From Shortest-path to All-path: The Routing Continuum Theory and its applications (Supplementary Materials)

Yanhua Li, Member, IEEE, Zhi-Li Zhang, Fellow, IEEE, and Daniel Boley, Member, IEEE

Abstract—This document provides numerical illustration results obtained by applying the routing continuum theory in two synthetic networks and a real network topology. Moreover, besides the generalizations of the mixed $L_1$- and $L_2$-norm network flow optimization problem discussed in the main file, in this document, we discuss one more application in analyzing network robustness, by introducing the generalized shortest path and random walk betweenness centrality measures.

Index Terms—Routing continuum, network flow, betweenness centrality.

1 NUMERICAL ILLUSTRATION OF ROUTING CONTINUUM THEORY

We use two synthetic networks and a real network to show how the routing continuum grows as the parameter $\theta$ changes. Fig. 1 shows an example topology, with three disjoint paths between the source 1 and the destination 5, and each link is with 1-unit weight. As the parameter $\theta \geq 0$ increases, the longer paths $P_3 = \{1 \rightarrow 3 \rightarrow 4 \rightarrow 5\}$ and $P_2 = \{1 \rightarrow 2 \rightarrow 5\}$ get truncated gradually, and the shortest path $P_1 = \{1 \rightarrow 5\}$ is obtained when $\theta$ increases to 1. Fig. 2 shows the routing continuum, i.e. the optimal flow distributions at each $\theta$. We see that within the interval $\theta \in [0, 0.4]$, the flows on the longer paths $P_2$ and $P_3$ get linearly redistributed to the shortest path $P_1$, and the longest path $P_3$ gets truncated when $\theta = 0.4$. Then the flow of the second longest path $P_2$ keeps decreasing as $\theta$ increases, until the second boundary condition $\theta = 1$ holds, where $P_2$ is truncated. During the routing evolution process, the network flows are always redistributed from longer paths to the shorter path, while increasing $\theta$. When $\theta > 1$, namely, the largest boundary condition, the routing solution is stabilized to the shortest path, i.e. $P_1$.

Fig. 3 shows another example with five connected nodes in the topology. Weights $w_{ij}$’s are marked on the links. The flow initiates at source 1 and is removed from destination 5. Fig. 4–Fig. 8 show the optimal flow distributions (marked on individual links) under five boundary conditions, $[\theta_0 = 0, \theta_1 = 0.0914, \theta_2 = 0.2850, \theta_3 = 0.5700, \theta_4 = 2]$. In Fig. 4($\theta_0 = 0$), every link is active and follows the potential based “all-path” routing. Then, as $\theta$ increases to $\theta_1 = 0.0914$ (in Fig.5), link $(1, 4)$ is truncated, and within the interval $\theta \in [0, \theta_1]$, only the flow on path $\{1 \rightarrow 4 \rightarrow 5\}$ decreases, and gets redistributed to other paths, because this path with total length 11 is the longest path in $P(0)$, i.e. the “all-path” routing graph. Then, when $\theta$ increases to $\theta_2 = 0.2850$, the flows on links $(2, 3)$ and $(3, 5)$ are truncated, because these two links are on the second longest path $\{1 \rightarrow 2 \rightarrow 3 \rightarrow 5\}$, with path length 5. Similarly, when $\theta$ keeps increasing to $\theta_3$ and $\theta_4$, the rest two longer paths $\{1 \rightarrow 2 \rightarrow 5\}$ and $\{1 \rightarrow 2 \rightarrow 4 \rightarrow 5\}$ get removed, respectively, and only the shortest path $\{1 \rightarrow 5\}$ is left at last.

