

# Random Walks on Digraphs: A Theoretical Framework for Estimating Transmission Costs in Wireless Routing

Yanhua Li and Zhi-Li Zhang

Department of Computer Science & Engineering, University of Minnesota

Email: {yanhua,zhzhang}@cs.umn.edu

**Abstract**—In this paper we develop a unified theoretical framework for estimating various transmission costs of packet forwarding in wireless networks. Our framework can be applied to the three routing paradigms—*best path routing*, *opportunistic routing*, and *stateless routing*—to which nearly all existing routing protocols belong. We illustrate how packet forwarding under each paradigm can be modeled as *random walks on directed graphs (digraphs)*. By generalizing the theory of random walks that has primarily been developed for undirected graphs to digraphs, we show how various transmission costs can be formulated in terms of hitting times and hitting costs of random walks on digraphs. As representative examples, we apply the theory to three specific routing protocols, one under each paradigm. Extensive simulations demonstrate that the proposed digraph based analytical model can achieve more accurate transmission cost estimation over existing methods.

**Index Terms**—Transmission cost, Digraph, Wireless routing, Random Walk, Spectral graph theory

## I. INTRODUCTION

Due to the unique characteristics of wireless technologies and the dynamics in the environments (e.g., mobility and interference) they operate in, wireless channels are known to be time-varying, unreliable, and asymmetric. Furthermore, wireless networks are often designed to support certain applications or missions, and deployed in specific environments. For these reasons, a plethora of wireless mechanisms—especially, routing algorithms and protocols—have been proposed and developed to achieve a range of different objectives such as throughput, latency, energy consumption, network life time, and so forth. Evaluating the efficacy of wireless protocols in terms of various transmission cost metrics, and deciding on which one to employ in a specific environment so as to attain certain performance objective, can be a challenging task in practice. The ability to analyze, estimate and quantify various transmission costs is therefore imperative in the design of wireless networks.

While experimentation and testing in realistic wireless environments are indispensable and provide the most definite and authoritative means to evaluate the efficacy of wireless routing protocols, they are in general very expensive and are typically utilized in the later stage of the network design and evaluation process. Simulation-based evaluation is also important and necessary; however, conducting realistic simulations

is hard, and simulation results often hinge on the settings and parameters used. We believe that analytical models and theories also play a critical role in the design of wireless networks, complementing the roles played by real-world experimentation and simulations. By generating performance bounds and theoretical limits, they provide important insights on what is achievable and under what conditions, and produce useful metrics for understanding the key design trade-offs. Such insights and understanding are particularly important in the early stage of wireless network design.

Guided by this belief, in this paper we develop an unified theoretical framework to quantify and estimate various transmission costs of wireless routing protocols. To account for the *stochastic* and *asymmetric* natures of wireless channels, we model a wireless network as a *directed graph* (in short, *digraph*), where each directed edge (link) is associated with a *packet delivery probability*. We consider three wireless routing paradigms, the (traditional) best path routing (e.g., AODV [21], DSR [14], and several energy-aware routing protocols [2], [5]), opportunistic routing (e.g., ExOR [3], MORE [4]), and stateless (stochastic) routing (e.g., as proposed in [7], [19])—nearly all existing routing protocols fall under one of these paradigms, or use a combination thereof. Under the (simplifying) assumption that packet delivery probabilities are independent, we demonstrate how packet forwarding under each paradigm can be modeled as a Markov chain on a digraph with an appropriately defined transition probability matrix capturing the specifics of the routing algorithm under consideration. In other words, the traversal of a packet being forwarded in a wireless network can be viewed as a *random walk on a digraph*. Consequently, various transmission costs of end-to-end packet delivery (e.g., the expected number of transmissions, end-to-end packet delivery ratio, throughput, latency, energy consumptions) can therefore be formulated using well-known notions such as *hitting times*, *sojourn times* associated with random walks.

The main contributions of this paper are summarized below.

- To our best knowledge, this is the first work that utilizes the random walk (Markov chain) model to formulate the end-to-end transmission costs for various types of wireless routing strategies.
- The theory of random walks (and the closely related *spectral graph theory* have been developed primarily for *undirected graphs* (see, e.g., [8], [18]). We successfully extend the theory of random walks on undirected graphs to *directed graphs*

This work is supported in part by the US National Science Foundation grant CNS-0626808, the DTRA HDTRA1-09-1-0050, and University of Minnesota DTC DTI grant.

(digraphs), with a more general definition of normalized (graph) Laplacian matrix  $\tilde{\mathcal{L}}_G$ . We also show how the hitting times, commute times and sojourn times (or hitting costs) can be computed using the Penrose-Moore pseudo-inverse of  $\tilde{\mathcal{L}}_G$ .

- Using three representative routing protocols as examples (one from each routing paradigm), we systematically illustrate how our proposed theoretical framework based on random walks on digraphs can be used to estimate various transmission costs. Our analysis subsumes earlier results obtained using more ad hoc methods. We also perform extensive simulations to show the relative errors in estimation when asymmetric links are artificially symmetrized and undirected graphs are used.

The remainder of this paper is organized as follows. The related works is briefly touched on below. In Section II, we first describe three wireless routing paradigms. We then illustrate how packet forwarding under each of them can be modeled using Markov chains/random walks, and various transmission costs can be formulated using hitting times and hitting costs associated with the random walks. We outline the theories of random walks on undirected and its generalization to directed graphs in Section III. In Section IV, we apply the theory of random walks on digraphs to three representative routing protocols. Simulation results are reported in Section V, and the paper is concluded in Section VI.

**Related Works.** We will present the three wireless routing paradigms in Section II and discuss some related wireless protocols in that context. Hence we do not touch on them here. Our paper is partly inspired by the work in [7] where results from random walks on *undirected graphs* are used to model and derive a delay estimation formula for stateless routing with heterogenous sojourn time. In contrast, our paper not only supersedes the results in [7] which essentially assume symmetric wireless links, but develops a general theoretical frame based on random walks on digraphs for estimating various transmission costs under all three wireless routing paradigms. As mentioned earlier, the theory of random walks has been developed primarily for undirected graphs. Relatively fewer attempts have been made to extend it to digraphs. In [9], Chung defines a *symmetrized* Laplacian matrix for directed graph, and successfully generalizes the well-known Cheeger Inequality to directed graphs. However, it is unclear whether this generalization can be used to compute hitting times and commute times for random walks on digraphs.

## II. WIRELESS ROUTING, TRANSMISSION COSTS AND RANDOM WALKS IN DIRECTED GRAPHS

In this section we briefly describe the three wireless routing paradigms. We then show how packet traversals under each routing paradigm can be modeled using Markov chains, and use the Markov models to estimate various transmission costs in a wireless network.

### A. Wireless Routing and Transition Costs

The existing (unicast) wireless routing schemes can be roughly classified into three categories: the traditional *best*

*path* routing, *opportunistic* routing, and *stateless (stochastic)* routing.

The traditional best path routing protocols (e.g., AODV [21], DSR [14] and their variations/extensions to multi-path or energy-aware routing) typically select a single *best* path, sometimes multiple paths, based on certain routing metric. Unlike wired networks, these best paths are selected typically on-demand, instead of pre-computed. Depending on the objective of the routing protocols, different routing metrics may be designed and used. For example, if the goal is to maximize the packet delivery probability and minimize the number of transmission, the ETX metric [11] may be used which captures the expected number of transmissions per link, and the best (least-cost) path is the path that minimizes the overall path ETX. If the objective is to minimize the energy consumption and maximize the network lifetime, an energy-aware metric should be used. For example, [6] has proposed a lifetime maximization algorithm for energy-aware routing in wireless sensor networks.

