An Efficient Data Aggregation Scheme for Multihop Wireless Sensor Network

Bhushan G. Jagyasi, Bikash K. Dey, S. N. Merchant, U. B. Desai
SPANN Laboratory, Electrical Engineering Dept., IIT Bombay, India
Email: {bhushan,bikash,merchant,ubdesai}@ee.iitb.ac.in

Abstract

We consider the problem of data fusion in wireless sensor network (WSN) for event detection application. For such an application, maximizing accuracy and network lifetime are the two primary requirements in the design of WSNs. In [1], an optimum one bit data fusion rule, called Chair-Varshney (CV) rule, has been proposed which is derived from an optimum likelihood ratio test, for a star topology. This requires one bit of local decision transmission from every node to the base station which in turn makes the one bit final decision about the occurrence of the event. Here, we propose an efficient way to extend CV rule for tree topology which is more realistic topology resulting in greater scalability and coverage.

I. Introduction

Wireless Sensor Network (WSN) [2] has come into existence with the development of low power and small size sensor nodes. These sensor nodes can sense, process and communicate information among themselves to collaboratively perform a particular task. We consider event detection using wireless sensor network (WSN) [3], [4]. The event of interest could be an intruder crossing the line of control, rise of ambient temperature above some threshold, camera detecting a burglar in a sensitive area, disaster control applications like detection and prediction of Landslide, Volcano, Tsunami or Earthquake [5]–[8]. The problem of event detection is a binary hypothesis testing problem in which the event hypothesis H takes two values $\{0, 1\}$ indicating the non-occurrence and the occurrence of the event.

For detecting an event accurately in a large area, we prefer the deployment of large number of inexpensive, albeit less precise sensor nodes in an application area rather than having a few expensive and more precise sensor nodes. For less precise sensors we assume that either 1 bit or small number of bits are used to quantize the information sensed by them. In the former case, the i^{th} sensor S_i observes $Y_i \in \{0, 1\}$, which contains information about event hypothesis as observed by the sensor. In the latter case, along with information about the occurrence of the event, sensors also have the information about the confidence on their observation about the event. For example, a sensor may use 2 bits to quantize the sensed data and thus it observes $Y_i \in \{0, 1, 2, 3\}$. Here, $Y_i = 0$ and $Y_i = 1$ indicate the non-occurrence of the event with more and less confidence respectively, and $Y_i = 2$ and $Y_i = 3$ indicate the observation of occurrence of the event with less and more confidence respectively.

As in our previous work [3], [4], here we focus on the problem of efficient data fusion while reducing the number of bits transmitted in the wireless sensor network. Distributed detection using multiple sensors with various network topologies were considered in [9], [10]. Communication is considered to require more power per bit as compared to computation and sensing [11]. Thus, to improve upon the network lifetime, in [1], [9], [10], [12]–[15] the observation of each sensor and the communication to its parent was considered to be one bit. In [1], [13]–[15], authors consider target detection application in which local sensors have different performance indices, measured in terms of probability of detection (pd) and the probability of false alarm (pf). The popular Chair-Varshney (CV) rule [1] gives an optimum way to fuse the one bit information received at fusion center from every sensor node. This requires knowledge of performance indices (pds and pfs) of all the sensors nodes. This gives excellent result as compared to the traditional counting rule for a star topology.

However, WSN with star topology has restricted scalability in terms of area covered by the nodes. This is because nodes deployed far apart from base station may not be within its communication range or they may require more power to transmit to base station. In tree topology, the nodes lying far from base station needs to transmit to their nearest node and not directly to the base station. This significantly reduces the communication cost. Here every node can be considered as a fusion center and can be used to do some processing on the received data before transmitting, while in centralized star topology, the processing capability of nodes is not leveraged. Therefore, in our work we concentrate on aggregation schemes for tree topology.

