Hidden Anchor: Providing Physical Layer Location Privacy in Hybrid Wireless Sensor Networks

Rania El-Badry, Moustafa Youssef*
Wireless Intelligent Networks Center (WINC)
Nile University, Egypt
Email: {rania.elbadry,mayoussef}@nileu.edu.eg

Mohamed Eltoweissy
The Bradley Dept. of Elect. and Comp. Eng.
Virginia Tech, USA
Email:toweissy@vt.edu

***Abstract*—In many hybrid wireless sensor networks (HWSNs) applications, sensor nodes are deployed in hostile environments where trusted and un-trusted nodes co-exist. In such hybrid networks, it becomes important to allow trusted nodes to share information, especially, location information and, at the same time, prevent un-trusted nodes from gaining access to this information. We focus on anchor-based localization algorithms in HWSNs, where a small set of specialized nodes, i.e. anchor nodes, broadcast their location to the network and other nodes can use the broadcast information to estimate their own location. The main challenge is that both trusted and un-trusted nodes can measure the physical signal transmitted from anchor nodes. Thus, un-trusted nodes can use the physical signal transmitted from an anchor node to estimate its location. In this paper, we propose *Hidden Anchor*, an algorithm that provides anchor physical layer location privacy. The *Hidden Anchor* algorithm exploits the inherently noisy wireless channel and uses identity cloning of neighboring trusted nodes to make anchors unobservable to un-trusted nodes while providing complete information to trusted nodes. Evaluation of the *Hidden Anchor* algorithm through analysis and simulation shows that it can hide the location and identity of anchor nodes with very low overhead.**

***Index Terms*—Anchor-based localization, location unobservability, physical layer location privacy.**

I. Introduction

Location discovery has been an active area of research in wireless sensor networks (WSN) due to its critical need in many applications including location-based routing [1], coverage [2], node identification, and information tagging. Localization algorithms can be categorized as either anchor-based or anchor-free [3]. Anchor-based algorithms, e.g. [4], [5], assume the existence of a small set of nodes with known locations, i.e. anchor nodes, that broadcast their location information to the network in special *beacon* frames. A node with an unknown location estimates its distance to the anchor node, in a process known as ranging, and combines the estimated distance to at least three anchor nodes with the broadcast anchors' locations in beacon packets to estimate its location in 2D (Fig.1). On the other hand, anchor-free localization algorithms e.g. [6], [7], do not assume the existence of anchor nodes and estimate the relative topology of the network, in which the coordinate system is established by a reference group of nodes. This paper focuses on anchor-based localization algorithms using Received Signal Strength (RSS) for ranging.

In many hybrid wireless sensor networks (HWSNs) applications, sensor nodes are deployed in hostile environments where trusted and un-trusted nodes co-exist. In such hybrid networks, it becomes important to allow trusted nodes to share information while, at the same time, prevent un-trusted nodes from gaining access to this information.

An anchor node may encrypt its beacon packets with a key shared only with trusted nodes. This will prevent un-trusted nodes from getting the information contained in the beacon packets. Although encryption can provide location information secrecy, it does not provide ***physical layer*** location privacy, where a group of un-trusted nodes can measure the received signal strength (RSS) of encrypted messages and cooperate to determine the anchor nodes' locations through trilateration. This paper proposes an algorithm, termed *Hidden Anchor*, that addresses the physical layer location privacy problem. In particular, the Hidden Anchor algorithm provides ***anchor nodes unobservability***, where un-trusted nodes cannot detect (observe) the existence of anchor nodes.

In [8], we proposed the *HyberLoc* algorithm for addressing the physical layer location privacy problem. *HyberLoc* depends on the anchor nodes to dynamically change their transmission power and to include the used transmission power in the encrypted beacon packet. However, *HyberLoc*'s advantage is limited because of the current limitations of the sensor hardware. We discuss this approach in details in section III.

