Collaborative Quickest Detection in Adhoc Networks with Delay Constraint - Part I: Two-node Network

Husheng Li, Chengzhi Li and Huaiyu Dai

Abstract—Collaborative quickest detection is used to detect a certain distribution change in an adhoc network, in which there is no data fusion center and local decisions need to be made at all nodes. In this part of our two-part paper, a two-thread Cumulative Sum (CUSUM) algorithm is proposed for a simple two-node network. The corresponding performance, particularly the performance gain induced by collaboration, is analyzed. Numerical results are provided to verify the theoretical analysis. In the sequel, we extend our study to the general multi-node case, and further consider block transmission mode and correlated observations.

I. INTRODUCTION

Quickest detection is a technique to detect the distribution change of observations as quickly as possible, which is useful in many applications, such as machinery monitoring, finance and seismology [1]. A useful algorithm for quickest detection, called Cumulative Sum (CUSUM) test, was proposed by Page in [9]. The asymptotic optimality of CUSUM test was shown by Lorden in [6] while Moustakides [7] proved its optimality for finite-sample case. The CUSUM test was extended to the case of exponential penalty for delay in [11]. There are also plenty of researches on the applications of quickest detection. In [16], quickest detection is used to detect the burst of a sinusoidal signal with unknown frequency, amplitude and phase while [8] applied quickest detection in quality control. An efficient way to analyze the performance of quickest detection is to apply Brownian motion approximation [11] [13] [17].

The above researches are all based on a single observer (called \textit{centralized quickest detection}). However, in many practical systems, a single observer may not be reliable. Therefore, people consider \textit{decentralized quickest detection} [18] [19] [20] [2], in which there are multiple sensors sampling the environment and a data fusion center collecting data from sensors and making final decision. Due to communication constraint, a key problem in decentralized quickest detection is the quantization of the data transmitted to the data fusion center. In [19], it has been proved that likelihood ratio quantizer (LRQ) is optimal and the structure of optimal quantizer has been discussed in the framework of the dynamic programming. [2] discussed a scheme in which a binary decision is made at each sensor and the data fusion center makes a final decision based on the binary decisions.

In this paper, we investigate quickest detection in an adhoc network, which has no data fusion center. In such networks, each node makes its own local decision, possibly with the collaboration of other nodes through information exchange. Such an assumption is valid in a wide range of scenarios. For instance, in distributed control, each node in the network can carry out local control when a distribution change is detected. A possible example is cognitive radio [22], in which a secondary radio needs to vacate from the corresponding frequency spectrum if a primary radio emerges. Essentially, this requires a quickest detection of the change of spectrum activity and quite often there is no centralized coordination in the secondary network if the secondary radio system is an ad hoc network.

Our approach is still based on the CUSUM test. Different from the data fusion center based network, there exists a communication delay for the information conveyed from other nodes. In this part, we consider a simple two-node network, propose a two-thread CUSUM test and analyze its performance by applying Brownian motion approximation. Extensions of this study will be given in a companion paper.

The remainder of this paper is organized as follows. The system model and known results in literatures are introduced in Sections II and III. In Section IV, a simple two-node network is discussed. Numerical results and conclusions are provided in Sections VI and VII. Due to limited space, all proofs in this paper are omitted.

The notations used in this paper are explained as follows.

- $\lceil x \rceil^+$ means $\max(x, 0)$.
- $E_\tau[\cdot]$ means the expectation conditioned that the change happens at time $\tau$. $V_\tau[\cdot]$ means the variance conditioned that the change happens at time $\tau$.
- $x \geq y$ means $\lim_{\gamma \to \infty} \frac{x}{y} \geq 1$, where $\gamma$ is the threshold for detection defined later. Similarly, we can define $\leq$ and $\approx$.

