Reliable Event-Detection in Wireless Visual Sensor Networks Through Scalar Collaboration and Game-Theoretic Consideration

Alexandra Czarlinska, *Student Member, IEEE*, and Deepa Kundur, *Senior Member, IEEE*

***Abstract*—In this work we consider an *event-driven* wireless visual sensor network (WVSN) comprised of untethered camera nodes and scalar sensors deployed in a hostile environment. In the event-driven paradigm, each camera node transmits a surveillance frame to the cluster-head *only if* an event of interest was captured in the frame, for energy and bandwidth conservation. We thus examine a simple image processing algorithm at the camera nodes based on difference frames and the chi-squared detector. We show that the test statistic of the chi-squared detector is equivalent to that of a robust (non-parametric) detector and that this simple algorithm performs well on indoor surveillance sequences and some, but not all, outdoor sequences. In outdoor sequences containing significant changes in background and lighting, this simple detector may produce a high probability of error and benefits from the inclusion of scalar sensor decisions. The scalar sensor decisions are, however, prone to attack and may exhibit errors that are arbitrarily frequent, pervasive throughout the network and difficult to predict. To achieve attack prediction and mitigation given an attacker whose actions are not known *a priori*, we employ game-theoretic analysis. We show that the scalar sensor error can be controlled through cluster-head checking and appropriate selection of cluster size n. Given this attack mitigation, we employ real-life sequences to determine the total probability of error when individual and combined decisions are utilized and we discuss the ensuing ramifications and performance issues.**

***Index Terms*—Actuation, event-detection, game theory, scalar-sensors, sensor network security, wireless visual sensor networks (WVSNs).**

I. Introduction

THE general popularity of mobile devices and the lure of innovative applications have continued to drive research in the area of wireless multimedia. Indeed much focus has been placed on the design of robust, efficient and secure schemes for delivering multimedia content over error-prone wireless channels [1], [2]. Recent years have also brought developments in another growing field of interest, that of wireless sensor networks (WSNs). Originally envisioned as simple devices for distributed environmental sensing, WSNs have continued to evolve in complexity to include autonomous mobility and actuation of elements in their surroundings [3], [4]. Visual-capability additions to WSNs are thus but a natural extension of this vibrant research and are in alignment with co-evolving mobile device research interests.

Emerging from these ideas, the nascent field of wireless visual sensor networks (WVSNs) considers battery-operated wireless (untethered) nodes equipped with cameras [5], [6]. Among other applications, it boasts importance to rapid-deployment surveillance and monitoring [7]. As all ambitious research ideas, WVSNs have many significant challenges to overcome. The energy limitations already encountered in WSNs collecting scalar data such as temperature are only exacerbated when nodes are deployed to collect, process and transmit visual data [8], [9]. The increased size of visual data also strains storage buffers and places a further burden on system design through increased transmission delay and bandwidth utilization [10]. Such challenges call for innovative solutions tailored to the unique characteristics of WVSNs [5].

Among possible approaches, *event-driven* WVSNs have emerged as an intriguing design possibility. In the event-driven approach, camera nodes buffer and transmit an image frame *only if* such a frame contains an *event of interest* as defined by the application [10], [11]. The event-driven paradigm hence aims to alleviate energy consumption and bandwidth use implicitly, via the local selection of relevant image frames by the nodes. This approach may be viewed as complementary to the joint optimization of video encoding and wireless transmission power [2], [8], or to efforts at providing higher bandwidth channels [7] and allocating their use fairly and efficiently [1]. Overall, WVSN systems will undoubtedly benefit from the incorporation of advances from the spectrum of these complementary strategies.

In this work we focus on event-driven WVSNs which critically rely on correct frame selection given the definition of an *event*. For surveillance nodes where unknown objects may or may not enter the camera's field of view, this definition may be largely motion-based, though in general the definition is application-dependent [12], [13]. Based on the event definition, a camera node should ideally achieve a high probability of event detection P_D to guarantee that important frames are transmitted. The camera nodes should also achieve an acceptably low probability of false alarm P_{FA} to avoid the wasteful transmission of non-event frames. Event detection in WVSNs must thus meet critical accuracy and reliability requirements while minimizing complexity and energy consumption. Though approaches based solely on image processing at the camera nodes are possible, they are generally expensive in terms of

Manuscript received July 31, 2007; revised November 2, 2007. Published July 9, 2008 (projected). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Eckehard Steinbach.

The authors are with the Electrical and Computer Department, Texas A&M University, College Station, TX 77843-3128 USA (e-mail: czlinska@ece.tamu.edu; deepa@ece.tamu.edu).

Digital Object Identifier 10.1109/TMM.2008.922775

1520-9210/$25.00 © 2008 IEEE

the computation required to achieve an acceptable probability of error [8], [9]. In this work we show how collaborative approaches exploiting available scalar data can be used in conjunction with low complexity image processing algorithms to achieve reasonable performance.

Achieving reliability in WSN scalar sensors for event detection is itself a challenging task due to the possibility of attack, miscalibration or sensor failure [14], [15]. Many such concerns regarding reliability may be addressed through component testing, node redundancy and node security mechanisms. The evolution of WSNs to include capabilities such as actuation however, renders the last factor particularly onerous due to the emergence of new attacks [15]. For instance, cryptographic and signal processing based mechanisms largely protect WSNs from false data injection due to node or key capture [16]. However, actuation attacks which perturb sensor readings directly have remained largely unexplored due to their physical nature [17]. Analysis of such attacks is generally complicated by considerations of the type of scalar sensor involved, such as temperature, sound or motion.

In this work, we describe a model which abstracts the details of the scalar sensor and focuses on the *interaction* between a sensor network collecting scalar data and a hostile actuating attacker. The effect of the attack is modeled as a change in a sensor's reported *decision* about the presence or absence of an event [11]. By considering the interaction between the network and the attacker as a competition, we are able to invoke the game-theoretic concept of a Nash equilibrium to assess the sensors' *reliability* [15], [18]. In this respect, our work continues the recent exploration of game-theoretic models for engineering applications such as wireless networks and multimedia communications [1], [19].

This paper thus analyzes the detection performance of event-driven WVSNs that utilize attack-prone scalar sensors in addition to low-complexity image processing. We now summarize the main focus and contributions of this paper.

1) **Image Processing at the camera nodes**: we show the mathematical equivalence between a known simple chi-squared detector and a theoretically robust detector. This is intended to illustrate that computationally simple algorithms *may* have desirable detection properties in the theoretical sense. Through experiments on real surveillance sequences, we show that in practice the chi-squared detector may perform well in indoor conditions and poorly in some, but not all, outdoor conditions. This leads to the question of *when* the low-complexity image processing should be trusted to provide accurate detection and whether this performance can be improved.
2) **Attack-Prone Scalar Sensors**: in contrast with occasional errors caused by malfunction, errors due to attack may be frequent and arbitrarily prevalent throughout the network. They may also be difficult to predict and mitigate since they are perpetrated by an active attacker whose actions are unknown *a prior*i. We show how a simple *type* (or count) detector drives an attacker's optimal attack probability to be *small* in the game-theoretic sense. We show how the number of nodes in a cluster and the probability of an event affect this attack probability and show that cluster-head feedback is not necessary for attack detection.
3) **Collaborative performance**: given attack-prone scalar sensors, it is not clear that combining them with the chi-squared detector improves detection performance. We employ real-life surveillance sequences to determine the total probability of error when individual and combined decisions are used. We determine that utilizing the simple chi-squared detector and combining its decisions with the scalar ones is generally advantageous, with further gains in performance when proper attack-mitigants are utilized.

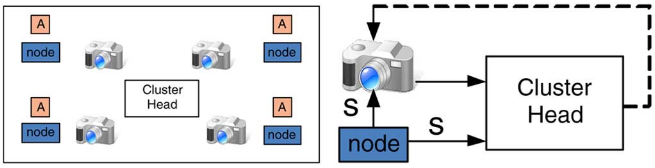

Fig. 1. (a) Each camera node receives support from one scalar sensor labeled "node" (1: 1 ratio). Attacking nodes denoted as A are also deployed by an opponent and may cause scalar sensor errors. (b) A scalar sensor sends its decision s to both the camera node and the cluster-head. Feedback from the cluster-head may or may not exist.

The remainder of this paper is organized as follows. In Section II, we detail the proposed event-driven WVSN, in Section III, we discuss the image processing algorithm at the cameras. In Section IV, we present the scalar-sensor attack model, its analysis & mitigation. In Section V, we present the performance of the collaborative approaches for utilizing both the camera decisions and the scalar decisions. Finally, in Section VI, we present our conclusions and discussion.

II. WVSN System Model

In this work, we consider an *event-driven* WVSN where camera nodes are deployed in an environment of interest for general surveillance purposes as shown in Fig. 1. Each camera node transmits image frames wirelessly to a collection center referred to as the cluster-head. Importantly, image frames are transmitted by a camera node to the cluster-head *only if* an event of interest was detected in those frames. To perform the event detection, each camera node utilizes a simple image processing algorithm (detailed in Section III) and each camera node also receives a detection decision from one supporting scalar-sensor node as shown in Fig. 1.[1] Each scalar sensor node transmits its decision about the presence or absence of an event both to its collaborating camera node and to the cluster-head (such as by appending this decision to packets that it may already be transmitting to these entities).

The scalar-sensors make their decisions based on readings from the environment such as sound, temperature or motion. For instance, a scalar sensor could collect audio input and alert

[1] The 1:1 ratio of scalar nodes to camera nodes can be seen as a worst-case scenario. For a ratio of $k:1$ where $k > 1$, the overall probability of error for the scalar sensors *might* improve, but is still not guaranteed to be small since each scalar sensor can be affected with non-zero probability by an attack.

TABLE I
KEY NOTATION

Notation	Description	Note
s	scalar-sensor decision	$s \in \{0,1\}$, $s=1$ signifies an event
I_m	visual node decision	$I_m \in \{0,1\}$, $I_m=1$ signifies an event
n	number of scalar-data nodes within cluster	$n \in \mathcal{Z}^+$
p	probability (at any given time) of event occurrence	$p \in [0,1]$
q'	probability of visual detection error in I_m	$q' \in [0,1]$, due to algorithm selection
q	probability of scalar sensor error in s	$q \in [0,1]$, due to attack
P_D	probability of detection	true positive
P_{FA}	probability of false alarm	false positive
P_S	probability of attack success	cluster-head misses attack

its corresponding camera node when this input crosses an (application-dependent) threshold due to the presence of a moving object of interest. For generality and tractability, in this work the specific type of sensor is not considered but rather modeled in terms of its error probability. Specifically, we consider the case where a scalar-sensor may suffer from errors due to the presence of a hostile attacking node, shown as A in Fig. 1, [14], [15]. In contrast with occasional malfunctions, errors caused by a hostile attack may occur with arbitrary probability. This error probability must thus be analyzed and controlled to guarantee reliable decision support for the camera nodes as detailed in Section IV.

