Power-Efficient Direct-Voting Assurance for Data Fusion in Wireless Sensor Networks

Hung-Ta Pai, *Member*, *IEEE*, and Yunghsiang S. Han, *Member*, *IEEE*

Abstract—Wireless sensor networks place sensors into an area to collect data and send them back to a base station. Data fusion, in which collected data are fused before they are sent to the base station, is usually implemented over the network. Since a sensor is typically placed in locations that are accessible to malicious attackers, information assurance of the data fusion process is very important. A witness-based approach [9] has been proposed to verify the fusion data. In this approach, the base station receives the fusion data and "votes" on the data from a randomly chosen sensor node. The vote comes from other sensor nodes, called "witnesses," to confirm the correctness of the fusion data. Since the base station receives the vote through the chosen node, this node could forge the vote if it is compromised. Accordingly, the witness node must apply cryptographic operations to the vote to prevent this forgery. The cryptographic operation requires more bits than the vote, increasing the transmission burden from the chosen node to the base station. The chosen node consumes too much power. This work improves the witness-based approach using a direct voting mechanism such that the proposed scheme performs better in terms of assurance, overhead, and delay. The witness node transmits the vote directly to the base station. Forgery does not pose a problem in this scheme. Moreover, fewer bits are necessary to represent the vote, significantly reducing the power consumption. Performance analysis and simulation results indicate that the proposed approach has a 40-times lower overhead than the witness-based approach.

Index Terms—Wireless sensor networks, data fusion assurance, power-efficient, voting mechanism, witness.

1 Introduction

Wireless sensor networks (WSNs) are comprised of many tiny low-cost battery-powered sensors in a small area [1], [5], [11], [13], [14], [24], [25]. The sensors detect environmental variations and then transmit the detection results to other sensors or a base station [2], [4], [6], [7], [27]. One or several sensors then collect the detection results from other sensors. The collected data must be processed by the sensor to reduce the transmission burden before they are transmitted to the base station. This process is called *data fusion* and the sensor performing data fusion is the *fusion node.* The fusion data may be sent from the fusion node to the base station through multiple hops [10] or through a direct link [21].

Although fusion markedly lowers the traffic between the fusion node and the base station, the fusion node is more critical and vulnerable to malicious attacks than the nonfusion sensors [16], [18], [23]. If a fusion node is compromised, then the base station cannot ensure the correctness of the fusion data that have been sent to it. This problem of fusion data assurance arises because the detection results are not sent directly to the base station, explaining why the fusion result usually cannot be verified.

This problem can be solved in two ways: One is hardware based [3], [15] and the other is software based [8], [9], [19], [28]. Since the hardware-based approach requires extra circuits to detect or frustrate the compromised node, the cost and continual power consumption of sensors are increased but protection against all attacks cannot be guaranteed. Conversely, most software-based methods require no or little extra hardware for data assurance. However, as has been mentioned elsewhere [8], [9], several copies of the fusion data or multiple detection results must be sent to the base station, so the power consumption for the data transmission is very high. Since power is very valuable in the WSN [1], [4], a power-efficient data assurance algorithm must be developed such that each sensor has a longer lifetime to perform its tasks in the WSN. That is, the power consumption overhead for performing data fusion assurance should be maintained as small as possible.

The witness-based approach that was presented by Du et al. [9] does not have this difficulty. Several fusion nodes are used to fuse the collected data and they can communicate with the base station. Only one node is chosen to transmit the fusion result to the base station. The other fusion nodes, serving as witnesses, hash the fusion results to message authentication codes (MACs). The MACs are then sent to the base station through the chosen fusion node. Finally, the base station utilizes the received MACs to verify the received fusion data. The verification may be wrong since the chosen node may be compromised and forge MACs. The correctness of the verification depends not only on the number of malicious fusion nodes but also on the length of the MAC. A long MAC increases the reliability of the verification. However, the transmission of a long MAC imposes a large communication burden. If the received fusion result at the base station cannot pass the verification, then a polling scheme is started to determine whether any valid fusion result is available at the other

• *The authors are with the Graduate Institute of Communication Engineering, National Taipei University, No. 151, University Rd., Sanhsia, Taipei County, 237 Taiwan. E-mail: {htpai, yshan}@mail.ntpu.edu.tw.*

Manuscript received 5 Apr. 2006; revised 28 May 2007; accepted 17 July 2007; published online 23 Aug. 2007.
Recommended for acceptance by S. Iyengar.
For information on obtaining reprints of this article, please send e-mail to: tc@computer.org, and reference IEEECS Log Number TC-0135-0406.
Digital Object Identifier no. 10.1109/TC.2007.70805.

fusion nodes. In addition to the fusion result that had been sent by the malicious fusion node, several copies of the correct fusion result may also have to be transmitted to the base station. The transmission of the correct fusion result consumes the power of the uncompromised fusion node.

Even though the witness-based approach developed in [9] is more attractive than previous approaches, it suffers from several drawbacks. First, several copies of the fusion result may be sent to the base station by uncompromised nodes, increasing the power consumed at these nodes. Second, a MAC mechanism must be implemented in each sensor node that occupies limited memory resources at each sensor. The MAC mechanism is designed solely for fusion data assurance; cryptographic operations are not otherwise needed for applications in which the fusion result need not be kept secret. Third, the voting information in the current polling round is not used in the next polling round if the verification has not been passed in the current polling round. All votes are collected in each polling round. If the voting can be used in any way, then the polling process should be shortened to save power and reduce the time delay. Finally, since all votes are collected by one node and sent to the base station, this node can forge the fusion result and the votes.[1] Such forgery must be prevented to increase security in the data fusion system.

This work develops a novel data fusion assurance mechanism to eliminate all of the aforementioned shortcomings in the witness-based method by Du et al. [9]. The correctness of the verification in the proposed scheme depends only on the number of compromised fusion nodes. As in the witness-based approach, a fusion node is selected to transmit the fusion result, while other fusion nodes serve as witnesses. Nevertheless, the base station obtains votes that contribute to the transmitted fusion result directly from the witness nodes. No valid fusion data are available if the transmitted fusion data are not approved by a preset number of witness nodes. Based on this voting mechanism, two schemes are described: One needs variant rounds of voting and the other requires only one round of voting. The key advantages of the variant-round (VR) scheme over that presented in [9] are summarized as follows:

- Only one copy of the correct (valid) fusion result, provided by one of uncompromised fusion nodes, is transmitted to the base station, regardless of whether the system is comprised of sufficient uncompromised nodes to support the fusion result. This single transmission saves the power of the uncompromised node. However, in the scheme in [9], when too few witness nodes are available to verify the correct fusion result, the polling continues until not enough votes to pass the verification can be collected to verify the fusion result. During the polling process, more than one uncompromised node may send the correct fusion result to the base station.
- The direct voting scheme is adopted and no MAC mechanism needs to be implemented at each node; therefore, no extra memory is needed to implement such a mechanism. Moreover, no communication is necessary between the sensors in this voting scheme. In contrast, the MAC message of each witness node must be collected at the fusion node in the scheme that is presented in [9].
- Early termination is achievable when the base station receives enough "agree" or "disagree" votes. In contrast, the scheme in [9] always collects all votes.
- A witness node may remain silent (without transmission) when it agrees with the transmitted fusion result. Only "disagree" votes need to be sent. This "silent assent" feature drastically reduces the transmission power consumption in the system. However, in [9], MACs are always sent and they cannot be too short to jeopardize verification of the fusion result.
- A compromised fusion node can be identified if it has been excluded by the base station during the polling process.[2] This "traitor exclusion" is useful for further verification of the fusion result. Even though the scheme in [9] also offers this "traitor exclusion" feature, it fails to exploit it when the fusion node can successfully forge the fusion result.
- No forged result can be accepted by the base station unless the number of compromised nodes reaches the number of support votes that is required to verify the fusion result and these nodes collude to forge the fusion result. In contrast, for the scheme in [9], the node that sends the fusion result and all votes may still successfully forge the fusion result, even when it is the only node to be compromised.
- Analytical and simulation results reveal that the proposed scheme has an up-to-40 times lower overhead than the scheme by Du et al. [9].

In the one-round OR scheme, the base station polls each sensor at most once. The maximum delay of the OR scheme is substantially less than that of the VR scheme.

2 Data Fusion Assurance Problem and Previous Work

Fig. 1 depicts a WSN for distributed detection with N sensors for collecting environment variation data and a fusion center to make a final decision concerning detections. This network architecture is similar to the architecture of the so-called SENsor with Mobile Access (SENMA) [26], [30], Message Ferry [31], and Data Mule [20]. Since the distance between the fusion node and the base station is usually long, the power consumed by the fusion node upon receiving data is much lower than the power associated with transmission [12], [22]. For example, the parameters given in [12], [22], where the energy consumed by the transmitter circuitry is 81 nJ/bit, and the antenna output energy to reach the destination at unit distance away is $0.1\ \mathrm{nJ/m^3}$, yield an energy consumption for data transmission of $81 + 0.1 \times 200^3 = 800,081\ \mathrm{nJ/bit}$ if the distance between the fusion node and the base station is 200 m and the path attenuation factor is 3. However, the energy

1. A malicious node can attempt forgery by first generating the forged fusion result and then randomly guessing the MAC of each witness. To reduce the probability of a successful forgery, each witness may use a different key to generate its MAC.