Our novel approach in the *Hidden Anchor* algorithm is to exploit the noisy characteristics of the wireless channel to hide the location of anchor nodes from un-trusted nodes, while providing complete information to trusted nodes. The idea is for anchor nodes to randomly use the identity of the nearby trusted nodes when broadcasting their beacon packets. As a result, un-trusted nodes will not be able to distinguish between anchor traffic and trusted node traffic. Shared information between the anchor and trusted nodes is used to give complete location information to trusted nodes. We evaluate the performance of the *Hidden Anchor* algorithm using analysis and simulation. The results show that the *Hidden Anchor* algorithm can hide the location and identity of anchor nodes while maintaining very low overhead.

The rest of this paper is organized as follows. Section II presents the problem statement. Section III details the *Hidden Anchor* algorithm and analyzes its performance. Section IV

* Also affiliated with Alexandria University, Egypt.

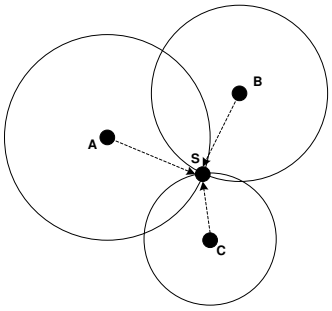

Fig. 1. Node S can estimate its location in 2D using the location messages received from the three anchor nodes A, B, and C and the estimated range to them. Similarly, three un-trusted nodes A, B, and C can cooperate to estimate the location of an anchor node S.

evaluates the *Hidden Anchor* algorithm through simulation. Finally, Section V concludes the paper.

II. Problem Statement

This section outlines the network model and security and privacy requirements. We also delineate the different ways an un-trusted node can use to identify the existence of an anchor node.

A. Network Model

We assume a hybrid wireless sensor network where anchor, trusted, and un-trusted nodes co-exist. We also assume that nodes use an RSS anchor-based localization algorithm. Thus, anchor nodes continuously broadcast beacon packets containing their position information.

Any node in the network can observe any frame transmitted by other nodes within its range. Sensor nodes are randomly distributed in the area of interest. Trusted nodes can use standard encryption algorithms to hide the anchor nodes' position information where both anchor nodes and trusted nodes share the required common information, e.g. cryptographic keys, prior to deployment.

Un-trusted nodes use the same radio hardware used by anchor nodes and trusted nodes. We further assume that there is no correlation between the frame information, such as size and content, with the frame type. Therefore, un-trusted nodes cannot differentiate between the frames from trusted nodes and those from anchor nodes. This can be achieved by encrypting the contents or padding the frames as needed.

We also assume that the goal of the un-trusted nodes is to estimate their range to the anchor nodes based on the physical signal transmitted by anchor nodes.

B. Security and Privacy Requirements

By considering the network model discussed in the previous section, we have two main requirements that should be considered.

1) Location Information Secrecy: Anchor nodes should be able to broadcast their position information periodically and trusted nodes should be able to use this information to estimate their position. On the other hand, un-trusted nodes should not be able to use anchor nodes' beacon packets to gain information about anchor nodes' locations. This can be achieved, for example, by encrypting the anchor nodes' beacon packets.

2) Physical Layer Location Privacy: Un-trusted nodes should not be able to exploit the measured physical signal to estimate the location of anchor nodes. This paper focuses on this privacy requirement.

C. Identifying Anchor Nodes

An un-trusted node may exploit one or a combination of the following vulnerabilities to identify the existence of an anchor node. Note that detecting the existence of the anchor node is a necessary prelude to determining its location.

1) Separate ID-space for anchor nodes: If the WSN is designed such that the ID-space for anchor nodes is separate from the ID-space of trusted nodes, an un-trusted node can identify the existence of an anchor node by its ID in the packet header.

2) Type of transmission - broadcast/unicast: Beacon packets are broadcast in the network. If only anchor nodes broadcast messages in the network, their packets can be distinguished from the packets of other nodes by noting the broadcast destination address. However, regular sensor nodes use both broadcast and unicast to transmit their own messages, making this way less useful for the un-trusted node.

3) Periodicity of packets: Even if other nodes send broadcast packets, the periodicity of the beacon packets make them easier to detect and hence expose the anchor node.

4) Packet size: Beacon packets usually have a fixed size. This can make them easily distinguishable by the un-trusted node.

5) Type field in packet header: If there is a field in the packet header to identify the different message types, this can be used to determine the anchor node by the type of messages it sends.