We employ the following notation in our discussion of event detection as summarized in Table I. A camera node's decision based on image processing techniques is denoted by I_m, where $I_m \in \{0,1\}$ such that $I_m = 1$ denotes the camera node decision "event is present" and $I_m = 0$ denotes "event is absent". We assume that the largest source of error affecting I_m stems from the selection of a simpler or more general image processing algorithm (i.e., one that is not tailored to the detection of specific objects and backgrounds). This assumption stems from the observation that attacks on camera nodes typically require physical tampering or proximity to the node which render the attack detectable (i.e., not stealthy as considered in this work). We employ the notation s to denote a scalar sensor decision based on environmental readings, where $s \in \{0,1\}$ such that $s = 1$ denotes the scalar decision "event is present". We assume that the main source of scalar sensor error is attack rather than harsh environmental conditions, miscalibrations or component defects. Unlike in the case of camera nodes, the decisions of scalar sensors may be attacked (i.e., perturbed or disrupted in a stealthy manner) by a hostile network co-deployed in the environment.

Each camera node thus faces a scenario where for each captured frame it has access to an image processing-based decision I_m and a scalar-sensor decision s.[2] The probability q' that I_m is in error may not be equal to the probability q that s is in error. For changing outdoor conditions and attack strategies, q' and q may be unknown to the camera node at any given time and furthermore, the relationship between them may be unknown (i.e., which error is smallest). In our analysis we thus consider different scenarios for using the I_m and s decisions at a camera node. As shown in Table II, the first approach is to always trust the scalar sensor s (including when a disagreement between I_m and s occurs). This approach may yield an acceptable performance if the scalar sensor is reliable, that is, if the predominant source of scalar sensor error is restricted to *occasional* malfunctions. The detection performance may however be *unacceptable* if the scalar sensors are also exposed to a hostile attack due to their unattended physical deployment in a given environment. In such a scenario, the number of incorrect scalar decisions is not restricted and depends on the varying actions of the attacker. Thus in the general case, it is difficult to assess the reliability of Method 1. In Section IV, we analyze the probability of scalar sensor error due to an attack and state some conditions under which this probability may be acceptably small to render the scalar sensors more reliable. Method 2 shown in Table II takes a more conservative approach by always marking a disputed frame as an "event". Thus, if *either* $s = 1$ or $I_m = 1$, a frame will be sent to the cluster-head. Finally, in approach 3, we always trust the I_m decision obtained through image processing, even if it differs from the scalar sensor decision.

TABLE II
DECISION METHODS AT THE CAMERA NODES

Method	Action
Method 1	Trust scalar-sensor s
Method 2	Always mark as "event" if disputed
Method 3	Trust Image Processing I_m

III. EVENT-DETECTION IN WVSN CAMERA NODES

A. Event-Detection and Sequence Characteristics

An extensive body of research exists examining image change and shot change detection in image sequences [20], [21]. Much of this research focuses on the processing, segmentation and classification of a stored movie or news-cast which is *already*

[2] In this work, we consider the overall detection decision at *each* time. Since objects may enter and exit a frame abruptly from frame to frame, the temporal relationships between decisions are currently not examined.

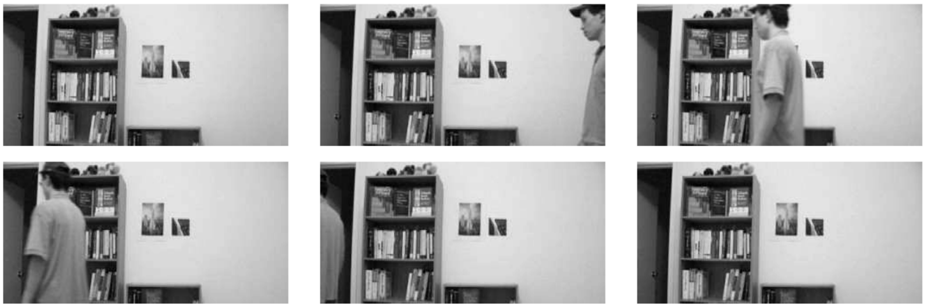

Fig. 2. Sequence 1 with frames (a)–(f) from top left to bottom right: indoor test conditions with constant lighting and no background changes.

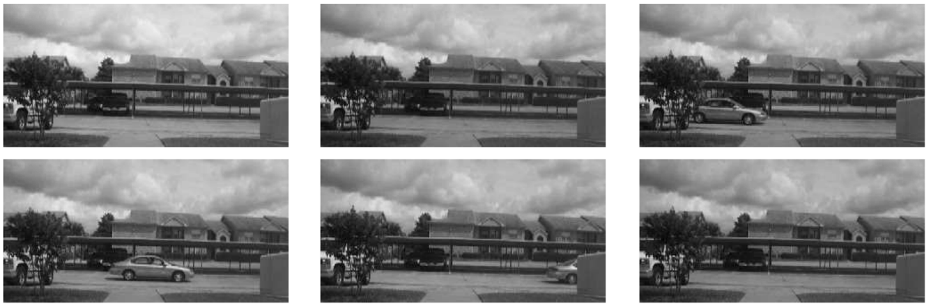

Fig. 3. Sequence 2 with frames (a)–(f) from top left to bottom right: outdoor variable lighting due to clouds. Ex: The light intensity changes by 70% between frames (a) and (b). Additional background movement due to shrub.

available in its entirety [22] or which obeys certain statistical assumptions. For instance, given an entire image sequence, it is often possible to obtain meaningful statistics via pre-processing of the frames [23]. The obtained statistics can then be utilized in detection and segmentation algorithms to distinguish between event and non-event frames. We note that the definition of event and non-event is largely application-dependent, as is the domain in which the processing is carried out (spatial, temporal or frequency-based) [24].

In contrast, a WVSN camera node collects *incoming* images containing unknown objects which may or may not enter into the frame at any time [6], [25]. Since we do not know ahead of time what objects will be encountered and under what lighting and changing background conditions, it is not possible to assume that we know the statistics of an event frame. Rather we make the weaker assumption that an event frame is one where "significant" motion has occurred [26]. Under certain deployment scenarios however, it may be reasonable to assume that we know the approximate statistics of a *non-event* frame (called the "null hypothesis" statistics) [12]. For instance, in this work we assume that a camera node is deployed or activated during a non-event time. The visual-node is thus able to process the initial frames it collects and determine some approximate null-hypothesis statistics.

The specific statistics utilized by the camera nodes may be dictated by the application and by the need to conserve energy [20]. The latter constraint is particularly worthy of mention in WVSNs since more sophisticated (but potentially more energy consuming) image processing might yield a better detection probability for a given false alarm rate [6]. Such improved performance translates into savings in transmission energy since fewer irrelevant (non-event) frames are sent [27]. However allocating more energy to image processing may drain the nodes too quickly, limiting their lifetime. The tradeoffs *among* various energy-allocation schemes for event-detection are a subject of ongoing study. In this work, we focus on achieving a lower probability of error given the use of scalar-sensors for a *fixed* chosen low-complexity image processing algorithm.

We consider a relatively simple and general event-detection algorithm (i.e., the detection statistics employed are not tailored to the detection of any *specific* object). The choice of algorithm (detailed below) is based in part on processing simplicity and in part on observations regarding the real-world image sequences used in our testing as shown in Figs. 2–5. The sequence of Fig. 2 is an idealized indoor test where the lighting and background conditions do not change appreciably over time and where the subject occupies most of the camera view. The only significant change comes from the event of interest in the form of a test subject entering the camera's field of view. The dominant source of noise in this case is internal camera noise and flicker.

The sequence of Fig. 3 shows outdoor parking-lot surveillance on a windy day, where the event of interest is the passing of an unidentified car. The detection task in this sequence is complicated by the presence of a nearby shrub which experiences significant swaying of its branches over time. Furthermore, the background lighting changes visibly with cloud movement (between frames 2(a) and 2(b), for example). The sequence of Fig. 4 also experiences changes due to swaying trees and variable light conditions. The event of interest is the appearance and small movement of a test subject which temporarily disappears behind a tree in frames 3(c) and 3(e).

Statistical analysis of image sequences 2 and 3 (such as Levine's Test and the t-test [28]) reveal that the mean and standard deviation are not reliable indicators of an event of

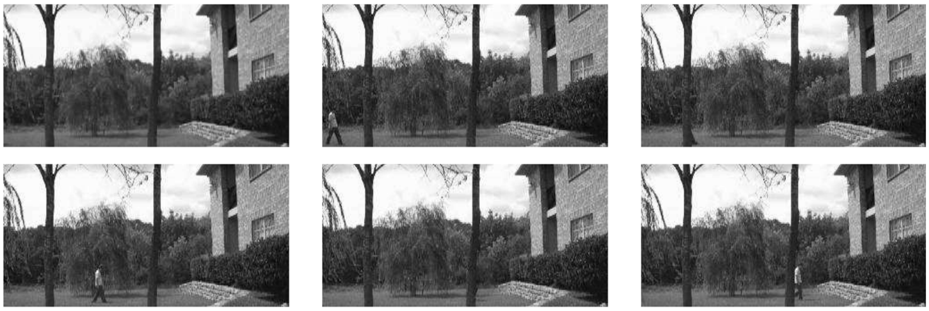

Fig. 4. Sequence 3 with frames (a)–(f) from top left to bottom right: changing outdoor light and background (swaying trees). The subject temporarily disappears behind a tree in frames (c) and (e).

Fig. 5. (a) Sequence 4(a) showing Sequence 2 modified to remove the shrub. (b) Sequence 4(b) showing Sequence 3 modified to remove the swaying trees.

interest occurring even after various filtering mechanisms are employed. This can be seen intuitively from the fact that the subjects of interest (person walking and car driving-by) do not occupy a much larger percent of a frame's pixels than the other randomly moving objects (shrub and trees). Hence, the mean and variance of the frames *do* change based on the appearance of the subject, but these differences are not statistically significant. In essence, the pixels corresponding to the person and car are getting dwarfed by the presence of many shrub and tree pixels which are also changing over time. Truncating the frames of sequences 2 and 3 as shown in Fig. 5 to exclude the vegetation does indeed improve the statistical difference between an event and non-event frame. However for the general WVSN deployment case (with cameras facing in various directions), we do not wish to select an event-detection technique which relies on the truncated assumption.

Based on the observed statistical similarity of event and non-event frames, we wish to determine an event detector suitable for WVSNs. In addition to its generality (detection not tailored to a specific type of object) and good detection performance, the chosen event-detector should be implementable in the simple WVSN devices. In addition to their hardware and general processing limitations, WVSNs process a large volume of surveillance frames which must in turn be transmitted wirelessly to the cluster-head if they contain an event of interest. Analysis of frames at the small block or pixel level is thus not a first-choice alternative for WVSNs.

