Bayesian inference based decision reliability under imperfect monitoring

Shameek Bhattacharjee and Mainak Chatterjee
Electrical Engineering & Computer Science
University of Central Florida
Orlando, Florida 32816
Email: {shameek, mainak}@eecs.ucf.edu

Kevin Kwiat and Charles Kamhoua
Air Force Research Laboratory
Information Directorate
Rome, NY 13441
Email: {kevin.kwiat,charles.kamhoua.1}@us.af.mil

***Abstract*—Reliability of a cooperative decision mechanism is critical for the proper and accurate functioning of a networked decision system. However, adversaries may choose to compromise the inputs from different sets of components that comprise the system. Often times, the monitoring mechanisms fail to accurately detect compromised inputs; hence cannot categorize all inputs into polarized decisions: compromised or not compromised.**

In this paper, we propose a Bayesian inference model based on multinomial evidence to quantify reliability for a cooperative decision process as a function of beliefs associated with observations from the imperfect monitoring mechanism. We propose two reliability models: an optimistic one for a normal system and a conservative one for a mission critical system. We also provide an entropy measure that reflects the certainty or uncertainty on the calculated reliability of the decision process. Through simulation, we show how the reliability and its corresponding entropy changes as the accuracy of the underlying monitoring mechanism improves. [1]

I. Introduction and Motivation

Oftentimes, reliability of a cooperative decision in a networked system depends on how well the individual components perform and how reliable they are [1]. The individual components could be a piece of hardware or software, or a link connecting two devices– all of which work together enabling the system as a whole to perform its tasks. To improve reliability of a decision making process in the presence of possible component failures, redundancy and voting schemes are often used [6]. Usually voting schemes are associated with a fusion rule (e.g., majority or plurality voting) which dictates how individual votes are combined to decide the final output.

A malfunction or a security breach due to an adversary might result in a faulty input (vote) to the central decision making entity. Such inputs may potentially have an adverse effect on the reliability of the decision process– the extent of which depends on the inputs in question. A failure monitoring mechanism is supposed to detect such faults and provide a 'feedback' on each component. However, this monitoring mechanism might not have the ability to identify such faults with certainty due to inherent imperfect temporal and spatial factors. Hence a binary decision on whether a fault occurred or not is not always possible for all inputs. With the adversary employing a wide variety of attacks to compromise a *dynamic set* of inputs (i.e., deciding the number of inputs and which ones), and imperfect monitoring conditions, it becomes difficult for a failure monitoring mechanism to quantify reliability of a collective decision based on inputs from individual components.

In the presence of adversaries i.e., intentional cyber-attacks, the compromised inputs may induce incorrect results in a cooperative decision making process. The set of inputs attacked usually vary over time. The damage is evident when some central entity of the system using some fusion rule fuses the inputs from all components– compromised or not. A simple example of a cooperative decision process could be a voting system where components vote to produce a binary outcome. If majority of the inputs from the components are compromised, then the simple majority voting rule may produce a wrong result [4].

In this paper, we investigate the quantification of reliability of a cooperative decision process based on the inputs from various components. We assume, each component provides a single input, all of which are prone to attacks. The underlying imperfect failure monitoring mechanism produces varying feedback over time. We consider that the outcome of the monitoring mechanism can be placed into three categories: those we know have *not been compromised*, those we know have *been compromised*, and those which cannot be inferred either way. Given these, we compute the reliability of a decision process in making a decision i.e., how reliable its output is. In this regard, we conceptualize the outcome of monitoring the input over time as a multinomial hypothesis of a Bayesian inference model with three parameters. We build a Bayesian inference based reliability model, where we assign a value to a decision that indicates how reliable the outcome of the decision is. The way reliability is computed also depends on how much risk a system can tolerate. For example, a mission critical system may need to have a more strict reliability model because the associated risks are too high. Therefore, we propose two ways of computing the reliability– first is an optimistic one and the second is a conservative one. The optimistic model could be applied to systems where some tolerance for wrong decisions are allowed. However, for a mission critical system where there is almost no room for erroneous

[1]Approved for Public Release; Distribution Unlimited: 88ABW-2014-4001, Dated 26 Aug 2014.

decisions, the conservative model could be used. To account for the confidence associated with the reliability computation, we propose an entropy based uncertainty value to represent how certain or uncertain the computed reliability is. A lower uncertainty associated with reliability value is an indication of being more confident about that value. We conduct extensive simulation experiments and show how reliability varies under a variety of system factors like attack intensity and inaccurate detection. We observe that with more inputs compromised, the reliability over the fused decision reduces. Low reliability may also be caused by temporal or initial lack of evidence due to uncertainty. However, as time evolves, the system adjusts itself towards more accurate monitoring and the reliability improves.

