Network Size Reduction Using Clustering Method for Bridge Transportation Network

Purba, Denissa
dpurba2, 667758459
University of Illinois at Urbana-Champaign, US

SUMMARY:

Bridges are critical assets in the transportation network to support economic growth, social development, daily traffic needs, emergency response, and as an indicator to define the success of a specific regional. Bridge Asset Management has emerged and studied extensively to manage bridges' risk and performance in its life-span. However, there are two main challenges found in almost every examination of the bridge's risk management, i.e., expandable size large-scale network due to a scenario-based approach and multiple performance simulations of the bridges; and highly connectivity redundancy in transportation network. Clustering techniques have been proposed in many literatures to address such problems of scale in general critical infrastructure network. However, there are few works of literature in clustering techniques focusing on the criticality of the bridge in a highly redundant network. The project proposes a clustering framework to reduce the complexity of the network and provide information on the criticality of bridges using the Girvan-Newman algorithm.

Keywords: Network Clustering, Edge Betweenness, Grivan-Newman, Bridge Criticality, Network Redundancy

1. INTRODUCTION

Bridges are critical assets in the transportation network to support economic growth, social development, daily traffic needs, emergency response, and as an indicator to define the success of a specific regional. However, they are not only continuously deteriorating throughout service, but also particularly vulnerable to extreme hazards. Severe structural damage and significant functionality loss could occur because of the collapse or closure of bridges. These losses disrupt daily economics activities, but also hinders the short-term emergency response and post-disaster recovery (Chang Nojima 1998, Nojima 1998), resulting in substantial socioeconomic losses (Sohn et.al 2003). Hence, civil engineers are challenged to provide bridge transportation network with a new targeted performance level that supports these roles in daily operational and minimize the risk to sustain its function during a hazard.

In civil engineering field, Bridge Asset Management has emerged and studied extensively to manage bridges' risk and performance in its life-span. This research includes and combines holistic and complex examinations, i.e., performance assessment and health monitoring; design planning for targeted performance; mitigation planning for emergency response and regional resiliency; recovery strategy; and operation and maintenance during the operational span. However, developing an effective algorithm for solving and handling each approach is challenging. There are two main challenges found in almost every examination of the bridge's risk management.

First, the evaluation of bridge asset management deals with an expandable size large-scale network due to a scenario-based approach and multiple performance simulations of the bridges. The network-based approach is necessary to describe the interdependency and sharing capacity of each infrastructure while serving its function. Sohn et al. (2003) declared that the most critical links are not always subject to the most significant economic consequences of the network. On the other hand, a scenario-based approach is also commonly used in risk assessment to evaluate the uncertainty of hazard to network performance. For instance, Chang et al. (2012) generated a retrofit plan for each scenario, which differs for each scenario, and evaluates the expected losses from each scenario to determine the final optimal bridge mitigation planning. Liu (2009) and Gomez and Baker (2019) introduced Two-stage Stochastic Programming to solve the optimal infrastructure risk investment through a finite number of scenarios. Also, in the practices, each examination would deal with distinct simulations. For instance, Chang et al., (2012) summarized three commonly used performance indicators in one mitigation planning evaluation, i.e., maximum flow, connectivity matrix, and travel delay cost. Based on the explanation, the network size increases as an increase in the number of scenarios and the number of simulation types.

Second, the topology in the most transportation network has highly connectivity redundancy that is overshadowed the bridge's criticality in the network. It is essential to understand that bridges are not the only asset in the transportation network. The infrastructure's asset in transportation network consists of road, other mode's infrastructure, and bridges. Even though bridges are the critical asset in the transportation network, the number of a bridge in the network is not as significant as the number of roads. Imagine, there are 603 bridges over 64,398 connected links in drive transportation network of Memphis, TN. Also, there are 787 bridges over 5,947 connected links in drive transportation network of Chicago, IL. Through this number, there may exist network redundancy in the transportation network depends on its topology. If there exists an alternative asset that provides robust connectivity and network's performance, then the bridge will lose its importance rank in the network.

