On Traffic-Aware Betweenness and Network Criticality

Ali Tizghadam
School of Electrical and Computer Engineering
University of Toronto, Toronto, Canada
Email: ali.tizghadam@utoronto.ca

Alberto Leon-Garcia
School of Electrical and Computer Engineering
University of Toronto, Toronto, Canada
Email: alberto.leongarcia@utoronto.ca

***Abstract*—This paper introduces a new network science metric to evaluate the relative load on different links and nodes of a communication network. Motivated by the definition of random-walk betweenness in graph theory, we define the notion of traffic-aware betweenness (TAB) for data networks, where usually an explicit (or implicit) traffic matrix governs the distribution of external traffic into the network. We study the relationship between TAB and average network utilization and show that the average normalized traffic-aware betweenness, which is referred to as traffic-aware network criticality (TANC) is a linear function of end-to-end effective resistances (or network criticalities) of the graph. As a result, TANC is in general is a (convex+concave) function of link weights. We find a subset of admissible traffic matrices, in which the problem is convex. We construct a semidefinite program to solve this convex optimization problem and evaluate its optimal solution on some well-known graph topologies.**

I. Introduction

Freeman [1] introduced a very useful metric in graph theory referred to as shortest-path betweenness centrality. For node k the shortest-path betweenness centrality with respect to flows from source node i to destination node j is defined as the proportion of instances of the shortest paths from node i to j that traverse node k. The overall shortest-path betweenness centrality of node k is the sum of the centralities over all source-destination pairs. Link betweenness is defined likewise. Later, a series of other related metrics defined by various researchers in the field of network science, including flow-betweenness [2], [3] and random-walk betweenness [4] each one addressing some limitations of shortest path betweenness. However, previous definitions of betweenness (shortest-path betweenness, flow-betweenness, random-walk betweenness) are purely topological and are oblivious about the traffic between different source-destination pairs, while in communication networks the external traffic is a major factor in analyzing the behavior of networks. Here we introduce a new notion of betweenness, Traffic-Aware Betweenness (TAB), to account for traffic in a network. This new definition can apply to all different versions of betweenness, but our derivations are for traffic-aware random-walk betweenness (TARWB). We assume that the traffic between every node pair is given by a traffic matrix $\Gamma = [\gamma_s(d)]$.

The paper is organized as follows. Section II reviews previous work on random-walk betweenness and network criticality, and introduces necessary notations. In section III the relationship of packet networks and betweenness centrality is studied. The proposed metrics traffic-aware betweenness and traffic-aware network criticality (TANC) are introduced and investigated in section IV. The optimization problem to minimize traffic-aware network criticality is discussed in section V, followed by two case studies of network planning for well-defined graphs in section VI. Conclusions are presented in section VII.

II. Random-Walk Betweenness and Network Criticality

In this paper, we define a network as an undirected weighted graph $G(N, E, W)$, where N and E are the set of nodes and links respectively, and W is the symmetric link weight matrix.

Our definition of random-walk betweenness follows Newman [4]. Consider a finite-state irreducible Markov Chain with transition probabilities p_{ij} of transitioning from state i at time t to state j at time $t+1$ (discrete time). The Markov chain can be represented by a state transition diagram with states as nodes in a graph and edges corresponding to allowable transitions, and labels associated with the edges denoting the transition probabilities. The Markov chain can be viewed as a random walk on the n-node graph with next-step transition probabilities p_{ij}.

We are interested in quantifying the betweenness of a node in the random-walk corresponding to the above Markov chain. Consider the set of trajectories that begin at node s and terminate when the walk first arrives at node d, that is, destination node d is an absorbing node. We define the betweenness $b_{sk}(d)$ of node k for the $s-d$ trajectories as the average number of times node k is visited in trajectories from s to d. Note that $b_{dk}(d) = 0$ for k not equal to d since such walks are terminated at step zero.

Let $B_d = [b_{sk}(d)]$ be the $n \times n$ matrix of betweenness metrics of node k for walks that begin at node s and end at node d. Note that the d^{th} row of the matrix is zero. It is shown in [5] that matrix B_d can be written as:

$$B_d = (I - P_d)^{-1}\Theta_d \tag{1}$$

$$\Theta_d = [\theta_{sk}(d)] = \begin{cases} 1 & if \ s = k \neq d \\ 0 & otherwise \end{cases}$$

Matrix P_d is also the same as P except that its d^{th} row and d^{th} column are zero vectors. In this paper we are interested in a special type of random-walks referred to as weight-based random-walk. The weight-based random-walk is defined on a Markov chain with transition probability matrix P according to the following rule:

$$p_{ij} = \begin{cases} \frac{w_{ij}}{\sum_{k \in A(i)} w_{ik}} & if \quad j \in A(i) \\ 0 & otherwise \end{cases} \tag{2}$$

where $A(i)$ is the set of adjacent nodes of i and $w_{ik} \geq 0$ is the weight of link (i, k).