The key idea behind opportunistic routing is to take advantage of the broadcasting nature of wireless communication channels, while at the same time addressing the probabilistic nature of packet reception. Instead of selecting one or multiple *fixed* best paths, opportunistic routing protocols (e.g., [3], [4]) specify a set of forwarders, often arranged in a prioritized list, referred to as a *forwarder list*. Using the (pre-specified) forwarder list, after each packet transmission, the “best” forwarder among those which happen to receive the packet is used to forward the packet towards the destination. Hence a packet may opportunistically traverse any path from the source, among the set of forwarders, to the destination, instead of a fixed path. Through experiments in the MIT Roofnet [22] testbed, ExOR [3]—one of the first practical opportunistic routing protocols—is shown to increase the throughput by a factor of two to four over traditional best path routing schemes. Further improvements to ExOR [4], [15], [17] have also been developed. For instance, in [15] the key problem of how to optimally select the forwarder list is addressed, and an optimal algorithm (MTS) that minimizes the expected total number of transmissions is developed.

By its name, stateless (stochastic) routing does not maintain any routing state (e.g., topology, routing tables), and performs packet forwarding in a purely “random” fashion. In contrast to opportunistic routing, no forwarder list is pre-specified in general; any node receiving a packet may decide to forward the packet (some mechanisms to avoid and reduce unnecessary duplicate transmissions are generally employed). Stateless (stochastic) routing is typically designed and best suited for resource-constrained, dynamically varying and highly unreliable wireless network environments (e.g., sensor or delay/disruption tolerant networks). For instance, a stateless routing protocol is developed in [19] for wireless sensor networks. Due to its stateless feature, stateless routing schemes can be highly scalable; however, due to the pure randomness employed in these schemes, their efficacy, e.g., in terms of end-to-end packet delivery and other performance

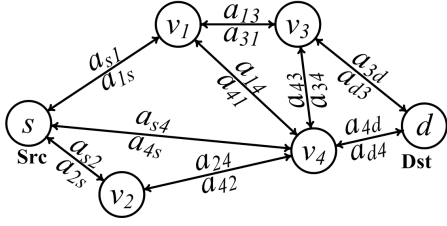


Fig. 1. An Example Wireless Topology

metrics, may suffer.

Thanks to widely disparate wireless network environments and diverse application objectives, no one routing paradigm always over-performs the others in practice. For instance, traditional best path routing may work very well in a static wireless environment with fairly stable and reliable wireless channels, while opportunistic routing may perform better where wireless channels are less reliable with frequently varying conditions. Hence in the design of practical routing protocols for wireless networks, which routing paradigm (or a hybrid combination thereof) to use will depend critically on the specific wireless environment. The ability to analyze, estimate and quantify various transmission costs (e.g., the expected number of transmissions, latency or energy consumption) is therefore imperative in the design of wireless networks. In the next subsection we illustrate how we can model the packet traversal in a wireless network under each of the wireless routing paradigms using Markov chains. Through these Markov chain models, transactions costs incurred by different routing schemes can then be computed using the notion of *hitting times* and other related quantities (e.g., sojourn times or hitting costs).

### B. Modeling Packet Traversal using Markov Chains

In this subsection, we illustrate how we can model packet forwarding under each of the three routing paradigms using Markov chains. Due to the probabilistic nature of wireless transmissions, when a packet is forwarded from one node, say  $i$ , to another node, say  $j$ , it only has some probability to “transit” from node  $i$  to node  $j$ . This suggests that we could model and trace the traversal of a packet when it is forwarded from one node to another in a wireless network as state transitions in a Markov chain. Before we proceed to describe how packet forwarding under each routing paradigm can be modeled using Markov chains, we first present some general notations and basic assumptions.

We model a wireless network as a (weighted) *directed* graph (i.e., a *digraph*)  $G = (V, E)$ , where  $V$  is the set of wireless nodes, and  $E$  is the set of *directed* wireless links. Here each directed link,  $\langle i, j \rangle$ , represents the relation that node  $j$  is within the transmission range of node  $i$ ; in other words, a packet transmitted by node  $i$  may be received by node  $j$  with some probability. We denote this probability by  $a_{ij}$ . Hence each link  $\langle i, j \rangle$  is associated with a link weight  $a_{ij}$ . We will simply refer to  $a_{ij}$  as the (link-level) packet delivery probability. More generally, we associate a weight  $a_{ij}$  to any (ordered) pair of nodes,  $\langle i, j \rangle$ . If  $\langle i, j \rangle \notin E$  (namely, node  $j$  is not within the

transmission range of node  $i$ ), we simply set  $a_{ij} = 0$ . Hence for any two distinct nodes  $i, j \in V, i \neq j$ , we have  $0 \leq a_{ij} \leq 1$ , and  $a_{ij} > 0$  if and only if  $\langle i, j \rangle \in E$ . Due to the asymmetric nature of wireless communications, in general we have  $a_{ij} \neq a_{ji}$ . In particular, we may have  $a_{ij} > 0$  but  $a_{ji} = 0$ . Further, for any node  $i \in V$ , we define  $a_{ii} = 0$ .

Let  $n = |V|$  denote the total number of nodes in the wireless topology. Then the  $n \times n$  matrix,  $A = [a_{ij}]$ , gives us a matrix representation of (one-hop or link-level) packet delivery probabilities of a wireless network. In general,  $A$  is *asymmetric*. We call  $A$  the adjacency matrix of the (weighted) directed graph  $G = (V, E)$ . In modeling packet forwarding using Markov chains, we assume that for any  $\langle i, j \rangle \in V$ , when a packet is forwarded by node  $i$ , the probability that the packet is received by node  $j$ , i.e.,  $a_{ij}$ , does not depend on where the packet was before reaching node  $i$ . Namely, except for node  $i$ ,  $a_{ij}$  does not depend on who and where the previous forwarders are. In other words, we assume that the Markov property holds. In modeling opportunistic routing and stateless routing, we will also make the simplifying assumption that the (link-level) packet delivery probabilities are independent. More precisely, let  $N(i) = \{j : \langle i, j \rangle \in E\}$  be the direct neighbors of node  $i$  that are within its transmission range. We assume that for any  $j_1, j_2 \in N(i)$ ,  $a_{ij_1}$  and  $a_{ij_2}$  are independent. We remark that to model the time-varying dynamics of a wireless network, we can introduce a series of time-dependent graphs  $G_t = (V_t, E_t, A_t)$  with time-varying node and edge sets as well as varying link-level packet delivery probabilities  $a_{ij}$ 's. For clarity and model simplicity, in this paper we focus only on *one instance* of such a time-varying graph, and assume that during this instance except for a few of them, the node/edge sets and  $a_{ij}$ 's are largely unchanged. Finally, we assume that the digraph graph  $G = (V, E)$  is *strongly* connected, namely, there is a (directed) path from any node to any other node in  $G$ .