The rest of the paper is organized as follows. In section II, we present the system model and state all the assumptions. In section III, we provide the Bayesian model for reliability. We present the reliability models in section IV. The simulation model and results are discussed in section V. Conclusions are drawn in the last section.

II. System Model and Assumptions

We consider a time-slotted system comprising N voting components

each of which provides only one input (i.e., the vote) on each time slot. The nature of the decision is generic; it could be as simple as a binary voting or it could be some complex decision metric. A centralized controller fuses all votes from each component through a fusion scheme (e.g., majority or plurality voting rule) to arrive at a *global decision*.

•**Adversarial model:** We assume that all the inputs from each component are exposed to an *adversary* whose goal is to disrupt the voting process at the central controller. The adversary has some predefined attack resources and can choose to attack different sets of inputs over time and also attack varying number of inputs in each time slot. However, it maintains a long term average of the fraction of the inputs it attacks which we call the probability of attack and denote as P_a. For example, $P_a = 0.6$ means that the adversary compromises 60% of the inputs over a large period of time. Hence a single observation (over one time slot) is not sufficient for characterizing the behavior of the adversary.

•**Imperfect component failure monitoring:** We assume that there is a component failure monitoring or anomaly detection mechanism in place that infers whether the input from each component has been compromised or not. Oftentimes, the monitoring mechanism cannot infer anomaly with certainty. Thus, it classifies the inputs into three categories: i) compromised, ii) not compromised, and iii) undecided. All three are a function of environmental parameters that may be dynamic over time. Note, the monitoring occurs over a period of time. Also system transients and noisy environments may increase temporal uncertainty. Therefore, reliability is computed over time– a larger time window of observation allows a more accurate estimation of the actual reliability.

•**Uniformly distributed prior inference:** Since there is no bias over any of the three possible outcomes of the monitoring process, we assume that the initial probabilities of each is equal.

•**Probability of detection:** We define the probability of detection as the percentage of 'components' inputs that can be *accurately* inferred as compromised or not compromised and denote it as P_{det}. Let us further illustrate the meaning of P_{det} using Fig. 1 that shows an input in reality could be either compromised or not compromised. If compromised, it can be inferred as either as compromised with a probability (say a_1) or undecided with probability $1-a_1$. Note, we assume that there is no way a compromised input will be inferred as not compromised. This assumption is because we argue that 'undecided' is inferred in absence of credible evidence else there will be a compromised feedback only when sure. Similarly, if an input is not compromised, it can be inferred as either 'not compromised' with a probability (say b_1) or 'undecided' with probability $1 - b_1$. Again, there is no way a not compromised component will be inferred as compromised. Thus, for the two real cases, detection occurs with probabilities a_1 and b_1. If an input has equal chances of being compromised and not compromised, then $P_{det} = \frac{a_1+b_1}{2}$. Else, a_1 and b_1 will have to be weighted with their corresponding probabilities. For all practical purposes, we consider P_{det} to be at least 0.5.

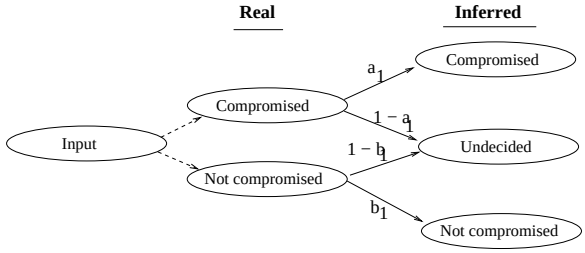

Fig. 1. Inference possibilities for detection probability

The above features make the problem of computing the reliability of the output decision a probabilistic concept. Hence, we compute the reliability as a continuous process based on observations over time slots. If the adversary uses the same attack strategy, then the reliability will converge sooner. On the other hand, if the adversary changes its attack strategy (i.e., dynamic attack strategy), the reliability will oscillate even for large time windows.