We now introduce *network criticality* [5], [6], to quantify the robustness of a network. We start by defining node/link criticality.

Definition 2.1: Node criticality is defined as the random-walk betweenness of a node over its weight (weight of a node is defined as the sum of the weights of its incident links). Link criticality is defined as the betweenness of a link over its weight.

Let η_k be the criticality of node k and η_{ij} be the criticality of link $l = (i, j)$. It is shown in [5], [6] that η_i and η_{ij} can be obtained by the following expressions:

$$\frac{b_{sk}(d)}{W_k} = l^+_{dd} - l^+_{sd} - l^+_{dk} + l^+_{sk} \tag{3}$$

$$\tau_{sd} = l^+_{ss} + l^+_{dd} - 2l^+_{sd} \quad or \quad \tau_{sd} = u^t_{sd} L^+ u_{sd} \tag{4}$$

$$\tau_{sd} = \frac{b_{sk}(d) + b_{dk}(s)}{W_k} \tag{5}$$

$$\eta_k = \frac{b_k}{W_k} = \frac{1}{2}\tau \quad , \quad \tau = \sum_s \sum_d \tau_{sd} \tag{6}$$

$$\eta_{ij} = \frac{b_{ij}}{w_{ij}} = \tau \tag{7}$$

where L^+ is the Moore-Penrose inverse of graph Laplacian matrix L [7], n is the number of nodes, and $u_{ij} = [0 \ldots 1 \ldots -1 \ldots 0]^t$ (1 and -1 are in i^{th} and j^{th} positions respectively). We define the average network criticality as $\bar{\tau} = \frac{1}{n(n-1)}\tau$.

Observation 2.2: Equations (3) to (7) show that node criticality (η_k) and link criticality (η_{ij}) are independent of the node/link position and only depend on τ (or $\bar{\tau}$) which is a global quantity of the network.

Definition 2.3: We refer to τ_{sd} as *end-to-end network criticality* and τ as *network criticality*.

End-to-end network criticality has a nice interpretation in electrical circuits. If we assume that our network is a resistive electrical network with link conductances equal to the weights of the corresponding links, then τ_{sd} is the *resistance distance* or *effective resistance* seen between two end points s and d [8], and τ is the total effective resistance of the network with many useful interpretations [9].

One can see that τ is a global quantity on the network graph. Equations (6) and (7) show that node (link) betweenness consists of a local parameter (weight) and a global metric (network criticality). τ can capture the effect of topology through the betweenness values. The higher the betweenness of a node/link, the higher the risk (criticality) in using the node/link. Furthermore, one can define node/link capacity as the weight of a node/link, then the higher the weight of a node/link, the lower the risk of using the node/link. Therefore network criticality can quantify the risk of using a node/link in a network which in turn indicates the degree of robustness of the network.

In this paper we extend the definition of betwenness, and network criticality by considering the effect of an explicit traffic matrix in the system. In our previous works ([5], [10]) we implicitly assumed that the average input traffic to all the nodes of the network are uniform. In this work we consider a general traffic matrix $[\gamma_s(d)]$ and will introduce *traffic-aware* versions of betweenness and network criticality.

III. Packet Networks and Random-Walk Betweenness

We now show that random-walk betweenness is closely related to packet network models. Consider a packet switching network in which packets arrive to packet switches from outside the network according to independent arrival processes. Each external packet arrival has a specific destination and the packet is forwarded along the network until it reaches said destination. We suppose that packet switches are interconnected by transmission lines that can be modeled as single-server queues. Furthermore, suppose that packet switches use a form of routing where the proportion of packets at queue i forwarded to the next-hop queue j is p_{ij}.

