Optimal Index Policies for Quickest Localization of Anomaly in Cyber Networks

Kobi Cohen[1], Qing Zhao[1], Ananthram Swami[2]
[1] Department of Electrical and Computer Engineering, University of California, Davis, CA 95616
{yscohen, qzhao}@ucdavis.edu
[2] Army Research Laboratory, Adelphi, MD 20783
a.swami@ieee.org

***Abstract*—We consider the problem of quickest localization of anomaly in a resource-constrained cyber network consisting of multiple components. Due to resource constraints, only one component can be probed at each time. The observations are random realizations drawn from two different distributions depending on whether the component is normal or anomalous. Components are assigned priorities. Components with higher priorities in an abnormal state should be fixed before components with lower priorities to reduce the overall damage to the network. The objective is to minimize the expected weighted sum of completion times of abnormal components subject to error probability constraints. We consider two different anomaly models: the independent model in which each component can be abnormal independent of other components, and the exclusive model in which there is one and only one abnormal component. We develop index policies under both models. Optimal low-complexity algorithms are derived for the simple hypotheses case, where the distribution is completely known under both hypotheses. Asymptotically (as the error probability approaches zero) optimal low-complexity algorithms are derived for the composite hypotheses case, where there is uncertainty in the distribution parameters. Simulation results then illustrate the performance of the algorithms.**

***Index Terms*—** Anomaly detection, intrusion detection, sequential hypothesis testing.

I. Introduction

An intrusion detection system (IDS) is a system that monitors the network to detect malicious activities (i.e., attacks) in the network. Once an IDS determines that a malicious activity has occurred, it then alerts the security administrator or initiates a proper response to the malicious activity. Good surveys of existing techniques for IDSs can be found in [1]–[14].

In this paper we address the problem of quickest localization of anomaly in a resource-constrained cyber network. We consider a network with K heterogeneous components which can be paths, routers, etc. Assume that an intrusion has been detected (by probing a subnet, for instance [14]). The goal here is to locate the infected components as quickly and as reliably as possible. We focus on a resource-constrained intrusion detection in cyber networks, as was done in [3], [14]. Due to resource constraints, only one component can be probed at each time. The completion time of component k is defined as the time where the IDS completes testing component k. Components are assigned priorities. Components with higher priorities in an abnormal state should be fixed before those with lower priorities to reduce the overall damage to the network.

Throughout this paper we use the theory of sequential detection. In sequential tests, after each observation has been collected, the detector decides whether to accept H_0, reject H_0 or to take another observation. The sample size achieved by sequential tests can be significantly reduced as compared to fixed-size tests. Therefore, it is a natural approach for quickest localization of anomaly. Sequential detection has been extensively studied in the literature. Sequential detection using ordered transmissions was introduced in [15]. In cases where the measurements can be collected sequentially according to a specific order, the number of measurements required for optimal detection can be significantly reduced [15]. Related works on this subject can be found in [15]–[18]. However, this is not the case in the IDS model. Change-point detection techniques can be applied to the problem of anomaly detection to identify a change in the probability distribution when a malicious activity occurs. Related works on this subject can be found in [7]–[9]. However, in this paper we consider a different problem. In our model, an intrusion has been detected. The goal here is to locate the infected components and not a change point. The observations are drawn from two different distributions depending on whether the component is normal or anomalous. The problem of sequentially testing the simple null hypothesis H_0 versus the simple alternative hypothesis H_1 was solved in [19]. It was shown that the Sequential Probability Ratio Test (SPRT) minimizes the expected sample size under given type I and type II error probability constraints. Related works on SPRT-based solutions for anomaly detection can be found in [2], [4], [5], [12], [13]. Various problems of sequentially testing the composite null hypothesis H_0 versus the composite alternative hypothesis H_1 were studied in [20]–[22]. In this case, asymptotically optimal performance can be obtained as the error probability approaches zero.

In the following, we summarize the main results of this paper. We formulate the anomaly localization problem as a constrained optimization problem. The objective is to minimize the expected weighted sum of completion times of abnormal components (since normal components do not cause damage to the network) subject to error probability constraints. We consider both independent and exclusive models. In the former, each component is abnormal, with some prior probability, independent of other components. Under the exclusive model, one and only one component is abnormal with some prior probability (which is a reasonable model when the probability of each component to be compromised is small). We develop index policies under both models. Optimal algorithms are derived for the simple hypotheses case, where the distribution is completely known under both hypotheses. Asymptotically (as the error probability approaches zero) optimal algorithms are derived for the composite hypotheses case, where there is uncertainty in the distribution parameters. In all cases, the algorithms have low-complexity.

