Social Network Analysis Road Way Network Interactions

Mohammad Kamil, Hamza Jamil[†] and Hammad Noukhez[†]

Department of Computer Science, Institute of Business Administration, Karachi, University Road, Karachi,Sindh, Pakistan.

*Corresponding author(s). E-mail(s): m.kamil.29464@khi.iba.edu.pk;
Contributing authors: h.jamil.29454@khi.iba.edu.pk;
h.noukhez.27225@khi.iba.edu.pk;
[†]These authors contributed equally to this work.

Abstract

Exploring social networks is like using a smart way to understand how lots of people connect, especially when these connections happen on the road. Our research focuses on comprehensive examination within the distinctive context of California's road network, with a primary focus on four critical aspects: Network Resilience, Roles Identification, Influence Mapping, and Network Structure Analysis. Our goal is to reveal the complex connections between these elements, providing a better understanding of the unique challenges and complexity inherent in California's road system. We introduce a novel methodology to efficiently discover critical nodes, termed network bottlenecks, crucial for information flow propagation. We identify relative checkpoints as probable sources of major inflows towards bottlenecks, enabling proactive surveillance and control over information outbursts. Network structure analysis involves closely examining how the road network is organized, revealing patterns and relationships between different nodes to understand the overall layout and connectivity. Assessing network resilience allows us to understand the network's ability to adapt and recover from disruptions, contributing to its robustness. Influence mapping explores the impact of specific nodes on information flow within the road network, providing insights into areas crucial for shaping information dynamics. Roles identification aims to uncover distinct functionalities within the network, identifying nodes with unique roles such as traffic hubs or controllers. Throughout our experiments, strategically chosen checkpoints serve as monitoring stations, contributing not only to improved traffic control but also to a deeper understanding of the network's dynamics in terms of structure, resilience, influence, and functional roles.

Keywords: Social Network Analysis, Transport Planning, Checkpoints, Centrality Measures,Network Resiliance, Traffic Flows

1 Introduction

Social Network Analysis (SNA) is like a tool that helps us understand social connections. It uses ideas from graph theory, a part of math that deals with representing and understanding networks. Imagine it as a way to draw and measure the relationships between people or things in a social group. SNA allows us to use math concepts to figure out how these connections work and what makes them unique. It's a helpful way to study and make sense of the different relationships in social systems or we can say that network representations are like maps that help us understand how things are connected. An Internet network or a website link network, they show us how different parts are linked together. In Biology and Chemistry, these networks are used to explain how proteins and atoms interact with each other [1].

In sociology, experts have always thought of social networks as ways to show how people in a community are connected to each other [2].Experts from fields like engineering, mathematics, biology, geography, physics, and other sciences have used analytical methods to understand social networks. These networks consist of nodes connected by links, representing interactions and relationships. Social Network Analysis (SNA) is a tool that helps researchers explore and make sense of the structures in networks, whether they are one-way or two-way connections. It's especially useful for studying networks with different types of relationships, like weighted, signed, and bipartite networks. In simpler terms, SNA allows experts from various fields to investigate and analyze the complexities within different connected systems [3].So in the case of transport network, Smart cities are getting a lot of attention, and Intelligent Transportation Systems (ITS) are a big part of making them work better. One key focus is dealing with sudden traffic problems, like during epidemics, that can lead to serious air pollution. To make transportation smarter, we need to use things like information fusion, computers, sensors, and wireless communication. For example, the whole transportation system—roads, airlines, metros, railways—uses a method called Social Network Analysis (SNA) to understand its complicated structure. This helps figure out things like traffic conditions, the best routes, sharing information, being aware of situations, detecting traffic issues, and understanding how people move around. Connecting SNA with ITS helps us solve these transportation challenges in smart and effective ways [4].

We introduce few main concepts in analysis of road network. The first one is the centrality of nodes in a road network, which provides a distribution. The centrality is widely used to detect important nodes and find nodal characteristics in networks.We compute three types of centralities; degree, closeness, and betweenness. Generally, the centralities explain how nodes play an important role in a social network structure. In a road network, degree of a node can be provides connectivity and popularity of an intersection with respect to spatially neighboring intersections. In similar fashion, the closeness and the betweenness of a node provide alternative routes with in the network.