In [16], aggregation scheme based on majority rule with a specific cost for routing was proposed for tree topology. This combination of aggregation scheme and routing results in a cross layer optimization. But this approach does not result in much gain in accuracy. In [3], we proposed a one bit weighted aggregation scheme (WAS) for tree topology which results in an improvement in accuracy as compared to the scheme presented in [16]. However, in WAS each node required additional knowledge of its children's number of descendants, and the results obtained were also suboptimal.

In this paper, we show the extension of CV rule requiring one bit transmission in a star topology [1], to a tree topology. Numerical results are presented which show a significant improvement in the accuracy of CV rule over other one bit aggregation schemes, for a tree topology.

In Section II, we extend CV rule to WSN with tree topology and present numerical results for the same. Here, we present numerical results to compare various one bit aggregation schemes for tree topology. Finally, in Section III we conclude the paper.

II. CV RULE FOR TREE TOPOLOGY

In our work we focus on an efficient multibit decision fusion rule for a tree topology. In order to get a better understanding of the multibit scenario, this section discusses the extension of CV rule, requiring one bit transmission, for tree topology.

We consider the occurrence of the equiprobable binary event with hypothesis $H \in \{0, 1\}$ in an application area. Let N number of sensor nodes be deployed with uniform distribution in the application area where an event may or may not occur. Consider that each sensor S_i is capable of observing one bit quantized data $Y_i \in \{0, 1\}$ about the occurrence of the event H with some known probability of detection pd_i and probability of false alarm pf_i. Here $pd_i = P(Y_i = 1|H = 1)$ and $pf_i = P(Y_i = 1|H = 0)$.

In [1], [14], CV rule is presented as optimum fusion rule for star topology (see Fig. 1). Fusion center S_0 makes one bit decision X_0 about the occurrence of event H based on one bit decision X_i received from every sensor node S_i for $i = 1, 2, \cdots, K$. Here $X_i = Y_i$, since decision of a node is its observation. To implement the CV rule, the fusion center needs to know the probability of detection Pd_i and probability of false alarm Pf_i of the received data X_i from each sensor S_i. Here, $Pd_i = P(X_i = 1|H = 1)$ and $Pf_i = P(X_i = 1|H = 0)$. Since every node transmits what it senses, i.e. ($X_i = Y_i$), the performance indices are also the same, i.e. $Pd_i = pd_i$ and $Pf_i = pf_i$. The C-V fusion rule is a threshold test for

$$\Lambda_{CV} = \sum_{i=1}^{K} [X_i \log \frac{Pd_i}{Pf_i} + (1 - X_i) \log \frac{1 - Pd_i}{1 - Pf_i}]. \tag{1}$$

The decision X_0 of S_0 can be obtained by

$$\Lambda_{CV} \underset{X_0=0}{\overset{X_0=1}{\gtrless}} T. \tag{2}$$

In this section, we show the extension of CV rule for the tree topology (see Fig. 2). Here, every leaf node has performance indices, Pd and Pf same as pd and pf respectively. In tree topology, starting from leaf node and going towards sink node, every node S_i will compute the probability of detection Pd_i and probability of false alarm Pf_i of its decision X_i based on the Pds, Pfs of its children nodes. It will then transmit this information to its parent which in turn calculates its decisions Pd and Pf, and