Except for the last two vulnerabilities, which can be easily mitigated by padding the packet with random data and encrypting the packet respectively, the *Hidden Anchor* algorithm mitigates the remaining three vulnerabilities to achieve anchor node unobservability as discussed in the next section.

III. The Hidden Anchor Algorithm

In this section, we start by a possible technique that addresses the physical layer location privacy problem. We show that this technique has shortcomings, thus motivating the need for a the *Hidden Anchor* algorithm.

A. Possible First-Cut Solution

Anchor nodes can confuse the un-trusted nodes using variable transmission power. For example in [8], we proposed a light-weight algorithm that provides secure anchor-based localization in hybrid wireless sensor networks. The idea is for anchor nodes to continuously and randomly change the transmit power and to include the transmit power encrypted in the frame using the shared genetic information between itself and trusted nodes. This change of transmit power will reduce the localization accuracy at the un-trusted nodes, achieving

location anonymity[1]. Trusted nodes can use the shared genetic information to extract the transmit power from the frames and, therefore, their localization accuracy is not affected by the transmit power change.

Based on the current sensor network hardware, e.g. [9], transmit power can be selected from a set of pre-specified discrete power levels. This has the disadvantage that after receiving a sufficient number of frames, an un-trusted node will be able to distinguish between the different discrete received power levels, thus removing the ambiguity introduced by the random change of transmit power. As a result, un-trusted nodes will be able to localize anchor nodes accurately. In addition, location anonymity provided by this technique is a weaker privacy notion than the unobseravability of anchor nodes provided by the Hidden Anchor algorithm.

B. Hidden Anchor Algorithm

The *Hidden Anchor* algorithm exploits the noisy wireless channel to hide the existence of anchor nodes. The idea is for anchor nodes to use the identity of the nearby trusted nodes when broadcasting their beacon packets. This way, un-trusted nodes cannot differentiate between anchor nodes and trusted nodes. On receiving a packet with ID a, an un-trusted node cannot determine whether this packet is from trusted node a or from an anchor node with the same ID. Note that since the anchor node chooses its ID from nearby nodes, its location is hidden within the noise of the wireless channel. The algorithm operates in two phases: The neighbor discovery phase and the location hiding phase.

1) Neighbor discovery phase: The purpose of this phase is for the anchor node to discover the IDs of the nearest neighbors so that the anchor node can select which IDs to use during the next phase.

The anchor node broadcasts a possibly encrypted identity-request message using a random ID to all its neighbors within a certain radius. This can be controlled by a hop count parameter[2]. All trusted nodes in the network that receive this message reply with an identify-reply message.

The anchor node waits for a certain time to collect the identify-reply messages along with their received signal strength. After that, the anchor node sorts the nodes in an ascending order with respect to their received signal strength and saves the identities of the k nearest trusted nodes in a set ($\mathbb{S}$).

Alternatively, the anchor node can discover the IDs of the first-hop neighbors by passively monitoring the network traffic and uses a random ID from this set for discovering neighbors that are further away.

2) Location hiding phase: In this phase, and when it is time to send a beacon packet, the anchor node chooses one ID from the set $\mathbb{S}$ randomly and uses it as its *unencrypted* identity. The true identity of the anchor node, along with the type of the message can be sent encrypted in the body of the message, if needed. Upon receipt of a broadcast beacon packet, a trusted node decrypts the packet and can determine the identity and location of an anchor node. On the other hand, an un-trusted node, not knowing the decryption key, cannot differentiate between the packets from the anchor nodes and trusted nodes, as they have the same identity and the difference in their signal strength is within the wireless transmission noise.

[1]Location anonymity refers to hiding the true location of an anchor node. This is a weaker notion than anchor node unobservability, where the anchor node existence is not detected at all.

[2]This controls both the size of the neighborhood set and the energy consumption.

C. Discussion

For the vulnerabilities identified in Section II-C, the *Hidden Anchor* algorithm eliminates any chance for un-trusted nodes to identify anchor nodes using their IDs as anchor nodes never use their real IDs in clear, which is equivalent to using only the trusted nodes ID-space. Also, the proposed algorithm removes the periodicity of beacon packets by using a different ID every time for the packets transmitted from anchor nodes. Finally, since anchor nodes clone the identity of trusted nodes, their traffic pattern, as seen by the un-trusted nodes, becomes indistinguishable.