III. Reliability Based Decision Making

Suppose that the three feedbacks of the monitoring mechanism– 'not compromised', 'compromised', or 'undecided' be denoted by α, β and μ respectively. Let n_α represent the number of inputs that have 'not' been compromised, n_β represent the number of inputs observed to have been compromised, and n_μ is the number of inputs for which we do not know either way. Of course, $n_\alpha + n_\beta + n_\mu = N$. Since the values of n_α, n_β and n_μ change over time, we represent these observations at time t as $n_\alpha(t)$, $n_\beta(t)$ and $n_\mu(t)$.

Since the system's underlying parameters of cooperative voting behavior are unknown, we propose to use Bayesian inference to update corresponding probability estimate for a hypothesis that the decision process is correct with a certain

TABLE I
NOTATIONS

Symbol	Meaning
α, β, μ	Not compromised, compromised or undecided events
n_α	Number of inputs detected as 'not compromised'
n_β, n_μ	No. of inputs det. as 'compromised' and 'undecided'
N	Total number of voting components
$\theta_\alpha, \theta_\beta, \theta_\mu$	Unknown probability for observing event α, β, μ
$\bar{\theta}$	Bayesian probability parameters of the three events
$X(\bar{\theta})$	Hypothesis of a event
$\hat{X}(\bar{\theta}$	Posterior hypothesis or belief
$D(N)$	Random vector denoting data hyperparameter
P_{det}	Probability of detection
R_α, R_β, R_μ	Posterior Bayesian belief of the three events
R_s^o	Reliability of an optimistic system
R_s^c	Reliability of an conservative system
E_s	Entropy or uncertainty associated with the system

reliability. The system is only as reliable as the individual inputs are. Therefore, we have to calculate the posterior probabilities associated with encountering each of the three feedbacks. The final decision reliability will be a function of these posterior probabilities which are also known as belief estimate in Bayesian inference.

To begin with, an uniform belief over the three possibilities is assumed as there is no initial information. As time progresses, we update the belief estimate based on the observed values of α, β, and μ which increases the accuracy of the estimate of the belief associated with each category.

We define θ_α, θ_β, and θ_μ as the probabilities for an input being 'not compromised, 'compromised', and 'undecided' respectively. Of course, $\theta_\alpha + \theta_\beta + \theta_\mu = 1$, since the outcomes are exhaustive and mutually exclusive. We define $X(\bar{\theta})$ as the hypothesis described by these underlying unknown *Bayesian probability parameters* where $\bar{\theta} = \{\theta_\alpha, \theta_\beta, \theta_\mu\}$.

Let D_α, D_β, and D_μ represent the random variables that represent the number of times the outcomes α, β and μ occur. The observation data can be represented as random observation vector $D(N) = \{D_\alpha, D_\beta, D_\mu\}$ having a multinomial distribution also known as *concentration hyperparameter* of the underlying 3-tuple probability parameter described by θ_α, θ_β, and θ_μ. The commonly used notations are tabulated in Table I.

A. Bayesian Inference

As mentioned earlier, there are N independently monitored components of a system whose parameters for voting behavior are unknown due to changing adversarial attack strategies and the imperfect monitoring mechanism. Given this, we calculate the Bayesian belief associated with 'not compromised'. Similarly, we will model Bayesian posterior belief for the other two cases as well viz. compromised and undecided.

We use the observation counts from the sequential observations over time to calculate the posterior Bayesian estimate of each of the parameters. Our objective is to estimate and update the probability parameters in $X(\bar{\theta})$, viz. θ_α, θ_β, and θ_μ based on observation evidence $D(N)$ and prior information on the hypothesis parameter, $\bar{\theta}$, itself.

Since there is no information about $\bar{\theta}$ initially, we consider the prior parameters of $\bar{\theta}$ to be uniformly distributed a-priori. Subsequent observations will decide how these parameters are updated. Our first step is to calculate the Bayesian estimate of $\bar{\theta}$.

First, we show the case of estimating belief that a 'not compromised' occurs (θ_α). Since in Bayesian inference, the assumption is that prior and posterior probability have the same distribution, we can formally define the probability parameters as:

$$\begin{aligned} P(X(\bar{\theta}) = \alpha|\bar{\theta}) &= \theta_\alpha \\ P(X(\bar{\theta}) = \beta|\bar{\theta}) &= \theta_\beta \\ P(X(\bar{\theta})) = \mu|\bar{\theta}) &= \theta_\mu \end{aligned} \tag{1}$$

This assumption is due to the well known fact that a Dirichlet distribution acts as a conjugate prior to multinomial distributions [7]. Hence prior and posterior preserve the same form.