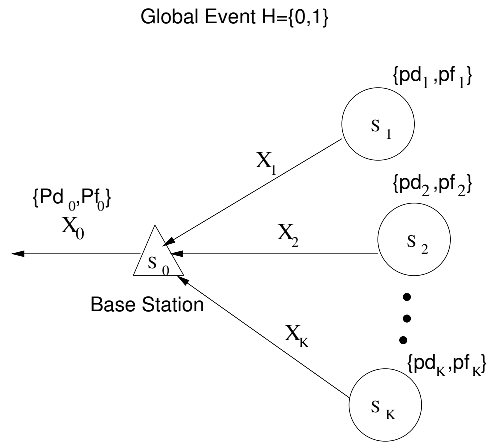

Fig. 1. Star Topology

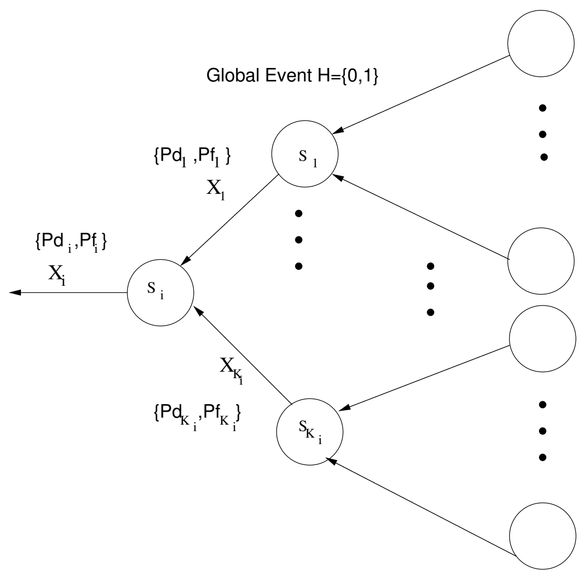

Fig. 2. A local view of the tree topology

transmits them to its parent. This repeats till the sink node. Thus for static topology these computations and transmissions happen only once during the initial setup. However, for a dynamic topology, computations of the performance indices and their transmissions to the parent node happens repeatedly depending upon how dynamic the network is.

Next we calculate Pd_i and Pf_i at any node S_i having K_i children $S_1, S_2, \cdots, S_{K_i}$. Here we assume S_i is aware of probability of detection Pd_j and probability of false alarm Pf_j of its every child S_j. Let $\underline{X} = [X_1, X_2, \cdots, X_{K_i}]$, and $\mathcal{X}$ be the set of all possible bit patterns of $\underline{X}$ for which $\Lambda_{CV} \geq T$. Now probability of detection Pd_i and probability of false alarm Pf_i of the decision X_i can be computed by

$$\begin{aligned} Pd_i &= P(\Lambda_{CV} \geq T | H = 1) \\ &= \sum_{\mathcal{X}} \prod_{j=1}^{K_i} (X_j Pd_j + (1 - X_j)(1 - Pd_j)) \end{aligned} \tag{3}$$

and

$$Pf_i = P(\Lambda_{CV} \geq T | H = 0)$$

$$= \sum_{\mathcal{X}} \prod_{j=1}^{K} (X_j Pf_j + (1 - X_j)(1 - Pf_j)). \quad (4)$$

Thus for the tree topology, every node acts as a fusion center. Every node makes its local decision based on CV rule using equation (1) and (2). It also computes the Pd and Pf of its local decision using equation (3) and (4). These decision and probabilities are then communicated to their parent node. The process of data (decision and probabilities) transfer starts from leaf node and completes at the sink node.

We are now interested to determine the *system level* probability of detection (PD) and probability of false alarm (PF) as obtained by sink node in the tree topology. Let the final decision made by sink node be X. We thus define $PD = P(X = 1|H = 1)$ and $PF = P(X = 1|H = 0)$.

PD vs. PF plots, also called as Receiver Operating Characteristic (ROC), for different fusion rules in star topology is shown in [14]. ROC is obtained by varying threshold T and it represents the tradeoff between PD and PF. We define $Accuracy = 0.5 * PD + 0.5 * (1 - PF)$ and consider it as a performance evaluating parameter. This is a fair parameter to evaluate the performance for the given problem, since we have assumed equiprobable event hypothesis. Thus an increase in PD or decrease in PF results in increase of accuracy by the same amount.

Here in tree topology as every node is the fusion center, we have two options either to fix the threshold for all the nodes or to have variable thresholds for different nodes.

