MAP Detection of Misbehaving Relay in Wireless Multiple Access Relay Networks

Sang Wu Kim, Taha Khalaf, and Sangmun Kim

***Abstract*—We propose the maximum *a posteriori* (MAP) detection of the misbehaving relay that injects false data or adds channel errors into the network encoder in multiple access relay networks. The proposed scheme does not require sending extra bits at the source and is optimal in the sense of minimizing the probability of incorrect detection. We derive the probability of false alarm and misdetection, taking into account the lossy nature of wireless links. The side information regarding the presence of relay misbehavior is exploited at the decoder to mitigate the relay misbehavior and enhance the reliability of decoding.**

***Index Terms*—False data injection, relay misbehavior, network coding, maximum *a posteriori* detection, multiple access relay network.**

I. Introduction

Network coding is a new relaying technique that replaces the traditional store and forward paradigm of network routing by a method that allows intermediate (relay) nodes to mix the received data before re-transmission [1], [2], [3]. This mixing has been shown to maximize throughput, as well as robustness against failures and erasures. While network coding can be an efficient means of information dissemination in networks, it also presents a new security challenge as the injection of even a single erroneous packet has the potential to corrupt every packet received by a given destination.

The problem of detecting misbehaving relays that inject false data in *single-source* networks has been studied in [4]-[7]. In [4] the authors consider a peer-to-peer (P2P) network in which the source generates a signature vector and broadcasts to all nodes where it is used to check the integrity of the received packets. In [5] and [6] several Information Theory-based algorithms are proposed for mitigating Byzantine modification attack. In [7] the authors consider inserting tracing bits in the data stream at the source in a cryptographically secure manner. The receiver then computes the ground truth of the tracing bits and compares them with the tracing bits received from a relay to determine whether it is malicious or cooperative. Extensions to multiple-source networks have been studied in [8], [9], where the tracing bits or polynomial hash functions are used in detecting the misbehaving relays. All these works, however, require sending extra reference data (overhead) at the source to detect the misbehaving relay.

In this letter, we propose the maximum *a posteriori* (MAP) approach in detecting the misbehaving relay that injects false data or adds channel errors into the network encoder in multiple access relay networks. The MAP detection scheme is based on the log-likelihood ratio (LLR) test which is optimal in the sense of minimizing the probability of incorrect decisions (false alarm and misdetection). The proposed scheme does not require sending extra bits at the source, such as hash function or message authentication check bits, and hence there is *no transmission overhead*. In addition, it makes an instantaneous decision about whether a relay is behaving properly without a long term observation. We derive the probability of false alarm and misdetection as a function of the signal-to-noise ratio, taking into account the lossy nature of wireless links.

The side information regarding the presence of relay misbehavior is exploited at the destination (decoder) to mitigate the relay misbehavior and enhance the reliability of decoding. We present the MAP decoding scheme that exploits the likelihood of the presence of relay misbehavior, and show that it provides a dramatic improvement over the conventional decoder that does not exploit the side information.

II. System Model

Consider a multi-access relay network composed of two sources, one relay, and one destination. The relay overhears the bits sent by the sources (possibly with some errors), linearly combines (encodes) them, and forwards the coded bit to the destination for every bit sent by the source nodes. We assume that all bits are sent through orthogonal Rayleigh fading channels with additive white Gaussian noise and path loss, and each node is equipped with a single antenna.

Let $x_i \in \{+1, -1\}$ denote the bit transmitted by the i-th source, $i = 1, 2$, and $x_i^r \in \{+1, -1\}$ denote the overheard bit by the relay, where $+1$ is the additive identity element under $\oplus$ (modulo-2) addition. The relay combines the overheard bits and produces a coded (parity) bit

$$p = x_1^r \oplus x_2^r \oplus f \tag{1}$$

where $f \in \{+1, -1\}$ denotes the injected bit by the relay to corrupt the communication. If $f = -1$, the false bit is injected, and if $f = +1$, no false bit is injected.

Let $e_i \in \{+1, -1\}$ be the error value between the i-th source and the relay, i.e. $x_i^r = x_i \oplus e_i$, where $e_i = -1$ means $x_i^r \neq x_i$, i.e. x_i is received in error at the relay, and $e_i = +1$ means $x_i^r = x_i$. Then, (1) can be written as

$$p \;=\; x_1 \oplus x_2 \oplus z \tag{2}$$

where $z = e_1 \oplus e_2 \oplus f$ captures the error events on the source-to-relay channels and the false data injection by the relay.

The received signals at the destination are given by

$$y_i \;=\; h_i x_i \sqrt{d_i^{-m} E_s} + n_i, \; i = 1, 2 \tag{3}$$

$$y_r \;=\; h_r p \sqrt{d_r^{-m} E_r} + n_r \tag{4}$$

Manuscript received July 23, 2010. The associate editor coordinating the review of this letter and approving it for publication was A. Burr.