**Best Path Routing.** Consider a specific source destination pair  $(s, d)$ . Let  $R(s, d) = \{u_0 = s, u_1, \dots, u_m, u_{m+1} = d\}$  denote the route (i.e., a best path) selected by a best path routing protocol for forwarding packets from  $s$  to  $d$ . We use  $G_R = (V_R, E_R) \subset G$  to denote the subgraph (a path or line subgraph) induced by  $R$ , where  $V_R = \{u_i, 0 \leq i \leq m+1\}$ , and  $E_R = \{\langle u_i, u_{i+1} \rangle, 0 \leq i \leq m\}$ . We can model the traversal of a packet being forwarded from  $s$  to  $d$  as a Markov chain with the state space  $V_R$  and the transition probability matrix  $P_R = [p_{ij}]$  defined as follows:

$$p_{ij} = \begin{cases} a_{i, i+1} & \text{if } j = i+1, i = 0, \dots, m \\ 1 - a_{i, i+1} & \text{if } j = i, i = 0, \dots, m \\ 1 & \text{if } j = i, i = m+1 \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Using the wireless topology shown in Fig. 1 as an example, let  $R_{s,d} = \{s, (u_1 =)v_2, (u_2 =)v_4, d\}$  be the best path (route) for the source-destination pair  $(s, d)$ . The corresponding Markov chain is schematically depicted in Fig. 2, where the arrows indicate the state transitions. The transition probability matrix  $P_R$  captures the fact that when a packet is forwarded by node

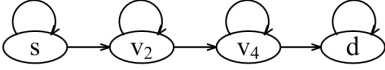


Fig. 2. Markov chain for Best-Path Routing

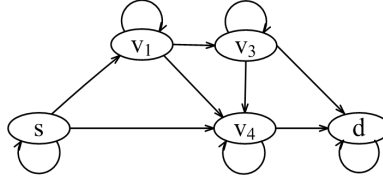
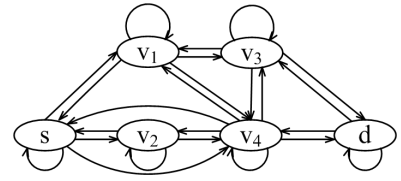
Fig. 3. Markov chain for Opportunistic Routing ( $FL = \{s, v_1, v_3, v_4, d\}$ )

Fig. 4. Markov chain for Stateless Routing

$u_i$ ,  $0 \leq i \leq m$ , with probability  $p_{i,i+1} = a_{i,i+1}$  the packet may be received by the next hop  $u_{i+1}$  (thus it transits or “walks” from node  $i$  to node  $i+1$  with probability  $p_{i,i+1}$ ), and with probability  $p_{ii} = 1 - a_{i,i+1}$  it is not received by node  $i+1$  (thus it stays with node  $i$ ). Hence packet forwarding under best path routing can be viewed as a *random walk* on the line subgraph  $G_R$  with  $P_R$  as the transition probability matrix. We note that this is an absorbing Markov chain, with node  $s$  as the starting state, and  $d$  the final absorbing state. As we will see later, using this Markov chain (or random walk on a digraph) we can formulate various transmission costs in terms of quantities associated with the Markov chain (random walk). For instance, the expected number of transmissions is the expected number of steps for a packet to “walk” from the source  $s$  to the destination  $d$ . Lastly, the above Markov chain model can be also easily generalized to (best-path-based) multi-path routing.

**Opportunistic Routing.** Given a source destination pair  $(s, d)$ , let  $FL(s, d) = \{u_0 = s, u_1, \dots, u_m, u_{m+1} = d\}$  denote the (*prioritized*) forwarder list selected by an opportunistic routing protocol, say, ExOR. We first note that unlike traditional best path routing, the forwarder list  $FL_{s,d}$  used in opportunistic routing represents not a path, but a subgraph  $G_{FL} = (V_{FL}, E_{FL})$  connecting the source  $s$  to the destination  $d$  (see Fig. 3 for an example, where  $FL_{s,d} = \{s, (u_1 = ) = v_1, (u_2 = )v_3, (u_3 = )v_4, d\}$ ). Within this subgraph  $G_{FL}$ , there are many (directed) paths from  $s$  to  $d$ ; which of them is actually traversed by a packet—during the packet forwarding process, depends on which nodes on the forwarder list receive the packet and which nodes forward the packet.

The priority of nodes is used in opportunistic routing to decide which node should forward a packet when several of them on the forwarder list receive the same packet. Here we use the convention that a node on the right has higher priority than a node to its left; namely, for any  $j' < j$ ,  $u_{j'}$  has higher priority than  $u_j$ . Using these priorities, we can describe the forwarding process of a single packet as follows: suppose node  $i$ , ( $0 \leq i \leq m$ ) is the current node to forward the packet. After its transmission, if the destination  $d$  receives it, then the forwarding process for this packet ends. Otherwise, suppose node  $j$ ,  $i < j \leq m$ , receives it. Node  $j$  will be the next forwarder if and only if no higher priority node,  $k > j$ , has received the packet. Hence to correctly capture the packet forwarding process under an opportunistic routing, we *must track which node is the next forwarder* instead of simply which nodes receive the packet. In other words, we say the packet has successfully “walked” from node  $i$  to node  $j$  if and only

if node  $j$  is the highest priority node that receives the packet. This happens with probability  $p_{ij} = a_{ij} \prod_{k>j} (1 - a_{ik})$ . The packet will stay with node  $i$ , if none of the higher priority nodes have received it. This happens with the probability  $p_{ii} = \prod_{k>i} (1 - a_{ik})$ . Hence we have a Markov chain defined on the state space  $V_{FL}$  with the transition probability matrix  $P_{FL} = [p_{ij}]$  given below,

$$p_{ij} = \begin{cases} a_{ij} \prod_{k>j} (1 - a_{ik}) & \text{if } 0 \leq i < j \leq m + 1 \\ \prod_{k>i} (1 - a_{ik}) & \text{if } j = i, 0 \leq i \leq m \\ 1 & \text{if } j = i, i = m + 1 \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

It is not too hard to verify that  $\sum_j p_{ij} = 1$ . Using the topology in fig.1 as an example, the corresponding opportunistic routing Markov chain is shown in Fig. 3. Again this is an absorbing Markov chain, with node  $s$  as the starting state, and  $d$  the final absorbing state. Using this Markov chain/random walk, we can again formulate various transmission costs using quantities associated with the chain/walk. As an aside, a key problem in opportunistic routing is to determine the “best” forwarder list  $FL$ , or subgraph  $G_{FL}$ , for a source and destination pair. This problem is addressed in [15], where an optimal algorithm is developed. In this paper we will assume that the (optimal) forwarder list is given and used.