We calculate the total arrival/departure rate of the traffic to/from each node. The total input rate of node k (internal plus external) is denoted by x_k. After receiving service at the i^{th} node, the proportion of customers that proceed to node k is p_{ik}. To find x_k we need to solve the following set of linear equations (see [11]):

$$x_k = \gamma_k + \sum_{i=1}^{n} x_i p_{ik} \tag{8}$$

where γ_k is the external arrival rate to node k. Note that equation 8 is essentially similar to KCL (Kirchhoff's Current Law). If we denote $\overrightarrow{x} = [x_1, x_2, ..., x_n]$ and $\overrightarrow{\gamma} = [\gamma_1, \gamma_2, ..., \gamma_n]$, then equation (8) becomes:

$$\overrightarrow{x} = \overrightarrow{\gamma} + \overrightarrow{x}P \tag{9}$$

Suppose we focus on traffic destined to node d, then node d is an absorbing node, and we suppose that the arrival rate at node d is zero (since said arrivals do not affect other nodes) and equation (9) can be written as:

$$\overrightarrow{x_d} = (\overrightarrow{\gamma_d} + \overrightarrow{x_d}P_d)\Theta_d \tag{10}$$

where $\overrightarrow{x_d}$ and $\overrightarrow{\gamma_d}$ are the same as $\overrightarrow{x}$ and $\overrightarrow{\gamma}$ except for the d^{th} element which is 0. Matrix P_d is also the same as P except that its d^{th} row and d^{th} column are zero vectors. Equation (10) can be solved for $\overrightarrow{x_d}$.

$$\overrightarrow{x_d} = \overrightarrow{\gamma_d} \times \Theta_d \times (I - P_d \times \Theta_d)^{-1} \tag{11}$$

To find the relationship of betweenness B_d and the input arrival rate x_k we notice that $p_{dk}(d) = 0$ which means that $P_d = \Theta_d \times P_d$.

Thus:

$$\begin{aligned} P_d \times \Theta_d &= \Theta_d \times P_d \times \Theta_d \\ \Theta_d - P_d \times \Theta_d &= \Theta_d - \Theta_d \times P_d \times \Theta_d \\ \Theta_d \times (I - P_d \times \Theta_d)^{-1} &= (I - P_d)^{-1} \times \Theta_d \end{aligned}$$

Using equation (1) we will have:

$$\Theta_d \times (I - P_d \times \Theta_d)^{-1} = B_d \tag{12}$$

We substitute equation (12) in (11) to find the relationship between the node traffic and node betweenness.

$$\overrightarrow{x_d} = \overrightarrow{\gamma_d} \times B_d \tag{13}$$

If we denote the k^{th} element of $\overrightarrow{x_d}$ and $\overrightarrow{\gamma_d}$ by $x_k(d)$ and $\gamma_k(d)$ respectively, we have:

$$x_k(d) = \sum_s \gamma_s(d) b_{sk}(d) \tag{14}$$

Now we can find the total load at node k by adding the effect of all destinations in equation (14).

$$x_k = \sum_d x_k(d) = \sum_{s,d} \gamma_s(d) b_{sk}(d) \tag{15}$$

IV. Traffic-Aware Betweenness

In this section we develop an expression for traffic-aware betweenness motivated by equation (15).

Definition 4.1: Let $B_d = [b_{sk}(d)]$ be the betweenness matrix and let the traffic matrix be $\Gamma = [\gamma_s(d)]$. We denote the total external traffic by $\gamma = \sum_{s,d} \gamma_s(d)$. We define traffic-aware betweenness (TAB) of node k as:

$$\begin{aligned} b'_{sk}(d) &= \frac{\gamma_s(d)}{\gamma} b_{sk}(d) && (16) \\ b'_k &= \sum_{s,d} \frac{\gamma_s(d)}{\gamma} b_{sk}(d) && (17) \end{aligned}$$

Note that definition 4.1 is generally applicable for the different types of betweenness. If we consider $b_{sk}(d)$ as the shortest-path betweenness, then definition 4.1 gives traffic-aware shortest-path betweenness, and so on.

A. Traffic-Aware Random-Walk Betweenness (TARWB)

In this section we develop traffic-aware random-walk betweenness, where $b_{sk}(d)$ denotes the random-walk betweenness (weight-based random-walk). In order to derive an appropriate expression for TARWB we notice that:

$$\begin{aligned} b'_k &= \sum_{s,d} \frac{\gamma_s(d)}{\gamma} b_{sk}(d) \\ &= \frac{1}{2} \sum_{s,d} \left(\frac{\gamma_s(d)}{\gamma} b_{sk}(d) + \frac{\gamma_d(s)}{\gamma} b_{dk}(s)\right) && (18) \end{aligned}$$

But, from equation (5) we know that:

$$b_{sk}(d) + b_{dk}(s) = W_k \tau_{sd} \tag{19}$$

Substituting equation (19) in (18) will result in:

$$\begin{aligned} b'_k &= \frac{1}{2} \sum_{s,d} \left(\frac{\gamma_s(d)}{\gamma} b_{sk}(d) + \frac{\gamma_d(s)}{\gamma} (W_k \tau_{sd} - b_{sk}(d))\right) \\ &= \frac{W_k}{2} \sum_{s,d} \frac{\gamma_d(s)}{\gamma} \tau_{sd} + \frac{1}{2} \sum_{s,d} \left(\frac{\gamma_s(d)}{\gamma} - \frac{\gamma_d(s)}{\gamma}\right) b_{sk}(d) && (20) \end{aligned}$$