Based on this explanation, it is essential to evaluate the criticality of bridges in a given transportation network and reduce the number of links in the network to evaluate the criticality of bridges. Clustering techniques have been proposed (Lim, Song, Kurtz 2015) to address such problems of scale. However, there are few works of literature in clustering techniques focusing on the criticality of the bridge in a highly redundant network. The project proposes a clustering framework to reduce the complexity of the network and provide information on the criticality of bridges using the Girvan-Newman algorithm. The algorithm would also be tested on a hypothetical low redundant network and hypothetical high redundant network.

2. NETWORK REDUNDANCY IN BRIDGE TRANSPORTATION NETWORK

Network redundancy (mesh) is a network that provides installed additional or alternative instances within the network infrastructure for ensuring network connectivity in case of path failure and unavailability. This installment is commonly found and implemented in a telecommunication network and an internet network. The transportation network might be classified as a redundant network depends on the performed assessment and definition in the network problem. As delivered previously, bridges are not the only asset in the transportation network. However, civil engineers need to evaluate the criticality of bridges and their contribution to the entire network's performance. For bridge asset management, it is vital to set the bridge as a critical component in the network. Hence, the knowledge to identify network redundancy is essential to be pre-determined in the initial phase of bridge risk management problems.

Even though this knowledge subjects to engineering judgment and experiences, the bridge transportation network could be simplified into two levels of redundancy, i.e., low redundant network, and highly redundant network. The low redundant bridge network is a network with a topology where the critically of bridges dominate the entire performance of the network directly. This type of network commonly represents the intercity transportation network and river area/city, e.g., New York, New Jersey, and California). Figure 1 shows the intercity transportation network where each city was zoomed out and represented with a big node. Bridges are the primary connectivity for each city. The high redundant bridge network is a network with a specific topology where there exists an alternative path in the network for not choosing bridges. This type of network represents most of the intracity transportation network's assessment. Figure 2 shows the drive transportation network in Memphis, TN, with 603 bridges over 64,398 connected links.

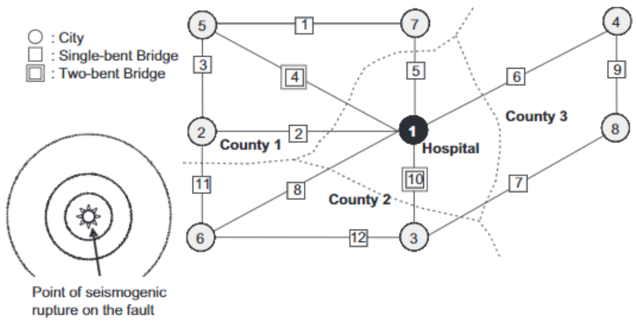


Figure 1 Hypothetical Intercity Network (Kang, Song, & Gardoni, 2008)

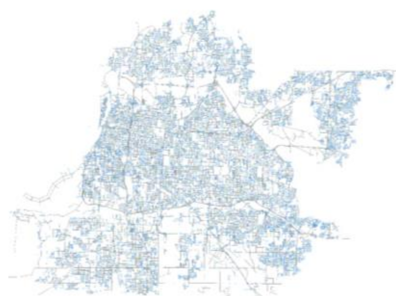


Figure 2 Drive Transportation Network in Memphis, TN (Boeing, 2017)

3. GIRVAN-NEWMAN CLUSTERING OF NETWORK

The Girvan-Newman algorithm (Girvan & Newman, 2002) is a hierarchical method used to detect the communities in complex systems. This method uses the centrality property to create clusters through edge betweenness. Edge betweenness centrality is the number of the shortest paths that go through an edge in a graph or network. Each edge in the network can be associated with an edge betweenness centrality value. An edge with a high edge betweenness centrality score represents a bridge-like connector between two parts of a network and the removal of which may affect the communication between many pairs of nodes through the shortest paths between them.

$$b(i,j) = \frac{\sum_{k,l \in V \setminus \{i,j\}} g^{e}_{k,l}}{\sum_{k,l \in V \setminus \{i,j\}} g_{k,l}}$$

where $b(i,j)$ is the edge betweenness value of arc (i,j); $g^{e}_{k,l}$ is shortest path from (k,l) and passed (i,j) in set of V; $g_{k,l}$ is all shortest path from (k,l) in set of V. Through this definition, the Girvan-Newman clustering algorithm is,

