This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

A New Measure of Wireless Network Connectivity

Soura Dasgupta, *Fellow IEEE*, Guoqiang Mao, *Senior Member IEEE* and Brian Anderson, *Life Fellow IEEE*

***Abstract*—Despite intensive research in the area of network connectivity, there is an important category of problems that remain unsolved: how to characterize and measure the *quality of connectivity* of a wireless network which has a realistic number of nodes, *not necessarily* large enough to warrant the use of asymptotic analysis, and which has unreliable connections, reflecting the inherent unreliability of wireless communications? The quality of connectivity measures how easily and reliably a packet sent by a node can reach another node. It complements the use of *capacity* to measure the quality of a network in saturated traffic scenarios and provides an intuitive measure of the quality of (end-to-end) network connections. In this paper, we introduce a probabilistic connectivity matrix as a tool to measure the quality of network connectivity. Some interesting properties of the probabilistic connectivity matrix and their connections to the quality of connectivity are demonstrated. We demonstrate that the largest magnitude eigenvalue of the probabilistic connectivity matrix, which is positive, can serve as a good measure of the quality of network connectivity. We provide a *flooding* algorithm whereby the nodes repeatedly flood the network with packets, and by measuring just the number of packets a given node receives, the node is able to asymptotically estimate this largest eigenvalue.**

***Index Terms*—Connectivity, network quality, probabilistic connectivity matrix**

I. Introduction

Connectivity is one of the most fundamental properties of wireless multi-hop networks [1], [2], [3], and is a prerequisite for providing many network functions, e.g. routing, scheduling and localization. A network is said to be *connected* iff there is a (multi-hop) path between any pair of nodes. A network is said to be k-connected iff there are k paths between any pair of nodes that do not share any node in common except the starting and the ending nodes. Of course, k-connectivity is often required for robust operations of the network. These notions however are essentially deterministic and do not allow straightforward reflection into a mathematical model of the fact that some links will successfully transmit some, but not necessarily all, of the time caused by the random and time-varying nature of wireless connections. To deal with probabilistic connections, in this paper we introduce the notion of a *probabilistic connectivity matrix*, and demonstrate that its largest magnitude eigenvalue, which is positive, quantifies the quality of network connectivity. The precise computation of the elements of this connectivity matrix, given the individual link transmission probabilities and the network topology, involves significant calculation; as an alternative we provide a *flooding algorithm*, that computes the largest magnitude eigenvalue in a decentralized fashion using experimental data (with multiple experiments to allow some averaging). The topology and link probabilities do not need to be known. As shown in the sequel, this new measure constitutes the first that can be determined for moderate to small size networks.

We note that there are two general approaches to studying the connectivity problem. The first, spearheaded by the seminal work of Penrose [3] and Gupta and Kumar [1], is based on an asymptotic analysis of large-scale random networks, which considers a network of n nodes that are *i.i.d.* on an area with an underlying uniform distribution. A pair of nodes are directly connected iff their Euclidean distance is smaller than or equal to a given threshold $r(n)$, independent of other connections. So the connection model is deterministic. Some interesting results are obtained on the value of $r(n)$ required for the above network to be *asymptotically almost surely* connected as $n \to \infty$. In [4], these results are extended to provide the radius for k-connectivity. In [5], [6], the authors extended the above results by Penrose and Gupta and Kumar from the unit disk model to a random connection model, in which any pair of nodes separated by a displacement $\boldsymbol{x}$ are directly connected with probability $g(\boldsymbol{x})$, independent of other connections (the well-known log normal model is a special case). The analytical techniques used in this approach have some intrinsic connections to continuum percolation theory [7] which is usually based on a network setting with nodes Poissonly distributed in an infinite area and studies the conditions required for the network to have a connected component containing an infinite number of nodes (in other words, the network *percolates*). We refer readers to [5] for a more comprehensive literature review.

The second approach is based on a deterministic setting and studies the connectivity and other topological properties of a network using algebraic graph theory. Specifically, consider a network with a set of n nodes. Its properties can be studied using its *underlying graph* $G(V,E)$, where $V \triangleq \{v_1, \ldots, v_n\}$ denotes the vertex set and E denotes the edge set. The underlying graph is obtained by representing each node in the network uniquely using a vertex and the converse. An undirected edge exists between two vertices iff there is a direct connection (or link) between the associated nodes[1]. Define an *adjacency matrix* A_G of the graph $G(V,E)$ to

Dasgupta is with the Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA 52242, USA. email: dasgupta@engineering.uiowa.edu. Most of this work was done while he was visiting National ICT Australia, Sydney. He is supported in part by, US NSF grants CCF-0830747, EPS-1101284, CCF-1302456 and CNS-1239509, ONR grant N00014-13-1-0202 and a grant from the Roy J. Carver Charitable Trust.

Mao is with the School of Computing and Communications, the University of Technology, Sydney and National ICT Australia. Email: g.mao@ieee.org. His research is supported by ARC Discovery projects DP110100538 and DP120102030. Mao is the correspondence author of the paper.

Anderson is with the Australian National University and NICTA Ltd, Canberra, A.C.T., Australia. His work is supported by the Australian Research Council through DP- 130103610 and DP-110100538 and by National ICT Australia. Email: brian.anderson@anu.edu.au.

[1] In this paper, we limit our discussions to a *simple graph* (network) where there is at most one edge (link) between a pair of vertices (nodes) and an undirected graph.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

be a symmetric $n \times n$ matrix whose $(i,j)^{th}, i \neq j$, entry is equal to one if there is an edge between v_i and v_j and is equal to zero otherwise. Further, the diagonal entries of A_G are all equal to zero. The *eigenvalues of the graph* $G(V,E)$ are defined to be the eigenvalues of A_G. The network connectivity information, e.g. connectivity and k-connectivity, is entirely contained in its adjacency matrix. Many interesting connectivity and topological properties of the network can be obtained by investigating the eigenvalues of its underlying graph. For example, let $\mu_1 \geq \ldots \geq \mu_n$ be the eigenvalues of a graph G. If $\mu_1 = \mu_2$, then G is disconnected. If $\mu_1 = -\mu_n$ and G is not empty, then at least one connected component of G is nonempty and bipartite [9, p. 28-6]. If the number of distinct eigenvalues of G is r, then G has a *diameter* of at most $r-1$ [10]. Some researchers have also studied the properties of the underlying graph using its Laplacian matrix [11], where the Laplacian matrix of a graph G is defined as $L_G \triangleq D - A_G$ and D is a diagonal matrix with degrees of vertices in G on the diagonal. Particularly, the *algebraic connectivity* of a graph G is the second-smallest eigenvalue of L_G and it is greater than 0 iff G is a connected graph. Further, the algebraic connectivity is also known to be a good indicator of the convergence rate of consensus algorithms [8]. We refer readers to [10] and [12] for a comprehensive treatment of the topic. Reference [9] provides a concise summary of major results in the area. The adjacency matrix, the Laplacian matrix and their associated parameters mainly focus on describing the connectivity between vertices with directed connections. As demonstrated later in this section, it is not trivial to use these tools to quantify the quality of end-to-end connections (especially when the existence of a direct connection between two nodes becomes probabilistic), which is of paramount concern in many communication applications. In this paper, we develop the probabilistic connectivity matrix, a concept defined later in the paper, to fill this theoretical gap.

The research most related to the work to be presented in this paper is possibly the more recent work of Brooks *et al.* [13]. In [13] Brooks *et al.* considered a probabilistic version of the adjacency matrix and defined a *probabilistic adjacency matrix* as a $n \times n$ square matrix M whose $(i,j)^{th}$ entry m_{ij} represents the probability of having a direct connection between distinct nodes i and j, and $m_{ii} = 0$. They observed that the probability that there exists at least one walk of length z between nodes i and j is m_{ij}^z, where m_{ij}^z is the $(i,j)^{th}$ entry of $M \otimes M \otimes \cdots \otimes M$ (z times). Here $C \triangleq A \otimes B$ is defined by $C_{ij} = 1 - \prod_{l \neq i,j} (1 - A_{il}B_{lj})$ where A_{ij}, B_{ij} and C_{ij} are the $(i,j)^{th}$ entries of the $n \times n$ square matrix A, B and C respectively and the operator $\otimes$ is associative, so that powers are well-defined. A *walk* of length z between nodes i and j is a sequence of z edges, where the first edge starts at i, the last edge ends at j, and the starting vertex of each intermediate edge is the ending vertex of its preceding edge. A *path* of length z between nodes i and j is a walk of length z in which the edges are distinct. Obviously, the existence of a walk implies the existence of a path and conversely. Further, the existence of a walk of length z implies the existence of a path of length smaller than or equal to z. Considering that in a walk, an edge may appear more than once whereas in a path, all edges are distinct, it is not trivial to use their result to derive the probability of existence of a path or the probability of existence of a path of a particular length.

An important category of problems remain unsolved: how to measure the *quality of connectivity* of a wireless multi-hop network which has a realistic number of nodes, *not necessarily* large enough to warrant the use of asymptotic analysis, and has unreliable connections, reflecting the inherent unreliable characteristics of wireless communications? The quality of connectivity measures how easily and reliably packets sent by a node can reach another. It complements the use of *capacity* to measure the quality of a network in saturated traffic scenarios and provides an intuitive measure of the quality of (end-to-end) network connections. The following paragraphs elaborate on the above question using two examples of networks with a fixed number of nodes and known transmission power.

Example 1. *Assume that the wireless propagation model of a network is known and its characteristics have been quantified through* a priori *measurements or empirical estimation. Further, a link exists between two nodes iff the received signal strength from one node at the other node, whose propagation follows the wireless propagation model and the signal strength is random, e.g. due to fading and shadowing, is greater than or equal to a predetermined threshold* and *the same is also true in the opposite direction. One can then find the probability that a link exists between two nodes at two fixed locations: It is determined by the probability that the received signal strength is greater than or equal to the pre-determined threshold. Two related questions can be asked: a) If these nodes are deployed at a set of known locations, what is the quality of connectivity of the network, measured by the probability that there is a path between any two nodes, as compared to node deployment at another set of locations? b) How can one optimize the node deployment to maximize the quality of connectivity?*

Example 2. *The transmission between a pair of nodes with a direct connection, say v_i and v_j, may fail with a known probability, say $1 - a_{ij}$, quantifying the inherent unreliable characteristics of wireless communications. There are no direct connections between some pairs of nodes because the probability of successful transmission between them is too low to be acceptable. How should one measure the quality of connectivity of such a network, in the sense that a packet transmitted from one node can easily and reliably reach another node via a multi-hop path. Will a single "good" path between a pair of nodes be preferable to multiple "bad" paths? These questions are illustrated in Fig. 1 and 2.*

In this paper, we introduce and explore the use of a *probabilistic connectivity matrix*, a concept to be defined later in Section II, as a tool to measure the quality of network connectivity. Some key properties of the probabilistic connectivity matrix and their connections to the quality of connectivity are demonstrated. Armed with certain inequalities derived in Section III, and assuming a symmetric network, in Section IV, we derive several properties of the eigenvalues of the probabilistic connectivity matrix. First we show that in a connected

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

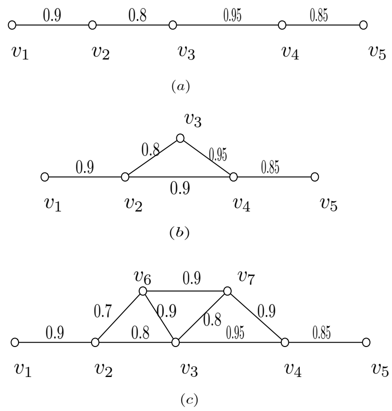

Fig. 1. An illustration of networks with different quality of connectivity. A solid line represents a direct connection between two nodes and the number beside the line represents the corresponding transmission successful probability. The networks shown in (a), (b), and (c) are all connected networks but not 2-connected networks, i.e. their connectivity cannot be differentiated using the k-connectivity concept. However intuitively the quality of the network in (b) is better than that of the network in (a) because of the availability of the additional high-quality link between v_2 and v_4 in (b). The quality of the network in (c) is even better because of the availability of the additional nodes and the associated high-quality links, hence additional routes, if these additional nodes act as relay nodes only. If these additional nodes also generate their own traffic, it is uncertain whether the quality of the network in (c) is better or not. Therefore it is important to develop a measure to quantitatively compare the quality of connectivity (for the networks in (a) and (b)) and to evaluate the benefit of additional nodes on connectivity (for the network in (c)).

