IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 3 December 2022, accepted 14 December 2022, date of publication 16 December 2022,
date of current version 21 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3230362

RESEARCH ARTICLE

Energy-Aware Cross-Layer Technique for Countering Traffic Analysis Attacks on Wireless Sensor Network

YOUSEF EBRAHIMI AND MOHAMED YOUNIS, (Fellow, IEEE)
Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250, USA
Corresponding author: Yousef Ebrahimi (yousef2@umbc.edu)

ABSTRACT The vital role of a base station (BS) in a wireless sensor network (WSN) has made it a favorable target in hostile environments. Despite attempts to physically make the BS hidden to prying eyes, traffic analysis would give an adversary insight into the network topology and the BS whereabouts. Evidence Theory (ET) is a prominent methodology for performing such an analysis. Unfortunately, all existing countermeasures not only overlook patterns of energy usage in WSNs, but also impose untamed overhead that shortens the WSN lifetime. In this paper, we first propose a novel energy-aware and multi-zone scheme to significantly reduce the overhead of countermeasures on highly overburdened nodes in the BS proximity, and hence significantly improve the WSN lifespan. We also show how our proposed scheme improves resilience against ET via diminishing the collected evidence by an adversary. We then propose a novel cross-layer technique that exploits transmission range adjustment to confuse the adversary about the data paths. This results in a versatile and effective countermeasure that significantly improves anonymity of the BS. The performance is validated through extensive simulation experiments.

INDEX TERMS Anonymity, evidence theory, location privacy, traffic analysis, wireless sensor networks.

I. INTRODUCTION

Thanks to lower production cost of electronic devices, a large deployment of sensor nodes has become possible; especially in remote and/or hostile environments. These nodes are equipped with radio transmitters to enable the establishment of communication links and forming a network [1]. A WSN constitutes an effective low-cost option for monitoring vast areas. Hence, a WSN is a suitable solution for applications such as border protection, security surveillance, fire detection, combat field reconnaissance, target tracking, etc. [2], [3]. In such WSN deployments, sensor nodes rely on multi-hop routing techniques to deliver their collected data to an in-situ unit – Base Station (BS) – for processing, analysis, and long-haul transmission off the field [2]; Fig. 1 shows an example. Since such a BS role is vital to the WSN operation, a motivated adversary would target the BS instead of the individual sensors. Therefore, it is crucial to protect the BS by concealing its location [4].

The associate editor coordinating the review of this manuscript and approving it for publication was Jiankang Zhang.

A. TRAFFIC ANALYSIS THREAT

Even if the BS is camouflaged, an adversary can intercept and analyze packets to distinguish the BS. While packet encryption and anonymous routing prevent identifying the BS by inspecting the sniffed packets [5], [6], [7], [8], they do not provide sufficient protection where an eavesdropper can intercept radio transmissions and apply traffic analysis techniques to uncover the network topology and locate the BS [9], [10], [11]. To mitigate such a threat, many countermeasures are proposed to prevent, or at least delay, an adversary from locating the BS. Many of these countermeasures aim to change the traffic pattern in the network to confuse the adversary [12], [13], [14], [15], [16], [17], [18], [19], [20], [21]. Others, e.g., [13], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], and [33] introduce extra fake packets to hide the data routes and hence lead the adversary to wrong locations. In almost all countermeasures, improving

the BS location anonymity comes at the expense of higher transmission rates by the individual nodes.

B. ENERGY CONCERN

A WSN is a wirelessly formed network of sensing devices that are typically deployed in hard-to-reach or harsh/hostile environments. A sensor node is typically powered by a battery, with a limited capacity and no practical means of replenishment. Therefore, conservative use of the onboard energy is of utmost importance to prolong the life of each node and as the result the effective WSN lifespan. Multi-hop data routing is a popular methodology for reducing energy consumption [2], [34], [35]; yet the route across the area converges towards the BS causing the nearby sensor nodes to experience large volumes of relaying traffic and deplete energy at a higher rate. For example, node A in Fig. 1 relays packets from nine other nodes and consequently its lifespan diminishes. Upon the death of its close nodes, the BS becomes increasingly unreachable and the WSN utility degrades. Traffic analysis countermeasures often worsen the energy concerns due to the increased transmissions.

C. CONTRIBUTION

Even though many studies in the literature have aimed to prolong WSN lifespan by energy-aware topology formation [36], routing [37], or multi-objective design optimization [38], the subject has not been analyzed from traffic analysis point of view. Therefore, this paper fills the gap by opting to overcome the shortcomings of contemporary traffic analysis countermeasures by devising energy-efficient mechanisms that not only boost the BS anonymity but also factor in the load on the individual nodes so that the WSN lifespan is sustained. First, the energy consumption profile of sensor nodes is studied to identify depletion rates for the k-hop neighbors of the BS. Then, a novel mechanism divides the network into multiple zones based on energy and anonymity metrics. Our energy-aware multi-zone mechanism increases the adversary's uncertainty about the BS whereabout without shortening the time for the first node to die in the network. We further enhance our zone-based mechanism by incorporating a link-layer anonymity boosting measure. The cross-layer design proves to be quite effective with respect to both anonymity and energy metrics. The simulation results validate the performance advantage of the proposed mechanisms.

D. ORGANIZATION

The rest of this paper is organized as follows. The contribution is distinguished from related work in Section II. The system and threat models are presented in Section III. The WSN energy overhead is analyzed in Section IV. Our energy-ware multi-zone management strategy is detailed in Section V. In Section VI, we present our cross-layer design. The simulation results are presented in Section VII. Finally, the paper is concluded in Section VIII.