II. Network Model and Problem Formulation

Consider a cyber network consisting of K components. Assume that an intrusion has been detected. The goal here is to locate

This work was supported by Army Research Lab under Grant W911NF1120086.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **DEC 2013**	2. REPORT TYPE	3. DATES COVERED **00-00-2013 to 00-00-2013**

4. TITLE AND SUBTITLE **Optimal Index Policies for Quickest Localization of Anomaly in Cyber Networks**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of California, Davis,Department of Electrical and Computer Engineering,Davis,CA,95616**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

in Proc. of IEEE Global Conference on Signal and Information Processing (GlobalSIP), 3-5 Dec 2013, Austin, TX.

14. ABSTRACT

We consider the problem of quickest localization of anomaly in a resource-constrained cyber network consisting of multiple components. Due to resource constraints, only one component can be probed at each time. The observations are random realizations drawn from two different distributions depending on whether the component is normal or anomalous. Components are assigned priorities. Components with higher priorities in an abnormal state should be fixed before components with lower priorities to reduce the overall damage to the network. The objective is to minimize the expected weighted sum of completion times of abnormal components subject to error probability constraints. We consider two different anomaly models: the independent model in which each component can be abnormal independent of other components, and the exclusive model in which there is one and only one abnormal component. We develop index policies under both models. Optimal low-complexity algorithms are derived for the simple hypotheses case, where the distribution is completely known under both hypotheses. Asymptotically (as the error probability approaches zero) optimal lowcomplexity algorithms are derived for the composite hypotheses case where there is uncertainty in the distribution parameters. Simulation results then illustrate the performance of the algorithms.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**4**	

the infected components. Due to resource constraints, only one component can be probed at each time. When component k is tested, a sequence of i.i.d. measurements $\{y_k(i)\}_{i\geq 1}$ is drawn in a one-at-a-time manner. If component k is in a healthy state, $\{y_k(i)\}_{i\geq 1}$ are drawn from distribution $f_k^{(0)}$; if component k is abnormal, $\{y_k(i)\}_{i\geq 1}$ are drawn from distribution $f_k^{(1)}$. Components are assigned priorities. Let w_k ($0 \leq w_k < \infty$) be the priority (or weight) of component k. Components with higher priorities in an abnormal state should be fixed before those with lower priorities to reduce the overall damage to the network.

We consider the case where the switching cost is high. Thus, switching between components is done only when testing the current component is completed. The advantages of this scheme are twofold. First, switching between components typically adds significant delay that should be avoided. Second, the IDS is required to store observations of only one component at each time. Thus, this scheme is applicable to limited-memory systems. For convenience, we define t_k as the time where the IDS has completed the $(k-1)^{th}$ test and starts the k^{th} test. After each observation has been collected, the IDS needs to decide whether to take more measurements from the current component or finalize the test on the current component by declaring its state (healthy or abnormal) and choose the next component to test. Let N_k be the random sample size required to make a decision regarding the state of component k. The random completion time C_k is defined as the time where the IDS completes testing component k. For example, if the IDS tests component 1 followed by component 2, then $C_1 = N_1$ and $C_2 = N_1 + N_2$.

Let τ_k be a stopping rule, which the IDS uses to decide whether to take more measurements from component k or to finalize the test by declaring its state. The vector of stopping rules for the K components is denoted by $\boldsymbol{\tau} = (\tau_1, ..., \tau_k)$. Let $\delta_k \in \{0, 1\}$ be a decision rule, where $\delta_k = 0$ if the IDS declares that component k is in a healthy state (i.e., H_0), and $\delta_k = 1$ if the IDS declares that component k is in an abnormal state (i.e., H_1). The vector of decision rules for the K components is denoted by $\boldsymbol{\delta} = (\delta_1, ..., \delta_K)$. Let $\phi(t_k) \in \{1, 2, ..., K\}$ be a selection rule, indicates which component is chosen to be tested at time t_k. The vector of selection rules for the K components is denoted by $\boldsymbol{\phi} = (\phi(t_1), ..., \phi(t_K))$. Finally, the set of all the abnormal components is denoted by $\mathcal{H}_1 = \{k : 1 \leq k \leq K \text{ , component } k \text{ is abnormal}\}$