Secondly, we have understanding Network Resilience. We use Social Network Analysis to figure out how tough a road network is. It's like checking how different parts (like intersections, roads, or nodes) are connected and how they affect the whole network. This helps us find important nodes or parts. If something happens to these critical points, it could really mess with how well the whole network works.

Next we have roles identification that involves using special algorithms to find groups or roles of nodes who discuss transportation and roads in similar ways. The goal is to uncover the connections and shared interests within these communities. It's like identifying different groups who have similar interest about transportation and roads, allowing us to understand their common interests and interactions in communities.

Further we have influence mapping that plays very important role in network. The goal is to identify and understand the nodes within the road network who exert a significant impact on netwprk or decision making. These influential nodes act as crucial points that shape the flow of information and decision paths within the road network. By mapping out these influential elements, we gain insights into how certain nodes drive discussions, affect network, and influence the overall dynamics of the road network.

For this project we choose california road network [5] so within the context of California's road network, we concentrate on four key aspects: Network Resilience, Roles identification, Influence Mapping, and Network Structure Analysis. The primary objective is to uncover intricate connections among these elements, offering a deeper insight into the distinct challenges and complexities within California's road system. We present an innovative methodology to effectively pinpoint crucial nodes, known as network bottlenecks, vital for the flow of information. Additionally, we identify relative checkpoints as potential sources of significant information inflows towards these bottlenecks, allowing for proactive surveillance and control over information dissemination.

The next part of our paper is set up like this: In second section, we will disuss related work or literature review of papers. Then, in third section we will discuss about proposed methodology and methods we're using, and where these methods can be helpful in context of road network. Fourth section will all about dataset discussion and finally we will have conslusion section.

2 Literature Review

This section is all about background of transport networks that how transportation networks, like roads and pathways, are put together. We're have read exisiting papers and got some ideas from how people connect in social networks to sort out the details of transportation systems. It is like solving a puzzle of connections in our roads. By combining different ways or techniques from studying social networks, we're getting a new perspective on how everything in transportation is linked. Recently, this mix of ideas has been helping us see the hidden structures and figure out the important parts that make transportation work well. It's like smart ideas from how people connect online to make our roads and transportation systems more efficient by identifying importnat nodes,checkpoints and roles that are very importance and has good influence in road network.

Paper [6] focuses on significantly contributes to the application of social network analysis (SNA) by focusing on the identification of critical nodes termed "network bottlenecks" within intricate social connections. These bottlenecks, pivotal in information flow, are extended into the extraction of "relative checkpoints." These checkpoints serve as sources for substantial inflows towards the bottlenecks, crucial for controlling information distribution. The proposed methodology showcases broad applications across domains, including road traffic monitoring, tracking extremist content, scrutinizing fake news, and unveiling online terrorist movements.

Paper identifies certain traffic hotspot checkpoints or regions and related to sets of nodes known as checkpoints. These checkpoints not only enhance traffic control but also contribute to fostering sustainable mobility. Additionally, the research delves into the utility of social network analysis for evaluating diverse network types and complex behaviors, employing metrics such as centrality, clustering-coefficient, and community detection.

In paper [6] practical experiment done on the California road network, emphasizing its scale-free property with significant hubs. Properties of the network like degree distribution, diameter, density, and average path length are analyzed.

In research paper [7] focuses on utilizing Social Network Analysis to evaluate and prioritize transportation intersections, aiming to identify those crucial for traffic flow and overall network efficiency. Going beyond traditional traffic studies, the study provides a fresh perspective on the significance of intersections within a network. Conducted in Louisiana, the research identifies critical intersections through SNA in three case studies, showcasing the model's effectiveness in pinpointing areas for infrastructure improvement. Notably, the study highlights successful upgrades, such as transforming a signal.

The paper [8] presents a framework integrating network theory methods for assessing transportation network robustness from a topological perspective. The main objectives of this paper focuses on developing a framework for measuring network performance under different stress levels, understanding constraints due to spatial embedding, and characterizing percolation transitions for evaluating urban resilience. The research investigates 13 city and 3 state transportation networks, revealing unique spatially embedded features that challenge certain theoretical methods. So, Critical percolation thresholds identify robustness features, and reducing network spatial complexity enhances empirical simulation alignment with theoretical results, providing insights into the spatial essence of infrastructure networks. This framework is validated across different transportation networks for robustness analysis to general network.