S. W. Kim and T. Khalaf are with the Dept. of Electrical and Computer Engineering, Iowa State University, Ames, Iowa 50011 (e-mail: {swkim, taha}@iastate.edu).

S. Kim is with the Dept. of Mathematics, Yonsei University, Seoul 120-749, Korea (e-mail: eMath@mit.edu).

Digital Object Identifier 10.1109/LCOMM.2011.012511.101323

1089-7798/11$25.00 © 2011 IEEE

where y_i and y_r are the received signals from the i-th source and the relay, respectively; h_i and h_r are the channel fading gain between the i-th source and the destination and that between the relay and the destination, respectively; d_i and d_r are the distance between the i-th source and the destination and that between the relay and the destination, respectively; m is the path loss exponent; E_s and E_r are the transmit energy per symbol at the source and the relay, respectively; n_i and n_r are the noise at the destination. It is assumed that h_i and h_r are the independent complex Gaussian random variables with mean zero and variance one, and n_i and n_r are the independent complex Gaussian random variables with mean zero and variance $N_0/2$ per dimension.

III. MAP Detection Scheme

The destination is interested in finding whether z is $+1$ (well-behaving) or -1 (misbehaving) and exploiting the information regarding z in decoding. The MAP detection rule which minimizes the probability of incorrect decisions is based on the LLR of z:

$$\begin{aligned} L(z|\mathbf{h},\mathbf{y}) &= \ln\frac{P(z=+1|\mathbf{h},\mathbf{y})}{P(z=-1|\mathbf{h},\mathbf{y})} && (5)\\ &= \ln\frac{P(p\oplus x_1\oplus x_2=+1|\mathbf{h},\mathbf{y})}{P(p\oplus x_1\oplus x_2=-1|\mathbf{h},\mathbf{y})} && (6)\\ &\approx \text{sign}(L_r)\text{sign}(L_1)\text{sign}(L_2)\\ &\quad\cdot\min\{|L_r|,|L_1|,|L_2|\} && (7)\end{aligned}$$

where $\mathbf{h}=(h_1,h_2,h_r)$, $\mathbf{y}(y_1,y_2,y_r)$, and

$$\begin{aligned} L_i &= \ln\frac{P(x_i=+1|h_i,y_i)}{P(x_i=-1|h_i,y_i)}\\ &= \frac{4\sqrt{d_i^{-m}E_s}}{N_0}Re\{h_i^* y_i\} && (8)\end{aligned}$$

is the LLR of x_i after knowing h_i and y_i, and

$$\begin{aligned} L_r &= \ln\frac{P(p=+1|h_r,y_r)}{P(p=-1|h_r,y_r)}\\ &= \frac{4\sqrt{d_r^{-m}E_r}}{N_0}Re\{h_r^* y_r\} && (9)\end{aligned}$$

is the LLR of p after knowing h_r and y_r. The approximation in (7) follows from [10]. Then, the MAP detection rule is

$$\begin{cases} \hat{z}=+1, & \text{if } L(z|\mathbf{h},\mathbf{y})\geq 0\\ \hat{z}=-1, & \text{if } L(z|\mathbf{h},\mathbf{y})<0\end{cases} \quad (10)$$

where $\hat{z}$ is the estimation of z.

IV. Probability of False Alarm and Misdetection

In this section we derive the probability of false alarm and misdetection. The false alarm is the event that the destination decides $\hat{z}=-1$ given $z=1$, i.e. $p=x_1\oplus x_2$, while the misdetection is the event that the destination decides $\hat{z}=1$ given $z=-1$, i.e. $p=x_1\oplus x_2\oplus(-1)$. It follows from (10) that the probability of false alarm is given by

$$P_{FA} = P(L(z|\mathbf{h},\mathbf{y},p=x_1\oplus x_2)<0) \quad (11)$$

Since (7) indicates that the sign of $L(z|\mathbf{h},\mathbf{y})$ becomes negative if one or three of the LLRs L_r, L_1, L_2 are negative, we obtain

$$\begin{aligned} P_{FA} &= P(L_r<0)\cdot P(L_1>0)\cdot P(L_2>0)\\ &\quad+P(L_r<0)\cdot P(L_1<0)\cdot P(L_2<0)\\ &\quad+P(L_r>0)\cdot P(L_1>0)\cdot P(L_2<0)\\ &\quad+P(L_r>0)\cdot P(L_1<0)\cdot P(L_2>0) && (12)\end{aligned}$$

where we assumed that the elements in $\mathbf{h}$ and $\mathbf{y}$ are independent, hence so are L_r, L_1, L_2. Applying the probability density function (pdf) of L_i, L_r [11]

