Poster: PCA-based statistical anomaly detection of reactive jamming in WiFi networks

Ni An
ECE Department
Drexel University
Philadelphia, PA 19104

Vinod Mishra
Computational and Information Sciences Directorate
U.S. Army Research Laboratory
Aberdeen Proving Ground, MD 21005

Steven Weber
ECE Department
Drexel University
Philadelphia, PA 19104

***Abstract*—Our work focuses on detecting reactive jamming attacks in a WiFi (802.11x) infrastructure wireless local area networks (WLANs) using the principal component analysis (PCA) based statistical anomaly detection algorithm. Simulations using the ns-3 network simulator are performed to evaluate the detection effectiveness. Simulation results demonstrate that the PCA-based approach can effectively identify reactive jamming attacks even when the reactive jamming activity is extremely stealthy, and its detection accuracy is superior to a previous approach. We also propose a novel variant on PCA, called cautious PCA, enabling an anomaly detector to self-assess the reliability of its estimated labels.**

I. Introduction

Denial of service (DoS) attack is one of the biggest threats against the normal operation of wireless networks. Reactive jamming is a type of DoS attack that can disrupt legitimate wireless communications, and is known for its detection difficulty, compared with other wireless jamming attacks (e.g., constant or random jamming) [1]. We study the performance of a statistical anomaly detection (SAD) method to detect reactive jamming in a 802.11g wireless local area network (WLAN).

Malicious attacks might change the correlation between various network features, such as the signal strength, packet counts, etc., and might be difficult to be detected by a simple thresholding technique of individual features. Principal component analysis (PCA) is superior to thresholding in that it utilizes feature correlation to identify anomalies, which are defined as events that deviate significantly from the normal patterns of network events. PCA-based SAD has been widely applied in detecting network-wide volume anomalies on (wired) network backbone links (e.g., [2]). The main idea of PCA is to learn a lower dimensional subspace, called the *principal subspace*, which captures most of the variance of a dataset, as the normal profile. Samples with large distances from the principal subspace are labeled as anomalies.

The WLAN consists of a single access point (AP), a single reactive jammer (RJ) and nine stationary wireless stations (STAs), illustrated in Fig. 1. The behavior of a RJ (Fig. 2), is defined by the triple (p_J, τ, T): it persistently listens to the wireless channel and, upon detection of a target transmission from a legitimate STA, decides with probability p_J to jam the transmission for duration T after an initial reaction time τ.

This research is supported by the National Science Foundation under award #CNS-1228847. S. Weber is the contact author: sweber@coe.drexel.edu.

II. Model

Consider an 802.11g WLAN in infrastructure mode. The AP acts as the anomaly detector, which monitors m stations and p distinct statistics (i.e., features) at each station. We fix $p = 5$ features: 1) the number of successfully received packets, 2) the total number of dropped packets, 3) the number of dropped packets due to corruption (i.e., packets that fail the CRC check, possibly on account of a collision), 4) the maximum duration between two consecutive successfully received packets [3], and 5) the average received signal strength (RSS) of dropped packets due to corruption. Therefore, for a network with m wireless stations, each sample has $n \equiv pm = 5m$ features. Let x_i^j be the length-p measurement vector of station j at time t_i and let $x_i = (x_i^j, j \in [m]) \in \mathbb{R}^n$ be the concatenation of the m station measurements at time t_i.

