Routing for Enhancing Source-Location Privacy in Wireless Sensor Networks of Multiple Assets

Yeonghwan Tscha

Abstract: **In wireless sensor networks, a node that reports information gathered from adjacent assets should relay packets appropriately so that its location context is kept private, and thereby helping ensure the security of the assets that are being monitored. Unfortunately, existing routing methods that counter the local eavesdropping-based tracing deal with a single asset, and most of them suffer from the packet-delivery latency as they prefer to take a separate path of many hops for each packet being sent. In this paper, we propose a routing method, greedy perimeter stateless routing-based source-location privacy with crew size w (GSLP-w), that enhances location privacy of the packet-originating node (i.e., active source) in the presence of multiple assets. GSLP-w is a hybrid method, in which the next-hop node is chosen in one of four modes, namely *greedy*, *random*, *perimeter*, and *retreat* modes. Random forwarding brings the path diversity, while greedy forwarding refrains from taking an excessively long path and leads to convergence to the destination. Perimeter routing makes detours that avoid the nodes near assets so that they cannot be located by an adversary tracing up the route path. We study the performance of GSLP-w with respect to crew size w (the number of packets being sent per path) and the number of sources. GSLP-w is compared with phantom routing-single path (PR-SP), which is a notable routing method for source-location privacy and our simulation results show that improvements from the point of the ratio of safety period and delivery latency become significant as the number of source nodes increases.**

Index Terms: **Active/dormant source, local eavesdropping-based packet-tracing (passive attack), location privacy in wireless sensor networks, multiple assets, routing for source-location privacy.**

I. INTRODUCTION

Over the past decades, progress in wireless communications and micro-electro-mechanical system (MEMS) technologies made it feasible to construct a wireless sensor network that is composed of hundreds to thousands of low-cost sensor nodes [1]. The resource-constrained sensor nodes can sense, measure and gather information from the underlying environment, and they can transfer the collected data to a base station (or sink) in a multi-hop manner by wireless communications. Applications are mainly military target tracking and surveillance, natural disaster prediction, habitats monitoring and traffic monitoring. Timely information gathering and dissemination is crucial to the success of all applications. Various power-efficient mechanisms are also considered in all aspect of hardware platforms, operating systems, protocols, and application services.

Expected to be widely deployed in the near future, sensor networks are highly vulnerable to packet-tracing attacks. Due to the open nature of wireless communication, it may be easy for adversaries to eavesdrop or inject packets into the networks [2]. Many networks are often deployed in outdoor areas. Thus, attackers may break up or replace the nodes of the networks. Furthermore, there exist some applications that need to consider *location privacy* of communicating nodes. For instance, wireless sensor networks deployed in battlefields to support snipers or in natural habitats to monitor rare wildlife may strongly need to protect the locations of assets (i.e., soldiers or rare wildlife) against the adversaries (i.e., enemies or poachers) [3], [4].

It is shown in [5] that adversaries can easily analyze the base station centric traffic by rate-monitoring attacks and deduce the location of a base station. If an adversary is around a base station, then the adversary may be able to eavesdrop all incoming packets. By repeatedly taking such hop-by-hop tracing, the adversary can approach the packet-originating node and may finally identify it. One of the popular countermeasures on the routing level is to make it difficult for the adversary to trace its way back to the origin of communications (i.e., *source* [3], [6], [7]). The preferred strategies usually adopt random walks to make the paths more irregular and longer, as opposed to the conventional routing that seeks the shortest or lowest-cost paths. Generally, each packet is sent over a separate path for path diversity. The goal of such approach is to send more packets before the source is located by the adversary, where the number of the packets delivered to the sink is known as *safety period* [3], [6]. [5], [8] are studies to protect the location of the sink but share a similar idea. However, as the direction of packet-forwarding is identical to that of the tracing by the opponent, *fake-packet injections* are usually deployed to direct the adversary to a wrong place.

A common problem in [3], [6]–[8] is that their schemes introduce *long* latencies when transferring packets, as they prefer to deliver more packets via long paths.[1] Another problem results from the fact that they regard a *single* asset in the network.[2] Many of their routing methods force to send each packet over a *separate* path. In this paper, we propose a routing scheme that deals with these problems. We consider the wireless sensor network that comprises many assets and sources. Our routing strategy may not only increase location privacy of the active source but also protect the dormant sources against the packet-

Manuscript received April 14, 2009.

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2007-521-D00409).

Y. Tscha is with the School of Computer, Information, and Communication Engineering, Sangji University, Kangwon, Korea, email: yhtscha@sangji.ac.kr.

[1] The work in [9] proposes a routing technique to reduce the latency while increasing the source privacy, but it also introduces many problems, which will be discussed in the next section.

[2] In [10], a scheme to protect multiple assets against the global eavesdropper is proposed. Unfortunately, it requires all nodes in the network to send packets periodically as fake sources.

tracing attack in the networks of a large number of assets.[3] The proposed routing technique, greedy perimeter stateless routing-based source-location privacy with crew size w (GSLP-w), is an extension of GPSR [12], [13]. Taking merit of the localized and distributed geographic (location)-aware routing, the perimeter routing concept of GPSR is adopted to detour the dormant sources so that their locations cannot be exposed to an adversary. With the probabilistic combination of random and greedy forwarding, the length of the route path can be controlled to avoid the excessive latency. Allowing more than one packet to be sent over each path, it may further reduce the packet-delivery latency, yet increasing location privacy of the active source. Through simulations, we evaluate the performance of GSLP-w based on two criteria: Normalized safety period (NSP) and normalized delivery latency (NDL).[4] We choose phantom routing-single path (PR-SP) [3], [6] for comparison because it is a notable routing protocol for source-location privacy and, to some extent, it may protect the location of the active source in the presence of multiple assets.

The rest of this paper is organized as follows. The related work is reviewed in the next section. Section III describes network and threat models used in this paper. In Section IV, we propose our routing scheme GSLP-w. Motivations underlying our approach and the routing method are presented. Section V gives the performance evaluation of GSLP-w through simulations. Comparisons with PR-SP are made in terms of the number of sources. Further improvements of GSLP-w are addressed. Section VI concludes this paper.

II. RELATED WORK

Phantom routing (PR) paved the way toward the *routing for location privacy* in wireless sensor networks [3]. In PR, each packet first takes a random walk of some fixed number of hops (say, 10 or 15 hops) to make it harder for the local eavesdropper to trace the movements of packets. Then, packets are delivered in the routing phase by using the shortest path routing (PR-SP) or flooding (PR-flooding (PR-F)). The privacy strength is directly proportional to the length of the random walk: Longer the walk, better the strength. The route can vary with time, thereby meaning that an adversary can get pulled to another portion to the network that might not see future packets after the route switches. In the subsequent work [6], a formal model for location privacy was introduced and the effect of the source mobility on location privacy was studied. This is the first study on routing-level location privacy in wireless sensor networks. However, the long-distance random walk may lead to excessive delay in the packet-delivery. The paths made during the random walk phase may lead to back-and-forth or zigzag movements, frequently. On the other hand, Kamat *et al.* [14] analyzed temporal privacy in delay tolerant sensor networks and proposed adaptive buffering at intermediate nodes. Clearly, delaying packets can address location privacy by allowing the asset to move.

In [7], the self-adjusted directed random walk was studied to overcome the routing *holes* and to maintain randomness throughout the random walk. The performance has also been studied by considering the ratio of the random walk length to the distance between the source and the sink. In [14], the algorithm greedy random walk (GROW) that makes a two-way random walk by both the source and the sink was proposed under the global eavesdropper model of having multiple monitoring points. However, route paths are now longer in order to increase source-location privacy (i.e., safety period), thus the schemes also suffer from the excessive delivery latency.

