Detecting Receiver Attacks in VRTI-Based Device Free Localization

Arijit Banerjee^{*}, Manas Maheshwari[†], Neal Patwari[‡] and Sneha Kumar Kasera^{*} *School of Computing, University Of Utah [†]System Software Engineer, Nvidia Corporation [§]Electrical & Computer Engineering, University of Utah

Abstract-Variance-based Radio Tomographic Imaging (VRTI) is an emerging technology that locates moving objects in areas surrounded by simple and inexpensive wireless sensor nodes. VRTI uses human motion induced variation in RSS and spatial correlation between link variations to locate and track people. An artificially induced power variations in the deployed network by an adversary can introduce unprecedented errors in localization process of VRTI and, given the critical applications of VRTI, can potentially lead to serious consequences including loss of human lives. In this paper, we tackle the problem of detecting malicious receivers that report false RSS values to induce artificial power variations in a VRTI system. We use the term "Receiver Attack" to refer to such malicious power changes. We use a combination of statistical hypothesis testing and heuristics to develop real-time methods to detect receiver attack in a VRTI system. Our results show that we can detect receiver attacks of reasonable intensity and identify the source(s) of malicious activity with very high accuracy.

I. INTRODUCTION

Variance-based Radio Tomographic Imaging (VRTI) [1] is an emerging technology that locates moving objects in areas surrounded by simple and inexpensive wireless sensor nodes. Figure 1 shows such a VRTI network setup. Human motion in vicinity of wireless links cause variation in the link Received Signal Strength (RSS). VRTI uses this motion induced variation in RSS and spatial correlation between link variations to locate and track people. The advantage of VRTI is that it is "device free", i.e., it does not require the moving object(s) to carry any electronic device and it can locate people even behind walls. Hence, VRTI can be used in many critical applications including emergency response, rescue operations, and security breaches. However, the VRTI-based device free localization (DFL) scheme assumes that the changes in the RSS measurements is caused by the movement of objects being localized and/or the environment noise only. An adversary can craftily introduce artificial variations in RSS measurements by maliciously programming some sensor nodes to vary their transmit power or report wrong RSS values as measured for some links. Such attacks can introduce significant errors in a VRTI-based DFL system and, given the critical applications of the system, can potentially lead to serious consequences including loss of human lives (e.g. in a hostage situation). Timely and accurate detection of adversarial nodes is of extreme importance for a VRTI system.



Figure 1. Sensor network layout in a VRTI system. The sensor nodes are depicted by (o). The area being monitored is bounded by the dashed line.

An adversary can make a make a person appear in the empty room by reporting false received RSS values for an otherwise static link. It can also suppress variation in a link, caused by human motion, by reporting constant RSS for that link. The importance of securing VRTI system, against adversaries, motivates us to explore the effect of power attacks in this system and develop efficient techniques to detect such attacks.

Notably, certain non-adversarial circumstances can also lead to unexpected changes in transmit power and/or reported RSS. These circumstances include faults in sensor nodes due to physical damage or software bugs. In our previous work [2], we presented methodologies to detect malicious transmitters who vary their transmit power to introduce errors in RSS-based localization system. In this paper, we tackle the problem of detecting unexpected changes in reported RSS, malicious or otherwise, in a VRTI system. We use the term *receiver attack* to denote unexpected changes in the reported RSS. We use a combination of statistical hypothesis testing and heuristics to develop realtime methods to detect the presence of receiver attack and identify the adversarial sensor node(s). In contrast to some existing methods (e.g., [3]), our method does not rely on any training data. This makes our method more usable in dynamic environments where training data may get obsolete frequently. In order to evaluate the effectiveness of our receiver attack detection method, we perform extensive experiments in indoor settings using a network of 802.15.4 compliant wireless sensor nodes (Zigbee). We find that using our methods, we can achieve close to zero probability of missed detections and the probability of false alarms in a VRTI setting. We achieve close to 100% detection rate for *receiver attacks*, intense enough to falsify the image in VRTI, by using only 3.5 seconds of measurement data. We are able to identify the source of malicious activity with more than 95% accuracy. To the best of our knowledge, we are the first one to study the effect of adversarial receiver nodes in VRTI-based DFL and provide methods to detect these. Our proposed methods are quite general and can be applied to RSS-based localization schemes other than VRTI.

