This paper was presented as part of the main technical program at IEEE INFOCOM 2011

# From Jammer to Gambler: Modeling and Detection of Jamming Attacks against Time-Critical Traffic

Zhuo Lu Wenye Wang
Department of Electrical and Computer Engineering
North Carolina State University, Raleigh NC 27606
Emails: {zlu3, wwang}@ncsu.edu

Cliff Wang
Army Research Office
Research Triangle Park, NC 27709
Email: cliff.wang@us.army.mil

***Abstract*—Time-critical wireless applications in emerging network systems, such as e-healthcare and smart grids, have been drawing increasing attention in both industry and academia. The broadcast nature of wireless channels unavoidably exposes such applications to jamming attacks. However, existing methods to characterize and detect jamming attacks cannot be applied directly to time-critical networks, whose communication traffic model differs from conventional models. In this paper, we aim at modeling and detecting jamming attacks against time-critical traffic. We introduce a new metric, message invalidation ratio, to quantify the performance of time-critical applications. A key insight that leads to our modeling is that the behavior of a jammer who attempts to disrupt the delivery of a time-critical message can be exactly mapped to the behavior of a gambler who tends to win a gambling game. We show via the gambling-based modeling and real-time experiments that there in general exists a phase transition phenomenon for a time-critical application under jamming attacks: as the probability that a packet is jammed increases from 0 to 1, the message invalidation ratio first increases slightly (even negligibly), then increases dramatically to 1. Based on analytical and experimental results, we further design and implement the JADE (Jamming Attack Detection based on Estimation) system to achieve efficient and robust jamming detection for time-critical wireless networks.**

## I. Introduction

Emerging time-critical wireless systems, such as wireless e-healthcare [1], [2] and wireless power networks [3]–[6], provide a new paradigm of modern wireless networks, whose primary goal is to achieve efficient and reliable message delivery for monitoring and control purposes, instead of providing data services for clients. Hence, a large amount of communication traffic is time-critical in such networks. For example, data messages in power substations are required to be delivered with specific latency constraints, ranging from 3 milliseconds (ms) to 1 second [7]. Due to their significance to human beings (e.g. e-healthcare [2]) and societies (e.g. power grids [3]–[6]), it is of crucial importance to guarantee network availability for such time-critical wireless networks. However, on the other hand, the shared nature of wireless channels inevitably exposes wireless networks to jamming attacks [8]–[10] that may severely degrade the performance of these time-critical networks. Although great progress has been made towards jamming characterization [8]–[10] and countermeasures [11]–[19] for conventional networks, little attention has been focused on time-critical wireless networks.

Indeed, time-critical networks pose challenging issues to existing research on jamming attacks. In conventional networks, the jamming impact is evaluated at packet level (e.g., packet send/delivery ratio [8], the number of jammed packets [11]) or network level (e.g., saturated network throughput [10]). However, packet-level or network-level metrics do not directly reflect the latency constraints of time-critical applications. Hence, conventional performance metrics cannot be readily adapted to measure the jamming impact on time-critical applications. Further, lack of the knowledge how jamming attacks affect time-critical traffic leads to a gray area in the design of jamming detection in time-critical networks: it becomes impractical to achieve efficient jamming detection since detectors are not able to accurately identify jamming attacks, which can cause potentially severe performance degradation of time-critical applications. Therefore, towards time-critical wireless applications, a fundamental question remains unsolved: *How to model, analyze, and detect jamming attacks against time-critical traffic?*

In this paper, we *study the problem of modeling and detecting jamming attacks against time-critical network applications*. Specifically, we consider a time-critical application whose messages must be successfully delivered with delay constraint $\sigma$, and a jammer who attempts to disrupt the message delivery of the time-critical application. There are two key observations that drive our modeling. (i) In such an application, a message becomes invalid as long as the message delay $D$ is greater than the threshold $\sigma$. Thus, we define a performance metric, message invalidation ratio, to quantify the impact of jamming attacks against the time-critical application. (ii) As a retransmission mechanism is adopted in the time-critical application, to successfully disrupt the delivery of a time-critical message, the jammer has to jam each physical transmission attempt of this message until the delay $D$ is greater than $\sigma$. As a result, such behavior of the jammer is exactly the same as the behavior of a gambler who intends to win each play in a game to collect enough fortune to achieve his gambling goal of $\sigma$ dollars.

Motivated by the two observations, we develop a gambling-based model to derive the message invalidation ratio of the time-critical application under jamming attacks. We set up real-time experiments to validate our analysis and further

The work is supported by Army Research Office (ARO) 53435-CS-SR and Secure Open Systems Initiative (SOSI).

evaluate the impact of jamming attacks on an experimental power substation network. Based on our theoretical and experimental results, we design and implement the JADE system (Jamming Attack Detection based on Estimation) to achieve efficient and reliable jamming detection for power networks. Our contributions in this paper are three-fold.

First, we introduce a metric, message invalidation ratio, to quantify the performance of time-critical applications. We show via both analytical and experimental results that the message invalidation ratio characterizes latency constraints of time-critical applications, and thus it is more appropriate than conventional performance metrics for time-critical applications.

Second, we develop a theoretical framework via a gambling game mapping to analyze the impact of jamming attacks on time-critical applications. We find that there exists a phase transition phenomenon for time-critical applications: when the jamming probability $p$ (the probability that a physical transmission is jammed) increases, the message invalidation ratio first increases slightly (and is negligible in practice), then increases dramatically to 1. The phenomenon indicates that there exists a critical probability $p^*$ for a time-critical application. If the jamming probability $p < p^*$, the performance degradation due to jamming attacks can be considered negligible in practice.

Third, we implement the JADE system for jamming detection in time-critical applications. Since the phase transition phenomenon implies that a jammer with jamming probability $p < p^*$ can only cause negligible performance degradation, JADE first estimates the jamming probability $\hat{p}$ and then compares $\hat{p}$ with $p^*$ to detect jammers that can cause non-negligible impact. JADE requires no profiling (training) step that is in general necessary in exiting methods [8], [11], [20]. We show via experiments that JADE achieves comparable detection performance with the statistically optimal likelihood ratio (LLR) test. We further show that JADE is more robust than the LLR test in the presence of a sophisticated time-varying jammer.

The rest of this paper is organized as follows. In Section II, we introduce preliminaries and define the metric of message invalidation ratio. In Sections III and IV, we map the jamming problem into a gambling problem, derive the message invalidation ratio and validate our analytical analysis by experiments. In Section V, we design and implement the JADE system. Finally, we conclude in Section VI.

## II. Preliminaries and Problem Statement

In this section, we first introduce the models of time-critical applications and jamming attacks, then define a performance metric, message invalidation ratio for later analysis.