**Stateless (Stochastic) Routing.** As no routing states are maintained or used, given a source-destination pair  $(s, d)$ , any node in  $G$  may be involved in the forwarding process of a packet. Suppose that node  $i$  is the current forwarder. After node  $i$ 's transmission, a subset of its direct neighbors,  $N(i)$ , may receive the packet. Unlike opportunistic routing where priorities are used to determine which node should be the next forwarder, any of these nodes may become the next forwarder with equal probability. For example, the next forwarder may be selected by using a random back-off mechanism where each node randomly sets a back-off timer value uniformly chosen from  $[0, t_0]$  where  $t_0$  is an appropriately chosen contention slot. Hence to track the packet traversals under stateless routing, we see that the packet stays with node  $i$  if and only if none of its neighbors receive the packet. This happens with probability  $p_{ii} = \prod_k (1 - a_{ik})$ . Otherwise, the packet transits or “walks” from node  $i$  to node  $j$ ,  $j \in N(i)$ , with probability  $p_{ij} = \frac{a_{ij}}{\sum_{k \in N(i)} a_{ik}} (1 - \prod_k (1 - a_{ik}))$ . Hence we have a Markov chain defined on the state space  $V$  (the entire node set) with the transition probability matrix  $P_G = [p_{ij}]$  given below,

$$p_{ij} = \begin{cases} \frac{a_{ij}}{\sum_k a_{ik}} (1 - \prod_k (1 - a_{ik})) & \text{if } i \neq j \\ \prod_k (1 - a_{ik}) & \text{if } i = j. \end{cases} \quad (3)$$

It is easy to verify that  $\sum_j p_{ij} = 1$ . If the graph is strongly connected, the Markov chain is irreducible. Especially, when the graph is symmetric, then the Markov chain will be reversible. The traversals of a packet under stateless routing are thus modeled as a random walk on the digraph  $G$  with the transition probability matrix  $P_G$ . Using the topology in Fig. 1 as an example with  $(s, d)$  as the source-destination pair, the resulting Markov chain is shown in Fig. 4.

**Modeling the Transmission Costs.** Given the Markov chain (or “random walk on a digraph”) models of wireless routing, we now briefly discuss how various transmission costs such as the expected number of transmissions, latency, duty cycle delay, or energy consumption, can be modeled using certain standard notions or quantities associated with the Markov chain/random walk.

We first use the expected number of transmissions as an example, and show how this cost can be formulated as the *hitting time*. In a Markov chain (or random walk), the hitting time  $H_{ij}$  is defined as the (expected) number of transitions (i.e., steps) for a random walker that starts from node (state)  $i$  to *first* reach (or hit) node  $j$ . The hitting time  $H_{ij}$  satisfies the following recursive relation

$$H_{ij} = \begin{cases} 1 + \sum_{k=1}^n p_{ik} H_{kj} & \text{if } i \neq j \\ 0 & \text{if } i = j. \end{cases} \quad (4)$$

Given the appropriately defined Markov chain for a wireless routing scheme, it is not too hard to see that the (expected) total number of transmissions needed to forward a packet from source  $s$  to destination  $d$  is exactly  $H_{sd}$ . The recursive relation eq.(4) plays a key role in computing the hitting time  $H_{sd}$ . The remainder of this paper is devoted to addressing this and other related computation problems.

To account for other transmission costs, we introduce a transition cost matrix  $T = [T_{ij}]$  associated with each one-hop transition,  $T_{ij} \geq 0, \forall i, j$ . For example, depending on the context and modeling objective,  $T_{ii}$  can be used to represent the per-node processing/transmission latency, duty cycle delay, or per-node energy consumption; where  $T_{ij}, j \in N(i)$  the one-hop forwarding latency, energy consumption, etc. Analogous to the notion of hitting time  $H_{ij}$ , we define the *hitting cost*,  $H_{ij}^s$ , (also referred as the *sojourn time* associated with  $T$ ) as the (expected) total cost (or “delay”) incurred by a random walk that starts at node  $i$  to first reach node  $j$ , where each state at any node  $k$  incurring a cost (delay)  $T_{kk}$  and each transition from node  $k$  to node  $l$  incurring a cost (delay) of  $T_{kl}$ . As in the case of  $H_{ij}$ ,  $H_{ij}^s$  satisfies the following recursive relation where  $s_i = \sum_j p_{ij} T_{ij}$  as the average transmission cost every time a packet visits.

$$H_{ij}^s = \begin{cases} \sum_{k=1}^n p_{ik} (T_{ik} + H_{kj}^s) = s_i + \sum_{k=1}^n p_{ik} H_{kj}^s & \text{if } i \neq j \\ 0 & \text{if } i = j. \end{cases} \quad (5)$$

Hence given the appropriately defined Markov chain for a wireless routing scheme and the transition cost matrix  $H^s$ , we can use  $H_{sd}^s$  to capture the (expected) total cost of transmission when forwarding a packet from source  $s$  to destination  $d$ . We

note that if  $T_{ij} = 1$  for all  $i, j$ , i.e.,  $T$  is the all-1 matrix, then  $H_{ij}^s = H_{ij}$ .

### III. RANDOM WALKS ON DIRECTED GRAPHS: HITTING, COMMUTE AND SOJOURN TIMES

In this section we briefly overview the random walk theory on *undirected graphs*, and show how important quantities such as hitting, commute and sojourn times can be computed. We then outline a generalization of the random walk theory to *directed graphs (digraphs)*, and show how the same quantities can be computed. The detailed description of this generalized *random walk theory on digraphs* and the associated proofs are reported in a separate paper [16].

#### A. Random Walks on Undirected Graphs

Given an undirected graph  $G = (V, E)$  that is finite connected (i.e., any node can reach any other node in  $G$ ), and let  $A$  be a *symmetric* weight (or adjacency) matrix appropriately defined on  $G$ , where  $n = |V|$ . For  $1 \leq i \leq n$ , define  $d_i = \sum_{j=1}^n a_{ij}$ , the (weighted) degree of node  $i$ , and  $d = \sum_{i=1}^n d_i$ , often referred to as the *volume* of  $G$ , denoted by  $vol(G)$ . Let  $D = \text{diag}[d_i]$  be a diagonal matrix of node degrees. Then  $P = D^{-1}A$  is a transition matrix associated with a Markov chain (a random walk) on  $G$ , where  $p_{ij} = a_{ij}/d_i$ . Let  $\pi = [\pi_i]_{1 \leq i \leq n}$  be its stationary distribution probability vector. It is well known (see, e.g., [1]) that this Markov chain (random walk) on  $G$  is *reversible*, namely

$$\pi_i p_{ij} = \pi_j p_{ji}. \quad (6)$$

where for  $1 \leq i \leq n$ ,

$$\pi_i = \frac{d_i}{\sum_k d_k} = \frac{d_i}{d}. \quad (7)$$

Given this random walk on an undirected graph, *hitting times*  $H_{ij}$  (cf. eq.(4)), *commute times*  $C_{ij} = H_{ij} + H_{ji}$  and the *hitting costs* or (*heterogeneous*) *sojourn times*  $H_{ij}^s$  (cf. eq.(5)) can be computed using a number of methods, through the well-known connections between the Markov chain/random walk theory, electrical resistance theory [12] and spectral graph theory [8], [10]. Here we present the results using the spectral graph theory.