Now, we apply the routing continuum theory to Internet2 Abilene Network [1]. The Abilene network was a high-performance backbone network established by the Internet2 community in the late 1990s. The Abilene Network was retired and became the “Internet2 Network” in 2007. Fig. 15 shows its 11 regional network aggregation points and backbone connections across them (primarily OC192 or OC48 backbone). We consider the transmission cost between two end points roughly proportional to their actual geographic distance, because the velocity of light in an optical fiber becomes 60-70% compared to it in vacuum [13], [26], [33]. Hence, in the numerical analysis, we simply use the geographical distance as the link weight for the transmission cost as marked in Fig. 9. We choose the flow demand from Sunnyvale to New York. As we increase $\theta$ from 0, we observe a sequence of five boundary $\theta$’s, i.e. $[\theta_0 = 0, \theta_1 = 0.1082, \theta_2 = 0.2498, \theta_3 = 0.4943, \theta_4 = 3.2108]$, in which order links $(4 \rightarrow 6)$, $(5 \rightarrow 1)$, $(10 \rightarrow 9 \rightarrow 3)$ and $(10 \rightarrow 7 \rightarrow 4 \rightarrow 1 \rightarrow 11 \rightarrow 8)$ get truncated in sequence, and the optimal flow distribution evolves from the “all-path” routing to “shortest-path” routing. When $\theta_0 = 0$, all paths are present in delivering the contents, whereas only the shortest path $\{10 \rightarrow 3 \rightarrow 6 \rightarrow 5 \rightarrow 2 \rightarrow 8\}$ is active for $\theta \geq 3.2108$.

2 NETWORK ROBUSTNESS ANALYSIS VIA GENERALIZED CENTRALITY MEASURE

The optimal flow distribution $X^*(\theta)$ to the mixed $L_1$– and $L_2$–norm network flow optimization problem indicates exactly the loads on each link (resp. node) for certain flow demands. When considering flow demands from all source

---

*This work was supported in part by the NSF grants CNS-0179092 and IIS-0916750, the DTRA grant HDTRA1-09-1-0050. An earlier version of this work appeared in the Proceedings of IEEE ICDCS [22], June 2011. Yanhua Li is with HUAWEI Noah’s Ark Lab, China. Zhi-Li Zhang and Daniel Boley are with the Department of Computer Science and Engineering, University of Minnesota, Twin Cities, Minneapolis, MN, 55455.
E-mail: {yanhua,zzhang,boley}@cs.umn.edu
destination pairs, the average network flow on each link (resp. node) infers the “importance” of the link (resp. node), namely, the influence of the link (resp. node) in case of failure or being attacked, which in turn reveals the robustness structure of networks, i.e., which area of the network is more vulnerable to attacks. The robustness centrality measure of links and nodes in the network has been extensively studied, and has been applied to design topology control algorithm and routing protocol in wireless sensor networks and delay tolerant networks [14], [19]. Below, we show how our routing continuum theory can be used to generalize various robustness centrality measures of links/nodes in networks, where the ranking of links/nodes in terms of their betweenness infer the network robustness structure, i.e., those areas with high betweenness links/nodes expose more risks to attacks or failures, as when removing these links/nodes, more flows have to be rerouted or failed.

2.1 Centrality measures for mixed network flow
Centrality measures were first developed in social network analysis [9], [25], for example, how influential a user is in a social network, with applications in robust community detection [18], [23], mobility prediction [5], and etc. There are four widely used centrality measures [25], that capture the relative importance of a vertex or an edge within a network from various aspects: degree 1, eigenvector centrality 2, betweenness [10], [16], and closeness [28]. Betweenness and closeness centrality measures are directly interpretable in

1. The node degree centrality is simply defined as the number of links associated with a node, which reflects locally (i.e., within one hop,) how well the node is connected to other nodes.

2. Eigenvector centrality takes the leading Eigenvector, i.e., the Eigenvector corresponding to the largest Eigenvalue, of the adjacent matrix A as relative scores to all nodes in the network, which follows the concept that connections to nodes with higher scores contribute more to the score of the node than connections to nodes with lower scores. PageRank [27] and Katz centrality [20] can be viewed as two variations of the Eigenvector centrality measure.
terms of the shortest path and all-path routing, thus can be
generalized using our routing continuum theory to account
for mixed network flows. In the following, we will introduce
the mixed-flow betweenness for nodes or edges, a natural
generalization of the existing betweenness centrality measures.
The mixed-flow betweenness measures indicate the importance
of nodes or edges in terms of the degree to which a node
or an edge is participating in the communication between
node pairs in the network, which has implications in network
resource relocations and detecting robust subgraphs that are
resilient to attacks and failures. Note that closeness centrality
can similarly be generalized, and we omit these results here
for brevity.