Our approach is semi-supervised, meaning only normal (non-jamming) samples are needed to train the anomaly detector. After N time instants $(t_1, ..., t_N)$, the AP can form a matrix $X \in \mathbb{R}^{N \times n}$ comprised of the N points (rows) $X = (x_1, ..., x_N)$. Assuming that X contains only clean samples, we can use X for training, and subtract the mean. The PCA-based anomaly detector learns the sample covariance matrix $S \equiv \frac{1}{N} X^\intercal X$, and then performs eigen-decomposition $S = \hat{V}\hat{\Lambda}\hat{V}^\intercal$, where the columns of $\hat{V}$ are the sample eigenvectors, and the diagonal matrix $\hat{\Lambda}$ holds the eigenvalues, sorted in order of decreasing magnitude. The sample *principal subspace* $\hat{\mathcal{S}}$ of dimension $k \leq n$ is defined by the leading k sample eigenvectors $\hat{V}^{(k)}$. PCA-based SAD is established on the assumption that normal patterns mainly lie in the principal subspace [2]. We use the *sample Q-statistic*, defined as the squared Euclidean distance to the principal subspace, to measure the abnormality of a sample $x \in \mathbb{R}^n$: $Q_{\hat{\mathcal{S}}}(x) \equiv \left\| x - \Pi_{\hat{\mathcal{S}}}(x) \right\|^2$, where $\Pi_{\hat{\mathcal{S}}} \equiv \hat{V}^{(k)}\hat{V}^{(k)\intercal}x$ is the projection of x onto $\hat{\mathcal{S}}$. Fixing a threshold $q > 0$, x is *labeled normal* if $Q_{\hat{\mathcal{S}}} \leq q$ or *labeled anomalous* if $Q_{\hat{\mathcal{S}}} > q$.

A. Cautious PCA

Since the sample principal subspace $\hat{\mathcal{S}}$ is learned from S, which is an estimate of the population covariance matrix, it might not accurately capture the normal patterns, i.e., anomalies detected using $\hat{\mathcal{S}}$ might not be true anomalies. Thus we propose a simple *cautious PCA* scheme wherein the anomaly

detector only assigns "*certain*" labels to points with distance $\sqrt{Q_{\hat{\mathcal{S}}}(x)}$ not within $\epsilon\sqrt{q}$ of $\sqrt{q}$, for a parameter $\epsilon > 0$ set by the detector. Points with distance $|\sqrt{Q_{\hat{\mathcal{S}}}(x)} - \sqrt{q}| \leq \epsilon\sqrt{q}$ are "*uncertain*" in the sense that the possibility of mislabeling such points, which are close to the PCA thresholding surface, is high. Cautious PCA thus assigns one of three labels to each point: i) (certain) normal if $\left\|x - \Pi_{\hat{\mathcal{S}}}(x)\right\|^2 \leq (1-\epsilon)^2 q$, ii) (certain) abnormal if $\left\|x - \Pi_{\hat{\mathcal{S}}}(x)\right\|^2 > (1-\epsilon)^2 q$, or iii) uncertain if $(1-\epsilon)^2 q < \left\|x - \Pi_{\hat{\mathcal{S}}}(x)\right\|^2 \leq (1+\epsilon)^2 q$.

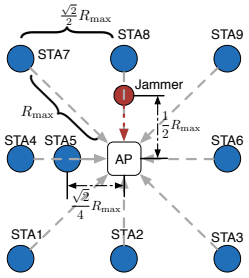

Fig. 1: WLAN topology.

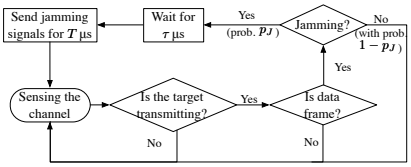

Fig. 2: Flowchart of the RJ.

III. Experiments

Simulations are performed using the ns-3 (version 3.21). The reactive jamming behavior is developed based upon the wireless jamming model [4] with two important modifications: i) instead of jamming all wireless stations, a RJ can be configured to only target a specific station or a set of stations, and ii) the RJ will ignore control frames (e.g., acknowledgements) and only jam data frames.[1] Simulation parameters are listed in Table I. By scaling the separation distance $R_{\max}$ between the STAs and keeping the transmission power constant, we effectively "scale" the likelihood of *hidden terminals* and thereby change the difficulty of accurate anomaly detection. Each simulation runs for 1000 seconds and the AP aggregates statistics for every STA over a time bin of $2s$. Due to space limitation, this abstract only shows the case of RJ without a specific target. Fig. 3 shows the false alarm rate (FAR) and true positive rate (TPR) tradeoffs with respect to various jamming probabilities p_J when $R_{\max} = 63$. With $p_J = 0.05$, we can correctly detect 98.57% of jamming instances at a FAR of only 0.6%. When $p_J \geq 0.2$, it reaches 100% detection rate without incurring any false alarms. When $R_{\max} = 63$, there are more hidden-terminal pairs compared with the case when $R_{\max} = 44$ or less. For example, STA1 and STA8 are hidden terminals when $R_{\max} = 63$, but they are not when $R_{\max} = 44$. We also compare the efficacy of our anomaly detector with the PDRRSS_Detect_Jam algorithm proposed by Xu *et al.* in [1]. Fig. 4 shows that the PCA-based anomaly detector significantly outperforms their method in terms of detection accuracy, especially when the jamming activity is not aggressive.