Cyclic entrapment method (CEM) [9] is unique in the sense that it aims at reducing the packet-delivery latency, while enhancing source-location privacy. The shortest path is created between the active source and the sink to deliver data packets, but the nodes that are on the path are attached to each *loop* of fixed hops (for instance, 10 or 15 hops). When a packet is routed along the shortest path, it encounters one of these pre-configured loops. Then, the encountered loops are activated and begin cycling fake packets around the loops. This makes it harder for the adversary to distinguish packets, whether they are from the real source or not. Both the shortest latency and the long safety period can be attainable as long as the attacker is enticed into the trap loop. However, there are several critical problems in the CEM approach.

First, the loops may not repeatedly or successfully entice the adversary because, with the help of some cache that stores the history of locations visited, the opponent may perceive that it is following the fake packet by analyzing the correlation among the locations on the traced path. Second, all nodes in the network must set up their own loops, but only a few of them lying on the shortest path (i.e., loop activation nodes) are activated for enticing. Thus, creating loops by all nodes introduce too many packets into the network. Third, each loop must be at least of length k, but finding such a loop reduces to the k-longest path problem, which is widely known as NP-complete [15]. Fourth, when the adversary gets near the loop activation node, the fake packets should be timely ejected from the loop to successfully entice the adversary away. Thus, *the fake synchronization problem*[5] should be resolved properly. Due to such drawbacks, the loop mechanism of CEM seems to be beyond a practical use.

On the other hand, there are studies on *destination-location privacy* in wireless sensor networks. In [3], several schemes to hide the location of the base station against the traffic-analysis attacks are proposed. In contrast to the local eavesdropper model in this paper, the adversary model is global as it monitors and gathers the traffic at several nodes called aggregators. Their strategies, combined with random walk, fake-packet injection, and multi-path routing, are applicable to a network of multiple base stations.

[3] Generally, a sensor node becomes a source upon detecting some asset that appears within its sensing range. We say that a source is active if it is in the process of sending packets to the sink and dormant otherwise. The dormant source may be involved in local monitoring of nearby assets or performing internal operations like compression of the gathered data [11].

[4] The terms safety period (SP) and delivery latency (DL) are introduced in [3] as metrics to evaluate the privacy strength and the efficiency of the routing protocol. The former is defined as the number of packets delivered to the sink before the location of the packet-originating node is exposed to the adversary, while the latter is the average length (in hops) of the path carrying the packets. In this paper, we normalize SPs and DLs with the least number of hops between the source and the sink, thus we use NSPs and NDLs.

[5] The problem arises as the adversary traces up the route path in the opposite direction of the packet delivery.

Location privacy routing (LPR) [8] is a mono-phase protocol that protects the packet destination against the packet-tracing attack started from the source. In LPR, each node divides its neighbors into two lists: $closer\_list$, consisting of neighbors closer to the destination and $further\_list$, comprising the rest of neighbors. Whenever a packet arrives, the node chooses its next-hop node from $further\_list$ with probability $p_f(0 < p_f < 1)$ and from $closer\_list$ with probability $1 - p_f$. Optionally, it emits a fake packet to one of its neighbors chosen from $further\_list$ LPR provides strong sink-location privacy. Unfortunately, it cannot be directly applicable for source-location privacy since the fake synchronization problem may arise, as in CEM. Besides, the path made by LPR may occasionally bring many oscillations, such as back-and-forth or zigzag movements. As addressed in [8], it usually introduces a long path of many hops for each packet and suffers from the excessive packet-delivery latency.

III. NETWORK AND THREAT MODELS

In this section, we briefly introduce the network and adversary models that will be used in this paper. Readers may refer to [3], [6] for more information. There exist N sensor nodes and multiple assets in the network. But the number of assets is much less than the number of nodes (for instance, less than 1.0% of N). Assets require their locations to be protected against the packet-tracing attack. Each node has the signal transmission (or sensing) range of r (> 0) and two nodes apart with distance more than r communicate via relay nodes in the multi-hop fashion. We do not consider any specific medium access control (MAC) protocol in our study. The link-layer transmission of each node is based on the omni-directional local (i.e., 1-hop) broadcast. We assume that neither collisions nor errors arise in packet transmissions.

The adversary considered in this paper is a *passive* attacker that can eavesdrop on the local traffic among nearby nodes. The adversary is able to perform the hop-by-hop tracing toward the packet-originating node (i.e., the active source), but neither injects any packets into the network nor interferes with node communications. Possibly, equipped with a GPS receiver, a spectrum analyzer, and an antenna, the adversary can measure the arrival angle of a packet transmission and the strength of the signal as well. The adversary is always in a listening mode and it chooses and moves to the immediate node that has transmitted the packet *first*. The adversary is also *patient* enough to wait at a location until it hears the new packet, i.e., the patient model in [6]. We assume that the adversary always starts its tracing from the base station.

Assume $L(v)$ denotes the coordinate of object v (for instance, a node, asset or adversary), i.e., $L(v) = (x_v, y_v)$. We say that the location of source s is *exposed to* or *captured by* adversary κ if and only if $|L(s) - L(\kappa)| \leq \alpha$ where, α is called the capture range [3] or disclosure distance, and a disk of radius α is said to be disclosure area. As in [3], [6]–[9], we assume that the hearing radius of the adversary is equal to that of the sensor node, i.e., $\alpha = r$

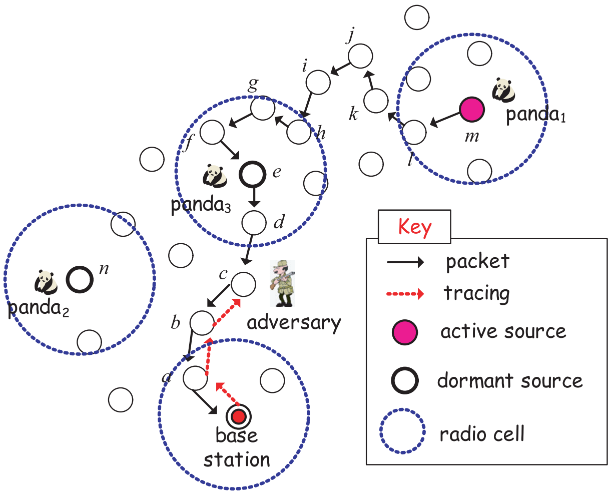


Fig. 1. Active source and dormant source in a wireless sensor network of multiple assets.

IV. GSLP-w APPROACH

This section describes the details of our proposed routing method, GSLP-w, that takes account of the presence of multiple assets and seeks to deliver more packets while avoiding excessively long paths. We first introduce the motivation behind our approach and then, the next-hop selection algorithm. We derive the expected length of the route path established by our algorithm.

A. Dormant Source

A node that senses assets appearing within its signal range is called a source. A source node usually gathers information from assets (and the environment) and sends it by a series of packets to the base station. A source node is said to be *active* if it is in the process of reporting gathered information to the sink and *dormant* otherwise. The dormant source may be involved in local monitoring of the nearby assets or internal operations like compression of the gathered data [11]. In Fig. 1, for instance, the source node m closest to panda$_1$ is active, and sources n and e, closest to panda$_2$ and panda$_3$, respectively, are dormant where, panda$_1$, panda$_2$, and panda$_3$ are assets. The active m makes use of a single path for the delivery of data packets to the base station, while the adversary tries to capture packets by tracing up the path. Considering one-hop tracing per packet, it can be inferred that three packets have delivered to the base station (by assuming that the attack always begins at the base station). Thus, the adversary has moved three hops closer to m. Since asset panda$_3$ is close to the node e that is two-hops away from the adversary, if one more packet is sent from m then, it will cause panda$_3$ to be in danger (the assumption is that the opponent can locate the asset within the capture distance α is such that $\alpha \leq r$ for signal transmission range r). In this case, panda$_3$ could be a victim of a routing method that takes the route path passing by near the asset. The asset panda$_3$ would be protected by using a routing strategy that makes a detour around the assets under protection. Therefore, we have to devise a next-hop selection strategy in which a node that comprises any asset within its sensing area of radius r is not chosen for routing.