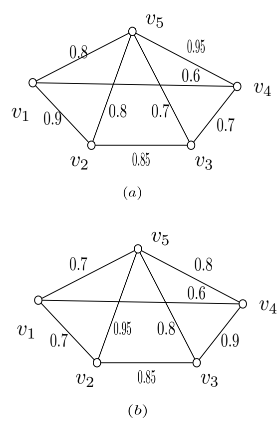

Fig. 2. The networks shown in (a) and (b) have the same topology but different link quality. It is difficult to compare the quality of the two networks.

network, i.e. where there is a path of non-zero probability between every pair of nodes, the largest magnitude eigenvalue, which is positive, does indeed quantify the quality of network connectivity. Should the network be disconnected, then we show that it naturally partitions into connected components. Specifically there is a path of nonzero probability between any two nodes in a connected component, but all inter-component paths have zero probability. In this case the probabilistic connectivity matrix is block diagonal, each diagonal block in turn being the connectivity matrix of a particular component. In this case the largest magnitude eigenvalue provides the connectivity measure of this component. We show also that the matrix is positive semidefinite, and is in fact positive definite, unless there is a path in the network that has probability one.

We also show that increasing a link probability increases the largest eigenvalue of the component to which the link belongs. In Section V, exploiting the positive semidefiniteness of this matrix we provide an algorithm that computes the largest eigenvalue in a decentralized fashion using experimental measurements on the network, including averaging over a number of experiments. Specifically this *flooding* algorithm requires the nodes to repeatedly flood the network with packets, and by measuring just the number of packets a given node receives, the node is able to asymptotically estimate this largest eigenvalue without knowing any element of the probabilistic connectivity matrix or the number of packets received by the other nodes. Section VI is the conclusion.

II. The Probabilistic Connectivity Matrix

In this section we define the network to be studied, its probabilistic adjacency matrix and probabilistic connectivity matrix, and gives an approach to computing the probabilistic connectivity matrix.

Consider a network of n nodes. For some pair of nodes, an edge (or link) may exist with a non-negligible probability. The edges are considered to be undirected. That is, if a node v_i is connected to a node v_j, then the node v_j is also connected to the node v_i. Further, as is commonly done in the area [1], [3], [7], [6], it is assumed that the event that there is an edge between a pair of nodes and the event that there is an edge between another distinct pair of nodes (which may include one node in common with the first pair) are independent. In addition to such spatial independence, we also assume *temporal independence*; specifically that each edge event is i.i.d. over time, e.g. due to fading and shadowing. This temporal independence is needed for the results of Section V, and is formalized in that section.

Denote the underlying graph of the above network by $G(V,E)$, where $V = \{v_1, \ldots, v_n\}$ is the vertex set and $E = \{e_1, \ldots, e_m\}$ is the edge set, which contains the set of *all possible* edges, i.e. all vertex pairs for which the probability of being directly connected is nonzero. Here the vertices and the edges are indexed from 1 to n and from 1 to m respectively. For convenience, in some parts of this paper we also use the symbol e_{ij} to denote an edge between vertices v_i and v_j when there is no confusion. We associate with each edge e_i, $i \in \{1, \ldots m\}$, an indicator random variable I_i such that $I_i = 1$ if the edge e_i exists; $I_i = 0$ if the edge e_i does not exist. The indicator random variables I_{ij}, $i \neq j$ and $i, j \in \{1, \ldots n\}$, are defined analogously. Furthermore, we use $(I_i, i \in \{1, 2, \ldots, m\})$ to denote a particular instance of the indicator random variables associated with an instance of the random edge set.

In the following, we give a definition of the probabilistic adjacency matrix, differing mildly from that of Brooks *et al*, [13] as described further below :

Definition 1. *The probabilistic adjacency matrix of* $G(V,E)$*, denoted by* A_G*, is a* $n \times n$ *matrix whose* $(i,j)^{th}$*,* $i \neq j$*, entry* $a_{ij} \triangleq \Pr(I_{ij} = 1)$ *and its diagonal entries are all equal to* 1.

Due to the undirected property of an edge mentioned above, A_G is a symmetric matrix, i.e. $a_{ij} = a_{ji}$. Note that the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

diagonal entries of A_G are defined to be 1, which is different from the usual convention in the literature, e.g. [13]. In [14] we have discussed the implication of this definition in the context of mobile *ad hoc* networks. This treatment of the diagonal entries reflects the fact that if a node in the network finds the wireless channel busy, it can store a packet (or equivalently transmit the packet to itself) until the channel is free. A pair of nodes v_i and v_j are said to be *directly connected* if the associated a_{ij} is greater than 0.

The probabilistic connectivity matrix is defined in the following way:

Definition 2. *The probabilistic connectivity matrix of $G(V,E)$, denoted by Q_G, is a $n \times n$ matrix whose $(i,j)^{th}$, $i \neq j$, entry is the probability that there exists a path between vertices v_i and v_j, and its diagonal entries are all equal to* 1.

As a ready consequence of the symmetry of A_G, Q_G is also a symmetric matrix. Further, the following property of Q_G can be easily obtained from the above definition. The Lemma refers to the *direct sum* between matrices, defined as $A \bigoplus B = \textbf{diag}\ \{A, B\}$.

Lemma 1. *Suppose A_G defined in Definition 1 is symmetric. Then the probabilistic connectivity matrix Q_G is a symmetric nonnegative matrix. If it has a zero element then there is an ordering of vertices under which Q_G is a direct sum of positive matrices.*

Proof. Symmetry of Q_G follows from the symmetry of A_G. Nonnegativity of Q_G follows from the fact that its diagonal elements are one and the rest are probabilities. Now suppose for some i, j, $q_{ij} = q_{ji} = 0$ but that for some k, $q_{ik} = q_{ki} \neq 0$. This indicates that all paths between v_i and v_j have zero probability (henceforth, v_i and v_j are not connected) but at least one between v_k and v_i has a nonzero probability (v_k and v_i are connected). Thus $q_{kj} = q_{jk} = 0$ as otherwise there is a path between v_k and v_j and consequently between v_i and v_k that has nonzero probability, violating the assumption that $q_{ij} = q_{ji} = 0$. Thus, one can partition the vertex set V into sets V_l, such that all nodes in V_l are connected to each other but are not connected to any node in V_m, $m \neq l$. Order the vertices so that for each l those of V_l are consecutive. The resulting Q_G is clearly a direct sum of positive matrices. □

Remark 1. *We call the network connected if Q_G is positive, as there is then a nonzero probability that a path exists between any two nodes. Lemma 1 and its proof also formalize the fact that a network that is not connected partitions into disjoint components, each of which is connected, but all paths between nodes from different components have probability zero (we are not distinguishing conceptually between the notion that a link or path may not exist, and the notion that a link or path always has zero probability).*

Given the probabilistic adjacency matrix A_G, the probabilistic connectivity matrix Q_G is fully determined. However the computation of Q_G is not trivial because for a pair of vertices v_i and v_j, there may be multiple paths between them and some of the paths may share common edges, i.e. paths are not *independent* or are *spatially correlated*. In the rest of this section, we give a method to compute Q_G.

A. Computation of the probabilistic connectivity matrix

We now indicate in rather formal language the conceptual basis of computing the probabilistic connectivity matrix Q_G.

Let $Q_G|\,(I_i, i \in \{1, 2, \ldots, m\})$ be the connectivity matrix of G *conditioned on* a particular instance of the indicator random variables $I_1, \ldots, I_m$ associated with an instance of the random edge set. The $(i,j)^{th}$ entry of $Q_G|\,(I_i, i \in \{1, 2, \ldots, m\})$ is either 0, when there is no path between v_i and v_j, or 1 when there exists a path between v_i and v_j (see also Lemma 4). The diagonal entries of $Q_G|\,(I_i, i \in \{1, 2, \ldots, m\})$ are always 1. Conditioned on $(I_i, i \in \{1, 2, \ldots, m\})$, $G(V,E)$ is just a deterministic graph. Therefore the entries of $Q_G|\,(I_i, i \in \{1, 2, \ldots, m\})$ can be efficiently computed using a search algorithm, such as breadth-first search. Given $Q_G|\,(I_i, i \in \{1, 2, \ldots, m\})$, Q_G can be computed using the following:

$$Q_G = E_{(I_i, i \in \{1,2,\ldots,m\})}\left(Q_G|\,(I_i, i \in \{1, 2, \ldots, m\})\right) \quad (1)$$

where the expectation is taken over all possible instances of $(I_i, i \in \{1, 2, \ldots, m\})$.

Using the technique introduced in the previous paragraph, the probabilistic connectivity matrix of the three networks in Fig. 1 and two networks in Fig. 2, denoted by Q_{1a}, Q_{1b}, Q_{1c}, Q_{2a} and Q_{2b} respectively, can be computed. For example,

$$Q_{2a} = \begin{bmatrix} 1.0000 & 0.9876 & 0.9744 & 0.9823 & 0.9880 \\ 0.9876 & 1.0000 & 0.9812 & 0.9856 & 0.9916 \\ 0.9744 & 0.9812 & 1.0000 & 0.9780 & 0.9827 \\ 0.9823 & 0.9856 & 0.9780 & 1.0000 & 0.9926 \\ 0.9880 & 0.9916 & 0.9827 & 0.9926 & 1.0000 \end{bmatrix} \quad (2)$$

$$Q_{2b} = \begin{bmatrix} 1.0000 & 0.9603 & 0.9571 & 0.9540 & 0.9614 \\ 0.9603 & 1.0000 & 0.9918 & 0.9854 & 0.9961 \\ 0.9571 & 0.9918 & 1.0000 & 0.9879 & 0.9936 \\ 0.9540 & 0.9854 & 0.9879 & 1.0000 & 0.9878 \\ 0.9614 & 0.9961 & 0.9936 & 0.9878 & 1.0000 \end{bmatrix} \quad (3)$$

A comparison of the entries of Q_{2a} and Q_{2b} leads to intuitive and quantitative conclusion on the quality of end-to-end paths between any pair of nodes in the two networks in Fig. 2.a and 2.b. In the rest of this paper, we will further establish properties of the probabilistic connectivity matrix that facilitates the analysis of network quality and connectivity.

The approach suggested in the last paragraph is essentially a brute-force approach to computing Q_G. More efficient algorithms can be possibly designed to compute Q_G. Indeed in Section IV we suggest an approach to simplify the computation of Q_G via a recursive procedure exploiting the property of Q_G. Since the main focus of the paper is on exploring the properties of Q_G that facilitate the connectivity analysis, an extensive discussion of designing computationally efficient algorithms to compute Q_G is left for future work.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

That said, the complications in computing Q_G are mitigated by the fact that a measure of connectivity developed in this paper can also be estimated using experimental data without explicitly obtaining the elements of Q_G. This measure is the largest eigenvalue of Q_G. As shown in the sequel it can be asymptotically estimated in a completely decentralized fashion without knowing the entries of Q_G or the link probabilities and network topology.

