Ad Hoc Networks 17 (2014) 60–70

Contents lists available at ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

A stochastic process model of the hop count distribution in wireless sensor networks ☆

Steffen Beyme *, Cyril Leung

Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada

ARTICLE INFO

Article history:
Received 25 June 2013
Received in revised form 10 January 2014
Accepted 13 January 2014
Available online 24 January 2014

Keywords:
Wireless sensor network
Localization
Hop count distribution
Stochastic jump process
Maximum entropy
Poisson distribution

ABSTRACT

We consider target localization in randomly deployed multi-hop wireless sensor networks, where messages originating from a sensor node are broadcast by flooding and the node-to-node message delays are characterized by independent, exponential random variables. Using asymptotic results from first-passage percolation theory and a maximum entropy argument, we formulate a stochastic jump process to approximate the hop count of a message at distance r from the source node. The resulting marginal distribution of the process has the form of a translated Poisson distribution which characterizes observations reasonably well and whose parameters can be learnt, for example by maximum likelihood estimation. This result is important in Bayesian target localization, where mobile or stationary sinks of known position use the hop count conditioned on the Euclidean distance, to estimate the position of a sensor node or event within the network, based solely on observations of the hop count. For the target localization problem, simulation results show that the proposed model provides reasonably good performance, especially for densely connected networks.

1. Introduction

Target localization in wireless sensor networks (WSNs) is an active area of research with wide applicability. Due to power and interference constraints, the vast majority of WSNs convey messages via multiple hops from a source to one or several sinks, mobile or stationary. Localization techniques which exploit the information about the Euclidean distance from a sensor node, contained in the hop count of a message originating from that node, are referred to as *range-free* [1]. Range-free localization is applicable to networks of typically low-cost, low-power wireless sensor nodes without the hardware resources needed to accurately measure node positions, neighbor distances or angles (for example using GPS, time or angle of arrival). It is therefore an attractive approach in situations where a compromise is sought between localization accuracy on one hand, and cost, size and power efficiency on the other.

Various hop-count based localization techniques for WSNs have been proposed; for a survey see e.g. [1]. Relating hop count information to the Euclidean distance between sensor nodes, exemplified by the probability distribution of the hop count conditioned on distance, remains a challenging problem. Except in special cases, such as one-dimensional networks [2], only approximations can be obtained; such approximations are often in the form of recursions, which tend to be difficult to evaluate [3–5]. Moreover, the hop count depends on the chosen path from the source to the sink and is therefore a function of the routing method employed by the network. An approximate,

☆ This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada under Grant OGP0001731 and by the UBC PMC-Sierra Professorship in Networking and Communications. Portions of this paper were presented at the 26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Regina, SK, May 5–8, 2013.

* Corresponding author. Tel.: +1 6046889022.
E-mail addresses: steffenb@ece.ubc.ca (S. Beyme), cleung@ece.ubc.ca (C. Leung).

http://dx.doi.org/10.1016/j.adhoc.2014.01.006
1570-8705/

closed-form hop count distribution is proposed in [6] and evaluated for nearest, furthest and random neighbor routing, in which a forwarding node selects the next node from a neighborhood oriented towards the sink, where it is assumed that this neighborhood is not empty. Some of the existing localization algorithms, such as the DV-hop algorithm [7] and its variants, define the hop count between nodes as that of the *shortest path* [3–5]. Other localization algorithms use the hop count of a path established through *greedy forward routing* [8–11]. In most cases, the overhead incurred by routing is not negligible. Simpler alternatives may be needed when sensor nodes impose more severe complexity constraints.