Now we write $b_{sk}(d)$ in terms of different elements of the matrix of end-to-end network criticalities $\Omega = [\tau_{sd}]$. Using equations (3) and (4) we have:

$$\begin{aligned} \tau_{sd} + \tau_{dk} - \tau_{sk} &= 2(l^+_{dd} - l^+_{sd} - l^+_{dk} + l^+_{sk}) \\ &= 2 \frac{b_{sk}(d)}{W_k} \end{aligned}$$

Thus

$$b_{sk}(d) = \frac{W_k}{2} (\tau_{sd} + \tau_{dk} - \tau_{sk}) \tag{21}$$

Substituting equation (21) in (20) will result in:

$$\begin{aligned} b'_k &= \frac{W_k}{4} \sum_{s,d} \left(\frac{\gamma_s(d)}{\gamma} + \frac{\gamma_d(s)}{\gamma}\right) \tau_{sd} \\ &+ \frac{W_k}{4} \sum_{s,d} \left(\frac{\gamma_s(d)}{\gamma} - \frac{\gamma_d(s)}{\gamma}\right)(\tau_{dk} - \tau_{sk}) && (22) \end{aligned}$$

Observation 4.2: According to equation (22), the normalized traffic-aware betweenness ($\frac{b'_k}{W_k}$) can be written as a linear function of τ_{sd}'s. Since τ_{sd} is a convex function of link weights ([9]), the normalized traffic-aware betweenness is also a convex function of link weights.

Observation 4.3: There are two special cases of interest in equation (22).

1) $\gamma_s(d) = \frac{\gamma}{n(n-1)} \quad \forall\ s - d\ pairs$
   When the average traffic between all source-destination pairs are equal, equation (22) is reduced to $b'_k = W_k \frac{\bar{\tau}}{2}$, where $\bar{\tau} = \frac{\tau}{n(n-1)}$ which is the original definition of random-walk betweenness just with a normalization coefficient $\frac{1}{n(n-1)}$.
2) $\gamma_s(d) = \gamma_d(s) \quad \forall\ s - d\ pairs$
   In the case of symmetric traffic demand matrix, equation 22 can be simplified as follows.

$$\frac{b'_k}{W_k} = \frac{1}{2} \sum_{s,d} \frac{\gamma_d(s)}{\gamma} \tau_{sd} \tag{23}$$

Equation (22) can be written in the following form:

$$\begin{aligned} \frac{b'_k}{W_k} &= \frac{1}{2}\left(\frac{1}{2} \sum_{s,d} \left(\frac{\gamma_s(d)}{\gamma} + \frac{\gamma_d(s)}{\gamma}\right) \tau_{sd}\right. \\ &+ \left.\frac{1}{2} \sum_{s,d} \left(\frac{\gamma_s(d)}{\gamma} - \frac{\gamma_d(s)}{\gamma}\right)(\tau_{dk} - \tau_{sk})\right) && (24) \end{aligned}$$

In analogy with the notion of $\frac{b_k}{W_k} = \frac{\tau}{2}$ and using equation (24) we can define Traffic-Aware Node Criticality (TANOC) τ'_k:

$$\begin{aligned} \tau'_k &= \frac{1}{2} \sum_{s,d} \left(\frac{\gamma_s(d)}{\gamma} + \frac{\gamma_d(s)}{\gamma}\right) \tau_{sd} \\ &+ \frac{1}{2} \sum_{s,d} \left(\frac{\gamma_s(d)}{\gamma} - \frac{\gamma_d(s)}{\gamma}\right)(\tau_{dk} - \tau_{sk}) && (25) \end{aligned}$$

We can also define point-to-point traffic-aware criticality (between points s and d) for a node k as follows:

$$\tau'_{sd}(k) = \frac{1}{2}(\frac{\gamma_s(d)}{\gamma} + \frac{\gamma_d(s)}{\gamma})\tau_{sd} + \frac{1}{2}(\frac{\gamma_s(d)}{\gamma} - \frac{\gamma_d(s)}{\gamma})(\tau_{dk} - \tau_{sk}) \quad (26)$$

Observation 4.4: Equation (25) shows that TANOC depends on the node position.