Instead we seek a simple form for the detector where a single frame statistic is compared to a threshold in order to determine the presence or absence of an event. However as discussed, event and non-event frames from real-world surveillance sequences have similar statistics. Furthermore it can be shown (Appendix I) that a difference image $D = B - A$ computed from two consecutive frames A and B is not perfectly Gaussian but rather contains significant outliers (this is shown for both event and non-event frames). An optimal non-parametric (robust) detector is thus more appropriate for this case of statistical similarity and presence of outliers. However we show that a simple chi-squared detector (relying on a comparison of a frame statistic to a threshold) is *equivalent in form* to the robust detector and can thus be used in WVSNs (Section III-B). Furthermore, through the use of composite hypothesis testing (Section III-C), we show that the chi-squared detector can be made uniformly most powerful (UMP) through proper threshold selection. The UMP property signifies that the detector achieves a probability of detection P_D *higher or equal to* the detection of all other detectors given the worst-case scenario probability of false alarm P_{FA}. In other words, no detector performs better given the same probability of false alarm.

B. Chi-Squared Detector

We describe the chi-squared detector as a simple adaptation of the detector proposed by Aach and Kaup [12], [29], where we use entire difference frames instead of blocks. We now overview the technique. In essence, a difference image $D = B - A$ between two consecutive frames A and B reveals all the pixels that have changed between these frames (containing both relevant and irrelevant changes such as the tree swaying). The mean squared error (MSE) of the difference image is computed as the relevant statistic, and it is compared to a theoretically-obtained robust threshold T. We now present the specific details of this detector.

In Aach and Kaup [12], [29] (and in Radke *et al.* [26]), the difference image D is computed and divided into smaller blocks. Importantly, each pixel of the *difference* image is modeled as a Gaussian random variable with 0 mean and variance σ_i^2, where $i = 0$ corresponds to a non-event frame and $i = 1$ corresponds to an event frame. In order to conserve computational energy, in this work we use the entire difference image instead of the block-based solution. The resulting detector hypothesis test can be summarized as

$$\mathcal{H}_0 : \quad \text{no event}, D_k \sim \mathcal{N}\left(0, \sigma_0^2\right) \ \forall k \tag{1}$$

$$\mathcal{H}_1 : \quad \text{event}, \ D_k \sim \mathcal{N}\left(0, \sigma_1^2\right) \ \forall k \tag{2}$$

with $\sigma_0^2 < \sigma_1^2$ and where D_k is the kth difference pixel in $D = B - A$. Since the entire difference image is utilized in the detection, instead of considering individual pixels we may consider a new random variable defined as

$$X = \sum_{k=1}^{n} D_k^2 = \sigma_j^2 \sum_{k=1}^{n} \frac{D_k^2}{\sigma_j^2} = \sigma_j^2 Y, \quad \text{for j} \in \{0, 1\} \tag{3}$$

where Y has distribution chi-squared with n degrees of freedom and where n is the total number of pixels in the difference frame. The new detection hypothesis test is thus given by

$$\mathcal{H}_0 \quad : \quad X \sim \frac{1}{\sigma_0^2} f_{\chi^2,n}\left(\frac{x}{\sigma_0^2}\right) \tag{4}$$

$$\mathcal{H}_1 \quad : \quad X \sim \frac{1}{\sigma_1^2} f_{\chi^2,n}\left(\frac{x}{\sigma_1^2}\right) \tag{5}$$

where $f_{\chi^2,n}(x)$ is the probability density function (pdf) of the chi-squared distribution with n degrees of freedom. Hence the hypothesis test is given by (6) for false alarm rate not exceeding α where $f_{\chi^2,n}^{-1}$ is the inverse of the chi-squared distribution and the threshold T is $\sigma_0^2 f_{\chi^2,n}^{-1}(1-\alpha)$

$$x \underset{H_0}{\overset{H_1}{\gtrless}} \sigma_0^2 f_{\chi^2,n}^{-1}(1-\alpha) \tag{6}$$

C. Detector Properties

We first show that the simple chi-squared detector can be made uniformly most powerful (UMP) [30]. We begin by showing that if there exists a real positive number γ, such that $\sigma_0^2 < \gamma$ and $\sigma_1^2 > \gamma$, where the actual σ_0^2, σ_1^2 are *unknown*, then there exists a UMP detector where a realization x from (3) is compared to a threshold T, such that the probability of false alarm $P_{\rm FA} = \alpha$ is given by

$$\alpha = \sup_{\sigma_0^2 < \gamma} \int_T^{\infty} \frac{1}{\sigma_0^2} f_{\chi^2,n}\left(\frac{x}{\sigma_0^2}\right) dx. \tag{7}$$

Proposition 1: Suppose there exists a $\gamma > 0$, such that $\sigma_0^2 < \gamma$ and $\sigma_1^2 > \gamma$ in (4) and (5). Then there exists a UMP test of the form

$$x \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma f_{\chi^2,n}^{-1}(1-\alpha) \tag{8}$$

for false alarm rate not exceeding α. Proposition 1 is a composite hypothesis test in which the parameters for the null and alternate hypotheses are unknown, but the regions for these parameters are divided by a threshold γ. The proposition says that if the parameter space is divided as thus, then a test that compares the actual x in (3) to a threshold achieves optimal detection when the worst case false alarm is considered [the use of $\sup$ in (7)].

Proof: If we can show that the likelihood ratio is monotonically increasing in x for $\sigma_1^2 > \sigma_0^2$, then the UMP test of the form in (8) follows from a theorem on composite hypothesis testing [30]. It can easily be shown that the log-likelihood ratio is given by $(1)/(2)((1)/(\sigma_0^2) - (1)/(\sigma_1^2))x + (n)/(2)\ln((\sigma_0^2)/(\sigma_1^2))$. Since $\sigma_1^2 > \sigma_0^2$, this ratio is strictly increasing in x. To show that T is as given on the left side of (8), we note that the probability of false alarm is given by $1 - f_{\chi^2,n}(T/\sigma_0^2)$ by applying an integration change of variable in (7). To get the $\sup$ in (7), it suffices to set $\sigma_0^2 = \gamma$. ■

The statistical similarity of event and non-event frames along with difference-frame distributions that are not quite Gaussian (Appendix I) render $\mathcal{H}_1$ and $\mathcal{H}_0$ almost indistinguishable when the entire frame is used. Thus we would like to maximize the event detection assuming that $\sigma_1^2 \approx \sigma_0^2$ rather than assuming that the statistics are significantly different. This can be re-phrased as

$$\max \left.\frac{\partial \beta}{\partial \sigma_1^2}\right|_{\sigma_1^2 = \sigma_0^2} \tag{9}$$

where $\beta = \Pr\{\text{declare } H_1 \,|\, H_1 \text{ occurs}\}$ is the probability of detection.

Proposition 2: The test

$$x \underset{H_0}{\overset{H_1}{\gtrless}} T \tag{10}$$

maximizes (9) for a false alarm rate not exceeding α, i.e., T is chosen so that

$$\alpha > \int_T^{\infty} \frac{1}{\sigma_0^2} f_{\chi^2,n}\left(\frac{x}{\sigma_0^2}\right) dx. \tag{11}$$

Proof: By the proof of the Neyman-Pearson lemma [30], the optimal test can be shown to be of the form

$$\frac{\left.\frac{\partial \frac{1}{\sigma_1^2} f_{\chi^2,n}\left(\frac{x}{\sigma_1^2}\right)}{\partial \sigma_1^2}\right|_{\sigma_1^2 = \sigma_0^2}}{\frac{1}{\sigma_0^2} f_{\chi^2,n}\left(\frac{x}{\sigma_0^2}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \tilde{T} \tag{12}$$

which is equivalent to

$$\frac{x - n\sigma_0^2}{2\sigma_0^4} \underset{H_0}{\overset{H_1}{\gtrless}} \tilde{T}. \tag{13}$$

Letting $T = 2\sigma_0^4 \tilde{T} + n\sigma_0^2$ proves the proposition. ■

In summary, given the actual statistics of the difference image, a non-parametric (robust) detector is appropriate to perform event detection. However the simple chi-squared detector is equivalent in form to the robust detector and can be made uniformly most powerful through threshold selection. The simple image difference test may thus be used at the camera nodes with acceptable performance within its class of algorithm complexity.

IV. Scalar Node Reliability

A. System Model

To serve as a reliable mechanism for event-driven WVSNs, scalar sensor decisions must not be prone to errors caused by failure, miscalibration or attack [10]. Component testing (pre or post-deployment) may guarantee an acceptable measure of confidence in the sensor components. Furthermore, cryptographic and signal processing based techniques may thwart a variety of sensor network attacks where nodes or their cryptographic keys are captured [16]. Such techniques thus largely eliminate the possibility of false data injection by a malicious entity. The envisioned extension of WSNs to include mobility and actuation however, opens up the possibility of new and largely unexplored attacks [15]. As we will discuss, at least one new form of attack is particularly relevant to sensor nodes performing event-detection.

We consider the presence of a rival network with hostile nodes as depicted in Fig. 1 and as studied in our prior work on sensor network attacks in [11], [15] and [17]. The rival network may be deployed by a foreign entity for seemingly legitimate data collection. Its very presence in the environment hence does not raise an initial alarm. We consider the scenario where the hostile nodes directly perturb the observations collected by the legitimate network from the environment. We refer to this type of attack as an *actuation* attack to distinguish it from cryptographic attacks. The type of perturbation is largely dependent on the type of scalar sensor, such as temperature, motion or sound.[3]

For purposes of generality and tractability we therefore model the scalar-sensor event-detection via a sensor's ultimate output *decision*, "event present" $(s = 1)$ or "event absent" $(s = 0)$. In the *absence of an attack*, the detection of an event of interest in the environment at scalar node i is modeled as a random variable X_i. Given our assumption that the sensor decisions are hard and not soft (i.e., $s_i = 0$ or $s_i = 1$ only), X_i is modeled using a Bernoulli distribution $\mathrm{Bern}(p_i) = \mathrm{Bern}(p)$ for all i as shown in (14). This model emerges out of the physical proximity of the nodes within a cluster. Based on such proximity, we assume that each node in a cluster experiences the same phenomenon probability mass function (PMF). Importantly, nodes within a different cluster may experience a different PMF. If additionally the nodes within the same cluster are homogeneous (configured with the same detection thresholds), then the resulting probability of event p is the same for each node within a cluster. We thus obtain the model shown in (14) where a realization $x_i = 1$ denotes a scalar-sensor i having detected an event of interest

$$X_i = \begin{cases} 1, & \text{w.p. } p \\ 0, & \text{w.p. } 1-p \end{cases} \tag{14}$$

An error at scalar-sensor i due to an actuation attack is correspondingly modeled as a random variable Y_i distributed with a Bernoulli distribution $\mathrm{Bern}(q)$. In this model $y_i = 1$ corresponds to an attack at node i while $y_i = 0$ corresponds to the case of no attack. The *effect* of the error due to attack is modeled as changing the decision of a scalar-sensor, from declaring a 1 (event occurred) to declaring a 0, or vice-versa. This effect is given by (15), where X_i is a decision based on the *true* event in nature, Y_i is the attack and Z_i is the resulting bit reported to the cluster-head. It can be shown that the distribution of the random variable Z_i is given by $\mathrm{Bern}(r)$ with $r = (1-p)q + (1-q)p = q + p - 2pq$. We note that although we started with the practical notion of an actuation attack, the abstract model reduced to the well-known binary symmetric channel model. This simplification brings the new problem into a well-understood framework and becomes a key enabler in the subsequent use of game-theoretic analysis

$$Z_i = X_i \oplus Y_i. \tag{15}$$

Assuming that there are n scalar sensors reporting to a cluster-head, the latter receives a data-vector $\mathbf{z}$ of length n which it uses to verify the reliability of the scalar sensors. The cluster-head must thus decide between two alternative hypotheses, as follows:

$$\begin{aligned} \mathcal{H}_0 &\quad : \quad \text{normal operation, PMF} \sim \mathrm{Bern}(p) \\ \mathcal{H}_1 &\quad : \quad \text{erroneous operation, PMF} \sim \mathrm{Bern}(r) \end{aligned}$$

where PMF stands for probability mass function. If the error q was caused by component failure instead of attack, prior quality testing may establish a reasonable estimate for q. Additionally if a similar probability estimate can be obtained for the occurrence of an event of interest (which realistically may not be available), then the probability p may also be obtained. In such a case, an optimal Neyman-Pearson (NP) binary detector may be used to distinguish between the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ at the cluster-head.[4] In the case of hostile conditions, however, it may not be possible to obtain an estimate for q given that it is controlled by an attacker who is free to change the attack statistics over time. Indeed the reason why an attacker may employ a random attack in lieu of a deterministic one is precisely to thwart attempts by the WVSN to estimate its attack patterns.