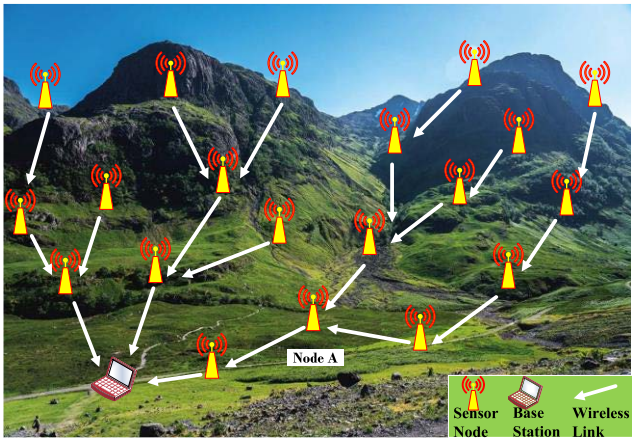


FIGURE 1. Multi-hop routing in WSN results in convergence of data paths as they get closer to the BS. Node A is an example of those sensors that experience high relaying traffic as it is a join point for multiple data paths.

II. RELATED WORK

An adversary would aim at identifying the role, location, and identity of WSN nodes to gain knowledge about the structure of the network and distinguish vital assets that could be attacked [39], [40]. Therefore, privacy and security in WSN has received a lot of attention from the research community [9], [10], [11], [12], [13], [14], [22], [26], [27], [28], [29], [30], [31], [41], [42], [43]. Particularly, concealing the role and location of BS is deemed to be more crucial [9], [44]. To eliminate the possibility of information leakage through packet sniffing, encryption of the packet header and payload, employing anonymous routing techniques are often applied [5], [7], [8], [45], [46], [47]. Nonetheless, the adversary may pursue traffic analysis where sensor transmissions are intercepted and correlated to extrapolate data routes and locate the BS [10], [11], [48]. ET is an information theory model [49], [50], that correlates transmissions to deduce link-relation among nodes. That has made ET a prominent traffic analysis methodology [9], [26], [27], [29], [41].

A. CONTEMPORARY COUNTERMEASURES

To boost the anonymity of BS against ET, researchers have proposed countermeasures that fall into three categories:

1) EXPLOITING BS MOBILITY

Given that an adversary opts to find the BS, a logical countermeasure is to frequently move the BS if feasible. Each relocation of BS directly impacts traffic patterns and hence invalidates the ET analysis. Liu et al. [51] propose moving the BS in a semi-random circular pattern. Some nodes are designated across the network to temporarily store data until BS moves close to them to collect the data. On the other hand, Kumar et al. [32] divide the network into multi-layer rings and let BS move freely and send its location to nodes in the central ring. Other nodes query the central ring to learn about the actual location of BS. Such design generates hotspot regions in the network to attract the adversary's attention

away from the BS. Similar approach is pursued in [52], but the possible BS locations are stored in pre-selected nodes. Acharya et al. in [26] assess the BS anonymity and relocate it when the threat level reaches a certain threshold. In [53], a combination of mobile and stationary BS are considered, in which mobile BSs move to lowest anonymity regions to lower the adversary's success.

2) ALTERATION OF DATA ROUTES

Some countermeasures try to disturb the adversary's analysis by altering the normal traffic pattern in the network. In [15], all packets are detoured to a dummy BS before being re-routed back to actual BS. The aim is to make the dummy BS a more interesting target for the adversary. Wright et al. [16] apply the same idea using multiple dummy sinks to not only introduce fake attractive attack-targets but also spread the traffic across the network to further confuse the adversary. El-Bardy et al. [17] pursue a similar approach using multiple real, rather than dummy, BS to eliminate the need for re-routing traffic back to a single BS. Unlike the aforenoted techniques which are centered around the BS, random walk (RW) and multi-parenting (MPR) introduce redundant data relaying [13]. In RW, each node randomly sends its data packet to a neighboring node. The goal is to disperse the traffic by randomly creating different routes to the BS. However, such an approach results in high packet delivery delay. MPR mitigates the delay issue by only sending the data packet to neighbors that are closer to the BS, while introducing controlled randomness in route creation. In [18], source and sink anonymity is achieved by having multiple streams, each carrying a portion of an image and reconstructing it at the destination. Obviously, breaking up a small sensing data increases the number of packets and the associated overhead.

Delaying the packet forwarding is used as the means of affecting normal routing in [19], especially with the focus on impacting temporal correlation of consecutive packets. Similarly, Chakraborty et al. [20] use forwarding delay as a means of blending traffic from different sources into one stream and confuse the adversary. Meanwhile, L-SRA [21] strives for having the same transmission rate among all nodes via buffering. It utilizes the number of active nodes at a given time to set the rate along the route to the BS. Generally, boosting the delays is not suitable for time sensitive applications. In [14], geographic routing is altered in the BS vicinity to give the illusion of a void. Route alteration is achieved through careful node selection on the boundary of the illusive void region so that their transmission reaches the BS. The data packets are routed on boundary nodes and further away from the BS to divert the adversary's attention.

3) GENERATION OF DECEPTIVE TRAFFIC

Many countermeasures aim to boost the BS anonymity by generating bogus packets. Deng et al. [13] create hot spots to divert attention away from the BS. In [32], a multi-layer ring is used to generate and route fake packets. A mobile BS triggers a fake flood by sending a packet to the farthest ring from its current position. Bicakci et al [33] also use a fake flood approach without the help of a mobile BS. For each data packet sent to BS, a fake packet is sent to every other node in the network. The approach opts to equalize the incoming and outgoing packets for every node and the BS to make the adversary see them alike. ATA [23] generates fake packets without routing them. Each node estimates the transmission rate of its parent and generates enough fake packets to match it. ATA aims to have uniform traffic volume across all nodes. Similar to ATA, PLAUDIT [28] strives to have uniform transmission rate across the network, but by using a corona-based load-balanced routing tree. Meanwhile, MSCLP [24] uses bogus packets in cluster-based topologies. Before a node within a cluster transmits its data packet, the cluster head sends a bogus packet to circulate within the cluster and de-signify the importance of upcoming data transmission. In [26], instead of nodes, BS selectively generates fake packets, called BAR, with varying time-to-live value to send out in random directions. The goal is to trick the adversary to assume that the BS is a relay node. A similar idea is pursued in [27] and [29] for two-tier routing topologies. Unlike the aforenoted work, DP [22], MoRF [31], MSI [30], and IATA [23] use fake sinks to route the generated bogus traffic to them. They strive to maximize the impact of bogus traffic through controlled routing while giving the illusion of multiple BS nodes in the network.

