Centrality-Based Approaches for Connected Dominating Set Formation in Complex Networks

Yeşim Aygül
Department of Computer Engineering, İzmir Bakırçay University
İzmir, Türkiye
yesim.aygul@bakircay.edu.tr
0000-0003-0605-9604

Melike Karatay
Department of Computer Engineering, İstanbul Esenyurt University
İstanbul, Türkiye
melikekaratay@esenyurt.edu.tr
0000-0001-6941-4752

Onur Ugurlu
Department of Computer Engineering, İzmir Bakırçay University
İzmir, Türkiye
onur.ugurlu@bakircay.edu.tr
0000-0003-2743-5939

Urfat Nuriyev
School of Engineering, Karabakh University, AZ2600 Khankendi, Azerbaijan
Institute of Control Systems the Ministry of Science and Education of the Republic of Azerbaijan, AZ1141
Baku, Azerbaijan
urfat.nuriyev@karabakh.edu.az
0000-0002-3337-5859

***Abstract*—Identifying a Connected Dominating Set (CDS) is a fundamental problem in complex networks, with significant applications in communication efficiency, infrastructure resilience, and social influence analysis. A CDS is a subset of nodes that dominates the entire network, ensuring that every node is either part of the CDS or adjacent to it while also maintaining connectivity. In this study, we evaluate the effectiveness of traditional centrality metrics—Degree Centrality, Closeness Centrality, Betweenness Centrality, and Eigenvector Centrality —in selecting the smallest CDS across different network types, including road networks, power networks, and social networks. By comparing the CDS sizes obtained using different centrality metrics, we identify which metric is most effective in minimizing the CDS size across various network topologies. Our findings provide helpful information on the role of centrality metrics in CDS formation and their implications across diverse real-world network structures.**

Keywords—Complex Network, Connected Dominating Set, Centrality Metrics, Network Analyses, Node Ranking

I. INTRODUCTION

Determining an effective CDS is a fundamental challenge in complex network research, as it plays a crucial role in enabling efficient communication, structural robustness, and the spread of influence. A CDS is a subset of nodes that is both dominating—ensuring that every node in the network is either part of the CDS or adjacent to a CDS node—and connected, facilitating information propagation and structural integration within the network [1]. The efficient formation of a CDS is particularly significant in social networks, where CDSs play a vital role in influence maximization, ensuring that information dissemination reaches a broad audience with minimal intervention. Similarly, in power networks, the identification of a CDS is crucial for ensuring the efficient distribution of electricity and maintaining network stability. By selecting a set of strategically placed nodes (such as substations or key transmission points), the network's resilience to failures can be improved while minimizing redundancy. This approach is especially valuable in preventing widespread outages and optimizing maintenance operations. Besides, CDSs are essential for efficient routing and traffic management in road networks. A CDS can help identify key intersections or transit hubs that control the flow of traffic across the network. This can lead to better planning of transportation systems, improving traffic efficiency, reducing congestion, and ensuring that traffic data or emergency responses can reach all parts of the network.

Several approaches have been proposed in the literature to identify the CDS with the smallest size in complex networks. Traditional heuristics often focus on the use of greedy algorithms, which iteratively select nodes to cover the network, but they can lead to suboptimal solutions in terms of set size [2-4]. Other studies have explored metaheuristic techniques such as genetic algorithms, simulated annealing, and ant colony optimization, which are capable of finding near-optimal solutions by exploring a wider search space [5-7]. These methods, while effective in specific scenarios, can be computationally intensive, making them less feasible for large-scale networks. Recent advancements have also introduced centrality-based approaches, where centrality metrics are utilized to select nodes that have the highest structural importance in the network, thereby forming efficient CDSs while minimizing the number of nodes required [8-11].

In this study, we perform a comparative analysis of different centrality metrics - Degree Centrality (DC), Closeness Centrality (CC), Betweenness Centrality (BC), and Eigenvector Centrality (EC) - to identify which metrics most effectively minimize the size of the CDS, while ensuring connectivity across different network types, including road networks, power networks, and social networks.

The rest of the paper is organized as follows. In Section II, we provide the necessary background. In Section III, we describe the methodology employed to evaluate the centrality metrics and the used network datasets. Section IV presents the comparative evaluation of the centrality metrics in minimizing CDS size across different network types. Finally, in Section V, we conclude the paper with a summary of the findings and discuss potential directions for future research.

II. Basic Definitions And Preliminaries

In this section, we provide fundamental definitions and preliminary concepts essential. A complex network can be formally represented as an undirected graph $G = (V, E)$, where V denotes the set of nodes (vertices) and E represents the set of edges (links) connecting the nodes [12]. The degree of a node in a graph is the number of edges incident to it. In other words, it represents the number of connections a node has to other nodes in the network. The degree of a node v is denoted as $deg(v)$. A path in a graph is a sequence of nodes in which each consecutive pair of nodes is connected by an edge. A path can be described by its sequence of nodes or edges.

