Robotic surveillance and Markov chains with minimal weighted Kemeny constant

Rushabh Patel, Pushkarini Agharkar, Francesco Bullo

***Abstract*—This article provides analysis and optimization results for the *mean first passage time*, also known as the *Kemeny constant*, of a Markov chain. First, we generalize the notion of the Kemeny constant to environments with heterogeneous travel and service times, denote this generalization as the *weighted Kemeny constant*, and we characterize its properties. Second, for reversible Markov chains, we show that the minimization of the Kemeny constant and its weighted counterpart can be formulated as convex optimization problems and, moreover, as semidefinite programs. Third, we apply these results to the design of stochastic surveillance strategies for quickest detection of anomalies in network environments. We numerically illustrate the proposed design: compared with other well-known Markov chains, the performance of our Kemeny-based strategies are always better and in many cases substantially so.**

I. Introduction

A. Problem description and motivation

The subject of this paper is the analysis, generalization and minimization of the mean first passage time for a random walk. This problem is of general mathematical and engineering interest in the study of Markov chains and random walks; similar to the fastest mixing Markov chain, the mean first passage time is a metric by which to gauge the performance of a random walk. In a robotic context, our motivating application is the design of surveillance algorithms for quickest detection of intruders and anomalies. Specific examples include the monitoring of oil spills [12], the detection of forest fires [27], the tracking of border changes [39], and the periodic patrolling of an environment [17], [35]. Other applications in single and multi-agent systems include minimizing emergency vehicle response times [5] as well as servicing tasks in robotic warehouse management [41]. In this paper we propose stochastic surveillance strategies based on Markov chains. This study is motivated by the desire to design surveillance strategies with pre-specified stationary distributions, that are easily implementable and inherently unpredictable. In areas of research outside of robotics, the study of the mean first passage time is potentially useful in determining how quickly information propagates in an online network [4] or how quickly an epidemic spreads through a contact network [40].

This work was supported in part by NSF grant CNS 1035917 and ARO grant W911NF-11-1-0092.

[1] Rushabh Patel, Pushkarini Agharkar, and Francesco Bullo are with the Mechanical Engineering Department and the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, CA 93106-5070, USA r_patel@engineering.ucsb.edu, agharkar@umail.ucsb.edu, bullo@enigneering.ucsb.edu

B. Literature review

For a random walk associated with a Markov chain, the mean first passage time, also known as the *Kemeny constant*, of the chain is the expected time taken by a random walker to travel from an arbitrary start node to a second randomly-selected node in a network. The Kemeny constant of a Markov chain first appeared in [26] and has since been studied by several scientists, e.g., see [25], [28] and references therein. Bounds on the mean first passage time for an arbitrary Markov chain over various network topologies appear in [25], [30]. To the best of our knowledge, no results are available so far on the optimization of the mean first passage time of a Markov chain, nor on a framework for this quantity which accounts for general weighted graphs.

The mean first passage time is closely related to other well-known metrics for graphs and Markov chains. We discuss two such quantities in what follows. First, the *Kirchhoff index* [29], also known as the *effective graph resistance* [15], is a related metric quantifying the distance between pairs of vertices in an electric network. The relationship between electrical networks and random walks on graphs is explained elaborately in [14]. For an arbitrary graph, the Kirchoff index and the Kemeny constant can be calculated from the eigenvalues of the conductance matrix and the transition matrix, respectively. The relationship between these two quantities for regular graphs is established in [34]. Second, the *mixing rate* of an irreducible Markov chain is the rate at which an arbitrary distribution converges to the chain's stationary distribution [13]. It is well-know that the mixing rate is related to the second largest eigenvalue of the transition matrix of the Markov chain. The influential text [31] provides a detailed review of the mixing rate and of other notions of mixing. Recently, [28] refers to the Kemeny constant as the "expected time to mixing" and relates it to the mixing rate.

In this paper we study strategies to surveil an environment, to provide a desired coverage frequency, and to detect an intruder in minimum time. The surveillance problem has appeared in the literature in various manifestations. The authors of [38] look at minimizing time of detection of noisy anomalies via persistent surveillance strategies, and in [32] wireless sensor networks are utilized for intruder detection in previously unknown environments. In [3], the authors explore strategies for surveillance using a multi-agent ground vehicle system which must maintain connectivity between agents. A non-cooperative game framework is utilized in [11] to determine an optimal strategy for intruder detection, and in [36] a similar framework is used to analyze intruder detection for ad-hoc mobile networks. In our setup we model the environment as a graph and design random walks on this graph imposing restrictions on the stationary distribution of the walk. In other works with a similar setup, Markov chain Monte Carlo methods [19], [37] are used to design surveillance strategies. In [19] convexity results for symmetric matrices are utilized to further optimize those strategies. Deterministic policies have been used to minimize the visit frequencies in a graph [16], [38], however, a main result of [37] shows that deterministic policies are ill-suited when designing strategies with arbitrary constraints on those visit frequencies. We take an alternate approach and design policies using Markov chains with minimal Kemeny constant.

C. Contribution

Before stating our contributions, it is worth mentioning that all work to date on the *mean first passage time* and mixing rate of a Markov chain on a graph has been complete under the assumption of homogeneous travel time along the edges of the graph. The contributions of this paper are then six fold. First, we provide a convex optimization framework to minimize the Kemeny constant of a reversible Markov chain given the underlying graph topology of the random walk and the desired stationary distribution. Second, using doubly-weighted graphs we extend the formulation of the *mean first passage time* to the network environments with non-homogeneous travel times, a generalization not yet looked at in the literature. We denote this extension the *weighted Kemeny constant*. Third, we derive a closed form solution for the *weighted Kemeny constant* and show its relation to the Kemeny constant. Fourth, we provide a convex optimization framework to minimize the weighted Kemeny constant of a Markov chain with desired stationary distribution. Fifth, we provide a semidefinite program (SDP) formulation for the optimization of the Kemeny constant and the weighted Kemeny constant. Finally, we look at two stochastic surveillance scenarios; in the first scenario we provide a setup in which minimizing the weighted Kemeny constant leads to the optimal Markov-chain strategy. In the second surveillance scenario we establish through numerical simulation that the Markov chain with the minimum weighted Kemeny constant performs substantially better compared with other well-known Markov chains (i.e., the fastest mixing chain and the Metropolis-Hastings Markov chain).

A preliminary short version of this article is to appear at the 2014 CDC conference as [1] and is currently available at `http://motion.me.ucsb.edu`. This article contains several addenda and updates not found in [1] including, but not limited to, detailed proofs of theorems that were presented in [1], the introduction of a new algorithm, and extensive simulation results.

D. Organization

The paper is organized as follows. In Section I-E we summarize the notation which we use throughout the paper and briefly review properties of Markov chains. In Section II we give relevant background for the Kemeny constant and present our results for its minimization. In Section III we introduce and provide detailed characterization of the weighted Kemeny constant as well as its minimization. In Section IV we provide practical surveillance applications of the weighted Kemeny constant. In the final Section we present our conclusions and future research directions.

E. Notation

We use the notation $A = [a_{ij}]$ to denote a matrix A with the element a_{ij} in its i-th row and j-th column and, unless otherwise indicated, use bold-faced letters to denote vectors. Letting δ_{ij} denote the Kronecker delta, $A_{\mathrm{d}} = [\delta_{ij}a_{ij}]$ represents the diagonal matrix whose diagonal elements are the diagonal elements of the matrix A. The column vector of all ones and length n is denoted by $\mathbb{1}_n \in \mathbb{R}^{n\times 1}$ and I represents the identity matrix of appropriate dimension. We use diag$[\boldsymbol{b}]$ to denote the diagonal matrix generated by vector $\boldsymbol{b}$ and $\mathbf{Tr}[A]$ to denote the trace of matrix A.

A *Markov chain* is a sequence of random variables taking value in the *finite* set $\{1,\dots,n\}$ with the Markov property, namely that, the future state depends only on the present state; see [23], [26] for more details. Let $X_k \in \{1,\dots,n\}$ denote the location of a random walker at time $k \in \{0,1,2,\dots\}$. We are now ready to summarize some terminology and results for Markov chains. (1) A Markov chain is *time-homogeneous* if $\mathbb{P}[X_{n+1} = j | X_n = i] = \mathbb{P}[X_n = j | X_{n-1} = i] = p_{ij}$, where $P \in \mathbb{R}^{n\times n}$ is the transition matrix of the Markov chain. (2) The vector $\boldsymbol{\pi} \in \mathbb{R}^{n\times 1}$ is a *stationary distribution* of P if $\sum_{i=1}^{n} \boldsymbol{\pi}_i = 1$, $0 \le \boldsymbol{\pi}_i \le 1$ for all $i \in \{1,\dots,n\}$ and $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$. (3) A time-homogeneous Markov chain is said to be *reversible* if $\boldsymbol{\pi}_i p_{ij} = \boldsymbol{\pi}_j p_{ji}$, for all $i,j \in \{1,\dots,n\}$. For reversible Markov chains, $\boldsymbol{\pi}$ is always a steady state distribution. (4) A Markov chain is *irreducible* if there exists a t such that for all $i,j \in \{1,\dots,n\}, (P^t)_{ij} > 0$. (5) If the Markov chain is *irreducible*, then there is a unique stationary distribution $\boldsymbol{\pi}$, and the corresponding eigenvalues of the transition matrix, λ_i for $i \in \{1,\dots,n\}$, are such that $\lambda_1 = 1$, $|\lambda_i| \le 1$ and $\lambda_i \ne 1$ for $i \in \{2,\dots,n\}$. For further details on irreducible matrices and about results (4) and (5) see [33, Chapter 8]. In this paper we consider finite irreducible time-homogeneous Markov chains.

II. The Kemeny constant of a Markov chain and its minimization

Consider a undirected weighted graph $\mathcal{G} = (\mathcal{V}, E, P)$ with node set $\mathcal{V} := \{1,\dots,n\}$, edge set $E \subset \mathcal{V}\times\mathcal{V}$, and weight matrix $P = [p_{ij}]$ with the property that $p_{ij} \ge 0$ if $(i,j) \in E$ and $p_{ij} = 0$ otherwise. We interpret the weight of edge (i,j) as the probability of moving along that edge. Therefore, element p_{ij} in the matrix represents the probability with which the random walk visits node j from node i. Throughout this document we assume that the underlying undirected graph $(\mathcal{V}, E)$ associated to the transition probabilities P is connected.