In [8], the *normalized Laplacian matrix* for undirected graph  $G$  is defined as

$$\mathcal{L} = D^{-\frac{1}{2}}(D - A)D^{-\frac{1}{2}} = D^{\frac{1}{2}}(I - P)D^{-\frac{1}{2}}, \quad (8)$$

where  $\mathcal{L}$  is symmetric and positive semi-definite. Let  $\lambda_k$  and  $\mu_k$ ,  $1 \leq k \leq n$ , be the eigenvalues and the corresponding eigenvectors of  $\mathcal{L}$ , where  $\lambda_k$ 's are arranged in the increasing order where  $\lambda_1 = 0 < \lambda_2 \leq \dots \leq \lambda_n$ . Then the hitting time  $H_{ij}$  can be computed as follows (see [18]):

$$H_{ij} = \sum_{k>1} \frac{d}{\lambda_k} \left( \frac{\mu_{kj}^2}{d_j} - \frac{\mu_{ki}\mu_{kj}}{\sqrt{d_i d_j}} \right), \quad (9)$$

and the commute time  $C_{ij} = H_{ij} + H_{ji}$  is equal to

$$C_{ij} = \sum_{k>1} \frac{d}{\lambda_k} \left( \frac{\mu_{ki}}{\sqrt{d_i}} - \frac{\mu_{kj}}{\sqrt{d_j}} \right)^2. \quad (10)$$

In [7] Chau and Basu introduce a (*diagonal*) *sojourn time* matrix<sup>1</sup>  $T = \text{diag}[T_i]$ , where  $T_i$  represents a *per-node* transition cost or “delay” incurred at node  $i$ , and define the following ( $T$ -extended) Laplacian matrix  $\mathcal{L}^s$

$$\mathcal{L}^s = T^{-\frac{1}{2}} \mathcal{L} T^{-\frac{1}{2}}. \quad (11)$$

Let  $\sigma_k$  (again arranged in the increasing order) and  $\gamma_k$ ,  $1 \leq k \leq n$ , be the eigenvalues and eigenvectors of  $\mathcal{L}^s$ . Chau and Basu [7] obtain the following solution for the hitting cost matrix  $H^s = [H_{ij}^s]$ , extending the above (homogeneous) result for the hitting matrix  $H = [H_{ij}]$ ,

$$H_{ij}^s = \sum_{k>1} \frac{d_o^s}{\sigma_k} \left( \frac{\gamma_{kj}^2}{d_j T_j} - \frac{\gamma_{ki} \gamma_{kj}}{\sqrt{d_i T_i d_j T_j}} \right), \quad (12)$$

where  $d_o^s = \sum_k d_k T_k$ .

### B. Random Walks on Directed Graphs (Digraphs)

We now outline the (generalized) random walk theory on *digraphs* we have developed. Due to the space limitation, the detailed description and proofs are omitted and can be found in [16]. Given a *directed graph*  $G = (V, E)$  which is assumed to be *strongly connected*. Let  $P$  be a transition matrix<sup>2</sup> associated with a Markov chain defined on  $G$ . Let  $\pi = [\pi_i]_{1 \leq i \leq n}$  be the stationary distribution probability vector, namely,  $\pi = \pi P$ . In general, the Markov chain (random walk) on  $G$  is *no longer reversible*, and eqs.(6) and (7) do not hold. The assumption that  $G$  is strongly connected (or equivalently,  $P$  is irreducible) yields the following properties [1]: i) the Markov chain is irreducible with no transient states; ii)  $P$  has a simple eigenvalue equal to 1 with right (column) eigenvector  $e = [1, 1, \dots, 1]^T$  and left (row) eigenvector  $\pi$ , the stationary probability vector; and iii)  $\pi_i > 0$ ,  $1 \leq i \leq n$ .

Let  $\Pi = \text{diag}[\pi_i]$  be a diagonal matrix with  $\pi_i$ 's on the diagonal. We introduce the *generalized normalized Laplacian* matrix for a strongly connected digraph  $G = (V, E)$  (and an associated Markov chain with the transition matrix  $P$ ) as follows:

#### DEFINITION 1 (Generalized Normalized Laplacian $\tilde{\mathcal{L}}$ ).

$$\tilde{\mathcal{L}} = \Pi^{\frac{1}{2}} (I - P) \Pi^{-\frac{1}{2}}, \quad (13)$$

or in the scalar form,

$$\tilde{\mathcal{L}}_{ij} = \begin{cases} 1 - p_{ii} & \text{if } i = j, \\ -\pi_i^{-\frac{1}{2}} p_{ij} \pi_j^{-\frac{1}{2}} & \text{if } \langle i, j \rangle \in E, i \neq j, \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

Let  $\tilde{\mathcal{L}}^+$  be the (*Penrose-Moore*) pseudo-inverse [13] of  $\tilde{\mathcal{L}}$ . We can show (in [16]) that  $\tilde{\mathcal{L}} \tilde{\mathcal{L}}^+ = I - \Pi^{\frac{1}{2}} J \Pi^{\frac{1}{2}}$ , where  $J = [J_{ij}]$  is the all-1 matrix and  $J_{ij} = 1$ . Using this relation and other properties of the generalized normalized Laplacian

<sup>1</sup>In contrast, in Section II-B we have introduced a general transition cost matrix  $T = [T_{ij}]$ ,  $T_{ij} \geq 0$ , with a transition cost associated with each node  $T_{ii}$  and each link  $T_{ij}$ , where  $T$  is not necessarily diagonal nor symmetric.

<sup>2</sup>For example,  $P$  can be derived from an appropriately defined (non-negative) weight (or “adjacency”) matrix  $A$  on  $G$ , where  $A$  is in general *asymmetric*.

matrix  $\tilde{\mathcal{L}}$ , in [16] we show that the hitting times  $H_{ij}$  as defined by the relation in eq.(4) can be computed using  $\tilde{\mathcal{L}}^+$  as stated in the following theorem.

**THEOREM 1 (Hitting and Commute Times for Random Walks on Digraphs).** *Given the generalized normalized Laplacian matrix  $\tilde{\mathcal{L}}$  defined in eq.(13) and eq.(14), and let  $\tilde{\mathcal{L}}^+$  be its pseudo-inverse. Then we have the hitting time  $H_{ij}$  as*

$$H_{ij} = \frac{\tilde{\mathcal{L}}_{jj}^+}{\pi_j} - \frac{\tilde{\mathcal{L}}_{ij}^+}{\sqrt{\pi_i \pi_j}}, \quad (15)$$

or in the matrix form,

$$H = J \cdot \text{diag}(\Pi^{-\frac{1}{2}} \tilde{\mathcal{L}}^+ \Pi^{-\frac{1}{2}}) - \Pi^{-\frac{1}{2}} \tilde{\mathcal{L}}^+ \Pi^{-\frac{1}{2}}. \quad (16)$$

The commute times,  $C_{ij} = H_{ij} + H_{ji}$ , can be computed as follows,

$$C_{ij} = \frac{\tilde{\mathcal{L}}_{jj}^+}{\pi_j} + \frac{\tilde{\mathcal{L}}_{ii}^+}{\pi_i} - \frac{\tilde{\mathcal{L}}_{ij}^+}{\sqrt{\pi_i \pi_j}} - \frac{\tilde{\mathcal{L}}_{ji}^+}{\sqrt{\pi_i \pi_j}}. \quad (17)$$

We note that if the underlying graph  $G$  is undirected, and the transition probability matrix  $P$  of the Markov chain is  $P = D^{-1}A$  (as defined in Section III-A, where  $A$  is *symmetric*), then  $\tilde{\mathcal{L}} = \mathcal{L}$ . Furthermore, one can show that eqs.(16) and (17) are equivalent to eqs.(9) and (10). Hence our theory of random walks on digraphs subsumes the existing theory of random walks on undirected graphs as a special case.