Note also that, even if un-trusted nodes cooperate to determine the location of all trusted nodes, they cannot determine the location of anchor nodes, as anchor nodes are unobservable. Statistical analysis is not useful here too as anchor nodes use the IDs of trusted nodes for sending their own traffic.

Compared to *HyberLoc*, the *Hidden Anchor* algorithm does not depend on changing the power levels and therefore, it is not affected by the current sensor network hardware limitations. Also, while the *HyberLoc* algorithm provides anchor node location anonymity, *Hidden Anchor* algorithm provides the stronger notion of anchor nodes unobservability.

D. Analysis

In this section, we derive an expression for the difference in received signal strength, which is directly related to the average difference in estimated distance, at the un-trusted receiver both when the *Hidden Anchor* algorithm is used and without using it in the presence of noise.

The difference in estimated distance metric represents the error in estimated distance when the trusted node is sending by itself on one hand and when both the trusted node and the anchor node share the same ID, i.e. using the *Hidden Anchor* algorithm. The distance estimate is based on the average received signal strength at the un-trusted node. The lower the value of this metric, the better in terms of privacy requirement as the un-trusted node will be unlikely to detect that both trusted node and anchor node are sharing the same ID.

1) Notation: We use the following notation:

- σ: The variance of the channel white gaussian noise. For a received power of V, the net power received in the presence of noise follows a rician distribution with a mean given by:

$$\mu = \sigma\sqrt{\frac{\pi}{2}}L_{1/2}(-V^2/2\sigma^2) \quad (1)$$

where σ^2 is the noise variance and $L_{1/2}$ denotes a Laguerre polynomial.

- v_{HA}: A random variable representing the average received power over n samples at the un-trusted receiver when the *Hidden Anchor* algorithm is used.
- $v_{\overline{HA}}$: A random variable representing the average received power over n samples at the un-trusted receiver when the *Hidden Anchor* algorithm is not used.
- P_r: A random variable representing the power received at the un-trusted node, whether from the trusted node or the anchor node.
- P_{r_t}: A random variable representing the power received at the un-trusted node from the the trusted node.
- P_{r_a}: A random variable representing the power received at the un-trusted node from the the anchor node.
- r: Traffic ratio, i.e. ratio of the traffic from the trusted node to the traffic from the anchor node.
- V_a: The received power at the un-trusted node from the anchor node in a noiseless environment.
- V_t: The received power at the un-trusted node from the trusted node in a noiseless environment.

2) Difference in Received Power: Using the above notation, the average received power over n samples when the *Hidden Anchor* algorithm is used, v_{HA}, is given by:

$$\begin{aligned} v_{HA} &= \frac{1}{n}\sum_{i=1}^{n} P_{r_i} \\ &= \frac{1}{n}\Big(\sum_{i=1}^{nr/(1+r)} P_{r_{t_i}} + \sum_{i=1}^{n/(1+r)} P_{r_{a_i}}\Big) \end{aligned} \tag{2}$$

where P_{r_i} represents the i^{th} sample from the corresponding random variable. From Eq. 2.

$$\begin{aligned} E[v_{HA}] &= \frac{1}{n}\Big[\sum_{i=1}^{nr/(1+r)} E(P_{r_{t_i}}) + \sum_{i=1}^{n/(r+1)} E(P_{r_{a_i}})\Big] \\ &= \frac{1}{r+1}(rE[P_{r_t}] + E[P_{r_a}]) \\ &= \frac{\sigma}{r+1}\sqrt{\frac{\pi}{2}}[rL_{1/2}(-V_t^2/2\sigma^2) + L_{1/2}(-V_a^2/2\sigma^2)] \end{aligned} \tag{3}$$

Similarly, when the *Hidden Anchor* algorithm is not in use, the expected average received power, $E(v_{\overline{HA}})$, over n samples is given by:

$$E[v_{\overline{HA}}] = \sigma\sqrt{\frac{\pi}{2}}L_{1/2}(-V_t^2/2\sigma^2) \tag{4}$$