Observation 4.5: By averaging over k in equations (25) and (26) we obtain a measure of average traffic-aware criticality. Let $\tau'_{sd} = \frac{1}{n}\sum_k \tau'_{sd}(k)$, then

$$\tau'_{sd} = \frac{1}{2}(\frac{\gamma_s(d)}{\gamma} + \frac{\gamma_d(s)}{\gamma})\tau_{sd} + \frac{1}{2n}(\frac{\gamma_s(d)}{\gamma} - \frac{\gamma_d(s)}{\gamma})(\tau_{d*} - \tau_{s*}) \quad (27)$$

where $\tau_{i*} = \sum_k \tau_{ik}$. Similarly, Let $\tau' = \frac{1}{n}\sum_k \tau'_k$, then:

$$\begin{aligned} \tau' &= \frac{1}{2}\sum_{s,d}(\frac{\gamma_s(d)}{\gamma} + \frac{\gamma_d(s)}{\gamma})\tau_{sd} \\ &+ \frac{1}{2n}\sum_{s,d}(\frac{\gamma_s(d)}{\gamma} - \frac{\gamma_d(s)}{\gamma})(\tau_{d*} - \tau_{s*}) \quad (28) \end{aligned}$$

We refer to τ' as Traffic-Aware Network Criticality (TANC).

Observation 4.6: If we suppose that the traffic matrix Γ is symmetric, then TANOC will be:

$$\tau'_k = \sum_{s,d} \frac{\gamma_d(s)}{\gamma}\tau_{sd} = \tau' \quad (29)$$

In this case, according to equation (29), TANOC is independent of the node position.

V. Minimization of Network Utilization

Node utilization is defined as the load of a node normalized by its capacity (or in a more general sense by its weight), the utilization of node k is equal to $V_k = \frac{x_k}{W_k}$. We denote the average network utility by $\bar{V} = \frac{\sum_k V_k}{n}$. Considering equations (15), (17), and using $\frac{b'_k}{W_k} = \frac{1}{2}\tau'_k$, one can see that $V_k = \frac{\gamma}{2}\tau'_k$ and $V = \frac{\gamma}{2}\tau'$. Therefore, minimizing average network utilization is equal to minimizing traffic-aware network criticality. Since τ' is a linear combination of convex functions (τ_{ij}), we arrive at the result that minimizing average network utilization can be formulated as a (convex+concave) optimization problem since the coefficients could be positive or negative.

We now consider minimization of a general linear function of effective resistances (or end-to-end network criticalities) as $\tau_\alpha = \sum_{s,d}\alpha_{sd}\tau_{sd}$. Traffic-aware network criticality is clearly one example of τ_α with appropriate selection of coefficients (according to equation (28)). In fact, on can show (by rearranging equation (28)) that for TANC we have:

$$\alpha_{sd} = \frac{\gamma_{sd} + \gamma_{ds}}{2\gamma} + \frac{\gamma_{*s} - \gamma_{s*}}{n\gamma} \quad (30)$$

In this paper we are interested in cases where τ_α is convex. Therefore we have to guarantee that the coefficients of $\tau_{sd} = \tau_{ds}$ are non-negative, that is $\alpha_{sd} + \alpha_{ds} \geq 0 \ \forall s,d \in N$. This provides a subset of admissible traffic sets given in the following:

$$\gamma_{sd} + \gamma_{ds} \geq \frac{1}{n}(\gamma_{s*} - \gamma_{*s} + \gamma_{d*} - \gamma_{*d}) \ \forall s,d \in N \quad (31)$$

In the rest of the paper we assume that our traffic matrices belong to this convex subset of all possible choices.

In order to construct the optimization problem, we add a maximum budget constraint to the problem. We assume that there is a cost z_{ij} to deploy each unit of weight on link (i,j). We also assume that there is a maximum budget of C to establish all network links. This constraint means that $\sum_{(i,j)\in E} w_{ij}z_{ij} = C$. Now we can write our optimization problem as follows:

$$\begin{aligned} &Minimize \quad \tau_\alpha \\ &Subject\ to \quad \textstyle\sum_{(i,j)\in E} z_{ij}w_{ij} = C \quad ,C\ is\ fixed \quad (32) \\ &w_{ij} \geq 0 \ \ \forall (i,j) \in E \end{aligned}$$

Theorem 5.1: The condition of optimality for optimization problem (32) can be written as:

$$\min_{(i,j)\in E} \frac{C}{z_{ij}}\frac{\partial \tau_\alpha}{\partial w_{ij}} + \tau_\alpha \geq 0$$

More specifically:

$$w^*_{ij}(C\frac{\partial \tau_\alpha}{\partial w_{ij}} + z_{ij}\tau_\alpha) = 0 \ \ \forall (i,j) \in E \quad (33)$$

where w^*_{ij} denotes the optimal weight for link (i,j).