### A. Modeling of Time-Critical Applications

Emerging wireless networks, such as wireless e-healthcare [1], [2] and wireless power systems [3]–[6], have been drawing increasing attention in both industry and academia. Compared with conventional networks, a large amount of messages in such networks have stringent timing requirements. For example, IEC 61850 [7] is a recent communication standard for power substation automation. IEC 61850 defines a variety of message types with specific timing constraints, in which the most time-critical message type, Generic Object Oriented Substation Event (GOOSE), has two end-to-end delay constraints[1]: 3ms and 10ms.

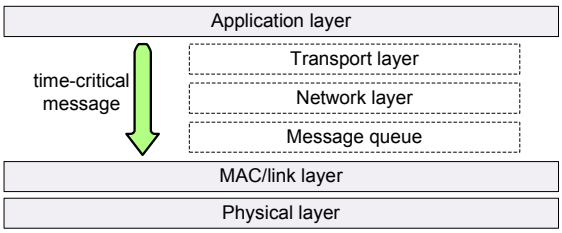


Fig. 1. The application layer sends a time-critical message directly to the MAC/link layer.

The nature of time-critical messages in such networks leads to several basic requirements for transmission protocol design: (i) time-critical messages must be processed with the highest priority; (ii) simple protocol processing and low communication overhead are required; (iii) packet queuing or buffering should be avoided.

For example, IEC 61850 maps time-critical GOOSE messages from the application layer directly to the MAC/link layer to reduce processing time and avoid tedious protocol headers. In this regard, since there is no transport layer to guarantee reliability, IEC 61850 defines that the application layer retransmits the same GOOSE message multiple times to ensure reliability.

Therefore, we assume in this paper that a time-critical message with end-to-end delay constraint $\sigma$ is passed from the application layer directly to the MAC/link layer, as shown in Fig. 1. There is no queuing, flow and congestion control for the transmission. The application layer has a simple processing function that retransmits the same message after the previous transmission fails. But the application layer will stop retransmission once the message delay exceeds the constraint $\sigma$, since the message becomes obsolete or invalid.

We also assume that a time-critical network is always unsaturated (i.e., the network bandwidth is greater than the overall traffic load). Otherwise, the timing requirement of a time-critical message may not be guaranteed since the message has to be queued before transmission. We note that the assumption is also valid in power networks. For example, IEC 61850 shows that the normal traffic load for a common power substation network ranges from 1.952Mbps to 7.592Mbps [7], which can be supported efficiently by either Ethernet (with 100Mbps) or IEEE 802.11g (with 54Mbps).

### B. Modeling of Jamming Attacks

The broadcast nature of wireless channels inevitably exposes time-critical wireless networks to jamming attacks that

[1]The end-to-end delay of a message is defined as the time interval from the instant that the transmitter's application layer generates the message to the instant that the receiver's application layer successfully receives it.

may severely degrade the network performance [8]–[10]. The jamming problem in conventional wireless network has been extensively studied regarding jamming strategies [8]–[10], jamming detection [11], [12], [20], and anti-jamming technologies [13]–[18]. According to [8], jamming attacks can be summarized into two major types: non-reactive and reactive jammers. Non-reactive jammers, including periodical, deceptive and random jammers [8], are not aware of any behavior of legitimate nodes and transmit the radio interference over the wireless channel following their own strategies. Reactive jammers [8], [13], [17], [18] are aware of the target communication systems. They stay quiet when the channel is idle, but start transmitting radio signals (or even meaningful signals [17]) to undermine ongoing communication as soon as they sense activity on the wireless channel. It has been shown in the literature (e.g. [8], [10]) that a reactive jammer is more efficient than a non-reactive jammer. Thus, we focus mainly on a reactive jammer and formally model the jamming strategy as follows.

*Definition 1:* A jamming strategy is represented by $\mathcal{J}(p)$, where $p \in (0, 1)$ is the jamming probability, defined as the probability that a physical transmission can be successfully jammed.

The jamming probability $p$ is sufficient to characterize the powerful level of a reactive jammer because of the following reasons. (i) When the jammer senses an ongoing packet transmission, he can jam the packet with a controllable probability $p$. (ii) If a transmission is protected by anti-jamming schemes, such as frequency hopping spectrum spread (FHSS) [13], [15] and direct sequence spectrum spread (DSSS) [14], [17], the jammer needs to guess (or deduce) the FHSS pattern or DSSS sequence in order to successfully jam the transmission. Therefore, a jammer can only disrupt the transmission with a certain probability, dependent on the jammer's computational ability [17].

### C. Definition of Message Invalidation Ratio

We have modeled the behavior of time-critical applications and the strategy of jamming attacks. We then define a performance metric to model the impact of jamming attacks on time-critical traffic.

In conventional networks, legitimate nodes usually request data services from service providers or exchange data among their neighbors. Hence, the throughput is one of the most important performance metrics in such networks. However, as stated earlier, the primary goal of time-critical wireless applications [1]–[6] is to achieve efficient message delivery for reliable monitoring and control of a variety of infrastructures and devices, instead of providing high throughput for clients. Therefore, the delay of a time-critical message is crucial to such applications. A time-critical message becomes invalid as long as its message delay $D$ is greater than the delay constraint $\sigma$. Therefore, we define a performance metric, message invalidation ratio, to evaluate the performance of time-critical applications.

*Definition 2:* For a time-critical message with end-to-end delay constraint $\sigma$, the message invalidation ratio is defined as

$$r = \mathbb{P}\{D > \sigma\}, \tag{1}$$

where $D$ is the end-to-end delay of the message.

With the definition of message invalidation ratio, we formally state our problem of quantifying the impact of jamming attacks against time-critical traffic as follows.

*Problem Statement:* In a time-critical wireless network under jamming attacks with strategy $\mathcal{J}(p)$, given a time-critical message with end-to-end delay constraint $\sigma$, find out the message invalidation ratio $r$.

In following sections, we will use analytical modeling to derive the message invalidation ratio and perform real-time experiments in a power substation network to validate our analytical analysis.

## III. Impact of Jamming Attacks against Time-Critical Traffic

In this section, we first formulate our jamming problem into a gambling problem, and then derive the message invalidation ratio of time-critical applications under jamming attacks.

### A. Gambling Game for A Jammer

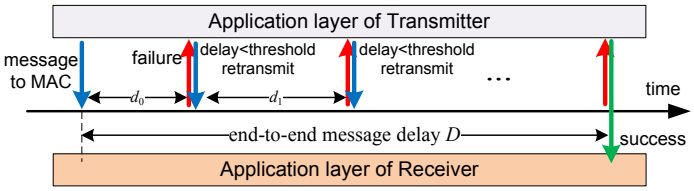


Fig. 2. The transmission process of a time-critical message at the application layer.