Remark 2. *For simplicity, the terms used in our discussion are based on the problems in Example 1. The discussion however can be easily adapted to the analysis of the problems in Example 2. For example, if a_{ij} is defined to be the probability that a transmission between nodes v_i and v_j is successful, the $(i,j)^{th}$ entry of the probabilistic connectivity matrix Q_G computed using* (1) *then gives the probability that a transmission from v_i to v_j via a multi-hop path is successful under the best routing algorithm, which can always find a shortest and error-free path between from v_i to v_j if it exists, or alternatively, the probability that a packet flooded from v_i can reach v_j where each node receiving the packet only broadcasts the packet to its directly-connected neighbors once. Therefore the $(i,j)^{th}$ entry of Q_G can be used as a quality measure of the end-to-end paths between v_i and v_j, which takes into account the fact that availability of an extra path between a pair of nodes can be exploited to improve the probability of successful transmissions.*

III. Some Key Inequalities for Connection Probabilities

The entries of the probabilistic connectivity matrix give an intuitive idea about the overall quality of end-to-end paths in a network. In this section, we provide some important inequalities that may facilitate the analysis of the quality of connectivity. Some of these inequalities are exploited in the next section to establish some key properties of the probabilistic connection matrix itself.

We first introduce some concepts and results that are required for the further analysis of the probabilistic connectivity matrix Q_G.

For a random graph with a given set of vertices, a particular event is *increasing* if the event is preserved when more edges are added into the graph. An event is *decreasing* if its complement is increasing.

The following theorems summarizing a relevant form of the FKG inequality and BK inequality respectively will be used:

Theorem 1. *[7, Theorem 1.4] (FKG Inequality) If events A and B are both increasing events or decreasing events depending on the state of finitely many edges, then*

$$\Pr(A\cap B)\geq \Pr(A)\Pr(B)$$

Theorem 2. *[15], [7, Theorem 1.5] (BK Inequality) If events A and B are both increasing events depending on the state of finitely many edges, then*

$$\Pr(A\Box B)\leq \Pr(A)\Pr(B)$$

where for two events A and B, $A\Box B$ denotes the event that there exist two disjoint *sets of edges such that the first set of edges guarantees the occurrence of A and the second set of edges guarantees the occurrence of B.*

Denote by ξ_{ij} the event that there is a path between vertices v_i and v_j, $i\neq j$. Denote by ξ_{ikj} the event that there is a path between vertices v_i and v_j *and* that path passes through the third vertex v_k, where $k\in\Gamma_n\backslash\{i,j\}$ and Γ_n is the set of indices of all vertices. Denote by η_{ij} the event that there is an edge between vertices v_i and v_j. Denote by π_{ikj} the event that there is a path between vertices v_i and v_k *and* there is a path between vertices v_k and v_j, where $k\in\Gamma_n\backslash\{i,j\}$. Obviously

$$\pi_{ikj}\Rightarrow\xi_{ij} \tag{4}$$

It is clear from the above definitions that

$$\xi_{ij}=\eta_{ij}\cup\left(\cup_{k\neq i,j}\xi_{ikj}\right) \tag{5}$$

Let q_{ij}, $i\neq j$, be the $(i,j)^{th}$ entry of Q_G, i.e. $q_{ij}=\Pr(\xi_{ij})$. The following theorem is obtained from the FKG inequality and the above definitions.

Theorem 3. *For two distinct indices $i,j\in\Gamma_n$ and $\forall k\in\Gamma_n\backslash\{i,j\}$*

$$q_{ij}\geq\max_{k\in\Gamma_n\backslash\{i,j\}}q_{ik}q_{kj} \tag{6}$$

$$q_{ij}\leq 1-(1-a_{ij})\prod_{k\in\Gamma_n\backslash\{i,j\}}(1-q_{ik}q_{kj}) \tag{7}$$

where $a_{ij}=\Pr(\eta_{ij})$.

Proof. We first prove inequality (6). It follows readily from the above definitions that the event ξ_{ij} is an increasing event. Due to (4) and the FKG inequality:

$$\Pr(\xi_{ij})\geq\Pr(\pi_{ikj})\geq\Pr(\xi_{ik})\Pr(\xi_{kj}) \tag{8}$$

The conclusion follows.

Now we prove the second inequality (7). We will first show that $\xi_{ikj}\Leftrightarrow\xi_{ik}\Box\xi_{kj}$. That is, the occurrence of the event ξ_{ikj} is a sufficient and necessary condition for the occurrence of the event $\xi_{ik}\Box\xi_{kj}$.

Using the definition of ξ_{ikj}, occurrence of ξ_{ikj} means that there is a path between vertices v_i and v_j *and* that path passes through vertex v_k. It follows that there exist a path between vertex v_i and vertex v_k *and* a path between vertex v_k and vertex v_j *and* the two paths do not have edge(s) in common. Otherwise, it will contradict the definition of ξ_{ikj}, noting that the definition of a path requires its edges to be distinct. Therefore $\xi_{ikj}\Rightarrow\xi_{ik}\Box\xi_{kj}$. Likewise, $\xi_{ikj}\Leftarrow\xi_{ik}\Box\xi_{kj}$ also follows directly from the definitions of ξ_{ikj}, ξ_{ik}, ξ_{kj} and $\xi_{ik}\Box\xi_{kj}$. Consequently

$$\Pr(\xi_{ikj})=\Pr(\xi_{ik}\Box\xi_{kj})\leq\Pr(\xi_{ik})\Pr(\xi_{kj}) \tag{9}$$

where the inequality is a direct result of the BK inequality.

Note that the event $\cup_{k\in\Gamma_n\backslash\{i,j\}}\xi_{ikj}$ and the event η_{ij} are independent because the existence of a direct connection between v_i and v_j has no impact on the event $\cup_{k\in\Gamma_n\backslash\{i,j\}}\xi_{ikj}$. Therefore using (5) and independence of edges (used in the second step)

$$q_{ij}\quad=\quad\Pr\left(\eta_{ij}\cup\left(\cup_{k\in\Gamma_n\backslash\{i,j\}}\xi_{ikj}\right)\right)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

$$\begin{aligned} &= 1-(1-a_{ij})\Pr\left(\cap_{k\in\Gamma_n\backslash\{i,j\}}\overline{\xi_{ikj}}\right) \\ &\leq 1-(1-a_{ij})\prod_{k\in\Gamma_n\backslash\{i,j\}}\Pr\left(\overline{\xi_{ikj}}\right) \quad (10) \\ &\leq 1-(1-a_{ij})\prod_{k\in\Gamma_n\backslash\{i,j\}}(1-q_{ik}q_{kj}) \quad (11) \end{aligned}$$

where in (10), FKG inequality and the obvious fact that $\overline{\xi_{ikj}}$ is a decreasing event are used and the last step is from (9). □

Remark 3. *Inequality* (6) *also provides another proof of a key relationship used in Lemma 1. Specifically, if $q_{ij}=0$ then this inequality implies that at least one among q_{ik} and q_{kj} must be zero. Likewise if neither q_{ik} nor q_{kj} is zero, then $q_{ij}>0$.*

When there is no edge between vertices v_i and v_j, the upper and lower bounds in Theorem 2 reduce to

$$\max_{k\in\Gamma_n\backslash\{i,j\}} q_{ik}q_{kj} \leq q_{ij} \leq 1-\prod_{k\in\Gamma_n\backslash\{i,j\}}(1-q_{ik}q_{kj}) \quad (12)$$

The above inequality sheds insight on how the quality of paths between a pair of vertices is related to the quality of paths between other pairs of vertices. It can be possibly used to determine the most effective way of improving the quality of a particular set of paths by improving the quality of a particular (set of) edge(s), or equivalently what can be reasonably expected from an improvement of a particular edge on the quality of end-to-end paths. Further, an immediate consequence of this inequality is that: If $q_{ij}=0$, then at least one of q_{ik} and q_{kj} must be 0 for all $k\neq i,j$.

The following lemma further shows that the occurrence of a certain relation among entries of the probabilistic connectivity matrix Q_G can be used to derive some topological information of the graph.

Lemma 2. *If $q_{ij}=q_{ik}q_{kj}$ for distinct vertices v_i, v_j and v_k, the vertex set V of the graph $G(V,E)$ can be divided into three non-empty and non-intersecting sub-sets V_1, V_2 and V_3 such that $v_i\in V_1$, $v_j\in V_3$ and $V_2=\{v_k\}$* and *any possible path between a vertex in V_1 and a vertex in V_2 must pass through v_k, and the converse. Further, for any pair of vertices v_l and v_m, where $v_l\in V_1$ and $v_m\in V_3$, $q_{lm}=q_{lk}q_{km}$.*

Proof. Using (8) in the second step, it follows that

$$\begin{aligned} q_{ij} &= \Pr\left(\xi_{ij}\backslash\pi_{ikj}\right)+\Pr\left(\pi_{ikj}\right) \\ &\geq \Pr\left(\xi_{ij}\backslash\xi_{ikj}\right)+q_{ik}q_{kj} \end{aligned}$$

Therefore $q_{ij}=q_{ik}q_{kj}$ implies that $\Pr\left(\xi_{ij}\backslash\pi_{ikj}\right)=0$ or equivalently $\xi_{ij}\Leftrightarrow\pi_{ikj}$

Further, $\Pr\left(\xi_{ij}\backslash\pi_{ikj}\right)=0$ implies that a *possible* path (i.e. a path with a non-zero probability) connecting v_i and v_k and a *possible* path connecting v_k and v_j cannot have any edge in common. Otherwise a path from v_i to v_j, bypassing v_k, exists with a non-zero probability which implies $\Pr\left(\xi_{ij}\backslash\xi_{ikj}\right)>0$. The conclusion follows readily that if $q_{ij}=q_{ik}q_{kj}$ for three distinct vertices v_i, v_j and v_k, the vertex set V of the underlying graph $G(V,E)$ can be divided into three non-empty and non-overlapping sub-sets V_1, V_2 and V_3 such that $v_i\in V_1$, $v_j\in V_3$ and $V_2=\{v_k\}$ *and* a path between a vertex in V_1 and a vertex in V_2, if exists, must pass through v_k.

Further, for any pair of vertices v_l and v_m, where $v_l\in V_1$ and $v_m\in V_3$, it is easily shown that $\Pr\left(\xi_{lm}\backslash\pi_{lkm}\right)=0$. Due to independence of edges and further using the fact that $\Pr\left(\xi_{lm}\backslash\pi_{lkm}\right)=0$, it can be shown that

$$\Pr\left(\xi_{lm}\right)=\Pr\left(\pi_{lkm}\right)=\Pr\left(\xi_{lk}\right)\Pr\left(\xi_{km}\right) \quad (13)$$

where (13) results due to the fact that under the condition of $\Pr\left(\xi_{lm}\backslash\pi_{lkm}\right)=0$, a path between vertices v_l and v_k and a path between vertices v_k and v_m cannot possibly have any edge in common. □

An implication of Lemma 2 is that for any three distinct vertices, v_i, v_j and v_k, if a relationship $q_{ij}=q_{ik}q_{kj}$ holds, vertex v_k must be a *critical* vertex whose removal will render the graph disconnected.

IV. The Largest Eigenvalue of Q_G

We now establish a measure of the quality of network connectivity. Just as the eigenvalues of the adjacency matrix provide a deterministic measure of connectivity, we now provide a series of arguments supporting the contention that a similar property can be ascribed to certain eigenvalues of the probabilistic connectivity matrix Q_G.

From Lemma 1, Q_G is a non-zero nonnegative matrix. Thus from the Perron-Frobenius Theorem, [24], its largest magnitude eigenvalue, known as the *Perron-Frobenius eigenvalue* is real and positive. Further as Q_G is symmetric, all its eigenvalues are real, and its largest magnitude eigenvalue $\lambda_{\max}(Q_G)$ is also its largest singular value. Also from the Perron-Frobenius Theorem, should the network be connected, i.e. Q_G is positive as opposed to just nonnegative, this eigenvalue is simple.

