j - B)$$

Here, $\lambda_j, j \in \{1,\ldots,k\}$ and μ are non-negative Lagrangian multipliers. Since $R_x(\mathbf{s})$ is convex in $\mathbf{s}$, the K.K.T conditions are sufficient for optimality of $\mathbf{s}$. The optimality conditions can be written as:

$$\frac{\partial \mathcal{L}(s,\lambda,\mu)}{\partial s_i} = \alpha \prod_{\substack{j=1\\ j\neq i}}^{k}(1-p_j(0)+\alpha s_j) - \alpha\lambda_i + \mu = 0 \tag{6}$$

$$\frac{\partial \mathcal{L}(s,\lambda,\mu)}{\partial \lambda_i} = \lambda_i(p_i(0)+\alpha s_i - 1) = 0, i \in \{1,\ldots,k\} \tag{7}$$

$$\frac{\partial \mathcal{L}(s,\lambda,\mu)}{\partial \mu_i} = \mu(\sum_{j=1}^{k} s_j - B) \tag{8}$$

Here Eqs. 7 and 8 represent the complementary slackness conditions associated with constraints Eqs. 3 and 4. From Eq. 6, we have,

$$\alpha(1-p_i(0)+\alpha s_i)\prod_{j\neq i}^{k}(1-p_j(0)+\alpha s_j)$$
$$-\alpha\lambda_i(1-p_i(0)+\alpha s_i) + \mu(1-p_i(0)+\alpha s_i) = 0 \tag{9}$$

Note that the first term in the LHS of the above equation, is equal to $\alpha R_x(\mathbf{s})$ and the second term is equal to zero due to the complementary slackness conditions Eq. 7. Hence, $\alpha R_x(\mathbf{s}) + \mu(1-p_i(0)+\alpha s_i) = 0$, and can be rewritten as: $R_x(\mathbf{s}) = -\frac{\mu}{\alpha}(1-p_i(0)+\alpha s_i),\ \forall i \in \{1,\ldots,k\}$. Thus, $R_x(\mathbf{s}) = -\frac{\mu}{\alpha}(1-p_i(0)+\alpha s_i) = -\frac{\mu}{\alpha}(1-p_j(0)+\alpha s_j)$.

Note that the budget constraint Eq. 4 must be binding since it is trivial to show that $R_x(\mathbf{s}') > R_x(\mathbf{s})$, for $\mathbf{s}' = \mathbf{s} + b\mathbf{e}$, for any feasible $\mathbf{s}$ and positive constant b. Here, $\mathbf{e}$ is the unit vector of appropriate dimension. Hence,

$$s_i = \frac{p_i(0)-p_j(0)}{\alpha} + s_j. \tag{10}$$

As a result, $s_i \geq s_j$, iff $p_i(0) \geq p_j(0)$.

Therefore, if we rank the nodes according to their probabilities, $p_j(0)$, in descending order, the optimal allocation is as follows:

$$s_1^* = min\{\frac{p_1(0)}{\alpha}, B\}, \ldots,$$
$$s_i^* = min\{\frac{p_i(0)}{\alpha}, B - \sum_{j=1}^{i-1} s_j^*\} \tag{11}$$

□

Note that Algorithm 2 has a computation complexity of $O(n \log n)$.

We solved the above optimization problem focusing on the example network depicted in Fig. 2. The optimal values of p_i, denoted as $p_i(s_i^*)$, are presented in Table 2. Notably, as the budget increases progressively from $B = 100$ to $B = 500$, the likelihood of a successful attack decreases significantly, leading to a substantial enhancement in network robustness. Note that we use a budget of zero as our baseline for comparison. Specifically, when considering betweenness centrality, we observe the most substantial increase in robustness, rising from 30.25% with a zero budget to 91.14% with a budget of 500. However, the use of closeness centrality results in the lowest improvement in robustness, albeit a notable one, rising from 33.91% with no budget to 76.48% when the budget reaches 500.

Table 2 Probabilities of attack ($p_i(s^*)$) for varied budgets (B) across different centrality measures