2.2 Node Betweenness centrality

Node betweenness has been studied in the past as a measure of
the centrality and influence of nodes in networks.

Shortest-path betweenness. A simple example of such a
betweenness measure initially proposed by Freeman [8],
[16], [17] is shortest-path betweenness. Given a node $i$, its
shortest-path betweenness is defined as the number of shortest
(geodesic) paths between pairs of all other nodes that run
through $i$. To be precise, given a graph $G = (V, E)$, node
$i$'s betweenness centrality [16], [17] $C_i^S$ is defined as3

$$C_i^S = \frac{2 \sum_{s < t \in V} g_{ij}^{(st)}}{n(n - 1)},$$

where $g_{ij}^{(st)}$ is the number of shortest paths from node $s$ to node
t $t$ that pass through $i$. Since the graph is undirected, $g_{ij}^{(st)} = g_{ji}^{(ts)}$ always holds, thus computing $g_{ij}^{(st)}$ for only half of all
node pairs (i.e., for $s < t$) is sufficient. If there is more than
one shortest path between a node pair, each path is given equal
weight such that the total weight of all of the paths is unity.
Since when $\theta$ is large enough, the optimal flow distribution
denoted by $X^*(\infty)$ represents the shortest path solution, the

shortest path betweenness centrality $C_i^S$ can be written as

$$C_i^S = \frac{2 \sum_{s < t \in V} \sum_{k \in V} X_{ij}^{(st)}^{*} k_i(\infty)}{n(n - 1)}. \quad (2)$$

Current-flow betweenness4. Considering that the circuit cre-
ated by placing a resister on each edge of the network and unit
current source and destination at a particular node pair. The
resulting current flow in the network will follow Kirchhoff’s
and Ohm’s laws, going from source to destination along a
for a node $i$ is defined as the absolute value of the currents
summed over all node pairs that run through $i$. The optimal
optimal flow distribution $X^{(st)}^*(\infty)$ of the $L_2$ norm flow
optimization problem represents exactly the current flow for
source destination pair $(s, t)$ with $\theta = 0$. The current-flow
betweenness $C_i^C$ of node $i$ can be written in terms of $X^{(st)}^*(\infty)$ as

$$C_i^C = \frac{2 \sum_{s < t \in V} \sum_{k \in V} X_{ij}^{(st)}^{*} k_i(\infty)}{n(n - 1)}. \quad (3)$$

Mixed-flow betweenness. Shortest-path betweenness and
current-flow betweenness present two extremes. One uses
only shortest paths, and the other favors all-path to deliver
network flow. Our routing continuum theory naturally leads
to a generalized mix-flow betweenness, $C_i(\theta)$, which captures
how much mixed flow $X^*(\theta)$ runs through a node given a
flow combination parameter $\theta$.

$$C_i(\theta) = \frac{2 \sum_{s < t \in V} \sum_{k \in V} X_{ij}^{(st)}^{*} k_i(\theta)}{n(n - 1)}, \quad (4)$$

with $\theta \geq 0$. Note that the shortest-path betweenness (eq.(2))
and the current-flow betweenness (eq.(3)) are two special
cases of mixed-flow betweenness, as $C_i^S = C_i(\infty)$ and
$C_i^C = C_i(0)$, respectively. Given a specific $\theta \geq 0$, $C_i(\theta)$
captures the importance of node $i$, in terms of the average
optimal flow going through node $i$ over all source destination
pairs.