We now argue that $\lambda_{\max}(Q_G)$ quantifies the quality of network connectivity. Indeed suppose that i-th node v_i transmits x_i number of packets in a time interval. This means that v_i floods the packet across the entire network and each node receiving the packet only broadcasts the packet once to its directly connected neighbors. If the same packet is received more than once by the same node, it is counted as one packet. Let $x=[x_1,\cdots,x_n]^\top$ and let y_i denote the expected number of packets received by the i-th node, $y=[y_1,\cdots,y_n]^\top$. Then by definition: $y=Q_Gx$. As the basic purpose of any network is to transport packets from some nodes in the network to some others, a measure of connectivity that naturally arises is *the largest size of y relative to x.* One measure of the size of y is its 2-norm, denoted by $\|y\|_2$. Then as Q_G is symmetric and non-negative,

$$\begin{aligned} \max_{\|x\|_2\neq 0}\frac{\|y\|_2}{\|x\|_2} &= \max_{\|x\|_2\neq 0}\frac{\sqrt{y^\top y}}{\sqrt{x^\top x}}=\max_{\|x\|_2\neq 0}\frac{\sqrt{x^\top Q_G^\top Q_G x}}{\sqrt{x^\top x}} \\ &= \max_{\|x\|_2\neq 0}\sqrt{\frac{x^\top Q_G^2 x}{x^\top x}}=\lambda_{\max}(Q_G). \end{aligned}$$

It is well known that for a symmetric Q_G, the maximum ratio is attained when x is the eigenvector associated with the eigenvalue $\lambda_{max}(Q_G)$. Observe also from Perron-Frobenius theory, [24], that as Q_G is nonnegative, the eigenvector associated with $\lambda_{max}(Q_G)$ has all entries of the same sign, without loss of generality nonnegative. Thus the largest value

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

of $\max_{\|x\|_2 \neq 0} \frac{\|y\|_2}{\|x\|_2}$ is itself attained by an x with nonnegative elements. Thus indeed one can strengthen the equality above to state that:

$$\max_{\|x\|_2 \neq 0, x_i \geq 0} \frac{\|y\|_2}{\|x\|_2} = \lambda_{\max}(Q_G).$$

Consequently, $\lambda_{\max}(Q_G)$ is a natural measure of network connectivity.

There are two other approaches to characterizing $\lambda_{\max}(Q_G)$: min-max and max-min flow gain:

$$\max_{x>0} \min_i \frac{y_i}{x_i} \text{ and } \min_{x>0} \max_i \frac{y_i}{x_i}.$$

Regardless of whether Q_G is symmetric, its largest magnitude eigenvalue, obeys min-max and max-min type relations through the Collatz-Wielandt equalities (see Corollary 8.1.31 in [17]). In particular,

$$\max_{x>0} \min_i \frac{y_i}{x_i} = \lambda_{\max}(Q_G) = \min_{x>0} \max_i \frac{y_i}{x_i}.$$

The case of using $\lambda_{\max}(Q_G)$ as a measure of connectivity is further supported by the following observation. When Q_G is positive as opposed to just nonnegative, $\lambda_{\max}(Q_G)$ strictly increases with increasing values of its off diagonal elements, [24]. If on the other hand, it has zero elements, then on the face of it, it is merely nondecreasing. However, recall from Lemma 1 and Remark 1, that if there are zero entries in Q_G, the network partitions into disjoint connected components represented by graphs $G_i(V_i, E_i)$, and Q_G itself can be expressed as $Q_G = \bigoplus_{i=1}^{l} Q_{G_i}$, with Q_{G_i} all positive. Should an element of a particular Q_{G_i} increase, then so must its largest eigenvalue. On the other hand for $v_i \in V_i$ and $v_j \in V_j$, $q_{ij} = 0$. Should now this become positive, then we argue that with $G'_{ij} = (V_i \bigcup V_j, E_i \bigcup E_j)$, $\lambda_{\max}(Q_{G'_{ij}})$ does indeed strictly increase. Indeed suppose the new $q_{ij} = q > 0$. Then from Lemma 1, for every $0 < q_{ij} < q$, the resulting $Q_{G'_{ij}}$ is positive and the result follows.

We next establish the remarkable fact that in fact Q_G is a *positive semidefinite* matrix. The implications of the positive semidefiniteness of Q_G will be explored later. At the core of the development leading to this result is the following fact.

Lemma 3. *Each off-diagonal entry of the probabilistic connectivity matrix Q_G is a multiaffine[2] function of a_{ij}.*

Proof. Consider an arbitrary off-diagonal entry, q_{kl} of Q_G. This is the probability that there is a path between vertices v_k and v_l. This event is ξ_{kl}. Enumerate the distinct events constituting a path between v_k and v_l, listing first those not containing edge e_{ij} as $\bar{\xi}_{1,kl}, \ldots \bar{\xi}_{s,kl}$ and then those containing edge e_{ij} as $\bar{\xi}_{s+1,kl} \wedge \eta_{ij}, \cdots, \bar{\xi}_{t,kl} \wedge \eta_{ij}$. Of course, the event that a path exists is the intersection of the events η_{pq} for the edges e_{pq} along the path. Evidently,

$$\xi_{kl} = \bar{\xi}_{1,kl} \vee \cdots \vee \bar{\xi}_{s,kl} \vee (\bar{\xi}_{1,kl} \wedge \eta_{ij}) \vee \cdots \vee (\bar{\xi}_{t,kl} \wedge \eta_{ij}) \quad (14)$$

[2]A multiaffine function is affine in each variable when the other variables are fixed.

Because every event η_{ij} is independent of all the other edge connection events, it is easy to verify that q_{kl} equals

$$Pr\left(\bar{\xi}_{1,kl} \vee \cdots \vee \bar{\xi}_{t,kl}\right) a_{ij} + Pr\left(\bar{\xi}_{1,kl} \vee \cdots \vee \bar{\xi}_{s,kl}\right)(1 - a_{ij}) \quad (15)$$

Since the probabilities multiplying a_{ij} and $1 - a_{ij}$ in (15) are probabilities of events independent of the event η_{ij}, they do not depend on a_{ij}. Thus if we hold a_{pq} with $\{i, j\} \neq \{p, q\}$ constant, q_{kl} is an affine function of a_{ij}. The same applies to every off-diagonal element of A_G. The result follows. □

Note that $Pr(\bar{\xi}_{1.kl} \vee \cdots \vee \bar{\xi}_{t,kl})$ is the probability of a connection between vertices v_k and v_l with the original network modified by eliminating any link between vertices $\{v_i, v_j\}$, while $Pr(\bar{\xi}_{s+1,kl} \vee \cdots \vee \bar{\xi}_{t,kl})$ is the probability of a connection between the same vertices with the original network modified by imposing a perfect connection ($a_{ij} = 1$) between vertices v_i and v_j (equivalently the two vertices are merged); the latter is obviously greater than or equal to the former. The associated matrices are themselves probabilistic connectivity matrices.

Due to this multiaffine property, for $k, l, i, j \in \{1, \ldots n\}$, where $k \neq l$ and $i \neq j$, the following holds:

$$q_{lk} = c_1 a_{ij} + c_2 \quad (16)$$

where c_1 and c_2 are in $[0, 1]$, are determined by the state of the set of edges in $E \setminus \{e_{ij}\}$ only, and are not affected by the state of e_{ij}; $c_2 = 0$ implies that v_l and v_k will be disconnected without the edge e_{ij}. Thus e_{ij} is a *critical* edge for the end-to-end paths between the vertices v_l and v_k. $c_1 = 0$ implies that the state of the edge e_{ij} is irrelevant for the end-to-end paths between v_l and v_k. In fact, c_1 measures the *criticality* of the edge e_{ij} to the end-to-end paths between v_l and v_k.

Using the multiaffine property, a more efficient algorithm for computing Q_G than the one suggested earlier using (1) can be constructed. Particularly, the probabilistic connectivity matrix of a network forming a tree can be easily computed. Therefore the algorithm may start by first identifying a spanning tree in $G(V, E)$ and computing the associated probabilistic connectivity matrix. Then, the edges in E but outside the spanning tree can be added recursively and the corresponding probabilistic connectivity matrix updated using (16). Since the computational complexity of Q_G depends on $2^{|E|}$, let l be the number of edges in the spanning tree, the computational complexity improves approximately by a factor of 2^l compared with the algorithm using (1) directly. We intend to explore in a forthcoming paper algorithms for computing Q_G from the a_{ij} and network topology. That said, a key purpose of this paper is to postulate and justify as valid, a measure of network connectivity and to formulate a procedure for estimating this measure, without having to explicitly obtain Q_G. The following remark is also instructive.

Remark 4. *Several papers have exploited multiaffine variations. These include the design of adaptive estimation algorithms, [20]-[22] and stability analysis [18], [19] and [25]. All exploit the fact that variations are individually affine in each variable as long as the other variables are fixed. The fact*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

that there is an increasing relationship between the elements of Q_G and $\lambda_{\max}(Q_G)$ and the latter depend multiaffinely on the probabilities a_{ij}, suggests the following obvious optimization. Modify one or more a_{ij} under suitable constraints to maximize $\lambda_{\max}(Q_G)$. The multiaffine dependence of the q_{ij} on the a_{ij} together with the fact that Q_G is positive semi-definite promise to provide several avenues for such optimization.

The basis for these calculations is likely to be the following observation. If $Q_G = a_{ij}Q_{1G}+Q_{2G}$ with Q_{1G}, Q_{2G} independent of a_{ij}, and if x is a positive eigenvector of Q_G associated with the maximum eigenvalue $\lambda_{max}(Q_G)$, then it is easily seen that $\frac{\partial \lambda_{max}}{\partial a_{ij}} = \frac{x^T Q_{1G} x}{x^T x}$.

We now establish that Q_G is positive semidefinite.

Theorem 4. *The matrix $Q_G = Q_G^\top \in \mathbb{R}^{n\times n}$, is a positive semi-definite matrix. It is not positive definite iff there exist $i \neq j$, such that $q_{ij} = 1$.*

We prove this theorem at the end of this section. For the moment we discuss its implications. One in particular is its use in the analysis of the flooding algorithm of the next section. There are also implications to the level of connectivity. Let $\lambda_{\max}(Q_G) \geq \lambda_2(Q_G) \geq \ldots \geq \lambda_{\min}(Q_G) \geq 0$ be the eigenvalues of Q_G. As all diagonal elements of Q_G are one, the trace of Q_G and hence $\lambda_{\max}(Q_G) + \lambda_2(Q_G) \ldots + \lambda_{\min}(Q_G)$ equals n. Thus as an easy consequence of Theorem 4, $n \geq \lambda_{\max}(Q_G) \geq 1$ and $1 \geq \lambda_{\min}(Q_G) \geq 0$. In the best case, Q_G is a matrix with all entries equal to 1. Then $\lambda_{\max}(Q_G) = n$ and $\lambda_2(Q_G) = \cdots = \lambda_{\min}(Q_G) = 0$. In the worst case, when no node is connected to any other, Q_G is an identity matrix. Then $\lambda_{\max}(Q_G) = \lambda_2(Q_G) = \ldots = \lambda_{\min}(Q_G) = 1$. Consider also the following consequence of Lemma 1.