The remainder of this paper is organized as follows. In Section II, we list our assumptions and describe our adversary model. In Section III, we formulate our method to detect receiver attacks and present the criteria used to evaluate our method. In Section IV, we present our experiments. In Section V, we present evaluation results of our detection method. In Section VI, we discuss previous research in detection of attacks in RSS based localization methods and conclude the paper and indicate directions for future work in Section VII.

II. ASSUMPTIONS AND ADVERSARY MODEL

We assume that faulty or malicious nodes are never present in majority in the network and all nodes have equal probability of developing fault or being targeted by an adversary. We allow multiple adversaries to be active at the same time but they do not collude with each other to carry out a coordinated receiver attack. Since faulty nodes are just a weaker form of the adversary being considered, all further discussions apply to both malicious and faulty nodes. We define *Receiver Attack* as follows:

 When an adversary maliciously programs one or more nodes to report false RSS values as measured for other nodes.

In the following subsection, we parameterize the action of malicious power change of the attackers.

A. Receiver Attack Parameters

In case of a receiver attack the adversary has control over the links she chooses to attack. We first define *minimum periodicity of attack*, w_{min}^r , as the smallest set of contiguous transmissions from a malicious receiver that contain at least one false RSS readings for the links being attacked.

Next, we define *minimum receiver attack amplitude*, a_{min}^r , as the minimum value by which a receiver should change the RSS measured for a link to perform an attack with significant effect on the VRTI accuracy. Changes with an amplitude less than a_{min}^r are not considered to be harmful to the application in any significant manner, and thus are not important to detect. Last, we define *minimum number* of links attacked, L_{min} , as the number of transmitter nodes for which a malicious receiver reports wrong RSS values. An attacker can inflict more damage if it concentrates on



Figure 2. Relative RMSE in VRTI Image (under receiver attack)

changing the RSS values of a few transmitter nodes that are placed close to it, i.e., for links of shorter lengths.

Figure 2 shows the impact of a receiver attack on VRTI accuracy. The root mean square error, RMSE, plotted on the y-axis, corresponds to the relative error in VRTI localization introduced by a receiver attack. Here, $L_{min} = 2$. We see from Figure 2 that RMSE increases with decreasing value of w_{min}^r and increasing value of a_{min}^r . For $a_{min}^r = 10dB$, the RMSE is negligible ($\approx 0.1m$). However, RMSE increases to more than 1m for $a_{min}^r = 13$ at low w_{min}^r . For moderate values of w_{min}^r (e.g., 5), we observe significantly high RMSE when the attack amplitude $\geq 15dB$. In the rest of this paper, we will focus on detecting receiver attacks that introduce high RMSE in VRTI.

III. ATTACK DETECTION METHODOLOGY

We consider a VRTI system built on a WSN with Ntransceiver nodes. Recall that human motion in vicinity of wireless links causes variation in the link Received Signal Strength (RSS). VRTI uses this motion induced variation in RSS and spatial correlation between link RSS variations to locate and track people. We define, for a transmitter k, a neighbor set given by $\mathcal{H}_k = \{n_0, n_1, ..., n_{M-1}\}$ consisting of M receivers capable of communicating with k. We make RSS measurements on each link between node pair (k, n_l) where $n_l \in \mathcal{H}_k$. A fully connected network is not required for our detection method, however, the neighbor set for each transmitter is assumed to be known at all nodes and remain constant. Detection in networks where \mathcal{H}_k can change with time will be considered in future work. Let $r_{k,j}(i)$ be the RSS measured for link $l_{k,j}$ at receiver j for transmission from node k at time i where $k \in \{1, ..., N\}$ and $j \in \mathcal{H}_k$. We define RSS vector as:

$$\mathbf{r}_{k}(i) = [r_{k,n_{0}}(i), ..., r_{k,n_{M-1}}(i)]^{\mathsf{T}}$$
(1)

and mean of RSS vector over a window of time T as:

$$\bar{\mathbf{r}}_{k}(i) = \frac{1}{T} \sum_{t=1}^{T} \mathbf{r}_{k}(i-t) = [\bar{r}_{k,n_{0}}(i), ..., \bar{r}_{k,n_{M-1}}(i)]^{\mathsf{T}} \quad (2)$$