2. The correct fusion result is assumed to be obtained after the whole polling process has been completed.

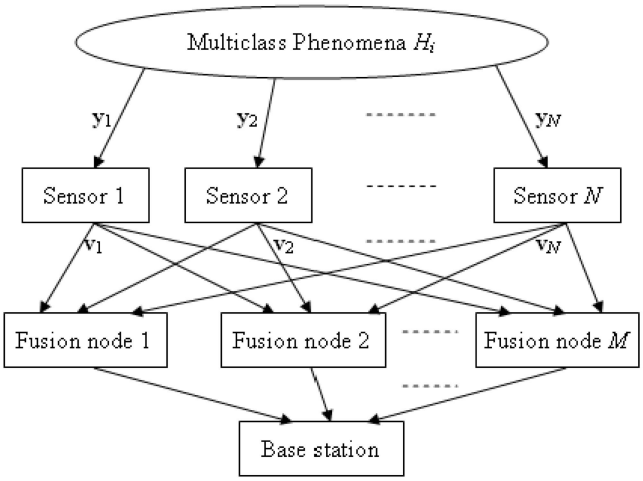

Fig. 1. Structure of a wireless sensor network for distributed detection using N sensors and M fusion nodes.

consumed in receiving data is 180 nJ/bit. Since the energy consumption for reception at a fusion node is 4,000 times less than that for transmission, only transmission energy is considered in the subsequent analysis and comparison.

At the jth sensor, one observation y_j is made for one of the phenomena H_i, where $i = 1, 2, \ldots, L$. If the detection (raw) data are transmitted to the fusion nodes without any processing, then the transmission imposes a very high communication burden. Hence, each sensor must make a local decision based on the raw data before transmission. The decisions,[3] v_j, $j = 1, 2, \ldots, N$, can be represented with fewer symbols than the raw data. The sensor then transmits the local decision to M fusion nodes by broadcasting. The fusion node combines all of the local decisions to yield a final result and it communicates directly with the base station. Finally, one of the fusion nodes is specified to send the final result to the base station. Unless all of the fusion nodes or all of the sensors fail, this detection and fusion scheme guarantees that the base station will receive the detection result. However, the accuracy of the result is uncertain.

Two problems must be solved to ensure that the base station obtains the correct result. First, every fusion node must correctly fuse all of the local decisions such that all of the fusion results must be identical. Several algorithms have been proposed to address this issue [7], [28], [29]. Among them, estimation theory was first employed to evaluate the security of fusion operations at a fusion node in [28]. Some of the fusion operations, such as "average" and "sum," are vulnerable according to this evaluation. Several methods, such as "truncation" and "trimming," were then proposed to increase the security level of the fusion operations. These methods focus on obtaining a correct fusion result. When they are applied to the assurance problem, all fusion results must be sent to the base station and too much power is consumed. We assume herein that this problem has been solved.

The second problem concerns the assurance of the fusion result. Transmission between the fusion node and the base station is assumed herein to be error-free. Since some fusion nodes may be compromised, the fusion node that is chosen by the base station to transmit the fusion result may be one of the compromised nodes. Malicious data may be sent by the compromised node and the base station will not be able to distinguish the compromised nodes from the normal fusion nodes since the data detected by the sensor are not sent directly to the base station. Consequently, the result obtained at the base station may be incorrect. That is, the base station can suffer from *stealthy attacks*, where an attacker tries to make the base station accept a forged result [19].

In an earlier study, a fusion node established a Merkle hash tree using collected detection results as leaves [19]. The base station requests one of the results and checks if it is consistent with the tree during the assurance process. The probability of detecting a cheating fusion node can be increased by transmitting fewer detection results to the base station. However, different assurance algorithms must be developed for various fusion operations. No general assurance approach is provided. Additionally, only one fusion node is assumed. When it is compromised, the base station can no longer receive correct fusion data.

Du et al. [9] presented a witness-based approach to ensure the correctness of the fusion result. All fusion nodes, other than the chosen node, act as witnesses to the transmitted fusion result. The witness nodes compute MACs on the fusion results with private keys that are shared with the base station and then send the MACs, as "votes," to the chosen node. The chosen node collects all of the MACs from the witness nodes and transmits them with its own fusion result to the base station. The base station determines from the received data whether the fusion result from the chosen node is accurate. In the $T+1$ out of M voting scheme, the fusion result of the chosen node needs support from at least T witness nodes, where M is the number of fusion nodes and T is a threshold. That is, the base station accepts the fusion result if the fusion result is supported by at least T MACs. Normally, $T > \lfloor M/2 \rfloor$. However, even when the number of compromised nodes C is less than T, the fusion result accepted by the base station is not always correct. If the chosen node is compromised, then it may forge the fusion result and the MACs. The probability that the base station accepts the forged fusion result is given by

$$P_e = \sum_{i=T}^{M-1} \binom{M-1}{i} \left(\frac{1}{2^{k_w}}\right)^i \left(1 - \frac{1}{2^{k_w}}\right)^{M-i-1},$$

where k_w is the size of each MAC. Since the number of the transmitted MACs is $M-1$,[4] the number of the transmitted bits, excluding the fusion result, is $(M-1)k_w$. For instance, consider the majority voting rule in which $T+1 \geq \lceil M/2 \rceil$. To ensure that $P_e \leq 2^{-10}$, set $k_w = \lceil 2(10/(M-1)+1) \rceil$. Although only one copy of the fusion result is sent to the base station by each chosen node in this witness-based approach, the witness nodes still require significant communication bandwidth because the MACs of the fusion results are transmitted. If the received fusion result at the

3. These decisions may be compressed data whose sizes depend on the application of the WSN.

4. Since the chosen node does not need to endorse the fusion result that it sends to the base station, only $M-1$ MACs are transmitted.

base station cannot pass the verification, then a polling scheme is started to determine whether any valid fusion result is available at the other fusion nodes. In addition to the fusion result sent by the malicious fusion node, several copies of the correct fusion result may also have to be transmitted to the base station.

In a fair comparison between the proposed scheme with the witness-based approach, the *overhead* is defined as the total number of bits, excluding the bits associated with *one* copy of the correct fusion result, that are transmitted to the base station by *uncompromised* nodes during the data assurance process. The power consumed at all compromised nodes is not considered since they are not useful to the WSN. Therefore, the overhead can be regarded as the *useful* power that is consumed for the data assurance by the sensor. Since the base station is generally powerful and not battery powered, its power consumption is not critical in a WSN. The *round delay* is defined as the number of rounds[5] that are required to collect all MACs (votes) from the witness nodes; the *polling delay* is defined as the number of votes (including all "agree" and "disagree" voting).[6] The overhead of the witness-based approach [9] is then derived in Appendix A and is summarized in (1), (2), and (3).

Notably, the maximum round delay is $C+1$, where C is the number of compromised nodes. When the fusion result is valid, including the case of no compromised node ($C=0$), 40 and 60 bits must be transmitted to the base station when $M=11$ ($k_w=4$, that is, $P_e \leq 2^{-10}$) and $M=21$ ($k_w=3$, that is, $P_e \leq 2^{-10}$), respectively, setting $K=0$, where K is the number of bits that represent the fusion result. (In practice, $K>0$.) A large amount of power must be consumed for this transmission, substantially reducing the lifetime of the fusion node. The problem of power consumption is even worse when the fusion result is invalid. For example, the maximum average overhead is about 109 and 314 bits for $M=11$ ($C=5$ and $T=6$) and $M=21$ ($C=10$ and $T=11$), respectively. Therefore, the witness-based approach must be enhanced.

3 Improved Voting Mechanism

The voting mechanism in the witness-based approach is designed according to the MAC of the fusion result at each witness node. This design is reasonable when the witness node does not know the fusion result at the chosen node. However, in practice, the base station can transmit the fusion result of the chosen node to the witness node. Therefore, the witness node can obtain the transmitted fusion result from the chosen node through the base station. The witness node can then compare the transmitted fusion result with its own fusion result. Finally, the witness node can send its vote (agreement or disagreement) on the transmitted result directly to the base station, rather than through the chosen node.

When a fusion node sends its fusion result to the base station, other fusion nodes serve as witness nodes. The witness node then starts to vote on the transmitted result. Two data fusion assurance schemes are proposed.

5. Or the number of fusion results sent to the base station.
6. The overall time delay can then be derived from these two delays.

3.1 Variant-Round Scheme

In this scheme, the base station must ask the witness node whether it agrees or disagrees with the transmitted fusion result. The witness node then sends its vote to the base station. No denial-of-service attack is assumed and the vote can be clearly identified at the base station [19], [28]. If the transmitted fusion result is not supported by at least T witness nodes, then the base station may have to select a witness node that does not agree with the transmitted result as the next chosen node. The steps of the scheme are given as follows:

Step 1. The base station chooses a fusion node. Other fusion nodes serve as witness nodes. Define a set of witness nodes that includes all witness nodes and let the nodes in the set be randomly ordered. Denote $M'=M-1$ as the size of the witness set in the current round.

Step 2. The chosen node transmits its fusion result to the base station.