We now extend the above results for hitting and commuting times to hitting and commute costs. Given an (asymmetric) transition cost matrix  $T = [T_{ij}]$ ,  $T_{ij} \geq 0$ , with  $T_{ij}$  as the per-link transition cost, and  $T_{ii}$  as the per-node transmission cost, define  $S = \text{diag}[s_i]$  a diagonal matrix with  $s_i = \sum_{j=1}^n T_{ij} p_{ij}$  as the average transmission cost every time a packet visits. We define the following *normalized cost Laplacian* matrix,  $\tilde{\mathcal{L}}^s$  as

#### DEFINITION 2 (Normalized Cost Laplacian $\tilde{\mathcal{L}}^s$ ).

$$\tilde{\mathcal{L}}^s = S^{-\frac{1}{2}} \tilde{\mathcal{L}} S^{-\frac{1}{2}}, \quad (18)$$

where the  $\tilde{\mathcal{L}}$  is the generalized normalized laplacian matrix.

The corresponding scalar form is

$$\tilde{\mathcal{L}}_{ij}^s = \begin{cases} s_i^{-1} (1 - p_{ii}) & \text{if } i = j, \\ -s_i^{-\frac{1}{2}} \pi_i^{-\frac{1}{2}} p_{ij} \pi_j^{-\frac{1}{2}} s_j^{-\frac{1}{2}} & \text{if } \langle i, j \rangle \in E, i \neq j, \\ 0 & \text{otherwise.} \end{cases} \quad (19)$$

Let  $\tilde{\mathcal{L}}^{s+}$  be the (Penrose-Moore) pseudo-inverse of  $\tilde{\mathcal{L}}^s$ . In [16] we show that the hitting costs  $H_{ij}^s$  as defined by the relation in eq.(5) can be computed using  $\tilde{\mathcal{L}}^{s+}$  as stated below

**THEOREM 2 (Hitting Costs for Random Walks on Digraphs).**

$$H_{ij}^s = d^s \left( \frac{\tilde{\mathcal{L}}_{jj}^{s+}}{\pi_j s_j} - \frac{\tilde{\mathcal{L}}_{ij}^{s+}}{\sqrt{\pi_i s_i \pi_j s_j}} \right), \quad (20)$$

where  $d^s = \sum_k \pi_k s_k$ .

Analogous to the commute times  $C_{ij}^s$ , we can also define the *commute costs* as  $C_{ij}^s = H_{ij}^s + H_{ji}^s$ , and they can be computed easily using eq.(20).

#### IV. ESTIMATING TRANSMISSION COST FOR DIFFERENT ROUTING STRATEGIES

In this section we apply the theory of random walks on digraphs to various routing strategies and see how it can be used to estimate various transmission costs. We consider three specific examples: the *keep-connect* routing, a class of energy-aware best-path routing algorithms proposed in [20], the *opportunistic routing* protocol as defined in [3] and analyzed in [15], and the *stateless routing* protocol introduced and analyzed in [7].

##### A. Keep Connect Routing for Network Life Time Maximization

Many energy-aware routing protocols have been developed for energy-constrained wireless networks, such as wireless sensor networks. These protocols take the energy cost as a key metric in selecting routes, and attempt to forward packets along a path that minimizes the energy consumption or maximizes the overall network life time. The *keep-connect* routing (thereafter referred to as *KC*) is a class of energy-aware routing algorithms proposed in [20] that take into account both the energy cost and the “importance” of nodes in the overall network connectivity so as to maximize the network life time. Here the network life time is defined as the time until the network becomes disconnected. In [20], the authors apply the spectral graph theory—in particular, use the Fiedler value (the second smallest eigenvalue of a graph Laplacian)—to the design and analysis of the *KC* routing algorithms. We use *KC* as a simple example to illustrate how to estimate the wireless transmission costs (hitting costs).

In *KC*, the “importance” of a node  $i$  in terms of its connectivity in a graph  $G = (V, E)$  is defined as follows:  $W(i) = 1/\lambda_2(L(G_{-i}))$ , where  $G_{-i}$  is a graph resulting from  $G$  with node  $i$  and its adjacent edges removed,  $L(G_{-i})$  is the graph Laplacian of  $G_{-i}$ , and  $\lambda_2(L(G_{-i}))$  is the second smallest eigenvalue (the Fiedler value) of  $L(G_{-i})$ . For each link  $\langle i, j \rangle \in E$ , let  $e_{ij}$  denote the one-hop transmission energy cost from  $i$  to  $j$ . We introduce the following transition cost matrix  $T = [T_{ij}]$ , where  $T_{ij} = e_{ij}W(i)$ , if  $\langle i, j \rangle \in E$ , and  $T_{ij} = 0$ , otherwise. Given a path or route  $R = \{u_0 = s, u_1, \dots, u_m, u_{m+1} = d\}$  for a source-destination pair  $(s, d)$ . Then using the Markov chain transition matrix  $P_R$  defined on the line (sub-)graph  $G_R$  (with  $V_R = \{u_0, u_1, \dots, u_{m+1}\}$ ) as given in eq.(1) of Section II-B, we can easily solve the hitting cost equation (5) and compute the hitting cost matrix  $H_R^s$  for  $G_R$  as follows:

$$H_{i,m+1}^s = \begin{cases} \sum_{k=i}^m e_{k,k+1} W(k) \frac{1}{a_{k,k+1}} & 0 \leq i < m+1 \\ 0 & \text{otherwise.} \end{cases} \quad (21)$$

In particular,  $H_{sd}^s = H_{0,m+1}^s$  is the expected total cost associated with the route  $R$ . Hence to find the best path in *KC* that minimizes the energy consumption and maximizes the network life time is equivalent to finding the best route  $R$  that minimizes  $H_{sd}^s$  for any given source-destination pair  $(s, d)$ .

##### B. Opportunistic Routing

Following the description in Section II-B, given a (prioritized) forwarder list  $FL(s, d) = \{u_0 = s, u_1, \dots, u_m, u_{m+1} = d\}$  for a source-destination pair  $(s, d)$ , let  $P_{FL}$  be the transition matrix for the Markov chain on the subgraph  $G_{FL}$  (with  $V_{FL} = \{u_0, u_1, \dots, u_{m+1}\}$ ), as given in eq.(2). Hence the total expected number of transmissions using this forwarder list  $FL$  can be computed using the hitting time matrix  $H$  associated with  $P_{FL}$  defined on  $G_{FL}$ . In particular,  $H_{sd} = H_{0,m+1}$  is the total expected number of transmissions from source  $s$  to destination  $d$ . We can apply the theory of random walks on digraphs to compute  $H_{sd}$ . In this case, we can explore the special structure of  $P_{FL}$  and find a *closed form* recursive formula for  $H$  directly. From eq.(2), we note that  $P_{FL}$  is an upper triangular matrix. Hence the hitting times,  $H_{i,m+1}$  satisfies the following recursive relations:

$$\begin{cases} H_{0,m+1} = 1 + \sum_{k \geq 0}^{m+1} p_{0k} H_{k,m+1}, \\ \dots \\ H_{i,m+1} = 1 + \sum_{k \geq i}^{m+1} p_{ik} H_{k,m+1}, \\ \dots \\ H_{m,m+1} = 1 + p_{m,m} H_{m,m+1} + p_{m,m+1} H_{m+1,m+1}, \\ H_{m+1,m+1} = 0. \end{cases} \quad (22)$$

We observe that these (upper triangular) linear recursive equations can be easily solved, starting from the bottom to the top. Plugging the transition probabilities  $p_{ij}$ 's in, using eq.(2), we can write the above recursive equation as follows:

$$H_{i,m+1} = \begin{cases} \frac{1 + \sum_{k > i}^{m+1} a_{ik} \prod_{j > k} (1 - a_{ij}) H_{k,m+1}}{1 - \prod_{k > i} (1 - a_{ik})} & 0 \leq i < m+1, \\ 0 & i = m+1. \end{cases} \quad (23)$$