A Dominating Set (DS) of G is a subset $S \subseteq V$ such that every node in V is either in S or has at least one neighbor in S. A CDS is a DS in which the induced subgraph $G[S]$ is connected. The minimum CDS problem involves finding the smallest possible subset S that satisfies the CDS property, which known to be a NP-hard problem [1].

Centrality metrics quantify the importance of nodes within a network. In this study, we evaluate the following traditional centrality metrics:

Degree Centrality: Measures the direct influence of node v based on its number of connections [13]:

$$DC(v) = \deg(v)$$

Closeness Centrality**:** Reflects how efficiently a node can access all other nodes in the network [14]:

$$CC(v) = \frac{1}{\sum_{u \in V} d(v, u)}$$

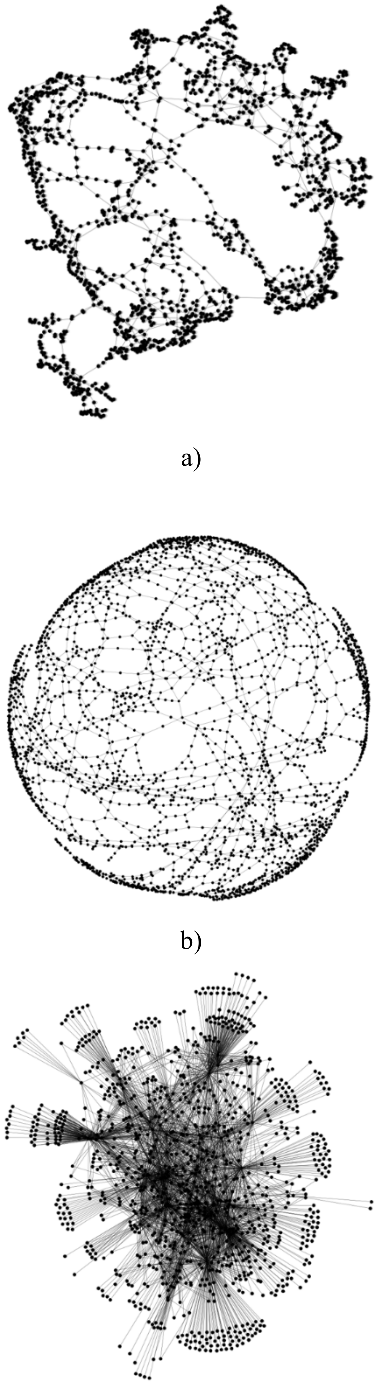
where $d(v, u)$ is the shortest path distance between nodes v and u.

Betweenness Centrality**:** Identifies nodes that act as intermediaries in shortest paths between other nodes [14]:

$$BC(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where σ_{st} is the number of shortest paths between nodes s and t, and $\sigma_{st}(v)$ is the number of those paths passing through v.

Eigenvector Centrality: Considers not only the direct connections of a node but also the importance of its neighbors. A node with high EC is connected to other well-connected nodes, amplifying its influence within the network [15].



a)

b)

c)

Fig. 1. Illustration of a) power-bcspwr09, b) road-minnesota, and c) soc-hamsterster.

III. Material And Method

This section outlines employed solution framework for constructing the CDS and the used dataset.

A. Dataset

The dataset used in this study consists of 15 real-world networks representing various domains, including power grids, road networks, and social networks [16]. These networks differ significantly in terms of size, structural complexity, and connectivity patterns. Fig. 1 illustrates representative real-world networks from three distinct domains: power (power-bcspwr09), road (road-minnesota), and social (soc-hamsterster).

Table I summarizes the fundamental properties of these networks. The number of nodes ranges from 685 to 129164, while the number of edges varies from 1417 to 165435, reflecting the differing scales and interconnections of the networks. The average path length (*APL*), indicating the typical number of hops between two nodes in the network, spans from 34 to 3790. Basic properties of 15 real-world networks.