1) CV with Fixed Threshold ($CV - FT$): Here the threshold T at every node is fixed and can be decided based on the point in the ROC curve in which we want to operate. Here, we would be basically interested in the threshold resulting in the best possible accuracy.
2) CV with Variable Threshold ($CV - VT$): Here every sensor node S_i throughout the tree uses different threshold T_i. The optimum thresholds $\underline{T} = [T_1, \cdots, T_N]$ are those which results in the maximum system accuracy. However, it is difficult to get such a set of optimal thresholds. Thus we do the local optimization at each parent node by finding a threshold which maximizes the accuracy of the decision made by that parent node. This also has an impact on overall increase in the system accuracy.

A. Numerical Results

We present next the results for different aggregation schemes for tree topology with $N = 100$ nodes deployed with uniform distribution. Here all sensors are considered to have same precision p. We vary the precision p of the nodes from 0.55 to 0.95 and look at the performance of various aggregation schemes. Fig. 3 and Fig. 4 show the computed results for accuracy and performance indices respectively for Counting rule, Weighted Aggregation Scheme (WAS) [3], Chair Varshney with Fixed Threshold ($CV - FT$) and Chair Varshney with Variable Threshold ($CV - VT$). It is clear from the accuracy plot (Fig. 3) that $CV - VT$ and $CV - FT$ both perform equally well in terms of accuracy as compared to the previously proposed one bit aggregation scheme for tree topology. Thus, we recommend the use of $CV - FT$, which is more easily implementable, instead of the more complicated and computationally extensive $CV - VT$, given that both have comparable accuracies. However, there is a slight difference in the PD, PF plots (Fig. 4) of these two schemes. This is because of the chosen threshold for $CV - FT$ which can be varied to get similar values of PD, PF for both the schemes.

III. Conclusion

We considered the problem of event detection using wireless sensor network. We suggested the use of tree topology instead of star topology for various reasons like scalability and coverage. The contribution of the paper was to extend the existing optimum one bit decision fusion rule called CV rule for the tree topology. We have shown that this can be done in two ways either by fixing the same threshold at every node ($CV - FT$) or by using variable thresholds at different nodes ($CV - VT$). We have shown by numerical results that $CV - FT$ performs almost as good as $CV - VT$ while they both perform better

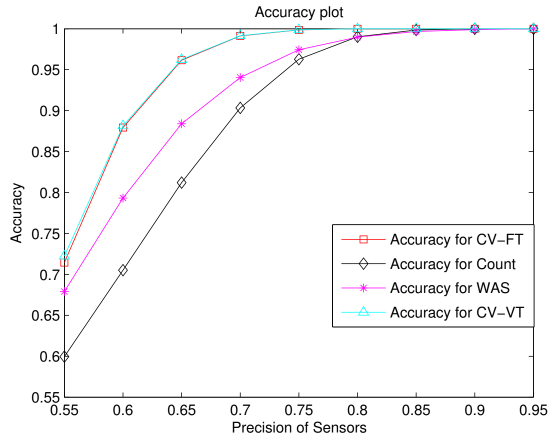

Fig. 3. Numerical plots for Accuracy vs. precision of sensors p for various aggregation schemes for tree topology

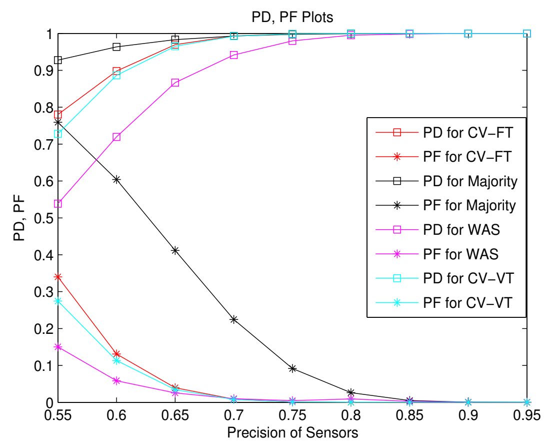

Fig. 4. Numerical plots for PD and PF vs. precision of sensors p for various aggregation schemes for tree topology

than other one bit fusion rules. And since $CV - VT$ is difficult to implement as compared to $CV - FT$, we recommend the use of $CV - FT$ given that both performs equally well. The CV rule thus outscored all the previous proposed aggregation scheme for tree topology.