Without information about q the cluster-head faces a difficult task in distinguishing between $\mathcal{H}_0$ and $\mathcal{H}_1$. However, for an event of interest (such as a person walking-by) occurring with some probability p and a fixed deployment of the nodes, we may expect that an average number c of nodes will record the event. For example, the object may appear in approximately $5 \pm \epsilon$ out of the 20 cameras in the cluster at any given time (which specific cameras capture the object depends on the subject's trajectory into the sensor field and is unknown). If p is the underlying probability of an event and n is the number of nodes, then c may be approximated as $c = np$ (this approximation improves as n increases). We thus propose the use of a *count* or *type* detector modified from [31]. The modification is based on the observation that in practice the sequence of observations collected by a single node over *time* is most likely not truly i.i.d (i.e., p may not be i.i.d over time). Furthermore, we wish to detect an attack early and to conserve storage space in a node's buffer. Thus we do not wish to collect a relatively long sequence of observations. Instead we assume that in a *given* time interval, p is i.i.d *spatially* within a cluster of nodes and q (which is unknown) is also i.i.d spatially within this cluster. Hence we propose the use of the detector $D(\mathbf{z})$ given in (16), where $c = np$ is the average number of 1s the cluster-head expects to receive from the n sensors, $w(\mathbf{z})$ is the *actual* count or *weight* (number of 1s) received in the data vector $\mathbf{z}$, and ϵ is a variance-related

[3] For example, hostile nodes may utilize micro-actuators to raise the acoustic readings of surrounding scalar sensors thus causing a false alarm regarding the presence of an object of interest to the WVSN.

[4] The NP detector forms the likelihood ratio of the two PMF's and compares it to a threshold determined by the desired probability of false alarm. If the likelihood ratio exceeds the threshold, the detector chooses hypothesis $\mathcal{H}_1$, otherwise it chooses $\mathcal{H}_0$.

"slack-factor" allowing the cluster-head to relax or tighten the detection constraint

$$D(\mathbf{z}) = |w(\mathbf{z}) - c| \underset{H_0}{\overset{H_1}{\gtrless}} \epsilon. \tag{16}$$

B. Game-Theoretic Analysis of Error Probability

Having established a model for the attack and its detection (at the cluster-head) via (15) and (16), respectively, we would like to determine the resulting probability of error q for the scalar sensors due to attack. We consider the case where the attacker's goal is to cause errors while minimizing the chance of getting detected (detection would prevent the attacker from continuing to misguide the network).

Given that attack detection is performed via the $D(\mathbf{z})$ detector in (16), to *avoid* detection the attacker wishes to maximize the probability $\Pr\{|w(\mathbf{X} \oplus \mathbf{Y}) - c)| < \epsilon\}$ where $\mathbf{Y}$ is the actuation vector. In other words, the attacker wishes to modify the sensor observations while remaining within ϵ of the data weight expected by the cluster-head on average for a given deployment. We thus denote the above probability by P_S or probability of *stealth* (attack avoidance). Through the use of combinatorial analysis based on our previous work in [15], we obtain the following expression for the probability of attack stealth P_S, which is also detailed in Appendix II.

Result 1: The attacking network wishes to maximize the probability P_S of a stealthy (thus successful) attack where P_S is given by (17) and where, a, b, and c are binomial coefficients given by (22), (23), and (24) in Appendix II

$$\begin{aligned} P_S &= \Pr\{|w(\mathbf{X} \oplus \mathbf{Y}) - c)| < \epsilon\} \\ &= \sum_{m=\lceil \frac{l-\epsilon}{2} \rceil}^{\lfloor \frac{l+\epsilon}{2} \rfloor} \sum_{k=1}^{n} \sum_{l=1}^{n} a(k,m)b(k,l,m)c(k) \\ &\quad \cdot p^k(1-p)^{n-k} q^l (1-q)^{n-l}. \end{aligned} \tag{17}$$

Having obtained an expression for P_S in terms of the probability of attack q and the probability of an event p, we now wish to determine the optimal q that maximizes P_S for a given cluster-size n and for a p that is not known a priori. Given the combinatorial form of (17), the best response of the attacker in terms of the selection of q is performed using asymptotic approximations as shown in Appendix III. We now state the result of this analysis.

Result 2: Given that the cluster-head performs a check on the scalar sensors utilizing the $D(\mathbf{z})$ detector, the optimal attack probability q for a p that is unknown *a priori* is forced to be *small* for *larger* clusters n. For smaller clusters (such as 10 nodes or less), this result does not hold and the attacker's optimal q may be quite large.

Finally, we wish to determine how the optimal q is affected by the probability of an event p for some fixed n, utilizing a best response analysis and asymptotic approximations as shown in Appendix IV. We now state the obtained result.

Result 3: The attacker's optimal q is affected by the true probability of an event p. Specifically, q is forced to be smaller for $p \ll 0.5$ (rare events) or for $p \gg 0.5$ (very common events).

TABLE III
OPTIMAL q FOR CLUSTER SIZE n AND PROBABILITY OF EVENT p

n	**Best q for $p = 0.5$**	**Best q for $p = 0.1$**
100	0.031	0.022
50	0.062	0.044
40	0.0780	0.055
30	0.103	0.074
20	0.152	0.110
10	0.990	0.205
5	0.505	0.405
3	0.655	0.655
2	0.999	0.999
1	0	0

Unlike Result 2 which has ramifications for choosing a cluster size n, Result 3 impacts WVSN design in a less direct way, since p cannot be controlled. However it provides an indication of conditions under which a scalar sensor's decision s may be less reliable and the camera decision I_m should be favored.

We now illustrate the above results with some specific examples. Table III shows the optimal value of q for $p = 0.1$ and for $p = 0.5$ (event occurring with probability of 10% and 50%, respectively). We see that for $n = 20$ and above, the attacker's optimal value of q is actually quite small $(q \approx 0.1)$ for both p's, in agreement with Result 2. However, for smaller clusters, such as $n = 10$ and below, the optimal value of q is quite large indicating that the scalar sensors are not reliable. In the degenerate case of a cluster of size 1, the only way for an attacker to evade detection is to actuate with a probability of 0 as we would expect. These results are depicted visually in Fig. 6 where the optimal value of q corresponds to a peak in the plot of P_S (P_S corresponds to the cluster-head *not* having detected the attack). We observe that for large n as in Fig. 6(d), the peak of P_S narrows and the optimal q becomes quite small. For small clusters such as in Fig. 6(b), the optimal q can be surprisingly large due to the combinatorial nature of P_S [(17) based on Result 1]. Finally, the behavior of probability P_S is also detailed in Table IV for $p = 0.1$ and $p = 0.5$. Again for clusters larger than $n = 20$, P_S decreases, especially for p away from 0.5 in agreement with Result 3.

Importantly, we note that a small value of q as the optimal for the attacker for large n emerges purely from the maximization of P_S. Thus *achieving* a small q does not actually *require* feedback from the cluster-head to the scalar nodes regarding the result of the $D(\mathbf{z})$ test. Rather, it is achieved by the mere fact of *having* a test at the cluster-head and further controlled via the selection of cluster size n. From the point of view of communication energy and delay, this is an advantageous property for wireless networks.

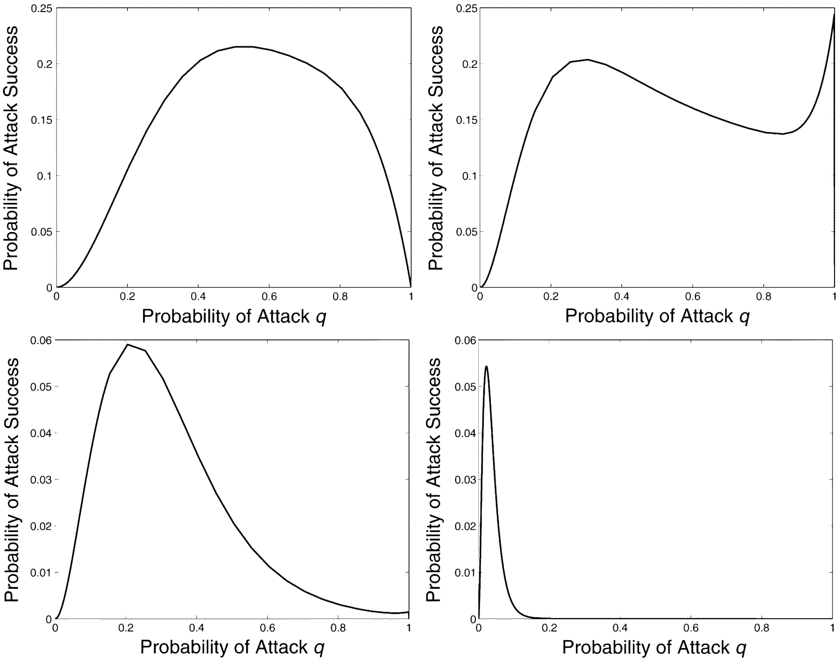

Fig. 6. Probability of attack success P_S versus probability of attack q for cluster size n and probability of event p. (a) $n = 5$ and $p = 0.5$. (b) $n = 10$ and $p = 0.5$. (c) $n = 10$ and $p = 0.1$ (d) $n = 100$ and $p = 0.1$.