2.3 Edge betweenness centrality

Analogically, the betweenness centrality can be defined for
edges, capturing how much network flow going through a
particular edge, summed over all node pairs in the network.
The shortest-path betweenness of an edge $(i, j)$ is the total
number of shortest paths running along $(i, j)$, which was
first introduced by Anthonisse in [6], and Newman formally
defined it in [23]. It can be written in terms of the optimal
shortest path flow distribution denoted by $X^*(\infty)$ as

$$C_{ij}^S = \frac{2 \sum_{s < t \in V} X_{ij}^{(st)}^*(\infty)}{n(n - 1)}. \quad (5)$$

Similarly, the current-flow betweenness of an edge $(i, j)$ is
the current flow running along $(i, j)$ [10], [23], which can be

4. Current-flow betweenness is proven to be equivalent to random walk
(RW) betweenness [24]. For a node $i$, we calculate the expected number
of times that a random walk between a particular node pair will pass through
$i$, and RW betweenness is the summation over all node pairs.
computed using the following eq.(6) in terms of the optimal $L_2$ network flow distribution denoted by $X^*(0)$ as
\[
C_{ij}^C = \frac{2 \sum_{s \in V} X^{(s)}_{ij}^*(0)}{n(n-1)} \tag{6}
\]

The mixed-flow betweenness of an edge $(i, j)$ is then a natural generalization of eq.(5) and eq.(6) for $\theta \geq 0$.
\[
C_{ij}(\theta) = \frac{2 \sum_{s \in V} X^{(s)}_{ij}^*(\theta)}{n(n-1)} \tag{7}
\]

As discussed earlier, the trade-off parameter $\theta \geq 0$ governs how much shortest path flow vs current flow is considered in the mixed flow optimization problem. The mixed-flow betweenness centrality measure for a link/node captures how crucial the link/node is in carrying the network flow for all possible node pairs. In communication networks, the link/node betweenness measures in fact indicate how much (mixed) network flow has to go through a particular link/node for all source-destination pairs. An attack or failure to the links/nodes with high betweenness leads to more influential impacts to the network traffic. Hence the ranking of the links/nodes in terms of their betweenness infer the network robustness structure, namely, areas with high betweenness links/nodes are more vulnerable to attacks or failures, since more flows have to be rerouted or failed if these links/nodes fail.

### Table 1

<table>
<thead>
<tr>
<th>Edge ranking</th>
<th>when $\theta \in [0.0, 0.002]$</th>
<th>when $\theta \in [0.002, 0.06]$</th>
<th>when $\theta \geq 0.06$</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>(1,2)</td>
<td>(1,2)</td>
<td>(2,4)</td>
</tr>
<tr>
<td>#2</td>
<td>(4,5)</td>
<td>(2,4)</td>
<td>(1,2)</td>
</tr>
<tr>
<td>#3</td>
<td>(2,4)</td>
<td>(4,5)</td>
<td>(2,4)</td>
</tr>
<tr>
<td>#4</td>
<td>(2,3)</td>
<td>(2,3)</td>
<td>(2,3)</td>
</tr>
<tr>
<td>#5</td>
<td>(3,5)</td>
<td>(3,5)</td>
<td>(3,5)</td>
</tr>
<tr>
<td>#6</td>
<td>(1,3)</td>
<td>(1,3)</td>
<td>(1,3)</td>
</tr>
<tr>
<td>#7</td>
<td>(2,5)</td>
<td>(2,3)</td>
<td>(2,5)</td>
</tr>
<tr>
<td>#8</td>
<td>(1,4)</td>
<td>(1,4)</td>
<td>(2,5)</td>
</tr>
</tbody>
</table>

#### 2.4 Numerical examples

Next, we use the topology in Fig. 3 and a real network topology, i.e., Internet2 Abilene Network [1], as examples, to show how the ranking of node/link in terms of betweenness changes over $\theta$.

When computing the betweenness centrality measures for the five node topology in Fig. 3, we observe that as increasing $\theta \geq 0$, the ranking of links in terms of their mixed-flow betweenness keeps relatively robust, namely, there are only three different link ranking orders (See Tab 1). The highest betweenness links are (1, 2), (4, 5) and (2, 4), which all have the smallest link weights. The link (2, 4) steps up to the highest ranking, when $\theta \geq 0.06$. The node betweenness ranking is more stable, which is unchanged over $\theta$’s for topology in Fig. 3 with nodes ranked as {2, 5, 4, 1, 3} in a decreasing order. Nodes with more links and lower link weights are ranked higher, since they are more likely to serve as hubs to carry more network flows.