II. SYSTEM MODEL

Suppose that a network containing $N$ nodes, whose node set is denoted by $\Theta$, is monitoring some property of the environment. Suppose that each node samples the environment using the same sampling frequency. We denote the observation of node $\theta$ at the sample period $t$ by $X_\theta(t)$ ($t$ begins from 0), whose distribution satisfies two possible hypotheses: $H_0$ and $H_1$. The corresponding probability density function under hypothesis $H_i$ at node $\theta$ is denoted by $f^{(i)}_\theta$. The corresponding log likelihood ratio is defined as

$$l_\theta(t) \triangleq \log \left( \frac{f^{(1)}_\theta(X_\theta(t))}{f^{(0)}_\theta(X_\theta(t))} \right).$$

($1$)
Suppose that the beginning distribution hypothesis is $H_0$ and it is changed to $H_1$ at a certain time, which is unknown to all nodes. The task of the nodes is to detect the change as quickly as possible. They can communicate with their neighbors to enhance the detection of distribution change. Suppose that the system is time slotted. Fig. 1 shows the structure of time slots. Each time slot has two periods: transmit period and sampling period. In the transmit period, each node transmits its own information to other nodes and receives information from others. In the sampling period, each node obtains one sample and prepares the transmission of the next time slot. We make the following assumptions:

- All nodes are perfectly synchronized.
- The communication channel is sufficiently good such that there is no transmission error.
- There is a delay of $D$ time slots for the communications between two neighboring nodes, which may be caused by signal propagation time and processing overhead, e.g., converting observation into log likelihood ratio, modulation and demodulation. Then, a node sees continuous error-free data streams with different delays from other nodes.
- $X_{s_{i}}(t_1)$ and $X_{s_{i}}(t_2)$ are mutually independent if $(i_1, t_1) \neq (i_2, t_2)$.

Note that the above assumptions may not be precise in practical systems. For instance, there always exist transmission errors due to noise or interference in practical communication channels. Moreover, in a practical communication system, data is transmitted in blocks; therefore, the data arrives at a node abruptly, instead of continuously. However, our analysis captures the essential impact of communication delay, which exists in almost all communication systems.

We use two average run lengths (ARL), which have been used in [7] [11], to measure the performance. For node $\theta$, these ARLs are defined as

$$D_{\theta} = \text{esssup} \left( E_\theta [T_{\theta} - t | \mathcal{F}_{t-1}] \right),$$

$$\mathcal{F}_{\theta} = E_{\infty} [T_{\theta}],$$

where $T_{\theta}$ is the stopping time that node $\theta$ determines that a distribution change has happened. Recall that $E_t$ denotes the expectation under the assumption that the change happens at time slot $t$, $\mathcal{F}_{t-1}$ is the filtration, namely the smallest $\sigma$-field with respect to $X_{0}(0), \ldots, X_{0}(t-1)$. The esssup means the worst case detection delay. By applying the same argument as in [7] [11], we can replace $t$ with 0 in the definition of $D_{\theta}$, namely we can assume that the change happens at time slot 0 without affecting the performance measure. $E_{\infty}$ is the expectation under the assumption that the change never happens (the change time is $\infty$). Obviously, we desire a small $D_{\theta}$ and a large $\mathcal{F}_{\theta}$.

### III. Known Results of Single Node

In this section, we list known results for the single node case, or equivalently centralized quickest detection. For simplicity, we ignore node indices in all notations.

#### A. CUSUM Test and Page’s Procedure

For a single node, the stopping time for the distribution change in CUSUM test [9] is given by

$$T^* = \inf \left\{ t \mid m(t) \geq \gamma \right\},$$

where $\gamma$ is a threshold and metric $m(t)$ is defined as

$$m(t) = \max \left( m(t-1) + l(t), 0 \right).$$

Intuitively, $T^*$ is the first time slot in which the metric $m(t)$ passes the threshold $\gamma$.