Degree centrality						
$B \rightarrow$	0	100	200	300	400	500
Node i	$p_i(s^*)$					
10	0.444	0.000	0.000	0.000	0.000	0.000
2	0.139	0.139	0.000	0.000	0.000	0.000
1	0.083	0.083	0.083	0.067	0.050	0.033
3	0.083	0.083	0.083	0.067	0.050	0.033
24	0.083	0.083	0.083	0.067	0.050	0.033
33	0.083	0.083	0.083	0.067	0.050	0.033
35	0.083	0.083	0.083	0.067	0.050	0.033
$R_d(\mathbf{s})(\%)$	30.96	55.73	64.72	70.82	77.38	84.41
Closeness centrality						
$B \rightarrow$	0	100	200	300	400	500
Node i	$p_i(s^*)$					
10	0.193	0.000	0.000	0.000	0.000	0.000
2	0.151	0.151	0.021	0.000	0.000	0.000
35	0.136	0.136	0.117	0.077	0.046	0.019
17	0.133	0.133	0.133	0.095	0.065	0.038
8	0.131	0.131	0.131	0.112	0.083	0.056
16	0.131	0.131	0.131	0.112	0.083	0.056
3	0.126	0.126	0.126	0.126	0.115	0.090
$R_c(\mathbf{s})(\%)$	33.91	42.00	49.48	57.50	66.46	76.48
Betweenness centrality						
$B \rightarrow$	0	100	200	300	400	500
Node i	$p_i(s^*)$					
10	0.471	0.000	0.000	0.000	0.000	0.000
2	0.148	0.148	0.000	0.000	0.000	0.000
12	0.110	0.110	0.110	0.000	0.000	0.000
17	0.095	0.095	0.095	0.095	0.000	0.000
1	0.085	0.085	0.085	0.085	0.085	0.000
28	0.060	0.060	0.060	0.060	0.060	0.060
35	0.030	0.030	0.030	0.030	0.030	0.030
$R_b(\mathbf{s})(\%)$	30.25	57.20	67.17	75.48	83.37	91.14
Eigenvector centrality						
$B \rightarrow$	0	100	200	300	400	500
Node i	$p_i(s^*)$					
10	0.320	0.000	0.000	0.000	0.000	0.000
2	0.144	0.144	0.000	0.000	0.000	0.000
3	0.112	0.112	0.112	0.063	0.027	0.000
35	0.111	0.111	0.111	0.076	0.040	0.012
33	0.110	0.110	0.110	0.083	0.048	0.020
24	0.104	0.104	0.104	0.104	0.100	0.074
8	0.100	0.100	0.100	0.100	0.100	0.100
$R_e(\mathbf{s})(\%)$	33.01	48.56	56.72	64.08	72.05	80.71

Figure 3 illustrates the correlation between network robustness and budget across all centrality measure methods, demonstrating a positive relationship. We also observe that network robustness, $R_x(\mathbf{s})$, increases quasilinearly as a function of the budget. This observation is particularly significant for real-world networks, as it allows for the exploitation of this quasilinear relationship to avoid solving large nonlinear optimization problems.

In the following, we focus on the resource allocation that led to improving the network robustness. Our findings, as summarized in Table 3, reveal that using degree centrality method for instance, in the absence of any budget allocated to safeguard critical node 10, the probability of a successful attack stands at 0.444. However, with an investment of 100 out of the available 300 budget, we managed to reduce the attack probability by 100% (from $p_i(0) = 0.444$ to $p_i(100) = 0$). These values are highlighted in bold in Table 3. In fact, for node 10 the probability of an attack has dropped to zero for $B > 0$ across all centrality methods. Similar trends are observed, for the other critical nodes, as the budget increases.

In cases with limited budget (e.g., $B = 100$), a majority of the resources are directed towards node 10. Conversely, in cases with a more generous budget (e.g., $B = 500$), substantial resources are allocated to nodes 10 and 2, effectively reducing the attack probability to zero for both nodes, while the remaining funds are distributed among less critical nodes.

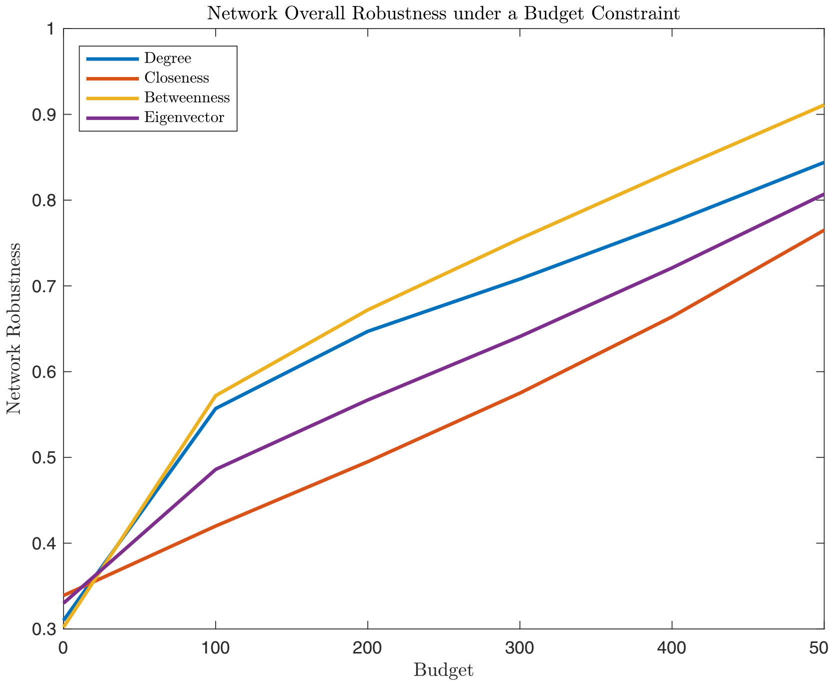
Our analysis yields analogous results for the other centrality methods.

Note that the problem described by Eqs. 2-5 can be formulated to minimize the security investment cost to ensure a certain level of network robustness.

6.2 Scenario 2: Enhancing Security of Common Top Critical Nodes

Ideally, comprehensive security measures should be implemented across all critical nodes as in scenario 1. However, practical scenarios often present a challenge due to budget constraints, particularly in the context of large networks. Consequently, it may not be sustainable to invest in the secu-

Fig. 3 Budget-dependent network robustness across different centrality measures - Scenario 1



rity of all critical nodes. Prioritizing the allocation of the available budget to protect a select subset of the most critical nodes can yield effective results. Therefore, in the subsequent discussion, we will assume our limited budget, B, will be allocated to enhance the security of specific number, denoted as a, nodes among the total pool of k critical nodes ($a > 0$).