The problem is to find a selection rule $\boldsymbol{\phi}$, a stopping rule $\boldsymbol{\tau}$ and a decision rule $\boldsymbol{\delta}$ that minimize the expected weighted sum of completion times of all the abnormal components subject to error probability constraints for each component:

$$\begin{aligned} &\inf_{\boldsymbol{\tau},\boldsymbol{\delta},\boldsymbol{\phi}} \quad \mathbf{E}\left\{\sum_{k\in\mathcal{H}_1} w_k C_k\right\} \\ &s.t. \quad P_k^{FA} \leq \alpha_k \qquad \forall k = 1, ..., K \ , \\ &\qquad\;\; P_k^{MD} \leq \beta_k \qquad \forall k = 1, ..., K \ , \end{aligned} \tag{1}$$

where P_k^{FA}, P_k^{MD} are the false-alarm and miss-detection probabilities at component k, respectively.
Higher penalties are assigned to higher-priority components in an abnormal state[1]. No penalty is associated with components in a healthy state since they do not cause damage to the network.

III. Two-stage Optimization Problem

Instead of solving (1) directly, we propose a two-stage optimization problem. At the first stage, the problem is to find a stopping rule τ_k and a decision rule δ_k for every component k that minimize the expected sample size given H_i subject to error probability constraints:

$$\begin{aligned} &\inf_{\tau_k,\delta_k} E(N_k|H_i) \ , \quad i = 0, 1 \\ &s.t. \quad P_k^{FA} \leq \alpha_k \ , \\ &\qquad\quad P_k^{MD} \leq \beta_k \ . \end{aligned} \tag{2}$$

For the simple hypotheses case, the solution to the first-stage optimization problem (2) is given by the SPRT [19]. Let $L_k(n) = \prod_{i=1}^{n} f_k^{(1)}(y_k(i)) / \prod_{i=1}^{n} f_k^{(0)}(y_k(i))$ be the Likelihood Ratio (LR) between the two hypotheses of component k at stage n and let A_k, B_k ($B_k > A_k$) be the boundary values used by the SPRT when testing component k, such that the error constraints are satisfied. According to the SPRT algorithm, at each stage n, the LR is compared to the boundary values as follows. If $L_k(n) \in (A_k, B_k)$, the IDS continues to take observations from component k. If $L_k(n) \geq B_k$, the IDS finalizes the test on component k and declares it as abnormal (i.e., H_1). If $L_k(n) \leq A_k$, the IDS finalizes the test on component k and declares it as normal (i.e., H_0). In general, the exact determination of the boundary values is very laborious. However, Wald's approximation to the boundary values can be applied to simplify the computation: $B_k \approx (1 - \beta_k)/\alpha_k, A_k \approx \beta_k/(1 - \alpha_k)$. Wald's approximation performs well for small α_k, β_k. Since type I and type II errors are typically small, Wald's approximation is widely used in practice [19].

At the second stage, the problem is to find a selection rule $\boldsymbol{\phi}$ that minimizes the objective function, given the solution to the K subproblems (2):

$$\begin{aligned} &\inf_{\boldsymbol{\phi}} \quad \mathbf{E}\left\{\sum_{k\in\mathcal{H}_1} w_k C_k\right\} \\ &s.t. \quad \text{solutions to (2) are given for } k = 1, ..., K \ . \end{aligned} \tag{3}$$

The solutions to the second-stage optimization problem for the independent and exclusive models are given in Section IV.

The formulation of the two-stage optimization problem allows us to decompose the original optimization problem (1) into $K+1$ subproblems (2) and (3). We use this formulation to design the solution to (1). In subsequent sections we show that the solution to the two-stage optimization problem solves the original optimization problem (1) under both the independent and exclusive models.

IV. Localization of Anomaly for the Simple Hypotheses Case

In this section we derive optimal solutions to both the independent and exclusive models under the simple hypotheses case, where the distribution under both hypotheses is completely known. Under the independent model, each component k is abnormal with a priori probability π_k independent of other components. Under the exclusive model, one and only one component is abnormal with a priori probability π_k, where $\sum_{k=1}^{K} \pi_k = 1$.

Based on the solution to the two-stage optimization problem, we propose Algorithms $1, 2$, presented in Tables I, II, to solve (1). It was shown in [23] that the optimal selection rule for the problem of minimizing the expected weighted sum of completion times given the expected testing time of each component is to select the components in decreasing order of $w_k/E(N_k)$. However, the problem in (3) is different. First, the objective is to minimize the expected weighted

[1] Note that the loss due to missed-detection events is negligible for small error probability, since $P_k^{MD} \in O(1/B_k)$ and $E(N_k) \in \Theta(\log B_k)$, where B_k is a boundary value of the sequential test [19], [20].