Consider a transmitter that has a time-critical message to send with delay constraint $\sigma$, and a jammer with strategy $\mathcal{J}(p)$ that attempts to disrupt message delivery in the network. The process for the transmitter to send the time-critical message is illustrated in Fig. 2: The time-critical message is initially generated at the application layer and is passed directly to the MAC layer to transmit. However, the transmission by the MAC layer may not succeed in the presence of the jammer. If transmission failure (e.g., ACK timeout) is reported by the MAC layer, the application layer will retransmit the same message as long as the cumulative message delay does not exceed the threshold $\sigma$. Therefore, the end-to-end message delay can be represented as

$$D = \sum_{i=0}^{N} d_i, \tag{2}$$

where $N$ is the number of retransmissions and $d_i$ is the MAC-layer delay during the $i$-th retransmission.

Note that the number of retransmissions $N$ and the MAC-layer delay $d_i$ are both random variables. If a message has no delay constraint, the application layer will keep transmitting the same message until it succeeds. In this case, the number of retransmissions $N$ follows the geometric distribution. Then,

the end-to-end delay $D$ in (2) becomes a geometric sum and it is not difficult to use asymptotic analysis to derive the distribution of $D$, similarly to existing work on computing the delay distribution for CSMA/CA networks (e.g., [10], [21]).

However, in our case with a specific delay threshold $\sigma$, jamming attacks can only lead to a finite number of retransmissions at the application layer. The number of retransmissions $N$ is in fact a bounded random variable dynamically coupled with the sum of MAC-layer delays $\{d_i\}$, since every time the application layer compares the accumulated message delay with the constraint $\sigma$ to check whether it should resend a transmission-failed message or drop it. Consequently, it is non-trivial to accurately model and derive the message invalidation ratio of the time-critical application under jamming attacks.

We then take a closer look at the transmission process for a time-critical message. There are two key observations.

1) Such a process has only two outcomes: the jammer either wins or loses. That is, either the jammer keeps successfully jamming every transmission until the delay is larger than the threshold, or the transmitter successfully delivers the message within the timing constraint.
2) In order to win, the jammer must cumulatively collect the reward, i.e., message delay. Every time he jams a physical transmission, a certain amount of delay contributes to the overall message delay.

Is there any process satisfying the two properties? Yes, it is *gambling*. In other words, if we consider the jammer as a gambler and the delay as money, we can exactly map our problem into a gambling game: a gambler attempts to win a game by consistently winning money to reach his goal. The probabilistic modeling of a gambling game, such as the *gambler's ruin* problem [22], has been well investigated by mathematicians. It has been shown that martingale theory [22], a branch of modern probabilistic measure theory, is an effective tool to solve the *gambler's ruin* problem. Therefore, we are motivated to map our problem into a gambling game and solve it by using martingale theory.

We first construct a game for a gambler shown in Fig 3. The gambler starts with $X_0 = d_0$ dollars. In the $n$-th play, when event $A$ happens (with probability $p_a$), the gambler wins $d_n$ dollars; when event $A^c$ happens (with probability 1-$p_a$), he loses $\frac{p_a}{1-p_a}\mathbb{E}(d_n)$ dollars.[2] His gambling goal is $\sigma$ dollars. The gambler quits when he either reaches his gambling goal or loses once (i.e., $A^c$ happens).

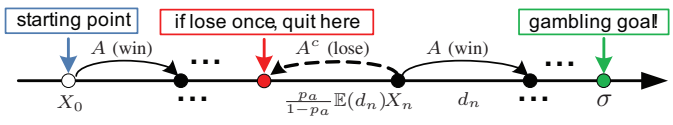


Fig. 3. Setups of our gambling game: the gambler either wins $d_n$ dollars (event $A$) or loses $\frac{p_a}{1-p_a}\mathbb{E}(d_n)$ dollars (event $A^c$) in the $n$-th play. The gambler quits when he either reaches his gambling goal or loses once.

[2]The value of $\frac{p_a}{1-p_a}\mathbb{E}(d_n)$ does not affect the interpretation of our gambling game mapping. It will be shown later that this value is essential to our martingale construction.

Let $\{X_n\}$ be the gambler's money in the $n$-th play. Specifically, we can write $X_n$ as follows.

$$X_0 = d_0,\ X_n = X_{n-1} + \xi_n,\quad (n \in \mathbb{N}), \tag{3}$$

where $\mathbb{N}$ is the set of positive integers, $\xi_n$ is the reward for the gambler in the $n$-th play. Since the gambler can either win or lose in the $n$-th play, the reward $\xi_n$ can be written as

$$\xi_n = d_n \mathbf{1}_A - \frac{p_a}{1-p_a}\mathbb{E}(d_n)\mathbf{1}_{A^c}, \tag{4}$$

where $\mathbf{1}_A$ is the indicator function, defined as

$$\mathbf{1}_A = \begin{cases} 1 & \text{event } A \text{ happens,} \\ 0 & \text{event } A^c \text{ happens.} \end{cases} \tag{5}$$

Then, we map our scenario of the time-critical transmission into the gambling game: the jammer is the gambler and the delay is money. Each transmission can be regarded as a play. Let event $A = \{$the gambler wins money in a play$\} = \{$transmission failure at the MAC layer$\}$. The goal of the jammer/gambler is to make the delay/money larger than the threshold $\sigma$. To achieve this goal, the jammer/gambler must keep jamming/winning successfully in each transmission/play (i.e., event $A$ always happens). However, once $A^c$ happens, the gambler/jammer loses/fails (i.e., the message is successfully delivered within the delay constraint $\sigma$). The message invalidation ratio, which denotes the probability that the cumulative delay is larger than the threshold, is equivalent to the probability that the gambler reaches his goal before he loses.

Note that $p_a$ denotes the transmission failure probability at the MAC layer. Since wireless MAC usually has its own retransmission mechanism due to CSMA/CA (e.g., the default long and short retry limits in IEEE 802.11g are 3 and 7, respectively), event $A$ happens only when every MAC-layer transmission attempt is disrupted by the jammer. Thus, given the number of MAC layer transmission attempts $N_{\text{mac}}$, we obtain $p_a = p^{N_{\text{mac}}}$. Since it has been shown (e.g., [23]) that the collision probability due to legitimate traffic is small if the network is unsaturated, we neglect the impact of legitimate traffic on the MAC-layer transmission failure in our analysis. (We will consider the impact in experiments later).

### B. Main Results

We have set up the rules for our gambling game. We then use the gambling-based model to derive the message invalidation ratio of time-critical applications under jamming attacks. Before we proceed, we first present the definition of a martingale according to [22].

