ASYMPTOTIC PERFORMANCE ANALYSIS FOR MINIMUM-HAMMING-DISTANCE FUSION

Po-Ning Chen[†], *Tsang-Yi Wang*[‡], *Yunghsiang S. Han*[*], *Pramod K. Varshney*[§] *and Chien Yao*[¶]

[†]Depart. of Communications Eng., National Chiao-Tung Univ., Taiwan, ROC
[‡]Graduate Institute of Communication Eng., National Chi Nan Univ., Taiwan, ROC
[*]Graduate Institute of Communication Eng., National Taipei Univ., Taiwan, ROC
[§]Electrical Engineering and Computer Science, Syracuse University, NY, USA
[¶]Depart. of Electronics Eng., National Chiao-Tung Univ., Taiwan, ROC

ABSTRACT

Distributed (M-ary) detection and fault-tolerance have been considered as two fundamental functions in the context of large-scale sensor networks. Distributed multiclass classification fusion using error correcting codes (DCFECC) has been proposed to provide good fault-tolerance capability in wireless sensor networks. Minimum Hamming distance fusion is an essential part of the DCFECC approach. In this paper, we study the asymptotic performance of minimum Hamming distance fusion for both fault-free and faulty situations when the number of sensors tends to infinity. We conclude that the error probability vanishes asymptotically as long as the minimum Hamming distance $d_{\min}$ of the DCFECC code approaches infinity, and the probabilities of correct local classification for all hypotheses are greater than one half. In case $d_{\min}/2$, normalized by the number of sensors, can be made larger than the largest local classification error, an explicit expression for the error exponent of the DCFECC system in terms of the Kullback-Leibler divergence can be established. A converse where the DCFECC decoding error is bounded away from zero is also addressed.

1. INTRODUCTION

There has been a great deal of recent interest in the notion of deploying large number of networked sensors for event or target classification in wireless sensor networks (WSNs). In WSNs, fault-tolerance capability is critical since sensors can be easily damaged or run out of battery energy [1].

Recently, a distributed classification fusion approach using error correcting codes (DCFECC) has been proposed to provide fault-tolerance capability [2, 3, 4]. In the proposed approach, an error-correcting code matrix is first designed by either simulated annealing or cyclic column replacement [3]. Each codeword forms a row in the code matrix, and corresponds to one of the hypotheses. Each local sensor employs a local classification rule to decide on the hypothesis present, and outputs the respective codebit according to the corresponding column in the code matrix. Upon receipt of the binary received vector, the fusion center makes a multiclass decision by performing minimum Hamming distance decoding (the fault-tolerant fusion rule). Unlike the conventional approach that employs the optimal fusion rule (MAP rule), such a fault-tolerant fusion rule can provide enough distance among all hypothesis acceptance regions by means of any feasible decoding algorithm. Therefore, the local decision vectors could still fall into the correct acceptance region even if several sensor faults are present.

In this paper, we characterize the performance of minimum Hamming distance fusion when the number of sensors is arbitrarily large. Asymptotic performance analysis for the distributed binary detection problem has been investigated in [5, 6]. However, these asymptotic results were considered from different perspectives. First, the MAP fusion rule was used for the parallel fusion network in [5, 6]. Secondly, the optimality under the identical sensor assumption was defined based on the error exponent. The results obtained in this paper show that the probability of error for minimum Hamming distance fusion can be upper-bounded by a quantity determined by the minimum Hamming distance $d_{\min}$ of the DCFECC code, and the largest local classification error $p_{\max}$ among all hypotheses. As a consequence, the DCFECC decoding error vanishes as long as $p_{\max} < 1/2$ and $d_{\min}$ approaches infinity. In situations where $d_{\min}/(2N)$ can be made larger than $p_{\max}$, where N is the number of sensors, the DCFECC decoding error approaches zero exponentially fast. The above analytical results are then used to characterize the relationship between $d_{\min}$ and the fault-tolerance capability of the DCFECC system.