Figure 3. Probability density function of Δ_{rdls}

We define the absolute mean deviation from mean RSS for a transmission of node k at time i as

$$\Delta \mathbf{r}_k(i) = \frac{1}{T} \sum_{t=1}^T |\mathbf{r}_k(i-t) - \bar{\mathbf{r}}_k(i)|$$
(3)

$$= [\Delta r_{k,n_0}(i), ..., \Delta r_{k,n_{M-1}}(i)]^{\mathsf{T}}$$
 (4)

We detect receiver using reciprocal difference in $\Delta \mathbf{r}_k(i)$ values in both directions of a wireless link and use statistical hypotheses testing to decide between an attacked and a normal environment as explained below:

A. Detecting Receiver Attack

Using Equation (3) we calculate,

$$\Delta_{rdls}(l_{k,j}(i)) = \Delta r_{k,j}(i) - \Delta r_{j,k}(i)$$
(5)

where $\Delta_{rdls}(l_{k,j}(i))$ is the reciprocal difference of absolute mean deviation in RSS for the wireless link between node k and node j at time i. In the rest of this paper, we use the term "RDLS" (Reciprocal Difference in Link Statistic) to refer to $\Delta_{rdls}(l_{k,j}(i))$.

Under normal circumstances, the multi-path properties and hence the RSS variations of a radio channel at any point in time are expected to be identical in both directions of a link. Some of the existing secret key extraction and antenna design techniques utilize this reciprocity property of wireless links (e.g., [4]-[7]). However, for the reciprocity property to hold good, the time gap between measurements of the link by its two ends points should be very small¹. In our system, in the absence of any power attack, all the nodes have almost identical variations in RSS in both directions of a link as these are measured within a very short time interval. Therefore, when no attack nodes are present, the mean absolute deviations are likely to be similar in both direction of a link. This implies that RDLS values will be very small for most of the links. Figure 3 shows that, in a no-attack VRTI scenario, the distribution of the RDLS values is mostly concentrated around the mean, the mean value being zero. The probability of RDLS values, as we move further away from the mean, decreases exponentially. A few moderately high RDLS values can be attributed to the movement in the environment and random noise.

However, in the case where a malicious node j reports wrong RSS for node k, the absolute mean deviation in measured RSS will be asymmetrical for links $L_{k,j}$ and $L_{j,k}$. Hence, the corresponding RDLS values will be higher. Figure 3 shows that under receiver attack, percentage of high RDLS values increases significantly in comparison to the normal case.

These observations suggest that one can detect receiver attacks by examining the extreme values in the *RDLS* distribution for all pairs of links, as long as the intensity of the attack is large enough to make the extreme values of the distribution in the attack case appear different from those in the non-attack case. Recall from Section II that, to have any significant impact on the VRTI accuracy, a receiver attack should have a_{min}^r and w_{min}^r values above 15dB and 5, respectively. We now present our methodology to detect attacks as follows.

1) Attack Detection: For detecting receiver attacks, we use a generic outlier detection technique [8] as follows: Let, at time i, $\tilde{\delta}(i)$ denote the median of $\Delta_{rdls}(l_{k,j}(i))$ values for all possible (k, j) pairs and MAD denote the median absolute deviation. Then,

$$MAD(i) = median(|\Delta_{rdls}(l_{k,j}(i)) - \tilde{\delta}(i)|)$$
 (6)

We define *z*-score [8] for each $\Delta_{rdls}(l_{k,j}(i))$ as follows:

$$z_{k,j}(i) = \frac{C(\Delta_{rdls}(l_{k,j}(i)) - \delta(i))}{MAD(i)}$$
(7)

where C is constant with value 0.6745 [8]. We consider an observation $\Delta_{rdls}(l_{k,j}(i))$ to be an outlier if the corresponding $z_{k,j}(i)$ score is greater than some pre-defined threshold. The rationale behind choosing this method is that the parameters (median absolute deviation) used to detect the outliers are minimally affected by the actual outlier values. Let $\delta_{max}(i) = max_{k,j} (\Delta_{rdls}(l_{k,j}(i)))$ denote the maximum RDLS value among all links at time *i*. We calculate

$$z_{max}(i) = \frac{C(\delta_{max}(i) - \delta(i))}{MAD(i)}$$
(8)

using (7). As in the case of transmitter attack detection, we consider deciding between two hypotheses:

- *H*^{*r*}₀: There is no receiver attack present in the environment
- H_1^r : There is at least one receiver attack present

We choose between H_0^r and H_1^r based on the value of $z_{max}(i)$ as follows:

$$z_{max}(i) \stackrel{H_1^r}{\underset{H_1^r}{>}} \alpha \qquad (9)$$

¹Other factors that also impact the reciprocity property [4], do not apply here.

where α is an experimentally determined threshold that separates the boundary between attacked and normal environment. A lower value of α decreases the probability of missed detections but at the same time increases the probability of false alarms (detecting an attack when there is none). Similarly, a high value of α decreases the probability of false-alarms (P_{FA}) but at the same time increases the probability of missed-detection (P_{MD}). We find that $\alpha = 10$ delivers a good trade-off in terms of P_{MD} and P_{FA} in our experiments.

2) Identifying malicious receiver nodes: Once we detect the presence of a receiver attack, our next important task is to identify the source(s) of the malicious activity. We do this by identifying links that contribute to extreme reciprocal differences. We compute the **z-score**s of reciprocal differences for all links $l_{k,j}$ using (7). If the **z-score** value is close to the value $z_{max}(i)$, we identify node j as a potential malicious node and put it in a "evidence list" *EL*. *EL* is a table which is indexed by the node-id. The table entry corresponding to a node j is incremented by the RDLS value corresponding to the link $l_{k,j}$. After examining all the links, we scan through *EL* and identity nodes with values above a certain threshold as malicious nodes.

The algorithm to update the evidence list EL and detecting malicious nodes is given in Algorithm 1. In the

Algorithm 1 Detecting Malicious Nodes
Let L be the set of all links
Initialize evidence list EL as an empty hash table
for each $l_{k,j} \in L$ do
$v \leftarrow \Delta_{rdls}(l_{k,j}(i))$ (using (5))
calculate $z_{k,j}(i)$ (using (7))
if $ z_{k,j}(i) - z_{max}(i) \leq \beta$ then
if j not in EL then
$EL[j] \leftarrow v$
else
$EL[j] \leftarrow EL[j] + v$
end if
end if
end for
for each key j in EL do
if $EL[j] \ge \rho$ then
flag node j as an malicious node
end if
end for

algorithm 1, β and ρ are parameters that must be determined experimentally. We use $\beta = 3$ and $\rho = 3$ for our experiments.

If there is only one attack link, $l_{k,j}$, our algorithm will flag both *i* and *j* as malicious nodes. Thus, for a single link attack case, we are able to identify the malicious link but it is not possible to identify the malicious node itself. As the number of links affected by an attacker node increases, the cumulative value of RDLS corresponding to that node also increases. This is because the evidence list gets updated for all instances where the malicious node impacts a link measurement. We can identify the malicious node by observing the increased cumulative RDLS value in the evidence list.

IV. EXPERIMENTS

In this section, we describe the testbed that we use and the experiments we perform for evaluating our methodology to detect receiver attacks on a VRTI system.

A. Experimental Setup

We deploy 20 TelosB wireless sensors nodes uniformly in a $6m \times 6m$ square area to form a WSN in a classroom (without any students) in our engineering building. These sensor nodes operate in the 2.4 GHz frequency band using the IEEE 801.2.15.4/ZigBee protocol. We use a round-robin token-passing protocol, called *spin* to schedule transmission of nodes in a manner that prevents packet collisions while still maintaining high data collection rate. When one node transmits, all other nodes receive the packet and measure RSS. These RSS measurements are transmitted to a base station 4.08 times/second by each node along with its unique ID (from 1-20). The base station collects all RSS measurements and forwards the data to a laptop for storage and processing. Each *spin cycle* [1] consists of RSS dataset with exactly one transmission from every transmitter node.

B. Experimental Details

We consider the following experimental scenarios:

1) No-Attack: During this experiment, no power attack is present in the network. A subject, timed by a metronome, walks in a known path at a constant speed of 0.6m/s in the deployed area for a 4 minute period.