*Definition 3 (Martingale):* A process $\{X_n\}$ is called a martingale relative to a filtration $\{\mathcal{F}_n\}$, (A sequence of $\sigma$-algebras $\{\mathcal{F}_n\}$ is called a filtration if $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for any $n \in \mathbb{N}$.) if (i) $X_n$ is $\mathcal{F}_n$-measurable, (ii) $\mathbb{E}|X_n| < \infty$ for any $n \in \mathbb{N}$, (iii) $\mathbb{E}(X_n|\mathcal{F}_{n-1}) = X_{n-1}$ almost surely.

We then show that the gambler's money $\{X_n\}$ is in fact a martingale due to our construction.

*Lemma 1:* The process $\{X_n\}$ defined in (3) is a martingale.

*Proof:* We prove $\{X_n\}$ is a martingale by verifying the definition.

(i) It is obvious from our construction that $\{X_n\}$ is relative to a filtration $\{\mathcal{F}_n\}$ and $X_n$ is $\mathcal{F}_n$-measurable.

(ii) For any $n \in N$, we have $\mathbb{E}|X_n| = \mathbb{E}|X_0 + \sum_{i=1}^{n} \xi_i| \leq \mathbb{E}|X_0| + n\mathbb{E}|\xi_i|$. Then, it suffices to show $\mathbb{E}|\xi_i| < \infty$. Observe that

$$\begin{aligned}\mathbb{E}|\xi_i| &= \mathbb{E}|d_i \mathbf{1}_A - \frac{p_a}{1-p_a}\mathbb{E}(d_i)\mathbf{1}_{A^c}| \\ &\leq \mathbb{E}|d_i| + \frac{p_a}{1-p_a}\mathbb{E}|d_i| < \infty. \end{aligned} \tag{6}$$

for $0 < p_a < 1$. We obtain $\mathbb{E}|X_n| < \infty$ for $0 < p_a < 1$.

(iii) Then, we prove $\mathbb{E}(X_n|\mathcal{F}_{n-1}) = X_{n-1}$. First, for any $i$, it holds that

$$\begin{aligned}\mathbb{E}(\xi_i) &= \mathbb{E}(d_i \mathbf{1}_A - \frac{p_a}{1-p_a}\mathbb{E}(d_i)\mathbf{1}_{A^c}) \\ &= p_a\mathbb{E}(d_i) - \frac{p_a}{1-p_a}(1-p_a)\mathbb{E}(d_i) = 0. \end{aligned} \tag{7}$$

Then, we have

$$\begin{aligned}\mathbb{E}(X_n|\mathcal{F}_{n-1}) &= \mathbb{E}(X_{n-1} + \xi_n|\mathcal{F}_{n-1}) = X_{n-1} + \mathbb{E}(\xi_n|\mathcal{F}_{n-1}) \\ &= X_{n-1} + \mathbb{E}(\xi_n) = X_{n-1}. \end{aligned} \tag{8}$$

From (i), (ii), and (iii), we obtain $\{X_n\}$ is a martingale. □

We then present our main result of the message invalidation ratio for time-critical traffic under jamming attacks.

*Theorem 1 (Message invalidation ratio for general cases):* Given a jamming strategy $\mathcal{J}(p)$, the message invalidation ratio $r$ is

$$r = \frac{\mathbb{E}(D_s) - \frac{c}{1-p_a}}{\mathbb{E}(D_s) - \frac{p_a c}{1-p_a} - \mathbb{E}(D_u)}, \tag{9}$$

where $p_a = p^{N_{\text{mac}}}$, $c = \mathbb{E}(d_i)$ is the mean of the i.i.d. MAC-layer delay $d_i$, $D_s \leq \sigma$ is the end-to-end delay of a successfully delivered message, and $D_u > \sigma$ is the delay of failed message delivery, defined as the interval from the instant that the transmitter starts transmitting a message to the instant that the transmitter stops retransmission due to message invalidation[3].

*Proof:* Let $n_1 = \inf_{n\in\mathbb{N}}\{X_n < X_{n-1}\}$. According to our construction, event $\{X_n < X_{n-1}\}$ happens if and only if $\xi_n < 0$ (i.e., event $A^c$ happens at the $n$-th play). Therefore, $n_1$ is the minimum time at which the gambler loses money (or, a transmission succeeds).

Let $n_2 = \inf_{n\in\mathbb{N}}\{X_n > \sigma\}$ . Then, $n_2$ is the minimum time at which the gambler reaches his goal (or, the message delay is larger than the threshold).

Thus, $\{n_1 > n_2\}$ means that event $\{X_n < X_{n-1}\}$ never happens prior to event $\{X_n > \sigma\}$, or the gambler reaches his gambling goal without any loss in each play. In other words, event $\{n_1 > n_2\}$ means that the jammer successfully delays the transmission of a message and leads to invalidation of the message.

Therefore, the message invalidation ratio $r = \mathbb{P}(n_1 > n_2)$.

Let $n_{\text{stop}} = \min(n_1, n_2)$. Then, $n_{\text{stop}}$ is a bounded stopping time and

$$X_{n_{\text{stop}}} = X_{n_1}\mathbf{1}_{\{n_1<n_2\}} + X_{n_2}\mathbf{1}_{\{n_1>n_2\}}, \tag{10}$$

where $X_{n_1}$ denotes the remaining money after the gambler loses money for the first time. Then, $X_{n_1-1}$ denotes the money before the gambler loses his money, which is exactly the end-to-end delay of successful message delivery $D_s$. Thus,

$$X_{n_1} = X_{n_1-1} - \frac{p_a\mathbb{E}(d_{n_1})}{1-p_a} = D_s - \frac{p_a c}{1-p_a}. \tag{11}$$

On the other hand, $X_{n_2}$ denotes the money after the gambler achieves his gambling goal of $\sigma$ dollars and quits. Thus, $X_{n_2}$ is exactly the delay of failed message delivery, i.e.,

$$X_{n_2} = D_u. \tag{12}$$

Since $\{X_n\}$ is a martingale (from Lemma 1) and $n_{\text{stop}}$ is a bounded stopping time, we obtain from Doob's optional sampling theorem [22, Ch.10] that the mean value of a martingale $\{X_n\}$ at a stopping time $n_{\text{stop}}$ is equal to the mean value at the starting point 0; i.e.,

$$\mathbb{E}(X_{n_{\text{stop}}}) = \mathbb{E}(X_0). \tag{13}$$

Then, it follows from (10) and (13) that

$$\begin{aligned}\mathbb{E}(X_{n_{\text{stop}}}) &= \mathbb{E}(X_{n_1})\mathbb{P}(n_1 < n_2) + \mathbb{E}(X_{n_2})\mathbb{P}(n_1 > n_2) \\ &= (1-r)(\mathbb{E}(D_s) - \frac{p_a c}{1-p_a}) + r\mathbb{E}(D_u) \\ &= \mathbb{E}(X_0) = \mathbb{E}(d_0). \end{aligned} \tag{14}$$

Therefore, we obtain from (14) that

$$r = \frac{\mathbb{E}(D_s) - \frac{c}{1-p_a}}{\mathbb{E}(D_s) - \frac{p_a c}{1-p_a} - \mathbb{E}(D_u)} \tag{15}$$

□

Theorem 1 shows that the message invalidation ratio can be analytically represented only by first-order statistics. The result in Theorem 1 is general since it does not make further assumptions on the distribution of the MAC-layer delay. To illustrate intuitive relations between message invalidation ratio $r$, jamming probability $p$, and delay threshold $\sigma$, we present our complementary analytical result as follows.