The CUSUM test has an intuitive explanation called Page’s Procedure [9]. It is easy to verify that the stopping time $T^*$ in (4) is also given by

$$T^* = \inf \left\{ T_k, k = 0, 1, 2, \ldots \right\},$$

where

$$T_k = \inf \left\{ t \mid \sum_{r=k}^{t} l(r) \geq \gamma \right\}.$$

Intuitively, $T_k$ means the hitting time of metric accumulated from time slot $k$ and $T^*$ is the earliest one among $T_k$’s.

#### B. Performance Analysis

It is difficult to obtain explicit expressions for performance analysis in the finite case. Brownian motion [14] is widely used to approximate the performance of finite systems having sufficiently large threshold (equivalently, sufficiently large amount of samples for detecting the change) since

$$\hat{B}(t) \triangleq \frac{1}{\sqrt{N}} \sum_{r=0}^{[Nt]} (l(r) - E[l(r)])$$

converges to a Brownian motion with zero drift (denoted by $\mu$) and variance $\sigma^2 \triangleq V[l(r)]$, as $N \to \infty$. Similarly the threshold is changed to $a = \frac{\gamma}{\sqrt{N}}$. Therefore, for sufficiently large $N$ and $\gamma$ of larger or equal order to $\sqrt{N}$, the random walk $\frac{1}{\sqrt{N}} \sum_{r=0}^{[Nt]} l(r)$ can be approximated by a Brownian motion $B(t)$ with drift $\mu = \sqrt{N} E[l(r)]$.

Note that the metric $m(t)$ in (5) can be rewritten as

$$m(t) = \sum_{r=0}^{t} l(r) - \min_{-1 \leq k \leq t} \sum_{r=0}^{k} l(r).$$

Then, the corresponding cumulated metric in the Brownian motion is given by

$$b(t) \triangleq B(t) - \min_{0 \leq s \leq t} B(s),$$

and the stopping time is given by

$$t^* = \inf \left\{ \left| b(t) \right| \geq a \right\}.$$
Hence, the performance analysis is converted to study the hitting time of random process $b(t)$ in (10). This has been studied in [17], in which the generating function for $t^*$ is obtained.

IV. TWO-NODE NETWORK

Let us first consider the simplest case: a two-node network. The two nodes, denoted by A and B, can exchange information for cooperation. Each node makes its own decision based on its local observations and exchanged information. For simplicity, we consider only the detection process at node A. The detection algorithm and performance analysis at node B can be derived similarly.

A. Page’s Procedure

To obtain a stopping rule of the change detection, we can adopt Page’s procedure in (6), in which the stopping time at node A is set to

$$T_A^* = \inf \{ T_A(k) | k = 0, 1, 2, \ldots \},$$

where

$$T_A(k) = \inf \{ t | m_k^{(A)}(t) \geq \gamma_A \},$$

and

$$m_k^{(A)}(t) = \sum_{r=k}^{t-D} l_A(r) + \sum_{r=k}^{t} l_B(r).$$

Note that $\gamma_A$ is a predetermined threshold at node A, which could be different from that of node B, denoted by $\gamma_B$. Obviously, $T_A(k)$ is the stopping time assuming that the change happens at time slot $k$, $m_k^{(A)}(t)$ is accumulated metric beginning from time slot $k$. Then, $T_A(k)$ is the hitting time of the accumulated metric $m_k^{(A)}(t)$ with respect to the threshold $\gamma_A$. The term $t - D$ in (14) is due to the communication delay $D$.

B. Two-thread CUSUM test

As stated in Section III, direct application of $T_A^*$ in the Page’s procedure in (12) requires infinite memory for storing the metrics of all time slots. However, we can compress this requirement to fixed amount of memory by using recursive computation of metrics, similar to CUSUM test in (4).