Lemma 4. *Suppose for all i, j, $a_{ij} \in \{0, 1\}$. Then there is a relabeling of vertices under which Q_G is a direct sum of matrices whose elements are all ones.*

Proof. From Lemma 1 under a reordering of vertices $Q_G = \bigoplus_i Q_{Gi}$, Q_{Gi} all positive. As all $a_{ij} \in \{0, 1\}$, there is an edge between v_i and v_j surely when $a_{ij} = 1$; or there is no edge between v_i and v_j surely when $a_{ij} = 0$. Thus either there is a path between v_i and v_j surely or there is no path between v_i and v_j surely, i.e. for all i, j, $q_{ij} \in \{0, 1\}$. Thus every element of every Q_{Gi} is 1. □

This lemma thus characterizes Q_G when $a_{ij} \in \{0, 1\}$ for all i, j, i.e. the network is effectively deterministic. In this case, there is an ordering of vertices for which Q_G is a direct sum of square matrices of all ones. If there are m such summands then $n-m$ eigenvalues of Q_G are 0. Of course, as noted above, in the extreme case where all $a_{ij} = 1$, there are $n - 1$ zero eigenvalues. *This also suggests that the proximity of $\lambda_{\min}(Q_G)$ to zero in a connected network, is a measure of connectivity, as is the number of eigenvalues that are close to zero when the network is not connected.*

Proof of Theorem 4: To prove Theorem 4 we prove in turn that (A) each Q_G is positive semidefinite (psd); (B) that should any $q_{ij} = 1$ for $i \neq j$ then Q_G cannot be positive definite (pd); and that (C) if for all $i \neq j$, $0 \leq q_{ij} < 1$, then Q_G is pd. First we recount Corollary 2.1 of [19] which exploits the facts that all convex combinations of psd matrices are psd; and that multiaffine functions are affine in each variable, if the others are fixed.

Lemma 5. *Suppose for integers n and N, $P(\alpha) \in \mathbb{R}^n$ is a multiaffine function of the elements of $\alpha = [\alpha_1, \cdots, \alpha_N]^\top$. Then $P(\alpha)$ is psd for all $\alpha_i \in [\alpha_i^-, \alpha_i^+]$ and $i \in \{1, \cdots, N\}$ iff it is psd for all $\alpha_i \in \{\alpha_i^-, \alpha_i^+\}$ and $i \in \{1, \cdots, N\}$.*

Proof of (A): As matrices of all ones are positive semidefinite, Lemma 4 proves that Q_G is psd whenever for all i, j, $a_{ij} \in \{0, 1\}$. The result follows from Lemmas 3 and 5.

Proof of (B): This follows from the following lemma and the fact that a matrix with two identical rows cannot be pd.

Lemma 6. *Suppose for some $i \neq j$, $q_{ij} = 1$. Then row i and row j of Q_G are identical, as are columns i and j.*

Proof. Note that Q_G is a symmetric matrix. Thus it suffices to show that the row property holds. One has

$$q_{ij} = q_{ji} = q_{ii} = q_{jj} = 1 \tag{17}$$

Now consider any $k \notin \{i, j\}$. Using Theorem 3 and (17): $q_{ik} \geq q_{ij}q_{jk} = q_{jk}$ and $q_{jk} \geq q_{ij}q_{ik} = q_{ik}$. Thus $q_{jk} = q_{ik}$. □

Proof of (C): Denote $N = \frac{n(n-1)}{2}$; $\mathcal{A} \in \mathbb{R}^N$ a vector whose elements are $0 \leq a_{ij} < 1$, $i > j$; $\mathcal{A}_l \in \mathbb{R}^N$ the vector whose first l elements equal the corresponding elements of $\mathcal{A}$ and the rest are zeros; $\mathcal{A}_l^+ \in \mathbb{R}^N$ the vector whose $(l + 1) - th$ element is one and the rest identical to $\mathcal{A}_l$; and $Q_G(\mathcal{A})$ the Q_G formed when the a_{ij} are the elements of $\mathcal{A}$. As $\mathcal{A}_N = \mathcal{A}$ it suffices to show that $Q_G(\mathcal{A}_l)$ is pd for all $l \in \{0, \cdots, N\}$.

Use induction on l. Note that for every l, there is an $\alpha_l \in (0, 1]$ such that $\mathcal{A}_{l+1} = \alpha_l\mathcal{A}_l+(1-\alpha_l)\mathcal{A}_l^+$. Because of Lemma 3, and the fact that only the $(l + 1)$-th element of the three vectors $\mathcal{A}_{l+1}$, $\mathcal{A}_l$ and $\mathcal{A}_l^+$ differ from each other, there holds:

$$Q_G(\mathcal{A}_{l+1}) = \alpha_l Q_G(\mathcal{A}_l) + (1 - \alpha_l)Q_G(\mathcal{A}_l^+), \quad \alpha_l \in (0, 1]. \tag{18}$$

As $\mathcal{A}_0 = 0$, $Q_G(\mathcal{A}_0) = I$ and is pd. Suppose for some $l \in \{0, \cdots, N - 1\}$, $Q_G(\mathcal{A}_l)$ is pd. From (A), $Q_G(\mathcal{A}_l^+)$ is psd. Thus (18) implies that $Q_G(\mathcal{A}_{l+1})$ is pd.

V. A DECENTRALIZED ALGORITHM FOR FINDING $\lambda_{\max}$

We now describe an algorithm for computing $\lambda_{\max}(G)$ in a decentralized fashion without having to know Q_G or even the individual link probabilities. We do require the ability to experiment by introducing packets repeatedly at nodes, and measuring how many arrive at their intended destinations. For this reason, we call the algorithm the *flooding algorithm*.

Section V-A provides a recursion and a theorem that provide the conceptual basis for the algorithm. Section V-B explains the theorem by exposing certain properties of positive matrices. Section V-C explains how the near convergence of this conceptual algorithm can be locally detected at each node. The recursion in principle requires that Q_G be known. Section V-D provides the flooding algorithm that under the temporal independence of the links, implements this algorithm in a completely decentralized fashion, without having to know Q_G.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

Section V-E discusses some practical issues and convergence rates. Section V-F has simulations. Section V-G proves a theorem in Section V-C.

A. A basic recursion.

We begin with a theorem on the conceptual recursion.

Theorem 5. *Suppose* $Q_G = Q_G^\top \in \mathbb{R}^{n\times n}$ *is positive. Consider* $z[k] = [z_1[k], \cdots, z_n[k]]^\top$ *and the recursion,*

$$z[k+1] = Q_G z[k] \tag{19}$$

with $z[0]$ *strictly positive. Then for all* $i \in \{1, \cdots, n\}$,

$$\lim_{k\to\infty} \frac{z_i[k+1]}{z_i[k]} = \lambda_{\max}(Q_G) \tag{20}$$

Thus $z[k]$ converges to a positive eigenvector of Q_G associated with its maximum eigenvalue. Further (19) induces $\frac{z_i[k+1]}{z_i[k]}$, *locally seen at each node*, to converge to $\lambda_{\max}(Q_G)$.

Many variations of this theorem appear in the literature, [26], [23], [27]. In most cases it is proved under an additional normalization, namely replacing (19) by:

$$z[k+1] \quad = \quad Q_G z[k]/\|z[k]\|. \tag{21}$$

Such a normalization militates against our eventual goal of decentralization as its implementation requires each node to know the state of all other nodes. We still omit the proof of Theorem 5. Instead we recount properties of positive matrices that explain this result and help derive an important refinement.

B. Properties of (19)

Consider the *projective metric* [27], $p(x, y)$, between two positive vectors x and y with elements x_i and y_i:

$$p(x, y) = \ln\left[\frac{\max_i \frac{x_i}{y_i}}{\min_i \frac{x_i}{y_i}}\right]. \tag{22}$$

Evidently $p(x, y) \geq 0$ with equality iff for a scalar α, $x = \alpha y$. This metric is *scale invariant*, i.e. for all positive scalar α, β

$$p(\alpha x, \beta y) = p(x, y). \tag{23}$$

For a strictly positive matrix such as Q_G there is a $0 \leq \tau < 1$ such that for all positive x, y, $p(Q_G x, Q_G y) \leq \tau p(x, y)$ [27]. *In fact* τ *is independent of* x *and* y *and depends only on* Q_G.

Call $\lambda_{\max}(Q_G)$ the Perron-Frobenius (PF) eigenvalue of Q_G and associated eigenvectors PF eigenvectors. Then for a positive Q_G, as PF eigenvectors are positive to within a scaling, with $\eta = [\eta_1, \cdots, \eta_n]^\top$ a positive PF eigenvector, using (23) in (19) one has:

$$p(z[k+1], \eta) = p(z[k+1], \lambda_{\max}(Q_G)\eta) \tag{24}$$
$$= p(Q_G z[k], Q_G \eta) \leq \tau p(z[k], \eta). \tag{25}$$

Thus as $0 \leq \tau < 1$,

$$\lim_{k\to\infty} p(z[k], \eta) = 0. \tag{26}$$

Thus, for every $\epsilon_n > 0$, there exists k_1 such that for all $k \geq k_1$,

$$0 \leq \ln\left[\frac{\max_i \frac{z_i[k]}{\eta_i}}{\min_i \frac{z_i[k]}{\eta_i}}\right] \leq \ln(1+\epsilon_n). \tag{27}$$

Then the following lemma connects (20) to (27).

Lemma 7. *Suppose the probabilistic connectivity matrix* $Q_G \in \mathbb{R}^{n\times n}$ *is symmetric and positive, and* $\eta = [\eta_1, \cdots, \eta_n]^\top$ *PF eigenvector of* Q_G *with all elements strictly positive. Consider (19) with positive* $z[0]$. *Suppose that for some* $\beta \geq 0$ *there exists a* k_0 *such that for all* $k \geq k_0$,

$$1 \leq \frac{\max_{i\in\{1,\cdots,n\}} \frac{z_i[k]}{\eta_i}}{\min_{i\in\{1,\cdots,n\}} \frac{z_i[k]}{\eta_i}} \leq 1+\beta. \tag{28}$$

Then for all $i \in \{1, \cdots, n\}$, *and* $k \geq k_0$

$$\left|\frac{z_i[k+1]}{z_i[k]} - \lambda_{\max}(Q_G)\right| \leq \beta\lambda_{\max}(Q_G).$$

Proof. As Q_G and $z[0]$ are positive so is $z[k]$. Consider any k for which (28) holds. At such a k define $\alpha = \min_i \frac{z_i[k]}{\eta_i}$. Then for all $i \in \{1, \cdots, n\}$, there holds

$$\alpha\eta_i \leq z_i[k] \leq (1+\beta)\alpha\eta_i. \tag{29}$$

Define $\xi[k] = z[k] - \alpha\eta$ and $(Q_G\xi[k])_i$ as the i-th element of $Q_G\xi[k]$. Because of (29), $\xi[k]$ is nonnegative. Thus, as Q_G is positive, $Q_G\xi[k]$ is nonnegative and for each $i \in \{1, \cdots, n\}$:

$$\begin{aligned} 0 &\leq (Q_G\xi[k])_i \\ &= (Q_G(z[k] - \alpha\eta[k]))_i \\ &\leq (Q_G((1+\beta)\alpha - \alpha))\eta[k])_i \\ &= \beta\alpha\lambda_{\max}(Q_G)\eta_i. \end{aligned} \tag{30}$$

As $z[k+1] = Q_G\xi[k] + \alpha\lambda_{\max}(Q_G)\eta$, and $\xi[k]$ is nonnegative, from (30) for all $i \in \{1, \cdots, n\}$, there thus holds:

$$\alpha\lambda_{\max}(Q_G)\eta_i \leq z_i[k+1] \leq \alpha(1+\beta)\lambda_{\max}(Q_G)\eta_i. \tag{31}$$

Hence (29) and (31) provide:

$$-\frac{\beta\lambda_{\max}(Q_G)}{1+\beta} \leq \frac{z_i[k+1]}{z_i[k]} - \lambda_{\max}(Q_G) \leq \beta\lambda_{\max}(Q_G).$$

□

Identify ϵ_n in (27) with β in Lemma 7. Then there exists a k_1 such that for all $k \geq k_1$ and all $i \in \{1, \cdots, n\}$:

$$\left|\frac{z_i[k+1]}{z_i[k]} - \lambda_{\max}(Q_G)\right| \leq \epsilon_n\lambda_{\max}(Q_G). \tag{32}$$

This constitutes an explanation if not a proof of Theorem 5. The convergence is monotonic in the sense of (25).