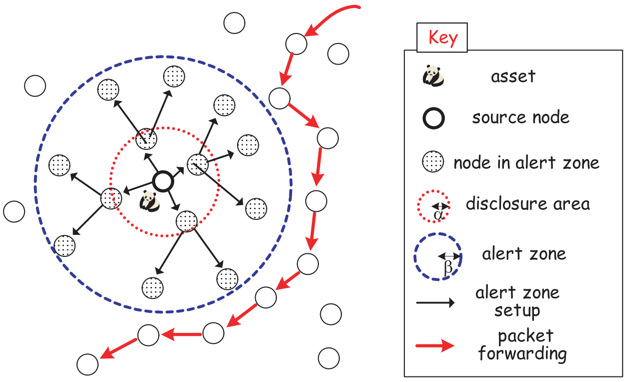


Fig. 2. Alert zone declaration by a source node and normal packet forwarding.

In the meantime, we take it for granted that the number of assets is quite small compared to the number of sensor nodes and that assets are very sparsely scattered over the network. For the sake of simplicity, it is further assumed that there exists a one-to-one relationship between each asset and its corresponding source such that no two or more assets lie within the same radio cell.[6] Hence, locating some specific source is equivalent to finding the corresponding asset, and vice versa. Equivalently, locating a source implies capturing its corresponding asset.

B. Alert Zone Setup

A simple procedure announcing a certain "be-aware-of" area is independently performed by each node that comprises an asset within its sensing range. That is, each source initially declares a circle of *alert range* β $(\geq \alpha)$ called alert zone, that is enough to hide itself and the corresponding asset from the tracing, and informs the nodes within the zone of it. As in Fig. 2, some control packets announcing the alert zone setup is diffused within the area by using geocasting [17]. Thus, every node within the zone is informed of the identity of the source node that is declared in the zone. This implies that there exist some assets within a disk of radius β. In routing, the nodes in the alert zones are not allowed to be chosen as the next-hop nodes. This prohibits the attacker that is tracing up the route path from coming into the zones, i.e., the adversary cannot approach any asset within the disclosure zones. Each alert zone can be regarded as a protective wall for the asset to block the adversary. The introduction of the alert zone may cause problem on protecting source location because a packet is going around the zone may reveal that there might be some asset near there. Under the condition that the number of assets in the network is quite less than the number of nodes, we can deploy many fake sources, as in [3], [6], or increase the capture range (i.e., disclosure distance) to make it difficult for the adversary to locate some assets.

C. Next-Hop Node Selection

In GSLP-w, every packet is forwarded to the next-hop node in one of four modes, $greedy$, $random$, $perimeter$, and $retreat$,

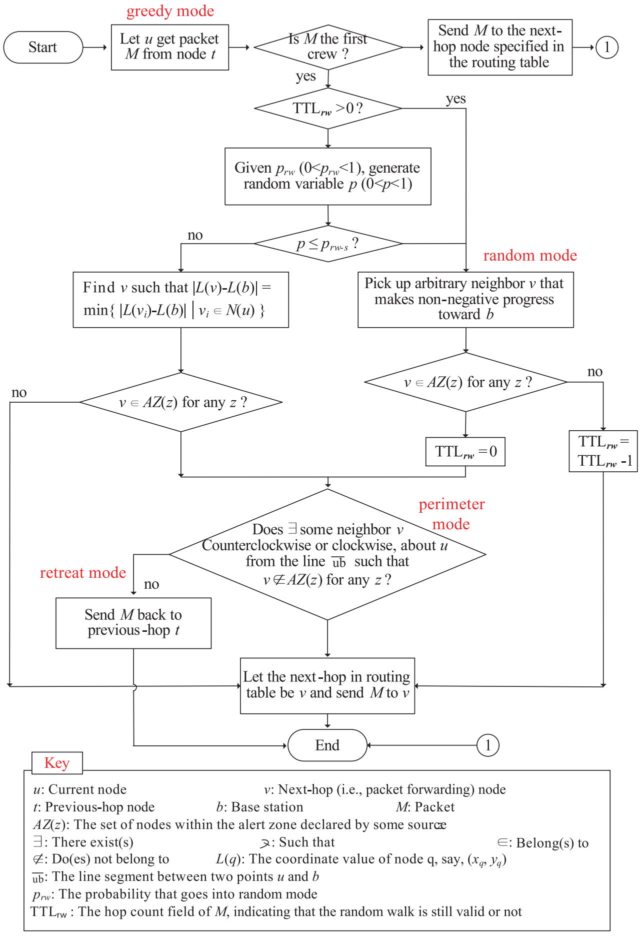

greedy mode
Start
Let u get packet M from node t
Is M the first crew ?
yes
Send M to the next-hop node specified in the routing table
1
TTL$_{rw}$ >0 ?
yes
Given p_{rw} (0<p_{rw}<1), generate random variable p (0<p<1)
no
$p \leq p_{rw\text{-}s}$?
random mode
Find v such that $|L(v)$-$L(b)|$ = min{ $|L(v_i)$-$L(b)|$ | $v_i \in N(u)$ }
Pick up arbitrary neighbor v that makes non-negative progress toward b
no
$v \in AZ(z)$ for any z ?
$v \in AZ(z)$ for any z ?
no
TTL$_{rw}$ =0
TTL$_{rw}$= TTL$_{rw}$ -1
perimeter mode
Does ∃ some neighbor v Counterclockwise or clockwise, about u from the line $\overline{\text{ub}}$ such that $v \notin AZ(z)$ for any z ?
retreat mode
no
Send M back to previous-hop t
Let the next-hop in routing table be v and send M to v
End
1
Key
u: Current node
v: Next-hop (i.e., packet forwarding) node
t: Previous-hop node
b: Base station
M: Packet
$AZ(z)$: The set of nodes within the alert zone declared by some source
∃: There exist(s)
∋: Such that
∈: Belong(s) to
∉: Do(es) not belong to
$L(q)$: The coordinate value of node q, say, (x_q, y_q)
$\overline{\text{ub}}$: The line segment between two points u and b
p_{rw}: The probability that goes into random mode
TTL$_{rw}$: The hop count field of M, indicating that the random walk is still valid or not

Fig. 3. GSLP-w, the proposed next-hop selection strategy.

as in Fig. 3. We assume that each node x knows of the coordinate of its neighbor $y \in N(x)$ and whether $y \in AZ(z)$ or not for any z, where $N(x)$ is the set of neighbor nodes of x and $AZ(z)$ denotes the set of nodes residing within the alert zone that is declared by some source z.

Greedy forwarding prevents the path getting very long and helps the path to converge to the destination. As u receives packet M from adjacent node t, it first checks if M is the first crew, i.e., the first packet that is to be sent over a path. If M is the first crew, the next-hop node is newly chosen otherwise, M is forwarded to the next-hop node specified in the routing table which is built for the geographic-aware routing like GPSR [12]. Assume p_{rw-s} denotes the probability of going into *random* mode from *greedy* mode and a random number p $(0 < p < 1)$ is generated. Then, a transition to *random mode* takes place if $p \leq p_{rw-s}$, the mode remains the same otherwise. The next-hop selection by u in *greedy* mode is as follows.