TABLE I
ALGORITHM 1 FOR THE INDEPENDENT MODEL

```
1.  arrange the components in decreasing
    order of $\pi_k w_k / E(N_k)$
2.  for $k = 1, ..., K$ components do:
3.      perform SPRT for component $k$,
        with $P_k^{FA} \leq \alpha_k$, $P_k^{MD} \leq \beta_k$
4.  end for
```

sum of completion times of abnormal components only. Second, the expected sample size depends on the component state. Furthermore, under the exclusive model, the state of each component depends on other components. Here, we derive optimal selection rules that solve the second-stage optimization problem (3) for the independent and exclusive models. These selection rules are given in step 1 in Tables I, II for the independent and exclusive models, respectively. Arranging the components in decreasing order of $\pi_k w_k / E(N_k)$ or $\pi_k w_k / E(N_k | H_0)$ in step 1 can be done in $O(K \log K)$ time via sorting algorithms. Next, by the optimal solution to (2), a series of SPRTs is performed according to this order until all the components are tested.

The index policies, described in Algorithms $1, 2$, are intuitively satisfying. The priority of component k in terms of testing order should be higher as the weight w_k increases, or the a priori probability to be abnormal π_k increases. Under the independent model, the priority of component k in terms of testing order should be higher as the expected sample size $E(N_k)$ decreases (since $E(N_k)$ is added to the completion time of every component which is tested after component k). On the other hand, under the exclusive model, the priority of component k in terms of testing order should be higher as $E(N_k | H_0)$ decreases. Note that under the exclusive model, we take into account the expected sample size under H_0 solely. The reason is that if component k is abnormal, there is no penalty to other components under the exclusive model (since only one component is abnormal). On the other hand, if component k is healthy, then $E(N_k | H_0)$ is added to the completion time of the components which are tested after component k (and may be abnormal). The SPRT is used in both models to minimize the expected sample size to reduce the completion times.

Note that the solution to the second-stage optimization problem (3) requires one to compute the expected sample size $E(N_k | H_i)$ for all $k = 1, 2, ..., K$, and for $i = 0, 1$ to select the components in decreasing order of $\pi_k w_k / E(N_k)$ or $\pi_k w_k / E(N_k | H_0)$. In general, it is difficult to obtain a closed-form expression for $E(N_k | H_i)$. However, Wald's approximation to the expected sample size can be applied to simplify the computation [19]. The approximation approaches the exact expected sample size for small α_k, β_k. Since type I and type II errors are typically small, Wald's approximation is widely used in practice [19]. For more details, the reader is referred to [24].

Theorem 1: Under the independent and exclusive models, Algorithms $1, 2$, *respectively, solve (1).*

The proof is given in the extended version of this paper [24].

TABLE II
ALGORITHM 2 FOR THE EXCLUSIVE MODEL

```
1.  arrange the components in decreasing
    order of $\pi_k w_k / E(N_k | H_0)$
2.  for $k = 1, ..., K$ components do:
3.      perform SPRT for component $k$,
        with $P_k^{FA} \leq \alpha_k$, $P_k^{MD} \leq \beta_k$
4.  end for
```

Note that Algorithms $1, 2$ use static selection rules (as stated in step 1), where the components order is predetermined at time t_1. However, Theorem 1 is not restricted to static selection rules. Theorem 1 shows that Algorithms $1, 2$ are optimal among the class of both static and dynamic selection rules (that update the selection dynamically at each time t_k).

V. Localization of Anomaly Under Uncertainty

In numerous cases under the adversary model, there is uncertainty in the distribution parameters (in particular when the component is in an abnormal state). Hence, in this section we discuss the extension of our results to the problem of localization of anomaly under uncertainty. Due to space limitation, we provide a general structure of the solution for this case. For more details, the reader is referred to [24]. Under uncertainty, asymptotically optimal sequential tests in terms of minimizing the expected sample size as the error probability approaches zero can be applied for testing every component [20]–[22]. Therefore, we modify step 3 in Algorithms $1, 2$, given in Tables I, II by performing an appropriate sequential test for composite hypotheses instead of the SPRT. In [24], we use the asymptotically optimal property of the sequential test, studied for a single process, to show that the modified algorithms asymptotically solve the original optimization problem (1) as the error probability approaches zero.