Remark 5. *Should* Q_G *be nonnegative as opposed to positive, the ratio for the* i*-th element of* $z[k]$ *will converge to the largest eigenvalue of the probabilistic connectivity matrix of the component to which the corresponding nodes belong.*

C. Local detection of convergence

Since the convergence in (20) is asymptotic, we now explore whether each node can detect near convergence locally. Indeed the next theorem states that should n successive ratios $\frac{z_i[k+1]}{z_i[k]}$ be close enough for *any given* i, then this ratio must be close to $\lambda_{\max}(Q_G)$ and will remain close in subsequent iterations.

Theorem 6. *Under the conditions of Theorem 5, consider, for some $c > 0$, $i \in \{1, \cdots, n\}$, $\delta > 0$ and k_0 the n inequalities:*

$$\left|\frac{z_i[k+1]}{z_i[k]} - c\right| \leq \delta, \quad \forall k \in \{k_0, k_0+1, \cdots, k_0+n-1\}. \quad (33)$$

Then for every $\epsilon > 0$, there exists a δ^* such that for all $0 < \delta \leq \delta^*$, (33) implies for all $k \geq k_0$

$$\left|\frac{z_i[k+1]}{z_i[k]} - \lambda_{\max}(Q_G)\right| \leq \epsilon. \quad (34)$$

We will prove this Theorem in Section V-G. While this theorem does permit the i-th node to conclude if its ratios are close to the postulated connectivity measure, the question remains, whether this node can also conclude that all other nodes are also close to convergence. We now argue that though this is not true in general, it is true for generic values of the probabilities a_{ij}, and hence also for generic networks.

To see this suppose for sufficiently small ϵ, (34) holds for $i = 1$. Were one to be able to conclude that this implied that $p(z[k_0], \eta)$ were small, η being a PF eigenvector of Q_G, then one can conclude that (34) would hold for all i, but possibly different, albeit small ϵ. So the issue boils down to whether (34) implies a correspondingly small $p(z[k_0], \eta)$?

Though a small $p(z[k_0], \eta)$, implies a small ϵ in (34), the reverse, is generically but not always true. For all $k \geq k_0$,

$$z_1[k] = e_1^\top Q_G^{k-k_0} z[k_0], \quad (35)$$

where $e_1 = [1, 0, \cdots, 0]^\top$. Should the pair $[Q_G, e_1^\top]$ be *completely observable (c.o.)*, [28], i.e.

$$W = \left[e_1^\top, e_1^\top Q_G, \cdots, e_1^\top Q_G^{n-1}\right]^\top \quad (36)$$

be nonsingular then the $z[k_0]$ leading to the n-successive samples in (35) is unique. In such a case a small ϵ in (34), with $i = 1$, forces a small $p(z[k_0], \eta)$. *Consequently, each node can detect near convergence of the ratios at all other nodes, from the near convergence of its own ratios.*

For every, $n > 2$, we now provide example networks, that (a) for a particular choice of the probabilities a_{ij} yield a Q_G for which $[Q_G, e_1^\top]$ is not c.o.; and (b) for a particular choice of the probabilities a_{ij} yield a Q_G for which $[Q_G, e_1^\top]$ is c.o..

In particular (a) shows that there are networks for which a single node cannot conclude that the near convergence of its ratios implies that other nodes are near convergence. What is more important from a practical point of view is (b), that shows that almost all choices of a_{ij} yield networks for which near convergence at one node implies near convergence at all. This is so as Q_G and hence W in (36) is polynomial in the a_{ij}. Thus, either W is singular for *all values of* a_{ij} or it is nonsingular for *generic values*. The network is as follows.

Example 3. *For $n > 2$, choose the $a_{ij} = a_{ji}$ as follows. For some $1 \geq r_i > 0$ and $i \in \{1, \cdots, n-1\}$ there holds: $a_{1,i+1} = r_i$. For all $i \in \{2, \cdots, n-1\}$ and $j \in \{i+1, \cdots, n\}$, $a_{ij} = 0$. Under the independence assumption,*

$$q_{ij} = \begin{cases} r_{j-1} & i = 1 \text{ and } j \in \{2, \cdots, n\} \\ r_i r_j & i \in \{2, \cdots, n-1\} \text{ and } j \in \{i+1, \cdots, n\} \end{cases}. \quad (37)$$

Thus, e.g. for $n = 4$ one has

$$Q = \begin{bmatrix} 1 & r_1 & r_2 & r_3 \\ r_1 & 1 & r_1 r_2 & r_1 r_3 \\ r_2 & r_1 r_2 & 1 & r_2 r_3 \\ r_3 & r_1 r_3 & r_2 r_3 & 1 \end{bmatrix}$$

The next Lemma proves both (a) and (b) above.

Lemma 8. *For $n > 2$, consider under $0 < r_i < 1$, the symmetric probabilistic connectivity matrix with diagonal elements $q_{ii} = 1$ and the remaining elements as in (37). Then with $e_1 = [1, 0, \cdots, 0]^\top$, the pair $[Q_G, e_1^\top]$ is completely observable iff the r_i are all distinct.*

Proof. By the Popov-Belevitch-Hautus (PBH) test, [28], $[Q_G, e_1^\top]$ is a c.o. pair iff for all scalar complex λ:

$$\textbf{rank}\ \left(\begin{bmatrix} e_1 & \lambda I - Q_G \end{bmatrix}\right) = n. \quad (38)$$

With $r = [r_1, \cdots, r_{n-1}]^\top$ and $R =$ **diag** $\{r_i^2\}_{i=1}^{n-1}$, (37) is

$$Q_G = \begin{bmatrix} 1 & r^\top \\ r & I - R + rr^\top \end{bmatrix}. \quad (39)$$

Suppose the r_i are distinct, but to establish a contradiction, $[Q_G, e_1^\top]$ is not c.o., i.e. (38) is violated. Then there exists a scalar complex λ and nonzero $f \in \mathbb{R}^{(n-1)}$ such that

$$r^\top f = 0 \quad (40)$$

and $\left((\lambda - 1)I + R - rr^\top\right) f = 0$; i.e.

$$((\lambda - 1)I + R) f = 0. \quad (41)$$

As $r_i > 0$, $\forall i$, from (40) at least two elements of f, without loss of generality f_1 and f_2, are non-zero. Thus (41) yields

$$\lambda = 1 - r_i^2 \quad \forall i \in \{1, 2\} \quad (42)$$

which is impossible as $r_1^2 \neq r_2^2$, establishing a contradiction.

Now suppose at least two elements of r, without loss of generality, r_1 and r_2, are equal. Choose $f = [0, 1, -1, 0, \cdots, 0]^\top$, and the scalar λ as in (42). Then clearly $e_1^\top f = 0$. Further,

$$(\lambda I - Q_G) f = \begin{bmatrix} 1 & r_1 & r_2 & 0_{n-3}^\top \end{bmatrix}^\top (r_1 - r_2) = 0,$$

where 0_{n-3} is the zero vector in $\mathbb{R}^{n-3}$ (empty if $n = 3$). Thus (38) is violated and $[Q_G, e_1^\top]$ is not c.o..

□

Note that for $n = 2$, $[Q_G, e_1^\top]$ is c.o. iff $q_{12} \neq 0$. We have effectively shown that for almost all networks, local detection of near convergence implies near convergence of all nodes.

D. The flooding algorithm

Observe, (19) requires that the i-th node knows all the q_{ij} as well as all elements of $z[k]$. We now provide an algorithm that sidesteps this need and can be used in our probabilistic network setting provided the transmissions at different time slots are i.i.d. Formally, we make the following assumption.

Assumption 1. *The indicator random variables I_i defined before Definition 1 are i.i.d. across transmission slots.*

This assumption permits us to postulate a flooding algorithm that asymptotically approximates (19) in a totally decentralized

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

fashion. Suppose, for some k, $z[k]$ has been obtained, and in a series of simultaneous experiments the i-th node *floods* the network with $x_i = z_i[k]$ number of packets. Then the expected number of packets received by this node is the i-th entry of $Q_G z[k]$. Now suppose for some K, each node repeats this flooding operation K times. Denote by $z[k,m], i \in \{1,...,n\}, m \in \{1,...,K\}$, the number of packets received by node v_i in the m-th repetition. Then, because of Assumption 1, by the law of large numbers, for sufficiently large K:

$$z[k+1] \approx \frac{1}{K}\sum_{m=1}^{K} z[k,m]. \tag{43}$$

There are clearly two approximations inherent in (43). First, implicitly for noninteger $z[k]$, we quantize to the nearest vector of integers. Secondly (43) represents a better approximation as K grows. We comment on the size of K in Section V-E.

Accordingly, the flooding algorithm we postulate is as follows: For some K, $l = 0$ and positive vector $y[l,K]$, let the i-th node flood the network with $y_i[l,K]$ number of packets. Every node repeats this experiment K times in the l-th iteration. The number of packets transmitted by the i-th node in the $(l+1)$-th iteration is the number of packets averaged over K transmissions, received by it in the l-th iteration. Then

$$\lim_{l\to\infty}\left\{\lim_{K\to\infty}\frac{y_i[l+1,K]}{y_i[l,K]}\right\} = \lambda_{\max}(Q_G). \tag{44}$$

In principle, the number of packets from a node increases by a factor approximately equal to $\lambda_{\max}(Q_G)$ in each iteration of (19). In a large network, this leaves open the risk that after a modest number of iterations, the number of packets becomes very large. As explained in Section V-E, this may require larger values of K for the approximation in (43) to be sufficiently good. The implementation of (21), rather than just (19) would avoid this difficulty. However, the normalization by $\|z[k]\|$ in (21), does not permit a decentralized implementation. Instead we propose an *optional* renormalization to combat this challenge. Specifically, should the $y_i[l,K]$ exceed a pre-specified threshold at a particular node i, then this node must divide the number of packets it transmits by a pre-specified factor. It can then piggyback this scaling information in every packet it transmits, so that all the other nodes are alerted of this scaling, and scale the number of packets they transmit by the *the same factor*. If the pre-designated threshold is chosen to be sufficiently large, the chance of missing this scaling information is negligible. As only the convergence of ratios are at issue, there is no resulting impact on convergence speed to speak of. As argued later, this option is rarely needed.

Despite quantization, and approximate averaging, simulations in Section V-F show that relatively small l and K, suffice for the ratios $\frac{y_i[l+1,K]}{y_i[l,K]}, i \in \{1,...,n\}$, to converge to a value that is very close to $\lambda_{\max}(Q_G)$.

E. Practical issues and convergence rates

To avoid the effect of network delays, packets must be accumulated over large intervals. The convergence speed of (19) is measured by $\lambda_2(Q_G)/\lambda_{max}(Q_G)$, where $\lambda_2(Q_G)$ is the second largest eigenvalue of Q_G. *Inter alia*, this suggests faster convergence in highly connected networks. To see why, observe that as Q_G is positive semidefinite, and its trace is always n, $\lambda_2(Q_G)$ is upper bounded by $n - \lambda_{max}(Q_G)$. Thus $\lambda_{\max}(Q_G)$ lower bounds the convergence rate.