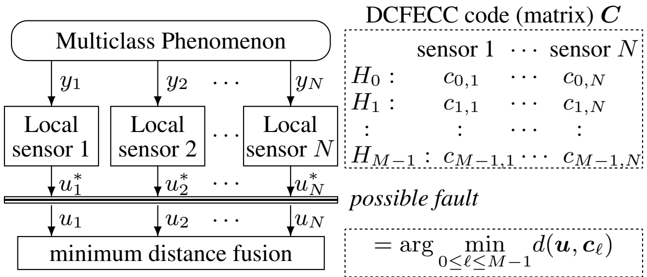

Fig. 1. Distributed classification system architecture

2. SYSTEM MODEL

The distributed M-ary classification system considered in this paper is depicted in Fig. 1.

In our system model, the local observations $\{y_j\}_{j=1}^N$ are assumed to be conditionally independent and identically distributed across sensors given each hypothesis. Also assume that each local sensor classifies its own observation, independent of all others, to one of the M hypotheses using an identical rule.[1] Thus, the probability of classifying H_j given that H_i is the true hypothesis is the same for all the sensors, and is denoted by $h_{j,i}$. Moreover, equally likely hypotheses assumption is made.

In the DCFECC approach, a code matrix $\boldsymbol{C}$ is specified in advance. This code matrix is an $M \times N$ matrix with element $c_{\ell,j} \in \{0,1\}$, where $\ell = 0,\ldots,M-1$ and $j = 1,\ldots,N$. Each hypothesis H_ℓ is associated with a row in the code matrix. Each column of $\boldsymbol{C}$ stands for the local binary outputs corresponding to the locally classified hypotheses at the respective sensor. Thus, sensor j transmits $c_{\ell,j}$, if H_ℓ is declared present locally. For notational convenience, $\boldsymbol{c}_\ell \triangleq (c_{\ell,1},\, c_{\ell,2},\, \ldots,\, c_{\ell,N})$ is used to denote the row of $\boldsymbol{C}$ corresponding to the hypothesis H_ℓ.

After the observation is locally processed, the local output codebit u_j^* is transmitted to the fusion center. Due to possible sensor faults (or channel transmission errors), the fusion center receives the word $\boldsymbol{u} = (u_1, u_2, \ldots, u_N)$ where u_j may or may not equal u_j^*. Finally, the minimum distance fusion rule, or specifically, $\omega = \arg\min_{0\le\ell\le M-1} d(\boldsymbol{u}, \boldsymbol{c}_\ell)$, where $d(\cdot,\cdot)$ is the Hamming distance, is employed to obtain the multiclass decision ω. The tie-break rule is to randomly pick a codeword from those with the same smallest Hamming distance to the received vector.

[1] For DCFECC given in [3], the classifier at each local sensor is jointly optimized with the global fusion rule (decoding rule) at the fusion center based upon the specified code matrix. Thus, classifiers at the local sensors might be different from each other. In this work, in order to simplify the analysis, we assume that all the local sensors use the same classifier.

3. ASYMPTOTIC PERFORMANCE ANALYSIS

In this section, the performance of the DCFECC approach is analyzed in a large scale sensor network environment. Results for both fault-free and faulty situations are provided.

3.1. Fault-free situation

Let $d_{\min} \triangleq \min_{i\neq j} d(\boldsymbol{c}_i, \boldsymbol{c}_j)$, $\delta_N \triangleq d_{\min}/(2N)$, $p_i \triangleq 1 - h_{i,i}$, and $p_{\max} \triangleq \max_{0\le i\le M-1} p_i$. The performance of minimum Hamming distance fusion can be bounded above according to the next theorem.

Theorem 1 If $p_{\max} < 1/2$, then the average probability of error satisfies

$$P_e \le (M-1)\left(\sqrt{4p_{\max}(1-p_{\max})}\right)^{d_{\min}},$$

where $P_e \triangleq \frac{1}{M}\sum_{i=0}^{M-1} \Pr(\text{decision} \neq H_i | H_i)$.