TABLE I. BASIC PROPERTIES OF 15 REAL-WORLD NETWORKS

Networks	*n*	*m*	*APL*	*Clus*
power-685-bus	685	1967	34	0,1725
power-1138-bus	1138	2596	361	0,0866
power-bcspwr09	1723	4117	357	0,0759
power-bcspwr10	5300	13571	268	0,0880
power-US-grid	4941	6594	1229	0,0801
road-euroroad	1174	1417	340	0,0167
road-luxembourg-osm	114599	119666	22426	0,0006
road-minnesota	2642	3303	129	0,0160
road-usroads-48	126146	161950	3049	0,0254
road-usroads	129164	165435	3790	0,0257
soc-advogato	6551	43427	601	0,1953
soc-fb-pages-government	7057	89455	312	0,4109
soc-fb-pages-politician	5908	41729	448	0,3851
soc-fb-pages-tvshow	3892	17262	552	0,3737
soc-hamsterster	2426	16630	170	0,5375

TABLE II. THE CDS SIZES FOUND BY DIFFERENT CENTRALITY METRICS

Networks	*DC*	*CC*	*BC*	*EC*
power-685-bus	589	680	523	683
power-1138-bus	945	1125	864	1133
power-bcspwr09	1456	1722	1337	1720
power-bcspwr10	5160	5293	4345	5289
power-US-grid	4247	4935	3809	4931
road-euroroad	1174	1174	1174	1174
road-luxembourg-osm	111967	114598	111301	114514
road-minnesota	2642	2642	2642	2642
road-usroads-48	124941	126142	117081	126142
road-usroads	129164	129164	129164	129164
soc-advogato	6551	6551	6551	6551
soc-fb-pages-government	6953	7042	3642	7046
soc-fb-pages-politician	5611	5905	3448	5905
soc-fb-pages-tvshow	3617	3887	2664	3891
soc-hamsterster	2426	2426	2426	2426

Furthermore, the clustering coefficient (*Clus*), which measures the tendency of nodes to form tightly connected local clusters, varies widely from 0.0006 to 0.5375. These differences in *Clus* suggest that some networks exhibit strong local connectivity, while others are more loosely connected, indicating diverse structural characteristics across the networks. These variations in network characteristics are influenced by various factors, such as geographical constraints, operational requirements, and the specific dynamics of each network's domain. The dataset's structural diversity makes it suitable for testing the performance of different centrality in forming a CDS.

B. Solution Framework

To investigate the effectiveness of different centrality metrics in constructing the CDS across various networks, a centrality-based solution approach is employed. The process involves calculating the CDS by iteratively selecting nodes based on their centrality scores. The goal is to identify a small subset of nodes that can dominate the entire network while ensuring connectivity. Specifically, for each centrality metric (e.g., DC, CC, BC, or EC), the centrality score of each node is calculated. Then, the nodes are sorted based on their scores from highest to lowest. Beginning with the node that has the highest centrality score, nodes are iteratively added to the CDS. The process continues until all nodes are covered and the CDS remains connected. Finally, the size of the resulting CDS is reported. By applying this procedure across different centrality metrics, the effectiveness of each metric in identifying the key nodes for forming the CDS is evaluated.

IV. RESULTS

The computational results on the 15 real-world networks are presented in this section. A comparative evaluation of the performance of various centrality metrics, including DC, CC, BC, and EC, in identifying a CDS is provided in Table II, which reports the CDS size for each network.

For the road networks, the results show that nearly all centrality metrics result in the same size of the CDS, which suggests that the structural characteristics of road networks may not significantly favor one centrality metric over another. This may be attributed to the relatively regular, grid-like structure of these networks, where node importance is uniformly distributed across the network.

This observation emphasizes that road networks, due to their predictable and structured layout, do not present the same level of complexity or variation in node relevance as more intricate network types, such as social or power networks. In contrast, the analysis of social networks reveals that the BC produces the smallest CDS size in most networks, such as soc-fb-pages-government, soc-fb-pages-politician, and soc-fb-pages-tvshow. This highlights BC's ability to identify nodes that serve as critical intermediaries in the network, facilitating efficient communication and control. However, for certain social networks such as soc-advogato and soc-hamsterster, all centrality metrics yield the same CDS size, implying that in these specific networks, the nodes hold relatively similar structural significance. These findings suggest that while BC generally outperforms its counterparts in social networks, the diversity in network structure across different social platforms can lead to varying performances of centrality metrics, with some networks presenting more uniform node centrality profiles.

When examining the power networks, BC consistently yielded the smallest CDS sizes across all power grid networks. Power grids are often large, highly complex networks that include numerous nodes with varying degrees of importance. In these networks, BC’s effectiveness stems from its ability to identify nodes that frequently lie on the shortest paths between other nodes, thus playing a critical role in the network's connectivity. This property makes BC the most effective centrality metric for minimizing the CDS size in power networks, as it ensures that only the most crucial nodes—those that connect different parts of the network—are selected, thus eliminating redundancy and maintaining essential network coverage.