B. INCURED OVERHEAD

Nothing comes for free [54] and security/privacy is no exception [55]. Anonymity boost achieved by any countermeasure has its own price tag, most notably:

1) NODE/NETWORK COMPLEXITY

Countermeasures that exploit mobility would clearly expect the BS to be able to physically move. Even though cost associated with this technique might be justifiable in certain setups, e.g., in a combat field, it may not be feasible or practical in many applications. Similarly, techniques that require the deployment of multiple BS units incur increased cost and logistical overhead. Usually, BS units are not as cheap and disposable as sensor nodes.

2) DATA DELIVERY LATENCY

Countermeasures that rely on altering the data route to boost the anonymity, delay packet arrival to the BS. In essence, pursuing inefficient routes often extends the length of data paths. The increased data delivery delay degrades the WSN responsiveness in time sensitive applications, such as target tracking.

3) ENERGY CONSUMPTION

In addition to delay, countermeasures that alter the data routes introduce additional packet relaying and increase energy consumption. Similarly, countermeasures that injects redundant packets in the network imposes a high energy overhead. Such

energy cost is particularly detrimental for heavily loaded nodes in the proximity to BS.

In this paper, we specifically consider the countermeasures from an energy cost perspective. Firstly, instead of looking at energy usage at the network level, we study energy consumption at the node level. Such an approach better reveals the actual overhead of countermeasures on nodes in the BS vicinity, and hence gauges their impact on the WSN lifespan. Unlike the work surveyed above, we promote a dynamic and adaptable design that does not impose high energy usage on the nodes in the BS proximity, and hence does not degrade network longevity. Our multi-zone scheme is geared for boosting anonymity while carefully distributing the energy overhead based on the typical involvement of nodes in normal WSN operation. To further enhance our multi-zone scheme, a cross-layer feature is added which uses transmission increase range to degrade the ET analysis.

Table 1 provides a comparative summary of the key distinguishing features of our work against related work in literature.

III. SYSTEM AND THREAT MODELS

A. NETWORK MODEL

This paper considers a WSN that consists of randomly deployed stationary sensor nodes that operate unattended in a hostile environment. The network is deemed to be fully functional as long as all initially deployed nodes are operational and can reach the BS to deliver their data. All sensor nodes have the same maximum radio range and processing capabilities and are deployed with the same initial battery energy. Energy replenishment is not feasible and node batteries cannot be manually replaced in the field. A more capable entity, namely, the BS, is also deployed in WSN to act as in-situ command and control unit to collect all the data from sensor nodes and process them locally or transmit through backend connection to an off-site center. The network lifetime is measured as the time for the first sensor node to die.

TABLE 1. Key features distinguishing the proposed approach from published work.

Countermeasure Category	Major imposed overhead	Energy-aware	Degrade WSN lifespan
BS Mobility	Node/network design complexity & Extra energy consumption	No	Yes
Route Alteration	Increased data delivery latency & Extra energy consumption	No	Yes
Deceptive packets	Extra energy consumption	No	Yes
Proposed approach	None	Yes	No

Data is routed over multi-hop paths to the BS to conserve energy by minimizing the sum of the distance squared between nodes on the paths [34]. Data packets have the same priority where a first-in, first-out (FIFO) queue is provisioned in all sensor nodes. Data compression and data aggregation are not used, and all incoming packets are forwarded without alteration. This approach is essential especially in event triggered applications for which raw data is required for processing. A shortest path route selection strategy is assumed in the discussion; yet any routing algorithm can be used for determining the data paths. The BS and sensor nodes are assumed to be physically camouflaged such that an adversary does not distinguish them from the environment. To eliminate the risk of information leakage through packet sniffing and header inspection, all data packets are encrypted via pairwise keys. Packet headers, including MAC address and IP, are also assumed to be encrypted [8], [45].

B. ADVERSARY MODEL

The adversary would prioritize locating and isolating, or physically damaging the BS to inflict the most impact on the WSN operation. A global and passive eavesdropper is assumed in this paper [44]. The adversary relies on deployment of multiple antennas across the area of interest to intercept all transmissions by three antennas or more. Such interception capabilities allow relatively accurate trilateration or triangulation of the transmission source [56], [57], [58], [59]. Adversary has sufficient processing and storage power to intercept and record all transmissions in WSN. By pursuing encrypted packet header and data payload the adversary is left with link-layer based traffic analysis as the only option. For that, spatial correlation of transmissions and traffic density are used [13], [49]. Evidence Theory (ET) is the dominant and advanced attack model for traffic analysis and is assumed to be used by the adversary in this paper [49]. In ET, each intercepted transmission is considered as evidence of direct link between a transmitter-receiver pair in the network. Even though the transmitter can be localized, the receiver can be at any point within the transmitter's range.