VI. Numerical Examples

In this section, we provide numerical examples to illustrate the performance of the algorithms. We compared three schemes: a Random selection SPRT (R-SPRT), where a series of SPRTs are performed until all the components are tested in a random order, and the proposed Algorithms $1, 2$, which are optimal for the independent and exclusive models, respectively.

We consider an intruder that tries to launch a Denial of Service (DoS) or Reduction of Quality (RoQ) attacks by sending a large number of packets to a component (which can be a relay node in this application). DoS attacks rely on overwhelming the component with useless traffic that exceeds its capacity so to make it unavailable for its intended use. On the other hand, RoQ attacks inflict damage on the component, while keeping a low profile to avoid detection. RoQ attacks do not cause denial of service. In order to detect such attacks, the IDS performs a traffic-based anomaly detection. It monitors the traffic at each component to decide whether a component is compromised. Similar traffic-based detection techniques were proposed in [6], [11] for different models, considering a single process without switching to other components. For each component k, we assume that packets arrive according to a Poisson process with rate $\theta^{(k)}$. When component k is tested, the IDS collects an observation $y_k(n) \in \mathbb{N}_0$ every time unit, which is the number of packets that arrived in the interval $(n-1, n)$. We consider the case where $\theta_k = \theta_k^{(0)}$ under normal state and $\theta_k = \theta_k^{(1)} = 1.5 \cdot \theta_k^{(0)}$ under abnormal state are known to the IDS. Let $\delta_K = (100-10)/(K-1)$. We set $w_k = \theta_k^{(0)} = 10 + (k-1)\delta_K$. By setting $w_k = \theta_k^{(0)}$, the objective function represents the total expected number of failed packets in the network during DoS attacks. Thus, the optimization problem can be observed as minimizing the maximal damage to the network in terms of packet-loss. Furthermore, this setting prioritizes components with higher normal traffic to reduce the delay caused by RoQ attacks. The error constraints were set to $P_k^{FA} = P_k^{MD} = 10^{-2}$ for all k. For the independent and exclusive models, we set $\pi_k = 0.8$ and $\pi_k = 1/K$ for all k, respectively. The performance of Algorithm 1 and Algorithm 2 are presented in Fig. 1, as compared to the R-SPRT. It can be seen that the proposed Algorithms save roughly 50% of the objective value as compared to the R-SPRT under both the independent and exclusive model scenarios.

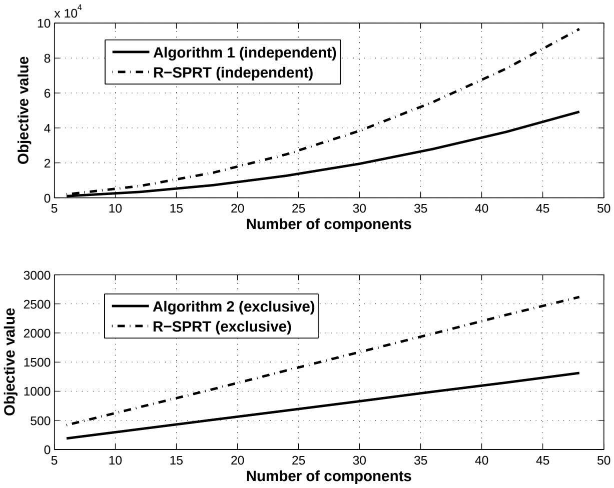

Fig. 1. Objective value as a function of the number of components for the independent and exclusive model scenarios.

VII. Conclusion

The problem of quickest localization of anomaly in a resource-constrained cyber network was investigated. Due to resource constraints, only one component can be probed at each time. The observations are random realizations drawn from two different distributions depending on whether the component is normal or anomalous. Components with higher priorities in an abnormal state should be fixed before components with lower priorities to reduce the overall damage to the network. The objective is to minimize the expected weighted sum of completion times subject to error probability constraints. For the simple hypotheses case, we derived optimal algorithms for both the independent and exclusive models. For the composite hypotheses case, we derived asymptotically (as the error probability approaches zero) optimal algorithms for both the independent and exclusive models. These optimal algorithms have low-complexity.

References

[1] T. F. Lunt, "A survey of intrusion detection techniques," *Computers & Security*, vol. 12, no. 4, pp. 405–418, 1993.

[2] K. C. Gross and W. Lu, "Early detection of signal and process anomalies in enterprise computing systems," in *Proc. of IEEE International Conference on Machine Learning and Applications*, 2002.