GIRVAN-NEWMAN ALGORITHM
$while\ G(N, A) is\ connected\ do$:
$\quad calculate\ edge\ betweenness\ of\ every\ edge$
$\quad remove(edge\ betweenness)\ in\ G(N, A)$

4. Proposed Network Size Reduction Algorithm

The project proposes a clustering framework to reduce the complexity of the network and provide information on the criticality of bridges using the Girvan-Newman algorithm. The transport network to be analyzed will be defined in the commonly adopted approach in network literature and graph theory, as G(N, A). That is, a graph defined by a set of nodes, N, and a set of arcs linking the nodes, A. In this project, the nodes mainly present all geographical position of road intersections. Then, the arcs represent the connectivity of each node. There are two types of links pre-determined in the network, i.e., bridge links and non-bridge links (e.g., roads). These arcs have weights to represent length by the distance between the connected nodes, or capacity/mean speed data as an embedded property. There are four main steps to execute the network reduction size algorithm.

NETWORK SIZE REDUCTION ALGORITHM
Define
$A = set\ of\ arc\ in\ the\ graph\ G(N, A)$
$B = set\ of\ arcs\ that\ are\ bridges$
$w_b = arc\ weight\ if\ arc\ (i, j) \in B$
$w_r = arc\ weight\ if\ arc\ (i, j) \in A \backslash B$
$c = set\ of\ component\ in\ G(N, A)$
$del = set\ of\ arc\ that\ is\ deleted$
$P = set\ of\ cluster\ pairs\ in\ G(N, A)$
$P_{ij} = G.subgraph(set(c_i) \& set(c_j))$
$A_{ij}^P = set\ of\ arcs\ in\ P_{ij}$
$Np_i = set\ of\ nodes\ in\ the\ perimeter\ of\ c_i$
$w_{sl} = arc\ weight\ for\ Super\ Link, average\ connectivity(n_{pi})\ in\ c_i$
Step 1: Link Weight
$set\ w_b \ll w_r$
Step 2: Modified Clutering Girvan − Newman
$int\ del = \{\}$
$while\ (u, v) \in B\ in\ any(G.subgraph(c))$:
$\quad do\ Girvan - Newman$
$\quad update\ c$
Step 3: Bridge Dominance

$for\ (c_i, c_j)\ in\ P$:
    $if\ (u,v) \in [del] \cap A_{ij}^{P} \cap B < 0$:
        $pick\ any\ one\ arc\ to\ connect\ c_i, c_j$
        $append(Np_i)$
    $elseif$:
        $if\ (u,v) \in [del] \cap A_{ij}^{P} \cap B > 1$:
            $do\ nothing$
            $append(Np_i)$
        $else$:
            $pick\ bridge\ arc\ to\ connect\ c_i,\ c_j$
            $append(Np_i)$

Step 4: Big Node, Perimeter Node, and Super Link

$if\ (u,v) \in B\ in\ c_i$:
    $do\ nothing$
$else$:
    G.add_node(Big Node)
    G.add_edge(Super Link)

4.1 Link Weight

The first step is defining the link weight to differentiate and emerge the criticality of bridge links in the network. As this algorithm implements edge betweenness value that uses minimization properties, the link weight in bridge links shall be smaller compared to weights in the road links. By applying smaller weights, bridge links are most likely to be the critical edges in edges betweenness centrality matrix. In the practices, we could illustrate these weights as the loss or drawback of using a particular link, e.g., travel time delay and risk/resilient matrix. In other interpretations, it shows that bridge links have lower drawbacks (higher benefit) in the functionality of the network than road links. This assumption forces the algorithm to identify the bridge as the most critical asset in the network.