The proposed paper sets itself apart from existing literature, as illustrated in Table 1. Our system utilizes the California road network [5] and is guided by specific key objectives. Within the scope of this paper, our primary emphasis lies on roles identification, influence mapping, network resilience, and an in-depth exploration of network structure. Notably, our approach fills a significant void in the current body of knowledge, as there is no single paper to date that comprehensively addresses all these aspects.

3 Proposed Methodology

In this section, we discussed all about that how we construct the road network, measure centrality, look at Network Resilience, Influence Mapping, and examine the network structure. To understand roads better, we start by creating a road network using data from California. Then, we take a close look at different techniques to analyze it. This process helps us to understand the details of the road network and the important techniques that are listed below:

Construction of the Road Network

We build road networks using the California dataset, where we pick intersections and junctions as nodes and the road segments connecting these points as links in the network. Once this is done, we convert our dataset into a graph, and from there, we delve into identifying nodes, edges, and implementing various techniques. Let's denote the

Table 1: Comparison of Techniques of Proposed and Existing Research

Techniques	Paper 1 [6]	Paper 2 [7]	Paper 3 [8]	Proposed Paper
Roles Identifications	×	×	×	✓
Network Resilience	×	✓	×	✓
Influence Mapping	✓	✓	✓	✓
Networ Structure Analysis	✓	✓	×	✓

graph of our road network as G=(V,E), where V represents a set of nodes and E represents a set of edges in G.he visual representation of our project's high-level architecture is presented in Figure 1. This framework guides our project implementation, ensuring we follow a structured approach in line with our proposed design.

Node Centrality

In the world of social network analysis, a key objective is pinpointing the most influential nodes within the network structure. Centrality is one of the measurements to rank nodes. Most widely used centrality are three: Degree, closeness, and betweennes centrality.

- **Degree Centrality**

  In social network, Degree centrality that tells us how connected a point is in a network. In simple words, It tells us or counts the number of direct links at some point has with other points in a road network. In our road networks, a higher degree means more connections at a junction mean at some point where it has certain counts. So it mean points with lots of connections might have more traffic compared to those with fewer connections as we see in our graph. In figure 2 we can see degree centrality of nodes and frequency in more proper way. Node with degree centrality around 0.0004 gave more frquence as we can see in our graph.So we can say that nodes with centrality have more frequency distribution.

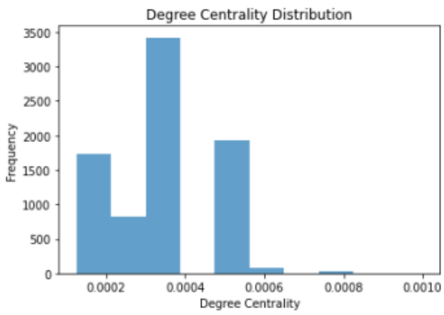

Fig. 2: Degree Centrality

- **Closeness Centrality**

  Closeness centerality measures how closely connected a node is to others, specifically by calculating the shortest maximum path between two nodes. So by identifying the shortest path across the entire network structure, closeness takes into account the network's global connectivity as we can see in Figure 3. It provides insights into how efficiently information or influence can spread from one node to any other node in the road network, highlighting the node's overall proximity of network. So in given graph we can see the closeness of nodes and their connection between nodes. In california road network, it looks like outlier where some nodes are more close and some are not too much.