Proof:

$$\begin{aligned}
&\Pr(\text{decision} \neq H_i | H_i) \\
&\quad\le \Pr\left(d(\boldsymbol{u}, \boldsymbol{c}_i) \ge \min_{0\le\ell\le M-1, \ell\neq i} d(\boldsymbol{u}, \boldsymbol{c}_\ell) \,\middle|\, H_i \right) \\
&\quad\le \sum_{0\le\ell\le M-1, \ell\neq i} \Pr\left(d(\boldsymbol{u}, \boldsymbol{c}_i) \ge d(\boldsymbol{u}, \boldsymbol{c}_\ell) \middle| H_i \right) \\
&\quad= \sum_{0\le\ell\le M-1, \ell\neq i} \\
&\qquad \Pr\left(\sum_{\{j\in[1,\cdots,N]\,:\,c_{\ell,j}\neq c_{i,j}\}} (z_{i,j} - \bar{z}_{i,j}) \ge 0 \,\middle|\, H_i \right),
\end{aligned}$$

where $z_{i,j} = u_j \oplus c_{i,j}$ and $\bar{z}$ represents the complement of the binary 0-1 variable z. Observe that

$$\Pr(z_{i,j} = 1 | H_i) = \sum_{k=0}^{M-1} (c_{i,j} \oplus c_{k,j}) h_{k,i} \le p_i < 1/2,$$

and $\{z_{i,j}\}_{j=1}^N$ are i.i.d. sequences given H_i is true. Therefore,[2]

$$\Pr\left(\sum_{\{j\in[1,\ldots,N]:c_{\ell,j}\neq c_{i,j}\}} (z_{i,j}-\bar{z}_{i,j})>0 \middle| H_i\right) \leq \left(\sqrt{4p_i(1-p_i)}\right)^{d(\mathbf{c}_\ell,\mathbf{c}_i)},$$

which results in

$$\begin{aligned} P_e &= \frac{1}{M}\sum_{i=0}^{M-1}\Pr(\text{decision}\neq H_i|H_i) \\ &\leq \frac{1}{M}\sum_{i=0}^{M-1}\sum_{0\leq\ell\leq M-1,\ell\neq i}\left(\sqrt{4p_i(1-p_i)}\right)^{d(\mathbf{c}_\ell,\mathbf{c}_i)} \\ &\leq (M-1)\left(\sqrt{4p_{\max}(1-p_{\max})}\right)^{d_{\min}}. \end{aligned}$$

□

Based on the above theorem, the average probability of error can be bounded above by a quantity depending on the accuracy of local classification. In other words, if local classification is such that $p_{max} < \frac{1}{2}$, one can make the DCFECC decoding error vanish as $d_{\min}$ approaches infinity. It is reasonable to expect a converse statement in that when the performance of local classifiers is poor ($p_{\max} > \frac{1}{2}$), it will be hard to compensate for them by sensor fusion even in the fault-free situation. This conclusion should also be applicable to the system that employs the optimal MAP fusion rule.

Theorem 1 also shows that the error exponent of minimum Hamming distance fusion is no smaller than $-\delta_N \log(4p_{\max}(1-p_{\max}))$. The next theorem indicates that another error exponent expression in terms of the Kullback-Leibler divergence can be given if a code with $\delta_N > p_{\max}$ is employed. Due to the page limitation, the proof of the next theorem is omitted.

Theorem 2 If $\delta_N > p_{\max}$, then

$$P_e \leq \exp\left\{-N\min_{0\leq i\leq M-1} D(\delta_N\|p_i)\right\},$$

where

$$D(x\|y) \triangleq x\log(x/y)+(1-x)\log[(1-x)/(1-y)]$$

is the binary Kullback-Leibler divergence function. In addition, if $\liminf_{N\to\infty}\delta_N > \delta > p_{\max}$, then for N sufficiently large,

$$P_e \leq \exp\left\{-N\min_{0\leq i\leq M-1} D(\delta\|p_i)\right\}.$$

It needs to be pointed out that although, according to Theorem 1, the probability of error can be bounded, and the bound goes to zero if $p_{\max} < 1/2$ and $d_{\min}$ approaches infinity, it is not necessarily true that if $p_{\max} > 1/2$, P_e is bounded away from zero. Indeed, it is only when the code happens to use poor local classifications infinitely many times that a bounded-from-below decoding error at the fusion center will occur. For instance, this will happen when $c_{i,j}\neq c_{\ell,j}$ for all $\ell\neq i$ occurs infinitely many times in j for some $p_i > 1/2$. A more specific condition under which the fusion error probalility is bounded from below is given in the next theorem.