TABLE IV
PROBABILITY OF ATTACK SUCCESS P_S

n	P_M **for** $p = 0.5$	P_M **for** $p = 0.1$
100	0.1945	0.0543
50	0.195	0.0548
40	0.196	0.055
30	0.1968	0.0555
20	0.1985	0.0563
10	0.2437	0.059
5	0.2150	0.0661
3	0.222	0.0799
2	0.499	0.1796
1	0	0

V. EVENT-DRIVEN WVSN PERFORMANCE RESULTS

The performance of event-driven WVSNs depends critically on correct frame selection based on the definition of an event. The camera nodes perform event detection based on a simple image processing algorithm with a resulting probability of error q'. To attain a high P_D for a chosen $P_{\rm FA}$, a camera node relies on support from a scalar sensor node. The scalar sensor node may itself however be in error due to attack with probability q. Since q and q' may not be equal or known, we are interested in determining how the camera nodes should utilize the decisions I_m and s in various scenarios.

TABLE V
VISUAL DETECTION BASED PERFORMANCE (I_m)

Image Sequence	$\mathbf{P_D}$	$\mathbf{P_{FA}}$	$\mathbf{P_{err}}$
Seq. 1. Indoor walking	1.0	0.13	0.03
Seq. 4a. Outdoor car, no trees	0.98	0.17	0.11
Seq. 4b. Outdoor walking, no trees	0.50	0.03	0.17
Seq. 2. Outdoor car, with trees	0.87	0.26	0.20
Seq. 3. Outdoor walking, with trees	0.05	0.23	0.90

The image processing algorithm described in Section III was implemented in Matlab and tested on the image sequences shown in Figs. 2–5.[5] The detection performance of this algorithm is shown in Table V where the results are listed in order of degrading event-detection performance. Table V shows the probability of detection P_D (which ideally should be as close to 1 as possible) and the probability of false alarm $P_{\rm FA}$ (which ideally should be as close to zero as possible). Since both probabilities are important to the correct functioning of the WVSN, Table V also shows the total probability of error $P_{\rm err}$ which is defined in (18). We see that the simple image processing algorithm performs best on image sequences with the least amount of lighting and background changes such as Sequence 1. It also performs well for cases where the subject occupies a significant portion of the frame such as in

[5] We note that the test sequences contain a single object of interest and as such, are a worst-case scenario for the event-detection algorithm. The presence of multiple moving objects increases the probability of detection by increasing the percent of relevant change pixels in a frame.

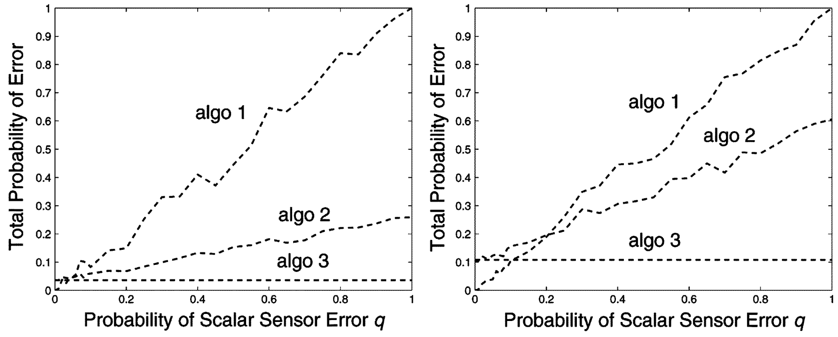

Fig. 7. Comparison of the P_{err} for the three methods. (a) Sequence 1 (indoor walking). (b) Sequence 4(a) (outdoor car no trees).

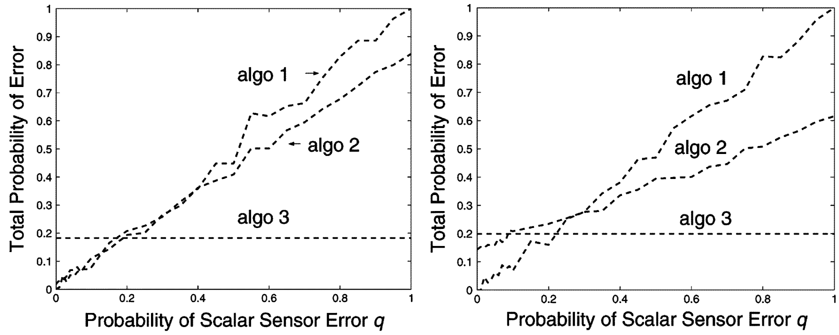

Fig. 8. Comparison of the P_{err} for the three methods. (a) Sequence 4(b) (outdoor walking, no trees). (b) Sequence 2 (outdoor car with trees).

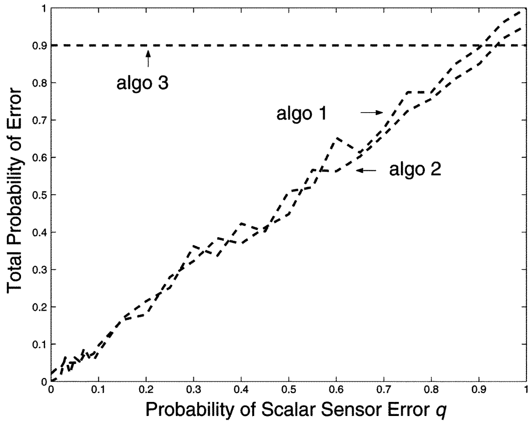

Fig. 9. Comparison of the P_{err} for the three methods Sequence 3 (outdoor walking with trees).

Sequence 4(a). For certain outdoor image sequences such as Sequence 3, P_{err} is unacceptably high at 0.9

$$P_{\text{err}} = P_{\text{FA}} \cdot P(\mathcal{H}_0) + (1 - P_D) \cdot P(\mathcal{H}_1). \qquad (18)$$

We thus wish to determine if the scalar sensors can help the camera nodes achieve an improvement in performance given their attack-prone nature. The methods incorporating scalar-sensor data into the decision process as listed in Table II were implemented using Matlab. For convenience, we recap these methods.

- Method 1: rely on scalar decisions (especially if I_m and s do no agree).
- Method 2: always send a frame that was marked as an event by *either* I_m or s.
- Method 3: rely on the image processing decision I_m. This is essentially the case with no scalar sensors. It is used for comparison with Methods 1 and 2.

The total probability of error P_{err} obtained for each image sequence using Methods 1, 2, and 3 for a camera node relying on one scalar sensor node is shown in Figs. 7–9, where algo_i refers to Method i and the scalar sensor error q is varied. We make the following observations.

1) Incorporating and trusting the scalar sensor decisions as in Method 1 greatly reduces P_{err} for most sequences, as long as $q < q'$. This is especially true for the more difficult sequences (Sequence 2 and 3) containing lighting and background changes such that q' is large. The condition for $q < q'$ may be met if q is due to a sensor network attack and cluster-head checking and cluster-size selection is performed as outlined in Section IV-B. As an example, Table VI summarizes P_{err} for $q = 0.1$ (corresponding to an optimal attack when the cluster has 20 nodes) for the three methods. We see that in this case Method 1 achieves the best performance except for the indoor sequence Sequence 1 where $q' < q$.

2) In the regime of larger sensor error q, Method 3 relying on I_m is superior to Method 1. This condition appears to be met for sequences with few significant background and lighting changes (such as indoor sequences). It may also be met if the largest source of scalar sensor error q is not due to attack and cannot be controlled.

3) Method 2 is generally not optimal but for certain sequences, it performs close to Method 1. Crucially, Method 2 becomes important for cases where we do *not* know the relationship between q and q' (i.e., which of the two

TABLE VI
COMPARISON OF P_{err} FOR $q = 0.1$ FOR EACH METHOD (ALGO_i)

Image Sequence	algo1	algo2	algo3
Seq. 1. Indoor walking	0.1515	0.0573	0.03
Seq. 4a. Outdoor car, no trees	0.1238	0.1531	0.11
Seq. 4b. Outdoor walking, no trees	0.11	0.09	0.17
Seq. 2. Outdoor car, with trees	0.12	0.20	0.20
Seq. 3. Outdoor walking, with trees	0.05	0.23	0.90

errors is smaller). For instance, though both sequences in Fig. 8(b) and 9 were obtained outdoors in variable conditions, the image processing error q' varies dramatically between them.

Point 3 highlights the important observation that for arbitrary conditions where the WVSN nodes are scattered for surveillance, the relationship between q and q' may not be known. Specifically, if the largest source of scalar sensor error q is due to attack as assumed in this work, the value of q may be determined from the results shown in Section IV.B. However, the error q' associated with arbitrary camera conditions remains unknown. If we do not know whether $q < q'$ or $q > q'$, we may be forced to utilize Method 2 or another comparable method. As can be seen from Figs. 7 and 8, this method generally achieves a performance which lies in between the performance of Methods 1 and 3. The loss in performance not only impacts the detection error, but also the energy use of WVSN camera nodes which now transmit a frame if *either* $s = 1$ or $I_m = 1$.

Finally, we note that to obtain an estimate of q' we may rely on feedback from the cluster-head which was *not* required to obtain an estimate of q. Recalling that the cluster-head also receives the image frames, we can envision a scenario where the cluster-head performs its own analysis of the frame, perhaps utilizing a different image processing algorithm. The detection-computation-delay trade-offs of the various possible approaches form the basis of ongoing research.

VI. DISCUSSION AND CONCLUSION

In the discussion of wireless systems, especially ones deployed for the purposes of image acquisition, energy and delay play a critical role. In terms of energy consumption, wireless data transmission is known to be a dominant source of energy drain [8]. This energy expenditure becomes even more significant in wireless image networks [5]. Much research has thus focused on energy-efficient compression and encoding of video, such that the required computational energy is itself minimized [2]. In comparison with this approach, event-driven systems aim to minimize transmission energy via local selection of relevant image frames.

The computational energy required for robust frame selection is also minimized via the use of a lower-complexity image processing algorithm with a robust threshold. The false alarm rate α of this detector is adjustable [(6)], which is important in security and limited-energy systems where α translates into the number of non-event frames sent to the cluster-head. This detector also does not require a training phase other than the original acquisition of a few non-event frames at the beginning of operation and is thus suitable for rapid-deployment applications. In such applications, the main focus of the WVSN cameras is on surveillance gathering rather than on data analysis which is performed at the cluster-head or system sink (which generally has more energy than the battery-operated camera nodes) [32]. A simple and general visual detection algorithm may thus be appropriate for such applications, especially if augmented by decisions from scalar sensors which may be deployed quickly in the environment.

The energy required to process, encode and transmit scalar decisions is generally considered much smaller than that of video data and also incurs lower processing and transmission delay [2]. Furthermore the scalar sensor infrastructure is flexible and many techniques have been developed to support its function following ad hoc deployment [14]. Based on its visual-scalar collaboration and feedback from the cluster-head, the event-driven WVSN approach also inherently provides fault recovery mechanisms and a trigger-mode of operation. As the cluster-head receives both the scalar-sensor data and the corresponding image frames, it is able to correlate the visual and scalar event-detection. Discrepancies may be communicated to the scalar-sensors, instructing them to move away from their current locations which may be under the influence of an actuation attack. Such feedback may also be provided to the camera nodes instructing them to temporarily or permanently disregard the scalar sensor decisions and to rely on their image processing results, or vice versa, to omit the image processing phase and rely on the scalar sensors.