The above equation is exactly the same formula obtained in [15], through an event-based direct probability analysis approach. The random walk method proposed here is far simpler. We can further extend the above analysis of the expected number of transmissions to other transmission costs. Let  $S = [s_i]$  be a diagonal average transition cost matrix in subgraph  $G_{FL}$ , as defined in section III-B. We can derive the following recursive formula for computing the expected total transmission cost matrix defined on  $G_{FL}$ :

$$H_{i,m+1}^s = \begin{cases} \frac{s_i + \sum_{k > i}^{m+1} a_{ik} \prod_{j > k} (1 - a_{ij}) H_{k,m+1}^s}{1 - \prod_{k > i} (1 - a_{ik})} & 0 \leq i < m+1, \\ 0 & i = m+1. \end{cases} \quad (24)$$

Hence for a given  $FL(s, d) = \{u_0 = s, u_1, \dots, u_m, u_{m+1} = d\}$ , the total expected transmission cost of  $FL$  is  $H_{sd}^s = H_{0,m+1}^s$ . In [15], using the recursive relations eqs.(23) and (24), an optimal algorithm (MTS, stands for *minimum transmission selection*) and generalized MTS algorithms have been developed to minimize respectively, the total expected number of transmission, total expected energy consumption, total end-to-end transmission latency, and so forth.

### C. Stateless routing

In [7], Chau and Basu analyze the stateless routing algorithm using the random walk theory for *undirected graphs*, and apply eq.(12) to derive an end-to-end delay estimation formula with heterogenous sojourn times. Since the undirected graph model, they assume that wireless links are symmetric. Given the theory of random walks on digraphs we have developed, we can now solve the same problem using eq.(20), where the transition matrix  $P$  for the Markov chain/random walk is given in eq.(3). A step-by-step procedure for constructing the Markov chain and then computing the hitting time matrix  $H$  is given in Algorithm 1.

---

**Algorithm 1** Delay Estimation Algorithm for Stateless routing, given the adjacency matrix  $A$ .

---

- 1: Form the transition probability matrix  $P$  using eq.(3).
  - 2: Compute the stationary probability matrix  $\Pi$ .
  - 3: Compute the generalized normalized Laplacian matrix  $\tilde{\mathcal{L}}$  (See DEFINITION 1).
  - 4: Compute the pseudo-inverse of  $\tilde{\mathcal{L}}$ .
  - 5: Compute the hitting time matrix  $H$  using eq.(15).
- 

Moreover, by introducing an arbitrary transmission matrix  $T = [T_{ij}], T_{ij} \geq 0$  to represent per-node and/or per-hop transmission cost, we can apply Algorithm 1 (with  $\tilde{\mathcal{L}}$  and  $H$  replaced by  $\tilde{\mathcal{L}}^s$  and  $H^s$ ) to compute various transmission costs associated with the stateless routing, such as latency, energy consumption, and so forth.

### V. PERFORMANCE EVALUATION AND COMPARISON

In this section, we compare the transmission cost estimation results obtained using the *random walks on digraphs* model vs. using the *random walks on undirected graphs* model, where the asymmetric packet delivery probabilities of a link,  $a_{ij}$  and  $a_{ji}$  are symmetrized using their average,  $\bar{a}_{ij} = \bar{a}_{ji} = (a_{ij} + a_{ji})/2$ . We use the stateless routing [7] as the example. Similar results are obtained for other routing schemes, which are not presented here due to space limitation.

The wireless topologies are generated in a  $300m \times 300m$  area, where nodes are randomly placed. The transmission radius of each wireless node is  $100m$ . The density (the number of nodes,  $n$ ) of the topology are varied in three settings,  $n = 15, 25, 40$ . In the first set of reported results (Section V-A), the link qualities between nodes are randomly assigned. In the second set of reported results (Section V-B), we control and vary both the degree of asymmetry (the percentage of asymmetric links) and the (asymmetric) link quality variation (the extent to which  $a_{ij}$  and  $a_{ji}$  differ), and randomly generate link qualities that meet these constraints. For each network density, we randomly generate 1000 topologies with a specific parameter setting, and compute the average of the results obtained for these 1000 topologies. Due to space limitation, we focus on the hitting time estimation (namely, the expected number of transmissions). Let  $H_{ij}^d$  represent the hitting time from node  $i$  to node  $j$  computed using the digraph model,

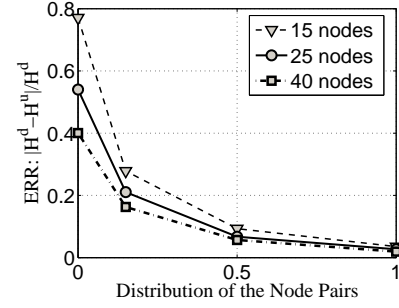


Fig. 5. Distribution of  $ERR$

and  $H_{ij}^u$  the hitting time computed using the undirected graph model. We use the following definition of *relative error* for comparing and measuring the inaccuracy in hitting time computation introduced by the symmetrization of link qualities (i.e., when using the undirected graph model).

$$ERR_{ij} = |H_{ij}^d - H_{ij}^u|/H_{ij}^d, \quad (25)$$

#### A. Effect of Network Density on $ERR$

Fig. 5 shows the results obtained for the three network density settings. Here, we sort the node pairs in each topology by their  $ERR$  values, so that we can plot the  $ERR$ 's in a monotonically decreasing order. We see that the relative errors are overall lower in denser networks. To better illustrate the effect of network density, we group the node pairs based on the ranges of their  $ERR$ 's:  $[0, 30\%)$ ,  $[30\%, 60\%)$ , and  $[60\%, \infty)$ , and compute the percentage of node pairs fall within each range, and the results are shown in Fig. 6. We see that using the undirected graph model, from 25% up to 50% of all node pairs have an average relative error at least 30%, and a few percentage have an average relative error of more than 60%. Both figures indicate that when the network density increases, the percentage of node pairs having large relative errors decreases. This is in fact not surprising: As the network is dense (e.g., 40 nodes in a  $300m \times 300m$  area, each having transmission radius of  $100m$ ), path diversity is high. In other words, the number of random (and shorter) paths between two nodes are typically high, thus reducing the hitting time between them. The asymmetric links thus likely have less impact on the overall results. On the other hand, when the network is relatively sparse, the asymmetric links have much higher impact, and therefore how to perform routing effectively becomes more critical.

#### B. Effect of Degree of Asymmetry and Link Quality Variation

In this second set of evaluation, we fix the network density to  $n = 25$ , and vary the degree of asymmetry (percentage of asymmetric links) and the (asymmetric) link quality variation (the extent to which  $a_{ij}$  and  $a_{ji}$  differ). For the former, we utilize the asymmetric distribution probability  $S\%$  to control the distribution of asymmetric links in the topology. For the latter, we use a parameter  $\epsilon$  defined below to control the link variation: for each node pair chosen to have a pair of asymmetric links (determined by a given  $S\%$ ), we first randomly generate  $a_{ij}$ , and then randomly generate  $a_{ji} (\neq$



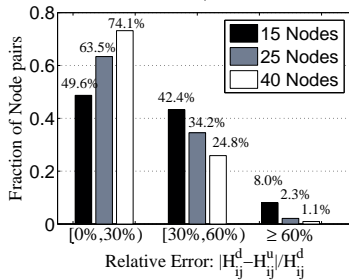


Fig. 6. ERR distribution for different network densities.