Our numerical results have revealed a consensus among four critical methods, all identifying nodes 10 and 2 as the most critical. Therefore, we will concentrate our efforts and allocate resources to secure these two nodes, with $a = 2$. When $a = k$, this scenario entails securing all critical nodes, which is described in scenario 1.

Let $A \subset S_x$ denote the set of the a nodes. We introduce the notation $\pi_{ax}(\mathbf{s})$, which represents the probability of a simultaneous attack on the nodes $i \in A$ considering the assumptions described above. We formulate the optimization problem given in Eqs. 12-15 to determine the optimal allocation $\mathbf{s}$ that minimizes the probability $\pi_{ax}(\mathbf{s})$. This probability quantifies the risk of a critical nodes being compromised within the specified budget constraint B, as expressed in constraint Eq. 14.

$$\min \pi_{ax}(\mathbf{s}) = \sum_{S_A \subset A} \prod_{i \in S_A} p_i(s_i) \prod_{j \in S_x \setminus S_A} (1 - p_j(s_j)) \quad (12)$$

$$\text{s.t.} \quad 0 \le p_i(s_i) = p_0 - \alpha s_i \le 1 \quad (13)$$

$$\sum_{i \in S_x} s_i \le B \quad (14)$$

$$s_i \ge 0, \forall i \in S_x \quad (15)$$

The formulation outlined above constitutes a nonlinear optimization problem and S_A is a subset of A.

Proposition 4 *The function $\pi_{ax}(\mathbf{s})$ is convex.*

Proof The proof follows the same steps as the one for proposition 2. □

We solved the aforementioned optimization problem by considering budgets within the range of 0 to 500. The numerical results, presented in Fig. 4, reveal that the betweenness centrality measure outperforms the other methods, with degree centrality closely following, while closeness centrality exhibits the lowest performance. All methods consistently exhibit a positive correlation between network robustness and budget.

It is important to highlight that the levels of robustness achieved in this scenario are slightly lower compared to those in scenario 1, albeit not significantly so. The difference becomes more pronounced, particularly with higher budgets. Specifically, the difference in performance is as follows: less than 7.5% for degree centrality, less than 11.8% for closeness centrality, 13.1% for betweenness centrality, and 9% for eigenvector centrality. However, for lower budgets ($B < 300$), the difference is less than 4%. This discrepancy can be justified since securing all critical nodes may entail additional costs (management, administration, etc.) beyond the scope of this study.

It is noteworthy that the results presented in Figs.3 and 4 show similar outcomes when the budget (B) is set to 100. In both scenarios, the $100 budget is only sufficient to reduce

Table 3 Optimal budget allocation (s_i^*) for critical nodes, across different centrality measures, leading to attack probability reduction ($p_\downarrow$)

Degree centrality											
	$B = 0$	$B = 100$		$B = 200$		$B = 300$		$B = 400$		$B = 500$	
Node i	$p_i(s_i^*)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$
10	**0.444**	100	100	100.000	100	**100**	**100**	100	100	100	100
2	0.139	0.000	0	100	100	100	100	100	100	100	100
1	0.083	0.000	0	0.001	0	20.002	20	40.002	40	60.001	60
3	0.083	0.000	0	0.001	0	20.002	20	40.002	40	60.001	60
24	0.083	0.000	0	0.001	0	20.002	20	40.002	40	60.001	60
33	0.083	0.000	0	0.001	0	20.002	20	40.001	40	60.001	60
35	0.083	0.000	0	0.001	0	19.993	20	39.995	40	59.997	60
Closeness centrality											
	$B = 0$	$B = 100$		$B = 200$		$B = 300$		$B = 400$		$B = 500$	
Node i	$p_i(s_i^*)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$
10	0.193	100	100	100	100	100	100	100	100	100	100
2	0.151	0.007	0	85.842	86	99.983	100	100	100	100	100
35	0.136	0.001	0	13.787	14	43.039	43	66.181	66	86.284	86
17	0.133	0.001	0	0.366	0	28.585	29	51.487	51	71.450	71
8	0.131	0.000	0	0.001	0	14.188	14	36.947	37	56.868	57
16	0.131	0.001	0	0.005	0	14.188	14	36.946	37	56.868	57
3	0.126	0.001	0	0.002	0	0.019	0	8.440	8	28.530	28
Betweenness centrality											
	$B = 0$	$B = 100$		$B = 200$		$B = 300$		$B = 400$		$B = 500$	
Node i	$p_i(s_i^*)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$
10	0.471	100	100	100	100	100	100	100	100	100	100
2	0.148	0.001	0	100	100	100	100	100	100	100	100
12	0.110	0.000	0	0.001	0	100	100	100	100	100	100
17	0.095	0.000	0	0.000	0	0.020	0	100	100	100	100
1	0.085	0.000	0	0.000	0	0.001	0	0.013	0	100	100
28	0.060	0.000	0	0.000	0	0.000	0	0.000	0	0.001	0
35	0.03	0.000	0	0.000	0	0.000	0	0.000	0	0.000	0
Eigenvector centrality											
	$B = 0$	$B = 100$		$B = 200$		$B = 300$		$B = 400$		$B = 500$	
Node i	$p_i(s_i^*)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$	s_i^*	$p_\downarrow(\%)$
10	0.320	100	100	100	100	100	100	100	100	100	100
2	0.144	0.001	0	100	100	100	100	100	100	100	100
3	0.112	0.001	0	0.001	0	43.902	44	76.250	76	100	100
35	0.111	0.001	0	0.001	0	31.681	32	63.733	64	89.460	90
33	0.110	0.001	0	0.003	0	24.399	24	56.287	56	81.829	82
24	0.104	0.001	0	0.001	0	0.013	0	3.729	4	28.721	29
8	0.100	0.000	0	0.000	0	0.006	0	0.001	0	0.009	0