4.2 Clustering Girvan-Newman

The second step is implementing a modified Girvan-Newman Algorithm. According to edges betweenness and the Girvan-Newman algorithm, the most critical links span between adjacent clusters. In this step, the algorithm would evaluate and provide information on the bridge's criticality in the network. The algorithm continuously generates clusters until there are no bridge links left inside the cluster. If the final clusters terminate with no bridge links in any clusters/subgraphs of G(N, A), then this condition indicates an ideal condition in which bridge links are dominant and crucial in the network. The maximum number of generated clusters is $(m_b + 1)$, where m_b is the total number of bridge links in the network. However, in general, the bridges transportation network subjects to redundant connectivity by another type of link. In most of the transportation case, bridge links might be inherently not the critical infrastructure due to topology-wise and connectivity redundancy. It also possible that link weight does not regularly provoke the criticality of the bridge.

4.3 Bridge Dominance in Adjacent Cluster

In the third step, the algorithm examines the removed links in each pair of adjacent clusters from the Girvan-

Newman algorithm to determine bridge dominance. The dominance is evaluated in every pair of adjacent clusters. In the first case, no bridge links are connecting a particular pair of clusters. Then, any connection between these clusters is not significant for bridge asset management applications. Thus, it is sufficient to pick one arc and delete the rest to indicate the existed connectivity between these adjacent communities. In the second case, there is one bridge link connecting a particular pair of clusters. Then, the entire connection within these two clusters would be represented solely by bridge links. Thus, it is necessary to pick the bridge link and delete the rest links to indicate the existed connection between these adjacent communities. In the third case, more than one bridge links are connecting a pair of clusters. In this case, there is no information to measure and compare the dominance level in every bridge link. In the bridge asset management approach, it is crucial to evaluate every bridge links in the network. It is necessary to track each bridge's links in the network and delete other types of links in between. Through this selection, the entire connection within these two clusters would be represented by multi bridge links.

4.4 Big Node, Perimeter Node, and Super Link

In the last step, this algorithm reduces each cluster size by generating three topology properties, i.e., big node, perimeter node, and super link. The big node illustrates the unity of all nodes in the cluster set. In this algorithm, we will zoom out the cluster as one node to reduce the complexity in the network. The perimeter node is a node in the perimeter of the cluster connected to adjacent outer links from the third step. The super link is a link connecting a big node and perimeter node. Each link is assigned with weight links as the average diameter of the perimeter node in the cluster. The diameter of a node i considers an average of all the shortest path lengths to adjacent nodes j, and therefore gives a first-order measure of how well connected to its surroundings a node is.

$$w_{i,sl} = \delta_i = \frac{1}{n-1}\sum_{(i,j)}^{n} d_{ij}$$

where, n is number of node and d_{ij} is the shortest path from node i to node j.

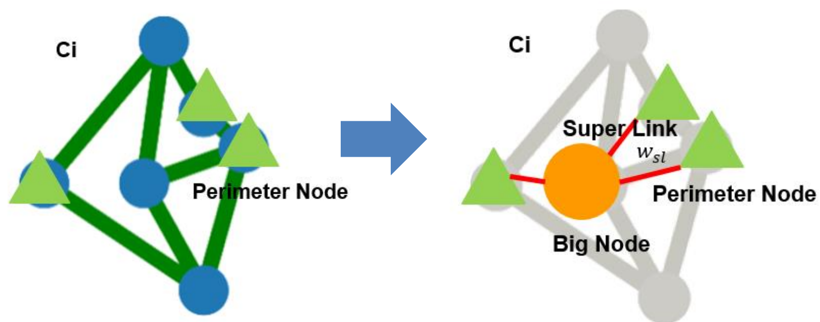
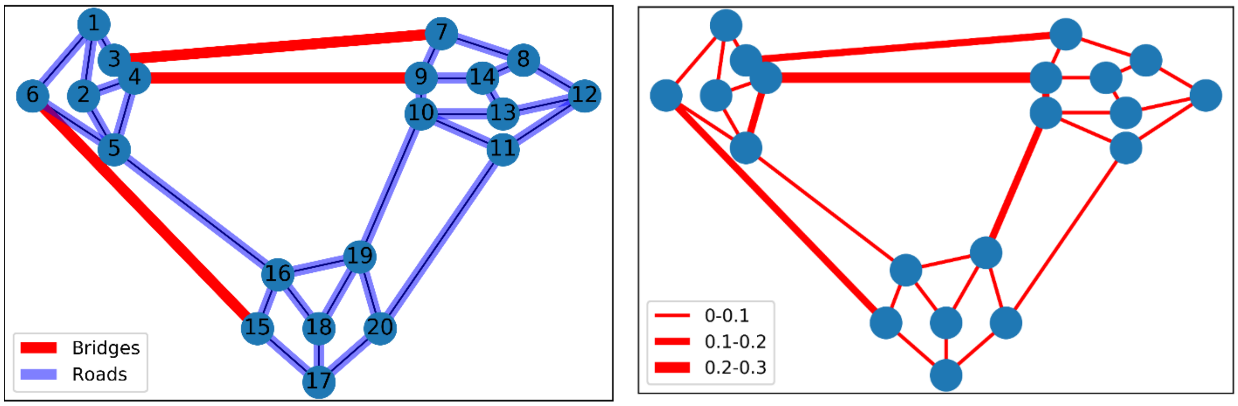