Step 3. The base station polls and sends the above fusion result to the node in the witness set by following the order of the witness nodes. The polling process does not stop until

- T witness nodes *agree* with the transmitted fusion result (agreeing nodes), where $1 \leq T \leq M'$,
- $M'-T+1$ witness nodes *disagree* with the transmitted fusion result (disagreeing nodes), or
- all witness nodes have been polled.

Step 4. A represents the number of polled witness nodes that agree with the transmitted fusion result. D denotes the total number of polled witness nodes that disagree with the transmitted fusion results **plus the number of unpolled witness nodes**. Notably, $A+D=M'$.

- If $A=T$, then the transmitted fusion result is verified. Stop the polling.
- If $A<T$ and $D<T+1$, then no reliable fusion result is valid. Stop the polling.
- If $D \geq T+1$, then exclude the A agreeing witness nodes from the witness set. Let the first node that disagrees with the transmitted fusion result be the chosen node to transmit its fusion result. Thus, the updated size of the witness set, M', is $D-1$.[7] Go to Step 2 for the next round of the polling.

In Step 1, the randomly generated order of witness nodes determines both the polling order and the order in which replacement fusion nodes are chosen. The random ordering facilitates the analysis of the scheme that is given in Section 3.2.

Step 3 lists three termination conditions for a round of polling. The first condition is that enough "agree" votes are obtained at the base station and the fusion result of the chosen node in Step 2 is accepted. (This condition should be compared with the first condition of Step 4.) The second condition is that the number of "disagree" votes is too large such that the transmitted fusion result cannot receive enough "agree" votes. Notably, the witness set in the

7. The number of nodes performing the polling in the next round becomes D.

current round has M' nodes and at least T "agree" votes are required such that the number of "disagree" votes cannot exceed $M' - T$. In this case, no further polling is necessary in the current round. Since early termination may occur under these two conditions, not all witness nodes have to transmit their votes; thus, the overhead is reduced.

Moreover, since a witness node can be silent when it agrees with the transmitted fusion result, the overhead may be greatly reduced, as discussed in Section 3.2. The "silent assent" mechanism can be implemented as follows: A time division multiple access (TDMA) protocol is assumed to be implemented between the base station and sensors to avoid transmission collision such that each sensor can identify the start and the duration of each time slot. At each polling slot, all sensors listen to the channel and wait for the polling signal. Since the base station has more power than the sensors, the polling signal is assumed not to be lost. In each round, the base station starts by polling the chosen node and requesting the fusion result. The polled sensor sends the result in the next time slot. Then, the base station polls the witness nodes that may not include excluded nodes in the next time slot according to the random order determined in Step 1. The polled node sends the disagree response to the base station or remains silent during the next time slot. If the base station cannot identify the response from a sensor such that the sensor is not silent but also does not send a clear signal of disagreement, then the base station polls the same sensor again. The maximum number of multiple pollings of one sensor can be limited to a predefined number to prevent stacking at one node forever when a sensor cannot clearly reply. After this number has been reached, the base station excludes this node from the polling process forever and moves on to the next sensor. In this case, the size of the witness set is immediately reduced by one. This dynamic time slot assignment reduces the time delay from that of the static time slot assignment in which each sensor is assigned a fixed time slot to transmit data or remain silent.

In Step 4, the scheme decides whether the whole polling process has to be terminated (under the first and second conditions) or an additional round of polling must be conducted (under the third condition). The first condition of Step 4 is the consequence of the first termination condition of Step 3 and, thus, the transmitted fusion result is accepted. The second condition of Step 4 refers to a situation in which enough "agree" votes cannot be obtained, even though the base station undertakes additional rounds of polling. The reason is that the number of witness nodes in the next round, $D - 1$, is less than T. Since the transmitted fusion results in the current round of polling differ from those in the next round, the agreeing nodes in the current round will not agree with the transmitted fusion result in the next round. Hence, the agreeing nodes in the current round are excluded under the third condition such that the base station polls fewer witness nodes in the next round.

The majority voting rule is assumed to be adopted in the VR scheme, where $T + 1 \geq \lceil M/2 \rceil$ and M is odd. The security strength of the VR scheme can be analyzed as follows: In this scheme, the base station accepts a forged fusion result only if the number of compromised fusion nodes, C, exceeds T and the compromised fusion nodes cooperate with each other. Based on the assumption that a fusion node may be compromised with an identically and independently distributed probability p_c, the probability that the base station accepts a forged fusion result is given by $\sum_{i=T+1}^{M} \binom{M}{i} p_c^i (1 - p_c)^{M-i}$. When more than T nodes are uncompromised, the base station can obtain the correct fusion result. Accordingly, the probability that the base station accepts a correct fusion result is $\sum_{i=T+1}^{M} \binom{M}{i} (1 - p_c)^i p_c^{M-i}$. Fig. 2 compares the probabilities that the base station accepts a forged fusion result and a correct result when $M = 11$ and $T = 5$. The probability that the base station accepts a correct result is much larger than the probability that the base station accepts a forged result when $p_c < 0.1$. In contrast, the probability that the base station accepts a forged result is larger than the probability that the base station accepts a correct result when $p_c > 0.5$.

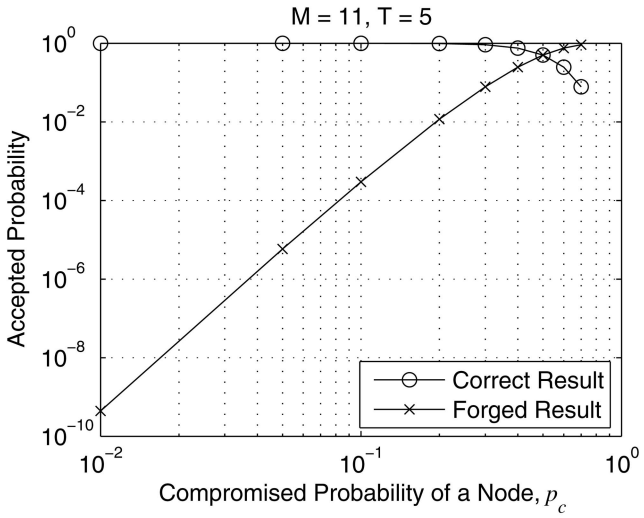

Fig. 2. Comparison of probabilities that the base station accepts a forged fusion result and a correct result when $M = 11$ and $T = 5$.

When the base station accepts a fusion result in the ith round, all chosen nodes and all of the nodes that were excluded before the ith round can be excluded from the next assurance process. When the accepted result is indeed the correct one, this "traitor exclusion" property can help the system to identify unreliable nodes in the network.

3.2 Analysis of Variant-Round Scheme

This analysis of the VR scheme assumes that the compromised node always transmits the forged fusion result when the compromised node is chosen to send its fusion result; therefore, the compromised node tries to make the base station accept a forged fusion result when it is chosen. When a compromised node serves as a witness node, it always disagrees with the correct fusion result and agrees with the forged fusion result with a probability P_f. If the compromised node attempts to make the base station accept the forged fusion result, then it always agrees with the fusion result that is transmitted by other compromised nodes, that is, $P_f = 1$, and at most two rounds of polling have to be run. Conversely, if the compromised node wants to make the polling process run for as long as possible, then

it always disagrees with the transmitted fusion result, that is, $P_f = 0$.

The performance of the VR scheme when $P_f = 0$ is analyzed in Appendix B and is summarized in (10) and (11). The performance of the VR scheme when $P_f = 1$ is given in [17]. The next section presents computer-simulated overheads when $P_f \neq 0, 1$.

An interesting property of the VR scheme is that, throughout the polling process, at most one fusion result is transmitted from all of the uncompromised nodes to the base station. Hence, the overhead of the scheme is independent of the size of the fusion result.[8] This claim is demonstrated by the following argument: In the case of a valid fusion result, when the fusion result from an uncompromised node is sent to the base station in a round of polling, the polling process stops in this round and the valid fusion result is obtained by the base station. Accordingly, only one valid fusion result is sent to the base station by all uncompromised nodes. In the case of no valid fusion result, fewer than T uncompromised nodes are witnesses. If a round of the polling process is the first in which the chosen node is uncompromised, then this round terminates when either all witness nodes have been polled or $M' - T + 1$ witness nodes disagree with the transmitted fusion result. In the former case, the polling process terminates. In the latter case, another polling round is required. Importantly, in the following round, all of the uncompromised nodes will be the last $T - 1$ nodes in the witness set and will not then be chosen to send any fusion result before the polling process is completed.[9] Therefore, only one fusion result is sent by all uncompromised nodes when no valid result can be obtained by the base station.

3.3 One-Round Scheme

The number of rounds in the above scheme is not fixed. Hence, the delay varies. A variable delay is undesirable in some applications, such as real-time systems. This work proposes another scheme that is based on the improved voting mechanism. In this scheme, the base station may receive different fusion results from the witness nodes. It requires that all received fusion results be stored. This scheme has a fixed delay and is summarized as follows:

Step 1. The base station randomly chooses a fusion node. Other fusion nodes serve as witness nodes. A set of witness nodes that includes all of the witness nodes is defined and the nodes in the set are randomly ordered.

Step 2. The chosen node transmits its fusion result to the base station. The base station sets the fusion result as the best temporary **voting** result and the number of votes for agreement with the fusion result is set to zero.