the probability of attack for node 10 (the most important node) to zero. Consequently, no resources were allocated to any other nodes. Therefore, the robustness remains the same under both scenarios. However, with a larger budget, scenario 1 allocates resources to multiple nodes, potentially covering all nodes. In contrast, scenario 2 restricts the budget to the top k critical nodes. As a result, the robustness behavior of the two scenarios diverges beyond a budget of \$100.

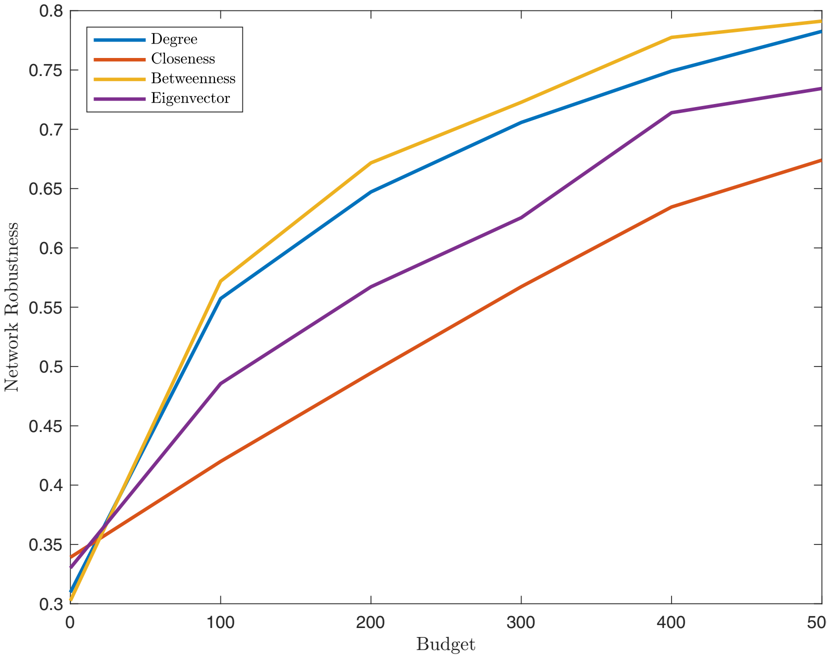


Fig. 4 Budget-dependent network robustness across different centrality measures - Scenario 2

7 Integrating Critical Node Centrality Methods to Enhance Network Security

When computing node centrality measures, results vary among different methods, each offering unique insights into network dynamics. Our numerical results in Table 1 consistently reveal that all critical methods identify the same top critical nodes, specifically nodes 10 and 2. Our analysis also reveals that among the k critical nodes considered in our study, there is an overlap in those identified by the four centrality methods. This consistency persisted across the various networks we analyzed. Moreover, as shown in Fig. 3, no single centrality method monotonically dominates the others in terms of network robustness, indicating that each method captures distinct network features. However, it is common for researchers to rely on a single critical node identification method tailored to their specific application.

These observations have motivated our investigation into using all methods to identify critical nodes, allowing us to encompass the full spectrum of intrinsic network characteristics through its critical nodes.

We focus on the k critical nodes derived from all four centrality methods, described in Section 4, aiming to harness both their shared characteristics and distinctions to bolster network security. Our objective is to assess their combined influence on network security. To achieve this, we start by forming the union of all nodes identified as critical by the four methods. We calculate the centrality of each node by aggregating its centrality values from all four methods. Subsequently, we determine the attack probability by normalizing the cumulative centrality measure for each node. Finally, we assess the network robustness using Eq. 1.

We outline our integration method for computing the cumulative centrality values, denoted as C_4, the probability of an attack, denoted as p, and the network robustness, denoted as R_4, in Algorithm 3.

Algorithm 3 Calculate Network Robustness R_4.

```
Require: Graph $G = (V, E)$, Centrality Methods $\mathcal{M}$.
Ensure: $V \neq \emptyset$
for all $x \in \mathcal{M}$ do
    $S_x \Leftarrow \emptyset$
    $C \Leftarrow compute\_centrality(V, x)$
    while $|S_x| \leq |V|$ do
        $i = \arg\max\{C\}, \forall i \in V \setminus S_x$
        $S_x \leftarrow S_x \cup \{i\}$
    end while
    $C \leftarrow$ Sort vector $C$ in descending order
    $S_x \leftarrow$ Select $k$ highest critical nodes
end for
$S = \cup_{x \in \mathcal{M}} S_x$
for all $i \in S$ do
    $C_4(i) = \sum_{x=1}^{4} C_x(i)$
end for
$p = C_4 / \sum_{i \in S} \sum_{x \in \mathcal{M}} C_x(i)$
$R_4 = \prod_{i \in S} (1 - p_i)$
return $R_4$
```

For the sample network depicted in Fig. 2, the application of the algorithm described above yielded an R_4 value of 34.5%. This represents a notable improvement over individual centrality methods: 10.26% over degree centrality, 1.71%

over closeness centrality, 12.31% over betweenness centrality, and 4.31% over eigenvector centrality, when each is used in isolation. We observed similar positive improvement considering various other networks. For brevity, those results are not shown here.