By subtracting Eq. 3 from Eq. 4, the expected difference in received power between using the *Hidden Anchor* algorithm and not using it is given by:

$$\begin{aligned} E[\text{Difference in Power Received}] = \frac{\sigma}{r+1}\sqrt{\frac{\pi}{2}}[L_{1/2}(-V_t^2/2\sigma^2) \\ - L_{1/2}(-V_a^2/2\sigma^2)] \end{aligned} \tag{5}$$

IV. Simulation Study

In this section we evaluate the performance of the *Hidden Anchor* algorithm and show the effect of changing different parameters on its performance. All simulation results match the analytical results in Section III-D, which validates the analysis.

A. Simulation Environment

The *Hidden Anchor* algorithm was implemented using Matlab. The sensor nodes were uniformly distributed over a square of 100×100 m^2. Results represent the average over five different network topologies where every network was randomly generated with a different seed. Without loss of generality, we show the results for only one un-trusted node. We use the "difference in estimated distance" metric, mentioned in Section III-D.

B. Simulation Parameters

We evaluate the effect of different parameters on the difference in estimated distance metric. The parameters that we considered in this simulation are:

1) Noise level: This parameter represents an additive white gaussian noise with zero mean and variance given by σ^2.

2) Traffic ratio: This parameter (r) represents the ratio of the traffic sent by the trusted node to the traffic sent by the anchor node.

3) Neighborhood radius: This parameter represents the maximum distance between the anchor node and the trusted nodes whose IDs it clones. Note that the anchor nodes picks its identity randomly from this set of neighbors.

4) Un-trusted distance: This parameter represents the distance between the anchor node and the un-trusted node.

C. Results

Fig. 2 shows the effect of changing the "neighborhood radius" parameter on the difference in estimated distance for different noise levels. The rest of the parameters were fixed at network density= 1/25 node/m^2, un-trusted distance= 70m, and traffic ratio= 1:1. The figure shows that as expected, as the neighborhood radius decreases, the difference in estimated distance decreases, leading to better security.

The figure also shows that as the noise level increases, the un-trusted node will not be able to distinguish between the physical signal sent by the anchor node and the physical signal sent by the trusted node.

Fig. 3 shows the effect of changing the neighborhood radius for different traffic ratios. The other parameters were fixed at network density= 1/25 node/m^2, un-trusted distance= 70m, and noise variance= 1e-4. The results show that as the traffic sent by the trusted node increases, relative to the traffic sent by the anchor node, the difference in the estimated distance decreases. Since typical anchor-based algorithms use low anchor traffic to trusted nodes traffic ratio, this shows the promise of the proposed *Hidden Anchor* algorithm.

Fig. 4 shows the effect of changing the neighborhood radius with different un-trusted distance values. The other parameters

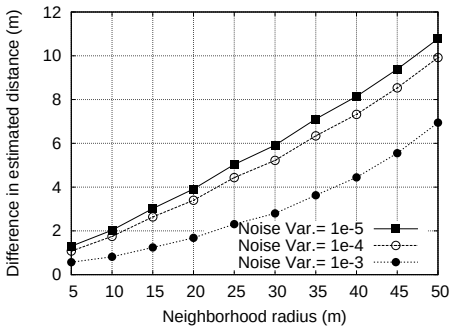

Fig. 2. Effect of changing the "noise level" and the "neighborhood radius" on the estimated distance at the un-trusted node.

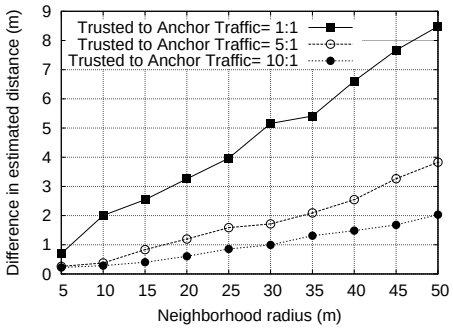

Fig. 3. Effect of changing the "traffic ratio" and the "neighborhood radius" on estimated distance at the un-trusted node.