Select v such that $|L(v) - L(b)| = \min\{|L(v_i) - L(b)| \mid v_i \in N(u)\}$, *where b stands for the sink. If* $v \notin AZ(z)$ *for any* z, *the current mode switches into perimeter mode.*

The mode *random* is to enhance the path diversity. Once the mode *random* is committed, certain subsequent next-hop nodes are supposed to be chosen under the same mode for further ran-

[6] When an asset is identified by more than one node then, a leader election algorithm like [16] can be applied for bidding. But, this is beyond the scope of this paper.

domization of the path. The number of such hops, defined as *random walk length*, is specified by field TTL_{rw} within the packet being sent. That is, node u that receives packet M first checks the field of it. If $\mathrm{TTL}_{rw} > 0$, by default the mode goes into *random* mode, then next-hop *v* is chosen and TTL_{rw} is decremented by one, i.e., $\mathrm{TTL}_{rw} = \mathrm{TTL}_{rw} - 1$. Otherwise (i.e., TTL_{rw} is 0), v is selected in the mode *greedy*. Field TTL_{rw} specifies an upper bound upon the random walk length because it is valid as long as *perimeter* mode does not happen. The next-hop selection in the mode *random* is as follows.

Select an arbitrary $v \in N(u)$ that makes non-negative progress (this is known as random progress in [18]). If $v \notin AZ(z)$ for any z, then the mode changes into perimeter mode

Perimeter routing was originally introduced in GPSR [12] to avoid the routing hole that might arise due to greedy forwarding. In GSLP-w, it is adopted to direct the path under development not to get into the alert zones of comprising assets. This brings detouring of alert zones encountered during the packet-delivery. Each source node that sets up its alert zone and the asset within it are protected from the packet-tracing attack as the adversary cannot approach within the disclosure or capture range α.

Assume $\bar{ub}$ denotes the line-segment from current node u to sink b. The routing in *perimeter* mode is performed as follows.

Select v that is the first neighbor counterclockwise (i.e., right-hand rule) or clockwise (i.e., left-hand rule) about u from $\bar{ub}$ such that $v \notin AZ(z)$ form any z. If such v does not exist, the mode goes into retreat mode.

The mode *retreat* is to get backtracking to the previous-hop node t when the path cannot be developed any longer at current node u, in the modes, *greedy*, *random*, and *perimeter*. We assume that the routing hole problem never happens in the networks and that there always exists a path between any pair of a packet-originating node and the sink. Readers may find more information on this topic in [19], [20].

Remarks: We now compare the proposed GSLP-w with other representative works in [3], [6]–[9], that are in the area of location privacy routing for wireless sensor networks. All approaches make use of random forwarding for the path diversity. They differ among each other on the heuristics applied to increase randomness. In PR-SP [3], [6] the first half of the routing is the random walk of some fixed length. LPR [8] performs probabilistic random forwarding through the whole process of the routing. In GSLP-w, the selection by using the random progress begins with probability p_{rw-s} and continues at the subsequent hops as long as $\mathrm{TTL}_{rw} > 0$, provided that *perimeter* mode does not arise. The probability that goes into *random* mode needs not be large (in simulations, it is given as $p_{rw-s} = 0.05$ for modest cases). Greedy forwarding is used only in the last phase of PR-SP [3], [6], while, in our method, every node at the mode *greedy* tries to perform it. In CEM [9], the route path for the packet-delivery is set up by using the shortest path discipline. Another feature of our approach is the *perimeter* mode that takes care of detouring the alert zones. As opposed to the single packet per path in other works [3], [6]–[8], in GSLP-w, w packets are allowed to go through a path, where different w values are randomly chosen for different paths (to be presented in the next section). Finally, back-and-forth or zigzag forwarding is not intended to enlarge the path in GSLP-w. □

D. Evaluation Criteria and Expected Path Length

Two criteria, namely safety period (SP) and delivery latency (DL) in [3] are used for the evaluation of the proposed routing method through simulations. We do not consider the number of packets carried into the network since our method and PR-SP [3], [6], both use a single path without fake-packet injections, thus the delivery latency (i.e., path length) is reduced to the number of packets carried per path. In the meantime, since this paper is also concerned with the dormant sources regarding multiple assets, we need to redefine the original SP as follows. The (modified) *safety period* is defined as the number of data packets successfully delivered to the sink from the active source before the source is captured by the adversary, *yet completely protecting location-privacy of other sources*. The metrics are valued under the assumption that the attacker begins his tracing at the sink. Therefore, we use *normalized safety period (NSP)* and *normalized delivery latency (NDL)* which are obtained by dividing SP and DL, respectively, with the least number of hops between the active source and the sink.

We now calculate the expected number of hops of the path taken by GSLP-w in Fig. 3. Our goal is to find how many times the path made by the proposed method is longer than the shortest path. Let p_g $(0 < p_g < 1)$ be the probability that greedy forwarding is committed in choosing the next-hop node. Denote by $E(k)$ the least number of hops remained toward the sink b after k (> 0)-consecutive *movement* (i.e., packet-forwarding) from the active source s. Let d be the number of hops of the shortest path between s and b. Initially, we have $E(0) = d$ at s, as there has been no movement yet. Then, 1-hop movement from s leads to the equality $E(1) = E(0) - p_g + (1 - p_g) = E(0) + (1 - 2p_g)$, because the movement directs the path under development to be shortest toward b with probability p_g, while to be non-shortest with probability $1 - p_g$. Thus, the recurrence relations for successive movements are given as follows: $E(0) = d$, $E(1) = E(0) + (1 - 2p_g)$, $E(2) = E(1) + (1 - 2p_g)$, $\cdots$, $E(k) = E(k-1) + (1 - 2p_g)$. This yields to a general equation $E(k) = d + k(1 - 2p_g)$. Suppose that the path converges to sink b after k movements. It implies that $E(k) = 0$, i.e., $d + k(1 - 2p_g) = 0$. Note here that k is the expected number of hops we want to get. Thus, it is given that *expected length of the path* (in hops) $= k = d/(2p_g - 1)$ and $\mathrm{NDL} = 1/(2p_g - 1)$. Put differently, the length of the path established by GSLP-w is $1/(2p_g - 1)$ times longer. That is, the safety period of GSLP-w is $1/(2p_g - 1)$ times longer than that of the shortest path between s and b. The inequality $p_g > 1/2$ must hold for any movement because the path should converge to the sink. Note that the equations above actually state the $upper\ bounds$ on the path lengths since the possibility that the next-hop node chosen in $random$ mode directs the path to be shortest to b has been ignored in the formulas above.

Remarks: Probability $1 - p_g$ that directs the next-hop movement at each intermediate node to be non-shortest to the sink can be expressed as $1-p_g = p_p + p_{rw-c} + p_r$, where p_p is the probability that $perimeter$ mode is committed, p_{rw-c} is the probability that $random$ mode is committed, and p_r is the probability that $retreat$ mode arises, respectively, in choosing the next-hop node. We note that p_{rw-c} differs from p_{rw-s}, the probability that mode $greedy$ switches into $random$ in general. Generally, $p_{rw-c} \leq p_{rw-s}$ holds because the next-hop node chosen at either $greedy$ or $random$ mode becomes void if it lies within some alert zone, and the new next-hop is chosen in $perimeter$ mode. Probability p_p depends on the number of dormant nodes in the network and both p_{rw-c} and p_r are also related to the number of neighbors of a node. When there exists only one asset as in [3], [6]–[9], it is given that $p_p = 0$ and $p_{rw-c} = p_{rw-s}$. Hence, making the path longer to increase privacy strength of the source-location more is quite simple because we can freely fix the values of system parameters p_g and p_{rw-s}. But, the case is different when $p_p > 0$ because the actual values of p_{rw-c}, p_g, and p_r are intrinsically *non-deterministic*. Next, we present the performance evaluation of the proposed GSLP-w through simulations. □

V. PERFORMACE EVALUATION

We begin with introducing the simulation configurations. Comparisons are made with PR-SP [3], [6], a routing strategy for source-location privacy in wireless sensor networks.