At time slot $t$, we consider two portions of observations (denoted by $\alpha$ and $\beta$), namely observations before $t - D + 1$ and observations from $t - D + 1$ to $t$. In portion $\alpha$, node A has observations of both nodes while it has only its own observations in portion $\beta$. We define two stopping times for the two stages, which are given by

$$T_A^\alpha = \inf \{ T_A(k) | k = 1, \ldots, t - D \},$$

and

$$T_A^\beta = \inf \{ T_A(k) | k = t - D + 1, \ldots, t \}.$$

Obviously, the stopping time $T_A^*$ in (12) can be expressed as

$$T_A^* = \min \left( T_A^\alpha, T_A^\beta \right).$$

Therefore, there are two simultaneously evolving threads and we call this algorithm two-thread CUSUM test.

Next, we discuss $T_A^\alpha$ and $T_A^\beta$, separately.

- Stopping Time $T_A^\alpha$: $T_A^\alpha$ can be obtained similarly to CUSUM test, namely

$$T_A^\alpha = \inf \left\{ t | m_A^\alpha(t - D) + \sum_{r=t-D+1}^{t} l_A(r) \geq \gamma_A \right\},$$

where

$$m_A^\alpha(t) = \max \{ m_A^\alpha(t - 1) + l_A(t) + l_B(t), 0 \}.$$

Note that $m_A^\alpha(-1) = 0$ since the test begins at time slot 0.

- Stopping Time $T_A^\beta$: For $T_A^\beta$, only observations at node A during time slots $t - D + 1$ to $t$ can be used. Then $T_A^\beta$ can be written as

$$T_A^\beta = \inf \left\{ t | \max_{t-D+1 \leq k \leq t} \sum_{r=k}^{t} l_A(r) \geq \gamma_A \right\}.$$

We can use the following procedure to compute $m_A^\beta(t)$

$$m_A^\beta(t) = \max_{t-D+1 \leq k \leq t} \sum_{r=k}^{t} l_A(r).$$

Note that the observations from time slots $t - D + 1$ to $t$ need to be stored at node A. We also need to store the following variable:

$$a(t) = \arg \max_{0 \leq k \leq D - 1} \sum_{r=t-k}^{t} l_A(r).$$

- Initialization: When the test begins at time slot 0, set $a(0) = 0$ and

$$m_A^\beta(0) = \max \{ l_A(0), 0 \}.$$

- Iterations: when $t$ is increased by one time slot:

* If $a(t) < D$, set

$$m_A^\beta(t) = \max \{ m_A^\beta(t - 1) + l_A(t), 0 \},$$

and

$$a(t) = a(t - 1) + 1.$$

* If $a(t) \geq D$, we have to search for a new $a(t)$ and recompute $m_A^\beta(t)$.

Iterate until $T_A^*$ is reached.

V. PERFORMANCE ANALYSIS OF TWO-NODE NETWORK

In this section, we analyze the performance of a two-node network. We first extend Wald’s Identity [21] to the case having two threads, then formulate the corresponding Brownian motion approximation and analyze $D_A$ and $F_A$. For performance analysis, we need the following definitions.

$$I_i \triangleq E \left[ l_i(r) \right], \quad i = A, B,$$

$$J_i \triangleq E \left[ l_i(\infty) \right], \quad i = A, B,$$

$$U_i \triangleq V \left[ l_i(r) \right], \quad i = A, B,$$

$$W_i \triangleq V \left[ l_i(\infty) \right], \quad i = A, B.$$
Obviously, $I_i$ is the Kullback-Leibler divergence $D\left(f_i^{(1)}||f_i^{(0)}\right)$ and is larger than 0. $J_i$ is equal to $-D\left(f_i^{(0)}||f_i^{(1)}\right)$ and is negative.

A. Extended Wald’s Identity

Based on the above definitions, we can obtain the following lemma which extends Wald’s Identity (Eq. (A.69), page 171, [21]), a fundamental tool for sequential analysis. Here, we assume that the observations at nodes A and B satisfy arbitrary constant distributions, or equivalently, there is no change in distributions.