To summarize, in this work we examine the detection performance of *event-driven* WVSNs where image frames are transmitted back to a cluster-head *only if* the frame contains an event of interest. The decision regarding a frame can be made purely based on image processing techniques or it can be made using assistance from a scalar-sensor that may itself contain errors due to hostile attack. Through game-theoretic analysis, we show that the scalar sensor error can be controlled through cluster-head checking and appropriate selection of cluster size n without feedback from the cluster-head. We also examine a simple image processing algorithm and show its equivalence in form to a robust detector with good performance in indoor and some outdoor surveillance sequences. For sequences containing changing background and lighting conditions, this detector may produce a high probability of error. We thus discuss approaches for utilizing the scalar sensor decisions collaboratively with the camera decisions and determine that such decision fusion provides a reasonable detection performance for WVSNs.

APPENDIX I
IMAGE SEQUENCE CHARACTERISTICS

In this section, we analyze in greater detail the statistics of a representative image sequence, that is Sequence 2 from Fig. 3, showing a moving car with trees. As pointed out in Aach and Kaup [12], [29], the difference pixels D_i generally do not obey a normal distribution model though this assumption is commonly

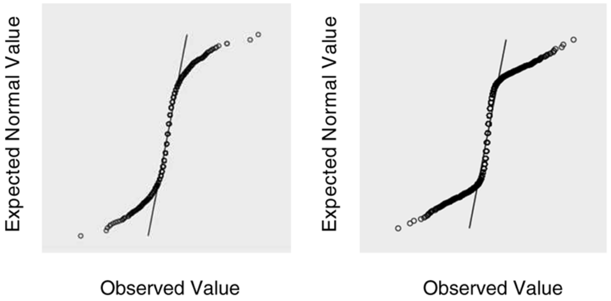

Fig. 10. Test for normality of difference images or q-q plot (car sequence) under (a) $\mathcal{H}_0$ and (b) $\mathcal{H}_1$.

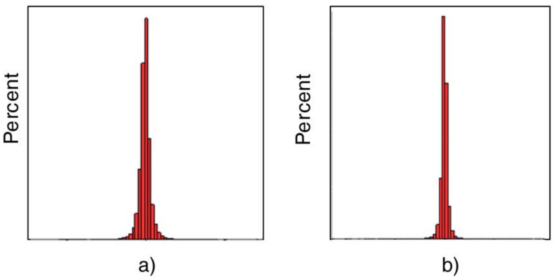

Fig. 11. Difference image histogram (car sequence) under (a) $\mathcal{H}_0$ and (b) $\mathcal{H}_1$.

made. To investigate the possible distribution of the difference pixels, we make use of a typical q-q plot as shown in Fig. 10 for the moving car sequence.

The q-q plot is used to test whether a set of samples are empirically Gaussian. The Gaussian assumption is rejected if the q-q plot returns a set of points that lie far off the straight line. As shown in Figs. 10(a) and (b), this is indeed the case for the car sequence under both hypotheses. The points lying off the main line may either indicate the presence of massive outliers in the data (the difference pixels), or may indicate that the distribution is not Gaussian. We thus call upon the use of histograms to provide further clues regarding the pixel distribution. The histograms for the difference images under the two hypotheses are shown in Fig. 11(c) and (d).

These histograms show that the distributions are most likely bell-shaped as can be confirmed through the use of Mardia statistics [33]. Based on the q-q plots and the histograms, we thus approximate the difference pixels as Gaussian. We note however that this assumption tends to fail for event frames (alternative hypothesis) where regions of interest that undergo change have a different distribution than regions of no change. Hence, strictly speaking, the Gaussian assumption only holds for certain non-event frames and certain regions of an event frame. It is nevertheless a helpful approximation that leads to a non-parametric (robust) detector.

In addition to the distribution assumption, we also consider the assumption that $\sigma_1^2 > \sigma_0^2$. Indeed this assumption is imperfect since in order for every difference pixel in an event frame to have σ_1^2, the object would have to move across the *entire* frame, thus covering all pixels. In Aach and Kaup [12], [29], smaller blocks are assumed to mitigate this flaw. However, through our use of entire difference frames, this assumption does not hold. To mitigate this issue we can assume that the effect of a small movement can be distributed across all pixels by decreasing σ_1^2. This assumption is more realistic since if we use the entire frame to estimate σ_1^2, then the estimate of σ_1^2 will mostly be of the non-moving regions. Thus the variance will be decreased to $\sigma_1^2 \approx \sigma_0^2$, leading to the use of a robust (non-parametric) detector as in Section III-C.

Appendix II
Game-Theoretic Analysis of Successful Attack

We show that the probability P_S of an unnoticed (hence successful) attack is given by

$$\begin{aligned} P_S &= \Pr\{|w(\mathbf{X} \oplus \mathbf{Y}) - c)| < \epsilon\} \\ &= \sum_{m=\lceil \frac{l-\epsilon}{2} \rceil}^{\lfloor \frac{l+\epsilon}{2} \rfloor} \sum_{k=1}^{n} \sum_{l=1}^{n} a(k,m) b(k,l,m) c(k) \\ &\quad \cdot p^k (1-p)^{n-k} q^l (1-q)^{n-l}. \end{aligned}$$

Since the cluster-head is employing the *type* detector $D(\mathbf{z})$ given by (16), we begin by examining the conditions required for the *weights* of the pristine data vector $\mathbf{x}$ and the potentially attacked data vector $\mathbf{z}$ to be the same (which causes the attack to be unnoticed).

Lemma 1: $w(\mathbf{x}) = w(\mathbf{x} \oplus \mathbf{y})$ if and only if the number of 1s in $\mathbf{x}$ and $\mathbf{y}$ coincide in exactly $m = (w(\mathbf{y}))/(2)$ positions.

Proof: Let $k = w(\mathbf{x})$, $l = w(\mathbf{y})$, and let A be the set of positions in $\mathbf{x}$ that have a value of 1, i.e., $A = \{i : x_i = 1\}$ and let B be the set of positions in $\mathbf{y}$ that have a value of 1, i.e., $B = \{i : y_i = 1\}$. Then $k = |A|$ and $l = |B|$. The set of positions in both A and B that coincide is $A \cap B$ so by definition $m = |A \cap B|$. The set of positions that do not coincide is $(A \cup B) - (A \cap B)$. In binary addition, bits that match always add to 0, and bits that do not match always add to 1, hence $w(\mathbf{x} \oplus \mathbf{y}) = |(A \cup B) - (A \cap B)| \overset{(a)}{=} |(A \cup B)| - |(A \cap B)| \overset{(b)}{=} |A| + |B| - 2|A \cap B| = k + l - 2m$, where (a) results because $A \cap B \subset A \cup B$, and (b) follows from the principle of inclusion and exclusion. If $w(\mathbf{x}) = w(\mathbf{x} \oplus \mathbf{y})$, then $k = k + l - 2m$, which implies $m = (l/2)$. Conversely, if $m = (l/2)$, then $w(\mathbf{x} \oplus \mathbf{y}) = k + l - 2(l/2) = k = w(\mathbf{x})$.

Lemma 2: Suppose $\mathbf{x}$ and $\mathbf{y}$ have exactly m overlapping 1s. Then $w(\mathbf{x}) \geq m$ and $w(\mathbf{y}) \geq m$ must satisfy

$$n - w(\mathbf{x}) \geq w(\mathbf{y}) - m \tag{19}$$

Proof: Define k, l, A, and B as in the proof of Lemma 1. The distinct 1-positions over both $\mathbf{x}$ and $\mathbf{y}$ are given by the set $A \cup B$. Since the total number of distinct 1-positions cannot exceed n (for either $\mathbf{x}$ or $\mathbf{y}$ would be of length greater than n), we have $n \geq |A \cup B| = |A| + |B| - |A \cap B| = k + l - m$.

The following definition is based on Lemma 2 and will be used in subsequent proofs.

Definition 1: Given a number m, the pair (k, l) is said to be *well-defined* if it satisfies $n - k \geq l - m$, and $n \geq k \geq m$, $n \geq l \geq m$.

In Lemma 3 we look at the conditional probability of $\mathbf{x}$ and $\mathbf{y}$ overlapping in exactly m positions. The technique used in the proof is to fix one of the vectors, i.e., $\mathbf{x}$, as shown in Fig. 12, and then choose $\mathbf{y}$s so that only m of their 1-s overlap with any

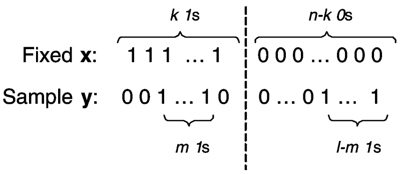

Fig. 12. Visualization of Lemma 3: choosing exactly m of $\mathbf{y}$'s 1s to overlap with $\mathbf{x}$'s 1s.

of the $\mathbf{x}$'s 1s. This probability turns out to be a hypergeometric distribution.

Lemma 3: Let E be the event that the number of 1s in $\mathbf{X}$ and $\mathbf{Y}$ overlap in exactly m positions. Then

$$\begin{aligned}\Pr(E|\{w(\mathbf{X}) = k \text{ and } w(\mathbf{Y}) = l\})\\ = \frac{\binom{k}{m}\binom{n-k}{l-m}}{\binom{n}{l}} \quad (20)\\ = \frac{\binom{l}{m}\binom{n-l}{k-m}}{\binom{n}{k}} \quad (21)\end{aligned}$$

when k and l are well-defined (as in Definition 1), otherwise the probability is 0.

Proof: We prove (20), and let the reader verify (21). We fix $\mathbf{x}$, and think of $\mathbf{y}$ as the binary string that we vary so that we may look at the event $E' = \{\mathbf{y}\,|\,\text{exactly } m \text{ of } \mathbf{y}'\text{s } l \text{ 1s overlap with } \mathbf{x}'\text{s } k \text{ 1s}\}$. We can count the number of $\mathbf{y}$ that satisfies E' by choosing m 1s in $\mathbf{y}$ from positions out of the k 1-positions in $\mathbf{x}$, which is $(\genfrac{}{}{0pt}{}{k}{m})$, and then choosing the remainder $l-m$ 1s in $\mathbf{y}$ from positions out of the $n-k$ 0-positions in $\mathbf{x}$, which is $(\genfrac{}{}{0pt}{}{n-k}{l-m})$. The multiplication counting rule gives $|E'| = (\genfrac{}{}{0pt}{}{k}{m})(\genfrac{}{}{0pt}{}{n-k}{l-m})$. Now if we vary $\mathbf{x}$, by the multiplication rule we have $|E| = (\genfrac{}{}{0pt}{}{n}{k})(\genfrac{}{}{0pt}{}{k}{m})(\genfrac{}{}{0pt}{}{n-k}{l-m})$. There are a total of $(\genfrac{}{}{0pt}{}{n}{k})(\genfrac{}{}{0pt}{}{n}{l})$ pairs of $(\mathbf{x}, \mathbf{y})$ of specified weights. Since all such pairs have the same probability, we can take the ratio of $|E|$ over the total number of pairs, giving us (20).