In conclusion, the utilization of centrality methods in isolation may inadvertently disregard valuable insights regarding node importance. Our findings underscore the significance of integrating all methods, as it provides a more holistic perspective on network security, uncovering crucial insights that might otherwise remain hidden.

8 Conclusion

Our research introduces a novel approach that leverages insights derived from computing node centrality scores using four widely-recognized centrality measures. The uniqueness of our work lies in its ability to convert these centrality measures into actionable insights for identifying network attack probabilities, offering an unconventional yet effective approach to strengthening network robustness.

Additionally, we propose a novel closed-form expression for network robustness, unveiling its direct correlation with node-centric features, including importance and attack probabilities.

These findings have transformed our understanding of network security, as we formulated the problem as a nonlinear optimization problem, under budget constraints. Within this optimization framework, we have successfully identified an optimal resource allocation strategy, with the goal of minimizing the probability of cyberattacks on critical nodes while maximizing network robustness.

Furthermore, our results highlight the significance of integrating centrality methods, as depending on them in isolation may lead to overlooking valuable information concerning node importance. We show that the integration of all methods yields a more comprehensive and holistic understanding of network security, ultimately resulting in enhanced network robustness.

One limitation of our proposed method is its dependence on the selected centrality method, which introduces a reliance on its time and space complexities. Moreover, it is known in the literature that computing centrality measures can be challenging for very large networks. Although Algorithm 2 remains computationally efficient for very large networks, addressing problems under scenario 2 remains challenging. Nevertheless, approximate solutions can be obtained using methods such as stochastic gradient descent.

Our future work will be directed towards validating our research outcomes in larger and more complex network environments, as well as refining and advancing our critical node identification techniques. We believe that these ongoing efforts will further solidify the practical applicability and effectiveness of our approach in safeguarding critical infrastructure and strengthening cybersecurity measures in an ever-evolving digital landscape.

Acknowledgements Not applicable

Author Contributions Dr. Hamouda and Dr. ElHafsi conceptualized the paper, developed mathematical models, and conducted numerical experiments and analysis. Dr. Son contributed to the paper's conception, and the editorial aspect of the paper. All authors have read and approved the content of the manuscript.

Funding The authors did not receive support from any organization for the submitted work.

Availability of data and materials We used data to generate the network in Fig. 2. We referenced the site from where the data was obtained in Section 3 per the publisher requirement posted on the website (https://networkrepository.com/networks.php). Here is the statement they posted:

Code Availibility It is available and it will be uploaded to the journal repository and to Github once the paper is accepted.

Declarations

Ethics approval Not applicable

Consent to participate Not applicable

Consent for publication Not applicable

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications, 60*, 19–31.

Alozie, G. U., Arulselvan, A., Akartunalı, K., & Pasiliao, E. L., Jr. (2021). Efficient methods for the distancebased critical node detection problem in complex networks. *Computers and Operations Research, 131*, 105254.

Alozie, G. U., Arulselvan, A., Akartunalı, K., & Pasiliao, E. L., Jr. (2022). A heuristic approach for the distance-based critical node detection problem in complex networks. *Journal of the Operational Research Society, 73*(6), 1347–1361.

Aringhieri, R., Grosso, A., Hosteins, P., & Scatamacchia, R. (2016). A general evolutionary framework for different classes of critical node problems. *Engineering Applications of Artificial Intelligence, 55*, 128–145.

Arulselvan, A. (2009). Network model for disaster management. PhD thesis, University of Florida Gainesville.

Arulselvan, A., Commander, C., Elefteriadou, L., & Pardalos, P. (2009a). Detecting critical nodes in sparse graphs. *Computers and Operations Research, 36*, 2193–2200.

Arulselvan, A., Commander, C. W., Elefteriadou, L., & Pardalos, P. M. (2009b). Detecting critical nodes in sparse graphs. *Computers and Operations Research, 36*(7), 2193–2200.

Ballinger, O. (2023). Insurgency as complex network: Image co-appearance and hierarchy in the pkk. *Social Networks, 74*, 182–205.

Berger, A., Grigoriev, A., & van der Zwaan, R. (2014). Complexity and approximability of the k-way vertex cut. *Networks, 63*(2), 170–178.

Cai, M., Liu, J., & Cui, Y. (2021). Network robustness analysis based on maximum flow. *Frontiers in Physics, 9*, 792410.

Commander, C. W., Pardalos, P. M., Ryabchenko, V., Uryasev, S., & Zrazhevsky, G. (2007). The wireless network jamming problem. *Journal of Combinatorial Optimization, 14*, 481–498.

Dang, F., Zhao, X., Yan, L., Wu, K., Li, S. (2023). Research on network intrusion response method based on bayesian attack graph. In *2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE)* (pp. 639–645). IEEE.