**Lemma 1:** Define random variable $R_A(n) = \sum_{k=1}^{n} l_A(k) + \sum_{k=1}^{n-D} l_B(k)$. For any stopping time $T$, we have

$$E\left[\frac{\exp(tR_A(T))}{M_1(t)M_2^{(T-D)+}(t)}\right] = 1,$$

where the expectation is with respect to an arbitrary distribution, $t$ is a continuous variable and

$$M_1(t) = E[\exp(tl_A(1))],$$

and

$$M_2(t) = E[\exp(tl_B(1))].$$

By taking derivative with respect to $t$ on both sides of (25) and letting $t \to 0$, we have, for any stopping time $T$,

$$E[R_A(T)] = I_A E[T] + I_B E\left[T - D\right].$$

By applying Jesen’s inequality and the fact that $f(x) = [x]^+$ is a convex function, which implies that $[E[T] - D]^+ \leq [E[T] - D]^+$, we have

$$E[R_A(T)] \geq I_A E[T] + I_B \left[E[T] - D\right].$$

B. Asymptotic Analysis and Brownian Motion Approximation

When it is difficult to obtain closed form expressions for analyzing finite systems, we can assume that $\gamma$ is sufficiently large, which is reasonable if we require a large $F_A$. Throughout this paper, we assume that $D$ is of the same order as $\gamma$ when applying asymptotic analysis. Otherwise, if $D$ is of smaller order than $\gamma$, both nodes obtain all observations with asymptotically vanishing delay; if $D$ is of larger order than $\gamma$, it is useless to communicate between the two nodes. In both cases, the performance degenerates to the canonical situation of centralized quickest detection [1] [9].

Similar to the single observer case in Section III, we can apply the Brownian motion approximation. Then, for finite but sufficiently large $\gamma$ and $D$, which means that the number of samples for test is sufficiently large, we use $B_A(t)$ to approximate the random walk $1, \ldots, \sum_{k=1}^{n} l_A(k)$, similarly to (8). We denote the drift by $\mu_A$ and variance by $\sigma_A^2$. Similarly, we can define $B_B(t)$ to approximate $\sum_{k=1}^{n-D} l_B(k)$.

We assume that time slot $D$ in the discrete system is equivalent to the continuous time $t = 1$ in the Brownian motion, which specifies a mapping between the continuous time in $B_A(t)$ and the discrete time slots (continuous time $t$ is equivalent to discrete time $[tD]$). When the underlying distribution hypothesis is $H_0$, it is easy to verify that that, for $B_A(t)$, the drift $\mu_A = \sqrt{D} I_A$ and the variance $\sigma_A^2 = \mu_A^2$. Similarly, $\mu_A = \sqrt{D} I_A$ and $\sigma_A^2 = \sigma_A^2$ for $H_1$. We assume that the stopping time in Brownian motion is a first order approximation for the stopping times in discrete systems, namely for a stopping time $T$ in the discrete system, random variable $\frac{T}{D}$ converges to the corresponding stopping time $t$ in the Brownian motion approximation in distribution. This assumption, which is called Brownian motion approximation assumption, will be validated with numerical simulation results.

We denote the corresponding stopping time in the Brownian motion approximation by $T_A^\ast$, which satisfies $T_A^\ast \approx D t_A^\ast$. Similarly, we use $t_A^\ast$, $T_B^\ast$, $\sigma_A$, and $\sigma_B$ to denote the quantities corresponding to $T_A^\ast$, $T_B^\ast$, $F_A$ and $D_A$ respectively. It is easy to check that the stopping times $t_A^\ast$ and $t_B^\ast$ in the Brownian motion approximation can be written as

$$t_A^\ast = \inf \left\{ \tau | B_A(\tau) + B_B(\tau - 1) \leq \min_{0 \leq s \leq \tau} (B_A(s) + B_B(s)) \geq a_A \right\},$$

and

$$t_B^\ast = \inf \left\{ \tau | B_A(\tau) - \min_{-1 \leq s \leq \tau} B_A(s) \geq a_A \right\},$$

where the threshold $a_A \triangleq \frac{\gamma A}{\sqrt{D}}$. The stopping time $t_A^\ast$ is given by

$$t_A^\ast = \min \left\{ t_A^\ast, t_B^\ast \right\}.$$

C. Analysis of $D_A$

Applying (28) and ignoring the overshoot at stopping time $T_A^\ast$, we can obtain an extended Wald’s approximation, which is given by

$$\gamma_A \approx I_A E_0[T_A^\ast] + I_B E_0 \left[\left[T_A^\ast - D\right]^+\right].$$