*Theorem 2 (General upper bound):* For the message invalidation ratio $r$ in Theorem 1, it satisfies that

$$r \leq \frac{p^{N_{\text{mac}}}c}{(1-p^{N_{\text{mac}}})(\sigma - c) + p^{N_{\text{mac}}}c}.$$

*Proof:* From Theorem 1, we have

$$\begin{aligned} r &= \frac{\mathbb{E}(D_s) - \frac{c}{1-p_a}}{\mathbb{E}(D_s) - \frac{p_a c}{1-p_a} - \mathbb{E}(D_u)} = 1 - \frac{\mathbb{E}(D_u) - c}{\mathbb{E}(D_u) + \frac{p_a c}{1-p_a} - \mathbb{E}(D_s)} \\ &\leq 1 - \frac{\mathbb{E}(D_u) - c}{\mathbb{E}(D_u) + \frac{p_a c}{1-p_a} - c} = \frac{\frac{p_a c}{1-p_a}}{\mathbb{E}(D_u) + \frac{p_a c}{1-p_a} - c}. \end{aligned} \tag{16}$$

Since the delay of failed message delivery $D_u$ is always larger than $\sigma$ ($D_u \geq \sigma$), it follows from (16) that

$$r \leq \frac{\frac{p_a c}{1-p_a}}{\sigma + \frac{p_a c}{1-p_a} - c} = \frac{p_a c}{(1-p_a)(\sigma - c) + p_a c}. \tag{17}$$

Since $p_a = p_{\text{mac}}^N$, we finally obtain from (17) that

$$r \leq \frac{p^{N_{\text{mac}}}c}{(1-p^{N_{\text{mac}}})(\sigma - c) + p^{N_{\text{mac}}}c}. \tag{18}$$

□

*Remark 1:* Theorem 2 provides a general upper bound of

[3]Note that the reason for $D_u > \sigma$ is that the MAC layer still needs to finish an ongoing transmission even though the application layer is aware that the cumulative delay exceeds the constant $\sigma$.

message invalidation ratio for time-critical applications. Note that when the jamming probability $p$ is sufficiently small, $(1-p^{N_{\text{mac}}})(\sigma-c)\approx\sigma-c\gg p^{N_{\text{mac}}}c$. We obtain that the upper bound of $r$ in (18) can be approximated as $p^{N_{\text{mac}}}c/(\sigma-c)$, indicating that the message invalidation ratio decays at least polynomially when $p$ is small and decreasing to 0. Consequently, a small jamming probability $p$ cannot lead to significant impact on the performance of time-critical applications.

Fig. 4 numerically illustrates the upper bound for a time-critical application with 10ms$<\sigma<$100ms, $0<p<1$, $N_{\text{mac}}$=3, and $c=\mathbb{E}(d_i)$=1ms. We observe from Fig. 4 that the message invalidation ratio, as a function of jamming probability $p$, has a phase transition phenomenon. That is, as $p$ increases, the message invalidation ratio has two distinct increasing phases: a slightly-increasing phase and a dramatically-increasing phase. For example, when $\sigma$=10ms, the transition point is approximately at $p$=0.7 and the corresponding upper bound of message invalidation ratio is $r$=5%. In other words, the upper bound only increases from 0% slightly to 5% as $p$ goes from 0 to 0.7 and increases from 5% dramatically to 100% as $p$ goes from 0.7 to 1.

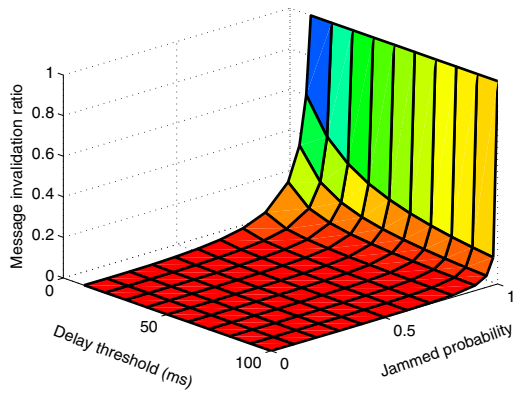


Fig. 4. Upper bound of message invalidation ratio for a time-critical application with 10ms$<\sigma<$100ms, $0<p<1$, $N_{\text{mac}}$=3, and $\mathbb{E}(d_i)$= 1ms.

## IV. Experimental Study in Power Networks

We have developed a gambling-based model to analytically derive the message invalidation ratio of a time-critical application under jamming attacks. As aforementioned, there are a variety of time-critical network applications in power systems. Recently, towards the smart grid vision, wireless technologies for power systems have attracted increasing attention in government [3], industry [6], and academia [4], [5]. Thus, in this section, we set up a WiFi-based power network to validate our analytical results and further evaluate the impact of jamming attacks on our experimental power system.

### A. Experimental Setups

*1) GOOSE Applications:* As stated earlier, IEC 61850 [7] is a digital communication protocol for modern power substation networks. The GOOSE message in IEC 61850 is a time-critical message with strict timing requirements. In our experiments, we use different GOOSE applications to evaluate the impact of jamming attacks on a power network. Specifically, we consider two protocol-defined GOOSE applications: types 1A/P1 and 1A/P2 with constraints of 3ms and 10ms [7], respectively. We also consider two GOOSE applications for transfer trip protection and anti-islanding with delay constraints of 8-16ms and 150-300ms [4], respectively.

The GOOSE application layer features an enhanced retransmission mechanism [7], in which the same message is retransmitted with increasing retransmission intervals. As shown in Fig. 5, the first retransmission interval is $T_1$, the second one is $T_2 \geq T_1$, and the interval keeps increasing up to $T_{\text{max}}$. How $T_1$ increases to $T_{\text{max}}$ is claimed to be a local issue and is not standardized in IEC 61850. Therefore, the interval of each retransmission is increased equally by $\delta$ in our implementation.


Fig. 5. The enhanced retransmission mechanism in GOOSE.

*2) Implementation:* We set up a WiFi-based wireless power network to evaluate the GOOSE performance under jamming attacks. Since GOOSE is mapped from the application layer directly to the MAC layer, we implement a GOOSE messaging module in the Linux kernel. Detailed setups are as follows.