Das, K., Samanta, S., & Pal, M. (2018). Study on centrality measures in social networks: a survey. *Social Network Analysis And Mining*, (13).

Davis, J. E., Kolozsvary, M. B., Pajerowska-Mukhtar, K. M., & Zhang, B. (2021). Toward a universal theoretical framework to understand robustness and resilience: from cells to systems. *Frontiers in Ecology and Evolution, 8*, 579098.

Deng, Y., Mo, H. (2019). Identifying node importance based on evidence theory in complex networks. *Physica A: Statistical Mechanics and its Applications*, 529.

Devkota, P., Danzi, M. C., & Wuchty, S. (2018). Beyond degree and betweenness centrality: Alternative topological measures to predict viral targets. *PloS one, 13*(5), e0197595.

Dinh, T. N., & Thai, M. T. (2011). Precise structural vulnerability assessment via mathematical programming. In *2011-MILCOM 2011 Military Communications Conference* (pp. 1351–1356). IEEE.

Dinh, T. N., Xuan, Y., Thai, M. T., Park, E. K., & Znati, T. (2010). On approximation of new optimization methods for assessing network vulnerability. In *2010 Proceedings IEEE INFOCOM* (pp. 1–9). IEEE.

Faramondi, L., Oliva, G., Pascucci, F., Panzieri, S., & Setola, R. (2016). Critical node detection based on attacker preferences. In *2016 24th Mediterranean Conference on Control and Automation (MED)* (pp. 773–778). IEEE.

Faramondi, L., Oliva, G., Setola, R., Pascucci, F., Amideo, A. E., & Scaparra, M. P. (2017). Performance analysis of single and multiobjective approaches for the critical node detection problem. In *Optimization and Decision Science: Methodologies and Applications: ODS, Sorrento, Italy, September 4-7, 2017 47* (pp. 315–324). Springer.

Fernandes, J. M., Suzuki, G. M., Zhao, L., & Carneiro, M. G. (2023). Data classification via centrality measures of complex networks. In *2023 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–8).

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Network, 1*, 215.

Ganguli, R., Mehta, A., Debnath, N. C., Aljahdali, S., & Sen, S. (2020). An integrated framework for friend recommender system using graph theoretic approach. *Proceedings of 35th International Confer, 69*, 242–255.

Ghayoori, A., & Leon-Garcia, A. (2013). Robust network design. In *2013 IEEE International Conference on Communications (ICC)* (pp. 2409–2414). IEEE.

Golbeck, J. (2013). Chapter 3 network structure and measures. *Morgan Kaufmann* (pp. 25–44).

Gouvy, N., Hamouda, E., Mitton, N., & Simplot-Ryl, D. (2012). Minimizing energy consumption through mobility with connectivity preservation in sensor networks. *International Journal of Parallel, Emergent and Distributed Systems, 27*(6), 521–540.

Gupta, B. B., Gaurav, A., Marín, E. C., & Alhalabi, W. (2023). Novel graph-based machine learning technique to secure smart vehicles in intelligent transportation systems. *IEEE Transactions on Intelligent Transportation Systems, 24*(8), 8483–8491.

Hamouda, E. (2024). A critical node-centric approach to enhancing network security. In *Lecture Notes in Computer Science*, vol 14321 (pp. 1–15). Springer Nature Switzerland AG.

Hamouda, E., Mitton, N., & Simplot-Ryl, D. (2011). Energy efficient mobile routing in actuator and sensor networks with connectivity preservation. (pp. 15–28) 01.

Helmi, R. A. A., Elghanuni, R. H., & Abdullah, M. I. (2021). Effect the graph metric to detect anomalies and non-anomalies on facebook using machine learning models. In *2021 IEEE 12th Control and System Graduate Research Colloquium (ICSGRC)* (pp. 7–12). IEEE.

Imran, M., Alnuem, M. A., Fayed, M. S., & Alamri, A. (2013). Localized algorithm for segregation of critical/non-critical nodes in mobile ad hoc and sensor networks. *Procedia Computer Science, 19*, 1167–1172. The 4th International Conference on Ambient Systems, Networks and Technologies (ANT 2013), the 3rd International Conference on Sustainable Energy Information Technology (SEIT-2013).

Invernizzi, L., Miskovic, S., Torres, R., Saha, S., Lee, S.-J., Kruegel, C., & Vigna, G. (2014). Nazca: Detecting malware distribution in large-scale networks. In *Proceedings of the 21st Symposium on Network and Distributed System Security Symposium*, February

Jain, A., & Reddy, B. V. R. (2013). Node centrality in wireless sensor networks: Importance, applications and advances. (pp. 127–131).

Johnson, J. R., & Hogan, E. A. (2013). A graph analytic metric for mitigating advanced persistent threat. In *2013 IEEE International Conference on Intelligence and Security Informatics* (pp. 129–133). IEEE.

Ke, L., Fang, X., & Fang, N. (2022). Pn-bbn: A petri net-based bayesian network for anomalous behavior detection. *Mathematics, 10*(20), 3790.

Kim, S. (2020). Anatomy on malware distribution networks. *IEEE Access, 8*, 73919–73930.

Kim, S., Kim, J., & Kang, B. B. (2018). Malicious url protection based on attackers' habitual behavioral analysis. *Computer Security, 77*, 790–806.