Fig. 2 illustrates how this analysis might be performed. After node S_3 transmits, its location and radio range are estimated by the adversary. The shaded area shows where the adversary suspects the receiver is residing. When the time comes for S_2 to relay forward the packet it received from S_3, the adversary locates and estimates its transmission range. Knowing the location of S_3, its range, and location of S_2, the adversary concludes that with high probability the earlier transmission from S_3 was sent to S_2. Hence, the adversary records a link-relation between the two nodes as $S_3 \rightarrow S_2$. As S_3's data is traversing the network to reach the BS, the adversary repeats the same steps and eventually forms the path $S_3 \rightarrow S_2 \rightarrow S_1 \rightarrow S_0 \rightarrow BS$. Since the BS is a sink for all data in a WSN, it will not relay incoming packets, and hence its existence at the end of each path is an educated guess. Through ET, the adversary infers more paths pointing to the same region that the BS is residing in, i.e., $S_5 \rightarrow S_4 \rightarrow BS$ in the example. Both paths point to the same region and hint at the BS presence in that vicinity.

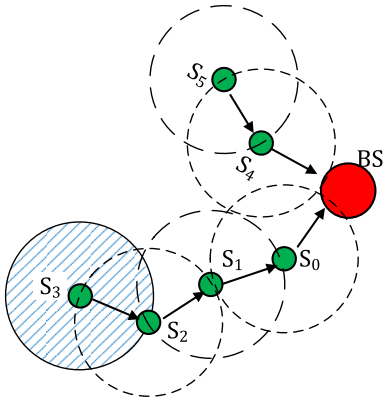


FIGURE 2. Illustrating how transmissions could be correlated to perform traffic analysis.

C. ANONYMITY ASSESSMENT

The example of Fig. 2 reflects a node level analysis, i.e., the adversary locates the individual sensor nodes and correlates their transmissions using ET. In [44] it was shown that such node-level analysis is not practical given the uncertainty about the source position, e.g., due to limited localization accuracy. Hence, a more suitable approach for the adversary would be grid-level analysis, where the area is divided into cells. All nodes within a cell are abstracted to one node represented by the cell. Such a grid approach significantly reduces the computational complexity and eliminates the need for an adversary to keep a record of past transmissions to correlate them with future ones. The latter is achieved because the grid cells are already predetermined, and the adversary knows which neighbors are reachable after each transmission without waiting for a future transmission. Fig. 3 shows an example of node-level analysis (Fig. 3-a) that is abstracted to the cell-level (Fig. 3-b). In Fig. 3-b, the adversary tracks only 4 cells rather than 8 sensor nodes in Fig. 3-a.

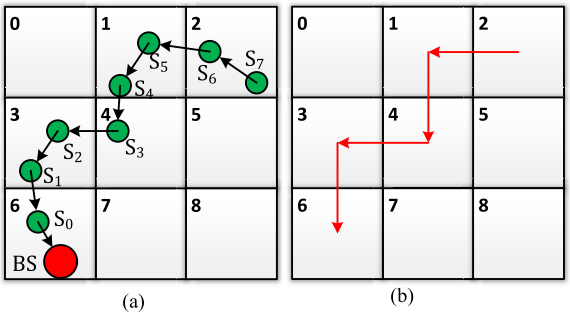


FIGURE 3. Node level vs. Grid level ET analysis.

As the adversary intercepts transmissions, it correlates them and builds up paths stretching from the source of transmission towards the BS. Each transmission is ET evidence, and each path is a potential proof of a route ending at the BS. Obviously, this process is not error prone, and mistakes could result in false evidence and wrong route identification. To process all collected evidence and overcome potential errors, ET defines path-based evidence (*PE*) as:

$$PE\,(L) = \min_{U \subseteq L} E\,(U)\,, \quad |L| \geq 2 \tag{1}$$

To express *PE* in proportion to all collected paths, a normalized value is calculated by dividing *PE* by *Total Evidence*. The latter is the sum of all the evidence collected by the adversary.

$$PE_{norm}\,(L) = \frac{PE(L)}{TotalEvidence} \tag{2}$$

D. Haung [49] has introduced a *Belief* function that represents the anonymity of a cell x based on the evidence of the set of paths P ending at it, Eq. (3)

$$Bel\,(x) = \sum_{L|LP} PE_{norm}\,(L) \tag{3}$$

The *Belief* represents the adversary confidence that cell x is the end point of a path P. A normalized *Belief* – that is result of dividing *Belief* to *Total Belief* of all cells – helps an adversary to weigh the findings in comparison to other *Belief* values, and is defined as follows:

$$Bel_{norm}\,(x) = \frac{Bel(x)}{TotalBeleif} \tag{4}$$

IV. ENERGY OVERHEAD ANALYSIS

To conduct traffic analysis and locate the BS, an adversary would use ET to spatially correlate the intercepted transmissions. As covered in Section II, published countermeasures rely on either introducing fake traffic or altering data routes. The former opts to disrupt the ET analysis and mislead the adversary; yet it imposes energy overhead. Similarly, route alteration causes data to be disseminated over energy-inefficient data paths. This section opts to analyze how such energy overhead becomes a real determent of the network lifetime. Such analysis motivates our novel optimization strategy presented in Section V. Our analysis focuses on the overhead distribution among network nodes; for that we consider proximity, in terms of hop count (level), to the BS. The rationale is that close nodes to the BS are already consuming energy at a high rate due to the pursuance of multi-hop routes. Since the data paths are in essence branches on a tree that is rooted at the BS and these branches tend to merge as they get closer to the BS. We use simulation to generate an energy profile of k-hop neighbors of the BS, where k is the depth of the routing tree. We consider two countermeasures, namely, Differential Fractal Propagation (DFP) and Random Walk (RW) [13]. The former is a prominent representative of countermeasures that introduce fake traffic, while the latter is an example of approaches that pursue route alteration strategies. As a baseline for comparison, we use a routing tree based on least-communication energy.