A. Simulation Setup

We are not aware of simulation tools, yet available in the public domain and dedicated to measuring location privacy strength and related performances. As in other works [3], [6], [7], [9] we developed our own software for simulations. Simulator software is developed in Java and consists of about 5,400 source lines and 541 kB of executable code. It covers the routing algorithms of PR-SP and GSLP-w without including the physical and MAC layers. All sources are assumed to be stationary throughout simulations. Packets are sent according to the low-duty cycle model [3], [6], i.e., the subsequent packet from the active source is not sent until its proceeding packet arrives at the destination. The tracing by the adversary is assumed to always begin at the sink.

Each simulation of GSLP-w uses 100 topologies of the network that is comprised of about 50,000 nodes with the average number of neighbors being 8, where the nodes are randomly placed. Excluding the least tens and the largest tens, 80 out of 100 results are averaged for evaluations. Routing holes and path looping are not taken into account in simulations. The number of dormant sources N_s is restricted within 0.8% of N because the competitor, PR-SP, hardly develops its own path in case it is beyond the bound. Initially, we considered only one active source but multiple dormant sources in the networks and later extended the number of active sources up to 8. Table 1 summarizes the simulation configurations. We define the *crew size* w as the number of packets to be sent over a path. Noting both privacy strength and performance of a routing method are proportional to h_{s-b} (i.e., the least number of hops between s and b), various fractions of it are considered for choosing w (hereafter, let us call all instances of GSLP-w as GSLP family).

Table 1. Simulation configurations.

Parameters			Values or ranges
Symbol		Meaning	
N		number of the nodes	50,000
average degree of the node			8
$h_{s\text{-}b}$		number of hops between active source s and sink b	30, 50, 70
N_s		number of dormant sources	0.2%, 0.4%, 0.8% of N
number of runs for each simulation			100
GSLP-w	w	crew size (number of packets being sent per path)	1, 1Q, 2Q, 3Q, 4Q, 2/3Q
	$p_{\text{rw-s}}$	probability that goes into *random* mode	0.05
	TTL$_{rw}$	random walk length (hops)	randomly chosen from [5%, 10%] of $h_{s\text{-}b}$ (min. 2)
	b	alert range	$2r$ (r: Transmission range)
	a	capture (disclosure) range	R
PR-SP		random walk length (hops)	randomly chosen from [25%, 50%] of $h_{s\text{-}b}$

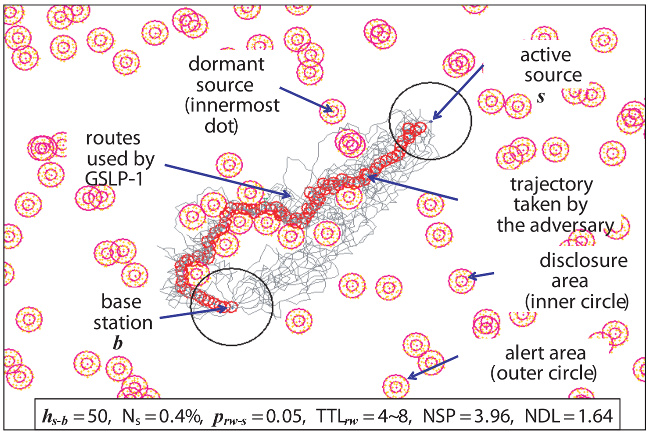

dormant source (innermost dot)
active source s
routes used by GSLP-1
trajectory taken by the adversary
disclosure area (inner circle)
base station b
alert area (outer circle)
$h_{s\text{-}b}$ = 50, N$_s$ = 0.4%, $p_{rw\text{-}s}$ = 0.05, TTL$_{rw}$ = 4~8, NSP = 3.96, NDL = 1.64

Fig. 4. Example of the paths established by GSLP-w (screenshot, w=1).

- ○ GSLP-1: $w = 1$, one packet
- ○ GSLP-1Q: $w \in 1Q$, i.e., chosen from $[2, h_{s-b}/4]$
- ○ GSLP-2Q: $w \in 2Q$, i.e., chosen from $[(h_{s-b}/4)+1, h_{s-b}/2]$
- ○ GSLP-3Q: $w \in 3Q$, i.e., chosen from $[(h_{s-b}/2)+1, (3h_{s-b})/4]$
- ○ GSLP-4Q: $w \in 4Q$, i.e., chosen from $[((3h_{s-b})/4)+1, h_{s-b}]$
- ○ GSLP-2/3Q: $w \in 2/3Q$, i.e., chosen from $[(h_{s-b}/4)+1, (3h_{s-b})/4]$

A screen shot of the paths established by using GSLP-w is shown in Fig. 4, where $w = 1$ and the ordinary nodes are not depicted for the simplicity of the figure. Each source is surrounded with two small circles: the inner one shows the disclosure area of capture range α $(= r)$ and the outer one outlines the alert zone of alert range β $(= 2r)$. Two big circles, respectively, on the top right and the bottom left emphasize the locations of active source s and sink (i.e., base station) b. The many paths established from s to b are drawn with thin ines. Following these in the reverse direction, the tracing taken by the attacker is represented with small thick circles. It is seen that s has been captured in the end, but all dormant sources (also their corresponding assets) remain “alive” from the packet-tracing attack, owing to detours taken by perimeter routing. The configured system parameters and obtained metrics are shown on the bottom of each figure.

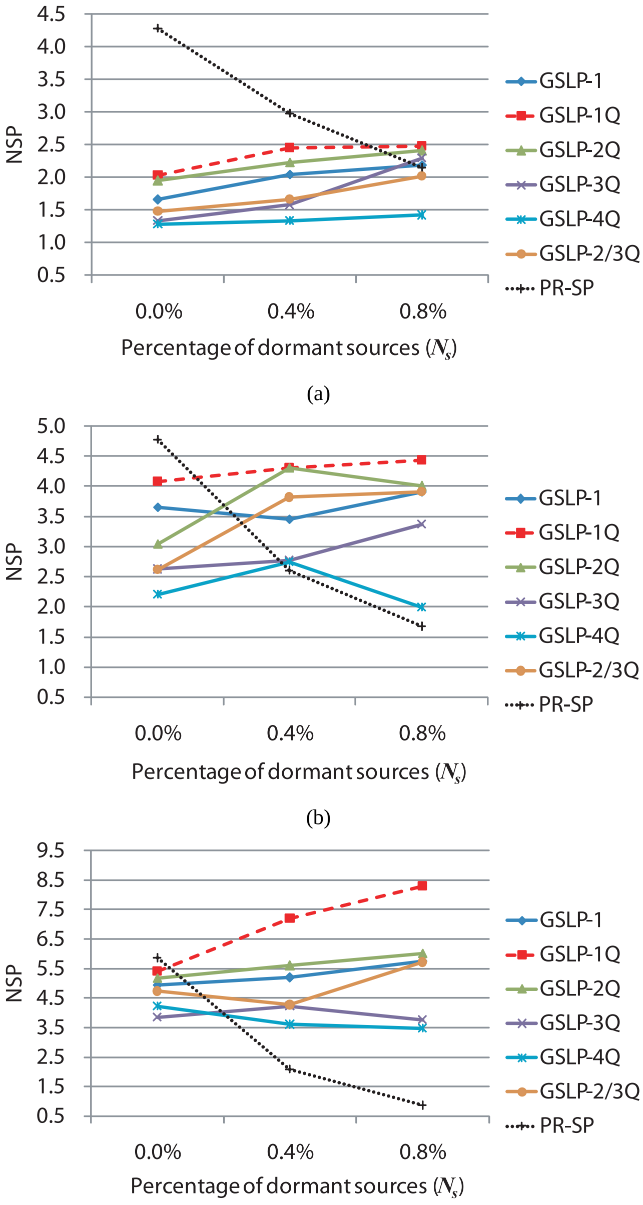
In Fig. 4, it can be inferred that the paths established by using GSLP-w are almost $evenly$ distributed from side to side in terms of the straight line-segment between s and b. This is due to the fact that the next-hop node selection in *perimeter* mode is committed by alternately taking the left-hand rule and in turn, the right-hand rule for route paths. Thus, if the attacker is enticed into one side, the subsequent path on the other side may carry w packets without being in danger of tracing. Paths $hardly$ introduce back-and-forth or zigzag movements, even though they may take place occasionally or pathologically in case of detouring the alert zones. Hence, the adversary tracing back the paths may rarely suspect that it has been enticed into the wrong place.