The observations data $D(N)$ can be treated as a multinomial distribution with probability parameter θ_α, θ_β, and $\theta\mu$, where the probability mass function is given by:

$$\begin{aligned} P(D_\alpha = n_\alpha, D_\beta = n_\beta, D_\mu = n_\mu|\bar{\theta}) &= P(D(N)|\bar{\theta}) \\ &= \frac{N!}{n_\alpha! n_\beta! n_\mu!} \theta_\alpha^{n_\alpha} \theta_\beta^{n_\beta} \theta_\mu^{n_\mu} \end{aligned} \tag{2}$$

Given this we can use Bayes theorem to calculate the posterior belief estimate on the event of a positive interaction $\hat{X}(\bar{\theta}) = \alpha$, given observation data $D(N)$ as:

$$P(\hat{X}(\bar{\theta}) = \alpha|D(N)) = \frac{P(\hat{X}(\bar{\theta}) = \alpha, D(N))}{P(D(N))} \tag{3}$$

Denominator of the above equation is the marginal probability that can be conditioned or marginalized on all possible outcomes for $\bar{\theta}$ and since probabilities are continuous

$$P(D(N)) = \int_{D(N)(\bar{\theta})} P(D(N)|\bar{\theta}) f(\bar{\theta}) d(\bar{\theta}) \tag{4}$$

Since there is no prior information on $\bar{\theta}$ (before any observations) in Eqn. (4), we can assume it to be uniformly distributed such that $f(\bar{\theta}) = 1$ and we can put Eqn. (2) in Eqn. (4), and get

$$P(D(N)) = \frac{N!}{n_\alpha! n_\beta! n_\mu!} \int_{D(N)(\theta_\alpha, \theta_\beta, \theta_\mu)} \theta_\alpha^{n_\alpha} \theta_\beta^{n_\beta} \theta_\mu^{n_\mu} d\theta_\alpha d\theta_\beta d\theta_\mu \tag{5}$$

For simplicity, let $\int_{D(N)(\theta_\alpha, \theta_\beta, \theta_\mu)} \theta_\alpha^{n_\alpha} \theta_\beta^{n_\beta} \theta_\mu^{n_\mu} d\theta_\alpha d\theta_\beta d\theta_\mu = I_1$

To solve for I_1 we use the multivariate generalization of the Eulerian integral of first kind. Note that $D(N)(\theta_\alpha, \theta_\beta, \theta_\mu)$ denotes a space and we know that a space of $m(= 3)$ parameters has only $m - 1(= 2)$ degrees of freedom due to the additivity constraint $\theta_\alpha + \theta_\beta + \theta_\mu = 1$. Therefore when we integrate over this space, the integration has $m - 1 = 2$ dimensions. Hence

$$I_1 = \int_0^1 \int_0^{1-\theta_\alpha-\theta_\beta} \theta_\alpha^{(n_\alpha+1)-1} \theta_\beta^{(n_\beta+1)-1} (1 - \theta_\alpha - \theta_\beta)^{(n_\mu+1)-1} d\theta_\alpha d\theta_\beta \tag{6}$$

Eqn. (6) is a known form for the multivariate extension of the Beta function which in this case is defined as $B(n_\alpha +$

$1, n_\beta + 1, n_\mu + 1)$. The proof can be found in Lemma 2.4.1 of [3]. In general $B(\alpha_1, \cdots, \alpha_m)$

$$= \int_{D(x_1,\cdots,x_{m-1})} x_1^{\alpha_1-1}...(1-\sum_{i=1}^{m-1} x_i)^{\alpha_m-1} dx_1...dx_{m-1}$$
$$= \frac{\prod_{i=1}^{m}\Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{m}\alpha_i)} = \frac{\Gamma(\alpha_1)\cdots\Gamma(\alpha_m)}{\Gamma(\alpha_1+\cdots+\alpha_m)} \quad (7)$$

Using the above result, we can write Eqn. (6) as

$$I_1 = B(n_\alpha+1, n_\beta+1, n_\mu+1) = \frac{\Gamma(n_\alpha+1)\Gamma(n_\beta+1)\Gamma(n_\mu+1)}{\Gamma(n_\alpha+1+n_\beta+1+n_\mu+1)} \quad (8)$$