Now we investigate how the link and node betweenness ranking vary over $\theta$ in Abilene network. As we increase $\theta \geq 0$, there are twelve boundary $\theta$’s, governing the different ranking of link betweenness in Abilene network as shown in Table 2. We observe that when $\theta$ changes, namely, the network flow evolves from “all-path” flow to “shortest-path” flow, the ranks of links with the highest betweenness keep high ranking over $\theta$, i.e., links at rank #1 to #5 are unchanged. On the other hand, the betweenness centrality link (3, 10) increases from the rank #11 to #6 gradually (as highlighted in Table 2), which happens because the high link weight of (3, 10) suppress the “all-path (current)” flow going through it, but it resides on more shortest paths among node pairs, thus generates higher shortest-path flow when $\theta$ is large. Moreover, the ranks of links such as (1, 11), (2, 8), and (8, 11) decrease as $\theta$ increases. The ranks of some other links, including (1, 5) and (4, 7), keep stable at rank #7-#9. When looking at the node betweenness, the ranking is more stable than links, which is unchanged for all $\theta$’s as shown in Table 3. The nodes placed in central US, such as Kansas City and Indianapolis posses highest node betweenness centrality, namely, being the busiest nodes in carrying network flow.

We also computed the ranking of link/node betweenness in other real networks, such as Roofnet [3] (with 38 nodes), CERNET [4] (with 36 nodes), GEANT [2] (with 23 nodes), where similar results are obtained and we omit them here for brevity. From all these results, the node betweenness centrality ranking is overall more stable than link betweenness centrality ranking, through the entire routing continuum, i.e., all $\theta \geq 0$.

### 3 Discussion

In a broader context, the mixed $L_1/L_2$ optimization formulation has been widely used, e.g., in the classical LASSO problems [30], namely, the least square optimization problems with a $L_1$-norm penalty term, and more recently, in compressive sensing [11], [31]. It is therefore well-known that the $L_1$-norm penalty forces the least-square solution, $X^*$, to meet certain sparsity constraints, i.e., $||X^*||_1 \leq \epsilon$. Compared with LASSO and compressive sensing settings, our setting has a set of additional flow conservation constraints — these are what makes the problem unique and leads to solutions that have interesting interpretations and consequences, where the solutions to the more general LASSO and compressive sensing settings may not have, apart from the sparsity of the solutions.

Indeed, our routing continuum theory and the mixed $L_1$- and $L_2$- flow optimization can be interpreted in terms of the “sparsity” of the solutions also: the optimal flow solution $X^*(\theta)$ to the mixed $L_1/L_2$-norm flow optimization leads to a sparser routing graph, where the path length of routes used for routing the optimal flow from a source to a destination can not be $(1 + \theta^{-1})$ longer than the shortest paths. More surprising and interesting is that we can generate the entire routing continuum from the mixed $L_1/L_2$-norm flow optimization. The flow conservation constraints in fact play a key role here: it leads to the duality of the optimal flows, $X^*(\theta)$, a function defined on the edges of a network, and the optimal (generalized) potential functions, $U^*(\theta)$, a function defined on the nodes of a network. This allows us to solve $U^*(\theta)$ through
TABLE 2

<table>
<thead>
<tr>
<th>Rank</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kansas City</td>
<td>Indianapolis</td>
<td>Denver</td>
<td>Atlanta</td>
<td>Sunnyvale</td>
<td>Houston</td>
<td>Chicago</td>
<td>Los Angeles</td>
<td>New York</td>
<td>Washington</td>
<td>Seattle</td>
<td></td>
</tr>
</tbody>
</table>

Node betweenness ranking in the Abilene network for all $\theta$'s.

### References

13. S. Cheshire. Latency survey results (for “it’s the latency, stupid”), 1996.