Figure 3 Illustration of Big Node, Perimeter Node, and Super Links

5. STUDY CASE 1: HYPOTHETICAL LOW REDUNDANT NETWORK

5.1 Network Parameter

The transport network to be analyzed will be defined in the commonly adopted approach in network literature and graph theory, as undirected G(N, A). The network consists of 20 nodes and 33 links with 3 bridges links. The link weight values are 1 for road links and 10 for bridge links. Then, the Network Reduction Size algorithm is performed to reduce and provoke the criticality of bridges in the given network.



(a) (b)

Figure 4 Low Redundant Network - Link Type Identification (a) and Initial Edge Betweenness Centrality (b)

5.2 Analysis of Size Reduction

Figure 4(b) shows the initial edge betweenness centrality value for the given network. The bridge links {(3,7), (4,9), and (6,15)} have high centrality and dominate the criticality of the network. This figure provides the initial important rank in the network in terms of centrality. From this rank and value, the bridge links are critical in the network. This result indicates that the network has low redundant connectivity in topology-wise.

Figure 5 presents the process in the proposed algorithm. In the clustering phase (Step 2), the Girvan-Newman algorithm gives a consistent indication of bridge dominance in the network as in the initial edge betweenness centrality matrix. This step is terminated with all bridge links selected as the inter-cluster link. In other interpretations, the bridge links are chosen as critical assets in the network.

In the final reduced network, this algorithm reduces node from 22 nodes into 11 nodes, and edges from 33 edged into 12 edges with three bridges links. Node {1,2,3} is merged out into one big node A; node {7,8,11,12,13,14} is merged out into one big node B; and node {16,18,17,19} is merged out into one big mode C. The final network results no connectivity redundancy by other infrastructure left in the network. This algorithm could reduce the complexity and maintain bridge importance in the network.

6. Study Case 2: Hypothetical High Redundant Network

6.1 Network Parameter

The transport network to be analyzed will be defined in the commonly adopted approach in network literature and graph theory, as undirected G(N, A). This study uses the hypothetical Sioux Falls network from Liu et.al. (2009). This network consists of 24 nodes and 38 links with 6 bridges links. The link weight values are 1 for road links and 10 for bridge links. Then, the Network Reduction Size algorithm is performed to reduce and provoke the criticality of bridges in the given network.