Putting Eqn. (8) in Eqn. (5) we get:

$$P(D(N)) = \frac{N!}{n_\alpha! n_\beta! n_\mu!} \frac{\Gamma(n_\alpha+1)\Gamma(n_\beta+1)\Gamma(n_\mu+1)}{\Gamma(n_\alpha+1+n_\beta+1+n_\mu+1)} \quad (9)$$

Since the parameters in gamma functions $n_\alpha+1$ etc. are all non-zero positive values, we can use the result $\Gamma(z) = (z-1)!$ to calculate Eqn. (9) as

$$P(D(N)) = \frac{N!}{(N+2)!} \quad (10)$$

Assuming conditional independence between the $\hat{X}(\bar{\theta})$, $D(N)$ and $\bar{\theta}$, we calculate the numerator of Eqn. (3), $P(\hat{X}(\bar{\theta}) = \alpha, D(N))$, as:

$$= \int_{D(N)(\bar{\theta})} P(X(\bar{\theta}) = \alpha, D(N)|\bar{\theta}) f(\bar{\theta}).d(\bar{\theta})$$
$$= \int_{D(N)(\bar{\theta})} P(X(\bar{\theta}) = \alpha|\bar{\theta}) P(D(N)|\bar{\theta}) d(\bar{\theta})$$
$$= \frac{N!}{n_\alpha! n_\beta! n_\mu!} \int_{D(N)(\theta_\alpha,\theta_\beta,\theta_\mu)} \theta_\alpha \theta_\alpha^{n_\alpha} \theta_\beta^{n_\beta} \theta_\mu^{n_\mu} d\theta_\alpha d\theta_\beta d\theta_\mu$$
$$= \frac{N!}{n_\alpha! n_\beta! n_\mu!} \int_{D(N)(\theta_\alpha,\theta_\beta,\theta_\mu)} \theta_\alpha^{n_\alpha+1} \theta_\beta^{n_\beta} \theta_\mu^{n_\mu} d\theta_\alpha d\theta_\beta d\theta_\mu \quad (11)$$

The above integral has the same form as Eqns. (5), (6), and (7). Hence this integral portion of Eqn. (11) can be rewritten as

$$= \int_0^1 \int_0^{1-\theta_\alpha-\theta_\beta} \theta_\alpha^{(n_\alpha+2)-1} \theta_\beta^{(n_\beta+1)-1} (1-\theta_\alpha-\theta_\beta)^{(n_\mu+1)-1} d\theta_\alpha d\theta_\beta$$

which can be solved using Eqn. (7).

Using the above result, Eqn. (11) can be simplified as

$$P(\hat{X}(\bar{\theta}) = \alpha, D(N)) = \frac{N!(n_\alpha+1)}{(N+3)!} \quad (12)$$

Thus, Eqn. (3), can be solved by dividing Eqn. (12) by Eqn. (10), which gives

$$P(\hat{X}(\bar{\theta}) = \alpha|D(N)) = \frac{n_\alpha+1}{N+3} \quad (13)$$

Similarly, $P(\hat{X}(\bar{\theta}) = \beta|D(N)) = \frac{n_\beta+1}{N+3}$ and $P(\hat{X}(\bar{\theta}) = \mu|D(N)) = \frac{n_\mu+1}{N+3}$. These equations are the expressions for posterior belief of 'not compromised', 'compromised', and 'undecided'. To simplify the notations of belief estimates of the three categories, we rewrite them as R_α, R_β, R_μ respectively. Of course, it can be verified that $R_\alpha + R_\beta + R_\mu = 1$. The above equations also satisfy the Cromwell's rule [5], which suggests that no prior belief unless logically impossible, should be assigned zero probability even if no events in that category has occurred so far.

We mentioned that in Bayesian inference posterior probabilities may be used as priors for future calculations. The Cromwell's rule leaves open the probability however small, to experience an event that has not occurred yet but may happen in future. Hence Bayesian estimates should have non-zero priors for an event that has not occurred yet. From the derived equation it is evident that even if $n_\beta = 0$, $R_\beta \neq 0$, is a very small number but not zero when N is large.

IV. Reliability Models

We propose two reliability models and also show how entropy can capture the uncertainty associated with the reliability calculation due to the undecided inputs.