A. SIMULATION ENVIRONMENT AND SETUP

1) SIMULATION PARAMETERS

The simulation environment is an event driven WSN written in Java. The network serves target tracking applications. To capture the performance of countermeasures at the network layer and nullify the impact of medium access collision on the results, we have assumed high link bandwidth. The network has 200 sensor nodes and one BS that are deployed randomly in an area of $1000 \times 1000\ m^2$. Sensor data is disseminated to the BS over the least-cost paths, where the communication energy is used as the link weight. All nodes have the same capabilities and initial on-board energy. A node generates data when a target is sensed in its vicinity. Targets randomly enter the WSN area and move across it. On average 8 targets are active in the area. To allow for fair comparison between experiments, target generation and motion patterns are recorded and replayed in experiments. The relevant sensor parameters are shown in Table 2. The communication and energy dissipation parameters are adopted from [60] and shown in Table 3. The size and type of packets are shown in Table 4. All payloads and headers are assumed to be encrypted. The adversary is assumed to use an 8×8 grid, i.e., 64 cells, to conduct ET base traffic analysis [44].

2) TOPOLOGY SET

A total of 50 unique topologies are used in each simulation run. The results are averaged over all 50 topologies. All topologies are generated randomly with two selection criteria:

- No two topologies have the BS in the same cell in an 8×8 grid.
- Ten topologies (i.e., 20% of the considered topologies) have the same maximum sensor level, where such a maximum varies between 9 and 13.

The level of a node represents the number of hops on the shortest path that a data packet generated by such a node takes to reach the BS. The notation of L_d represents all nodes that have level d. The first criterion ensures the diversity of BS location. Meanwhile the second criterion prevents dense topologies, i.e., with small routing tree depth, from dominating the results.

TABLE 2. Sensor parameters.

Sensing Range	120 m
Base Transmission Range	120 m
Max Transmission Range	240 m
Buffer Size	15 packets
Duty Cycle	1 sec

B. ENERGY OVERHEAD

We have used the average energy per node as a metric and distinguished among the nodes based on their level on the routing tree. Fig. 4 shows how the considered countermeasures significantly increase energy consumption across all sensor levels. Any extra energy overhead shortens the operation lifespan of the involved nodes and expedites the time until the first node dies, which is commonly used as a metric to gauge the lifetime. Depending on the location of a node in the network, such early death of a node could result in (i) network segmentation in which network would not be able to operate in its full capacity, (ii) complete isolation of BS when routing paths cannot reach BS, or (iii) little impact if the node is at the end of routing path far away from BS. As seen in Fig. 4, nodes in the lower levels (closer to BS) are already experiencing high traffic volume of relaying packets. The extra burden introduced by a traffic analysis countermeasure further taxes their energy reserves and accelerates their death. As was mentioned earlier, death of nodes closer to BS could lead to BS isolation and disruption of the network operation. In Fig. 4, sensors in level 1 consume 1.7 and 2 times more energy than the baseline case under RW and DFP, respectively. The peak overhead for both RW and DFP is for sensors at level 2, where the energy consumption rate is 2.3 and 3.7 times the baseline. In other terms, RW and DFP shorten the WSN lifetime by 230% and 370%, respectively.

TABLE 3. Communication and energy model parameters.

Energy dissipated in transceiver electronics	50 nJ/bit
Energy dissipated by the transmitter amplifier	100pJ.bit/m^2
Path loss factor	2
Data transmission bit rate	2 Mbps

TABLE 4. Details of packets.

Data packet size	10 Kbits
Routing packet size	2 Kbits
Dummy/Fake packets	10 Kbits

It should be mentioned that the countermeasures aim to improve the BS anonymity. Using the *Success Rate* of an adversary's finding of the BS as a metric [44], [25], our experiments show that such a rate is 60%, 48% and 20%, for the baseline, RW and DFP, respectively. Thus, the 12% and 40% reductions achieved by RW and DFP come at a high price in terms of network lifetime. In the next section, we propose an optimization strategy to mitigate such a side effect on the network lifespan. The basic idea is to utilize nodes to boost anonymity without raising their energy consumption higher than the max value that a normal WSN operation dictates, which in essence is the peak value (level 1) of the baseline curve in Fig. 4.

V. MULTI-ZONE MANAGEMENT STRATEGY

As confirmed by Fig. 4, contemporary traffic analysis countermeasures impose uneven energy overhead in the network where the already overloaded lower-level nodes experience more overhead than higher-level nodes, which diminishes the WSN lifespan. Ideally, a countermeasure should be

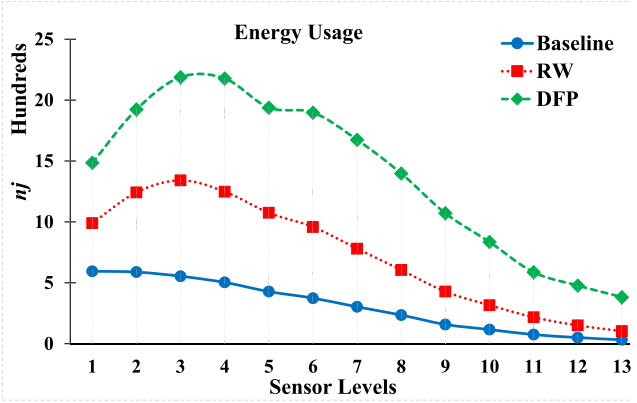


FIGURE 4. Assessing the energy consumption imposed by the traffic analysis countermeasures.