B. Simulation Results

B.1 Normalized Safety Period (NSP)

The impact of the number of dormant sources N_s on NSP is shown in Fig. 5. The key point is that, as N_s increases, NSPs of most of the GSLP family slightly increase, while those of PR-SP (shown in dotted lines) drop sharply. This trend stems from the fact that the GSLP family possesses the perimeter routing capability that detours the alert zones throughout the packet-delivery, but PR-SP does not have such capability. In PR-SP, the possibility that packet-forwarding gets into the alert zones increases as h_{s-b} increases. This shortens the safety periods of PR-SP. Among the GSLP family, GSLP-1Q (drawn in thick broken line) provides the highest NSPs for all cases. Concerning the crew size w, more w implies less NSPs in general. But, it is worth noting that one packet for each path (i.e., $w = 1$) is not so good as much as GSLP-1Q (i.e., $w \in 1Q$), and it ranks roughly in the middle among the GSLP family. Too large as $w \in 3Q$ or $4Q$ and too small as $w = 1$, cases are not good choices. This can be explained as follows. The large crew size forces the tracing by the adversary to more quickly capture the active source as many packets are sent over a path. In case of $w = 1$ each packet takes a separate path, thus more paths are established. But, because of small values of parameters TTL_{rw} ($= 5\%{\sim}10\%$ of h_{s-b}) and $p_{rw-s}(= 0.05)$, the path diversity is not sufficient and the path is also not sufficiently longer. However, the case of crew size $w \in 1Q$ entices the adversary into further random places and delays the tracing, compared to the case of $w \in 3Q$ or $w \in 4Q$, and it allows more packets per path (i.e., increases NSPs), compared with the case of $w = 1$. GSLP-$2/3Q$ yields NSPs, as expected, roughly between NSPs given by GSLP-2Q and GSLP-3Q.

Interestingly, PR-SP is better than the GSLP family when h_{s-b} is relatively small as 30 or N_s is near zero (Fig. 5(a) and (b)). Since $p_{rw}-s$ is 0.05 and TTL_{rw} is 2 or 3 when $h_{s-b} = 30$, neither the path diversity nor the path length cannot be enlarged sufficiently. On the contrary, the random walk length of PR-SP varies from 7 to 15, thus PR-SP yields relatively longer and more randomized paths and it provides more increased NSPs. Later, we address how to improve NSPs of the GSLP family further in such a case.



(a)

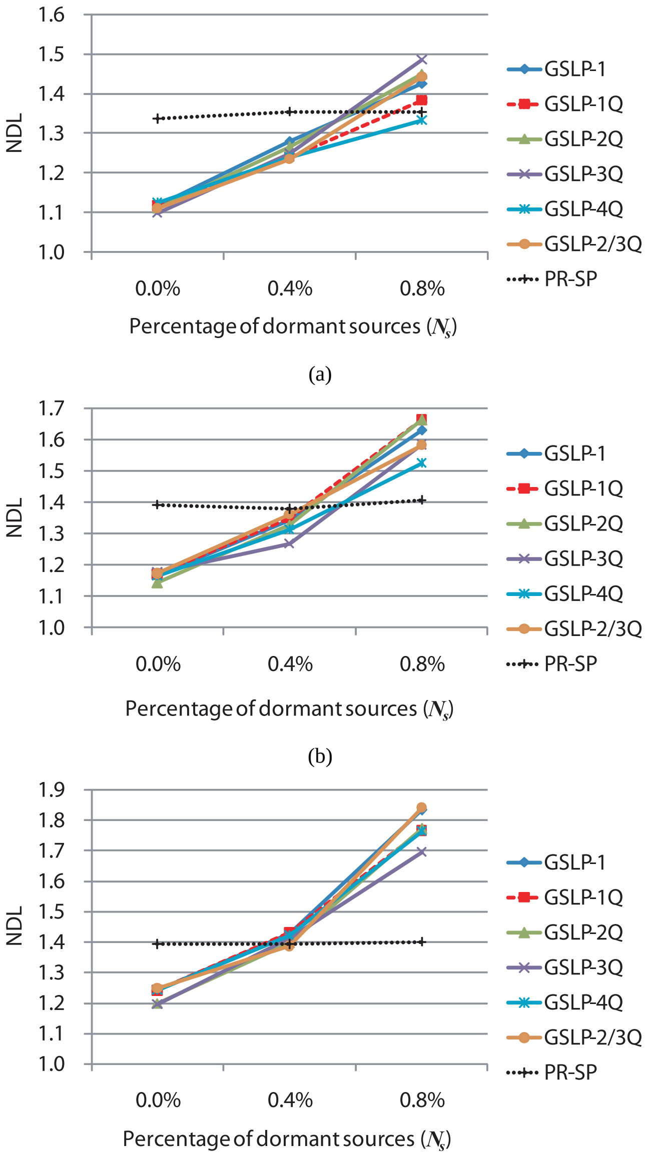
(b)

(c)

Fig. 5. The impact of N_s on NSP: (a) $h_{s-b} = 30$, (b) $h_{s-b} = 50$, and (c) $h_{s-b} = 70$.

B.2 Normalized Delivery Latency (NDL)

Fig. 6 shows NDLs measured through simulations. Since PR-SP does not take into account the alert zones during the path development, it gives nearly invariant delivery latencies. NDLs of PR-SP remain below 1.4 for all cases. The GSLP family takes 1.53 on average and a maximum 1.85. Detouring the alert zones in the GSLP family usually makes their paths longer than PR-SP. As N_s increases, the deviations among NDLs of the GSLP family slightly increase but still remain within 0.2. From the point of the ratio of NSP to NDL, the GSLP family offers better results than PR-SP for $Ns \geq 0.4$ at $h_{s-b} = 50$ and $N_s \geq 0.2$ at $h_{s-b} = 70$.