A. Decision Reliability for an Optimistic System

For a system, we assumed that adversary has uniformly chosen the inputs it chooses to attack i.e., there is no reason for preferential attack on a certain component's input. Hence we can account for the undecided R_μ by splitting it in the ratio of $R_\alpha : R_\beta$, and adding it to the value R_α to provide the optimistic reliability denoted by R_s^o. Of course, when the proportion of R_μ is high, we may not be as confident on the reliability value than when we have lower values of R_μ. Thus, R_s^o is computed as:

$$R_s^o = R_\alpha + \frac{R_\alpha}{R_\alpha + R_\beta} R_\mu \quad (14)$$

B. Decision Reliability for a Conservative System

Unlike the optimistic approach, where the undecided ones are split in a ratio, the conservation model treats the undecided ones as if they were compromised. In other words, only the 'not compromised' event (R_α) is used for computing the reliability. Hence,

$$R_s^c = R_\alpha \quad (15)$$

This conservative way of computing the reliability is more appropriate for mission-critical systems where the decisions can only be made based on the 'not compromised' inputs. No risk is taken on the undecided inputs even if there could be some that were not compromised.

C. Uncertainty associated with Reliability

System reliability as computed by Eqn. (14) can yield the same value for different sets of R_α, R_β, and R_μ. For example, consider two scenarios.

Scenario 1: $R_\alpha = R_\beta = 0.5$, and $R_\mu = 0$

Scenario 2: $R_\alpha = R_\beta = 0.3$, and $R_\mu = 0.4$

For both scenarios, the optimistic decision reliability as given

TABLE II
RELIABILITY-ENTROPY TUPLE; N=1000

Scenario	R_α	R_β	R_μ	R_s^o	E_s
1	0.5	0.5	0.0	0.5	1.00
2	0.3	0.3	0.4	0.5	1.57
3	0.2	0.4	0.4	0.33	1.51

by Eqn. (14) is 0.5 as shown in Table II. It can be noted that though scenario 2 has higher R_μ than scenario 1, R_s^o for both are the same. However, intuitively we ought to trust scenario 1 more than scenario 2 because more certain decisions have been made when R_μ is less. We know that higher values of R_μ reduces the chances of being closer to the real value of reliability. As shown in the two scenarios, the optimistic reliability can yield the same value for different sets of R_α, R_β, and R_μ

In order to illustrate the uncertainty associated with the R_μ, we use entropy which is a measure of uncertainty inherent in a system. Usually, entropy of a system uses the steady state probabilities that the system could be in. We define the entropy of the decision reliability as:

$$E_s = -[R_\alpha log_2(R_\alpha) + R_\beta log_2(R_\beta) + R_\mu log_2(R_\mu)] \quad (16)$$

The entropy, E_s, captures the uncertainty which is shown in Table. II. Scenario 1 with $R_\mu = 0$ has an uncertainty measure of 0.69 which is lower than scenario 2 with 1.08. Thus, the reliability of scenario 1 can be trusted more than that of scenario 2.

Now, let us consider scenario 3 which has the same R_μ as of scenario 2; however, R_β is higher than scenario 2. Hence scenario 3 has lesser reliability value than scenario 2. Since scenario 2 has been established to be less trustworthy than 1, we can also verify that from R_s^o of scenario 3 is also less than that of scenario 1. However, these arguments hold true only when at least 50% of the observations are accurate observations, i.e., $R_\mu < 0.5$. This is a reasonable assumption as it is impractical to have a detection mechanism which cannot detect or decide majority of the time.

V. SIMULATION MODEL AND RESULTS

We simulate a generic centralized system with 100 components. Inputs from all components are monitored by an imperfect monitoring mechanism that produces three possible outcomes. The probability of detection, P_{det}, is varied to capture its effects on decision reliability.

An adversary attacks and compromises different sets of inputs over time. The number of inputs compromised vary over time slot; although the long-term average of the number of inputs compromised, denoted by P_a, remains the same. We study the decision reliability for different values of P_a, and plot instantaneous and moving average of decision reliability. We first present the results for an optimistic system and then for a more conservative system.

A. Optimistic Decision Reliability: Instantaneous and Average

In Fig. 2, we plot both the instantaneous and average decision reliability for the optimistic reliability model when the adversary launches attacks with $P_a = 0.5$. We observe that the decision reliability fluctuates over time. As expected, with sufficient observations, the moving average of decision reliability converges to a steady state reliability equal to $1-P_a$ which in this case is 0.5.