Theorem 1: Let E be the event that the number of 1s in $\mathbf{X}$ and $\mathbf{Y}$ overlap in exactly m positions. Define

$$a(k,m) = \begin{cases} \binom{k}{m}, & \text{if } k \geq m \\ 0, & \text{o.w} \end{cases} \quad (22)$$

$$b(k,l,m) = \begin{cases} \binom{n-k}{l-m}, & \text{if } n-k \geq l-m \\ 0, & \text{o.w} \end{cases} \quad (23)$$

$$c(k) = \begin{cases} \binom{n}{k}, & \text{if } n \geq k \\ 0, & \text{o.w} \end{cases} \quad (24)$$

In addition, we define $(\genfrac{}{}{0pt}{}{a}{b})$ to be equal to 0 if either a or b are not integers. Then

$$\Pr(E) = \sum_{k=1}^{n}\sum_{l=1}^{n} a(k,m)b(k,l,m)c(k)p^k \times (1-p)^{n-k}q^l(1-q)^{n-l}. \quad (25)$$

Proof:

$$\begin{aligned}&\Pr(\{w(\mathbf{X}) = k \text{ and } w(\mathbf{Y}) = l \text{ and}\} \cap E)\\ &= \Pr(E|\{w(\mathbf{X}) = k \text{ and } w(\mathbf{Y}) = l\})\\ &\quad \cdot \Pr\{w(\mathbf{X}) = k\}\Pr\{w(\mathbf{Y}) = l\} \quad (26)\\ &= \frac{\binom{k}{m}\binom{n-k}{l-m}}{\binom{n}{l}} \binom{n}{k} p^k(1-p)^{n-k}\\ &\quad \times \binom{n}{l} q^l(1-q)^{n-l}. \quad (27)\end{aligned}$$

Equation (27) follows from Lemma 3, where again we assume k and l are well-defined, or the probability is 0. Finally, we can extract the desired marginal distribution:

$$\Pr(E) = \sum_{k=1}^{n}\sum_{l=1}^{n} \Pr(\{w(\mathbf{X}) = k \text{ and } w(\mathbf{Y}) = l \text{ and}\} \cap E). \quad (28)$$

Corollary 1:

$$\begin{aligned}&\Pr\{w(\mathbf{X}) = w(\mathbf{X} \oplus \mathbf{Y})\}\\ &= \sum_{k=1}^{n} \sum_{l\,\mathrm{iseven}}^{n} a\left(k, \frac{l}{2}\right) b\left(k, l, \frac{l}{2}\right) c(k)\\ &\quad \cdot p^k(1-p)^{n-k}q^l(1-q)^{n-l} \quad (29)\end{aligned}$$

where $a(k,m)$, $b(k,l,m)$, $c(k)$ are as defined in Theorem 1.

Proof: Apply Lemma 1 to Theorem 1. ■

Corollary 2: Let ϵ be a positive integer. Then

$$\begin{aligned}&\Pr\{|w(\mathbf{X}) - w(\mathbf{X} \oplus \mathbf{Y})| < \epsilon\}\\ &= \sum_{m=\lceil \frac{l-\epsilon}{2} \rceil}^{\lfloor \frac{l+\epsilon}{2} \rfloor} \sum_{k=1}^{n}\sum_{l=1}^{n} a(k,m)b(k,l,m)c(k)\\ &\quad \cdot p^k(1-p)^{n-k}q^l(1-q)^{n-l} \quad (30)\end{aligned}$$

where $a(k,m)$, $b(k,l,m)$, and $c(k)$ are as defined in Theorem 1.

Proof: Given realizations $\mathbf{x}$ and $\mathbf{y}$, let $k = w(\mathbf{x})$ and $l = w(\mathbf{y})$

$$|w(\mathbf{x}) - w(\mathbf{x} \oplus \mathbf{y})| < \epsilon \quad (31)$$

$$\Rightarrow \begin{cases} k < k + l - 2m + \epsilon \\ k > k + l - 2m - \epsilon \end{cases}. \quad (32)$$

The second line relies on the proof for Lemma 1. Since the events of strings having 1-s overlapping in exactly m' and m'' positions are disjoint, the probability of the union of the events is the sum of the probabilities of the individual events. ■

APPENDIX III
GAME-THEORETIC OPTIMAL ATTACK PROBABILITY

We determine the optimal probability q that a hostile actuating network should use to maximize its probability of attack success P_S [18].

Theorem 2: For n sufficiently large, $q^* \downarrow 0$ as $n \to \infty$, where $q^* = \arg\max_{q \in [0,1]} \Pr\{w(\mathbf{X}) = w(\mathbf{X} \oplus \mathbf{Y})\}$.[6]

Proof: Let $\varphi(q) = \Pr\{w(\mathbf{X}) = w(\mathbf{X} \oplus \mathbf{Y})\}$ where X_i is $\mathrm{Bern}(p)$.[7] As $n \to \infty$, $(w(\mathbf{X}))/(n) \overset{\text{a.s.}}{\to} p$, and similarly $(w(\mathbf{Y}))/(n) \overset{\text{a.s.}}{\to} q$. Using this idea, it can be shown that $w(\mathbf{x}) \approx np$ and $w(\mathbf{y}) \approx nq$ for all realizations $\mathbf{x}$ and $\mathbf{y}$ for sufficiently large n. Hence, substituting $k = np$ and $l = nq$ into (29), we obtain

$$\varphi(q) \approx \binom{np}{nq/2}\binom{n(1-p)}{nq/2} \times \binom{n}{np} p^{np}(1-p)^{n(1-p)} q^{nq}(1-q)^{n(1-q)} \quad (33)$$

when n is sufficiently large. The conditions for being well-defined imply $q \leq \min\{2p, 2(1-p)\}$, which can readily be verified. Again for sufficiently large n, we apply Stirling's approximation ($n! \approx ((n/e))^n$) to (33). In (34), shown at the bottom of the page, we have removed the non-negative constants (independent of q). Next we examine the derivative of $\varphi(q)$

$$\begin{aligned} \frac{\partial \varphi(q)}{\partial q} & \\ \propto\, & 2^{nq-1} n^{1-nq} (1-q)^{n(1-q)} \\ & \cdot [n(1-p-q/2)]^{-n(1-p-q/2)} [n(p-q/2)]^{-n(p-q/2)} \\ & \cdot \ln\left(\frac{4(p-q/2)(1-p-q/2)}{(1-q)^2}\right). \end{aligned} \quad (35)$$

We can verify that the point $q = 0$ is a salient point (i.e., a "corner" since $\varphi(q)$ "turns on" at $q = 0$) so the derivative does not exist at this point. Hence our discussion of the derivative is restricted to $0 < q \leq \min\{2p, 2(1-p)\}$. Under the implications for being well-defined, it can be verified that the terms outside $\ln$ are always non-negative. However, the $\ln$ term can be shown to be non-positive by showing that the argument inside $\ln$ is always ≤ 1. Therefore, $(\partial \varphi(q))/(\partial q) \leq 0$ over the interval of $q \in (0, \min\{2p, 2(1-p)\}]$ implies $\varphi(q)$ is monotonically decreasing in this interval. The maximum q^* must then be the left boundary, i.e., $q^* \downarrow 0$. ■

Corollary 3: Suppose the malicious sensor network plays q^* according to Theorem 2. Let $l = E[w(\mathbf{Y})]$ be the average (or expected) number of legitimate bits that are flipped. Then $l \ll n$.

Proof: Since $l = nq^*$, therefore $(l/n) = (nq^*)/(n) = q^* \ll 1$, which implies $E[w(\mathbf{Y})] \ll n$. ■

APPENDIX IV
EFFECT OF THE PROBABILITY OF AN EVENT

In this work, we assume that an event of interest occurs with some probability p according to a Bernoulli distribution. *If* we consider nature as a "player" in a game against the actuating network (this is only conceptual), we are able to determine how the probability p affects the hostile nodes' choice of q. In what follows we refer to nature as "Player" 1 and to the hostile actuating network as Player 2.

For a static game, the pure strategy Nash equilibria are the strategy vectors (p^*, q^*) ($p^* \in [0,1]$, $q^* \in [0,1]$) such that Player 1 would not find it beneficial to deviate from p^* given Player 2 plays q^*, and vice versa [18]. This problem is generally difficult, but if we look at the asymptotic case, that is for n sufficiently large, we may determine the following result:

Theorem 3: Suppose that Player 1 can only play p^* from a closed subinterval of $[0,1]$, denoted $P = [a,b]$, $a < b$, while $q^* \in [0,1]$. For n sufficiently large, the pure strategy Nash equilibrium is given by

$$p^* = \begin{cases} a, & \text{if } a < |1-b| \\ b, & \text{if } a > |1-b| \end{cases} \quad (36)$$

and $q^* = \delta$, where $\delta \downarrow 0$. If $a = |1-b|$, then there are two equilibria at (a, δ) and (b, δ).

Remark 1: The expression in (36) refers to choosing the left-most boundary if it is closer to 0 than the right-most boundary is closer to 1, and choosing the right-most boundary if it is closer to 1 than the left-most boundary is closer to 0.

Proof: First we find the best response of Player 2 to Player 1's p. We have already shown in Theorem 2 that $q^* \downarrow 0$ as $n \to \infty$. For q^* is sufficiently small, we can assume it is irrespective of p. Next we examine the best response of Player 1 to Player 2's q, where we know that q will always be approaching 0. With this in mind, we define $\phi(p)$ for q fixed at q^*

$$\phi(p) = \sum_{k=1}^{n} \sum_{l \text{ is even}}^{n} \binom{l}{l/2}\binom{n-l}{k-l/2} \times \binom{n}{l} q^l (1-q)^{n-l} p^k (1-p)^{n-k} \quad (37)$$

where we have used (21) instead of (20) which we have been using so far. Now since q^* is small, $nq^* \ll n$ so $\binom{n}{l} q^l (1-q)^{n-l}$ can be approximated by a Poisson distribution $(\lambda^l)/(l!)e^{-\lambda}$, where $\lambda = nq^*$

$$\phi(p) \approx \sum_{k=1}^{n} \sum_{l \text{ is even}}^{n} \binom{l}{\frac{l}{2}}\binom{n-l}{k-\frac{l}{2}} \frac{\lambda^l}{l!} e^{-\lambda} p^k (1-p)^{n-k} \quad (38)$$

$$\approx \sum_{k=1}^{n-1} \binom{2}{1}\binom{n-2}{k-1} \frac{\lambda^2}{2!} e^{-\lambda} p^k (1-p)^{n-k} \quad (39)$$

$$\varphi(q) \propto \frac{(1-q)^{n(1-q)}}{\left(\frac{n}{2}\right)^{nq} [n(p-q/2)]^{n(p-q/2)} [n(1-p-q/2)]^{n(1-p-q/2)}}. \quad (34)$$

[6] Although q^* *approaches* 0, setting $q^* = 0$ is incorrect, as this results in $\varphi(0) = 0$.