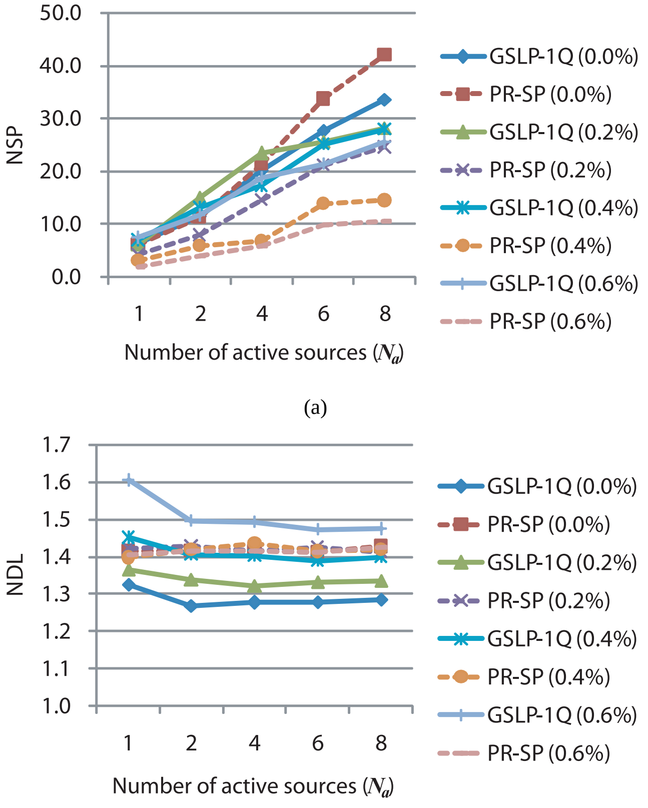


(a)

(b)

(c)

Fig. 6. The impact of N_s on NDL: (a) $h_{s-b} = 30$, (b) $h_{s-b} = 50$, and(c) $h_{s-b} = 70$.



(a)

(b)

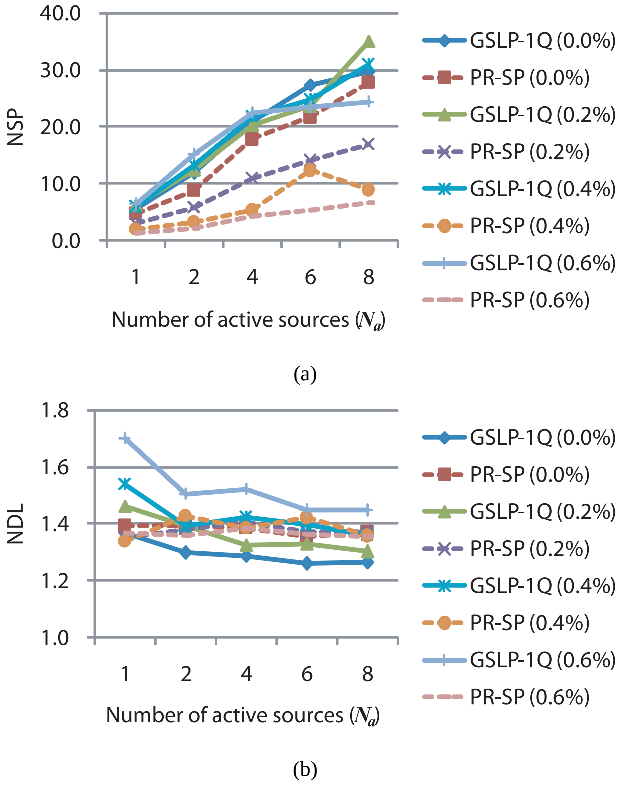
Fig. 7. The impact of the number of active sources ($h_{s-b} = 70$): (a) NSP and (b) NDL.

B.3 Impact of Number of Active Sources

So far we have considered the case when there is only one active source even though there exist multiple dormant sources in the networks. Fig. 7 shows the results of experiments when the number of active sources, N_a, is 1, 2, 4, 6, and 8, respectively, where the distance of the shortest path between each pair of an active source and the sink is about 70-hop and active sources are evenly placed into omni-directionally. Percentages in parentheses in legends of Fig. 7 refer to, N_s's, the rates of dormant sources to the total number of nodes. Overall, NSPs of both GSLP-1Q and PR-SP increase as N_a increases. This is because the adversary usually traces up the active source that send a packet *for the first time* to the sink thus, the other sources may send packets with little danger of location-exposure. Since PR-SP does not possess perimeter routing capability that avoids alert zones in the networks, NSPs of PR-SP gradually decrease as N_s increases. On the other hand, NDLs of PR-SP are almost invariant to both N_a and N_s, but those of GSLP-1Q slightly increase as N_s does due to perimeter routing. Fig. 8 shows the simulation results when each active source is randomly chosen out of [50, 100]-hop with respect to the sink.

B.4 Further Improvements by Increasing p_{rw-s} and TTL_{rw}

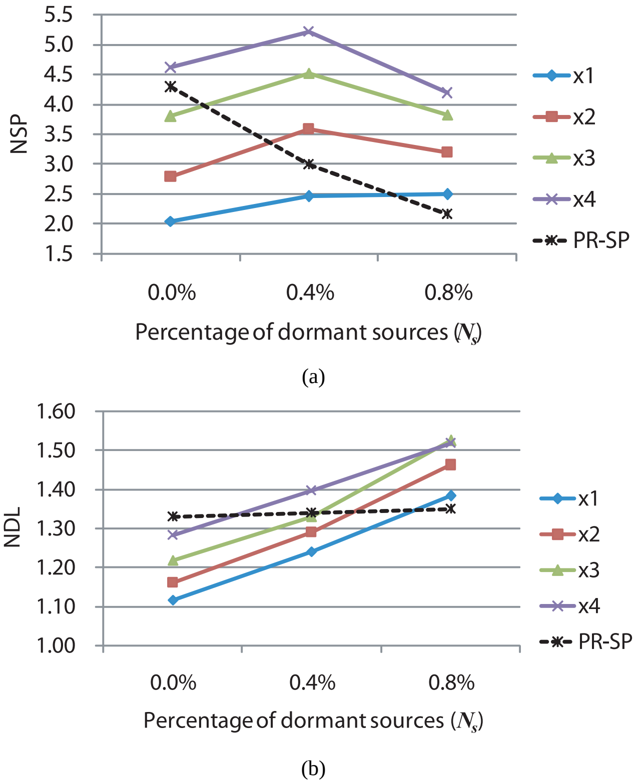
In the previous experiments, small values were assigned to system parameters p_{rw-s}, h_{s-b}, and TTL_{rw} (see Table 1). We have seen that the path randomization and enlargement by the GSLP family may not be so in effect at such values. We *intentionally* consider the case of $h_{s-b} = 30$ because, at this relatively small value, NSPs of the GSLP family are lower than those of PR-SP as in Fig. 5 (a). We want to observe the impact of parameters p_{rw-s} and TTL_{rw} on NSPs. First, the random walk length (i.e., TTL_{rw}) is increased as one-folded ($\times 1$), two-folded ($\times 2$), three-folded ($\times 3$), and four-folded ($\times 4$), respectively. Thus, new TTL_{rw} ranges we consider are [2, 3], [4, 6], [6, 9], and [8, 12], respectively, but still less than 15, the half of h_{s-b}. As in Fig. 9 (a), NSPs of GSLP-1Q increases until $N_s \leq 0.4\%$, but decreases after that. The reason is as follows. During the first half, the effect of the path diversity by the extended random walk continues, because there still exist a few spaces to hold the paths that can make detours to avoid the alert zones. But for $N_s \leq 0.4\%$, there may exist many



(a)

(b)

Fig. 8. The impact of the number of active sources when each source is randomly chosen out of [50, 100]-hop with respect to the sink: (a) NSP and (b) NDL.