This paper is motivated by the range-free target localization problem in networks of position-agnostic wireless sensor nodes, which broadcast messages using flooding under the assumption that node-to-node message delays can be characterized by independent, exponential random variables. This is a reasonable assumption in situations where sensor nodes enter a dormant state while harvesting energy from the environment and wake up at random times, or when the communication channel is unreliable and retransmissions are required. Under these conditions, a *first-passage path* emerges as the path of minimum passage time from a source to a sink. Networks of this type can be described in terms of *first-passage percolation* [12,13]. Localization of the source node may be performed by mobile or stationary sinks able to fuse hop count observations $Z \in \mathbb{N}$ to infer the location $X \in \mathbb{R}^2$ of the source node, where $p_X(x)$ denotes the *a priori* pdf of X. By Bayes' rule, the *a posteriori* pdf of the source location is $p_X(x|z) \sim p_Z(z|x)p_X(x)$, conditioned on observing the hop count z at the sink position. Knowledge of the observation model $p_Z(z|x)$, that is, the conditional pdf of the hop count, given the source location hypothesis x, is essential for the success of this approach, which may be complicated further by the presence of model parameters whose values are not known *a priori* and must be learnt on- or off-line. Bayesian localization involves a large number of numerical evaluations of the observation model, due to the typically large space of location hypotheses. This creates a need for observation models with low computational complexity, which may outweigh the need for high accuracy in some applications.

The *main contribution* of this paper is the formulation of a *jump stochastic process* whose marginal distribution has a simple analytical form, to model the hop count of the first-passage path from a source to a sink, which is at a distance r. In contrast to earlier works, which use geometric arguments to construct a hop count distribution, our approach utilizes the abstract model of a jump stochastic process and attempts to describe the hop count in its terms. Starting with a process of stationary increments satisfying a strong mixing condition, we make a simplifying independence assumption which allows the hop count to be modeled as a jump Lévy process with drift [14]. We show that, consistent with our assumptions about the hop count, the maximum entropy principle leads to the selection of a *translated Poisson distribution* as the marginal distribution of the hop count model process.

The paper is structured as follows. In Section 2, we review relevant concepts from stochastic geometry and first-passage percolation and introduce our network model. Our main result, the stochastic process $\{Z_r\}$ which models the observed hop count distribution at distance r from a source node, is derived in Section 3. We describe how the parameters of this process can be learnt using maximum likelihood estimation. In Section 4, we present simulation results which show that in sensor networks of the type considered in this paper, the marginal distribution of the model process approximates the empirical hop count quite well. Extending work presented in [15], we study the localization error due to the approximation by comparison with a ficticious, idealized network where observations are generated as independent draws from our model. Especially for more densely connected WSNs, we observe a good localization error performance. In the Appendices, we provide proofs of propositions used to derive our model, which were omitted in [15].

2. System model

The geometry of randomly deployed WSNs is commonly described by Gilbert's disk model [16]. Given a spatial Poisson point process $\mathcal{P}_\lambda = \{X_i : i \in \mathbb{N}\}$ of density λ on $\mathbb{R}^2$, two sensor nodes are said to be linked if they are within communication range R of each other. Gilbert's model induces a random geometric graph $G_{\lambda,R} = \{\mathcal{P}_\lambda, \mathcal{E}_\mathcal{R}\}$ with node set $\mathcal{P}_\lambda$ and edge set $\mathcal{E}_R$. The node density λ and the communication range R are related through the mean node degree $\delta = \pi\lambda R^2$, so that the graph can be defined equivalently in terms of a single parameter as G_δ. Without loss of generality, it is convenient to condition the graph on there being a node at the origin. By Slivnyak's Theorem, if we remove the point at the origin from consideration, we have a Poisson point process [17].

Of key interest in the study of G_δ is continuum percolation, that is, the conditions under which a cluster of infinitely many connected nodes emerges in such a graph. Percolation is widely known to exhibit a phase transition from a subcritical state characterized by the existence of only clusters with a finite number of nodes, to the formation of one infinite cluster with probability 1, as δ increases above a critical threshold δ_c. The *percolation probability* is defined as the probability that the origin belongs to the infinite cluster. It is zero for values of δ below the threshold, and an increasing function of δ above. Just above the threshold, the incipient infinite cluster has characteristics generally undesirable for communication networks, such as large minimum path lengths. Therefore, the mean node degree δ is generally chosen so that the network operates sufficiently deep in the supercritical regime.

The dissemination of messages in some broadcast WSNs has been described in terms of *first-passage percolation* [12,13,18,19], where every node forwards the broadcast message to all of its neighbors. We assume that the node-to-node message delays can be characterized by *independent, exponentially distributed random variables* with common means. This is a reasonable assumption when the nodes enter a dormant state while replenishing their energy storage from the environment, and wake up randomly. Messages convey the cumulative number of hops,