However, it is difficult to decouple $T_A^\ast$ from $E_0[T_A^\ast - D]^+$ in (33). We need to know the probability that $T_A^\ast$ is larger (or smaller) than $D$, which is difficult to obtain. Therefore, we seek help from Brownian motion approximation. The following proposition states that, for sufficiently large $D$, it is highly probable that, if $I_A D \geq \gamma_A$ ($I_A D \leq \gamma_A$), the threshold $\gamma_A$ can (cannot) be reached during time slots 0 to $D$ (or equivalently, during time interval [0, 1] in the Brownian motion approximation)$^1$.

**Proposition 1:** In the Brownian motion approximation, we have

$$P(t_A^\ast < 1) \to \begin{cases} 1, & \text{if } \mu_A > a_A \\ 0, & \text{if } \mu_A < a_A \end{cases},$$

as $a_A \to \infty$.

Based on Prop. 1, we can obtain the following corollary:

**Corollary 1:** For sufficiently large $D$, we have

$$\frac{[T^* - D]^+}{D} = \begin{cases} 0, & \text{if } I_A D > \gamma_A \\ \frac{t_A^\ast}{D}, & \text{if } I_A D < \gamma_A \end{cases},$$

$^1$ We ignore the discussion of $I_A D = \gamma_A$ since it rarely holds. Simulation results show that $P(t_A^\ast < 1) = 0.5$ in this situation.
Applying (1) to (33), for sufficiently large $D$, we have
\[ D_A \approx \begin{cases} \frac{2A}{2A+D_B}, & \text{if } DI_A > \gamma_A \\ \frac{\gamma_A + Di_B}{\gamma_A + Di_B}, & \text{if } DI_A < \gamma_A \end{cases}. \tag{36} \]

Qualitatively speaking, the information conveyed from node B does not help to improve $D_A$ if the threshold $\gamma_A$ is small. However, this by no means implies that it is useless to exchange information between nodes A and B since we can adjust the threshold to achieve a better tradeoff between $F_A$ and $D_A$.

D. Analysis of $F_A$

It is difficult to analyze $F_A$ in finite systems. Again, we use Brownian motion approximation for performance analysis when $\gamma_A$ is sufficiently large.

The following lemma shows that we need not consider $T_A$ since it is rarely attained when $\gamma_A$ is sufficiently large, with the assumption that $D$ is of the same order as $\gamma_A$. Or equivalently, the false alarm seldom happens before the information from node B arrives. Note that $\mu_A < 0$ under $H_0$.

**Lemma 2:** When $\mu_A < 0$, as $a \to \infty$, we have
\[ P \{ I_A^* > 1 \} \to 1. \tag{37} \]

Based on Lemma 2, we can obtain the following proposition for $F_A$. The proof is omitted due to limited space.

**Proposition 2:** As $a_A \to \infty$, we have
\[ \log f_A \geq \frac{2|\mu_A|}{\sigma^2} - \frac{2|\mu_A|^2}{\sigma^2} - 2\mu^2\sigma_A^2 \tag{38} \]
where
\[ \mu = \mu_A + \mu_B, \]
and
\[ \sigma^2 = \sigma_A^2 + \sigma_B^2. \]

Based on the Brownian motion approximation, a corollary of Prop. 2 is given as follows.