Step 3. The base station polls the nodes with the best temporary voting result, which currently has the maximum number of votes, following the order of the witness nodes. The witness node compares its fusion result with the best temporary voting result.

- If the witness node agrees with the best temporary voting result, it sends an agreeing vote to the base station. The base station increases the number of agreeing votes for the best temporary voting result by one.
- If the witness node does not agree with the best temporary voting result, it transmits its fusion result to the base station.
  - If the fusion result has been stored in the base station, then the base station increases the number of agreeing votes for the fusion result by one.
  - If the fusion result has not been stored in the base station, then the base station stores the fusion result and the number of agreeing votes for the fusion result is set to zero.

The base station sets the best temporary voting result to the received fusion result that had received the maximum number of agreeing votes to poll the next witness node. If two or more fusion results receive the maximum umber of votes, then the temporarily best voting result is set to the result that had most recently been voted for. The polling stops when any received fusion result receives T votes or when the number of unpolled nodes plus the maximum number of votes for the results recorded at the base station is less than T.

From Step 3, we know that the base station keeps only one best temporary voting result when it is polling a witness node. Therefore, the witness node may be silent when it agrees with the best temporary voting result. The same "silent assent" mechanism given in Section 3.1 can be applied to the OR scheme. The analysis of the OR scheme will be left to future work.

4 Performance Evaluation

In this section, numerical and computer simulations are conducted to evaluate the performance of the proposed schemes. The performance of the proposed VR scheme is numerically calculated by the results given in Section 3 and the Appendix when $P_f = 0$ and 1. The performances of the VR scheme when $P_f \neq 0$ or 1 and that of the OR scheme are evaluated using Monte Carlo computer simulations. The proposed schemes are compared using the witness-based approach in terms of overhead, average round delay, and average polling delay. In the witness-based approach [9], the size of each MAC, k_w, is assumed to be four bits. Notably, this approach still suffers the risk that the chosen node may create a forged fusion result. In the evaluation of the overhead of the VR scheme, the size of the fusion result is zero (that is, $K = 0$) such that the witness-based approach in [9] performs best. As stated at the end of Section 3.2, the overhead of the VR scheme is independent of the size of the fusion result. Under this assumption, the VR scheme is inferred to have a better overhead performance than the witness-based approach for all sizes of fusion result whenever it outperforms with a zero-sized fusion result. All results are presented for the number of nodes $M = 11$.[10]

8. Note that the overhead is defined as the total number of bits that are transmitted minus the number of bits associated with the correct fusion result.

9. In the next round, all uncompromised polled nodes are deleted from the witness set according to the scheme.

10. Similar results can be obtained for $M = 21$ but omitted to save space.

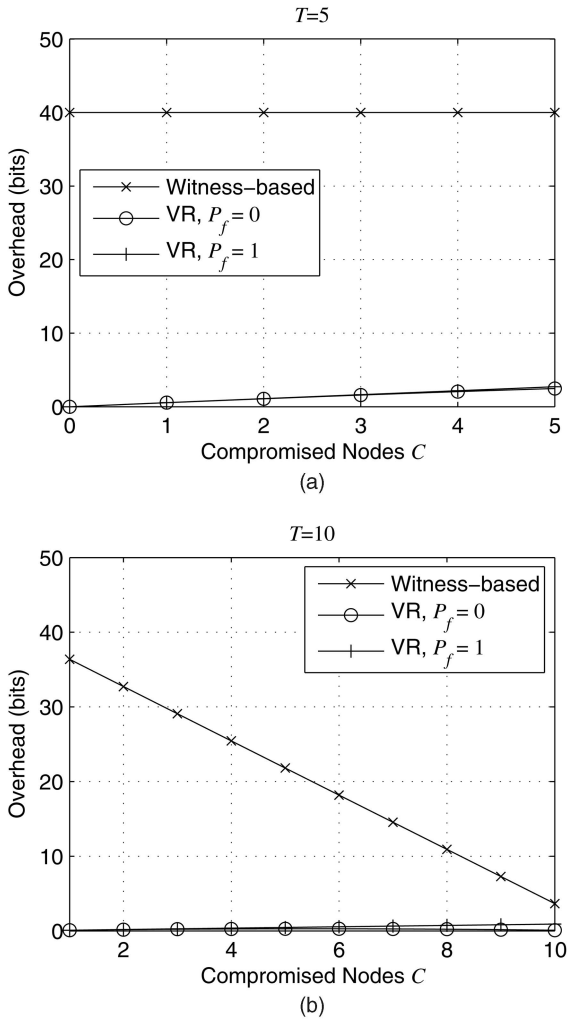

Fig. 3. Overhead comparison between the VR scheme and the witness-based approach [9] for $M = 11$ and $P_f = 1, 0$ (a) when $T = 5$ (valid fusion result) and (b) when $T = 10$ (invalid fusion result).

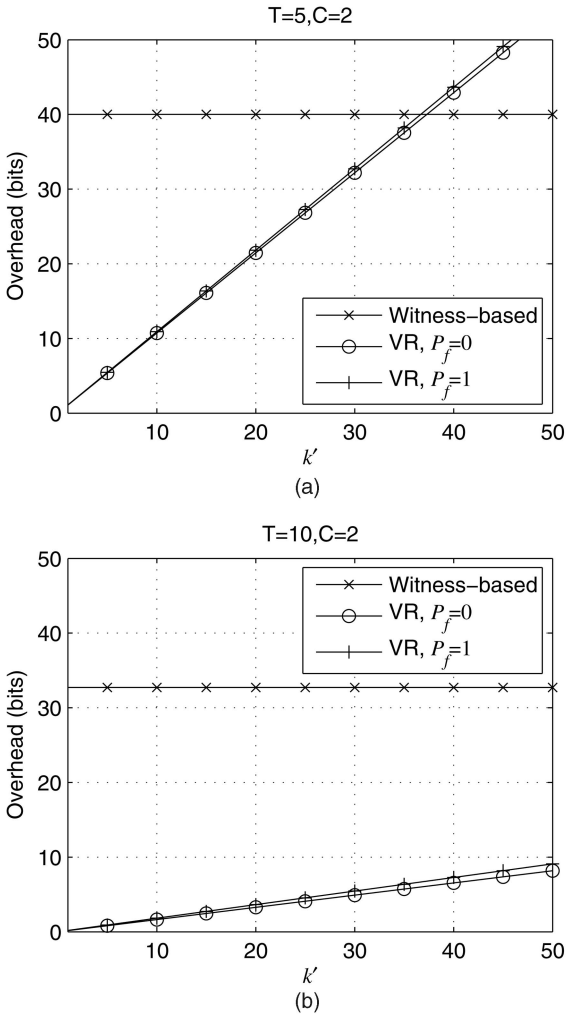

Fig. 4. The overheads of the VR scheme and the witness-based approach [9] for $M = 11$, $C = 2$, and $P_f = 1, 0$ (a) when $T = 5$ (valid fusion result) and (b) when $T = 10$ (invalid fusion result) are compared. The packet headers transmitted by the disagreeing nodes are considered.

Furthermore, the overhead for collecting votes from witness nodes at the chosen node is also not counted in the witness-based approach [9].

A certain fixed packet header may be needed when the fusion node transmits a nonzero number of bits to vote. That is, when k or k', where k (k') is the number of bits that must be sent by a witness to the base station when it agrees (disagrees) with the transmitted fusion result, does not equal to zero, the packet header may be necessary. Here, the packet header is ignored. The effect of the packet header will be discussed later. The VR scheme is compared with the witness-based approach. In the VR scheme, $k = 0$ and $k' = 1$ are set. Fig. 3 compares overheads. The VR scheme substantially outperforms the witness-based approach, regardless of the fusion result at the base station. For example, according to Fig. 3, when $T = 5$ and $M = 11$, the VR scheme is almost 40 times better than the witness-based approach given in [9] in terms of overhead for $C = 1$ and 2.

If a packet header is required to transmit a "disagree" vote when $k = 0$, k' may exceed 1. Fig. 4 depicts the overhead of the VR scheme for various k' when $C = 2$ and $T = 5, 10$. The VR scheme outperforms the witness-based approach until $k' = 36$ when $T = 5$ (valid fusion result). The VR scheme outperforms the witness-based approach even for the maximum k' simulated when $T = 10$ (invalid fusion result). Recall that the overhead for collecting votes from witness nodes at the chosen node, potentially increasing the overhead of the witness-based approach, is again not counted in the witness-based approach.