1) Operating system: Linux (kernel version 2.6.32).
2) GOOSE parameters: we set $T_1$=1ms, $T_{\text{max}}$=5ms, and $\delta$=1ms. For the most time-critical (3ms) case, we set $T_1$=$T_{\text{max}}$=1ms. During the experiments, the application layer is set to stop retransmission once the message delay exceeds the threshold.
3) MAC layer: IEEE 802.11g (basic service set) at 2.462 GHz. As GOOSE requires the highest priority for processing, we use Madwifi driver [24] to set minimum and maximum 802.11 contention windows to be 4 and 8, respectively. We also set the retry limit to be 3.
4) Jammer: We use the USRP system with GNU radio (version 3.3) to set up a low-power jammer to disrupt the GOOSE messaging. The length of jamming signals is set to be 22 microseconds as given in [10].

*3) Performance Metric:* We use the message invalidation ratio to measure the jamming impact. We transmit 1000 GOOSE messages for every GOOSE application in each experiment, We then measure the delay of each GOOSE message, compare the delay with the threshold and compute the message invalidation ratio.

### B. A Two-Node-and-One-Jammer Scenario

Our first experiment is to evaluate a simple communication scenario that commonly exists in power systems: an electronic device observes an event (e.g., an abnormal status) and transmits a GOOSE message to inform the other of this event. The goal of this experiment is to show how a jammer can affect time-critical GOOSE transmissions between a single transmitter-receiver pair.

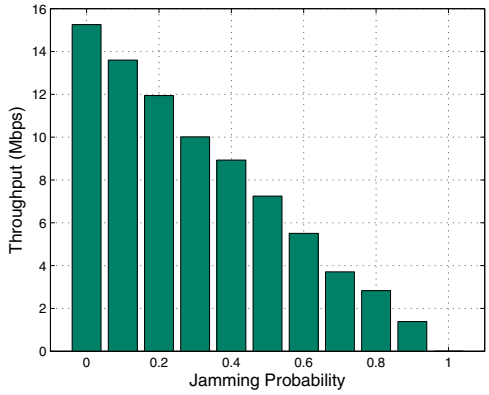


Fig. 6. The 802.11g saturated throughput of a single transmitter-receiver pair under jammer attacks with jamming probability $p \in [0, 1]$.

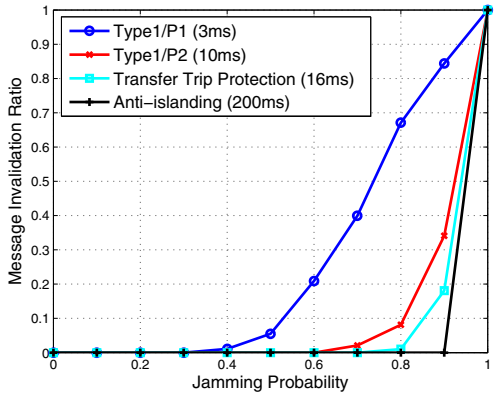


Fig. 7. The message invalidation ratios of four different GOOSE applications.

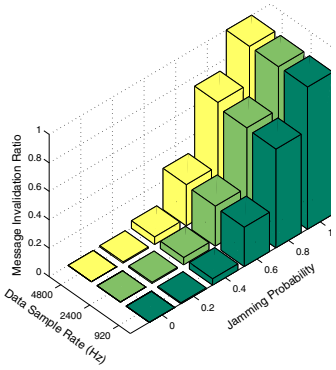


Fig. 8. Message invalidation ratio (Type-1A/P1 GOOSE with 3ms limit) as a function of jamming probability $p$ and transmission rate of the MU IED.

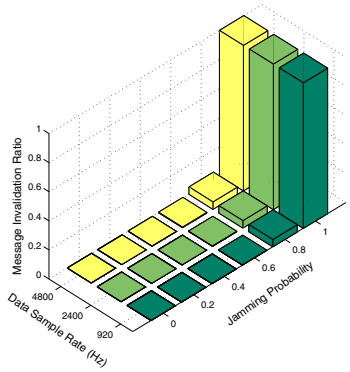


Fig. 9. Message invalidation ratio (Type-1A/P2 GOOSE with 10ms limit) as a function of jamming probability $p$ and transmission rate of the MU IED.

We first show in Fig. 6 the conventional 802.11g saturated throughput of the transmitter under jamming attacks. We can see that the throughput decreases approximately linearly as the jamming probability $p$ increases. Thus, the jammer must choose a large jamming probability $p$ to significantly degrade the throughput performance in a WiFi network.

We then show in Fig. 7 the message invalidation ratios for different GOOSE applications with delay limits of 3ms, 10ms, 16ms, and 200ms, respectively. It can be seen from Fig. 7 that every GOOSE application exhibits a phase transition phenomenon: when the jamming probability $p$ is small, the message invalidation ratio is 0; and as $p$ increases, the message invalidation ratio becomes non-zero and increases dramatically to 1. For example, Fig. 7 illustrates that when $p$ goes from 0 to 0.6, the Type-1A/P2 (10ms limit) message invalidation ratio always remains zero, which implies that a small jamming probability $p$ cannot lead to significant performance degradation. Fig. 7 also shows that some GOOSE applications are not extremely vulnerable to jamming attacks, especially for less delay-sensitive ones. For example, for the anti-islanding application, the message invalidation ratio is 0.1% at $p = 0.9$.

Note that Figs. 6 and 7 indicate that severe throughput degradation under jamming attacks does not necessarily lead to a large message invalidation ratio. For example, when $p = 0.9$, the throughput is degraded by 88% in Fig. 6, but the message invalidation ratio is 0.1% for the anti-islanding application in Fig. 7. Thus, the message invalidation ratio is an application-oriented performance metric and is more appropriate than the saturated throughput to quantify the performance of time-critical applications.

### C. A Small-Scale Network Scenario

We then consider a WiFi-based power network scenario [25]: a transformer bay in a Type D2-1 power substation has two breaker intelligent electronic devices (IEDs), two protection-and-control (P&C) IEDs, and one merging-unit (MU) IED. All breaker IEDs and P&C IEDs send updated meter values to a station server at a fixed rate of 20Hz. The MU IED sends raw data messages to P&C IEDs at a rate of 920Hz, 2400Hz, or 4800Hz. (All setups are from [25].) Our goal is to not only investigate the impact of jamming attacks but also evaluate the effect of legitimate traffic on GOOSE messaging in a small-scale power network over WiFi access.