**Corollary 2:** As $\gamma_A \to \infty$, we have
\[ \log F_A \geq \frac{2|\gamma_A|}{W_A + W_B} - \frac{2|\gamma_A|^2}{(W_A + W_B)^2} - 2|\gamma_A|^2 W_A \tag{39} \]

E. Special Case: Gaussian Distributions

Suppose that, for node $i$, $H_0 \sim N(\mu_i, 1)$ and $H_1 \sim N(\mu_i, 1)$, namely both hypotheses are Gaussian distributions with different expectations. Then, it is easy to check [11]
\[ I_A = -J_A = (\mu_A - \mu_0)^2, \tag{40} \]
\[ I_B = -J_B = (\mu_B - \mu_0)^2, \tag{41} \]
\[ U_A = 2I_A, \tag{42} \]
\[ U_B = 2I_B. \tag{43} \]

Then, substituting (40) into (38) for sufficiently large $a_A$ (or $\gamma_A$), we have
\[ \log (f_A) \geq a_A, \tag{44} \]
or equivalently
\[ \log (F_A) \geq a_A. \tag{45} \]

Combining the above results, we can reach the following proposition which states that the inequality in (45) is actually an equality in asymptotic sense.

**Proposition 3:** When Brownian motion approximation holds and the distributions are Gaussian, we have
\[ \log (F_A) = a_A. \tag{46} \]

The proof of Prop. 3 makes use of the following lemma.

**Lemma 3:** For any $c \in (0, 1)$ and any stopping time $T_A$, there exists a $C(c) \in \mathbb{R}$ such that
\[ \log E\left[ T_A \right] \leq a_A. \tag{47} \]

Then, based on Brownian motion approximation and (47), we have
\[ I_A D_A + I_B [D_A - D]^+ \geq \log F_A. \tag{48} \]

However, based on (36) and (45), we have
\[ I_A D_A + I_B [D_A - D]^+ \leq \log F_A. \tag{49} \]

Combining the above two inequalities yields the conclusion of Prop. 3. The argument also yields the following proposition which characterizes the tradeoff between $F_A$ and $D_A$.

**Proposition 4:** For node A, the following equality holds asymptotically:
\[ I_A D_A + I_B [D_{1A} - D]^+ \equiv \log F_A. \tag{50} \]

VI. NUMERICAL RESULTS

We now illustrate the analytical results in this paper via simulation. Note that the performance metric $F$ is not simulated since it is too large for a numerical simulation.

We first simulate the performance of two-node networks. Fig. 2 shows the curves of $D$ as a function of communication delay $D$ while the threshold $\gamma$ is fixed at 400 for both nodes. We assume that both nodes share the distributions: $H_0 \sim N(1, 1)$ and $H_1 \sim N(1, 1)$. We observe a good match of analytical and simulation results. In Fig. 3, we assume that node A has worse observations, namely $H_0 \sim N(-0.1, 1)$ and $H_0 \sim N(0, 1)$ while those of node B are the same as in Fig. 2. The figure
Fig. 3: Two-node network: $D$ versus different $D$’s: different distributions.

Fig. 4: The probability that the collaboration is beneficial versus different $D$’s.

shows the ARL of detection delay of node $A$. We observe that the cooperation can substantially improve the performance of node having low signal-to-noise ratio (SNR) observations. Fig. 4 shows the probability that the collaboration is beneficial (or equivalently, $T^* > D$) versus different communication delays, where the radio is defined as $\frac{D}{T^*}$. We observe that, as $D$ increases, the probability converges to 0 or 1, which coincides with the conclusion in Prop. 1.

VII. Conclusions

In this paper, we have extended the traditional quickest detection problem to an adhoc network without a data fusion center. We have considered the transmission delay as communication constraint on the collaborative quickest detection. For a simple two-node network, we have proposed a two-thread algorithm as an extension of CUSUM test. We have analyzed the corresponding detection performance, measured by the ARL of detection delay and ARL of false alarm, by applying Brownian motion approximation. Numerical results have shown that the collaboration among different nodes can substantially improve the system performance.

REFERENCES