2) Receiver-Attack: We use the data collected from the No-Attack experiment to simulate receiver-attack scenarios. We choose one node at random from the set of deployed nodes and change some of the values in its transmitted data that contains the measured RSS values of other nodes. The number of values changed in the transmitted vector and the amount of change in the RSS values correspond to the L_{min} and a_{min}^r parameters, respectively, as described in Section II-A. We change the values in such a way that each set of w_{min}^r transmissions from the malicious node contains at least one malicious transmission. We carry out the experiments for different combination of values of w_{min}^r , L_{min} and a_{min}^r that can cause significant errors in localization in VRTI. Specifically, we experiment with $w_{min}^r =$ 3, 5, and 7; $L_{min} = 1$, 2, and 4 and $a_{min}^r = 10 dB$, 15 dB, and 20dB. We repeat the same experimental procedure by selecting more number of nodes (up to ten nodes) to act as adversaries to test the scalability of our method.



Figure 4. ROC curves for different (a) a_{min}^r , (b) w_{min}^r and (c) L_{min}



Figure 5. Identifying malicious receivers

V. RESULTS

At any time instant i, we differentiate between an attack scenario and a normal scenario using the maximum RDLS values for T measurement cycles as described in Section III. We choose T = 15. Figure 4 shows the effect of the three different parameters a_{min}^r , w_{min}^r , and L_{min} on detector performance. We observe that the detection performance improves as the amplitude of receiver attack (a_{min}^r) increases. This is because as the value of a_{min}^r increases, it contributes to more asymmetrical RSS measurements in both directions of an attacked link. Similarly, as the value of w_{min}^r decreases, more number of malicious transmissions fall under the measurement window thereby improving the detection performance. Probability of detecting an attack also increases with increasing values of L_{min} as more number of links contribute to the higher RDLS values. These results show that we can achieve more than 80% probability of detection with less than 1% probability of false alarms when $a_{min}^r = 15 dB$ and $w_{min}^r = 5$. For $a_{min}^r = 20 dB$ and $w_{min}^r = 5$, close to 100% probability of detection is achieved with negligible false alarm rate. We also see that the probability of detection is low for $a_{min}^r = 10 dB$ and $w_{min}^r = 7$. However, at these values there is no significant degradation in VRTI performance and hence we can ignore these.

Once a receiver attack is detected using the hypotheses testing (as described in Section III-A), we apply the al-

gorithm (1) to identify the source(s) of malicious activity in a certain measurement window. For each node in the evidence list EL, we plot (Figure 5) its corresponding value from the table. This value is the cumulative RDLS (CRDLS) contribution of the node in the chosen measurement window. We also plot the experimentally determined threshold ρ . If a node's CRDLS value lies above the threshold, we identify it as part of some malicious link. In figure 5(a), node 0 is a malicious receiver that reports false RSS for node 1. We see from this figure that the values for node 0 and node 1 lie above the chosen threshold, and hence we identify link $l_{0,1}$ as an attack link. Note that in this case, node 6 is also in the evidence list, possibly due to noise and/or movement in the environment, but its CRDLS contribution is less than the chosen threshold. Therefore, node 6 it is not identified as a malicious node. Figure 5(b) shows the case where node 0 is malicious and reports false RSS for four of its neighboring nodes (node 1, node 2, node 18 and node 19). This figure shows that our chosen threshold identifies all the attack links. Moreover, the high L_{min} value for node 0 implies that its cumulative contribution in the evidence list is much more compare to the other nodes and thus, node 0 can be singled out as a malicious node. Figure 5(c) shows the scenario when there are 5 malicious nodes - node 0 (reports false RSS for node 1 and node 2), node 2 (reports false RSS for node 3 and node 4), node 6 (reports false RSS for node 7 and node 8), node 8 (reports false RSS for node 9 and node 10) and node 14 (reports false RSS for node 15 and node 16). As before, we are able to identify all the attack links using our detection algorithm.

Our algorithm enables us to identify the attack links even when an attacker reports false RSS for a few number of nodes (small L_{min}). As the value of L_{min} increases, we can single out the attacker node by increasing the value of the threshold ρ . We set our thresholds to detect the lowest intensity receiver attacks that can cause significant error in the VRTI system II-A. Our method ensures that any higher intensity attack will be caught with a high success rate. We use only 15 spin cycles, which corresponds to approximately 3.75 seconds of measurement data to detect an attack.