In this section we look into a discrete-time random walk defined by a Markov chain on a graph $\mathcal{G}$. At each time step (*hop*) of the random walk we move to a new node or stay at the current node according to the transition probabilities described by a transition matrix P as discussed above. We do this with three objectives in mind. The first objective is to analyze the random walk and characterize the average visit time between nodes in the graph. The second objective is to minimize the average visit time between any two nodes and the final is to achieve a long term (infinite horizon) visit frequency $\boldsymbol{\pi}_i$ at node i. Here, the frequency $\boldsymbol{\pi}_i$ is the ratio of visits to node i divided by the total number of visits to all nodes in the graph. Throughout the paper, we describe the random walk using realizations of a Markov chain with transition matrix $P = [p_{ij}]$.

A. The mean first passage time for a weighted graph

Let $X_k \in \{1, \dots, n\}$ denote the location of the random walker at time $k \in \{0, 1, 2, \dots\}$. For any two nodes $i, j \in \{1, \dots, n\}$, the *first passage time from i to j*, denoted by T_{ij}, is the first time that the random walker starting at node i at time 0 reaches node j, that is,

$$T_{ij} = \min\{k \geq 1 \mid X_k = j \text{ given that } X_0 = i\}.$$

It is convenient to introduce the shorthand $m_{ij} = \mathbb{E}[T_{ij}]$, and to define the *mean first passage time matrix M* to have entries m_{ij}, for $i, j \in \{1, \dots, n\}$. The *mean first passage time from start node i*, denoted by $\boldsymbol{k}_i$, is the expected first passage time from node i to a random node selected according to the stationary distribution $\boldsymbol{\pi}$, i.e.,

$$\boldsymbol{k}_i = \sum_{j=1}^{n} m_{ij} \boldsymbol{\pi}_j.$$

It is well known [25] that the mean first passage time from a start node is independent of the start node, that is, $\boldsymbol{k}_i = \boldsymbol{k}_j$ for all $i, j \in \{1, \dots, n\}$. Accordingly, we let $K = \boldsymbol{k}_i$, for all $i \in \{1, \dots, n\}$, denote the *mean first passage time*, also known as the Kemeny constant, of the Markov chain.

Next, we provide formulas for these quantities. By definition, the first passage time from i to j satisfies the recursive formula:

$$T_{ij} = \begin{cases} 1, & \text{with probability } p_{ij}, \\ T_{kj} + 1, & \text{with probability } p_{ik}, k \neq j. \end{cases}$$

Taking the expectation, we compute

$$m_{ij} = p_{ij} + \sum_{k=1, k\neq j}^{n} p_{ik}(m_{kj} + 1) = 1 + \sum_{k=1, k\neq j}^{n} p_{ik} m_{kj},$$

or in matrix notation,

$$(I - P)M = \mathbb{1}_n \mathbb{1}_n^T - P M_{\mathrm{d}}, \qquad (1)$$

where P is the transition matrix of the Markov chain. If the Markov chain is irreducible with stationary distribution $\boldsymbol{\pi}$, then one can show $M_{\mathrm{d}} = \operatorname{diag}[\{1/\boldsymbol{\pi}_1, \dots, 1/\boldsymbol{\pi}_n\}]$, and

$$\boldsymbol{\pi}^T M \boldsymbol{\pi} = \sum_{i=1}^{n} \boldsymbol{\pi}_i \sum_{j=1}^{n} \boldsymbol{\pi}_j m_{ij} = \sum_{i=1}^{n} \boldsymbol{\pi}_i \boldsymbol{k}_i = K.$$

Clearly, the Kemeny constant can be written as the function $P \mapsto K(P)$, however, to ease notation we simply write K and use $K(P)$ only when we wish to emphasize the constant's dependence on P.

The Kemeny constant $K = \boldsymbol{\pi}^T M \boldsymbol{\pi}$ can be computed from equation (1) or can be expressed as a function of the eigenvalues of the transition matrix P as is stated in the following theorem [25].

Theorem 1: (Kemeny constant of an irreducible Markov chain): Consider a Markov chain with an irreducible transition matrix P with eigenvalues $\lambda_1 = 1$ and λ_i, $i \in \{2, \dots, n\}$. The Kemeny constant of the Markov chain is given by

$$K = 1 + \sum_{i=2}^{n} \frac{1}{1 - \lambda_i}.$$

Using Theorem 1, we derive the following equivalent expression for *reversible* Markov chains in terms of the trace of a symmetric positive definite matrix. Before stating our result, we first introduce some notation. Given a stationary distribution vector $\boldsymbol{\pi} \in \mathbb{R}^{n\times 1}$ for a Markov chain with transition matrix $P \in \mathbb{R}^{n\times n}$, we define the matrix $\Pi \in \mathbb{R}^{n\times n}$ as $\Pi = \operatorname{diag}[\boldsymbol{\pi}]$ and the vector $\boldsymbol{q} \in \mathbb{R}^{n\times 1}$ as $\boldsymbol{q}^T = (\sqrt{\boldsymbol{\pi}_1}, \dots, \sqrt{\boldsymbol{\pi}_n})$. We are now ready to state our first result.

Theorem 2: (Kemeny constant of a reversible irreducible Markov chain): The Kemeny constant of a reversible irreducible Markov chain with transition matrix P and stationary distribution $\boldsymbol{\pi}$ is given by

$$K = \mathbf{Tr}\left[(I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)^{-1}\right]. \qquad (2)$$

Proof: We start by noting that P is an irreducible row-stochastic matrix therefore the eigenvalues of P are $\{\lambda_1 = 1, \lambda_2, \dots, \lambda_n\}$, where $|\lambda_i| \leq 1$ and $\lambda_i \neq 1$ for $i \in \{2, \dots, n\}$. It follows that the eigenvalues of $(I - P)$ are $\{0, 1 - \lambda_2, \dots, 1 - \lambda_n\}$. Since P is irreducible and reversible, there exists a stationary distribution $\boldsymbol{\pi} \in \mathbb{R}^n_{>0}$ implying Π is invertible and that $\Pi^{1/2}(I - P)\Pi^{-1/2} = I - \Pi^{1/2} P \Pi^{-1/2}$ is symmetric. It can easily be verified that $I - P$ and $I - \Pi^{1/2} P \Pi^{-1/2}$ have the same eigenvalues and that $\boldsymbol{q}$ is the eigenvector associated with the eigenvalue at 0. Next, notice the matrix $(I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)$ is symmetric and that $(I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)\boldsymbol{q} = \boldsymbol{q}$. Therefore, $(I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)$ has an eigenvalue at 1 associated with the vector $\boldsymbol{q}$. Since $(I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)$ is symmetric, the eigenvectors corresponding to different eigenvalues are orthogonal; implying for eigenvector $\boldsymbol{v} \neq \boldsymbol{q}$ that $(I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)\boldsymbol{v} = (I - \Pi^{1/2} P \Pi^{-1/2})\boldsymbol{v}$ since the eigenvalue at 1 is simple. Therefore, the eigenvalues of $(I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)$ are $\{1, 1 - \lambda_2, \dots, 1 - \lambda_n\}$. Thus, $K = \mathbf{Tr}\left[(I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)^{-1}\right] = 1 + 1/(1 - \lambda_2) + \dots + 1/(1 - \lambda_n) = K$. ∎

Given the above result, we are now ready to state our first problem of interest.

Problem 1: (Optimizing the Kemeny constant of a reversible Markov chain): Given the stationary distribution $\boldsymbol{\pi}$ and graph $\mathcal{G}$ with vertex set $\mathcal{V}$ and edge set E, determine the transition probabilities $P = [p_{ij}]$ solving:

$$\begin{aligned} \text{minimize} \quad & \mathbf{Tr}\left[(I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)^{-1}\right] \\ \text{subject to} \quad & \sum_{j=1}^{n} p_{ij} = 1, \text{ for each } i \in \{1, \ldots, n\} \\ & \boldsymbol{\pi}_i p_{ij} = \boldsymbol{\pi}_j p_{ji}, \text{ for each } (i,j) \in E \\ & 0 \leq p_{ij} \leq 1, \text{ for each } (i,j) \in E \\ & p_{ij} = 0, \text{ for each } (i,j) \notin E. \end{aligned} \tag{3}$$

Remark 3: All feasible solutions P to Problem 1 are inherently irreducible transition matrices: when P is not irreducible, the matrix $(I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)$ is not invertible. Moreover, a feasible point always exists since the Metropolis-Hastings algorithm applied to any irreducible transition matrix associated with $\mathcal{G}$, generates a reversible transition matrix which is irreducible and satisfies the stationary distribution constraint [21].

The following theorem establishes the convexity of the Kemeny constant for transition matrices with fixed stationary distribution.

Theorem 4 (Convexity of Problem 1): Let $\mathcal{P}_{\boldsymbol{\pi}}$ denote the set of matrices associated to irreducible reversible Markov chains with stationary distribution $\boldsymbol{\pi}$. Then, $\mathcal{P}_{\boldsymbol{\pi}}$ is a convex set and $P \mapsto K(P)$ is a convex function over $\mathcal{P}_{\boldsymbol{\pi}}$.

Proof: Let $\mathcal{S}$ denote the set of symmetric positive definite matrices, for any stationary distribution $\boldsymbol{\pi} \in \mathbb{R}^n_{>0}$, denote the set $\mathcal{S}_{\mathcal{P},\boldsymbol{\pi}} := \{I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T \mid P \in \mathcal{P}_{\boldsymbol{\pi}}\}$. We begin by showing that $\mathcal{P}_{\boldsymbol{\pi}}$ is a convex set. Let $P_1, P_2 \in \mathcal{P}_{\boldsymbol{\pi}}$, then $\mathcal{P}_{\boldsymbol{\pi}}$ is convex if for an arbitrary $\theta \in [0, 1]$ that

$$\theta P_1 + (1 - \theta)P_2 = P_3 \in \mathcal{P}_{\boldsymbol{\pi}}. \tag{4}$$

Pre and post multiplying (4) by $\Pi^{1/2}$ and $\Pi^{-1/2}$, respectively, we have that $\theta\Pi^{1/2}P_1\Pi^{-1/2} + (1-\theta)\Pi^{1/2}P_2\Pi^{-1/2} = \Pi^{1/2}P_3\Pi^{-1/2}$. Then $\Pi^{1/2}P_3\Pi^{-1/2}$ is symmetric since $\Pi^{1/2}P_1\Pi^{-1/2}$ and $\Pi^{1/2}P_2\Pi^{-1/2}$ are symmetric. Pre multiplying (4) by $\boldsymbol{\pi}^T$ we easily verify that the stationary distribution P_3 is $\boldsymbol{\pi}^T$ and similarly, post multiplying by $\mathbb{1}_n$ verifies that P_3 is row stochastic. Finally taking the left hand side of (4) to the n-th power gives $(\theta P_1 + (1-\theta)P_2)^n = \theta^n P_1^n + (1-\theta)^n P_2^n + \zeta$, where ζ denotes the sum of all other terms in the expansion and has the property $\zeta_{ij} \geq 0$ for all i, j since P_1 and P_2 are non-negative element-wise matrices. Moreover from irreducibility, there exists a sufficiently large N such that for $n > N$, $(P_1^n)_{ij} > 0$ and $(P_2^n)_{ij} > 0$ for all i, j, which implies $(P_3^n)_{ij} > 0$, therefore $P_3 \in \mathcal{P}_{\boldsymbol{\pi}}$ and $\mathcal{P}_{\boldsymbol{\pi}}$ is convex.