Overall, the results demonstrate that BC outperforms other centrality metrics (DC, CC, BC) in minimizing the CDS size, especially in power and social networks. This can be attributed to BC's feature to capture nodes that frequently lie on the shortest paths between other nodes. By prioritizing structurally critical nodes regarding shortest paths, BC enables the formation of a compact CDS, effectively minimizing redundancy while preserving network coverage. The ability of BC to capture "bottleneck" nodes is crucial for optimizing CDS in these large-scale, intricate systems, which often feature highly interconnected structures. The performance of BC in social and power network types underscores its suitability for networks with complex topologies, such as social and power networks, where node importance is often dependent on connectivity and interaction with other nodes. However, it is important to note that all centrality metrics perform similarly for the road network. This indicates that in road networks, factors such as geographical layout and connectivity may outweigh the contribution of centrality measures in determining node importance. Therefore, while BC remains the most effective metric for complex networks like power grids and social networks, its advantage is less pronounced in road networks.

V. Conclusion

The selection of a CDS significantly influences the structural robustness and efficiency of various real-world networks. In this study, we have evaluated the effectiveness of traditional centrality metrics—DC, CC, BC, and EC—for construing a CDS across different network types, including road networks, power networks, and social networks. Our results show that BC performs the best in minimizing the CDS size across most network types. Specifically, in power and social networks, BC leads to the smallest CDS sizes, highlighting its ability to identify key nodes that provide both coverage and connectivity effectively.

The findings have significant implications for optimizing network design and improving efficiency in applications such as communication, traffic management, and influence maximization. Future research could further explore the combination of multiple centrality measures to improve the performance of CDS formation in more complex or dynamic networks.

Acknowledgements

Y. Aygul thanks TUBITAK for its scholarship support under the BIDEB 2211-A program.

References

[1] Du, D. Z., & Wan, P. J. (2012). *Connected dominating set: theory and applications* (Vol. 77). Springer Science & Business Media.

[2] Butenko, S., Cheng, X., Oliveira, C. A., & Pardalos, P. M. (2004). A new heuristic for the minimum connected dominating set problem on ad hoc wireless networks. *Recent developments in cooperative control and optimization*, 61-73.

[3] Ugurlu, O., Tanir, D., & Nuri, E. (2016). A better heuristic for the minimum connected dominating set in ad hoc networks. In *2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT)* (pp. 1-4). IEEE.

[4] Li, Y., Thai, M. T., Wang, F., Yi, C. W., Wan, P. J., & Du, D. Z. (2005). On greedy construction of connected dominating sets in wireless networks. *Wireless Communications and Mobile Computing*, *5*(8), 927-932.

[5] Hedar, A. R., Ismail, R., El-Sayed, G. A., & Khayyat, K. M. J. (2019). Two meta-heuristics designed to solve the minimum connected dominating set problem for wireless networks design and management. *Journal of Network and Systems Management*, *27*, 647-687.

[6] Jovanovic, R., & Tuba, M. (2013). Ant colony optimization algorithm with pheromone correction strategy for the minimum connected dominating set problem. *Computer Science and Information Systems*, *10*(1), 133-149.

[7] Dahmri, H., Bouamama, S., & Balbal, S. (2024). A Greedy Simulated Annealing-based Multiobjective Algorithm for the Minimum Weight Minimum Connected Dominating Set Problem. *Engineering, Technology & Applied Science Research*, *14*(5), 16686-16691.

[8] Meghanathan, N. (2015). Use of centrality metrics to determine connected dominating sets for real-world network graphs. In *2015 12th International Conference on Information Technology-New Generations* (pp. 243-248). IEEE.

[9] Meghanathan, N., Rahman, M. A., & Akhter, M. (2019). Centrality Metrics-Based Connected Dominating Sets for Real-World Network Graphs. In *Strategic Innovations and Interdisciplinary Perspectives in Telecommunications and Networking* (pp. 1-29). IGI Global.

[10] Meghanathan, N. (2014). Centrality-based connected dominating sets for complex network graphs. *International Journal of Interdisciplinary Telecommunications and Networking (IJITN)*, *6*(2), 1-24.

[11] Rehm, H., Kassouf-Short, R., & Rombach, P. (2025). Generating Dominating Sets Using Locally Defined Centrality Measures. IEEE Transactions on Networking.

[12] Diestel, R. (2024). Graph theory. Springer (print edition); Reinhard Diestel (eBooks).

[13] Freeman, L. C. (1978). Centrality in social networks conceptual clarification, Social Networks 1 (3), 215–239.

[14] Freeman, L. C. (1977). A set of measures of centrality based on betweenness. Sociometry.

[15] Solá, L., Romance, M., Criado, R., Flores, J., García del Amo, A., & Boccaletti, S. (2013). Eigenvector centrality of nodes in multiplex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *23*(3).

[16] Network Repository, March 2025 [online] Available: https://networkrepository.com/power.php