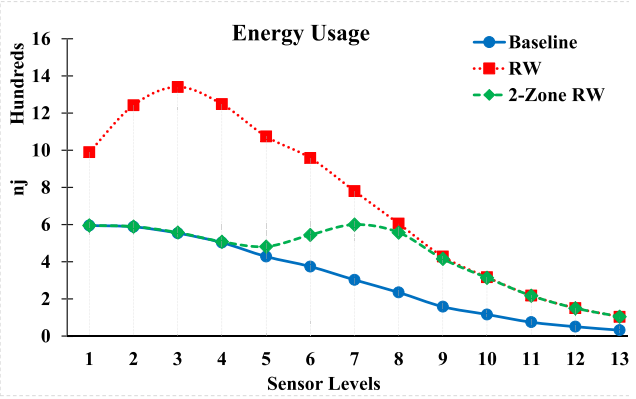


FIGURE 5. Comparing the energy impact of differentiating the application of RW based on two zones to the baseline case.

energy-aware and engage nodes according to their load in the network. In other words, the energy overhead should be disproportionate where farther nodes from the BS are engaged more, and closer nodes experience minimal additional overhead. In this section we present a novel zone-based strategy for shaping the distribution of the additional packet traffic that a countermeasure introduces. We first illustrate the idea through a simple example, where RW is applied to boost the BS anonymity.

A. ILLUSTRATIVE TWO-ZONE EXAMPLE

To illustrate the idea, let us consider an example involving two zones. An inner zone includes closer nodes to the BS and is configured to not participate in the RW activities, i.e., packets are routed using the least-cost paths. On the other hand, the outer zone consists of nodes in higher levels, i.e., farther from the BS, and applies the RW countermeasure. We have used the same simulation parameters and configuration discussed in Section IV to see the impact on energy consumption. The results are shown in Fig. 5 in comparison with the case where RW is applied throughout the network and with the baseline case (no countermeasure is applied). Fig. 5 shows no increase in the energy consumption for nodes in levels 1 to 4. The two-zone division is based on max level in each topology, i.e., a topology with max level of 9, would have levels 1-4 in inner zone, and 5-9 in outer zone. Fig. 5 reports the average of results across 50 topologies with max levels in the range [9], [13]; hence, the diversion on energy consumption in two-zone starts at 5. The two-zone curve in Fig. 5 also show that even though the energy overhead imposed on nodes in the outer zone is in line with RW, it does not exceed beyond the peak value of the baseline (level 1). In other words, the two-zone RW has negligible impact on the WSN lifespan.

RW achieves its anonymity improvement by relying on route alteration which results in a higher number of transmissions. Therefore, one may think that the lower energy consumption by the two-zone implementation of RW comes at the expense of degraded WSN anonymity. In contrast, our experiment reveals that the adversary's *Success Rate* in the two-zone RW is reduced from 48% in RW to 38%; i.e., a 10% boost in anonymity. Since such an improvement might appear counterintuitive, let us analyze the impact of the two-zone design on ET and network anonymity.

ET spatially correlates the intercepted transmissions; a larger transmission count means more instances of link relations that ET will factor in. RW randomizes the routes that packets take to reach the BS, and by doing so, RW strives to inject false link relations (evidence) to the ET analysis. As effective as RW is, not only the existence of those altered routes in the BS vicinity provides spatial evidence to an adversary, but also, they may nudge the analysis in favor of the adversary. Fig. 6 visualizes this side effect clearly. In Fig. 6-a, at t_1 data packet from *Cell*6 is directly delivered to the BS in *Cell*9, enabling the adversary to note $6 \rightarrow 9$ as link-relation evidence, along with links to all other cells reachable by the transmission, shown with green circle in Fig. 6-a; that is one valid evidence out of the 9 being considered.

Fig. 6-b shows the altered route that RW has chosen, namely, $6 \rightarrow 5 \rightarrow 4 \rightarrow 9$. For transmission at times t_1, the adversary collects the evidence $6 \rightarrow 5$, and $6 \rightarrow 9$, i.e., considers the presence of one link to the BS in addition to the actual link. Even though the link $6 \rightarrow 9$ is not along the actual path that RW has chosen, and it should be counted as false evidence, the fact that it points to the BS cell renders it valuable, and therefore results in two, rather than one, valid evidence. The transmissions at t_2 has the same outcome, i.e., yields two valid pieces of evidence, while at t_3 only one valid evidence is collected by the adversary. Adding those up, the altered route provides ET with 5 valid pieces of evidence. Clearly, RW is not achieving what it aims for. Altering the route in the vicinity of the BS negatively impacts the BS anonymity. The two-zone RW, on the other hand, addresses this issue and improves the BS anonymity.

B. MULTI-ZONE DESIGN

The dynamic and evolving nature of event-triggered networks do not permit an offline assessment of the load on the individual nodes and consequently make it infeasible

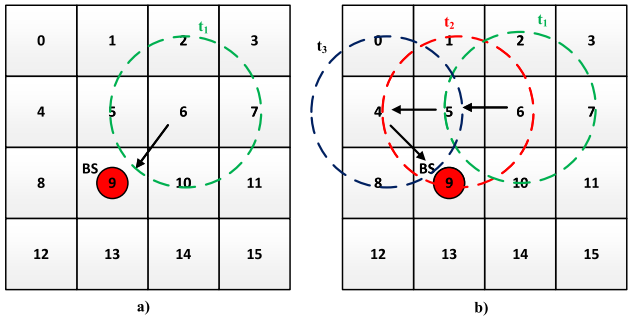


FIGURE 6. Comparing the impact of RW and two-zone RW on the collected evidence, where: (a) in the baseline network data is sent from cell 6 to the BS, and (b) RW randomly chooses a longer data path.

to pre-determine the optimal engagement of the network nodes in the countermeasure activities. Hence, a dynamic and adaptive strategy is more suitable for event triggered WSNs. To achieve that, on one end of the spectrum, each node on each level could be evaluated and tasked individually. Yet, the overhead of such a fine-grained approach can be overwhelming and impractical, especially for large WSNs. On the other hand, not distinguishing among nodes would result in the unfavorable energy overhead pattern that we aim to address. Therefore, we are proposing a middle ground in which a WSN is divided into zones, each is configured independently. A zone defines the intensity of a node involvement in countermeasure duties, and hence controls the energy overhead. As shown above, considering two zones could diminish the overhead, avoid overburdening loaded nodes (sustain the network lifetime) and even increase the BS anonymity.