References

[1] Z. Chair and P. K. Varshney, “Optimum data fusion in multiple sensor detection systems,” *IEEE Trans. Aerospace Electro. Sys.*, vol. 27, pp. 98–101, Jan 1986.

[2] L. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” *IEEE Communication Magazine*, August 2002.

[3] B. G. Jagyasi, B. K. Dey, S. N. Merchant, and U. B. Desai, “An mmse based weighted aggregation scheme for event detection using wireless sensor network,” *14th European Signal Processing Conference, EUSIPCO 2006*, Sep 2006.

[4] ——, “Weighted aggregation scheme with lifetime-accuracy tradeoff in wireless sensor network,” *in Proc. 4th International Conference on Intelligent Sensing and Information Processing 2006, ICISIP 2006*, Dec 2006.

[5] G. Werner-Allen, J. Johnson, M. Ruiz, J. Lees, and M. Welsh, “Monitoring volcanic eruptions with a wireless sensor network,” *in Proc. 2nd European Workshop on Wireless Sensor Networks (EWSN 05)*, January-Febuary 2005.

[6] G. Werner-Allen, K. Lorincz, M. Welsh, O. Marcillo, J. Johnson, M. Ruiz, and J. Lees, “Deploying a wireless sensor network on an active volcano,” *IEEE Internet Computing Online*, vol. 10, March-April 2006.

[7] C. Hartung, R. Han, C. Seielstad, and S. Holbrook, “Firewxnet: A multitiered portable wireless system for monitoring weather conditions in wildland fire environments,” *in Proc. 4th international conference on Mobile systems, applications and services MobiSys 2006*, vol. 10, June 2006.

[8] A. Sheth, K. Tejaswi, P. Mehta, C. Parekh, R. Bansal, S. N. Merchant, U. B. Desai, C. Thekkhath, K. Toyama, and T. Singh, “Poster abstract-senslide: A sensor network based landslide prediction system,” *in Proc. 3rd ACM Conference on Embedded Networked Sensor Systems, SenSys*, Nov 2005.

[9] R. Viswanathan and P. K. Varshney., “Distributed detection with multiple sensors: Part i - fundamentals,” *In Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, January 1997.

[10] R. Blum, S. Kassam, and H. V. Poor, “Distributed detection with multiple sensors: Part ii - advanced topics,” *In Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, January 1997.
[11] W. Heinzelmal, A. Chandrakasam, and H. Balakrishnan, “An application specific protocol architecture for wireless microsensor networks,” *IEEE transaction on wireless communications*, vol. 1, no. 4, pp. 660–670, October 2002.
[12] Z. Chair and P. K. Varshney, “Distributed bayesian hypothesis testing with distributed data fusion,” *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 18, no. 5, pp. 695–699, Oct 1988.
[13] B. Chen, R. Jiang, T. Kasetkasam, and P. K. Varshney, “Channel aware decision fusion in wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 52, no. 12, Dec 2004.
[14] R. Niu and P. K. Varshney, “Distributed detection and fusion in a large wireless sensor network of random size,” *EURASIP Journal on Wireless Communication and Networking*, vol. 4, pp. 462–472, 2005.
[15] R. Niu, P. K. Varshney, M. H. Moore, and D. Klamer, “Decision fusion in a wireless sensor network with a large number of sensors,” *in Proc. 7th IEEE International Conference on Information Fusion (ICIF’04), Stockholm, Sweden*, June-July 2004.
[16] S. Borbash, “Design consideration in wireless sensor networks,” Ph.D. dissertation, Institute for system research (ISR), University of Maryland, 2004.