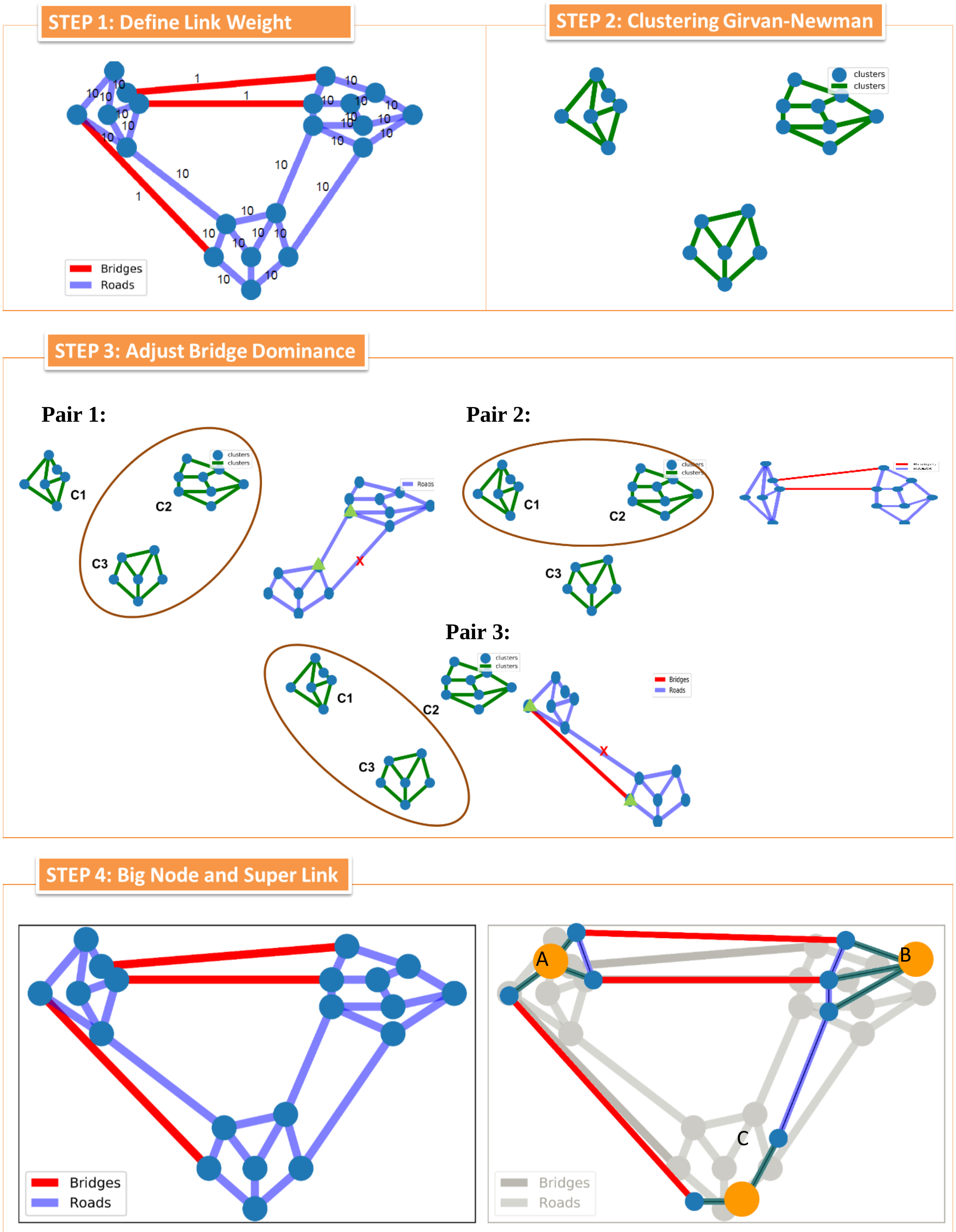


Figure 5 Low Redundant Network – Algorithm Procedure and Result

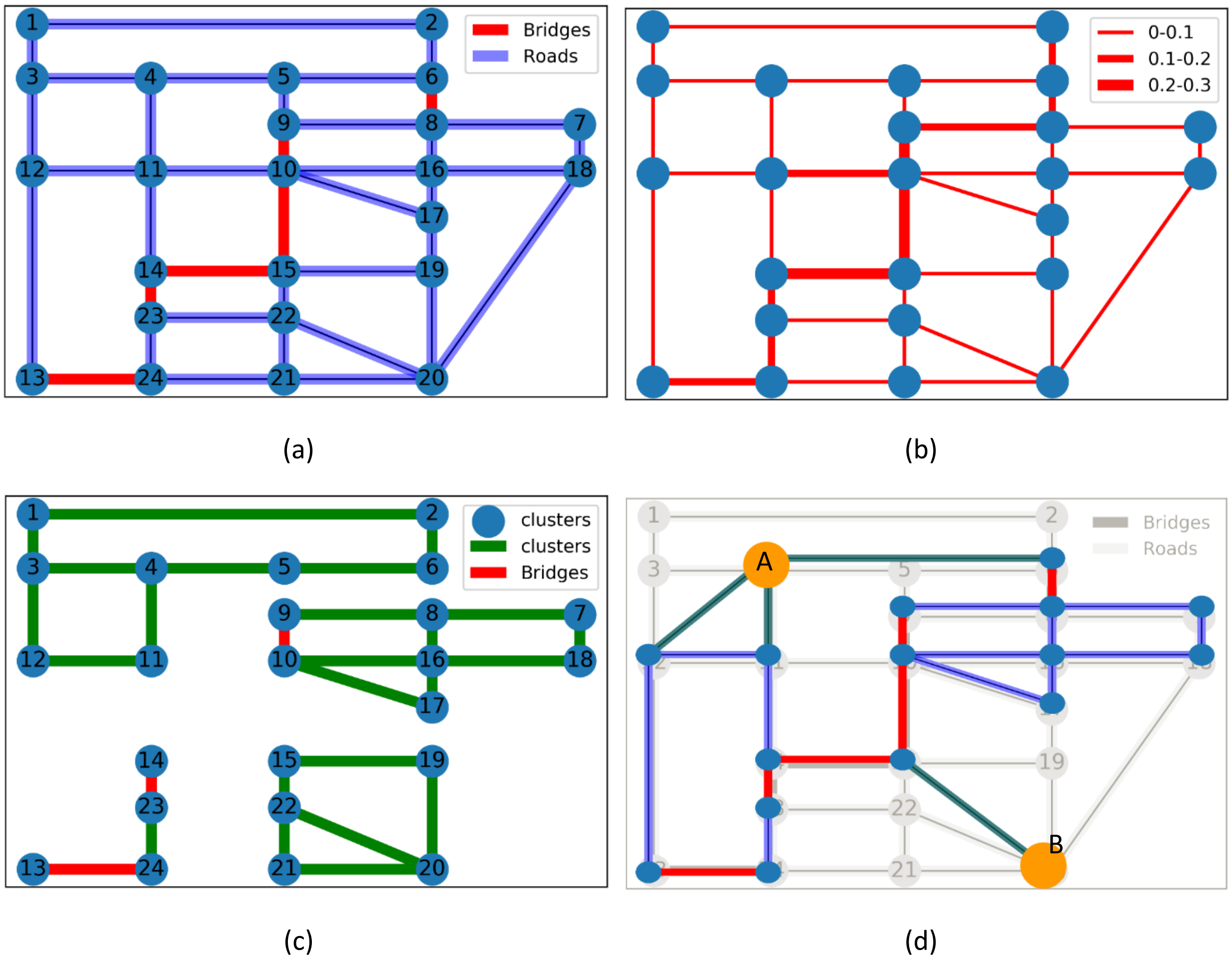


Figure 6 High Redundant Network - Link Type Identification (a); Initial Edge Betweenness Centrality (b); Clusters (c); and Final Reduced Network (d)

6.2 Analysis of Size Reduction and Bridge Criticality

Figure 6(b) shows the initial edge betweenness centrality value for the given network. The bridge links {(10,15), and (14,15)} have high centrality and dominate the criticality of the network. Then, it is followed by several road links and bridge links {(6,8), (9,10), (14,23), (13,24)}. As the initial important centrality rank in the network, the result indicates that not all bridge links dominates the criticality in the network. In this step, we gained information that bridge links {(6,8), (14,23), (13,24)} remain subjected to connectivity redundancy. Even with the assigned link weight, the inherent connectivity redundancy remains to appear in network. It indicates that the link weight is not effective in highlighting the criticality of bridge in the network. The current weight is important parameter to assure the criticality of the network. Therefore, the calibration of link weight is needed in the preliminary steps. It is necessary to have a link weight that account the inherent redundancy of the given network uniquely. Having this type of link might provide better understanding of the criticality of the bridges.