Next we show that $\mathcal{S}_{\mathcal{P},\boldsymbol{\pi}} \subset \mathcal{S}$. From the proof of Theorem 2 we have for $P \in \mathcal{P}_{\boldsymbol{\pi}}$ that $I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T$ has eigenvalues $\{1, 1 - \lambda_2, \ldots, 1 - \lambda_n\}$, where λ_i for $i \in \{1, \ldots, n\}$ are the eigenvalues of P, where $\lambda_i \leq |1|$ for all i and $\lambda_i \neq 1$ for $i \in \{2, \ldots, n\}$. Therefore, all eigenvalues of $I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T$ are strictly greater than zero. Finally, since P is reversible $\Pi^{1/2}P\Pi^{-1/2}$ is symmetric implying $(I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)^T = I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T$ and so $\mathcal{S}_{\mathcal{P},\boldsymbol{\pi}} \subset \mathcal{S}$.

Finally, define the mapping $g : \mathcal{P}_{\boldsymbol{\pi}} \mapsto \mathcal{S}_{\mathcal{P},\boldsymbol{\pi}}$ by $g(X) = I - \Pi^{1/2}X\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T$. This is an affine mapping from the convex set $\mathcal{P}_{\boldsymbol{\pi}}$ to a subset of $\mathcal{S}$. From [18] we know that $\mathbf{Tr}[X^{-1}]$ is convex for $X \in \mathcal{S}$, therefore the composition with the affine mapping $g : \mathcal{P}_{\boldsymbol{\pi}} \mapsto \mathcal{S}_{\mathcal{P},\boldsymbol{\pi}} \subset \mathcal{S}$, $\mathbf{Tr}[g(X)^{-1}]$ is also convex [8, Chapter 3.2.2]. ∎

Problem 1 includes constraints on the stationary distribution of the transition matrix, a notion which has not been looked at in the literature before. The author of [28] provides bounds to determine the set of transition matrices such that K is minimized and [25] gives special matrices for which the optimal Kemeny constant can be found, but these are all approached for the general setting with no constraint on the actual stationary distribution. In real-world settings, constraints on the stationary distribution are important and have many practical interpretations. For example, it is often desirable to visit certain regions more frequently than other based on each region's relative importance.

B. SDP framework for optimizing the Kemeny constant

Here we show how Problem 1 can be equivalently rewritten as an SDP by introducing a symmetric slack matrix.

Problem 2: (Optimizing the Kemeny constant of a reversible Markov chain (SDP)): Given the stationary distribution $\boldsymbol{\pi}$ and graph $\mathcal{G}$ with vertex set $\mathcal{V}$ and edge set E, determine $X = [x_{ij}]$ and the transition probabilities $P = [p_{ij}]$ solving:

$$\begin{aligned} \text{minimize} \quad & \mathbf{Tr}[X] \\ \text{subject to} \quad & \\ & \begin{bmatrix} I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T & I \\ I & X \end{bmatrix} \succeq 0 \\ & \sum_{j=1}^{n} p_{ij} = 1, \text{ for each } i \in \{1, \ldots, n\} \\ & \boldsymbol{\pi}_i p_{ij} = \boldsymbol{\pi}_j p_{ji}, \text{ for each } (i,j) \in E \\ & 0 \leq p_{ij} \leq 1, \text{ for each } (i,j) \in E \\ & p_{ij} = 0, \text{ for each } (i,j) \notin E. \end{aligned}$$

The first inequality constraint in Problem 2 represents a linear matrix inequality (LMI) and denotes that the matrix is positive semidefinite. Since the matrix in the LMI has off-diagonal entries equal to the identity matrix, it holds true if and only if X is positive definite and $(I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T) - X^{-1}$ is positive semidefinite [2, Theorem 1]. This implies $(I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)$ is positive definite and that $X \succeq (I - \Pi^{1/2}P\Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)^{-1}$. Therefore, the SDP given by Problem 2 minimizes the Kemeny constant.

III. The weighted Kemeny constant of Markov chain and its minimization

In most practical applications, distance/time traveled and service costs/times are important factors in the model of the system. We incorporate these concepts by allowing for an additional set of weighted edges in our graph in addition to the edge weights which describe the transition probabilities. Such a system can be represented by the doubly-weighted graph $\mathcal{G} = (\mathcal{V}, E, P, W)$, where $W = [\omega_{ij}]$ is a weight matrix with the properties that: if $(i,i) \in E$, then $\omega_{ii} \geq 0$; if $(i,j) \in E$, $i \neq j$, then $\omega_{ij} > 0$; and if $(i,j) \notin E$, then $\omega_{ij} = 0$. The weighted adjacency matrix $P = [p_{ij}]$ has the

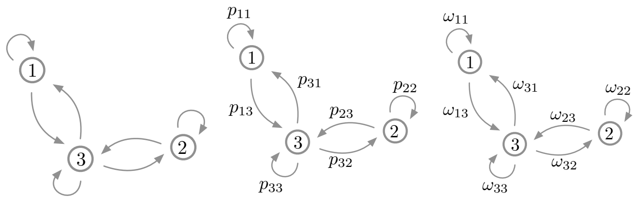

Fig. 1. Example of a doubly-weighted graph $\mathcal{G} = (\mathcal{V}, E, P, W)$ with three nodes: (a) shows the edge set, E, allowed for the graph with three nodes, (b) shows the probabilities, p_{ij} to move along each edge, and (c) shows the time (i.e., distance traveled), ω_{ij} to move along each edge.

same interpretation as before as an irreducible row-stochastic transition matrix P which governs the random walk on the graph. An example of a doubly-weighted graph is shown in Figure 1. In the following, we will interpret ω_{ij}, $i \neq j$ as the time to travel between two nodes, i and j, in the graph and ω_{ii} as the service time at node i. We discuss another motivating example and interpretation for ω_{ij} in a later section.

Recall that $X_k = i$ denotes that the random walker is at node i at time k. If a sample trajectory of the random walk is $X_0 = i$, $X_1 = j$, $X_2 = k$, then the time instant at which a random walker arrives in state X_2 is $\omega_{ij} + \omega_{jk}$. Thus the time interval between two consecutive steps of this random walk depends on the weighted adjacency matrix, W, of the graph and is not constant.

In the following analysis, we look at several characterization and optimization objectives: The first involves extending the notion of the Kemeny constant to doubly-weighted graphs and providing a detailed characterization of this extension. The second involves the minimization of the mean first passage time of a doubly-weighted graph and the third involves characterization and minimization of the mean time to execute a single *hop*. The first and second objectives are motivated by the need to minimize visit times to nodes in the graph, and the third is motivated by the desire to minimize resource consumption when moving between nodes. We seek to design transition matrices P with stationary distribution $\boldsymbol{\pi}$ which optimize each problem. We start with the first objective.

A. The mean first passage time for a doubly-weighted graph

The mean first passage time for the Markov chain on a weighted graph $\mathcal{G} = (\mathcal{V}, E, P)$ by definition, is simply its Kemeny constant. Recall that the mean first passage time for node i, defined by $\boldsymbol{k}_i$, is determined by taking the expectation over the first passage times m_{ij}, from node i to all other nodes j. We present an analogous notion of the first passage time between two nodes on a doubly-weighted graph. We start with defining the first passage time random variable for a random walk on a doubly-weighted graph and provide a recursive formulation for its expectation.

As in Section II-A, for any two nodes $i, j \in \{1, \dots, n\}$, the *first passage time from i to j* is the first time that the random walker starting at node i at time 0 reaches node j, that is,

$$T_{ij} = \min\Big\{ \sum_{n=0}^{k-1} w_{X_n, X_{n+1}}, \text{ for } k \geq 1 \mid X_k = j \quad \text{given that } X_0 = i \Big\}.$$

Lemma 5: (First passage time for a doubly-weighted graph): Let $n_{ij} = \mathbb{E}[T_{ij}]$ denote the mean first passage time to go from i to j for a graph with weight matrix W and transition matrix P. Then

$$n_{ij} = p_{ij}(\omega_{ij}) + \sum_{k \neq j} p_{ik}(n_{kj} + \omega_{ik}), \tag{5}$$

or, in matrix notation,

$$(I - P)N = (P \circ W)\mathbb{1}_n \mathbb{1}_n^T - P N_{\mathrm{d}}, \tag{6}$$

where $(P \circ W)$ is the element-wise product between P and W and where $N_{\mathrm{d}} = [\delta_{ij} n_{ij}]$.

Proof: By its definition, the first passage time satisfies the recursive formula:

$$T_{ij} = \begin{cases} \omega_{ij}, & \text{with probability } p_{ij}, \\ \omega_{ik} + T_{kj}, & \text{with probability } p_{ik}, k \neq j. \end{cases} \tag{7}$$

Therefore, the results follows from taking the expectation:

$$\mathbb{E}[T_{ij}] = \omega_{ij} p_{ij} + \sum_{k \neq j} p_{ik}(\mathbb{E}[T_{kj}] + \omega_{ik}).$$

■

The matrix N, which we call the *mean first passage time matrix for a doubly-weighted graph* thus satisfies an equation similar to (1) for the passage time matrix M of a graph with a single weight matrix, the transition matrix P. In fact, we see that equation (6) is equivalent to (1) when $w_{ij} = 1$ for all $(i, j) \in E$ (i.e., for an unweighted graph).