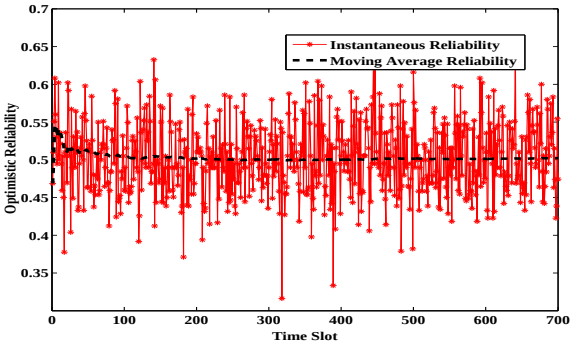

Fig. 2. Instantaneous and average decision reliability with $P_a = 0.50$ and $P_{det} = 0.8$

B. Decision Reliability and Entropy

In Fig. 3, we plot steady state decision reliability with increasing P_a and for two different values of P_{det}. The plots show a steady decrease of decision reliability values for P_{det} 0.5 and 0.9. Recall, with different P_{det} the values of average decision reliability may differ but the relation between R_s^o and P_a does not change with change in P_{det} unlike R_s^c. This is because inputs chosen by the adversary is uniformly random. The conservative reliability R_s^c, also falls linearly with increasing P_a, but the slope or rate of this change varies for the different values of P_{det}. This is because the conservative model does not account for the undecided ones. Hence a lower value of P_{det} yields a lower reliability value than a higher value of P_{det}.

Furthermore, Fig. 4 shows the entropy values of the system under the same set of P_{det}. This shows that a system with higher P_{det} has a lower uncertainty on the decision reliability estimate for all values of P_a. This is useful when we use optimistic system reliability which uses a fraction of undecided ones to contribute to the final reliability value.

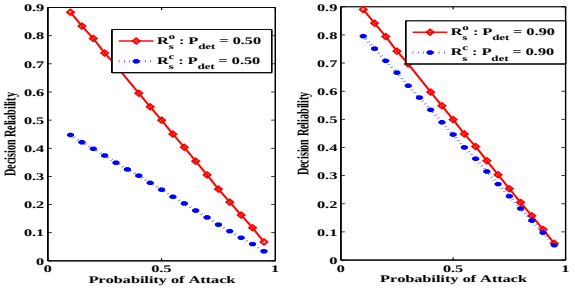

Fig. 3. Optimistic and Conservative Reliability over P_a for different P_{det}

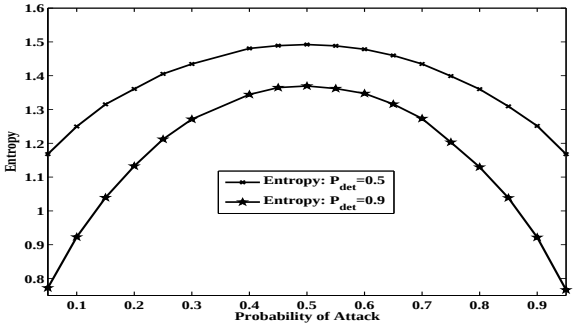

Fig. 4. Entropy over different values of P_a

In Fig. 5, we show how entropy changes with increasing probability of attack. As time evolves and P_{det} improves, the uncertainty (i.e., entropy) associated with the decision reliability decreases indicating increased confidence on the reliability values.

C. Decision Reliability with Time Variant P_a

Some adversaries may choose to attack with short-term low or high attack probability; however, maintaining the long-term average value of P_a. For example, an adversary attacks less initially conserving its attack resources for future and eventually attacking more (under favorable conditions). In Fig. 6, we investigate how the proposed model behaves in such cases. We consider an adversary with $P_a = 0.5$. The first 500 slots are attacked less and next 500 slots are compensated by attacking more. We observe that decision reliability progressively moves towards the expected reliability, although no strict convergence on the decision reliability is achieved.