Proof: The steps of the proof are similar to what we have done for the case of uniform traffic in [5]. We omit the details of the proof due to the lack of space.

■

A. Semidefinite Program (SDP) to Minimize τ_α

In this section we find a semidefinite program to solve optimization problem (32). To this end we notice that:

$$\begin{aligned} \tau_\alpha &= \sum_{s,d}\alpha_{sd}\tau_{sd} \\ &= \sum_{ij}\alpha_{sd}u^t_{sd}L^+u_{sd} \\ &= \sum_{ij}\alpha_{sd}Tr(u_{sd}u^t_{sd}L^+) \\ &= Tr(U_\alpha L^+) \quad (34) \end{aligned}$$

where $U_\alpha = \sum_{sd}\alpha_{sd}U_{sd}$ and $U_{sd} = u_{sd}u^t_{sd}$. It is easy to see that U_α is a symmetric matrix with the sum of the entries of its rows equal to zero, and for $\alpha_{sd} + \alpha_{ds} \geq 0 \ \forall s,d \in N$, it is a positive semidefinite matrix (sufficient condition). As we stated earlier, in this paper we only consider the case of $\alpha_{sd}+\alpha_{ds} \geq 0 \ \forall s,d \in N$, and the general case will be reported in our future papers. In terms of the traffic matrix we consider a subset of admissible set of traffic matrices that guarantee the positive semi-definiteness of U_α. More precisely, we consider those traffic matrices which satisfy equation (31).

For the set of traffic matrices given by (31), optimization problem (32) can be converted to the following semidefinite programming problem (SDP):

$$\begin{aligned} &Minimize \quad \textstyle\sum_{s,d\in N}\alpha_{sd}t_{sd} \quad (35) \\ &Subject\ to \quad Tr(Z^tW) = C \quad ,C\ is\ fixed \\ &diag(Vec(W)) \succeq 0 \\ &\begin{pmatrix} t_{sd} & u^t_{sd} \\ u_{sd} & L + \frac{J}{n} \end{pmatrix} \succeq 0 \quad \forall s,d \in N \end{aligned}$$

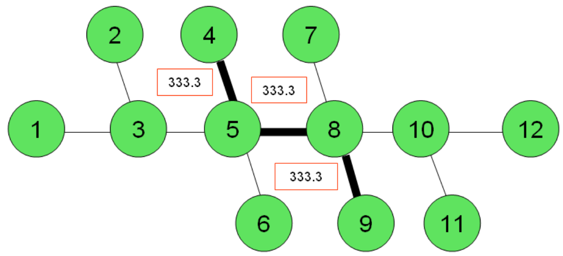

Fig. 1. Optimal Weight Assignment for Parking-Lot to Minimize τ_{49}

Link	Optimal Link Weight
(1,3)	77.2654
(2,3)	77.2654
(3,5)	121.0498
(4,5)	77.2656
(5,6)	77.2653
(5,8)	**139.7768**
(7,8)	77.2655
(8,9)	77.2656
(8,10)	121.0496
(10,11)	77.2655
(10,12)	77.2655

TABLE I
OPTIMAL WEIGHTS FOR PARKING-LOT IN CASE OF EQUAL TRAFFIC

where $Vec(W)$ is a vector obtained by concatenating all the rows of weight matrix W and $diag(x)$ means a diagonal matrix with main diagonal equal to vector x.

Optimization problem (35) provides a framework for designing appropriate link weights to minimize the average network utilization. One can then derive different algorithms for traffic engineering in communication networks to minimize average network utilization. In this paper we concentrate on the first goal and try to solve optimization problem (35) for some representative networks. Optimization problem (35) can be solved with standard methods for solving semidefinite programs. There are also various commercial and academic software tools to solve semidefinite programs. We used open-source CVX package [12], [13] for our examples in this paper.

VI. CASE STUDY

In this section we solve optimization problem (35) for two different network topologies: parking-lot and general tree topology.

A. Parking-Lot Network

We consider a parking-lot network on 12 nodes as shown in Fig. 1. First we study the case where there is traffic only between two specific nodes a and b (i.e. $\gamma_{ij} = 0 \quad \forall i,j \in N$ except for $\gamma_{ab} = \gamma_{ba}$). In this case using equation (28), traffic-aware network criticality (and average network utilization) can be written as: $\tau' = \gamma_{ab}\tau_{ab}$. Therefore, we effectively minimize end-to-end effective resistance or network criticality between two end points a and b by allocating the weights along the links between a and b. In order to be more specific let's assume $a = 4$ and $b = 9$ in parking-lot topology for Fig. 1. We assume that the cost of all links are equal to 1 and the maximum budget is set to $C = 2000$. Then the optimal weight assignment to minimize τ_{49} in the parking-lot topology is highlighted in Fig. 1 by thick lines.