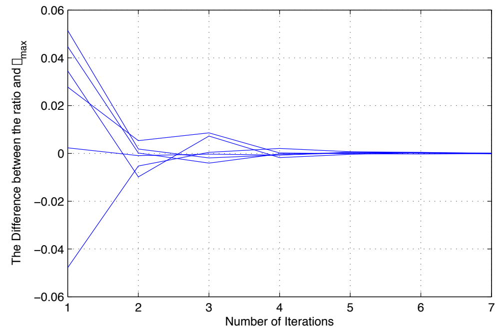

Fig. 3. An illustration of the convergence of the ratio $\frac{y_i[l+1]}{y_i[l]}$ to $\lambda_{\max}(Q_G)$. The simulation result is obtained from a random network with six nodes. a_{ij}s, $1 \le i < j \le 6$, are drawn uniformly, randomly and independently from $[0,1]$. K is chosen to be 10. The horizontal axis is the number of iterations and the vertical axis is the difference between $\frac{y_i[l+1]}{y_i[l]}$ and $\lambda_{\max}(Q_G)$. Since there are six nodes, six curves are shown in the figure corresponding to the value of $\frac{y_i[l+1]}{y_i[l]}$ for each of the six nodes.

The slowest part of the convergence is determined by the law of large numbers. In fact K is proportional to the variance of the i.i.d. variables being averaged. As Q_G is positive semidefinite and has trace n, $\lambda_{\max}(Q_G) \ge 1$. Thus, in (19) $z_i[k]$ is potentially unbounded *though ratios of successive values is not.* Nonetheless the flooding algorithm does not estimate these ratios directly, but rather estimates the $z_i[k]$.

Just as the $z_i[k]$, $y_i[l,K]$ grow in size with l. Larger they are, the larger their initial variance. This in turn correspondingly increases the required K, thus slowing convergence. This underscores the importance of the renormalization proposed in Section V-D, and used in the simulations. There are other mechanisms of renormalization one may invoke. For example, for some predetermined integer m all nodes scale down $y_i[l,K]$ by a factor C whenever l is a multiple of m.

Actually, in practice renormalization is rarely needed. As shown in the simulations in Section V-F, in networks with even moderate connectivity, convergence is so rapid that it can be detected well before packet growth becomes unmanageable. In networks with low connectivity, $\lambda_{\max}(Q_G)$ is relatively small, and larger number of iterations can be sustained before packet growth becomes so large as to require normalization.

F. Simulations

The simulation shown in Fig. 3 and Fig. 4 involves six nodes, and $K = 10$. Within just seven iterations, the ratio (44) converges to within half a percent of the true $\lambda_{\max}(Q_G)$.

Fig. 5, considers a network with 50 nodes where a_{ij}s, $1 \le i < j \le 50$, are drawn uniformly from $[0,P]$. Varying P, which controls network connectivity, illustrates the effect of connectivity to convergence speed. Note that when the number of nodes equals 50, the number of edges equals 1225. It becomes computationally prohibitive to compute Q_G

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

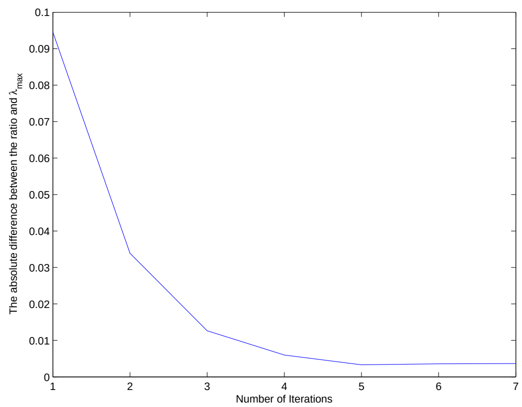

Fig. 4. A further illustration of the convergence of the ratio $\frac{y_i[l+1]}{y_i[l]}$ to $\lambda_{\max}(Q_G)$. The simulation result is obtained from a random network with six nodes. a_{ij}s, $1 \le i < j \le 6$, are drawn uniformly, randomly and independently from $[0,1]$. K is chosen to be 10. The horizontal axis is the number of iterations and the vertical axis is the average absolute difference between $\frac{y_i[l+1]}{y_i[l]}$ and $\lambda_{\max}(Q_G)$, i.e. $|\frac{y_i[l+1]}{y_i[l]} - \lambda_{\max}(Q_G)|$ averaged over six nodes. Further the simulation is repeated 50 times and each point in the curve corresponds to the average value over 50 simulations.

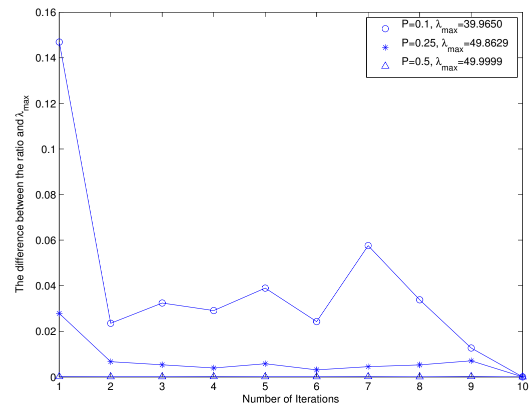

Fig. 5. A further illustration of the convergence of the ratio $\frac{y_i[l+1]}{y_i[l]}$ to $\lambda_{\max}(Q_G)$. The simulation result is obtained from a random network with 50 nodes. a_{ij}s, $1 \le i < j \le 50$, are drawn uniformly, randomly and independently from $[0,P]$. K is chosen to be 10. The horizontal axis is the number of iterations and the vertical axis is the average absolute difference between $\frac{y_i[l+1]}{y_i[l]}$ and $\lambda_{\max}(Q_G)$, i.e. $|\frac{y_i[l+1]}{y_i[l]} - \lambda_{\max}(Q_G)|$ averaged over 50 nodes. Further the simulation is repeated 10 times and each point in the curve corresponds to the average value over 10 simulations. As P increases above 0.5, the ratio converges to the true value of $\lambda_{\max}(Q_G) = 50$ immediately in the first iteration.

and $\lambda_{\max}(Q_G)$ whose computational complexity increases approximately with the number of edges according to $2^{|E|}$ with $|E|$ being the number of edges. Therefore in the figure we use $\frac{y_i[10]}{y_i[9]}$ averaged over 50 nodes as an approximation of $\lambda_{\max}(Q_G)$. Further, as explained in Section V-D, to make the algorithm more efficient, whenever the number of packets flooded by a node in an iteration exceeds 5000, the number of packets flooded by all nodes in the next iteration is divided by a common factor equal to the number of nodes.

A feature of note is that foreshadowed at the end of Section V-E. Observe in Figure 5, that even with $P = 0.5$, representing a network of moderate connectivity, convergence is virtually immediate. When $P > 0.5$, this convergence occurs by $l = 1$, obviating the need for renormalization.

G. Proof of Theorem 6

We conclude this Section by proving Theorem 6 which requires the following lemma.

Lemma 9. *Suppose $F = F^\top \in \mathbb{R}^{n \times n}$ is positive and $h \in \mathbb{R}^n$ is nonnegative. Suppose also that there exists a $\psi \in \mathbb{R}^n$ such that:*

$$\left[h^\top, h^\top F, \cdots, h^\top F^{n-1}\right]^\top \psi = 0 \tag{45}$$

Consider any eigenvector ω_i of F, other than the PF eigenvector, and a nonzero $\gamma \in \mathbb{R}^n$ that is given by $\alpha\psi + \beta\omega_i$ for some constants α, β. Then γ must have at least one element negative and another positive.

Proof. As $F = F^\top$, its eigenvalues are real and *the eigenvectors can be chosen to form an orthonormal basis*. Suppose $\lambda_1 > \lambda_2 \ge \cdots \ge \lambda_n$, where the strictness of the first inequality is a consequence of F being positive. Suppose ω_i is a unit norm eigenvector corresponding to λ_i, with at least one element positive. From the PF Theorem ω_1 is positive.

Suppose γ is a linear combination of ψ with some ω_i, $i \in \{2, \cdots, n\}$. To establish a contradiction suppose all elements of $\gamma \ne 0$ are nonnegative. Define the orthogonal matrix $U = \begin{bmatrix}\omega_1 & \Omega\end{bmatrix}$ with $\Omega = \begin{bmatrix}\omega_2, & \cdots, & \omega_n\end{bmatrix}$. Observe: $\psi = UU^\top\psi = \sum_{i=1}^n \omega_i \left(U^\top\psi\right)_i$, where $\left(U^\top\psi\right)_i$ denotes the i-th element of $U^\top\psi$. Now consider two cases.

Case I $\left(U^\top\psi\right)_1 = 0$: Then ψ is in the range space of Ω. Then as γ is a linear combination of ψ and a column of Ω, γ is in the range space of Ω as well. Now as every column of Ω is orthogonal to ω_1, so must be γ. Then as ω_1 is positive, γ cannot be nonnegative and nonzero, establishing a contradiction.

Case II $\left(U^\top\psi\right)_1 \ne 0$: Observe that $F = U\Lambda U^\top$, with $\Lambda =$ **diag** $\{\lambda_1, \cdots, \lambda_n\}$. Thus, [28], (45) implies for all t:

$$0 = h^\top e^{Ft}\psi = h^\top U e^{\Lambda t} U^\top \psi = \sum_{i=1}^N \left(h^\top U\right)_i \left(U^\top\psi\right)_i e^{\lambda_i t}$$

As $\lambda_1 \ne \lambda_i$, for all $i \in \{2, \cdots, n\}$, this in particular implies that $\left(h^\top U\right)_1 \left(U^\top\psi\right)_1 = 0$, i.e. $0 = \left(h^\top U\right)_1 = h^\top\omega_1$. As $h \ne 0$ is nonnegative and ω_1 is positive, this cannot be true. □

We now prove Theorem 6 by showing in turn the following: For small enough δ, (A) c in (33) is close to an eigenvalue of Q_G; (B) that this is $\lambda_{\max}(Q_G)$; and (C) that subsequent ratios $\frac{z_i[k+1]}{z_i[k]}$ remain close to $\lambda_{\max}(Q_G)$.

Proof of (A): With e_i a vector with i-th element 1 and rest 0,

$$z_i[k] = e_i^\top Q_G^k z[0] \;\; \forall k \ge 0. \tag{46}$$

Because of (33), there exist $|\delta_i| < \delta$, such that for all $k \in \{k_0, \cdots, k_0 + n\}$, there holds:

$$z_i[k] = \left\{\prod_{j=k_0}^{k} (c + \delta_j)\right\} z_i[k_0] \tag{47}$$

Suppose the characteristic polynomial of Q_G is given by: $\det(\lambda I - Q_G) = \lambda^n - \sum_{i=0}^{n-1} \alpha_i \lambda^i$. Then

$$Q_G^n = \sum_{i=0}^{n-1} \alpha_i Q_G^i. \tag{48}$$

From (47), (46) and (48), there obtains,

$$\begin{aligned}
&\left\{\prod_{j=k_0}^{n} (c+\delta_j)\right\} z_i[k_0] = z_i[k_0+n] \\
=\ & e_i^\top Q_G^{k_0+n} z[0] = e_i^\top \left(\sum_{l=0}^{n-1} \alpha_l Q_G^l\right) Q_G^{k_0} z[0] \\
=\ & \sum_{l=0}^{n-1} \alpha_l e_i^\top Q_G^{k_0+l} z[0] = \sum_{l=0}^{n-1} \alpha_l z_i[l+k_0] \\
=\ & \left(\sum_{l=0}^{n-1} \alpha_l \left\{\prod_{j=k_0}^{k_0+l} (c+\delta_j)\right\}\right) z_i[k_0].
\end{aligned}$$

A positive $z[0]$ implies $z[k]$ is positive for all $k > 0$. Thus:

$$\left\{\prod_{i=k_0}^{k_0+n} (c+\delta_i)\right\} = \left(\sum_{l=0}^{n-1} \alpha_l \left\{\prod_{i=k_0}^{k_0+l} (c+\delta_i)\right\}\right). \tag{49}$$

As the roots of a monic polynomial vary continuously with its coefficients, with $\lambda_{\max}(Q_G) = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$ the eigenvalues of Q_G, for every $\epsilon > 0$ there exists a δ^* such that for all $0 < \delta \leq \delta^*$ under (33)

$$c \in \bigcup_{i=1}^{n} [\lambda_i - \epsilon, \lambda_i + \epsilon] \tag{50}$$

Proof of (B): We will now show that in fact for every $\epsilon > 0$ there exists a δ^* such that for all $0 < \delta \leq \delta^*$ under (33) $c \in [\lambda_1 - \epsilon, \lambda_1 + \epsilon]$.