The random variable tracking the time interval between consecutive visits to a node has been referred to as the *refresh time* of that node [35] and n_{ii} is the expected value of the refresh time for the random walk. We now obtain a relation between $\boldsymbol{\pi}$ and the refresh times n_{ii}.

Theorem 6 (Refresh times for doubly-weighted graphs): Consider a Markov chain on a doubly-weighted graph $\mathcal{G} = (\mathcal{V}, E, P, W)$ with stationary distribution $\boldsymbol{\pi}$ and associated mean first passage time matrix N. The refresh time for node i is given by $n_{ii} = (\boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n)/\boldsymbol{\pi}_i$, implying that

$$N_{\mathrm{d}} = \boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n M_{\mathrm{d}}.$$

Proof: The stationary distribution of the transition matrix P is $\boldsymbol{\pi} \in \mathbb{R}^{n \times 1}$. Therefore, premultiplying equation (6) by $\boldsymbol{\pi}^T$, we obtain

$$0 = \boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n \mathbb{1}_n^T - \boldsymbol{\pi}^T N_{\mathrm{d}},$$

where the left hand side of equation (6) is zero since $\pi^T(I - P) = \pi^T - \pi^T = 0$. Now we have that $(\boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n)\mathbb{1}_n^T = \boldsymbol{\pi}^T N_{\mathrm{d}}$. Since N_d is a diagonal matrix and $\boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n$ is a scalar, we get that $\boldsymbol{\pi}_i n_{ii} = \boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n$. Thus, dividing by $\boldsymbol{\pi}_i$ we have that $n_{ii} = \boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n / \boldsymbol{\pi}_i$, and in matrix form $N_d = \boldsymbol{\pi}^T(P \circ W)\mathbb{1}_n \mathrm{diag}(\{1/\boldsymbol{\pi}_1, \ldots, 1\ \boldsymbol{\pi}_n\}) = \boldsymbol{\pi}^T(P \circ W)\mathbb{1}_n M_d$. ■

This theorem implies that the refresh time n_{ii} of the random walk is directly proportional to the visit frequencies $1/\boldsymbol{\pi}_i$. Therefore, the relative visit frequencies of one node compared to another are not a function of the weight matrix W and only depend on the stationary distribution of the transition matrix P.

We now investigate the properties of the mean first passage time of the weighted random walk. The mean first passage time for a doubly-weighted graph $\mathcal{G} = (\mathcal{V}, E, P, W)$ with associated passage times matrix N is given by $K_W = \boldsymbol{\pi}^T \boldsymbol{k}_W$, where $\boldsymbol{k}_W = N\boldsymbol{\pi}$ is the vector of *first passage times* and the $i-$th entry $\boldsymbol{k}_{W,i}$ in $\boldsymbol{k}_W$ denotes the mean time to go from i to any other node. We refer to the mean first passage time for a doubly-weighted graph, K_W, as the *weighted Kemeny constant*. We now provide an analytic expression for the vector $\boldsymbol{k}_W \in \mathbb{R}^{n\times 1}$.

Lemma 7: (First passage times for a doubly-weighted graph): Given a Markov chain on a doubly-weighted graph $\mathcal{G} = (\mathcal{V}, E, P, W)$ with stationary distribution $\boldsymbol{\pi}$ and associated first passage time matrix N, the following equality holds:

$$(I - P)\boldsymbol{k}_W = (P \circ W)\mathbb{1}_n - \mathbb{1}_n \boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n, \qquad (8)$$

where $\boldsymbol{k}_W = N\boldsymbol{\pi}$.

Proof: Post multiplying equation (6) on both sides by $\boldsymbol{\pi}$ gives

$$\begin{aligned}(I - P)N\boldsymbol{\pi} &= (P \circ W)\mathbb{1}_n \mathbb{1}_n^T \boldsymbol{\pi} - P N_{\mathrm{d}} \boldsymbol{\pi},\\ (I - P)\boldsymbol{k}_W &= (P \circ W)\mathbb{1}_n - P(\boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n)\mathbb{1}_n\\ &= (P \circ W)\mathbb{1}_n - \mathbb{1}_n \boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n.\end{aligned}$$

■

The right hand side of (8) gives the insight that, in general, $\boldsymbol{k}_{W,i} \neq \boldsymbol{k}_{W,j}$ on the doubly-weighted graph, unlike the counterpart for the single-weighted graph (where $\boldsymbol{k}_i = K$ for all $i \in \{1, \ldots, n\}$). Interestingly enough however, there does exist a relation between the weighted Kemeny constant K_W and the Kemeny constant K as is stated in the following theorem, whose proof is postponed to the Appendix.

Theorem 8: (Weighted Kemeny constant of a Markov chain): For the doubly-weighted graph $\mathcal{G} = (\mathcal{V}, E, P, W)$, the weighted Kemeny constant K_W of the Markov chain is given by

$$K_W = \boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n K, \qquad (9)$$

where K is the Kemeny constant associated with the irreducible transition matrix P with stationary distribution $\boldsymbol{\pi}$.

Remark 9: The expected number of *hops* to go from one node to another in a Markov chain with transition matrix P is its Kemeny constant. The expected distance travelled (and hence time taken) executing one *hop* is $\sum_i \boldsymbol{\pi}_i \sum_j p_{ij}\omega_{ij} = \boldsymbol{\pi}(P \circ W)\mathbb{1}_n$. Hence, it is consistent with intuition that the expected time to travel from one node to another should be $K\boldsymbol{\pi}^T(P \circ W)\mathbb{1}_n$ as is formally shown in the appendix.

Given the above results, we are now ready to state another problem of interest.

Problem 3: (Optimizing the weighted Kemeny constant of a reversible Markov chain): Given the stationary distribution $\boldsymbol{\pi}$ and graph $\mathcal{G}$ with vertex set $\mathcal{V}$, edge set E and weight matrix W, determine the transition probabilities $P = [p_{ij}]$ solving:

$$\begin{aligned}&\text{minimize} \\ &\qquad (\boldsymbol{\pi}^T (P \circ W)\mathbb{1}_n)\left(\mathbf{Tr}\left[(I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)^{-1}\right]\right)\\ &\text{subject to} \quad \sum_{j=1}^n p_{ij} = 1, \text{ for each } i \in \{1, \ldots, n\}\\ &\qquad\qquad \boldsymbol{\pi}_i p_{ij} = \boldsymbol{\pi}_j p_{ji}, \text{ for each } (i,j) \in E\\ &\qquad\qquad 0 \le p_{ij} \le 1, \text{ for each } (i,j) \in E\\ &\qquad\qquad p_{ij} = 0, \text{ for each } (i,j) \notin E.\end{aligned}$$

The following theorem establishes the convexity of the weighted Kemeny constant for transition matrices with fixed stationary distribution.

Theorem 10 (Convexity of Problem 3): Given the graph $\mathcal{G}$ with vertex set $\mathcal{V}$, edge set E and weight matrix W, let $\mathcal{P}_{\mathcal{G},\boldsymbol{\pi}}$ denote the set of matrices associated with $\mathcal{G}$ that are irreducible reversible Markov chains with stationary distribution $\boldsymbol{\pi}$. Then, $\mathcal{P}_{\mathcal{G},\boldsymbol{\pi}}$ is a convex set and $P \mapsto \boldsymbol{\pi}^T(P \circ W)\mathbb{1}_n K(P)$ is a convex function over $\mathcal{P}_{\mathcal{G},\boldsymbol{\pi}}$.

Proof: Let $\mathcal{S}$ denote the set of symmetric positive definite matrices, for any stationary distribution $\boldsymbol{\pi} \in \mathbb{R}^n_{>0}$, denote the set $\mathcal{S}_{\mathcal{G},\mathcal{P},\boldsymbol{\pi}} := \{I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T \mid P \in \mathcal{P}_{\mathcal{G},\boldsymbol{\pi}}\}$. The proof of convexity of the set $\mathcal{P}_{\mathcal{G},\boldsymbol{\pi}}$ is similar to that of the proof of $\mathcal{P}_{\boldsymbol{\pi}}$ in Theorem 4 and so is omitted for brevity. Then from the proof of Theorem 4 we know there exists an affine mapping $g(X) : \mathcal{P}_{\mathcal{G},\boldsymbol{\pi}} \mapsto \mathcal{S}_{\mathcal{G},\mathcal{P},\boldsymbol{\pi}}$ given by $g(X) = I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T$. We know from [18] that $f(X) = \mathbf{Tr}[X^{-1}]$ is convex, therefore the perspective function $h(X, t) = \{tf(X/t) \mid t > 0\}$ is also convex [8, Chapter 3.2.6]. Moreover the composition of $h(X, t)$ with the affine mapping $g(X)$, $h(g(X), t)$ is also convex. Let $t = (\boldsymbol{\pi}^T (X \circ W)\mathbb{1}_n)^{1/2}$, and notice that $\boldsymbol{\pi}^T (X \circ W)\mathbb{1}_n > 0$ for $X \in \mathcal{P}_{\mathcal{G},\boldsymbol{\pi}}$ and therefore $t > 0$. Also notice that for a constant $k \in \mathbb{R}^n_{>0}$ and matrix $X \in \mathbb{R}^{n\times n}$ that $\mathbf{Tr}[(\frac{X}{k})^{-1}] = k\mathbf{Tr}[X^{-1}]$. Then $h(g(X), t) = t\mathbf{Tr}[(\frac{g(X)}{t})^{-1}] = t^2\mathbf{Tr}[(g(X)^{-1}] = \boldsymbol{\pi}^T (X \circ W)\mathbb{1}_n \mathbf{Tr}[(I - \Pi^{1/2} X \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)^{-1}]$ for $X \in \mathcal{P}_{\mathcal{G},\boldsymbol{\pi}}$. ■

B. SDP framework for optimizing the weighted Kemeny constant

In a similar fashion to Problem 1, we can formulate Problem 3 as an SDP by introducing the symmetric slack matrix $X \in \mathbb{R}^{n\times n}$ and the scalar variable t as is shown in the following.