Fig. 1 shows that the optimal distribution of weights is along the shortest path from source node (node 4) to the destination node (node 9) and the weights are evenly distributed among the links on the shortest path. In fact, in the minimization of point-to-point network criticality (effective resistance), the resulting graph is not continuous.

As another example consider the same network, but now suppose that all the elements of traffic matrix are equal to 1. Then solving the optimization problem (35) will result in the optimal distribution of weights as given in Table I.

Finally, consider a nonuniform traffic matrix as defined in the following equation: $\gamma_i(j) = \frac{1}{(i+j)^2} \quad \forall i,j \in N, \quad i \neq j$. Table II gives the optimal weight assignment in this case. It is seen

Link	Optimal Link Weight
(1,3)	138.3109
(2,3)	125.3322
(3,5)	**144.1829**
(4,5)	93.0090
(5,6)	74.2826
(5,8)	119.2677
(7,8)	67.6480
(8,9)	57.5601
(8,10)	82.9718
(10,11)	50.2098
(10,12)	47.2251

TABLE II
OPTIMAL WEIGHTS FOR PARKING-LOT: $\gamma_i(j) = \frac{1}{(i+j)^2}$

that link $(5,8)$ is not the most critical one anymore, because of the nonuniformity in traffic. The highest load is now shifted to the link $(3,5)$, therefore, this links need special attention. For example, in case we want to design backup mechanisms, for this traffic matrix link $(3,5)$ needs highest shared bandwidth.

B. A General Tree Network

In this section we derive a general formula for WNC of a tree with a nonuniform traffic matrix $[\gamma_{ij}]$. We note that a tree is an acyclic simple graph, which means that there is exactly one path between every two nodes of a tree. It follows that network criticality of a tree can be found from the following equation.

$$\tau_\alpha = \sum_{(i,j)\in E} \frac{\lambda_{ij}}{w_{ij}} \tag{36}$$

where λ_{ij} denotes the total traffic passing through link (i,j) (λ_{ij} is the sum of those components of the traffic matrix whose end-to-end path traverses link (i,j)). Using equation (36) we have:

$$\frac{\partial \tau_\alpha}{\partial w_{ij}} = -\frac{\lambda_{ij}}{w_{ij}^2} \tag{37}$$

Considering the condition of optimality given by equation (33) and using equation (37) we get:

$$\frac{\partial \tau_\alpha}{\partial w_{ij}} = -\frac{\lambda_{ij}}{w_{ij}^2} = -\frac{z_{ij}\tau_\alpha}{C}$$

Therefore

$$\Rightarrow w_{ij} = (\frac{\lambda_{ij} C}{z_{ij}\tau_\alpha})^{\frac{1}{2}} \tag{38}$$

From the constraint of the optimization problem we have $\sum_{(i,j)\in E} z_{ij} w_{ij} = C$, hence:

$$\sum_{(i,j)\in E} (\frac{\lambda_{ij} z_{ij} C}{\tau_\alpha})^{\frac{1}{2}} = C \tag{39}$$

$$\tau_\alpha = (\sum_{(i,j)\in E} (\frac{\lambda_{ij} z_{ij}}{C})^{\frac{1}{2}})^2 \tag{40}$$

Now it is enough to substitute τ from equation (40) in equation (38) to have optimal weight for tree.

$$w_{ij} = (\frac{\lambda_{ij} C}{z_{ij}})^{\frac{1}{2}} \times \frac{1}{\sum_{(i,j)\in E} (\frac{\lambda_{ij} z_{ij}}{C})^{\frac{1}{2}}}$$

Finally

$$w_{ij} = \frac{C}{z_{ij}} \times \frac{(\lambda_{ij} z_{ij})^{\frac{1}{2}}}{\sum_{(i,j)\in E} (\lambda_{ij} z_{ij})^{\frac{1}{2}}} \tag{41}$$

Equation (41) shows that the optimal weight of a link in a tree is proportional to the square root of λ_{ij}.