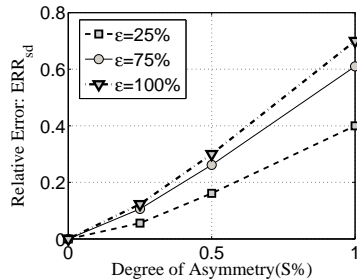


Fig. 7. Effect of degree of asymmetry.

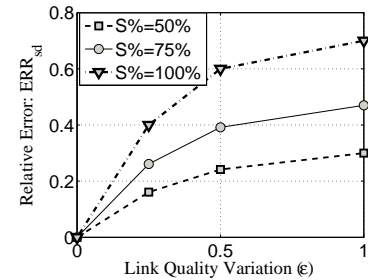


Fig. 8. Effect of (asymmetric) link quality variation.

$a_{ij}$ ) such that  $a_{ji} \in [(1 - \epsilon)a_{ij}, \min((1 + \epsilon)a_{ij}, 1)]$ . In the simulations, we select  $S\% = 0, 25\%, 50\%, 75\%, 100\%$ , and  $\epsilon = 0, 25\%, 50\%, 75\%, 100\%$ . We randomly generate 1000 topologies for each combination  $(S\%, \epsilon)$  and compute the average of relative error  $ERR$ 's for each node pair. In the following, the node pairs with the worst average performance are used.

Fig. 7 shows the effect of the degree of link asymmetry,  $S\%$  on the worst-case  $ERR$  performance. Here we fix  $\epsilon$  to 25%, 75% and 100%, and vary the degree of link asymmetry,  $S\%$ . We see that as the degree of asymmetry  $S\%$  increases, the (worst-case)  $ERR$  increases rapidly, up to more than 0.6 when  $S\% = 100\%$  and  $\epsilon = 75\%$  or 100%.

Fig. 8 shows the effect of the (asymmetric) link quality variation on the worst-case  $ERR$  performance. Here we fix  $S\%$  to 50%, 75% and 100%, and vary  $\epsilon$  from 0 to 1. Similarly, we see that when we allow a larger extent that the qualities of asymmetric link pair,  $a_{ij}$  and  $a_{ji}$ , can differ, the higher the (worst-case)  $ERR$  is. All in all, we conclude that when the degree of asymmetry is high, and the link qualities of asymmetric link pairs can differ to a larger extent, it is important to take into account the asymmetric link qualities in routing decision making and estimation of various transmission costs. This is especially the case in a relatively sparse network.

## VI. CONCLUSION

In this paper we have developed a unified theoretical framework for estimating various transmission costs of packet forwarding. We illustrated how packet forwarding under each of three routing paradigms—best routing, opportunistic routing, and stateless routing—can be modeled as *random walks on digraphs*. By generalizing the theory of random walks that has primarily been developed for undirected graphs to digraphs, we showed how various transmission costs can be formulated in terms of hitting times and hitting costs of random walks on digraphs. As representative examples, we applied the theory to three specific routing protocols, one under each paradigm. Extensive simulations demonstrate that the proposed digraph based analytical model can achieve more accurate transmission cost estimation over existing methods. As part of future work, we plan to apply the theory of random walks on digraphs for analysis of the transmission costs under various wireless *broadcasting* strategies, by considering the *cover times* of random walks. We are also interested in extending the theory

for studying the dynamics and evolution of wireless networks, incorporating the time dimension into the model.

## REFERENCES

- [1] D. Aldous and J. A. Fill. Reversible markov chains and random walks on graphs. In <http://www.stat.berkeley.edu/~Aldous/RWG/book.html>.
- [2] S. Banerjee and A. Misra. Minimum energy paths for reliable communication in multi-hop wireless networks. In *Proceedings of MOBIHOC'02*, pages 146–156, 2002.
- [3] S. Biswas and R. Morris. ExOR: Opportunistic routing in multi-hop wireless networks. In *Proceedings of SIGCOMM'05*. ACM, 2005.
- [4] S. Chachulski, M. Jennings, S. Katti, and D. Katabi. Trading structure for randomness in wireless opportunistic routing. In *Proceedings of SIGCOMM'07*. ACM, 2007.
- [5] J.-H. Chang and L. Tassiulas. Energy conserving routing in wireless ad-hoc networks. In *Proceedings of INFOCOM'00*, pages 22–31, 2000.
- [6] J.-H. Chang and L. Tassiulas. Maximum lifetime routing in wireless sensor networks. *IEEE/ACM Transaction on Networking*, 12(4):609–619, 2004.
- [7] C.-K. Chau and P. Basu. Exact analysis of latency of stateless opportunistic forwarding. In *Proceedings of INFOCOM '09*. IEEE, 2009.
- [8] F. R. K. Chung. Spectral graph theory. <http://www.math.ucsd.edu/~fan/research/revise.html>.
- [9] F. R. K. Chung. Laplacians and the cheeger inequality for directed graphs. *Annals of Combinatorics*, 9:1–19, sep 2005.
- [10] F. R. K. Chung and S. T. Yau. Discrete green's functions. *Journal of Combinatorial Theory, Series A*, pages 191–214, July 2000.
- [11] D. D. Couto, D. Aguayo, J. Bicket, and R. Morris. High-throughput path metric for multi-hop wireless routing. In *Proceedings of MobiCom '03*. ACM, 2003.
- [12] P. G. Doyle and L. J. Snell. Random walks and electric networks. *Mathematical Association of America*, Jan 1984.
- [13] R. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1st edition, 1985.
- [14] D. B. Johnson and D. A. Maltz. Dynamic source routing in ad hoc wireless networks. In *Mobile Computing*. Kluwer Academic Publishers, 1996.
- [15] Y. Li, W. Chen, and Z.-L. Zhang. Optimal forwarder list selection in opportunistic routing. In *Proceedings of IEEE MeshTech'09*, 2009.
- [16] Y. Li and Z.-L. Zhang. Green's function and generalized normalized laplacian for directed graph. Technical Report, <http://www.cs.umn.edu/>, CSE Department of University of Minnesota, 2009.
- [17] Y. Lin, B. Li, and B. Liang. CodeOR: Opportunistic routing in wireless mesh networks with segmented network coding. In *Proceedings of ICNP '08*. IEEE, 2008.
- [18] L. Lovász. Random walks on graphs: A survey. *Combinatorics, Paul Erdős is Eighty*, 2:1–46, 1993.
- [19] I. Mabrouki, X. Lagrange, and G. Froc. Random walk based routing protocol for wireless sensor networks. In *Proceedings of ValueTools '07*, pages 1–10, 2007.
- [20] C. Pandana and K. Liu. Robust connectivity-aware energy-efficient routing for wireless sensor networks. *IEEE Transactions on Wireless Communications*, 7(10):3904–3916, October 2008.
- [21] C. E. Perkins and E. M. Royer. Ad hoc on-demand distance vector routing. In *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications*, pages 90–100, 1999.
- [22] Roofnet. MIT Roofnet. <http://pdos.csail.mit.edu/roofnet>.