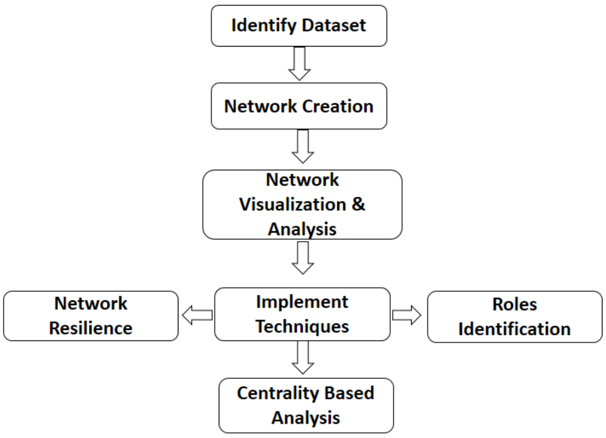

Fig. 1: Proposed High Level Architecture

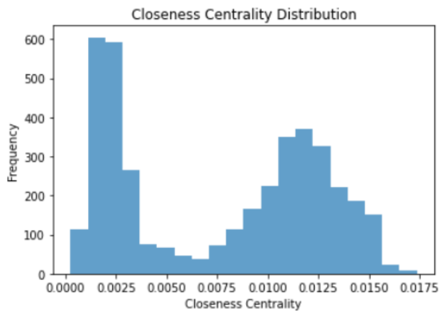

Fig. 3: Closeness Centrality

- **Betweenness Centrality**

  Betweenness centrality helps us understand how a node can bring together other nodes that aren't directly connected. It counts how many times a node acts as a bridge between others. This is a global measure because it looks at the whole network to find these connecting paths. Essentially, it shows us which nodes play a key role in linking different parts of the network, making them important for overall connectivity. In the context of a road network, betweenness centrality provides insights into how certain intersections act as crucial connectors between roads that don't directly meet. It counts how often an intersection serves as a vital link between other intersections or road segments.

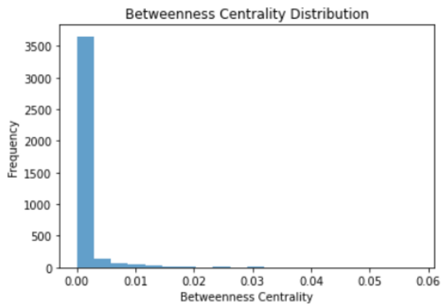

Fig. 4: Betweenness Centrality

Roles Identifications

Roles identification in a road network are very important that is used to identify the distinct functions that specific intersections or nodes fulfill within the overall network structure. Figure 11 shows roles identification in california road network while figure 12 shows random natwork roles identification so in the context of road network various roles emerge like hub junctions act as central meeting points, connecting multiple roads and facilitating significant traffic movement. Connector nodes play a critical role in linking disparate parts of the network, enhancing overall connectivity. Traffic distributors efficiently manage the flow of vehicles, optimizing traffic distribution. High-degree nodes, with a multitude of connections, often signify areas of heightened activity. Peripheral nodes at the network's outskirts may have specific roles in managing incoming or outgoing traffic. Bottleneck intersections, where traffic tends to accumulate, become focal points for analysis.

Understanding Network Resilience

Network resilience in a road network is to grasp how its components, especially critical intersections, are interconnected. This involves evaluating the contributions of individual nodes to the overall performance of the network. By understanding the network's resilience, we aim to assess its ability to endure disruptions and challenges. This insght is essential for ensuring the road network remains robust, reliable, and capable of maintaining optimal performance even in adverse conditions.

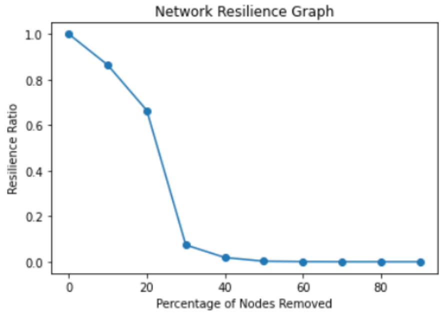

Fig. 5: Network Resilience

Influence Mapping

One of the key objective of our project, Influence mapping examining the impact of nodes and edges within the network. This includes mapping the influence patterns to understand how individual nodes and edges contribute to the overall network dynamics. The focus is on assessing both node influence and edge influence, providing valuable insights into the factors shaping traffic flow and connectivity in the road network.

Network Structure Analysis

In this section we explore the overall structure and organization of the network. This involves employing various graph metrics like degree distribution and network connectivity to understand the intricacies of the network's layout. Degree distribution provides insights into how nodes are connected, while network connectivity assesses the overall connectedness of the road network. Network properties values are explicitly listed in Table 2, offering a comprehensive overview of the california road network's and random Network properties. Also there are graphs visualization of degree distributions as well where we can see in Figure 6 that nodes with degree three are more than others.