Fig. 5 compares the average round delays of the proposed VR scheme and the witness-based approach. This figure demonstrates that the average round delays of the proposed scheme are smaller than those of the witness-based approach when the base station can obtain the valid fusion results; however, they perform equally when the base station obtains invalid fusion results. Fig. 6 compares the average polling delays of the proposed VR scheme and the witness-based approach. In Fig. 6, the proposed scheme has much smaller average polling delays than the witness-based approach. This difference is not evident in the average round delay performance. For example, when $T = 10$ and $M = 11$, the proposed VR scheme is almost five times better than the witness-based approach in terms of

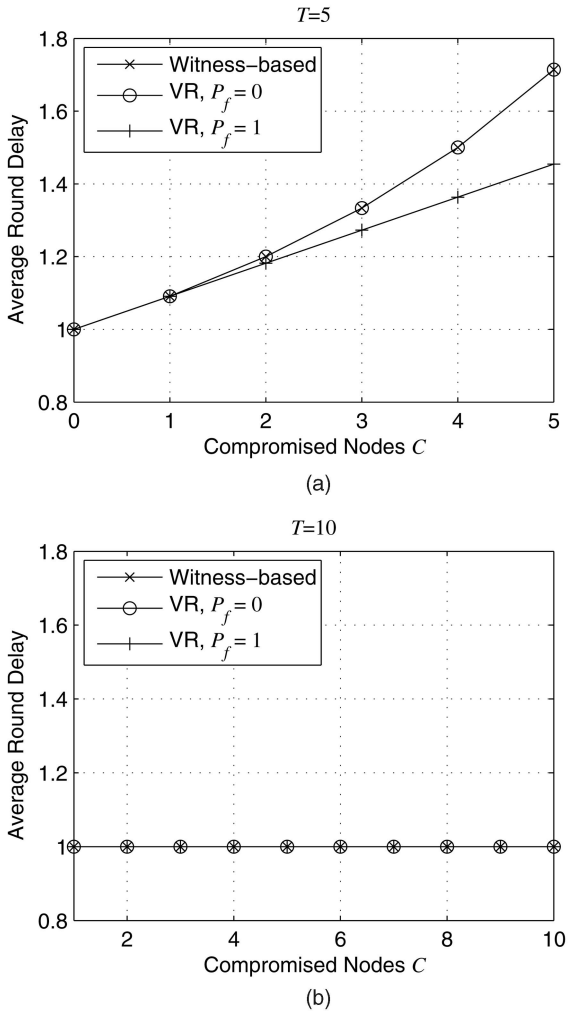

Fig. 5. Average round delay comparison between the VR scheme and the witness-based approach [9] for $M = 11$ and $P_f = 1, 0$ (a) when $T = 5$ (valid fusion result) and (b) when $T = 10$ (invalid fusion result).

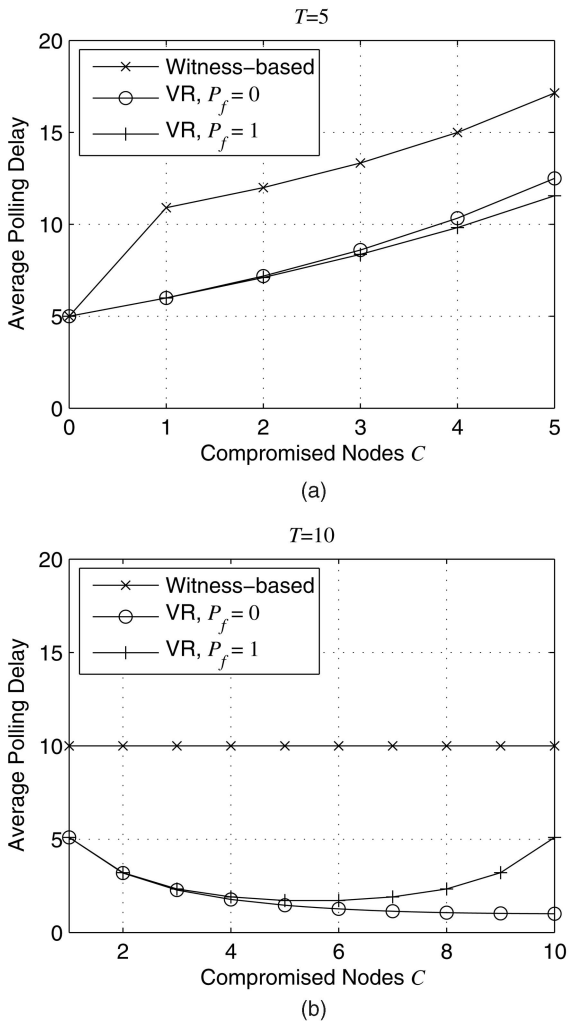

Fig. 6. Average polling delay comparison between the VR scheme and the witness-based approach [9] for $M = 11$ and $P_f = 1, 0$ (a) when $T = 5$ (valid fusion result) and (b) when $T = 10$ (invalid fusion result).

average polling delay for $C = 4$ and 5. Accordingly, the proposed scheme outperforms the witness-based approach [9] in terms of overhead and delay.

The following computer simulations evaluate the VR scheme when $P_f = 0.25, 0.5$, and 0.75 by performing 10,000 Monte Carlo tests for each simulation. In the first set of simulations, the witness-based approach [9] is compared with the proposed VR scheme when $P_f = 0.25, 0.5, 0.75$. Fig. 7 presents the results for $M = 11$, $k = 0$, and $k' = 1$. The VR scheme outperforms the witness-based approach in every P_f simulated. For example, in Fig. 7, when $T = 5$, the proposed VR scheme is almost 40 times better than the witness-based approach [9] in terms of overhead performance for $C = 1, 2$ and $P_f = 0.25, 0.5, 0.75$.

In the second set of simulations, the average numbers of bits sent by uncompromised nodes in the OR scheme is evaluated. When the compromised node does not agree with the best temporary voting result, the fusion result transmitted by the compromised node differs from the other fusion results and the size of the fusion result is $K = 48$. Fig. 8 plots the results for $M = 11$, when $P_f = 0$, 0.5, and 1. In Fig. 8a, when the base station can obtain a valid fusion result, the number of bits that are transmitted by the uncompromised nodes to the base station in the OR scheme increases with the number of compromised nodes C, as expected. However, they are fewer than those in the witness-based approach. Notably, in this case, the number of bits transmitted by uncompromised nodes in the witness-based approach is constant since, once an uncompromised node has been polled, the polling process is completed. For small C, such as $C = 0$, 1, 2, or 3, the number of bits transmitted by uncompromised nodes in the OR scheme is about half that of those in the witness-based approach. Additionally, the performance of the OR scheme when $P_f = 1$ is worst in all three simulations since any compromised node that agrees with the forged result sometimes makes the forged result have the largest number of votes and forces the base station to use it as the best temporary voting result to poll the next node. Then, the next uncompromised node needs to transmit its fusion result to the base station instead of sending only an agreeing vote, increasing the total number of bits transmitted by the uncompromised nodes.

According to that in Fig. 8b, when the base station cannot obtain a valid fusion result, the number of bits transmitted

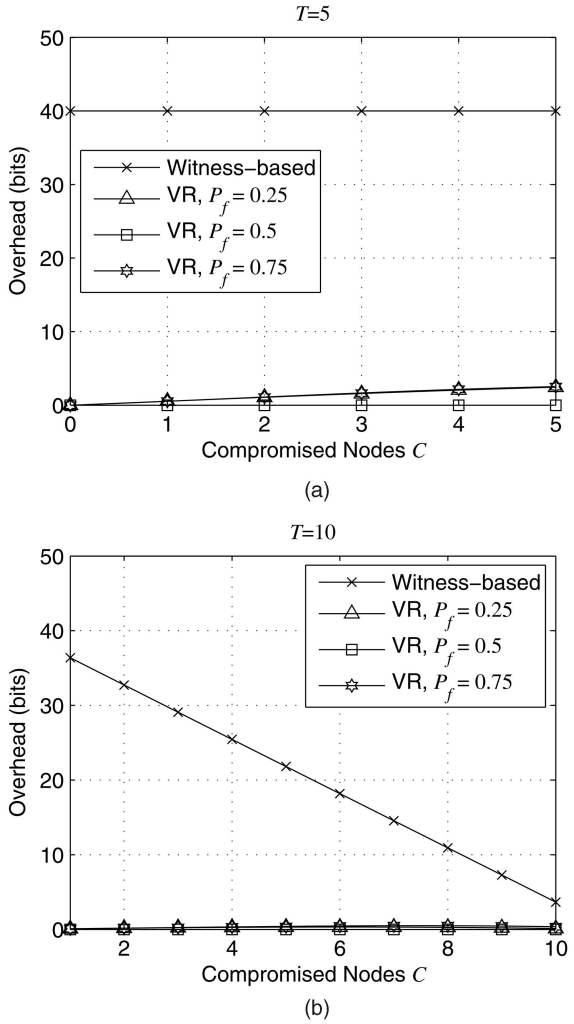

Fig. 7. Overhead comparison between the VR scheme and the witness-based approach [9] for $M = 11$ and $P_f = 0.25, 0.5, 0.75$ (a) when $T = 5$ (valid fusion result) and (b) when $T = 10$ (invalid fusion result).