[1]The RJ might alternatively focus on jamming certain control messages (e.g., ARPs, beacons, etc.), which can significantly degrade the channel throughputs and disrupt data transmissions [5].

We also propose three metrics to quantify the effectiveness of a RJ: i) $\eta_J^{(1)}(t_s, t_f)$ is the fraction of time over $[t_s, t_f]$ that the RJ transmitted while a legitimate station transmitted; ii) $\eta_J^{(2)}(t_s, t_f)$ is time the RJ and a legitimate station transmitted over the time that a legitimate station transmitted; iii) $\eta_J^{(3)}(t_s, t_f)$ is time the RJ and a legitimate station transmitted over the time that the RJ transmitted. Notice that the value of $\eta_J^{(1)}$ first increases then decreases as T grows in Fig. 5. The reason is that increasing T by a reasonable amount increases the chance of collision, however the transmission of reactive jammer occupies the channel when T becomes excessively large and prevent legitimates STAs from transmitting.

Parameter	Value
Physical Data rate	24 Mbps
Propagation model	Log-distance path loss model
Transmission power	0.04 Watts (16.0206 dBm)
Energy detection threshold	-83 dBm
Cca mode 1 threshold	-86 dBm
Traffic type	Constant bit rate (CBR)
Rate (per station)	1.5 Mbps
Packet size	1024 Bytes

TABLE I: Simulation parameters

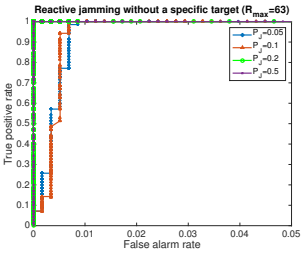

Fig. 3: ROC curves.

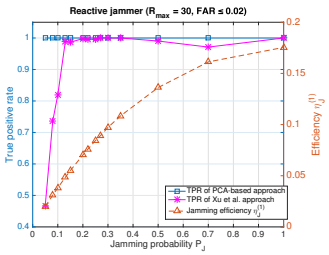

Fig. 4: Comparision with [1].

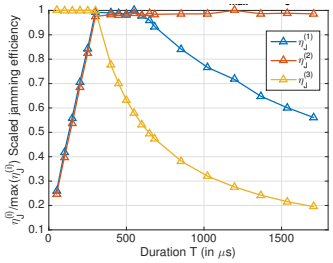

Fig. 5: Scaled $\eta_J^{(i)}$ vs. T ($R_{\max} = 44$,, $p_J = 0.5$, $\tau = 50$).

References

[1] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," in *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Urbana-Champaign, IL, April 2005, pp. 46–57.

[2] A. Lakhina, M. Crovella, and C. Diot, "Mining anomalies using traffic feature distributions," *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 4, pp. 217–228, 2005.

[3] O. Puñal, I. Aktaş, C.-J. Schnelke, G. Abidin, K. Wehrle, and J. Gross, "Machine learning-based jamming detection for IEEE 802.11: Design and experimental evaluation," in *Proc. of the IEEE International Symp. World of Wireless, Mobile, and Multimedia Networks*. IEEE, 2014, pp. 1–10.

[4] ns 3 Consortium. Wireless jamming model. [Online]. Available: https://www.nsnam.org/wiki/Wireless_jamming_model

[5] K. Pelechrinis, M. Iliofotou, and S. V. Krishnamurthy, "Denial of service attacks in wireless networks: The case of jammers," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 2, pp. 245–257, 2011.