[7] The difficulty in this proof lies not in taking the derivative of $\varphi(q)$, but rather in solving the first-order condition $(\partial \varphi(q))/(\partial q) = 0$; hence we resort to the asymptotic case as well as approximations.

where we have kept only the smallest $l = 2$ in the series. Next we use the identity $\binom{n-2}{k-1} + \binom{n-2}{k} = \binom{n-1}{k}$

$$\begin{aligned}\phi(p) &\approx \lambda^2 e^{-\lambda}\left\{\sum_{k=1}^{n-1}\binom{n-1}{k}p^k(1-p)^{n-k-1}(1-p)\right.\\ &\quad\left. - \sum_{k=1}^{n-2}\binom{n-2}{k}p^k(1-p)^{n-k-2}(1-p)^2\right\}\\ &= \lambda^2 e^{-\lambda}\left\{(1-p)\left[\sum_{k=0}^{n-1}\binom{n-1}{k}\right.\right.\\ &\quad\left.\times\, p^k(1-p)^{(n-1)-k}\right] - (1-p)^n\\ &\quad - (1-p)^2\left[\sum_{k=0}^{n-2}\binom{n-2}{k}p^k(1-p)^{(n-2)-k}\right]\\ &\quad\left. + \, (1-p)^n\right\} \qquad (41)\end{aligned}$$

where in the last line we have extended k to start at 0 in the series, and hence must subtract/add the $k = 0$ term to maintain equality. We have chosen to do this because each of the series sum to 1 as both series represent the total sum of a binomial distribution. The resulting expression $\phi(p) \approx \lambda^2 e^{-\lambda}\{(1-p) - (1-p)^2\}$ then simplifies to $\lambda^2 e^{-\lambda}(p - p^2)$.

This shows that $\phi(p)$ is approximated as a concave function with peak at $p = (1/2)$ when n is sufficiently large. If p can be chosen from the entire interval $[0, 1]$, then the minima of $\phi(p)$ would be at $p^* = 0$ and $p^* = 1$. If instead we have to choose p from the closed subinterval $P \subset [0, 1]$, then we would take either the left or right boundary, whichever is closer to 0 or 1, respectively. ■

ACKNOWLEDGMENT

The authors would like to thank the reviewers and to acknowledge William Luh from the ECE Department at Texas A&M University for his insights into robust detectors.

REFERENCES

[1] F. Fu and M. V. D. Schaar, "Noncollaborative resource management for wireless multimedia applications using mechanism design," *IEEE Trans. Multimedia*, vol. 9, no. 4, pp. 851–868, Jun. 2007.

[2] Z. He and D. Wu, "Resource allocation and performance analysis of wireless video sensors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 5, pp. 590–599, May 2006.

[3] I. Akyildiz and I. Kasimoglu, "Wireless sensor and actor networks: Research challenges," *Ad Hoc Netw. J.*, vol. 2, no. 4, pp. 3351–3677, 2004.

[4] H. Ma and Y. Liu, "Correlation based video processing in video sensor networks," in *Proc. IEEE Int. Conf. Wireless Networks, Communications and Mobile Computing*, Maui, HI, Jun. 2005, pp. 987–992.

[5] I. Akyildiz, T. Melodia, and K. Chowdhury, "A survey on wireless multimedia sensor networks," *Comput. Netw.*, vol. 51, no. 4, pp. 921–960, Mar. 2007.

[6] K. Veeraraghavan, D. Peng, and H. Sharif, "Energy efficient multi-resolution visual surveillance on wireless sensor networks," in *Proc. IEEE Int. Conf. Electro Information Technology*, Lincoln, NE, May 22–25, 2005.

[7] D. Kundur, W. Luh, and U. N. Okorafor, "Emerging security paradigms for distributed multimedia sensor networks," *Proc. IEEE Special Issue on Distributed Multimedia*, 2007, to appear.

[8] D. Maniezzo, K. Yao, and G. Mazzini, "Energetic trade-off between computing and communication resource in multimedia surveillance sensor network," in *Proc. 4th IEEE Conf. Mobile and Wireless Communications Networks (MWCN)*, Stockholm, Sweden, Sep. 2002, pp. 373–376.

[9] C. Margi, V. Petkov, K. Obraczka, and R. Manduchi, "Characterizing energy consumption in a visual sensor network testbed," in *Proc. 2nd Int. Conf. Testbeds and Research Infrastructures for the Development of Networks and Communities*, Barcelona, Spain, Mar. 2006.

[10] A. Basharat, N. Catbas, and M. Shah, "A framework for intelligent sensor network with video camera for structural health monitoring of bridges," in *Proc. 3rd IEEE Int. Conf. Pervasive Computing and Communications Workshops (PerCom 2005)*, Mar. 2005, pp. 385–389.

[11] A. Czarlinska and D. Kundur, "Attack vs. failure detection in event-driven wireless visual sensor networks," in *ACM Multimedia & Security Workshop (MM&Sec'07)*, Dallas, TX, Sep. 20–21, 2007 [Online]. Available: www.ece.tamu.edu/-czlinska/ACM07AVVSN.pdf

[12] T. Aach and A. Kaup, "Statistical model-based change detection in moving video," *Signal Process.*, vol. 31, pp. 165–180, Mar. 1993.

[13] L. Hongliang, L. Guizhong, Z. Zhongwei, and L. Yongli, "Adaptive scene-detection algorithm for vbr video stream," *IEEE Trans. Multimedia*, vol. 6, no. 4, pp. 624–633, Aug. 2004.

[14] L. Buttyan and J.-P. Hubaux, "Report on a working session on security in wireless ad hoc networks," *Mobile Comput. and Commun. Rev.*, vol. 6, no. 4, pp. 1–17, 2002.

[15] A. Czarlinska, W. Luh, and D. Kundur, "Attacks on sensing in hostile wireless sensor-actuator environments," *Proc. IEEE Globecom* Nov. 26–30, 2007 [Online]. Available: www.ece.tamu.edurczlinska/Glob07/AttacksSenAct.pdf, Washington, DC

[16] Y. Mao and M. Wu, "A joint signal processing and cryptographic approach to multimedia encryption," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 2061–2075, Jul. 2006.

[17] A. Czarlinska and D. Kundur, "Distributed actuation attacks in wireless sensor networks: Implications and countermeasures," in *Proc. 2nd IEEE Workshop on Dependability and Security in Sensor Networks and Systems*, Apr. 2006, pp. 3–12.

[18] M. Osborne and A. Rubinstein*, A Course in Game Theory*. Cambridge, MA: MIT Press, 1994.

[19] V. Rodriguez, "Resource management for scalably encoded information: The case of image transmission over wireless networks," in *Proc. IEEE Int. Conf. Multimedia and Expo (ICME 2003)*, Baltimore, MD, Jul. 6–9, 2003, pp. 813–816.

[20] R. Joyce and B. Liu, "Temporal segmentation of video using frame and histogram space," *IEEE Trans. Multimedia*, vol. 8, no. 1, pp. 130–140, Feb. 2006.

[21] R. Ford, C. Robson, D. Temple, and M. Gerlach, "Metrics for scene change detection in digital video sequences," in *Proc. IEEE Int. Conf. Multimedia Computing and Systems*, Ottawa, ON, Canada, Jun. 3–6, 1997, pp. 610–611.

[22] D. Lelescu and D. Schonfeld, "Statistical sequential analysis for real-time video scene change detection on compressed multimedia bitstream," *IEEE Trans. Multimedia*, vol. 5, no. 1, pp. 106–117, Mar. 2003.

[23] P. L. Rosin, "Thresholding for change detection," *Comput. Vis. and Image Understand.*, vol. 86, no. 2, pp. 79–95, May 2002.

[24] P. Rosin and E. Loannidis, "Evaluation of global image thresholding for change detection," *Pattern Recognit. Lett.*, vol. 24, no. 14, pp. 2345–2356, Oct. 2003.

[25] K.-Y. Chow, K.-S. Lui, and E. Lam, "Efficient on-demand image transmission in visual sensor networks," *EURASIP J. Advances in Signal Process.*, Article ID 95076 2007.

[26] R. Radke, S. A. O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.

[27] K.-Y. Chow, K.-S. Lui, and E. Lam, "Balancing image quality and energy consumption in visual sensor networks," in *Proc. IEEE Int. Symp. Wireless Pervasive Computing*, Phuket, Thailand, Jan. 16–18, 2006.

[28] R. Ott and M. Longnecker*, An Introduction to Statistical Methods & Data Analysis*. Pacific Grove, CA: Duxbury Press, 2001.

[29] T. Aach and A. Kaup, "Bayesian algorithms for adaptive change detection in image sequences using markov random fields," *Signal Process.: Image Commun.*, vol. 1, no. 2, pp. 147–160, Aug. 1995.

[30] H. L. Van Trees*, Detection, Estimation, and Modulation Theory Part I*. New York: Wiley, 2001.

[31] K. Liu and A. Sayeed, "Asymptotically optimal decentralized type-based detection in wireless sensor networks," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Montreal, QC, Canada, May 2004, pp. 873–876.

[32] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, Aug. 2000.

[33] K. V. Mardia and I. L. Dryden, "The statistical analysis of shape data," *Biometrika*, vol. 76, pp. 271–281, 1989.

Alexandra Czarlinska (S'05) received the B.A.Sc. degree in engineering science (electrical option) in 2002 from the University of Toronto, Toronto, ON, Canada. She is currently pursuing the Ph.D. degree in the Wireless Communications Group (WCL), Department of Electrical and Computer Engineering, Texas A&M University, College Station.

Her current research focuses on the identification and prevention of new security attacks in mobile multimedia wireless sensor and actuator networks and on applications of game theory.

Ms. Czarlinska was the recipient of the National Scholarship Award from the University of Toronto.

Deepa Kundur (SM'03) received the B.A.Sc., M.A.Sc., and Ph.D. degrees in electrical and computer engineering in 1993, 1995, and 1999, respectively, from the University of Toronto, Toronto, ON, Canada.

In January 2003, she joined the Department of Electrical and Computer Engineering, Texas A&M University, College Station, where she is currently an Associate Professor and leads the Sensor Media Algorithms and Networking for Trusted Intelligent Computing (SeMANTIC) Research Group of the Wireless Communications Laboratory. Her research interests include security and privacy for scalar and broadband sensor networks, multimedia security, digital rights management, and steganalysis for computer forensics.

She has given tutorials in the area of information security at ICME-2003 and Globecom-2003, and was a Guest Editor of the June 2004 Proceedings of the IEEE Special Issue on Enabling Security Technologies for Digital Rights Management. She currently serves as the Vice Chair for the Security Interest Group of the IEEE Multimedia Communications Technical Committee and is an Associate Editor for the IEEE TRANSACTIONS ON MULTIMEDIA, IEEE COMMUNICATION LETTERS, and the EURASIP *Journal on Information Security*.