Network Properties	California Network Values	Random Network Values
Nodes	1965206	1000
Edges	2766607	4000
Average Clustering Coefficient	0.0463	0.0081
Average Degree	2.63	8.0
Diameter of Component	849	6
Giant Component	1957027	1000

Table 2: Comparison of California Network and Random Network

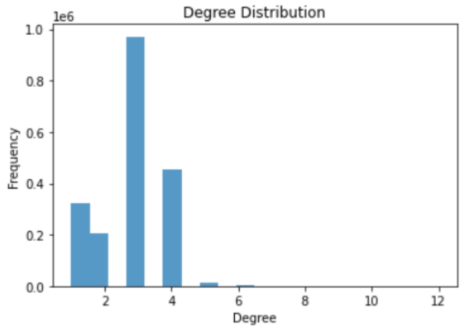

Fig. 6: Degree Distribution

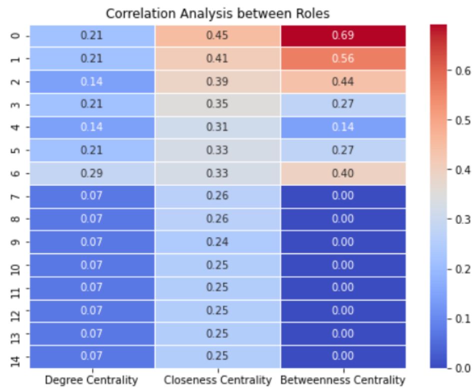

Fig. 7: Analysis of Correlation Roles

Analysis of Correlation Roles

In this section we discusses the correlation of roles that identified in network. Figure 7 shows initially 15 roles correlation where we have correlation matrix and values, we can see different values in matrix and values close to 1 indicate a strong positive correlation, values close to -1 indicate a strong negative correlation, and values close to 0 indicate a weak correlation.

The correlation matrix helps understand how these centrality measures are related to each other in the road network graph. For example, if degree centrality and closeness centrality have a high positive correlation, it suggests that nodes with high degree centrality tend to be centrally located in the network.

Similarly, correlations between other centrality measures provide insights into the network's structure and characteristics

4 Results and Discussion

In this section, we present the application of our proposed model to reveal noteworthy findings within the complex transportation network, serving as the cornerstone of our empirical study. To conduct a quantitative analysis, we specifically identify certain bottlenecks. Our dataset focuses on the California road network, represented as an undirected graph comprising 1,965,206 nodes and 2,766,607 edges [5]. This dataset forms the foundation of our empirical exploration, allowing us to extract meaningful insights into the dynamics and performance of the transportation network. We extract Network properties and summarized in Table 2.

In our study, we had specific goals. We looked into roles identification, which is like figuring out the important nodes in the network. Our results showed that some nodes and edges have more influence than others, playing important roles in how the network works. We also checked how well the network can handle challenges, and we found out that it can be impacted quite a bit. We did a lot of calculations and also created a random network to compare with our original one. The results were different between them as shown below figures.

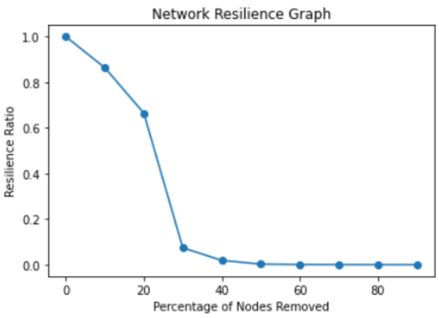

Fig. 8: Random Network Resilience

Figure 10 gives a visual comparison, showing that the way connections are spread out in the random network is different from our original one. This helps us see that the design of the network, how things are connected, has a big impact on the roles and overall structure of the network.

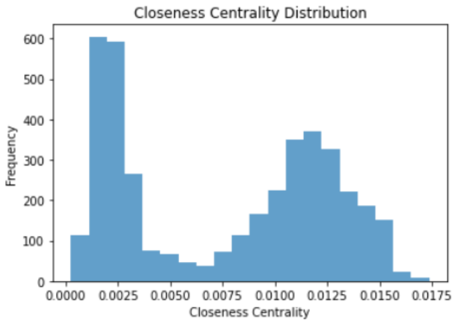

Fig. 9: closeness centrality distribution

Looking at Figures 11 and 12, it's clear that the roles and their distribution in our original network are quite different from those in the randomly generated network.

In the Role Identification Graph, there are clusters or groups of nodes that collaborate closely, each having a specific role in shaping the network. It's like these nodes team up to influence how the network functions. On the other side, the Random Network Role Identification doesn't display such clear patterns as we can see in figure 12.