[3] A. Agah, S. K. Das, K. Basu, and M. Asadi, "Intrusion detection in sensor networks: A non-cooperative game approach," in *Proc. IEEE International Symposium on Network Computing and Applications (NCA)*, pp. 343–346, 2004.

[4] J. Jung, V. Paxson, A. W. Berger, and H. Balakrishnan, "Fast portscan detection using sequential hypothesis testing," in *Proceedings IEEE Symposium on Security and Privacy*, pp. 211–225, 2004.

[5] J. Jung, S. E. Schechter, and A. W. Berger, "Fast detection of scanning worm infections," in *Proceeding Inter. Symposium on Recent Advances in Intrusion Detection (RAID)*, pp. 59–81, 2004.

[6] I. Onat and A. Miri, "An intrusion detection system for wireless sensor networks," in *IEEE International Conference on Wireless And Mobile Computing, Networking And Communications*, vol. 3, pp. 253–259, 2005.

[7] A. G. Tartakovsky, B. L. Rozovskii, R. B. Blazek, and H. Kim, "A novel approach to detection of intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods," *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 3372–3382, 2006.

[8] G. Androulidakis, V. Chatzigiannakis, S. Papavassiliou, M. Grammatikou, and V. Maglaris, "Understanding and evaluating the impact of sampling on anomaly detection techniques," in *IEEE Military Communications Conference (MILCOM)*, pp. 1–7, 2006.

[9] T. Van Phuong, L. Hung, S. Cho, Y. K. Lee, and S. Lee, "An anomaly detection algorithm for detecting attacks in wireless sensor networks," *Intelligence and Security Informatics*, pp. 735–736, 2006.

[10] T. He and L. Tong, "Detecting encrypted stepping-stone connections," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1612–1623, 2007.

[11] V. B. Misic and J. Begum, "Evaluating the feasibility of traffic-based intrusion detection in an 802.15.4 sensor cluster," in *IEEE International Conference on Advanced Information Networking and Applications*, pp. 619–624, 2007.

[12] A. A. Cárdenas, S. Radosavac, and J. S. Baras, "Evaluation of detection algorithms for mac layer misbehavior: theory and experiments," *IEEE/ACM Transactions on Networking*, vol. 17, no. 2, pp. 605–617, 2009.

[13] G. Thatte, U. Mitra, and J. Heidemann, "Parametric methods for anomaly detection in aggregate traffic," *IEEE/ACM Transactions on Networking*, vol. 19, no. 2, pp. 512–525, 2011.

[14] K. Liu and Q. Zhao, "Intrusion detection in resource-constrained cyber networks: A restless multi-armed bandit approach," *submitted to IEEE/ACM Transactions on Networking. Avialable at http://arxiv.org/abs/1112.0101*.

[15] R. S. Blum and B. M. Sadler, "Energy efficient signal detection in sensor networks using ordered transmissions," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3229–3235, 2008.

[16] K. Cohen and A. Leshem, "Energy-efficient detection in wireless sensor networks using likelihood ratio and channel state information," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1671–1683, 2011.

[17] Y. R. Tsai and L. C. Lin, "Sequential fusion for distributed detection over BSC channels in an inhomogeneous sensing environment," *IEEE Signal Processing Letters*, vol. 17, no. 1, pp. 99–102, 2010.

[18] P. Braca, S. Marano, and V. Matta, "Single-transmission distributed detection via order statistics," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 2042–2048, 2012.

[19] A. Wald, "Sequential analysis," *New York: Wiley*, 1947.

[20] T. L. Lai, "Nearly optimal sequential tests of composite hypotheses," *The Annals of Statistics*, pp. 856–886, 1988.

[21] I. V. Pavlov, "Sequential procedure of testing composite hypotheses with applications to the Kiefer-Weiss problem," *Theory of Probability and Its Applications*, vol. 35, no. 2, pp. 280–292, 1990.

[22] A. G. Tartakovsky, "An efficient adaptive sequential procedure for detecting targets," in *IEEE Aerospace Conference Proceedings, 2002*, vol. 4, pp. 1581–1596, 2002.

[23] W. E. Smith, "Various optimizers for single-stage production," *Naval Research Logistics Quarterly*, vol. 3, no. 1-2, pp. 59–66, 1956.

[24] K. Cohen, Q. Zhao, and A. Swami, "Optimal index policies for anomaly localization in resource-constrained systems," *technical report TR-13-01, available at http://www.ece.ucdavis.edu/ qzhao/Report.html*, 2013.