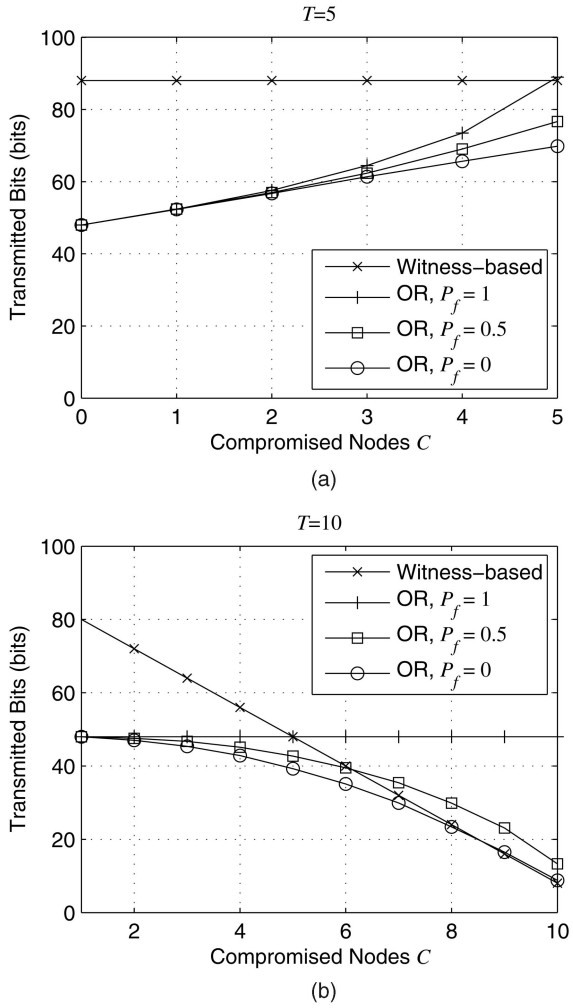

Fig. 8. Comparison of transmitted bits between the OR scheme and the witness-based approach [9] for $M = 11$, $K = 48$, and $P_f = 0, 0.5, 1$ (a) when $T = 5$ (valid fusion result) and (b) when $T = 10$ (invalid fusion result).

by the uncompromised nodes to the base station in the OR scheme decreases as the number of compromised nodes C increases, except for $P_f = 1$. This phenomenon is caused by the fact that the scheme stops when the number of unpolled nodes plus the maximum number of votes for any result that is recorded at the base station is less than T. When $T = M - 1$ as simulated, the recording of the two results at the base station stops the polling process. Recall that, when $P_f = 1$, the only way to stop the polling process is for one fusion result to be sent by an uncompromised node and the other to be sent by a compromised node and for the number of transmitted bits of the uncompromised nodes to be the same for all C. If $P_f \neq 1$, then the two compromised nodes may yield different results and no bit is transmitted by the uncompromised node. This concludes the simulation results. This subfigure reveals that the OR scheme outperforms when C is small but not when C exceeds $T/2$.

5 Conclusions

This work proposes a power-efficient scheme for data fusion assurance in which the base station in the WSN collects the fusion data and the votes on the data directly from the fusion nodes. The proposed scheme is more reliable with less assurance overhead and delay than the witness-based approach. That is, the power and delay associated with the transmission of the fusion result and the votes are significantly decreased. Notably, the proposed schemes are designed for systems whose receiving power is much less than the transmission power. For systems in which the receiving power is on the same order of magnitude as the transmission power, the proposed schemes may not be superior. Table 1 summarizes the effect of receiving power on the VR scheme in which the

TABLE 1
Increase in Overhead Due to Receiving in the VR Scheme When $C = 2$, $M = 11$, $T = 5$, and $P_f = 0$, a 10-Bit Packet Header and a 20-Bit Fusion Result

Ratio of receiving power to transmission power (per bit)	$\frac{1}{4000}$	$\frac{1}{400}$	$\frac{1}{40}$
Increase in overhead (%)	0.4%	4%	41%

average number of transmitted bits for is 11. This table reveals that the increase in overhead due to receiving in the VR scheme is insignificant, even for a receiving power to a transmission power ratio of only 1/400.

APPENDIX A

THE OVERHEAD OF THE WITNESS-BASED APPROACH GIVEN IN [9]

If the received fusion result is not accepted, then the base station may start a polling mechanism to seek the correct fusion result. The base station randomly specifies another fusion node. The new chosen node then sends its fusion result and all MACs from the witness nodes to the base station.[11] When the number of compromised fusion nodes, C, exceeds T, the compromised nodes can cooperate with each other and successfully forge a wrong fusion result that will be accepted by the base station. When $C > M - T - 1$, the number of uncompromised witness nodes is less than T and the correct fusion result cannot then receive enough support. Accordingly, if $T + 1 > C > M - T - 1$, then the fusion result is invalid despite the fact that $M - T$ fusion nodes are chosen to transmit their fusion results to the base station in the polling process.

Since the overhead, as defined, considers only the power consumption of uncompromised nodes, the number of uncompromised nodes, denoted as i, among the $M - T$ chosen nodes must be determined to compute the overhead. However, since the base station randomly specifies the fusion node, i is a random number, where $0 \le i \le M - T$. The probability for each value of i must be calculated and the average overhead is computed from this. In the probability calculation, all compromised nodes (uncompromised nodes) are assumed to behave identically. The calculation can be treated as the problem of counting C black balls (compromised nodes) and $M - C$ white balls (uncompromised nodes) together. First, the total number of possible combinations of all fusion nodes that are chosen by the base station is given by $\binom{M}{C}$. Next, the number of possible combinations of first $M - T$ fusion nodes, including i uncompromised nodes, is given by $\binom{M-T}{i}$. Finally, the number of possible combinations of last T fusion nodes, including $M - C - i$ uncompromised nodes, is determined from $\binom{T}{M-C-i}$. Therefore, the probability that i of the $M - T$ chosen nodes are uncompromised can be calculated by

$$P_w(i) = \frac{\binom{M-T}{i}\binom{T}{M-C-i}}{\binom{M}{C}}.$$

Let K be the number of bits that represent the fusion result. The average overhead is thus

$$O_w = \sum_{i=1}^{M-T} P_w(i)[(M-1)ik_w + K(i-1)](\text{bits}), \quad (1)$$

where $i - 1$ is used, instead of i, because the overhead is defined such that one copy of the correct fusion result is not counted. Equation (1) indicates that the number of the correct fusion results that are transmitted by the uncompromised fusion nodes may be up to $M - T$. Restated, the uncompromised nodes wasted a significant power when the correct fusion result cannot be obtained by the base station. Moreover, since each chosen node must collect all MACs from the witness nodes, the average round delay, R_w, and the average polling delay, D_w, are

$$R_w = M - T \quad \text{and } D_w = (M-T)(M-1), \quad (2)$$

respectively.

Conversely, when the number of uncompromised nodes exceeds T, the base station obtains the correct fusion result. If the base station receives the correct result in round i such that the chosen fusion nodes from round 1 to $i - 1$ are compromised, then the average round delay, the average polling delay, and overhead are given by

$$R_w = \sum_{i=1}^{C+1} i \frac{\binom{M-i}{C-i+1}}{\binom{M}{C}},\ D_w = R_w(M-1), \quad \text{and } O_w = (M-1)k_w, \quad (3)$$

respectively.

APPENDIX B

PERFORMANCE ANALYSIS OF THE VARIANT-ROUND SCHEME WHEN $P_f = 0$

If the compromised node always disagrees with the transmitted fusion result, then no forged fusion result is accepted. Two cases must be addressed.

Case 1. $C \ge M - T$.

Case 2. $C < M - T$.

Notably, the valid fusion result is not available in Case 1.

Assume that the chosen node in the first round **is** compromised. The probability that the chosen node is compromised in the first round is given by C/M. The first round of polling finishes when $M - T$ witness nodes do not agree with the transmitted fusion result, as described in Step 3. Thus, the polling order (which is the order of witness nodes, as described in Step 1) determines the number of uncompromised witness nodes that the base station must poll in this round of polling. The number of possible polling orders, in the sense of the black-white-ball model that was presented in Appendix A, is given by

$$\Pi_{v1}^{c1} = \frac{(M-1)!}{(C-1)!(M-C)!} = \binom{M-1}{C-1},$$

where the subscript, $v1$, denotes the first case of the VR scheme and the superscript, $c1$, represents the first round of polling when the chosen node is compromised. Since the chosen node in the first round of polling is compromised, the polling stops after $M - T$ witness nodes have been polled.

11. All MACs must be sent to the base station again to avoid denial of service, since the previously chosen compromised fusion node might have modified the MACs before forwarding them to the base station. This action is not clearly presented in [9].

This early termination is due to the fact that all of the polled nodes disagree with the transmitted fusion result and the remaining $M-1-(M-T)=T-1$ unpolled nodes are not enough to verify the transmitted fusion result, even when they all agree with the result. Since the number of unpolled nodes is $T-1$, the number of uncompromised nodes among the unpolled nodes is $M-C-i$ if i uncompromised nodes are polled. Hence, the probability that i of the $M-T$ polled nodes are uncompromised, where $0 \le i \le M-T$[12] is then given by

$$\frac{1}{\Pi_{v1}^{c1}} \binom{M-T}{i} \binom{T-1}{M-C-i},$$

where $\binom{M-T}{i}$ is the number of ways to choose i uncompromised nodes from $M-T$ nodes and $\binom{T-1}{M-C-i}$ is the number of ways to choose the remaining $M-C-i$ uncompromised nodes from the $T-1$ unpolled nodes in this round. No node is excluded from the witness set because no node agrees with the transmitted fusion result. The number of compromised nodes, the number of fusion nodes, and the size of the witness set in the second round become $C-1$, $M-1$, and $M-2$, respectively, and only the chosen node is excluded. Let $O_{v1}(M-1,T,C-1,0)$ be the average overhead when the number of fusion nodes is $M-1$ and the number of compromised nodes is $C-1$, where T represents the number of votes required for verification and 0 refers to the case in which $P_f=0$. Then, the average overhead, when the chosen node in the first round is compromised, is expressed recursively by

$$O_{v1}^{c}(M,T,C,0) = \sum_{i=0}^{M-T} \frac{1}{\Pi_{v1}^{c1}} \binom{M-T}{i} \binom{T-1}{M-C-i} ik' + O_{v1}(M-1,T,C-1,0), \quad (4)$$

where k (k') is the number of bits that must be sent by a witness to the base station when it agrees (disagrees) with the transmitted fusion result.[13] The first term in the above formula counts the number of bits that are transmitted to the base station by polled uncompromised nodes when they all disagree with the transmitted fusion result. Moreover, the average round delay and the average polling delay under the same conditions are represented by

$$R_{v1}^{c}(M,T,C,0) = 1 + R_{v1}(M-1,T,C-1,0) \quad (5)$$

and

$$D_{v1}^{c}(M,T,C,0) = M-T+D_{v1}(M-1,T,C-1,0), \quad (6)$$

where $R_{v1}(M-1,T,C-1,0)$ and $D_{v1}(M-1,T,C-1,0)$ are the average round delay and polling delay, respectively, when the number of fusion nodes that are polled for votes is $M-1$ and the number of compromised nodes among them is $C-1$.