Figs. 8 and 9 show the message invalidation ratios of Type-1A/P1 (3ms limit) and Type-1A/P2 (10ms limit) GOOSE messages transmitted from a breaker IED to a P&C IED, respectively. Note that the WiFi-based network is always unsaturated even when the transmission rate of the MU IED is 4800Hz. We can see from Figs. 8 and 9 that unsaturated traffic load has nearly negligible effect on the message invalidation ratio. For example, when the jamming probability $p$ is fixed to be 0.8, the message invalidation ratio of Type-1A/P2 (10ms limit) GOOSE messages increases from 4.9% to 5.2% as the MU IED transmission rate goes from 920Hz to 4800Hz. Thus, we conclude that the increasing of unsaturated traffic load can only slightly degrade the performance of time-critical transmissions. It is also noted from Figs. 8 and 9 that legitimate traffic does not affect the phase transition phenomenon of the message invalidation ratio.

## V. JADE: Jamming Attack Detection based on Estimation

In previous sections, we have modeled the impact of jamming attacks on time-critical applications and validated our

analysis by performing experiments in a power network. Our analytical and experimental results provide a prerequisite to the design of jamming detectors for time-critical applications. In this section, we implement a jamming detection system, JADE (Jamming Attack Detection based on Estimation) for power systems. We show that JADE achieves both efficiency and reliability for jamming detection in power networks.

### A. Design and Implementation

Due to the importance of power networks, a jamming detector should yield a reliable output within a very short decision time to notify network operators of potential threats. Existing methods in general require a profiling step, which estimate parameters [8], [11] or infer statistical models [12], [20] from measured data, to provide empirical knowledge for jamming detection. For example, a sequential jamming detector proposed in [11] needs to estimate the transmission failure probabilities in both non-jamming and jamming cases before performing jamming detection. However, such profiling-based methods face several practical issues for time-critical systems: (i) the profiling phase inevitably increases the detection time; (ii) it is unclear in practice how much reliability the profiling phase can provide for later jamming detection.

As we can see, existing profiling-based detectors may not be directly used in practical power systems. Thus, we are motivated to design a new jamming detection system, JADE, to achieve both efficiency and reliability for jamming detection in power systems. The intuition of JADE is as follows.

First, the profiling-based methods are used in ad-hoc or sensor networks where network parameters for a node (e.g., number of nodes, background traffic) are usually considered unknown. However, nodes in a power network are usually static and have nearly predictable traffic (e.g., the raw data sampling rate and meter update rate of IEDs). Thus, the profiling phase for jamming detection is not necessary in a power network.

Second, as we observe in previous sections, the phase transition phenomenon indicates that when the jamming probability $p$ is sufficiently small, the jamming impact is nearly negligible. This means that in order to detect the presence of a harmful jammer, a detection system only needs to estimate the jamming probability $\hat{p}$, and then to compare the estimation with a critical jamming probability $p^*$, with which a jammer can cause non-negligible impact on power networks. If $\hat{p}$ is small, whether it is induced by channel collision, fading, or even jamming, it cannot lead to significant performance degradation. Otherwise, the detection system should raise an alarm.

Accordingly, we implement the JADE system at a MU IED that periodically transmits raw data samples at the rate of 920Hz [4]. JADE observes the transmission result of each data sample and estimates the jamming probability $\hat{p}$ by

$$\hat{p} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{S_i}, \tag{19}$$

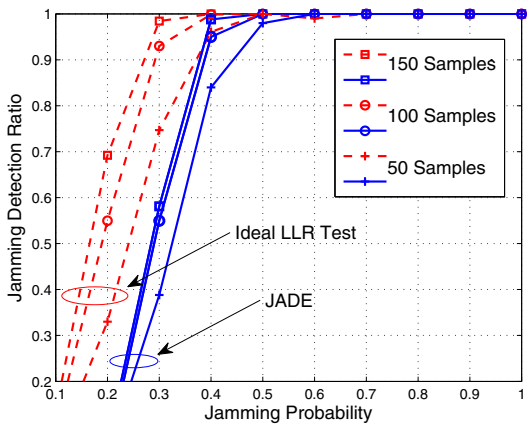


Fig. 10. Jamming detection ratios of both JADE and the likelihood ratio test in the presence of a jammer with different jamming probabilities.

where $N$ is the number of observations, and $S_i$ denotes the event that the $i$-th transmission succeeds.

After the estimation in (19), JADE raises a jamming alarm if $\hat{p} > p^*$. Note that given all setups of a power network, the threshold $p^*$ can be chosen via either analytical analysis or experiments.

### B. Experimental Results

We then use the experimental power network in Section IV-C to assess the performance of JADE. As the lowest bound of GOOSE delay is 3ms, we choose the corresponding critical jamming probability (detection threshold) $p^*$=0.3 from experimental results in Figs. 7 and 8. We also implement the statistically optimal likelihood ratio (LLR) test in our experiments for performance comparison. (A sequential version of the LLR test is used in [11].) The LLR test first requires a profiling step to estimate the packet jammed probability. During our experiments, we assume that the LLR test knows the information perfectly; i.e., we set exactly the same jamming probability in the LLR test as that used by the jammer. Thus, we refer to this detector as the ideal LLR test. Given the raw data transmission rate of 920 Hz, we set $N$=50, 100 and 500 samples such that the corresponding decision time for detection is 54 ms, 109 ms and 163 ms, respectively.

Fig. 10 shows the jamming detection ratios (i.e. the probability that a detector issues an alarm when there indeed exists jamming) of both JADE and the ideal LLR test. We can see that the ideal LLR test outperforms JADE significantly when the jamming probability $p<0.3$. This is because JADE does not target jamming attacks with jamming probability $p<p^*=0.3$. However, the phase transition phenomenon has shown that less aggressive jammers cannot dramatically affect the performance of time-critical traffic. Hence, with jamming probability $p<0.3$, even a jammer evades the detection of the JADE system, he fails to cause noticeable message invalidation ratios. It is further observed from Fig. 10 that when the jamming probability is greater than 0.3, the ideal LLR test and JADE achieve comparable performance especially when the number of samples $N$ is large. For example, when $N$=150 and

TABLE I
DETECTION RATIOS OF BOTH JADE AND LIKELIHOOD RATIO TEST IN THE PRESENCE OF A TIME-VARYING JAMMER.

| Number of Samples: | 50 | 100 | 150 | 200 |
|---|---|---|---|---|
| JADE: | 98.6% | 99.1% | 100% | 100% |
| LLR Test: | 91.3% | 92.1% | 92.5% | 91.6% |

$p$=0.4, the detection ratios of JADE and the ideal LLR test are $98.4\%$ and $99.1\%$, respectively. Thus, JADE is able to detect harmful jamming attacks with nearly optimal performance.