Suppose instead that for some $l \in \{2, \cdots, n\}$, $c \in [\lambda_l - \epsilon, \lambda_l + \epsilon]$. As (46) holds for all $k \in \{k_0, \cdots, k_0+n-1\}$, under (36) we have $\begin{bmatrix} z_i[k_0] & \cdots & z_i[k_0+n-1] \end{bmatrix} = z^\top[k_0] W^\top$. Suppose χ is an eigenvector of Q_G corresponding to λ_l, and χ_i is its i-th element. Then: $\begin{bmatrix} \chi_i & \cdots & \lambda_l^{k_0+n-1} \chi_i \end{bmatrix} = \chi^\top W^\top$. Then a standard continuity argument shows that for every ϵ there exists a δ^* such that for all $0 < \delta \leq \delta^*$ under (33)

$$z[k_0] = \psi + \chi + e, \ \ \|e\| \leq \epsilon, \text{ and } W\psi = 0. \tag{51}$$

As $z[0]$ is positive, so is $z[k_0]$. Yet, because of Lemma 9, $\psi + \chi$ has at least one negative element. Thus, because of (51) for sufficiently small ϵ, $z[k_0]$ has at least one negative element.

Proof of (C): Thus with η a PF eigenvector of Q_G, and ψ obeying (51), for every ϵ there exists a δ^* such that for all $0 < \delta \leq \delta^*$ under (33), (51) holds. Thus

$$e_i^\top Q_G^m \psi = 0; \ \ \forall\, m. \tag{52}$$

Now consider the alternative recursion: $s[k+1] = Q_G s[k]; \ \ s[k_0] = \eta + e$. Because of (52) for all $k \geq k_0$,

$$z_i[k] = s_i[k]. \tag{53}$$

Further, as η is a PF eigenvector, for every ϵ there exists a δ^* such that for all $0 < \delta \leq \delta^*$ under (33) $p(s[k_0], \eta) \leq \ln(1+\epsilon)$. Consequently from (25) and Lemma 7 for every ϵ there exists a δ^* such that for all $0 < \delta \leq \delta^*$ under (33) the following holds for all $j \in \{1, \cdots, n\}$ and $k \geq k_0$: $\left|\frac{s_j[k+1]}{s_j[k]} - \lambda_{\max}(Q_G)\right| \leq \epsilon \lambda_{\max}(Q_G)$. The result follows as this also holds for $j = i$, $\lambda_{\max}(Q_G)$ is finite and (53).

VI. Conclusions and Further Work

We have considered the probabilistic connectivity matrix Q_G as a tool to measure the quality of network connectivity. Key properties of this matrix and their relation to the quality of network connectivity have been demonstrated. In particular, the off-diagonal entries of the probabilistic connectivity matrix provide a measure of the quality of end-to-end connections. We have provided theoretical analysis supporting the use of the largest eigenvalue of Q_G as a measure of the quality of overall network connectivity. Our analysis compares networks with the same number of nodes. For networks with different number of nodes, the largest eigenvalue of Q_G, normalized by the number of nodes may be used as the quality metric. A flooding algorithm is presented for experimentally estimating the largest eigenvalue in a decentralized fashion, without knowledge of the individual link probabilities or the network topology.

Inequalities between the entries of the probabilistic connectivity matrix have been established. These may provide insights into the correlations between quality of end-to-end connections. We have also shown that Q_G is positive semidefinite and its off-diagonal entries are multiaffine functions of link probabilities. These two properties should facilitate optimization and robust network design, e.g. determining the link that maximally impacts network quality, and determining quantitatively the relative criticality of a link to either a particular end-to-end connection or to the entire network.

We assume that the links are symmetric *and* independent. We expect that our analysis can be extended with nontrivial work to the case where the assumption on symmetric links is removed. We conjecture that the largest singular value, as opposed to the largest eigenvalue of Q_G is a more appropriate measure of connectivity. Relaxing the independence assumption requires more work. Yet, we are encouraged by the fact that the elements of Q_G, being probabilities of union of edge events, are multiaffine functions of the a_{ij} and the conditional link probabilities, as $P(A \bigcup B) = P(A) + P(B) - P(B|A)P(A)$. Thus we still expect all the results in Section IV to hold, though the proof may be non-trivial. In real applications link correlations may arise due to both physical layer correlations and correlations caused by traffic congestion.

Another implicit assumption in the paper is that traffic is uniformly distributed and traffic between every source-destination pair is equally important. If this is not the case, a weighted version of the probabilistic connectivity matrix can be contemplated. Whether our results can be extended to a weighted probabilistic connectivity matrix is an open issue.

References

[1] P. Gupta and P. R. Kumar, “Critical power for asymptotic connectivity in wireless networks,” in *Stochastic Analysis, Control, Optimization and Applications*. Boston, MA: Birkhauser, 1998, pp. 547–566.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2014.2366106, IEEE Transactions on Mobile Computing

[2] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, 2009.

[3] M. D. Penrose, *Random Geometric Graphs*, ser. Oxford Studies in Probability. Oxford University Press, USA, 2003.

[4] P. J. Wan and C. W. Yi, "Asympototic critical transmission radius and critical neighbor number for k-0connectivity in wireless ad hoc networks", *MobiHoc '04*, pp. 1-8, May 2004, Roppongi, Japan.

[5] G. Mao and B. D. O. Anderson, "Towards a Better Understanding of Large Scale Network Models," *IEEE/ACM Transactions on Networking*, vol. 20, no. 2, pp. 408 - 421, 2012.

[6] ——, "Connectivity of Large Wireless Networks under a General Connection Model," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1761 - 1772, 2013.

[7] R. Meester and R. Roy, *Continuum Percolation*, ser. Cambridge Tracts in Mathematics. Cambridge University Press, 1996.

[8] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215-233, 2007.

[9] L. Hogben, *Handbook of Linear Algebra*, ser. Discrete Mathematics and Its Applications. Chapman & Hall/CRC, 2007.

[10] N. L. Biggs, *Algebraic Graph Theory*. Cambridge University Press, 1974.

[11] B. Mohar, "Laplace eigenvalues of graphs - a survey," *Journal of Discrete Mathematics*, vol. 109, pp. 171–183, 1992.

[12] C. Godsil and G. F. Royle, *Algebraic Graph Theory*. Springer, 2001.

[13] R. R. Brooks, B. Pillai, S. Racunas, and S. A. R. S. Rai, "Mobile network analysis using probabilistic connectivity matrices," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 4, pp. 694–702, 2007.

[14] G. Mao and B. D. O. Anderson, "Graph theoretic models and tools for the analysis of dynamic wireless multihop networks," in *IEEE WCNC*, 2009, pp. 1–6.

[15] J. v. d. Berg and H. Kesten, "Inequalities with applications to percolation and reliability," *Journal of Applied Probabilities*, vol. 22, pp. 556–569, 1985.

[16] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, ser. Computer Science and Applied Mathematics. Academic Press, 1985.

[17] R. A. Horn and C. R. Johnson,*Matrix Analysis*, Cambridge University Press, 1985.

[18] L. Zadeh and C. A. Desoer, *Linear Systems Theory*. McGraw Hill, New York, 1963.

[19] S. Dasgupta, G. Chockalingam, M. Fu, and B. D. O. Anderson, "Lyapunov functions for uncertain systems with applications to the stability of time varying systems," *IEEE Transactions on Circuits and Systems-I, Fundamental Theory and Applications*, vol. 41, no. 2, pp. 93–105, 1994.

[20] S. Dasgupta and B. D. O. Anderson, "Physically based parameterizations for designing adaptive algorithms", *Automatica*, pp. 469-477, 1987.

[21] S. Dasgupta, B. D. O. Anderson, and R. J. Kaye, "Identification of physical parameters in structured systems", *Automatica*, pp. 217-225, 1988.

[22] S. Dasgupta, B. D. O. Anderson, and R. J. Kaye, "Output error identification methods for partially known systems", *International Journal of Control*, pp. 177-191, 1985.

[23] R. A. Varga, *Matrix Iterative Analysis,* pp. 34, Prentice Hall, Englewood Cliffs, N.J., 1962.

[24] F. Gantmacher, *The Theory of Matrices, Volume 2*, AMS Chelsea Publishing, 2000.

[25] B. D. O. Anderson, F. Kraus, M. Mansour and S. Dasgupta, "Easily testable sufficient conditions for the robust stability of systems with multilinear parameter dependence", *Automatica*, pp. 25-40, 1995.

[26] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins Studies in Mathematical Sciences, 3rd Ed., 1996.

[27] D. J. Hartfiel, *Nonhomogeneous Matrix Products*, World Scientific, 2002.

[28] C.T. Chen, *Linear System Theory and Design*, 4th Ed. 2013.

Soura Dasgupta received the B.E. degree in Electrical Engineering from the University of Queensland (Australia), in 1980, and the Ph.D. in Systems Engineering from the Australian National University in 1985. He is currently Professor of Electrical and Computer Engineering at the University of Iowa, U.S.A.

He has held visiting appointments at the University of Notre Dame, University of Iowa, Université Catholique de Louvain-La-Neuve, Belgium and the Australian National University.

Between 1988 and 1991, 1998 to 2009 and 2004 and 2007 he respectively served as an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE Control Systems Society Conference Editorial Board, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-II. He is a corecipient of the Gullimen Cauer Award for the best paper published in the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS in the calendar years of 1990 and 1991, a past Presidential Faculty Fellow, a subject editor for the International Journal of Adaptive Control and Signal Processing, and a member of the editorial board of the EURASIP Journal of Wireless Communications. He was a recipient of the University Iowa Collegiate Teaching award in 2012. In the same year he was selected by the graduating class for an award on excellence in teaching and commitment to student success.

His research interests are in Controls, Signal Processing and Communications. He was elected as a Fellow of the IEEE in 1998.

Guoqiang Mao received PhD in telecommunications engineering in 2002 from Edith Cowan University. Between 2002 and 2013, he was an Associate Professor at the School of Electrical and Information Engineering, the University of Sydney. He currently holds the position of Professor of Wireless Networking, Director of Center for Real-time Information Networks at the University of Technology, Sydney. He has published more than 100 papers in international conferences and journals, which have been cited more than 2000 times. His research interest includes intelligent transport systems, applied graph theory and its applications in networking, wireless multihop networks, wireless localization techniques and network performance analysis.

Brian D. O Anderson (M'66-SM'74-F'75-LF'07) was born in Sydney, Australia, and educated at Sydney University in mathematics and electrical engineering, with PhD in electrical engineering from Stanford University in 1966. He is a Distinguished Professor at the Australian National University and Distinguished Researcher in National ICT Australia. His awards include the IEEE Control Systems Award of 1997, the 2001 IEEE James H Mulligan, Jr Education Medal, and the Bode Prize of the IEEE Control System Society in 1992, as well as several IEEE and other best paper prizes. He is a Fellow of the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, the Royal Society, and a foreign associate of the US National Academy of Engineering. He holds honorary doctorates from a number of universities, including Catholic University of Louvain, Belgium, and Swiss Federal Institute of Technology Zurich. He is a past president of the International Federation of Automatic Control and the Australian Academy of Science. His current research interests are in distributed control, sensor networks and econometric modelling.