The roles are scattered, and it's not straightforward to pinpoint specific groups or nodes that hold significant influence in this case. In simpler terms we can say that, our original network that mean california network has organized in such a way where groups working together, while the randomly generated network lacks this clear structure network, making it harder to identify specific roles or influential nodes. So the visual representation of graphs helps us understand how visual representation of original network leads in more organized roles and their relationships among nodes and can be identified easily.

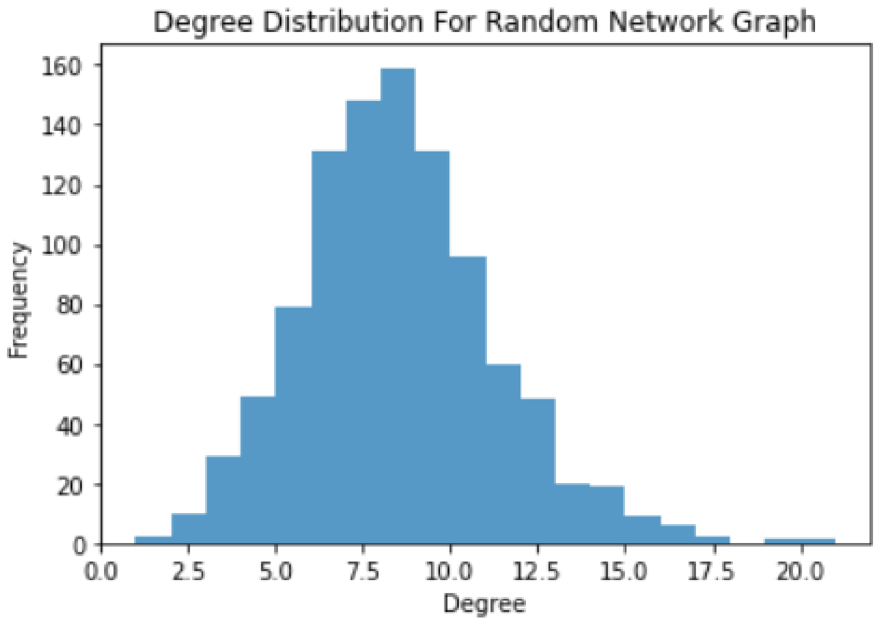

Fig. 10: Degree Distribution Graph For Random Network

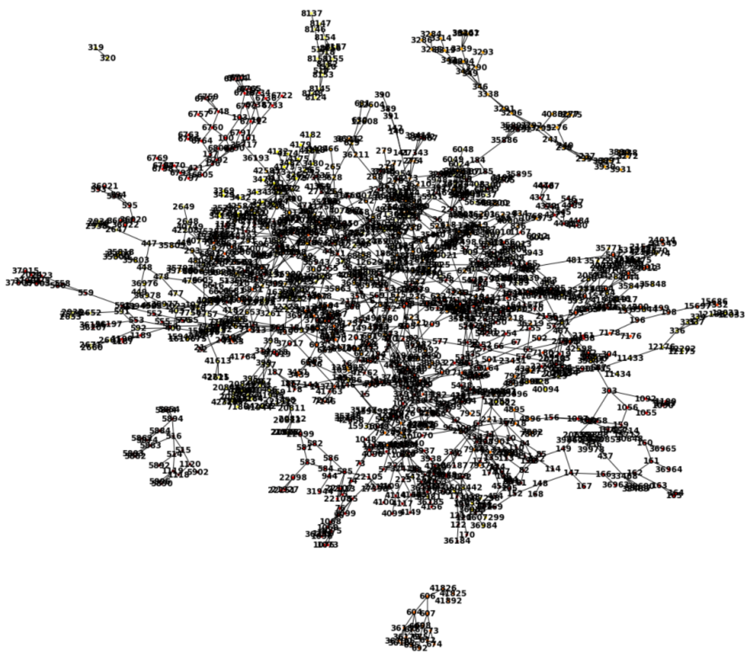

Fig. 11: Role Identification Graph

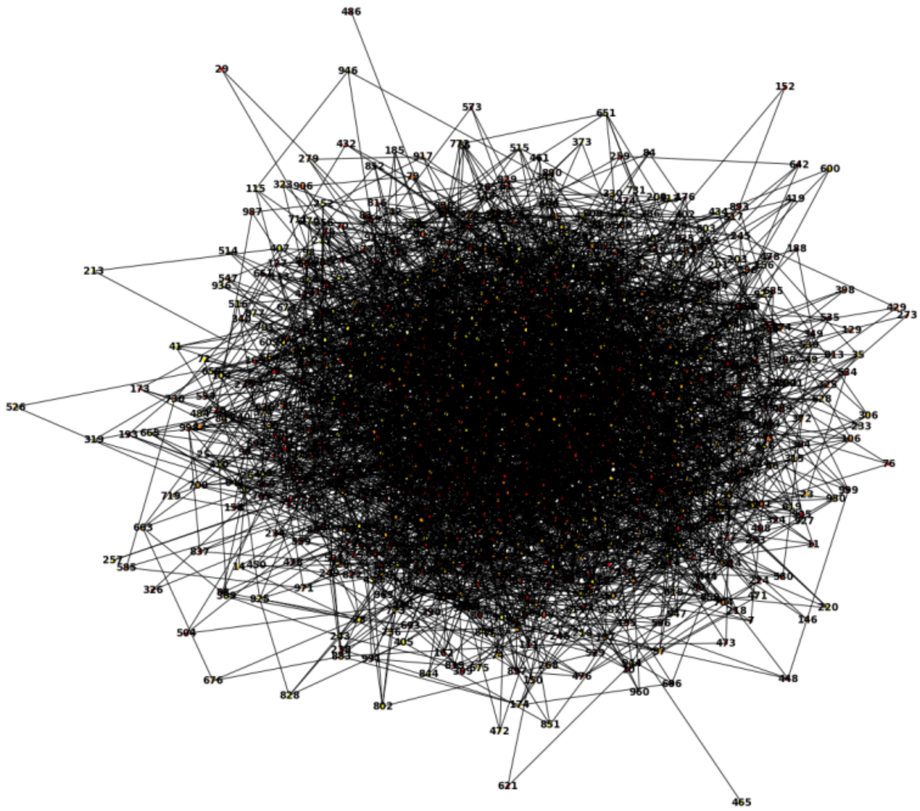

Fig. 12: Random Network Role Identification

5 Conclusion

We propose road network analysis applying social network analysis methods which are widely used in the other areas as well. First we analyz ethe network structure in which we discussed degree distribution, Diamaer,Giant Component, Avergae clustering coefficients and Average degree after that we discussed the three types of centralities, degree centrality, closeness centrality, and betweenness centralities. Those node centralities describe how a node important in a network, and we extract the histogram of those centrality values respectively.