Theorem 3 P_e is bounded away from zero, if

$$\liminf_{N\to\infty}\max_{0\leq i,\ell\leq M-1}\sum_{\{j\in[1,\ldots,N]\,:\,c_{\ell,j}\neq c_{i,j}\}}\left(2\sum_{k=0}^{M-1}(c_{i,j}\oplus c_{k,j})h_{k,i}-1\right)>0.$$

Proof: The condition given in Theorem 3 implies that for N sufficiently large, there exists an $i=i(N)$ and an $\ell=\ell(N)$,[3] such that

$$\sum_{\{j\in[1,\ldots,N]\,:\,c_{\ell,j}\neq c_{i,j}\}}\left(2\sum_{k=0}^{M-1}(c_{i,j}\oplus c_{k,j})h_{k,i}-1\right)>0,$$

which implies that

$$\begin{aligned} m_{\ell,i} &\triangleq E\left[\sum_{\{j\in[1,\cdots,N]\,:\,c_{\ell,j}\neq c_{i,j}\}}(z_{i,j}-\bar{z}_{i,j})\right] \\ &= \sum_{\{j\in[1,\ldots,N]\,:\,c_{\ell,j}\neq c_{i,j}\}}\left(2\sum_{k=0}^{M-1}(c_{i,j}\oplus c_{k,j})h_{k,i}-1\right) \\ &> 0. \end{aligned}$$

[2] Assume independent $\{Y_j\}_{j=1}^m$ with $\Pr(Y_j=+1)=1-\Pr(Y_j=-1)=q_j$. Let $q_{\max}\triangleq\max_{1\leq j\leq m}q_j$ and assume $q_{\max}<1/2$. Then

$$\begin{aligned} \Pr(Y_1+\cdots+Y_m\geq 0) &= \Pr\left(e^{s(Y_1+\cdots+Y_m)}\geq 1\right) \quad \text{for } s>0 \\ &\leq \prod_{j=1}^m E\left[e^{sY_j}\right]=\prod_{j=1}^m\left(q_je^s+(1-q_j)e^{-s}\right) \\ &\leq \left(q_{\max}e^s+(1-q_{\max})e^{-s}\right)^m. \end{aligned}$$

Taking $s>0$ to satisfy $e^{2s}=(1-q_{\max})/q_{\max}>1$, we obtain:

$$\Pr(Y_1+\cdots+Y_m\geq 0) \leq \left(\sqrt{4q_{\max}(1-q_{\max})}\right)^m.$$

[3] We omit their dependence of i and ℓ on N for notational convenience.

As a result,

$$\begin{aligned}
&\Pr(\text{decision} \neq H_i|H_i) \\
&\quad \geq \quad \Pr\left(d(\boldsymbol{u},\boldsymbol{c}_i) > \min_{0\leq \ell \leq M-1, \ell \neq i} d(\boldsymbol{u},\boldsymbol{c}_\ell) \middle| H_i \right) \\
&\quad \geq \quad \Pr\left(d(\boldsymbol{u},\boldsymbol{c}_i) > d(\boldsymbol{u},\boldsymbol{c}_\ell) \middle| H_i \right) \\
&\quad = \quad \Pr\left(\sum_{\{j\in[1,\cdots,N]\,:\,c_{\ell,j}\neq c_{i,j}\}} (z_{i,j}-\bar{z}_{i,j}) > 0 \middle| H_i \right) \\
&\quad \geq \quad \Pr\left(\sum_{\{j\in[1,\cdots,N]\,:\,c_{\ell,j}\neq c_{i,j}\}} (z_{i,j}-\bar{z}_{i,j}) - m_{\ell,i} > 0 \middle| H_i \right) \\
&\quad \to \quad \frac{1}{2}, \text{ if } d(\boldsymbol{c}_\ell,\boldsymbol{c}_i) \text{ approaches infinity,}
\end{aligned}$$