Figure 6(c) shows the final generated clusters of the network. The Girvan-Newman algorithm gives a consistent indication of bridge dominance in the network as in the initial edge betweenness centrality matrix. Initially, we gained information that bridge links {(6,8), (9,10), (14,23), (13,24)} remain subjected to connectivity redundancy. The clustering termination provides us another information that the least critical bridge links are {(9,10), (14,23), (13,24)}. First, tt is important to understand that the clustering is sensitive to the assigned link weight. Using different link weight might provide different clusters. Second, for this given link weight, it is interesting that bridge links {(9,10) (14,23), (13,24)} are remained in the cluster not because the criticality issue, but the centrality issues. Bridge link {(9,10)} in cluster C2 fail to be released from the clustering because the link's position is not in the center of the cluster. The centrality failed to detect the criticality from the assigned weight in every perimeter links. We could detect the same characteristic in Bridge link {(14,23), (13,24)} in cluster C3. The algorithm failed to detect the criticality of the perimeter links because the topology-wise. For every given weight in outer/perimeter links to provoke criticality, the centrality properties alone could not differentiate those links, not because link weight issues, but the topology or centrality issue.

Despite the issue in link weight and centrality, this algorithm still capable to reduce the network size from 24 nodes into 18 nodes, and from 38 edged into 20 edges with 6 bridges links. Node {1,2,3,4,5} is merged into one big node A; and node {19,20,21,22} is merged into one big node B. The final network reveals that there still exists connectivity redundancy by other infrastructure in the network, however, improved. This algorithm still performs well to provide a good the complexity reduction and maintain bridge importance in the network.

7. Conclusion

Based on the evaluation in the project, there are several conclusions, as following:

1. The proposed network reduction size allows the reduction of connectivity redundancy and adequately highlight the bridge criticality in the given transportation network.
2. The link weight is important to differentiate and emerge the criticality of bridge links in the network. The value of link weight needs to be calibrated in the preliminary steps of the algorithm to account the inherent redundancy in the given network. It is necessary to have a link weight that account the inherent redundancy of the given network uniquely. Having this type of link might provide better understanding of the criticality of the bridges.
3. There is a limitation in the centrality approach. This approach only allows criticality detection of any given central links in the given network. The centrality properties alone could not differentiate those links, not because link weight issues, but the topology or centrality issue. The accurate link weight might reduce the probability of having bridge links in perimeter/outline of the cluster. However, in a worst-case scenario, link weight itself could not guarantee a set of clusters where there is no bridge links in the perimeter.

8. Future Developments

In future development, this project could be developed by improving several issues and assumptions for the current model, as following:

- The value of link weight needs to be calibrated in the preliminary steps of the algorithm to account the inherent redundancy in the given network. It is necessary to have a link weight that account the inherent redundancy of the given network uniquely.

- This approach only allows criticality detection of any given central links in the given network. It is possible that we will have high redundant network with bridge links positioned in the perimeter of the network.
- It is important to study and verify the reduction size significant of this algorithm with any given network with varios cluster coefficient.In the flood modelling, the project only consider the depth of inundation in the network. This project neglect direction, dynamic movement, and high-speed of the water in simulation.

REFERENCES

Chang, S. E., and Nojima, N. (1998). "Measuring lifeline system performance: Highway transportation system in recent earthquakes." The 6th National Earthquake Conf. on Earthquake Engineering, Paper No. 70, Earthquake Engineering Research Institute, Oakland, CA.

Kang, W. H., Song, J., & Gardoni, P. (2008). Matrix-based system reliability method and applications to bridge networks. *Reliability Engineering and System Safety*, *93*(11), 1584-1593.

Lim H-W, Song J, Kurtz N. Seismic reliability assessment of lifeline networks using clustering-based multi-scale approach. Earthquake Eng Struct Dyn 2015;44(3):355–69.

Liu, C., Fan, Y., and Ordonez, F. (2009). "A two-stage stochastic programming model for transportation network protection." Comput. Oper. Res., 36(5), 1582–1590.

Nojima, N. (1998). "Prioritization in upgrading seismic performance of road network based on system reliability analysis." Proc., 3rd China- Japan-US Trilateral Symp. on Lifeline Earthquake Engineering, State Seismological Bureau, Beijing.