Problem 4: (Optimizing the weighted Kemeny constant of a reversible Markov chain (SDP)): Given the stationary distribution $\boldsymbol{\pi}$ and graph $\mathcal{G}$ with vertex set $\mathcal{V}$, edge set E and weight matrix W, determine $Y = [y_{ij}]$, X and t solving:

$$\begin{aligned}
&\text{minimize} \quad \mathbf{Tr}[X] \\
&\text{subject to} \\
&\qquad \begin{bmatrix} t(I + \boldsymbol{q}\boldsymbol{q}^T) - \Pi^{1/2} Y \Pi^{-1/2} & I \\ I & X \end{bmatrix} \succeq 0 \\
&\qquad \sum_{j=1}^{n} y_{ij} = t, \text{ for each } i \in \{1, \ldots, n\} \\
&\qquad \boldsymbol{\pi}_i y_{ij} = \boldsymbol{\pi}_j y_{ji}, \text{ for each } (i,j) \in E \\
&\qquad 0 \le y_{ij} \le t, \text{ for each } (i,j) \in E \\
&\qquad y_{ij} = 0, \text{ for each } (i,j) \notin E \\
&\qquad \boldsymbol{\pi}^T (Y \circ W) \mathbb{1}_n = 1 \\
&\qquad t \ge 0.
\end{aligned}$$

Then, the transition matrix P is given by $P = Y/t$.

As in Problem 2, the first inequality constraint in Problem 4 represents an LMI. Before noting when the LMI holds, first note that the constraints in Problem 4 imply that $Pt = Y$ and that $t = \frac{1}{\boldsymbol{\pi}^T (P \circ W) \mathbb{1}_n}$. Hence, using a similar argument as in Problem 2, the LMI constraint holds true if and only if $X \succeq \boldsymbol{\pi}^T (P \circ W) \mathbb{1}_n (I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)^{-1}$ where X and $\pi^T (P \circ W) \mathbb{1}_n (I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)^{-1}$ are both positive definite, and therefore the SDP given by Problem 4 minimizes the weighted Kemeny constant.

C. Minimizing single hop distance

We now look at the objective of minimizing the mean time for a single *hop* of a random walk. At time k, let S_{ij} be the time required to transition from i to j in a single hop along an edge of length ω_{ij}. Then,

$$\begin{aligned}
\mathbb{E}[S] &= \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} S_{ij} \\
&= \sum_{i=1}^{n} \mathbb{P}[X_k = i] \sum_{j=1}^{n} \omega_{ij} \mathbb{P}[X_{k+1} = j] \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \boldsymbol{\pi}_i \omega_{ij} p_{ij} = \boldsymbol{\pi}^T (P \circ W) \mathbb{1}_n. \qquad (10)
\end{aligned}$$

The function $\boldsymbol{\pi}^T (P \circ W) \mathbb{1}_n$ is clearly convex in P. If one assumes that $\omega_{ii} = 0$ for all $i \in \{1, \ldots, n\}$, then minimizing (10) over P yields the trivial solution $P = I$. We can take into account both the single hop distance as well as the Kemeny constant to design a Markov chain as follows.

Problem 5: (Optimizing Kemeny constant and mean distance): Given the stationary distribution $\boldsymbol{\pi}$ and graph $\mathcal{G}$ with vertex set $\mathcal{V}$, edge set E and weight matrix W, and given user specified constant $\alpha \in [0, 1]$, determine the transition probabilities $P = [p_{ij}]$ solving:

$$\begin{aligned}
\text{minimize} \quad & \alpha \mathbf{Tr}\left[(I - \Pi^{1/2} P \Pi^{-1/2} + \boldsymbol{q}\boldsymbol{q}^T)^{-1}\right] \\
& + (1-\alpha) \boldsymbol{\pi}^T (P \circ W) \mathbb{1}_n \\
\text{subject to} \quad & \sum_{j=1}^{n} p_{ij} = 1, \text{ for each } i \in \{1, \ldots, n\} \\
& \boldsymbol{\pi}_i p_{ij} = \boldsymbol{\pi}_j p_{ji}, \text{ for each } (i,j) \in E \\
& 0 \le p_{ij} \le 1, \text{ for each } (i,j) \in E \\
& p_{ij} = 0, \text{ for each } (i,j) \notin E.
\end{aligned} \qquad (11)$$

This problem is convex since the sum of two convex problems is convex, moreover, it can be extended to an SDP utilizing the LMI defined in Problem 2. In the context where $\omega_{ii} = 0$ for all $i \in \{1, \ldots, n\}$, varying the parameter α can be used to control the connectivity of the Markov chain; the choice $\alpha = 1$ ensures connectivity, and the choice $\alpha = 0$ minimizes the single hop distance while making the graph disconnected.

IV. Applications of the mean first passage time to surveillance

The results on mean first passage time for doubly-weighted graphs (i.e., the weighted Kemeny constant) presented in this work provide a general framework which can potentially be applied to the analysis and design in a myriad of fields. We focus on one application in particular; the intruder detection and surveillance problem. We look at two variations of this problem:

a) Scenario I

In practical stochastic intruder detection and surveillance scenarios, there is often a desire to surveil some regions more than others (i.e.,have a pre- specified stationary distribution) while simultaneously minimizing the time any one region has to wait before it is serviced. For this setup, in every step of the random walk, the agent must move to a new region and execute its surveillance task. Assuming we are working on a doubly-weighted graph described by $\mathcal{G} = (\mathcal{V}, E, P, W)$, let us also assume there is a fixed persistent intruder in the environment and it takes s_i time for an agent to determine if the intruder is in region $i \in \mathcal{V}$. Denote the time to move from region i to region j by d_{ij}, where we can assume, without loss of generality, that $d_{ii} = 0$. Then, we can define the weight corresponding to the edge from i to j as $\omega_{ij} = d_{ij} + s_j$. In this scenario we wish to minimize the expected time to capture the persistent intruder when no prior knowledge of their position is known.

b) Scenario II

In this scenario we consider the intruder detection problem and adopt a similar setup to Scenario I, however, we now assume a set of intruders are distributed throughout the environment. Each intruder performs a malicious activity in its host region for a fixed duration of time, which we call the *intruder life-time*, followed instantaneously by another intruder. The intruder is caught only if the agent is in the same region as the intruder for any duration of the *intruder life-time*. For simplicity only a single intruder appears at a time.

In the following subsections we analyze the performance of various stochastic surveillance policies as applied to Scenario I and Scenario II described above. More specifically, we gauge the performance of other well-known Markov chain design algorithms against the algorithms presented in this paper.

A. Optimal strategy for Scenario I

In the context of Scenario I the weighted Kemeny constant of the agent's transition matrix, P, corresponds to the average time it takes to capture an intruder regardless of where the agent and intruder are in the environment. Therefore by definition of the Kemeny constant, we have the following immediate corollary for Scenario I.

Corollary 11 (Optimal surveillance and service strategy): The transition matrix P which has minimal mean first passage time is the optimal strategy for the intruder detection problem described by *Scenario I.*

This tells us that if we restrict ourselves to reversible Markov chains, then not only is the chain with minimal mean first passage time optimal, but given the results from Section II and III, we can also optimally design this chain.

B. Numerical analysis of Scenario II

In Scenario II the transition matrix with minimum mean first passage time is not guaranteed to be the optimal policy, and thus to gauge its performance compared to other policies we analyze both homogeneous (uniform service/travel times) and heterogeneous environment cases. To compare performance we generate a random walk for the environment using the Metropolis-Hastings, fastest mixing Markov chain (FMMC) [7], and Kemeny constant algorithms. While game theoretic frameworks [11], [6] also generate stochastic policies, they are based on assumptions on the intruder behavior. We avoid such assumptions here and, therefore, omit them from our comparative analysis.

We first look at the homogeneous case which is described by the discretized environment shown in Figure 2. We assume that a single surveillance agent executes a random walk in the environment, spending 1 time unit in each region, and that the agent transitions between two connected regions instantaneously. Also, we assume a uniform stationary distribution on the transition matrix (each node in the region must be visited with equal likelihood). The Markov chain generated by the Metropolis-Hastings algorithm is generated by applying the algorithm to the random walk described by $p_{ij} = 1/d_i$ for $i \neq j$ and $p_{ij} = 0$ for $i = j$, where d_i is the degree of a node i (excluding self-loops) [21]. The *intruder life-time* is set to 66 time units and 500 intruders appear per simulation run (the sequence in which the intruders appear is determined before each simulation run), for a total simulation run of $33,000$ time units. As stated in the scenario description, the intruder is caught if the surveillance agent is in the same region as the intruder for any portion of the *intruder life-time*. Table I summarizes the statistical performance of each algorithm after 200 runs of the simulation and justifies our use of the Kemeny constant algorithm as a valid surveillance strategy; the Kemeny constant algorithm captures intruders more frequently than the other two algorithms, and its worst case performance is still better than the worst case performance of the other two algorithms. Although results for an intruder life-time of only 66 time units are presented here, we have found that the Kemeny constant algorithm always outperforms the other two algorithms or is equivalent; the algorithms become equivalent in the limiting case, when the intruder life-times are so low that no intruder can be caught, or when the intruder-life times are so large that the intruder is always caught.

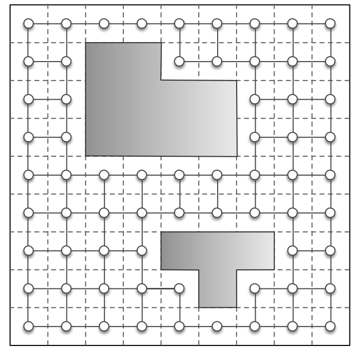

Fig. 2. Environment with two obstacle represented by an unweighted graph.

Algorithm	**Min**	**Mean**	**Max**	**StdDev**	K
Kemeny constant	26.6	32.4	38.2	2.1	207
FMMC	24.6	29.8	34.4	1.9	236
Metropolis-Hastings	24.8	31.1	37	2.1	231

TABLE I
STATISTICS ON THE PERCENTAGE OF INTRUDERS CAUGHT IN 200 SIMULATION RUNS FOR THE ENVIRONMENT IN FIG. 2.