It is well known that the performance of the LLR test could be degraded by model mismatch due to imperfect estimation or insufficient profiling. To compare the robustness of JADE with that of the LLR test, we design a reactive jammer that keeps changing its jamming probability randomly and uniformly within $[0.4, 0.9]$. In this case, the LLR test first estimates the jamming probability and then performs jamming detection based on the estimation output. Table I shows the detection ratios of both JADE and the LLR test for $N$=50, 100, 150, and 200. We can see that JADE is more robust than the LLR test to detect such a time-varying jammer. Because of the model mismatch problem, we observe from Table I that increasing the number of samples cannot improve the performance of the LLR test.

Note that the false alarm ratio, which is the probability that a detector issues an alarm when there is no jamming, is also an important metric to evaluate the performance of jamming detectors. During our experiments, neither JADE nor the LLR test issues a jamming alarm when there exists no jamming, since the wireless network is unsaturated and transmission failure rarely happens.

### C. Discussions

Our experimental results show that JADE achieves efficient and robust jamming detection for aggressive and harmful jammers, at the cost of low detection ratio for less-aggressive jammers. We note that JADE is an application-oriented detector that can be applied directly to practical wireless power systems.

It is also worth noting that JADE is implemented at the application layer for jamming detection. A cross-layer detection mechanism that combines application-layer information and physical-layer information can further improve the detection accuracy, such as the signal strength consistency check [8]. Thus, our future work includes the implementation of a cross-layer detector to further improve the detection accuracy of the JADE system.

## VI. CONCLUSIONS

In this paper, we provided an in-depth study on the impact of jamming attacks against time-critical network applications by theoretical modeling and system experiments. We introduced a performance metric, message invalidation ratio, to quantify the impact of jamming attacks. We showed via both analytical analysis and real-time experiments that there in general exists a phase transition phenomenon in time-critical applications under jamming attacks. Based on our analysis and experiments, we designed and implemented the JADE system to achieve efficient and robust jamming detection for power networks.

## REFERENCES


[1] H. Su and X. Zhang, "Battery-dynamics driven TDMA MAC protocols for wireless body-area monitoring networks in healthcare applications," *IEEE J. Select. Areas in Commun.*, vol. 27, pp. 424–434, May 2009.

[2] M. Li, W. Lou, and K. Ren, "Data security and privacy in wireless body area networks," *IEEE Wireless Commun.*, vol. 17, pp. 51–58, Feb. 2010.

[3] Office of the National Coordinator for Smart Grid Interoperability, "NIST framework and roadmap for smart grid interoperability standards, release 1.0," *NIST Special Publication 1108*, 2009.

[4] P. M. Kanabar, M. G. Kanabar, W. El-Khattam, T. S. Sidhu, and A. Shami, "Evaluation of communication technologies for IEC 61850 based distribution automation system with distributed energy resources," in *Proc. of the IEEE Power & Energy Society General Meeting (PES '09)*, July 2009.

[5] B. Akyol, H. Kirkham, S. Clements, and M. Hadley, "A survey of wireless communications for the electric power system," in *Tech. Report, Pacific Northwest National Laboratory*, Jan. 2010.

[6] Wi-Fi Alliance, "WiFi for the smart grid: Mature, interoperable, security-protected technology for advanced utility management communications," Sept. 2009.

[7] IEC Standard, "IEC 61850: Communication networks and systems in substations," 2003.

[8] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," in *Proc. of ACM MobiHoc '05*, 2005, pp. 46–57.

[9] L. Sang and A. Arora, "Capabilities of low-power wireless jammers," in *Proc. of IEEE INFOCOM '09 mini-conference*, Apr. 2009.

[10] E. Bayraktaroglu, C. King, X. Liu, G. Noubir, R. Rajaraman, and B. Thapa, "On the performance of IEEE 802.11 under jamming," in *Proc. of IEEE INFOCOM '08*, Apr. 2008, pp. 1265–1273.

[11] M. Li, I. Koutsopoulos, and R. Poovendran, "Optimal jamming attacks and network defense policies in wireless sensor networks," in *Proc. of IEEE INFOCOM'07*, May 2007, pp. 1307–1315.

[12] A. L. Toledo and X. Wang, "Robust detection of MAC layer denial-of-service attacks in CSMA/CA wireless networks," *IEEE Trans. Info. Forensics and Security*, vol. 3, pp. 347–358, Sept. 2008.

[13] M. Strasser, S. Capkun, C. Popper, and M. Cagalj, "Jamming-resistant key establishment using uncoordinated frequency hopping," in *Proc. of IEEE Symposium on Security and Privacy*, May 2008, pp. 64–78.

[14] M. Strasser, C. Popper, and S. Capkun, "Efficient uncoordinated FHSS anti-jamming communication," in *Proc. of MobiHoc '09*, 2009.

[15] J. T. Chiang and Y.-C. Hu, "Dynamic jamming mitigation for wireless broadcast networks," in *Proc. of IEEE INFOCOM '08*, Apr. 2008.

[16] V. Navda, A. Bohra, S. Ganguly, and D. Rubenstein, "Using channel hopping to increase 802.11 resilience to jamming attacks," in *Proc. of IEEE INFOCOM '07*, May 2007, pp. 2526–2530.

[17] Y. Liu, P. Ning, H. Dai, and A. Liu, "Randomized differential DSSS: Jamming-resistant wireless broadcast communication," in *Proc. of IEEE INFOCOM '10*, Mar. 2010.

[18] C. Popper, M. Strasser, and S. Capkun, "Jamming-resistant broadcast communication without shared keys," in *Proc. of USENIX Security Symposium (Security '09)*, Aug. 2009.

[19] M. Cagalj, S. Capkun, and J. P. Hubaux, "Wormhole-based antijamming techniques in sensor networks," *IEEE Trans. Mobile Computing*, vol. 6, no. 1, pp. 100–114, Jan. 2007.

[20] A. Hamieh and J. Ben-Othman, "Detection of jamming attacks in wireless ad hoc networks using error distribution," in *Proc. of IEEE ICC '09*, Jun. 2009.

[21] D. Malone, K. Duffy, and D. Leith, "Modeling the 802.11 distributed coordination function in nonsaturated heterogeneous conditions," *IEEE/ACM Trans. Networking*, vol. 15, pp. 159–172, Feb. 2007.

[22] W. David, *Probability with Martingales*. Cambridge University Press, 1991.

[23] I. Aad, J.-P. Hubaux, and E. W. Knightly, "Impact of denial of service attacks on ad hoc networks," *IEEE Trans. Networking*, vol. 16, no. 4, pp. 791–802, Aug. 2008.

[24] Madwifi, http://madwifi.org.

[25] T. S. Sidhu and Y. Yin, "Modelling and simulation for performance evaluation of IEC61850-based substation communication systems," *IEEE Trans. Power Delivery*, vol. 22, no. 3, pp. 1482–1489, July 2007.