(a)

(b)

Fig. 9. Further improvement by increasing p_{rw-s} and TTL_{rw} ($h_{s-b} = 30$): (a) NSP and (b) NDL.

dormant sources and alert zones, hence the mode *perimeter* is more frequently committed. So the path diversity effect gradually diminishes and it results in the decrease of NSPs. Nonetheless, compared with PR-SP, GSLP-1Q provides higher NSPs for $N_s \geq 0.3\%$ at two-folded TTL_{rw} and for $N_s \geq 0.1\%$ at three-folded TTL_{rw}. And it offers always higher NSPs at four-folded TTL_{rw}. NDLs under the same simulation settings are shown in Fig. 9 (b). As N_s increases, so does the number of alert zones. Thus, the path made by GSLP-1Q lengthens and NDLs of it, as well. At $N_s = 0.4\%$, two- or three-folded TTL_{rw} suffices to make GSLP-1Q offer higher NSPs with lower NDLs.

Remarks: We further measured NSPs by increasing p_{rw-s} as one-folded ($\times 1{=}0.05$), two-folded ($\times 2{=}0.10$), three-folded ($\times 3{=}0.15$), and four-folded ($\times 4{=}0.20$), respectively (the results are not depicted here). But the effects are not as good as the case of increased TTL_{rw}. It is expected that NSPs will further increase at the cost of lengthening NDLs if TTL_{rw} and p_{rw-s} are simultaneously increased. □

VI. CONCLUSION

In this paper, we have proposed a new routing technique GSLP-w that enhances location privacy of the packet-originating node in the presence of multiple assets. With combining three modes, *greedy*, *random*, and *perimeter* during the packet-delivery, GSLP-w can render more path diversity, yet allowing location-privacy of multiple assets. The paths developed by the proposed method hardly possess back-and-forth or zigzag. Thus, the adversary tracing up the paths may suspect less that it might be enticed into the wrong place. Through simulations, we found that GSLP-1Q among the GSLP family shows the best results regarding both the safety strength and the packet-delivery latency. As N_s increases, improvements of NSPs compared to PR-SP become significant in terms of the ratio of safety period to delivery latency. Higher NSPs can be achieved by increasing TTL_{rw} or p_{rw}.

There is still open research issue to make a proper tradeoff among the metrics, safety period, delivery latency, communication cost (power consumption), and control overhead, in the networks of multiple assets and mobile sources. The issue on location privacy in non delay-tolerant networks under the high-duty cycle model is also a challenging topic. More powerful adversary models and evaluation of routing strategies under non-uniform distributions of multiple assets are also future work.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their valuable and constructive comments, which helped greatly in improving the quality and presentation of this paper. Special thank goes to Prof. Ufuk Çağlayan (Boğaziçi University, Istanbul, Turkey) for his insightful advice and comments on this paper. The author also thanks the members of the study group INQ in the School of Computer, Information, and Communication Engineering at Sangji University for their running simulation and collecting data.

REFERENCES

[1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Int. J. Comput. Netw.*, vol. 52, no. 12, pp. 2292–2330, 2008.

[2] S. Guizani, H.-H. Chen, and P. Muellerc (eds.), "Special issue on security on wireless ad hoc and sensor networks," *Int. J. Comput. Commun.*, vol. 30, no. 12, pp. 2311–2518, 2007.

[3] C. Ozturk, Y. Zhang, and W. Trappe, "Source-location privacy in energy-constrained sensor network routing," in *Proc. SASN*, 2004, pp. 88–93.

[4] P. Kamat, W. T. W. Xu, and Y. Zhang, "Temporal privacy in wireless sensor networks," in *Proc. IEEE ICDCS*, 2007, p. 23.

[5] J. Deng, R. Han, and S. Mishra, "Countermeasures against traffic analysis attacks in wireless sensor networks," in *Proc. IEEE SecureComm*, 2005, pp. 113–126.

[6] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk, "Enhancing source-location privacy in sensor network routing," in *Proc. IEEE ICDCS*, 2005, pp. 599–608.

[7] L. Zhang, "A self-adjusting directed random walk approach for enhancing source-location privacy in sensor network routing," in *Proc. ACM IWCMC*, 2006, pp. 33–38.

[8] Y. Jian, S. Chen, Z. Zhang, and L. Zhang, "Protecting receiver-location privacy in wireless sensor networks," in *Proc. IEEE INFOCOM*, 2007, pp. 1955–1963.

[9] Y. Ouyang, Z. Le, G. Chen, and J. Ford, "Entrapping adversaries for source protection in sensor networks," in *Proc. IEEE WoWMoM*, 2006, pp. 23–32.

[10] K. Mehta, D. Lie, and M. Wright, "Location privacy in sensor networks against a global eavesdropper," in *Proc. IEEE ICNP*, 2007, pp. 314–323.

[11] S. Puthenpurayil, R. Gu, and S. S. Bhattacharyya, "Energy-aware data compression for wireless sensor networks," in *Proc. IEEE ICASSP*, vol. 2, 2007, pp. 45–48.

[12] B. Karp and H.-T. Kung, "Greedy perimeter stateless routing for wireless networks," in *Proc. ACM/IEEE MobiCom*, 2000, pp. 243–254.

[13] I. S. P. Bose, P. Morin, and J. Urrutia, "Routing with guaranteed delivery in ad hoc wireless networks," in *Proc. ACM DIALM*, 1999, pp. 48–55.

[14] H. Frey and I. Stojmenovic, "On delivery guarantees of face and combined greedy-face routing in ad hoc and sensor networks," in *Proc. ACM/IEEE MobiCom*, 2006, pp. 390–401.

[15] Y. Xi, L. Schwiebert, and W. Shi, "Preserving source location privacy in monitoring-based wireless sensor networks," in *Proc. IEEE IPDPS*, 2006, pp. 25–29.

[16] M. Garey and D. Johnson, *Computer and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, San Francisco, 1979.

[17] G. Chen, J.-W. Branch, and B.-K. Szymanski, "Local leader election signal strength aware flooding, and routeless routing," in *Proc. IEEE IPDPS*, vol. 13, 2005, pp. 241–249.

[18] Y.-B. Ko and N. Vaidya, "Geocasting in mobile ad hoc networks: Locationbased multicast algorithms," in *Proc. IEEE WMCSA*, 1999, pp. 101–110.

[19] N. R and L. Kleinrock, "The spatial capacity of a slotted aloha multihop packet radio network with capture," *IEEE Trans. Commun.*, vol. 32, p. 32, 1984.

[20] N. Ahmed, S. Kanhere, and S. Jha, "The hole problem in wireless sensor networks: A survey," *Mobile Comput. and Commun. Review*, vol. 9, pp. 4–18, 2005.

[21] P. Kamat, W. Xu, W. Trappe, and Y. Zhang, "Temporal privacy in wireless sensor networks, in *Proc. IEEE ICDCS*, 2007, pp. 25–29.

Yeongwhan Tscha received the B.Sc. degree from Inha University in 1983, the M.Sc. degree from KAIST (Korea Advanced Institute of Science and Technology) in 1985, and the Ph.D. degree from Inha University in 1993, respectively, all in Computer Science. He joined ETRI (Electronics and Telecommunications Research Institute), Korea, in 1985, for R&D in the areas of ISDN, CCS No.7 Systems, and Protocol Engineering. During 1986–1987, he was a Guest Scientist at NIST, USA and was participated in the project on automated protocol development methodology. He left ETRI in 1990 and completed his Ph.D. degree in 1993. Since 1994, he has been with the School of Computer, Information, and Communication Engineering, Sangji University, Korea, where currently he is a tenured Professor. In 2004–2005, he spent his sabbatical year as a Visiting Professor in the Department of Computer Engineering at Boğaziçi University, Istanbul, Turkey. His research includes network architectures and their protocols, layerless switching and routing, networking protocols for ad hoc and sensor networks, and network security.