For the heterogeneous case, we work with the environment shown in Figure 3. In this environment the time taken by the agent to travel an edge is no longer instantaneous and has travel weights as shown in the figure. Once in a new region, the agent is required to spend 1 time unit examining the region for malicious activities. We again assume that each node in the region must be visited with equal likelihood. We again also assume an intruder is caught if the surveillance agent is in the same region as a intruder for any portion of the *intruder life-time*, but now set the *intruder life-time* to 11 time units with a intruder appearing 500 times (total of 5500 time units per simulation run). Since the design of the FMMC and Metropolis-Hastings algorithms do not inherently account for non-uniform travel and service times, we also compare the performance of the random walk generated by the weighted Kemeny constant algorithm against the random walk generated by solving Problem 5 with $\alpha = 0.1$ (smaller α emphasizes minimizing the length of the average edge traveled in the graph). Table II summarizes the statistical performance of each algorithm after 200 runs of the simulation. The weighted Kemeny constant algorithm's performance compared to the FMMC and Metropolis-Hastings stochastic policies in this scenario is significantly better than what was seen in the first scenario. We also note that the random walk policy determined by solving Problem 5 performs comparably to the weighted Kemeny constant policy. This is to be expected since the Metropolis-Hastings and FMMC stochastic policies do not account for heterogeneous travel/service times on the graph.

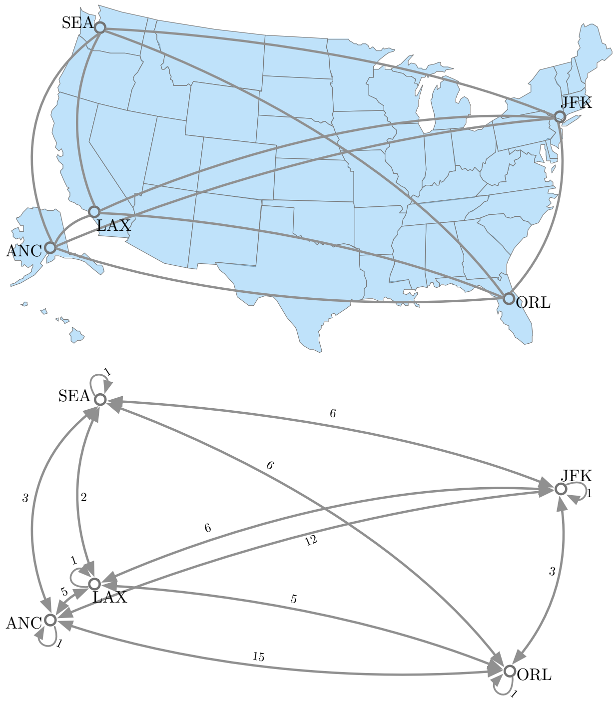

Fig. 3. Various airport hub locations (top), and the corresponding weight map (bottom). Edge weights between two hubs account for travel time between hubs plus required service time once at hub. Self loops have no travel time so encompass only service time required at hub.

To get a better understanding of each algorithm's performance in this intruder scenario, the simulation is run for different intruder life-times, the results of which can be seen in Figure 4. There are several key items worth noting from the simulation. First, we see that the weighted Kemeny constant algorithm significantly outperforms the other algorithms for a large range of intruder life-times. This matches our intuition since the algorithm inherently minimizes average travel time between nodes. Second, notice that the Markov chain generated by solving Problem 5 (with $\alpha = 0.1$) performs well for small intruder life-times but its performance plateaus quickly. This is due to the fact that the transition matrix generated by solving Problem 5 forces agents to stay at a given node rather than jump nodes; as one would suspect, once intruder life-times increase, a strategy which places emphasis on an agent that moves between regions will begin to perform relatively better. Finally, observe that as intruder life-time increases, the algorithms' capture rates start to converge. As in the homogeneous case, this is due to the fact that once the intruder's life-time is long enough, the agent will almost surely reach the intruder regardless of the policy it follows.

Algorithm	**Min**	**Mean**	**Max**	**StdDev**	K_W
Weighted Kemeny	44	50.1	56	2.2	19.5
Kemeny+Mean Dist.	40.6	47.1	53	2.2	23.1
FMMC	29.8	35.4	40.4	2.2	26.2
Metropolis-Hastings	30.4	36	41.6	2.1	26.5

TABLE II
STATISTICS ON THE PERCENTAGE OF INTRUDERS CAUGHT IN 200 SIMULATION RUNS FOR THE ENVIRONMENT IN FIG. 3.

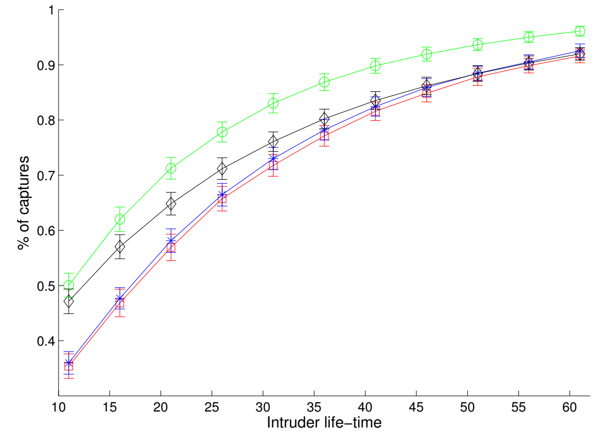

Fig. 4. Percentage of intruders detected for varying intruder life-times by a surveillance agent executing a random walk according to the Markov chain generated by the mean first-passage time algorithm (circle), FMMC algorithm (square), M-H algorithm (asterisk), and the Markov chain generated by solving Problem 5 (diamond). Average points and standard deviation error bars are taken over 200 runs, where the intruder appears 500 times for each run.

To solve for the Markov chains with minimal Kemeny constant (Problem 2 and Problem 4) and with fastest mixing rate, we use CVX, a Matlab-based package for convex programs [20]. The execution time to solve each Markov chain for the examples described above takes on the order of a couple seconds using a computer with a 1.3 GHz processor.

V. CONCLUSIONS

We have studied the problem of how to optimally design a Markov chain which minimizes the mean first passage time to go from one region to any other region in a connected environment. We have presented the first formulation of the mean first passage time for a doubly-weighted graph, which we refer to as the weighted Kemeny constant, and have also provided a provably correct convex formulation for the minimization of both the Kemeny constant and the weighted Kemeny constant. Finally, we have shown that both problems can be written as SDPs and, moreover, have demonstrated the effectiveness of using a Markov chain with minimal mean first passage time as a surveillance policy as compared to other well-known Markov chain policies.

This work leaves open various directions for further research. First, we designed surveillance policy only for single agent systems and it would be of practical interest to consider the case where there are multiple agents: [37], [9], [3], [10] are examples of work in this direction. Second, it would be useful to understand bounds on the design of of the mean first passage time for general graph topologies. Finally, we treat only the optimization of the transition matrix of the graph. It would be of interest to study how we can optimize the weight matrix W in conjunction with the transition matrix. This can have the interpretation of optimizing the "capacity" or "resistance" of the graph, a topic in optimization which is of independent interest [18].

Appendix A
Proof of Theorem 8

Proof of Theorem 8: Let $\beta = \boldsymbol{\pi}^T(P \circ W)\mathbb{1}_n$, then from Theorem 6 we have that N_d from (6) can be written as βM_d. Now from Theorem 13 the general solution to (6) is

$$N = G((P \circ W)\mathbb{1}_n\mathbb{1}_n^T - \beta P M_\mathrm{d}) + (I - G(I-P))U, \quad (12)$$

where G is a generalized inverse of $(I-P)$ (see Theorem 15) and U is an arbitrary matrix as long as the consistency condition

$$\big(I - (I-P)G\big)\big((P \circ W)\mathbb{1}_n\mathbb{1}_n^T - \beta P M_\mathrm{d}\big) = 0 \quad (13)$$

holds. Substituting (18) from Lemma 16 in for $(I-P)G$ in (13) gives that

$$\begin{aligned}
&\big(I - (I-P)G\big)\big((P \circ W)\mathbb{1}_n\mathbb{1}_n^T - \beta P M_\mathrm{d}\big) \\
&\quad = \frac{\boldsymbol{t}\boldsymbol{\pi}^T}{\boldsymbol{\pi}^T\boldsymbol{t}}((P \circ W)\mathbb{1}_n\mathbb{1}_n^T - \beta P M_\mathrm{d}), \\
&\quad = \frac{\boldsymbol{t}}{\boldsymbol{\pi}^T\boldsymbol{t}}(\boldsymbol{\pi}^T(P \circ W)\mathbb{1}_n\mathbb{1}_n^T - \beta\boldsymbol{\pi}^T P M_\mathrm{d}), \\
&\quad = \frac{\boldsymbol{t}}{\boldsymbol{\pi}^T\boldsymbol{t}}(\beta\mathbb{1}_n^T - \beta\mathbb{1}_n^T) = 0,
\end{aligned}$$

and so we have that the system of equations is consistent. This implies we can design U to reduce (12). We start by seeing how the second term in (12) can be reduced. Using (19) from Lemma 16 we have that $(I - G(I-P))U = \frac{\mathbb{1}_n\boldsymbol{u}^T}{\boldsymbol{u}^T\mathbb{1}_n}U = \mathbb{1}_n\boldsymbol{h}^T$, where $\boldsymbol{h}^T = \frac{\boldsymbol{u}^T U}{\boldsymbol{u}^T\mathbb{1}_n}$. Hence, we can re-write (12) as

$$N = G((P \circ W)\mathbb{1}_n\mathbb{1}_n^T - \beta P M_\mathrm{d}) + \mathbb{1}_n\boldsymbol{h}^T, \quad (14)$$

designing U reduces to designing the n elements of $\boldsymbol{h}$. Let $H = \mathrm{diag}[\boldsymbol{h}]$, then $\mathbb{1}_n\boldsymbol{h}^T = \mathbb{1}_n\mathbb{1}_n^T H$. Also, let $\Xi = \mathbb{1}_n\boldsymbol{\pi}^T$, where $\Xi_\mathrm{d} = \mathrm{diag}[\boldsymbol{\pi}]$. Utilizing these expressions in (14) and taking the diagonal elements gives

$$\begin{aligned}
&\big(N = G((P \circ W)\mathbb{1}_n\mathbb{1}_n^T - \beta P M_\mathrm{d}) + \mathbb{1}_n\mathbb{1}_n^T H\big)_\mathrm{d}, \\
&\implies \beta M_\mathrm{d} = (G(P \circ W)\Xi)_\mathrm{d} M_\mathrm{d} - \beta(GP)_\mathrm{d} M_\mathrm{d} + H, \\
&\implies H = \beta M_\mathrm{d} - (G(P \circ W)\Xi)_\mathrm{d} M_\mathrm{d} + \beta(GP)_\mathrm{d} M_\mathrm{d},
\end{aligned}$$