Kivimäki, I., Lebichot, B., Saramäki, J., & Saerens, M. (2016). Two betweenness centrality measures based on randomized shortest paths. *Scientific reports, 6*(1), 1–15.

Laha, M., Kamble, S., & Datta, R. (2020). Edge nodes placement in 5g enabled urban vehicular networks: A centrality-based approach. In *2020 National Conference on Communications (NCC)* (pp. 1–6). IEEE.

Lalou, H. K. M., & Tahraoui, M. A. (2018). The critical node detection problem in networks: A survey. *Computer Science Review, 28*, 92–117.

Lalou, M., Tahraoui, M. A., & Kheddouci, H. (2018). The critical node detection problem in networks: A survey. *Computer Science Review, 28*, 92–117.

Lalou, M., Tahraoui, M. A., & Kheddouci, H. (2018a). The critical node detection problem in networks: A survey. *Computer Science Review, 28*, 92–117.

Liu, X., Hong, Z., & Rodríguez-Patón, A., Zou, Q., Zeng, X., Liu, J., Lin, Y. (2020). Computational methods for identifying the critical nodes in biological networks, briefings in bioinformatics. 21, 486–497.

Liu, J., Zhou, M., Wang, S., & Liu, P. (2017). A comparative study of network robustness measures. *Frontiers of Computer Science, 11*, 568–584.

Li, Y., Yang, X., Zhang, X., Xi, M., & Lai, X. (2022). An improved voterank algorithm to identifying a set of influential spreaders in complex networks. *Frontiers in Physics, 10*, 955727.

Lou, Y., Wang, L., & Guanrong, C. (2023). Structural robustness of complex networks: A survey of a posteriori measures [feature]. *IEEE Circuits and Systems Magazine, 23*(1), 12–35.

Lozano, M., Garcia-Martinez, C, Rodriguez, F. J., & Trujillo, H. M. (2017). Optimizing network attacks by artificial bee colony. *Information Sciences, 377*, 30–50.

Majeed, A., & Rauf, I. (2020). Graph theory: A comprehensive survey about graph theory applications in computer science and social networks. *Inventions, 5*(1), 10.

Mazlumi, S. H. H., & Kermani, M. A. M. (2022). Investigating the structure of the internet of things patent network using social network analysis. *IEEE Internet of Things Journal,9*(15), 13458–13469.

Megzari, A., Pravija Raj, P. V., Osamy, W, & Khedr, A. M. (2023). Applications, challenges, and solutions to single-and multi-objective critical node detection problems: a survey. *The Journal of Supercomputing* (pp. 1–39).

Mitchell, C., Agrawal, R., & Parker, J. (2019). The effectiveness of edge centrality measures for anomaly detection. (pp. 5022–5027).

Mitton, N., Pavkovic, B.,-Simplot-Ryl, D., & Hamouda, E. (2009). Energy-aware georouting with guaranteed delivery in wireless sensor networks with obstacles. *International Journal of Wireless Information Networks, 16*, 142–153.

Powell, J., & Hopkins, M. (2015). A librarian's guide to graphs, data and the semantic web.

Proselkov, Y., Herrera, M., Parlikad, A. K., & Brintrup, A. (2021). Distributed dynamic measures of criticality for telecommunication networks. In *Service Oriented, Holonic and Multi-Agent Manufacturing Systems for Industry of the Future: Proceedings of SOHOMA 2020* (pp. 421–432). Springer.

Rains, H. (2022). *Dark Networks: An Exploration of the Ties that Bind Insurgent Groups and Shape Illicit Behavior*. PhD thesis, University of Kansas.

Rajalakshmi, K., Sambath, M., Joseph, L., Ramesh, K., & Surendiran, R. (2023). An effective approach for improving data access time using intelligent node selection model (insm) in cloud computing environment.

Riquelme, F., & Vera, J.-A. (2022). A parameterizable influence spread-based centrality measure for influential users detection in social networks. *Knowledge-Based Systems, 257*, 109922.

Rodrigues, F. A. (2019). Network centrality: an introduction. *A mathematical modeling approach from nonlinear dynamics to complex systems* (pp. 177–196).

Ryan, A. (2015). *Rossi and Nesreen K*. In AAAI: Ahmed. The network data repository with interactive graph analytics and visualization.

Sariyüce, A. E., Kaya, K., Saule, E, & Çatalyiirek,Ü. V. (2013). Incremental algorithms for closeness centrality. In *2013 IEEE International Conference on Big Data*, (pp. 487–492). IEEE.

Shen, Y., Dinh, T. N., & Thai, M. T. (2012a). Adaptive algorithms for detecting critical links and nodes in dynamic networks. In *MILCOM 2012-2012 IEEE Military Communications Conference* (pp. 1–6). IEEE.

Shen, Y., Nguyen, N. P., Xuan, Y., & Thai, M. T. (2012b). On the discovery of critical links and nodes for assessing network vulnerability. *IEEE/ACM Transactions on Networking, 21*(3), 963–973.

Shen, Y., Nguyen, N., Xuan, Y., & Thai, M. (2013). On the discovery of critical links and nodes for assessing network vulnerability. *Networking, IEEE/ACM Transactions on, 21*, 963–973.