1) FUNDAMENTAL DESIGN ISSUES

To apply a dynamic zoning strategy (DZS), the following fundamental issues ought to be addressed: (i) what network state is needed, (ii) what criteria the zones are defined based on, (iii) how many zones are to be formed and how big each zone is, and (iv) how the nodes are notified about the zone-related actions. We advocate a DZS that is orchestrated by the BS to eliminate the overhead, and complexity imposed on nodes. First, the BS estimates the energy profile of nodes based on the routes used in data dissemination. The energy consumption rate of a node S_i can be inferred by the BS based on the routes that S_i serves on. The data paths can also be used by the BS to determine the depth of the routing tree and find out the level of each node. The energy profile of nodes can then be aggregated to predict the energy consumption rate at the various levels, i.e., plotting the baseline curve in Fig. 4. Based on the analysis in Section III, the energy consumption rate per level is used as the criterion for zoning. Basically, the level that consumes the most energy sets the bound on the overhead allowed by the traffic analysis countermeasure in other levels in the network. Such a bound will guide the formation of zones, as explained next. Meanwhile, the determination of the number and size of zones is fundamentally a set partitioning problem that is known to be NP-hard. Therefore, DZS pursues a heuristic approach.

2) ZONE SIZE

Since the countermeasure generates redundant traffic or alters the packet routes, managing the overhead at the granularity of a node level would be a viable option only for static networks where data sources and routes do not change, e.g., when data is collected periodically. On the other hand, event-triggered data generation and dissemination causes the load to vary over time and an aggregate zone-based assessment would be more practical. Moreover, the node density in the area is not usually uniform, especially when random node deployment schemes are pursued. Hence, a zone should not include only one level to enable flexibility in balancing the anonymity and energy conservation goals. Our approach factors in the node density at each level in determining the width of a zone. Basically, a zone should include sufficient nodes to facilitate the application of the countermeasure while keeping the average consumed energy in check. The latter naturally depends on the node density per level.

The other factors that affect the zone size is the node degree, which reflects connectivity among nodes. Particularly, a higher node degree will increase the complexity of ET due to the increased link relations. Similarly, a zone with low node degree will have limited options for route alteration and generation of fake traffic. Meanwhile, the number of zones would naturally depend on the number of levels, i.e., the depth of a balanced routing tree, d. Intuitively if each node can reach the BS through few hops, both the zone size and the number of zones should be small. Our zone formation algorithm does not determine the number of zones, but rather group levels into zones that meet the aforementioned criteria.

Let L_i denote the set of nodes in level #i, and $E_{i,i-1}$ denote the set links (edges) between the nodes of two consecutive levels L_i and L_{i-1}. The level that experiences the highest average energy consumption is referred to by L_P, which tends to be level 1. We define η_p as the average energy consumed by a node in L_P. Since the levels next to L_P often experience relatively high energy consumption, as also shown in Fig. 4, we consider the group of levels $\leq p + m$ to form the first zone, Z_P, where m defines number of additional levels beyond L_P to be included in the first zone. We define m based on energy such that $(\eta_p - \eta_i)/\eta_p < \varepsilon$, $\forall p < i < p + m$, where ε is a threshold and should be relatively small, e.g., 0.5, reflecting how η_i is close to η_p. Based on the baseline curve in Fig. 4, a possible setting for the parameter m is 1 or 2. In addition to energy conservation, having $m \geq 1$ also would boost anonymity; as shown in Fig. 6, unnecessary traffic in the BS vicinity can give the adversary valuable evidence/insight.

3) ZONE FORMATION

In summary, Z_P naturally reflects the zone with the highest energy consumption in the network and is spared from injecting redundant traffic, similar to the two-zone example above. This decision is to avoid: (a) the negative effect of the

countermeasure. i.e., evidence provided to the adversary, and (b) the potential energy consumption increases for nodes in L_P. Meanwhile, the goal of DZS is to form other zones such that: (i) the average energy per node within each zone does not exceed η_p, which in essence sustains the WSN lifespan of the baseline case, and (ii) the evidence (link relations) contributed by a level within a zone exceeds that pointing to nodes of levels outside the zone. Our approach focuses on the second condition and addresses the first condition by controlling the volume of redundant transmissions within a zone.

The zone formation steps are as follows. We progressively form zones from the outer nodes (with d being the highest level). Noting that a zone in our approach consists of more than one level, we start by grouping levels d and d-1 to form the outermost zone Z_{out}. Next, we check whether it is beneficial to incorporate L_{d-2} in Z_{out} by comparing $E_{d-2,d-1}$ and $E_{d-2,d-3}$. If $E_{d-2,d-1} \geq E_{d-2,d-3}$, L_{d-2} is added to Z_{out}. Otherwise, a new zone is created. The process is repeated until reaching L_{P+m}. We note that L_{P+m+1} cannot be in a zone by itself, i.e., if L_{P+m+2} is included in an outer zone, by default L_{P+m+1} will be added to it as well. Upon determining the zone boundaries, the BS will broadcast the zone information to the sensor nodes. To avoid exposing the BS through long range transmissions, our approach constrains the BS transmission power to that of a sensor node. Hence, the zone announcement will take the form of multicast where the BS will inform the sensors in L_1, which in turn inform those in L_2 and so on. A pseudo code summary of the DZS approach is provided in Algorithm 1. The overall zone formation process is $O(d)$, which is quite scalable for larger networks.