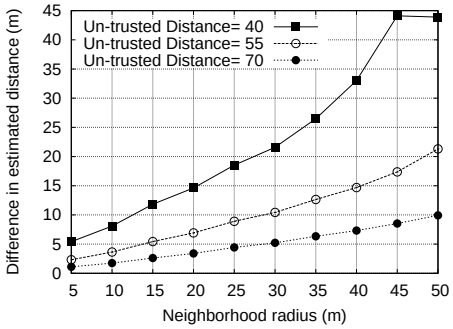

Fig. 4. Effect of changing the "un-trusted distance" and the "neighborhood radius" on estimated distance at the un-trusted node.

were fixed at network density= 1/25 node/m^2, traffic ratio= 1:1, and noise variance= 1e-4. The figure shows that as the distance between the anchor node and the un-trusted node increases, the difference in the estimated distance decreases. Consequently, the un-trusted node will not be able to differentiate between the physical signal transmitted by the anchor node and the physical signal transmitted by the trusted node. The reason is that the effect of the distance difference on signal strength between the anchor node and the neighboring trusted nodes diminishes as we go away from the anchor node.

D. Summary

In this section, we have evaluated the performance of the *Hidden Anchor* algorithm for different parameters. The noise level, traffic ratio, the maximum distance between the anchor node and the neighboring trusted nodes, and the distance between the anchor node and the un-trusted node.

The designer of the network can only control the traffic ratio and the maximum distance between the anchor node and the neighboring trusted nodes. Reducing the anchor node to trusted nodes traffic ratio enhances security but may decrease the localization accuracy at trusted nodes.

The results show that we can make the performance metric arbitrary small, indicating better privacy, by controlling these two parameters.

V. Conclusion

In this paper, we focused on the physical layer location privacy problem, where an anchor node wants to hide its physical signal information from un-trusted nodes, while at the same time allows trusted nodes to benefit from this information. We proposed the *Hidden Anchor* algorithm for solving the physical layer location privacy and evaluated its performance through analysis and simulation experiments. Our results show that the *Hidden Anchor* algorithm can effectively hide the location and identity of anchor nodes without limiting the localization accuracy for trusted anchor nodes, thus providing anchor nodes' unobservability. Currently, we are investigating the performance of the *Hidden Anchor* algorithm under different metrics applicable to a wider range of localization algorithms. In addition, we are investigating the performance on the algorithm in indoor environments under other possible attacks.

Acknowledgement

This work is supported in part by QNRF under grant number NPRP-1-7-7-3.

References

[1] B. Karp and H. Kung, "Greedy perimeter stateless routing for wireless networks," in *International Conference on Mobile Computing and Networking*, 2004.

[2] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. B. Srivastava, "Coverage problems in wireless ad-hoc sensor networks," in *IEEE INFOCOM*, 2001, pp. 1380–1387.

[3] A. Youssef and M. Youssef, "A taxonomy of localization schemes for wireless sensor networks," in *The International Conference on Wireless Networks*, 2007.

[4] P. Liu, X. Zhang, S. Tian, Z. Zhao, and P. Sun, "A novel virtual anchor node-based localization algorithm for wireless sensor networks," in *The International Competition Network Conference*, 2007.

[5] K. Langendoen and N. Reijers, "Distributed Localization in Wireless Sensor Networks: A Quantitative Comparison," *Computer Networks (Elsevier), special issue on Wireless Sensor Networks*, pp. 374–387, August 2003.

[6] S. Capkun, M. Hamdi, and J.-P. Hubaux, "Gps-free positioning in mobile adhoc networks," in *Hawaii International Conference on System Sciences (HICSS-34)*, January 2001, pp. 3481–3490.

[7] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less Low-cost Outdoor Localization for Very Small Devices," *IEEE Personal Communications*, vol. 7, no. 5, pp. 28–34, October 2000.

[8] M. Adel, M. Ibrahim, K. Abulmakarem, M. Youssef, and M. Eltoweissy, "Hyberloc:demonstrating secure localization in hybrid sensor networks," in *The International Conference on Mobile Computing and Networking*, 2008.

[9] "Telosb mote platform." [Online]. Available: http://www.xbow.com