VI. RELATED WORK

Significant work has involved securing WSNs using traditional key based authentication and encryption protocols [9] [10]. These methods provide some level of security as long as the adversary is assumed not to gain physical control over the sensor nodes. Other works on secure localization include SPINE [11], ROPE [12], SeRLoc [13] and HirLoc [14]. However, these methods are vulnerable to capture of critical nodes by the adversary. Perhaps the work that comes closest to ours is by Chen et al. [15] which proposes a generic method for two broad range of active localization methods: multilateration based and RSS based. Their method, however, can not be applied to VRTI-based DFL. In contrast to their work, our method does not use any training data which makes it more suitable for dynamic environments. In our previous work [2], we developed a preliminary framework to detect transmitter-attacks and applied it to a VRTI setting. In this paper, we complement this previous work to study the effect of receiver attacks in VRTI based localization systems and include methodologies to detect receiver-attacks and identify source of malicious activity.

VII. CONCLUSION

In this paper, we present statistical hypotheses based detection models and heuristics to detect the presence of receiver attacks and for identifying the attacked links in a VRTI-based DFL system. Our experimental results show that using our methods, we can achieve close to zero percent probability of false alarms and missed detections for attacks intense enough to cause significant impact on the results on VRTI based DFL systems. In the future, we will examine the effect of combined transmitter and receiver attack problem in the presence of colluding adversaries.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under CPS Grant No. 1035565, and by the ONR/ARL MURI Grant W911NF-07-1-0318.

REFERENCES

- J. Wilson and N. Patwari, "See Through Walls: Motion Tracking Using Variance-Based Radio Tomography Networks," *IEEE TMC*, 2010.
- [2] M. Maheshwari, S. Ananthanarayanan, A. Banerjee, S. K. Kasera, and N. Patwari, "Detecting malicious nodes in rssbased localization,," in DSPAN '11.
- [3] Y. Chen, K. Kleisouris, X. Li, W. Trappe, and R. Martin, "The robustness of localization algorithms to signal strength attacks: a comparative study," *Distributed Computing in Sensor Systems*, pp. 546–563, 2006.
- [4] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "On the effectiveness of secret key extraction from wireless signal strength in real environments,," in ACM Mobicom '09.
- [5] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "PARADIS: Wireless device identification with radiometric signatures," in *ACM Mobicom* '08.
- [6] B. Azimi-Sadjadi, A. Kiayias, A. Mercado, and B. Yener, "Robust key generation from signal envelopes in wireless networks," in ACM CCS '07.
- [7] W. Stutzman and G.A.Theile, *Antenna Theory and Design*. John Wiley and Sons, 1981.
- [8] B. Iglewicz and D. C. Hoaglin, *How to detect and handle outliers*. ASQC Quality Press (Milwaukee, Wis.), 1993, vol. 16.
- [9] K. Jamshaid and L. Schwiebert, "Seken (secure and efficient key exchange for sensor networks)," in *IEEE International Conference on Performance, Computing, and Communications*, 2004.
- [10] A. Khalili, J. Katz, and W. Arbaugh, "Toward secure key distribution in truly ad-hoc networks," in *Applications and the Internet Workshops*, 2003.
- [11] S. Capkun and J. Hubaux, "Secure positioning of wireless devices with application to sensor networks," in *Proceedings* of 24th Annual Joint Conference of the IEEE Computer and Communications Societies., vol. 3, 2005, pp. 1917–1928.
- [12] L. Lazos, R. Poovendran, and S. Čapkun, "ROPE: robust position estimation in wireless sensor networks," in *Proceedings* of the 4th international symposium on Information processing in sensor networks, 2005, p. 43.
- [13] L. Lazos and R. Poovendran, "SeRLoc: Secure rangeindependent localization for wireless sensor networks," in *Proceedings of the 3rd ACM workshop on Wireless security*, 2004, pp. 21–30.
- [14] —, "HiRLoc: High-resolution robust localization for wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 2, pp. 233–246, 2006.
- [15] Y. Chen, W. Trappe, and R. Martin, "Attack detection in wireless localization," in *IEEE INFOCOM 2007*.