12. Actually, $\max\{M-T-C+1,0\} \le i$. Since $\binom{a}{b}$ is 0, when $b>a$, 0 is adopted as the lower bound of i.

13. The bits that are sent to the base station when a node agrees with the fusion result are separated from those that are sent when the node disagrees with the result, since the node can be silent when it agrees with the result, and only a few bits are sent when it disagrees.

Suppose that the node that is chosen in the first round **is not** compromised, the situation for which has a probability of $(M-C)/M$. The number of the possible polling orders is given by

$$\Pi_{v1}^{u1} = \binom{M-1}{C},$$

where the superscript, $u1$, denotes the first round of polling when the chosen node is uncompromised. Since no valid fusion result is available, when the polling stops at witness node j, the node does not agree with the transmitted result and the base station polled $M-T$ disagreeing nodes (including witness node j) (the second termination condition in Step 3 when $M'=M-1$). Furthermore, $M-j-1$ nodes are unpolled and $C-(M-T)=T+C-M$ of these are compromised since $M-T$ compromised nodes must be polled. Since the witness set has $M-C-1$ uncompromised nodes, the maximum number of polled witness nodes is $(M-C-1)+(M-T)$, where all uncompromised nodes that agree with the transmitted fusion result and additional $M-T$ disagreed nodes are polled. Thus, the probability that the polling stops at the jth witness node, where

$$M-T \le j \le (M-C-1)+(M-T) = 2M-T-C-1$$

is given by

$$P_{v1}^{u1}(j) = \frac{1}{\Pi_{v1}^{u1}} \binom{j-1}{M-T-1} \binom{M-j-1}{T+C-M},$$

where $\binom{j-1}{M-T-1}$ denotes that $M-T-1$ disagreeing nodes are polled in the first $j-1$ positions to stop the polling at the jth position.

Since $j \ge M-T$, the number of unpolled nodes $M-j-1$ is less than T. In the following round, the $M-T$ disagreeing nodes, along with these unpolled nodes, are the fusion nodes, which are involved in the polling process. Since $P_f=0$, the first $M-T-1$ will disagree with the transmitted fusion. Now, $M'=M-T-1+M-j-1 \le M-2$ and $M'-T+1$ in the second condition of Step 3 will be equal to or less than $M-T-1$. Therefore, these unpolled nodes will never be polled in the following rounds before the voting mechanism stops. Accordingly, only compromised nodes are polled after the first round and no uncompromised nodes are polled further. Notably, the number of uncompromised polled nodes is $j-(M-T)=j-M+T$. Therefore, the average overhead, when the chosen node in the first round **is not** compromised, is given by

$$O_{v1}^{u}(M,T,C,0) = \sum_{j=M-T}^{2M-T-C-1} P_{v1}^{u1}(j)(j-M+T)k. \quad (7)$$

The size of the witness set after the first round becomes $M'=(M-j-1)+(M-T)-1=2M-T-j-2$ (which is the total number of unpolled and disagreeing nodes minus one). The polling process stops if

$$M' \le T-1 \Leftrightarrow 2M-T-j-2 \le T-1 \Leftrightarrow j \ge 2M-2T-1.$$

Otherwise, the size is decreased by 1 in each following round until it becomes T. Consequently, the total number of the following rounds is $2M-T-j-2-T+1 = 2M-2T-j-1$ and the total number of rounds is $2M-2T-j-1+1 = 2M-2T-j$. The average round delay is then represented by

$$R^u_{v1}(M,T,C,0) = \sum_{j=M-T}^{2M-2T-2} P^{u1}_{v1}(j)(2M-2T-j) + \sum_{j=2M-2T-1}^{2M-T-C-1} P^{u1}_{v1}(j). \tag{8}$$

Since the polling process ends when $M'-T+1$ nodes are polled in each round, the polling delay is

$$D^u_{v1}(M,T,C,0) = \sum_{j=M-T}^{2M-T-C-1} P^{u1}_{v1}(j)j + \sum_{j=M-T}^{2M-2T-2} P^{u1}_{v1}(j)\frac{(2M-2T-j)(2M-2T-j-1)}{2}. \tag{9}$$

Equations (4) to (9) and the initial conditions then give the average overhead and the average delays of Case 1 as

$$O_{v1}(M,T,C,0) = \begin{cases} 0 & M \le T \\ \frac{C}{M} O^c_{v1}(M,T,C,0) + \frac{M-C}{M} O^u_{v1}(M,T,C,0) & \text{else,} \end{cases}$$

$$R_{v1}(M,T,C,0) = \begin{cases} 0 & M \le T \\ \frac{C}{M} R^c_{v1}(M,T,C,0) + \frac{(M-C)}{M} R^u_{v1}(M,T,C,0) & \text{else,} \end{cases} \tag{10}$$

$$D_{v1}(M,T,C,0) = \begin{cases} 0 & M \le T \\ \frac{C}{M} D^c_{v1}(M,T,C,0) + \frac{(M-C)}{M} D^u_{v1}(M,T,C,0) & \text{else.} \end{cases}$$

The second case, $C < M-T$, produces a valid fusion result. Similarly, if the chosen node in the first-round polling is compromised, then the polling stops after $M-T$ witness nodes have been polled. The average overhead and the average delays, when the chosen node in the first round is compromised, are expressed, respectively, as

$$O^c_{v2}(M,T,C,0) = \sum_{i=M-T-C+1}^{M-T} \frac{1}{\Pi^{c1}_{v1}} \binom{M-T}{i} \binom{T-1}{M-C-i} ik' + O_{v2}(M-1,T,C-1,0),$$

$$R^c_{v2}(M,T,C,0) = 1 + R_{v2}(M-1,T,C-1,0),$$

$$D^c_{v2}(M,T,C,0) = M-T+D_{v2}(M-1,T,C-1,0).$$

Only one round of polling is needed when the chosen node is uncompromised in the first round. When the polling stops at witness node j, the node agrees with the transmitted result and the base station has polled T agreeing nodes (including witness node j). Moreover, $M-j-1$ nodes are unpolled, of which $M-1-C-T$ are uncompromised. The probability that the polling process ends at the jth witness node, where $T \le j \le T+C$, is given by

$$P^{u1}_{v2}(j) = \frac{1}{\Pi^{u1}_{v1}} \binom{j-1}{T-1} \binom{M-j-1}{M-C-T-1}.$$

The number of polled uncompromised nodes is T. The average overhead and the average polling delay when the chosen node is uncompromised in the first round are represented as

$$O^u_{v2}(M,T,C,0) = \sum_{j=T}^{T+C} P^{u1}_{v2}(j)Tk, \; D^u_{v2}(M,T,C,0) = \sum_{j=T}^{T+C} P^{u1}_{v2}(j)j.$$

Consequently, the average overhead $O_{v2}(M,T,C,0)$, the average round delay $R_{v2}(M,T,C,0)$, and the average polling delay $D_{v2}(M,T,C,0)$ can be represented as

$$O_{v2}(M,T,C,0) = \begin{cases} 0 & M \le T \\ Tk & M > T \text{ and } C = 0 \\ \frac{C}{M} O^c_{v2}(M,T,C,0) + \frac{M-C}{M} O^u_{v2}(M,T,C,0) & \text{else,} \end{cases}$$

$$R_{v2}(M,T,C,0) = \begin{cases} 0 & M \le T \\ 1 & M > T \text{ and } C = 0 \\ \frac{C}{M} R^c_{v2}(M,T,C,0) + \frac{M-C}{M} & \text{else,} \end{cases}$$

$$D_{v2}(M,T,C,0) = \begin{cases} 0 & M \le T \\ T & M > T \text{ and } C = 0 \\ \frac{C}{M} D^c_{v2}(M,T,C,0) + \frac{M-C}{M} D^u_{v2}(M,T,C,0) & \text{else.} \end{cases} \tag{11}$$

ACKNOWLEDGMENTS

The preliminary results of this work were presented at the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC '06) held in Taiwan, Republic of China. The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financially supporting this research under Contract Nos. NSC 94-2213-E-305-002 and NSC 94-2213-E-305-001.