Shi, W., Shi, X., Wang, K., Liu, J., & Xiong, Q. (2016). Evaluating the importance of nodes in complex networks. *Physica A: Statistical Mechanics and its Applications, 452*, 209–219.

Shukla, S. (2023). Angle based critical nodes detection (abcnd) for reliable industrial wireless sensor networks. *Wireless Personal Communications, 130*(2), 757–775.

Si, W., Mburano, B., Zheng, W. X., & Qiu, T. (2022). Measuring network robustness by average network flow. *IEEE Transactions on Network Science and Engineering, 9*(3), 1697–1712.

Ugurlu, O. (2022). Comparative analysis of centrality measures for identifying critical nodes in complex networks. *Journal of Computational Science, 62*, 101738.

Ventresca, M., & Aleman, D. (2014). A derandomized approximation algorithm for the critical node detection problem. *Computers and Operations Research, 43*, 261–270.

Veremyev, A., Prokopyev, O. A., & Pasiliao, E. L. (2015). Critical nodes for distance-based connectivity and related problems in graphs. *Networks, 66*(3), 170–195.

Walteros, J. L., Veremyev, A., Pardalos, P. M., & Pasiliao, E. L. (2019). Detecting critical node structures on graphs: A mathematical programming approach. *Networks, 73*(1), 48–88.

Wang, B., Jia, J., Zhang, L., & Gong, N. Z. (2018). Structure-based sybil detection in social networks via local rule-based propagation. *IEEE Transactions on Network Science and Engineering, 6*, 523–537.

Xing, Y., Shu, H., & Kang, F. (2023). Peerremove: An adaptive node removal strategy for p2p botnet based on deep reinforcement learning. *Computers and Security, 128*, 103129.

Yan, G., Chen, G., Eidenbenz, S. J., & Li, N. (2011). Malware propagation in online social networks: nature, dynamics, and defense implications. In *ACM Asia Conference on Computer and Communications Security*.

Yang, H., & An, S. (2020). Critical nodes identification in complex networks. *Symmetry, 12*(1), 123.

Yen, C. -C., Yeh, M. -Y., & Chen, M. -S. (2013). An efficient approach to updating closeness centrality and average path length in dynamic networks. In *2013 IEEE 13th International Conference on Data Mining* (pp. 867–876).

Yin, R., Yin, X., Cui, M., & Yinghan, X. (2019). Node importance evaluation method based on multi-attribute decision-making model in wireless sensor networks. *EURASIP Journal on Wireless Communications and Networking, 2019*(1), 1–14.

Yong, Y., Zhou, B., Chen, L., Gao, T., & Liu, J. (2022). Identifying important nodes in complex networks based on node propagation entropy. *Entropy, 24*(2), 275.

Zaki, A. A., Saleh, N. A., & Mahmoud, M. A. (2023). Performance comparison of some centrality measures used in detecting anomalies in directed social networks. *Communications in Statistics-Simulation and Computation, 52*(7), 3122–3136.

Zhang, S., Yu, H. et al (2022). Modeling and simulation of tennis serve image path correction optimization based on deep learning. *Wireless Communications and Mobile Computing*, 2022.

Zheng, H., Xue, M., Lu, H., Hao, S., Zhu, H., Liang, X., & Ross, K. W. (2017). Smoke screener or straight shooter: Detecting elite sybil attacks in user-review social networks. arXiv:1709.06916

Zverovich, V. (2021). Modern applications of graph theory. Oxford University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Dr. Essia Hamouda is an Associate Professor of Information and Decision Sciences at California State University San Bernardino. She holds a BSc and an MS degree in Industrial and Systems Engineering from The Ohio State University and the University of Florida, respectively. She also holds a Ph.D. in Engineering and Computer Science from the University of California, Riverside. Her research interests include networking, security, and privacy. She is also interested in resource optimization in sensor networks, the Internet of Things, and Smart Healthcare. She is the author of numerous research papers published in leading journals including the European Journal of Operational Research.

Dr. Mohsen ElHafsi is a Professor and renowned researcher in the field of Operations and Supply Chain Management at the University of California, Riverside. His research focuses on the design and management of complex and global supply chains. He is the author of numerous research papers published in leading journals including Management Science, Operations Research, Production and Operations Management, European Journal of Operational Research, Naval Research Logistics, and IIE Transactions. Dr. ElHafsi is a recipient of several fellowships and awards, including the prestigious distinguished Fulbright Fellow award for the year 2006/2007. He has been a Visiting Professor at several international institutions, including the "Ecole Centrale de Lille," one of France's elite engineering schools. Dr. ElHafsi has served and continues to serve as an editor, associate editor, and board member for several journals in his area of research. He has also served as Associate Dean of Graduate Programs and area coordinator. Dr. ElHafsi is ΦΚΦ Honor Graduate of the University of Florida, holding Ph.D. and M.Sc. degrees in Industrial and Systems Engineering.

Dr. Joon Son is an Associate Professor in the department of Information and Decision Sciences at California State University San Bernardino (CSUSB). Prior to joining CSUSB, he worked as a software engineer in Silicon Valley and as an information security scientist at Johns Hopkins Applied Physics Lab (JHAPL). His research interests include security policy specification, formal methods, ontology-based applications, and IT & Cyber Security education. He teaches a wide range of topics including cybersecurity, networking, database, programming, semantic web and business analytics.

Terms and Conditions