We applied these methods using data of California roads network to show how our approach works in real situations. The simulation results revealed statistical and structural features of the actual network. Our experiments provided valuable insights into an information space, uncovering bottlenecks and checkpoints through graphical representations. This practical application helps justify the effectiveness of our proposed techniques in a real-world context.

References

[1] D.D. Wu, X. Hu, in *Proceedings of the 2006 ACM symposium on Applied computing* (2018), pp. 185–189

[2] L.C. Freeman, et al., Centrality in social networks: Conceptual clarification. Social network: critical concepts in sociology. Londres: Routledge **1**, 238–263 (2019)

[3] K.K. Kapoor, K. Tamilmani, N.P. Rana, P. Patil, Y.K. Dwivedi, S. Nerur, Advances in social media research: Past, present and future. Information Systems Frontiers **20**, 531–558 (2018)

[4] N.K. Jain, R. Saini, P. Mittal, A review on traffic monitoring system techniques. Soft computing: Theories and applications: Proceedings of SoCTA 2017 pp. 569–577 (2019)

[5] SNAP: Network datasets: California road network — snap.stanford.edu. https://snap.stanford.edu/data/roadNet-CA.html. [Accessed 08-12-2023]

[6] S. Jain, A. Sinha, Social network sustainability for transport planning with complex interconnections. Sustainable Computing: Informatics and Systems **24**, 100351 (2019)

[7] I.H. El-Adaway, I. Abotaleb, E. Vechan, Identifying the most critical transportation intersections using social network analysis. Transportation planning and technology **41**(4), 353–374 (2018)

[8] S. Dong, A. Mostafizi, H. Wang, J. Gao, X. Li, Measuring the topological robustness of transportation networks to disaster-induced failures: A percolation approach. Journal of Infrastructure Systems **26**(2), 04020009 (2020)