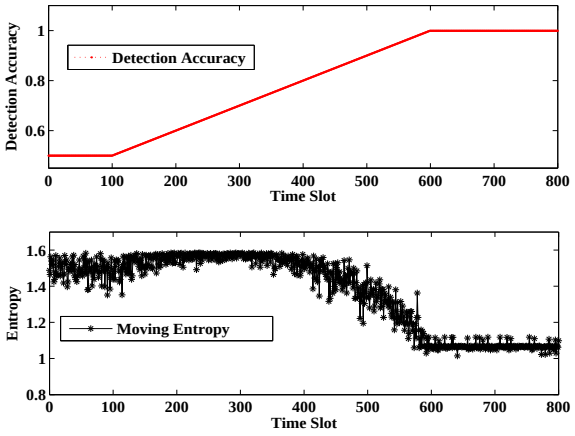

Fig. 5. Conservative decision entropy with incremental increase of P_{det}

D. Conservative Reliability Model

In Fig. 7, we plot the changes in conservative decision reliability, over time with gradual increase in P_{det}. As mentioned earlier, if the system evolves into more accurate monitoring, the decision reliability also improves, although $R_s \neq 1 - P_a$. Thus, having a conservative decision reliability is not unfair, given that risk associated with it is high, and there is scope for improving reliability when detection accuracy increases.

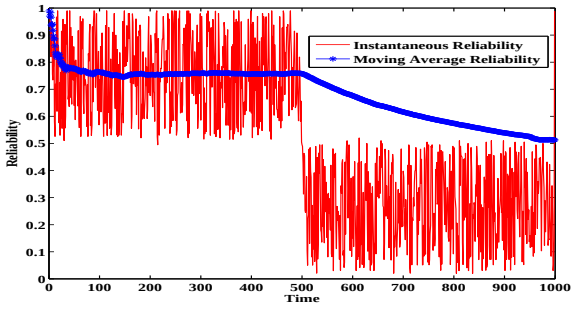

Fig. 6. Decision Reliability under Non-Uniformly distributed $P_a = 0.50$

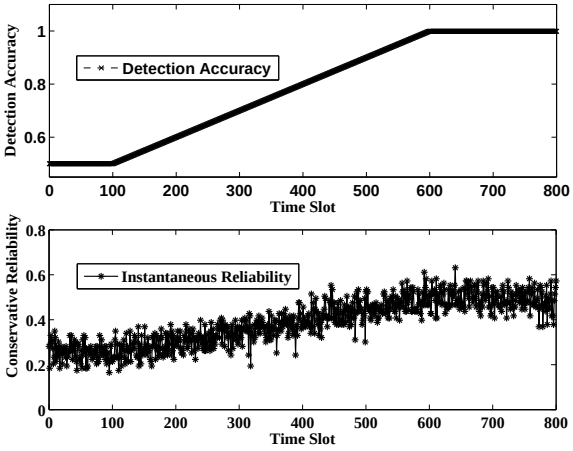

Fig. 7. Conservative decision reliability with increasing P_{det}

VI. Conclusions

In this paper, we presented a Bayesian inference based model to translate multinomial observations for any voting system under imperfect monitoring. We computed the posterior belief for the decision reliability based on the number of compromised, not compromised, and undecided inputs. We proposed two models for decision reliability– optimistic and conservative. Through simulation experiments, we showed how the decision reliability changes with attack probability which also affects the detection accuracy of the underlying monitoring mechanism. Using entropy, we provided a way to evaluate the certainty of the reliability calculation. Finally, we showed how the decision reliability decreases when the attack probability increases.

References

[1] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," *IEEE Trans. Dependable Secure Computing*, vol. 1, no. 1, pp. 11–33, Jan. 2004.
[2] D. Lindley "Making Decisions", *Wiley 2nd Edition*, pp. 104, 1991.
[3] K. Fang and Y.T. Zhang, "Generalized Multivariate Analysis, *Springer-Verlag*, pp. 47-49, 1990.
[4] K. Kwiat, A. Taylor, W. Zwicker, D. Hill, S. Wetzonis, and S. Ren, "Analysis of binary voting algorithms for use in fault-tolerant and secure computing," ICCES, 2010, pp. 269-273.
[5] D. Lindley "Making Decisions", *Wiley 2nd Edition*, pp. 104, 1991.
[6] D. P. Siewiorek and R. S. Swarz, *Reliable Computer Systems: Design and Evaluation*, 3rd Edition, A. K. Peters/CRC Press, 1998.
[7] Y.L. Sun, W. Yu, Z. Han and K.J.R. Liu, "Information Theoretic Framework of Trust Modeling and Evaluation for Ad-Hoc Networks, *IEEE J. Sel. Areas in Commn*, Vol. 24, Issue 2, 2006.