REFERENCES

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks," *IEEE Comm. Magazine*, vol. 38, no. 8, pp. 102-114, Aug. 2002.

[2] S.A. Aldosari and J.M.F. Moura, "Detection in Decentralized Sensor Networks," *Proc. Int'l Conf. Acoustics, Speech, and Signal Processing*, pp. 277-280, May 2004.

[3] R. Anderson and M. Kuhn, "Tamper Resistance—A Cautionary Note," *Proc. Second Usenix Workshop Electronic Commerce*, pp. 1-11, Nov. 1996.

[4] J.-F. Chamberland and V.V. Veeravalli, "Asymptotic Results for Decentralized Detection in Power Constrained Wireless Sensor Networks," *IEEE J. Selected Areas Comm.*, vol. 2, no. 6, pp. 1007-1015, Aug. 2004.

[5] D. Culler, D. Estrin, and M. Srivastava, "Overview of Sensor Networks," *Computer,* vol. 37, no. 8, pp. 41-49, Aug. 2004.
[6] L. Dan, K.D. Wong, H.H. Yu, and A.M. Sayeed, "Detection, Classification, and Tracking of Targets," *IEEE Trans. Signal Processing,* vol. 19, no. 3, pp. 17-29, Mar. 2002.
[7] A. D'Costa, V. Ramachandran, and A.M. Sayeed, "Distributed Classification of Gaussian Space-Time Sources in Wireless Sensor Networks," *IEEE J. Selected Areas Comm.,* vol. 22, no. 6, pp. 1026-1036, Aug. 2004.
[8] J. Deng, R. Han, and S. Mishra, "A Performance Evaluation of Intrusion-Tolerant Routing in Wireless Sensor Networks," *Proc. Int'l Workshop Information Processing in Sensor Networks,* pp. 349-364, 2003.
[9] W. Du, J. Deng, Y.S. Han, and P.K. Varshney, "A Witness-Based Approach for Data Fusion Assurance in Wireless Sensor Networks," *Proc. IEEE Global Telecomm. Conf.,* vol. 3, pp. 1435-1439, Dec. 2003.
[10] Y. Lin, B. Chen, and P.K. Varshney, "Decision Fusion Rules in Multi-Hop Wireless Sensor Networks," *IEEE Trans. Aerospace and Electronic Systems,* vol. 41, no. 2, pp. 475-488, Apr. 2005.
[11] J.N. Al-Karaki and A.E. Kamal, "Routing Techniques in Wireless Sensor Networks: A Survey," *IEEE Wireless Comm. Magazine,* vol. 11, no. 6, pp. 6-28, Dec. 2004.
[12] G. Mergen, Q. Zhao, and L. Tong, "Sensor Networks with Mobile Access: Energy and Capacity Considerations," *IEEE Trans. Comm.,* to appear.
[13] D. Niculescu, "Positioning in Ad Hoc Sensor Networks," *IEEE Network,* vol. 18, no. 4, pp. 24-29, July-Aug. 2004.
[14] D. Niculescu, "Communication Paradigms for Sensor Networks," *IEEE Comm. Magazine,* vol. 43, no. 3, pp. 116-122, Mar. 2005.
[15] S. Olariu, A. Wadaa, L. Wilson, and M. Eltoweissy, "Wireless Sensor Networks: Leveraging the Virtual Infrastructure," *IEEE Network,* vol. 18, no. 4, pp. 51-56, Apr. 2004.
[16] S. Olariu and Q. Xu, "Information Assurance in Wireless Sensor Networks," *Proc. IEEE Int'l Parallel and Distributed Processing Symp.,* p. 236a, Apr. 2005.
[17] H.-T. Pai and Y.S. Han, "Power-Efficient Direct-Voting Assurance for Data Fusion in Wireless Sensor Networks," http://arxiv.org/abs/0705.3683, May 2007.
[18] A. Perrig, J.A. Stankovic, and D. Wagner, "Security in Wireless Sensor Networks," *Comm. ACM,* vol. 47, no. 6, pp. 53-57, June 2004.
[19] B. Przydatek, D. Song, and A. Perrig, "SIA: Secure Information Aggregation in Sensor Networks," *Proc. Conf. Embedded Networked Sensor Systems,* pp. 255-265, Nov. 2003.
[20] R.C. Shah, S. Roy, S. Jain, and W. Brunette, "Data Mules: Modeling a Three-Tier Architecture for Sparse Sensor Networks," *Proc. IEEE Workshop Sensor Network Protocols and Applications,* pp. 30-41, May 2003.
[21] C.-C. Shen, C. Srisathapornphat, and C. Jaikaeo, "Sensor Information Networking Architecture and Applications," *IEEE Personal Comm. Magazine,* vol. 8, no. 4, pp. 52-59, Aug. 2001.
[22] E. Shih, S.-H. Cho, N. Iches, and R. Min, "Physical Layer Driven Protocol and Algorithm Design for Energy-Efficient Wireless Sensor Networks," *Proc. MobiCom '01,* pp. 272-286, 2001.
[23] P. Sholander, A. Harris, and J. Brown, "Intersensor Information Assurance for DOD Tactical Networks," *Proc. Military Comm. Conf.,* vol. 2, pp. 1456-1461, Oct. 2002.
[24] F. Sivrikaya and B. Yener, "Time Synchronization in Sensor Networks: A Survey," *IEEE Network,* vol. 18, no. 4, pp. 45-50, July-Aug. 2004.
[25] K. Sohrabi, W. Merrill, J. Elson, L. Girod, F. Newberg, and W. Kaiser, "Methods for Scalable Self-Assembly of Ad Hoc Wireless Sensor Networks," *IEEE Trans. Mobile Computing,* vol. 3, no. 4, pp. 317-331, Oct.-Dec. 2004.
[26] L. Tong, Q. Zhao, and S. Adireddy, "Sensor Networks with Mobile Agents," *Proc. IEEE Military Comm. Conf.,* pp. 688-693, Oct. 2003.
[27] J.N. Tsitsiklis, "Decentralized Detection by a Large Number of Sensors," *Math. Control, Signals, and Systems,* vol. 1, no. 2, pp. 167-182, 1988.
[28] D. Wagner, "Resilient Aggregation in Sensor Networks," *Proc. ACM Workshop Security of Ad Hoc and Sensor Networks,* pp. 78-87, Oct. 2004.
[29] T.-Y. Wang, Y.S. Han, B. Chen, and P.K. Varshney, "A Combined Decision Fusion and Channel Coding Scheme for Distributed Fault-Tolerant Classification in Wireless Sensor Networks," *IEEE Trans. Wireless Comm.,* vol. 5, no. 7, pp. 1695-1705, 2006.
[30] Z. Yang and L. Tong, "Cooperative Sensor Networks with Misinformed Nodes," *IEEE Trans. Information Theory,* vol. 51, no. 12, pp. 4118-4133, Dec. 2005.
[31] W. Zhao and M.H. Ammar, "Message Ferrying: Proactive Routing in Highly-Partitioned Wireless Ad Hoc Networks," *Proc. IEEE Workshop Future Trends in Distributed Computing Systems,* pp. 308-314, May 2003.

Hung-Ta Pai received the BSc degree in electrical engineering from National Tsing Hua University, Taiwan, in 1992, and the MSc and PhD degrees in electrical engineering from the University of Texas at Austin in 1996 and 1999, respectively. He served as an Army officer from 1992 to 1994. He was a research assistant at the University of Texas at Austin from 1995 to 1999. He worked as a senior design engineer at Silicon Integrated Systems Corp from 1999 to 2001. From 2001 to 2002, he was an assistant professor in the Department of Electrical Engineering at Tatung University, Taiwan. He was an assistant professor in the Department of Electrical Engineering at National Taiwan University of Science and Technology, Taiwan, from 2002 to 2004. Since August 2004 he has been an assistant professor in the Graduate Institute of Communication Engineering at National Taipei University, Taiwan. His research interests include signal processing and communication systems. He is a member of the IEEE.

Yunghsiang S. Han received the BSc and MSc degrees in electrical engineering from the National Tsing Hua University, Hsinchu, Taiwan, in 1984 and 1986, respectively, and the PhD degree from the School of Computer and Information Science, Syracuse University, Syracuse, New York, in 1993. From 1986 to 1988, he was a lecturer at Ming-Hsin Engineering College, Hsinchu, Taiwan. He was a teaching assistant from 1989 to 1992 and a research associate from 1992 to 1993 in the School of Computer and Information Science at Syracuse University. From 1993 to 1997, he was an associate professor in the Department of Electronic Engineering at Hua Fan College of Humanities and Technology, Taipei Hsien, Taiwan. He was with the Department of Computer Science and Information Engineering at National Chi Nan University, Nantou, Taiwan, from 1997 to 2004. He was promoted to a professor in 1998. He was a visiting scholar in the Department of Electrical Engineering at University of Hawaii, Manoa, Hawaii, from June to October 2001 and the SUPRIA visiting research scholar in the Department of Electrical Engineering and Computer Science and CASE Center at Syracuse University from September 2004 to January 2004. He is now with the Graduate Institute of Communication Engineering at National Taipei University, Taipei. His research interests are in wireless networks, security, and error-control coding. He was a recipient of the 1994 Syracuse University Doctoral Prize. He is a member of the IEEE.