where the last step follows from the central limit theorem for the sum of independent bounded variables. Thus, the claim of the theorem holds for the case that $d(\boldsymbol{c}_\ell,\boldsymbol{c}_i)$ tends to infinity. In the situation when $d(\boldsymbol{c}_\ell,\boldsymbol{c}_i)$ is bounded, the theorem is trivially valid. □

3.2. Fault-tolerance capability with possible faulty sensors

The network considered is likely to contain faulty sensor nodes due to harsh environmental conditions. Faults may include all types of misbehaviors, ranging from simple sensor crash faults, stuck-at faults, to sensors that behave arbitrarily or maliciously. Thus, by assuming that there are μ misbehaving sensors, we obtain that the new $\bar{d}_{\min}$, defined as the resultant effective minimum Hamming distance at the fusion center after considering the faulty sensors, should satisfy:

$$d_{\min} \geq \bar{d}_{\min} \geq d_{\min} - \mu.$$

Therefore, as long as $p_{\max} < 1/2$ and $\bar{d}_{\min}$ approaches infinity linearly, the probability of error goes to zero exponentially by directly following the result of Theorem 1.

Similarly, the resultant new $\bar{\delta}_N \triangleq \bar{d}_{\min}/(2N)$ should also satisfy:

$$\delta_N \geq \bar{\delta}_N \geq \delta_N - \nu_N,$$

where $\nu_N \triangleq \mu/(2N)$. By assuming that $\nu_N \to \nu$, we have

$$\delta \geq \limsup_N \bar{\delta}_N \geq \liminf_N \bar{\delta}_N \geq \delta - \nu.$$

Thus, following from Theorem 2, we know that if $p_{\max} < \delta - \nu$, then the probability of DCFECC decision error approaches zero exponentially as N goes to infinity. Therefore, $(\delta - p_{\max})$ can be viewed as a *sufficient* fault-tolerance capability of the DCFECC code in the sense that under vanishing error requirement, the system can tolerate around $2N(\delta - p_{\max})$ misbehaving sensors.

4. CONCLUSIONS

In this paper, the performance analysis of the minimum-Hamming-distance fusion rule, when the number of sensors becomes arbitrarily large, is investigated. The results show, based on identical local decision rules, that error probability vanishes asymptotically when the Hamming distance of the DCFECC code approaches infinity, and the probabilities of correct local classification for all the hypotheses are greater than one half. Furthermore, the probability of error approaches zero asymptotically as long as the ratio of the minimum Hamming distance of the employed code and twice the number of sensors is greater than the largest probability, of the sensor deciding on the wrong hypothesis given a particular hypothesis. Moreover, the minimum Hamming distance of employed code in this fusion rule is shown to have a relation with the fault-tolerance capability of the system.

5. REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, pp. 102–114, August 2002.

[2] T.-Y. Wang, P. K. Varshney, and Y. S. Han, "Distributed fault-tolerant classification fusion using error correcting codes," in *Proceedings of SPIE on Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2003*, April 2003, pp. 47–57.

[3] T.-Y. Wang, Y. S. Han, and P. K. Varshney, "Fault-tolerant classification in multisensor networks using coding theory," in *The 6th International Conference on Information Fusion*, vol. 2, July 2003, pp. 772–779.

[4] ——, "A combined decision fusion and channel coding scheme for fault-tolerant classification in wireless sensor networks," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing 2004*, vol. 2, May 2004, pp. 1073 – 1076.

[5] P.-N. Chen and A. Papamarcou, "New asymtotic results in parallel distributed detection," *IEEE Trans. Inform. Theory*, vol. 39, no. 6, pp. 1847–1863, November 1993.

[6] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Conrol, Signals, and Systems*, vol. 1, no. 2, pp. 167–182, 1988.