1) Capacity Planning for a Tree: The capacity assignment problem for a tree when the link loads are known can be solved by applying the following changes in equation (41):

$$\begin{aligned} w_{ij} &\rightarrow c_{ij} - \lambda_{ij} \\ C &\rightarrow C - \sum_{(i,j)\in E} z_{ij}\lambda_{ij} \end{aligned}$$

The optimal capacity assignment for a tree would be:

$$c_{ij} = \gamma_{ij} + \frac{C - \sum_{(i,j)\in E} z_{ij}\gamma_{ij}}{z_{ij}} \times \frac{(\lambda_{ij} z_{ij})^{\frac{1}{2}}}{\sum_{(i,j)\in E} (\lambda_{ij} z_{ij})^{\frac{1}{2}}} \tag{42}$$

There is a close analogy between our result and Kleinrock's result for capacity assignment. In [11] Kleinrock showed that under the independence assumption the optimal capacity (to minimize average delay of the network) of a link is proportional to the square root of the link rate. Note that λ_{ij} is the link load, as a result, equation (42) is similar to the Kleinrock's equation for optimal capacity ([11], §5.7, equation 5.26). This result is not surprising because the network criticality of a tree according to equation (36) is equal to $\tau = \sum_{(i,j)\in E} \frac{\lambda_{ij}}{c_{ij} - \lambda_{ij}}$ (considering $w_{ij} = c_{ij} - \lambda_{ij}$). This is the same expression that is used in [11] to find the average delay of a network ([11], §5.6, equation 5.19), therefore, the minimization of network criticality is equal to the minimization of the average network delay when the network is a tree.

VII. Conclusion

In this paper we defined a new metric, random-walk traffic-aware betweenness (TAB) for the nodes and links of a graph as an extension to the previous definition of random-walk betweenness. We investigated the properties of TAB and showed that for the case of generic random-walk, where the transition probabilities are proportional to the link weights, the TAB normalized by the weight can be written as a linear function of end-to-end network criticalities. This led to the definition of traffic-aware network criticality (TANC) as an extention of the effective resistance. We established a semidefinite program for the optimization problem with some appropriate constraints and studied its behavior for parking-lot and general tree topologies.

We need to investigate the properties of the optimization problem to minimize TANC in detail. We plan to use traffic-aware network criticality to evaluate different existing data-center topologies at the presence of a traffic matrix and propose suggestions and new directions to design data-center topologies in general.

References

[1] L. C. Freeman. Centrality in Networks: I. Conceptual Clarification. *Social Networks*, (1):215–239, 1978/79.

[2] L.C. Freeman, S.P. Borgatti, and D.R. White. Centrality in valued graphs: a measure of betweenness based on network flow. *Social Networks*, 13:141–154, 1991.

[3] L. S. P. Borgatti. Centrality and Network Flow. *Social Networks*, 27(1):55–71, 2005.

[4] M. Newman. A Measure of Betweenness Centrality Based on Random Walks. *arXiv cond-mat/0309045.*, 2003.

[5] A. Tizghadam and A. Leon-Garcia. A Graph Theoretical Approach to Traffic Engineering and Network Control Problem. In 21^{th} *International Teletraffic Congress (ITC21)*, Paris, France, September 2009.

[6] A. Tizghadam and A. Leon-Garcia. Autonomic Traffic Engineering for Network Robustness. *IEEE Journal of Selected Areas in Communications (J-SAC)*, 28(1):39 – 50, January 2010.

[7] C. R. Rao and S. K. Mitra. *Generalized Inverse of Matrices and its Applications*. John Weily and Sons Inc., 1971.

[8] D. J. Klein and M. Randic. Resistance Distance. *Journal of Mathematical Chemistry*, 12(1):81–95, December 1993.

[9] A. Ghosh, S. Boyd, and A. Saberi. Minimizing Effective Resistance of a Graph. *SIAM Review, problems and techniques section*, 50(1):37–66, February 2008.

[10] A. Tizghadam and A Leon-Garcia. Survival Value of Communication Networks. In *INFOCOM Workshop on Network Science for Communications (NetSciCom)*, pages 1–6, Rio de Janeiro, Brazil, April 2009.

[11] L. Kleinrock. *Queueing Systems*, volume II. John Wiley & Sons, 1975.

[12] M. Grant and S. Boyd. CVX: Matlab Software for Disciplined Convex Programming (Web Page and Software). $http://stanford.edu/ boyd/cvx$. September 2008.

[13] M. Grant and S. Boyd. Graph Implementations for Nonsmooth Convex Programs, Recent Advances in Learning and Control (a tribute to M. Vidyasagar), V. Blondel, S. Boyd, and H. Kimura, editors, $http://stanford.edu/ boyd/graph_dcp.html$. *Lecture Notes in Control and Information Sciences, Springer*, 2008.