where we use Lemma 14 to get the initial diagonal reduction. Substituting the expression for H into (14), and recalling that $\mathbb{1}_n\boldsymbol{h}^T = \mathbb{1}_n\mathbb{1}_n^T H$ gives

$$\begin{aligned}
N = \big(&G(P \circ W)\Xi - \mathbb{1}_n\mathbb{1}_n^T(G(P \circ W)\Xi)_\mathrm{d} \\
&+ \beta(\mathbb{1}_n\mathbb{1}_n^T(GP)_\mathrm{d} - GP + \mathbb{1}_n\mathbb{1}_n^T)\big)M_\mathrm{d},
\end{aligned} \quad (15)$$

where we use the fact that $\mathbb{1}_n\mathbb{1}_n^T = \Xi M_\mathrm{d}$. Now from (19) we have that $I - G - GP = \frac{\mathbb{1}_n\boldsymbol{u}^T}{\boldsymbol{u}^T\mathbb{1}_n}$. Notice that $\mathbb{1}_n\mathbb{1}_n^T(I - G - GP)_\mathrm{d} = \mathbb{1}_n\mathbb{1}_n^T(\frac{\mathbb{1}_n\boldsymbol{u}^T}{\boldsymbol{u}^T\mathbb{1}_n})_\mathrm{d} = \frac{\mathbb{1}_n\boldsymbol{u}^T}{\boldsymbol{u}^T\mathbb{1}_n}$ and so this implies that $\mathbb{1}_n\mathbb{1}_n^T - \mathbb{1}_n\mathbb{1}_n^T G_\mathrm{d} + \mathbb{1}_n\mathbb{1}_n^T(GP)_\mathrm{d} = I - G + GP$, which implies that $\mathbb{1}_n\mathbb{1}_n^T + \mathbb{1}_n\mathbb{1}_n^T(GP)_\mathrm{d} - GP = I - G + \mathbb{1}_n\mathbb{1}_n^T G_\mathrm{d}$. Substituting this into (15) gives the following reduced form.

$$\begin{aligned}
N = \big(&G(P \circ W)\Xi - \mathbb{1}_n\mathbb{1}_n^T(G(P \circ W)\Xi)_\mathrm{d} \\
&+ \beta(\mathbb{1}_n\mathbb{1}_n^T G_\mathrm{d} + I - G)\big)M_\mathrm{d}.
\end{aligned} \quad (16)$$

Observing the definition of the generalized inverse, G, given by Theorem 15 part (ii) and recalling that $\Xi = \mathbb{1}_n\boldsymbol{\pi}^T$, we can rewrite the first term on the right hand side of (16) as $G(P \circ W)\Xi = (I - P + \boldsymbol{t}\boldsymbol{u}^T)^{-1}(P \circ W)\mathbb{1}_n\boldsymbol{\pi}^T$. Substituting (20) in for the right hand side with $\boldsymbol{t} = (P \circ W)\mathbb{1}_n$ gives $G(P \circ W)\Xi = \frac{\mathbb{1}_n\boldsymbol{\pi}^T}{\boldsymbol{u}^T\mathbb{1}_n} = \frac{1}{\boldsymbol{u}^T\mathbb{1}_n}\Xi$, and so $\mathbb{1}_n\mathbb{1}_n^T(G(P \circ W)\Xi)_\mathrm{d} = \mathbb{1}_n\mathbb{1}_n^T(\frac{1}{\boldsymbol{u}^T\mathbb{1}_n}\Xi)_\mathrm{d} = \frac{1}{\boldsymbol{u}^T\mathbb{1}_n}\Xi = G(P \circ W)\Xi$. Therefore, the first two terms in (16) cancel giving the equality

$$N = \beta(\mathbb{1}_n\mathbb{1}_n^T G_\mathrm{d} + I - G)M_\mathrm{d}. \quad (17)$$

We have already defined $\boldsymbol{t}$ in the generalized inverse G but not $\boldsymbol{u}$. Let $\boldsymbol{u} = \boldsymbol{\pi}$ and multiply the right hand side of (17) by $\boldsymbol{\pi}$ and the left hand side by $\boldsymbol{\pi}^T$. Utilizing equality (21) from Lemma 16 gives

$$\begin{aligned}
\boldsymbol{\pi}^T N \boldsymbol{\pi} &= \boldsymbol{\pi}^T\beta(\mathbb{1}_n\mathbb{1}_n^T G_\mathrm{d} + I - G)M_\mathrm{d}\boldsymbol{\pi} \\
&= \beta(\mathbb{1}_n^T G_\mathrm{d} + \boldsymbol{\pi}^T - \frac{\boldsymbol{\pi}^T}{\beta})\mathbb{1}_n \\
&= \beta(\mathbb{1}_n^T G_\mathrm{d}\mathbb{1}_n + 1 - \frac{1}{\beta}) \\
&= \beta(\mathbf{Tr}[G] + 1) - 1.
\end{aligned}$$

Noting that the eigenvalue at 1 for an irreducible row-stochastic matrix is unique, it can be easily verified using the orthogonality property of left and right eigenvectors that the eigenvalues of G^{-1} are $\bar{\lambda}_i = (1 - \lambda_i)$ for $i \in \{2, \ldots, n\}$, where λ_i are eigenvalues of P and $\lambda_i \neq 1$. Therefore, it only remains to find $\bar{\lambda}_1$. Taking the trace of G^{-1} gives $\mathbf{Tr}[I - P + (P \circ W)\mathbb{1}_n\boldsymbol{\pi}^T] = \mathbf{Tr}[I - P] + \mathbf{Tr}[(P \circ W)\mathbb{1}_n\boldsymbol{\pi}^T] = \sum_{i=1}^n(1 - \lambda_i) + \boldsymbol{\pi}^T(P \circ W)\mathbb{1}_n$, which implies that $\bar{\lambda}_1 = \boldsymbol{\pi}^T(P \circ W)\mathbb{1}_n = \beta$. Therefore, $\beta(\mathbf{Tr}[G] + 1) - 1 = \beta(\frac{1}{\beta} + \sum_{i=2}^n \frac{1}{1-\lambda_i} + 1) - 1 = \beta(1 + \sum_{i=2}^n \frac{1}{1-\lambda_i})$. ■

Appendix B
Supplemental Material

For completeness, we include the following results which are needed in the proof of Theorem 8. We begin with some standard results on generalized inverses. For more details refer to [23, Chapter 4] or [22] .

Definition 12 (Generalized inverse): A *generalized inverse* of an $m \times n$ matrix A is defined as any $n \times m$ matrix A^- that has the property

$$AA^-A = A.$$

It should be noted that a generalized inverse always exists, although it is not always unique. However, for non-singular matrices the generalized inverse is unique and corresponds to the usual notion of a matrix inverse. The following theorems summarize practical considerations when working with generalized inverses.

Theorem 13: The equation $A\boldsymbol{x} = \boldsymbol{b}$ admits a solution if and only if every generalized inverse A^- satisfies

$$AA^-\boldsymbol{b} = \boldsymbol{b}.$$

Then, we say $A\boldsymbol{x} = \boldsymbol{b}$ is *consistent* and all its general solutions are given by

$$\boldsymbol{x} = A^-\boldsymbol{b} + (A^-A - I)\boldsymbol{z},$$

where $\boldsymbol{z}$ is an arbitrary vector. Moreover, a necessary and sufficient condition for the equation $AX = C$ to be consistent is that $(I - AA^{-})C = 0$, in which case the general solution is given by

$$X = A^{-}C + (I - A^{-}A)U,$$

where U is an arbitrary matrix.

The next two results come from [24, Chapter 7].

Lemma 14 (Diagonal matrix properties): For $\boldsymbol{\pi}$ with positive non-zero elements, let $\mathbb{1}_n\boldsymbol{\pi}^T = \Xi$, where $\Xi_{\mathrm{d}} = \mathrm{diag}[\boldsymbol{\pi}]$. Also, let Λ be any diagonal matrix, X any square matrix of same dimensions as Λ, and $D = (\Xi_{\mathrm{d}})^{-1}$, then

(i.) $(X\Lambda)_{\mathrm{d}} = (X_{\mathrm{d}})\Lambda$, and
(ii.) $(X\mathbb{1}_n\mathbb{1}_n^T)_{\mathrm{d}} = (X\Xi)_{\mathrm{d}}D$, and
(iii.) $\mathbb{1}_n\mathbb{1}_n^T\Xi_{\mathrm{d}} = \Xi$.

Theorem 15 (Generalized inverse of $I - P$): Let $P \in \mathbb{R}^{n\times n}$ be the transition matrix of a irreducible Markov Chain with stationary distribution $\boldsymbol{\pi}$. Let $\boldsymbol{u}, \boldsymbol{t} \in \mathbb{R}^n$ be any vectors such that $\boldsymbol{u}^T\mathbb{1}_n \neq 0$ and $\boldsymbol{\pi}^T\boldsymbol{t} \neq 0$, then

(i.) $I - P + \boldsymbol{t}\boldsymbol{u}^T$ is nonsingular, and
(ii.) $(I - P + \boldsymbol{t}\boldsymbol{u}^T)^{-1}$ is a generalized inverse of $I - P$.

Lemma 16 (Properties of the generalized inverse of $I - P$): Let $\mathcal{G} = (\mathcal{V}, E, P, W)$ be a doubly-weighted graph with associated weight matrix W and irreducible transition matrix P with stationary distribution $\boldsymbol{\pi}$. Also let $G = (I-P+\boldsymbol{t}\boldsymbol{u}^T)^{-1}$ denote the generalized inverse of $(I - P)$, then the following relations hold.

$$(I - P)G = I - \frac{\boldsymbol{t}\boldsymbol{\pi}^T}{\boldsymbol{\pi}^T\boldsymbol{t}}, \tag{18}$$

$$G(I - P) = I - \frac{\mathbb{1}_n\boldsymbol{u}^T}{\boldsymbol{u}^T\mathbb{1}_n}, \text{ and} \tag{19}$$

$$\frac{\mathbb{1}_n}{\boldsymbol{u}^T\mathbb{1}_n} = G\boldsymbol{t}. \tag{20}$$

If $\boldsymbol{t} = (P \circ W)\mathbb{1}_n$ and $\boldsymbol{u}^T = \boldsymbol{\pi}^T$ then

$$\boldsymbol{\pi}^T G = \frac{\boldsymbol{\pi}^T}{\boldsymbol{\pi}^T(P \circ W)\mathbb{1}_n} \tag{21}$$

Proof: First, notice that $(I-P+\boldsymbol{t}\boldsymbol{u}^T)(I-P+\boldsymbol{t}\boldsymbol{u}^T)^{-1} = I$ implies that

$$(I - P)G = I - \boldsymbol{t}\boldsymbol{u}^T G. \tag{22}$$

Multiplying both sides on the left by $\boldsymbol{\pi}^T$ and noting that $\boldsymbol{\pi}^T(I-P) = 0$ gives that $\boldsymbol{\pi}^T = (\boldsymbol{\pi}^T\boldsymbol{t})\boldsymbol{u}^T G$. Dividing through by $(\boldsymbol{\pi}^T\boldsymbol{t})$ gives

$$\frac{\boldsymbol{\pi}^T}{\boldsymbol{\pi}^T\boldsymbol{t}} = \boldsymbol{u}^T G, \tag{23}$$

and substituting (23) into (22) gives (18).