$$\begin{aligned} f_{L_i}(x) &= \frac{\exp((x-|x|\sqrt{1+\gamma_{s,i}^{-1}})/2)}{4\sqrt{\gamma_s(1+\gamma_s)}} && (13)\\ f_{L_r}(x) &= \frac{\exp((x-|x|\sqrt{1+\gamma_r^{-1}})/2)}{4\sqrt{\gamma_r(1+\gamma_r)}} && (14)\end{aligned}$$

in (12) yields

$$P_{FA} = \frac{1}{2}\left[1-\left(\frac{\gamma_r\gamma_{s,1}\gamma_{s,2}}{(1+\gamma_r)(1+\gamma_{s,1})(1+\gamma_{s,2})}\right)^{\frac{1}{2}}\right] \quad (15)$$

where $\gamma_{s,i} = d_i^{-\alpha}E_s/N_0$ and $\gamma_r = d_r^{-\alpha}E_r/N_0$ are the received symbol SNR at the destination from the i-th source and relay, respectively. Similarly, it can be shown that the probability of misdetection is given by

$$\begin{aligned} P_{MD} &= P(\hat{z}=+1|z=-1)\\ &= P_{FA} && (16)\end{aligned}$$

V. Decoding Schemes

We consider three decoding schemes, depending on the availability of side information regarding z.

A. MAP Decoder

The MAP decoder calculates the LLR of x_1 (or x_2) given $\mathbf{h}$ and $\mathbf{y}$, $L(x_1|\mathbf{h},\mathbf{y})$, and decide $\hat{x}_1=+1$ if $L(x_1|\mathbf{h},\mathbf{y})>0$ and $\hat{x}_1=-1$ otherwise. It can be shown that $L(x_1|\mathbf{h},\mathbf{y}) = L(x_1|h_1,y_1)+L(x_1|h_r,y_r,h_2,y_2)$ where

$$\begin{aligned} &L(x_1|h_r,y_r,h_2,y_2) =\\ &\ln\frac{e^{L(x_2\oplus p|h_r,y_r,h_2,y_2)}P(z=+1)+P(z=-1)}{e^{L(x_2\oplus p|h_r,y_r,h_2,y_2)}P(z=-1)+P(z=+1)} && (17)\end{aligned}$$

and

$$P(z=+1) = E_{\mathbf{h},\mathbf{y}}\left[e^{L(z|\mathbf{h},\mathbf{y})}/(1+e^{L(z|\mathbf{h},\mathbf{y})})\right] \quad (18)$$

The statistical average in (18) may be replaced by the time average. Note that the MAP decoder requires the side information $L(z|\mathbf{h},\mathbf{y})$ in (5) which represents the likelihood of the presence of relay misbehavior.

B. Genie-aided Decoder

The genie-aided decoder assumes perfect information regarding z, and chooses the most probable codeword from the set of transmitted codewords. Although the genie-aided decoding may not be realistic, it may serve as a reference for performance comparison with other decoders.

C. Conventional Decoder

The conventional decoder does not utilize the side information regarding z. Hence, the decoder considers (x_1, x_2, p_t), where $p_t = x_1 \oplus x_2$, as a valid codeword no matter what z is. Therefore, if $z = -1$, the conventional decoder considers (x_1, x_2, p_t) as a valid codeword while $(x_1, x_2, -p_t)$ is valid. If $z = 1$, it considers (x_1, x_2, p_t) as a valid codeword while (x_1, x_2, p_t) is valid.

VI. Numerical Results and Discussions

Fig. 1 shows the probability of false alarm P_{FA} and misdetection P_{MD} versus the fraction of energy allocated to each source, $\alpha = \gamma_s/\gamma_b$, where $\gamma_b = (2\gamma_s + \gamma_r)/2$ is the received SNR per information bit. We find that the optimum α that minimizes P_{FA} and P_{MD} is 2/3. In general, for K sources and R relays, the optimum α can be shown to be $K/(K + \sqrt{R})$.

Fig. 2 shows the probability of bit error versus the received SNR per information bit, γ_b. We notice that the MAP decoder that exploits the side information $L(z|\mathbf{h}, \mathbf{y})$ provides a dramatic improvement over the conventional decoder that does not exploit the side information and that the probability of bit error for the latter exhibits an error floor. This shows that it is crucial to exploit the side information regarding the presence of relay misbehavior at the decoder to effectively mitigate the relay misbehavior. Also shown in the figure is the probability of bit error with the genie-aided decoder that knows z exactly. We can see that the diversity order is two, and therefore, the relay misbehavior can completely be mitigated with the genie-aided decoder.