Following a similar procedure as before with $(I - P + \boldsymbol{t}\boldsymbol{u}^T)^{-1}(I - P + \boldsymbol{t}\boldsymbol{u}^T) = I$, where we now multiply both sides on the right by $\mathbb{1}_n$ and note that $(I - P)\mathbb{1}_n = 0$ results in (20), which after appropriate substitution gives (19).

For the proof of equality (21), first we check that $\boldsymbol{t} = (P \circ W)\mathbb{1}_n$ and $\boldsymbol{u}^T = \boldsymbol{\pi}^T$ satisfy the conditions of Theorem 15. The definition of W guarantees that $P \circ W$ has at least one non-zero element which implies $\boldsymbol{\pi}^T\boldsymbol{t} = \boldsymbol{\pi}^T(P \circ W)\mathbb{1}_n \neq 0$. Also, $u^T\mathbb{1}_n = \boldsymbol{\pi}^T\mathbb{1}_n = 1$. Now substituting $\boldsymbol{u}$ and $\boldsymbol{t}$ into (23) gives (21). ■

References

[1] P. Agharkar, R. Patel, and F. Bullo. Robotic surveillance and Markov chains with minimal first passage time. In *IEEE Conf. on Decision and Control*, pages 6603–6608, Los Angeles, CA, USA, December 2014.

[2] A. Albert. Conditions for positive and nonnegative definiteness in terms of pseudoinverses. *SIAM Journal on Applied Mathematics*, 17(2):434–440, 1969.

[3] D. A. Anisi, P. Ögren, and X. Hu. Cooperative minimum time surveillance with multiple ground vehicles. *IEEE Transactions on Automatic Control*, 55(12):2679–2691, 2010.

[4] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic. The role of social networks in information diffusion. In *Int. Conference on World Wide Web*, pages 519–528, Lyon, France, 2012.

[5] T. H. Blackwell and J. S. Kaufman. Response time effectiveness: Comparison of response time and survival in an urban emergency medical services system. *Academic Emergency Medicine*, 9(4):288–295, 2002.

[6] B. Bošanský, V. Lisý, M. Jakob, and M. Pěchouček. Computing time-dependent policies for patrolling games with mobile targets. In *International Conference on Autonomous Agents and Multiagent Systems*, pages 989–996, 2011.

[7] S. Boyd, P. Diaconis, and L. Xiao. Fastest mixing Markov chain on a graph. *SIAM Review*, 46(4):667–689, 2004.

[8] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[9] G. Cannata and A. Sgorbissa. A minimalist algorithm for multirobot continuous coverage. *IEEE Transactions on Robotics*, 27(2):297–312, 2011.

[10] D. W. Casbeer, D. B. Kingston, R. W. Beard, T. W. Mclain, S.-M. Li, and R. Mehra. Cooperative forest fire surveillance using a team of small unmanned air vehicles. *International Journal of Systems Sciences*, 37(6):351–360, 2006.

[11] L. Chen and J. Leneutre. A game theoretical framework on intrusion detection in heterogeneous networks. *IEEE Transactions on Information Forensics and Security*, 4(2):165–178, June 2009.

[12] J. Clark and R. Fierro. Mobile robotic sensors for perimeter detection and tracking. *ISA Transactions*, 46(1):3–13, 2007.

[13] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *Annals of Applied Probability*, 1(1):36–61, 1991.

[14] P. G. Doyle and J. L. Snell. *Random Walks and Electric Networks*. Mathematical Association of America, 1984.

[15] W. Ellens, F. M. Spieksma, P. Van Mieghem, A. Jamakovic, and R. E. Kooij. Effective graph resistance. *Linear Algebra and its Applications*, 435(10):2491–2506, 2011.

[16] Y. Elmaliach, N. Agmon, and G. A. Kaminka. Multi-robot area patrol under frequency constraints. *Annals of Mathematics and Artificial Intelligence*, 57(3-4):293–320, 2009.

[17] Y. Elmaliach, A. Shiloni, and G. A. Kaminka. A realistic model of frequency-based multi-robot polyline patrolling. In *International Conference on Autonomous Agents*, pages 63–70, Estoril, Portugal, May 2008.

[18] A. Ghosh, S. Boyd, and A. Saberi. Minimizing effective resistance of a graph. *SIAM Review*, 50(1):37–66, 2008.

[19] J. Grace and J. Baillieul. Stochastic strategies for autonomous robotic surveillance. In *IEEE Conf. on Decision and Control and European Control Conference*, pages 2200–2205, Seville, Spain, December 2005.

[20] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx, October 2014.

[21] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[22] J. J. Hunter. Generalized inverses and their application to applied probability problems. *Linear Algebra and its Applications*, 45:157–198, 1982.

[23] J. J. Hunter. *Mathematical Techniques of Applied Probability*, volume 1 of *Discrete Time Models: Basic Theory*. Academic Press, 1983.

[24] J. J. Hunter. *Mathematical Techniques of Applied Probability*, volume 2 of *Discrete Time Models: Techniques and Applications*. Academic Press, 1983.

[25] J. J. Hunter. The role of Kemeny's constant in properties of Markov chains. *Communications in Statistics - Theory and Methods*, 43(7):1309–1321, 2014.

[26] J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. Springer, 1976.

[27] D. B. Kingston, R. W. Beard, and R. S. Holt. Decentralized perimeter surveillance using a team of UAVs. *IEEE Transactions on Robotics*, 24(6):1394–1404, 2008.

[28] S. Kirkland. Fastest expected time to mixing for a Markov chain on a directed graph. *Linear Algebra and its Applications*, 433(11-12):1988–1996, 2010.

[29] D. J. Klein and M. Randić. Resistance distance. *Journal of Mathematical Chemistry*, 12(1):81–95, 1993.

[30] M. Levene and G. Loizou. Kemeny's constant and the random surfer. *The American Mathematical Monthly*, 109(8):741–745, 2002.

[31] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2009.

[32] Y. Y. Li and L. E. Parker. Intruder detection using a wireless sensor network with an intelligent mobile robot response. In *IEEE SoutheastCon*, pages 37–42, Huntsville, AL, USA, April 2008.

[33] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, 2001.

[34] J. L. Palacios. Closed-form formulas for Kirchhoff index. *International Journal of Quantum Chemistry*, 81(2):135–140, 2001.

[35] F. Pasqualetti, A. Franchi, and F. Bullo. On cooperative patrolling: Optimal trajectories, complexity analysis and approximation algorithms. *IEEE Transactions on Robotics*, 28(3):592–606, 2012.

[36] A. Patcha and J. Park. A game theoretic approach to modeling intrusion detection in mobile ad hoc networks. In *Information Assurance Workshop. Proceedings from the Fifth Annual IEEE SMC*, pages 280–284, June 2004.

[37] K. Srivastava, D. M. Stipanovìc, and M. W. Spong. On a stochastic robotic surveillance problem. In *IEEE Conf. on Decision and Control*, pages 8567–8574, Shanghai, China, December 2009.

[38] V. Srivastava, F. Pasqualetti, and F. Bullo. Stochastic surveillance strategies for spatial quickest detection. *International Journal of Robotics Research*, 32(12):1438–1458, 2013.

[39] S. Susca, S. Martínez, and F. Bullo. Monitoring environmental boundaries with a robotic sensor network. *IEEE Transactions on Control Systems Technology*, 16(2):288–296, 2008.

[40] X. Wu and Z. Liu. How community structure influences epidemic spread in social networks. *Physica A: Statistical Mechanics and its Applications*, 387(2):623–630, 2008.

[41] P. R. Wurman, R. D'Andrea, and M. Mountz. Coordinating hundreds of cooperative, autonomous vehicles in warehouses. *AI Magazine*, 29(1):9–20, 2008.

Rushabh Patel (IEEE S'12) received the B.S. and M.S. degrees in aerospace engineering from California Polytechnic State University, San Luis Obispo, California, in 2007. From 2007 through 2011 he worked as a performance analyst at Northrop Grumman Aerospace Systems. He is currently pursuing the Ph.D. degree in control and dynamical systems at the University of California, Santa Barbara. His research interests include coordination and surveillance in networked systems.

Pushkarini Agharakar (IEEE S'12) received the B.Tech. and M.Tech. degrees in aerospace engineering from Indian Institute of Technology, Madras, in 2011. She is currently pursuing the Ph.D. degree in mechanical engineering at the University of California at Santa Barbara. Her research interests include vehicle routing problems, robotic surveillance and motion coordination for robotic systems.

Francesco Bullo (IEEE S'95-M'99-SM'03-F'10) is a Professor with the Mechanical Engineering Department and the Center for Control, Dynamical Systems and Computation at the University of California, Santa Barbara. He was previously associated with the University of Padova, the California Institute of Technology and the University of Illinois at Urbana-Champaign. His main research interests are network systems and distributed control with application to robotic coordination, power grids and social networks. He is the coauthor of "Geometric Control of Mechanical Systems" (Springer, 2004, 0-387-22195-6) and "Distributed Control of Robotic Networks" (Princeton, 2009, 978-0-691-14195-4). He received the 2008 IEEE CSM Outstanding Paper Award, the 2010 Hugo Schuck Best Paper Award, the 2013 SIAG/CST Best Paper Prize, and the 2014 IFAC Automatica Best Paper Award. He has served on the editorial boards of IEEE Transactions on Automatic Control, ESAIM: Control, Optimization, and the Calculus of Variations, SIAM Journal of Control and Optimization, and Mathematics of Control, Signals, and Systems.