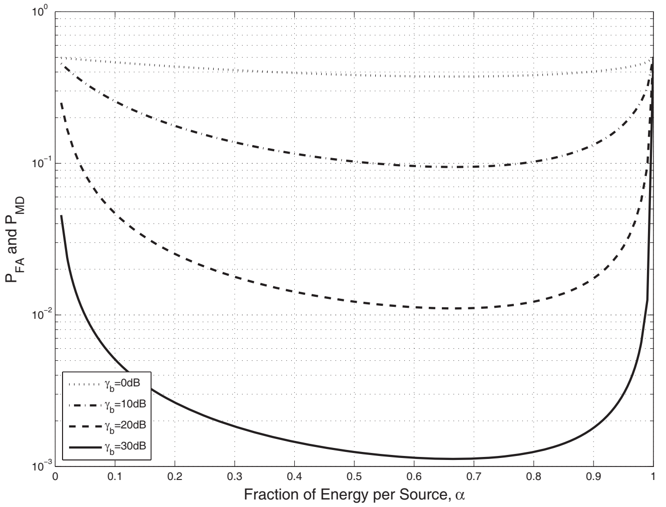

Fig. 1. Probability of false alarm and misdetection versus fraction of energy per source.

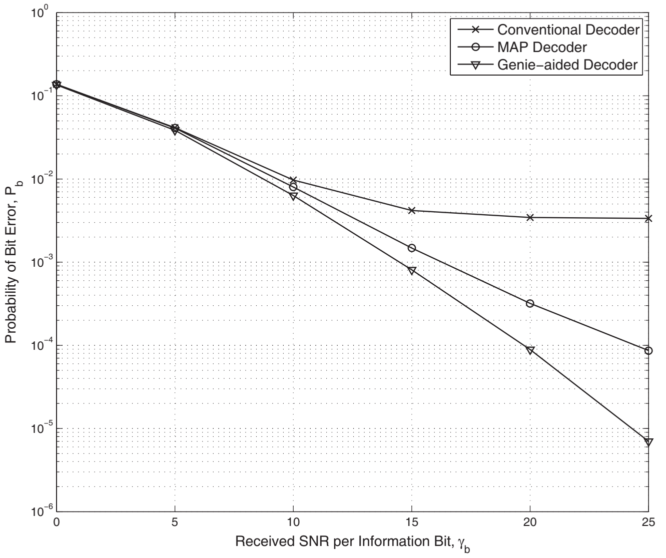

Fig. 2. Probability of bit error versus received SNR per information bit (dB); $\alpha = 2/3$, $P(z = -1) = 10^{-2}$.

VII. Conclusions

We proposed the MAP approach in detecting the misbehaving relay that injects false data or adds channel errors into the network encoder in multi-access relay networks. The proposed scheme does not require sending extra bits at the source, and therefore, there is no transmission overhead. In addition, it is optimal in the sense of minimizing the probability of false alarm and misdetection. We derived the probability of false alarm and misdetection, taking into account the lossy nature of wireless links. We also presented the MAP decoding scheme that exploits the side information regarding the presence of relay misbehavior and showed that it can effectively mitigate the relay misbehavior.

Acknowledgement

This work was supported in part by the U.S. National Science Foundation (NSF) under grant ECCS 0824237.

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory* , vol. 46, pp. 1204–1216, 2000.

[2] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Networking*, vol. 11, pp. 782–795, 2003.

[3] C. Fragouli, J.-Y. Le Boudec, and J. Widmer, "Network coding: an instant primer," *ACM SIGCOMM Comput. Commun. Rev.*, pp. 63–68, 2006.

[4] F. Zhao, T. Kalkert, M. Medard, and K. J. Han, "Signatures for content distribution with network coding, in *Proc. IEEE International Symposium on Information Theory*, June 2007.

[5] T. Ho, B. Leong, R Koetter, M. Medard, M. Effros, and D. R. Karger, "Byzantine modification detection in multicast networks with random network coding, *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2798-2803, June 2008.

[6] S. Jaggi, M. Langberg, S. Katti, T. Ho, D. Katabi, M. Medard, and M. Effros, "Resilient network coding in the presence of Byzantine adversaries, *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2596-2603, June 2008.

[7] Y. Mao and M. Wu, "Tracing malicious relays in cooperative wireless communications, *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 2, pp. 198-212, June 2007.

[8] T. Khalaf and S. W. Kim, "Error analysis in multi-source, multi-relay, multi-destination networks under falsified data injection attacks," in *Proc. IEEE MILCOM*, 2008.

[9] M. Kim, M. Medard, J. Barros, and R. Kotter, "An algebraic watchdog for wireless network coding, in *Proc. IEEE International Symposium on Information Theory*, June 2009.

[10] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 429–445, 1996.

[11] S. W. Kim and E. Y. Kim, "Optimum receive antenna selection minimizing error probability," in *Proc. IEEE Wireless Communications and Networking Conference*, vol. 1, pp. 441–447, Mar. 2003.