4) COUNTERMEASURE PARTICIPATION CONTROL

The BS also sets parameters that control the volume of redundant traffic for each zone based on: (i) the energy consumption profile, and (ii) the employed countermeasure. As shown earlier in Fig. 4 and Fig. 5, the baseline energy consumption rapidly decreases as the node level increases; using curve fitting, such an energy consumption profile can be well represented with:

$$\eta_p e^{-x^2} \tag{5}$$

Ideal distribution of redundant traffic should equalize the energy consumed at each level. Thus, we control participation with the same exponential trend but in the reverse order. We define a level participation factor (LPF) in Eq. (6), where i and d reflect the level number and the highest level, respectively.

$$LPF_i = (1 - e^{-(i/d)^2}) \tag{6}$$

We further define the zone participation factor (ZPF) as the average LPF of the levels within a zone. ZPF can be applied to key parameters of the employed countermeasure to control the associated node participation in each zone. For example, in RW the probability of picking a node farther or closer to the BS as the next hop is determined by a ''factor of randomness'', which in turn influences energy consumption. In [12] and [13], such randomness is controlled via a threshold value set by (1- pr). By using ZPF as pr, the node participation grows as the level number increases. As part of the zone formation process, the BS calculates ZPF for each zone and informs the associated sensors.

Algorithm 1 A Pseudo Code Summary of the DZS Algorithm

```
//Dynamically constructs zones.
//When finished, WSN_Zones has list of zones
//Each zone has list of levels.
DZS()
{
    p = HighestEnergy( node_Ls)
    //find m
    i = p;
    while ( Energy( p) - Energy( i) / Energy( p) < e)
      ++ i;
    m = i - p;
    //Form Zones
    foreach L in node_Levels in decending order
        if L == p + m
          End;
        // "L" is last remaining level.
        if L == p + m+1
          current_zone.add( L);
          WSN_Zones.Add( current_zone);
          End;
        //each zone has to be at least 2 levels.
        if current_zone.size() <2
          current_zone.add( L);
          continue;
        if ConnectivityToHigherLevel ( L) >=
                        ConnectivityToLowerLevel ( L)
          currrent_zone.add ( L);
        else //currrent_zone is concluded;
          current_zone.add ( L);
          WSN_Zones.add ( current_zone);
          currrent_zone =new zone ();
          current_zone.add ( L);
}
```

C. DETAILED DZS EXAMPLE

We illustrate the application of DZS using one of the topologies that we have used in our experiments. We demonstrate the dynamic zone formation and how ZPF is used by RW to set its randomness value for each zone. The considered topology has 10 levels, i.e., d =10. Table 5 shows the number of nodes in each level, average energy consumption per node in such a level, and connectivity to previous level ($E_{i,i-1}$) and next level ($E_{i,i+1}$). Table 4 also shows the energy slope of each level relative to L_P. As Table 5 shows, nodes in L_1 have the highest energy consumption, and hence level 1 is set as L_P by DZS. In our experiments, we set $\varepsilon = 0.5$ and thus the value of m is set to 2. That means, the innermost zone, Z_P, stretches

from level 1 up to level $p + m = 1 + 2 = 3$. DZS then forms the other zones starting from the outermost level.

TABLE 5. Node distribution, connectivity, and energy per level for the considered example topology.

Level (L_i)	# node ($\|L_i\|$)	$E_{i,i-1}$	$E_{i,i+1}$	Avg. consumed energy per node, η_i, in nj	$(\eta_p - \eta_i)/\eta_p$
10	14	31	0	4.22	0.93
9	19	36	31	7.67	0.87
8	19	50	36	11.24	0.81
7	21	63	50	12.32	0.79
6	28	65	63	11.95	0.79
5	28	51	65	15.29	0.74
4	29	64	51	18.03	0.69
3	21	37	64	31.41	0.5
2	11	27	37	52.11	0.1
1	10	0	27	58.20	0

Each zone must include at least two levels. Therefore, the outermost zone must include at minimum L_{10} and L_9. In the next step, DZS checks the connectivity of L_8 to L_9, i.e., $E_{8,9}$ versus $E_{8,7}$. Based on Table 5, $E_{8,9}$ =36, and $E_{8,7} = 50$. Since $E_{8,9} < E_{8,7}$, DZS decides that L_8 should not be part of the zone of L_9 and should instead be bundled with L_7 to form a new zone. Next, DZS evaluates the connectivity of L_6 to L_7 and L_5. As $E_{6,7} = 63 < E_{6,5} = 65$, DZS would not include L_6 with L_7 in the same zone, meaning that a new zone is to be formed consisting of L_6 and L_5. Finally, L_4 is the only remaining level, and hence DZS includes it in the newly formed zone alongside L_6 and L_5. In summary, DZS forms 4 zones, {(1, 2, 3), (4, 5, 6), (7,8), (9, 10)}.

After dynamic zone formation, ZPF is calculated for each zone. Table 6 shows how those values are calculated for this example. As discussed in the previous subsection, *ZPF* is a gradient factor that is applied to key parameters of the employed countermeasure. In the case of RW, the value of randomness, $1 - pr$, can be mapped directly to the ZPF values. Other countermeasures might multiply their key parameter by ZPF to incorporate the gradient increase in participation as moving from inner zones to outer zones. Such a method is discussed in the next section when presenting our proposed cross-layer energy-efficient anonymity boosting mechanism.

Using the simulation parameters of Section IV.A, we ran experiments to evaluate the effectiveness of the DZS configuration while RW countermeasure is applied; depicted as DZS+RW. Fig. 7 shows the energy consumption per level for DZS+RW compared to RW, two-zone RW, and the baseline case. The two-zone RW reflects statically splitting the topology into two zones with equal number of levels and enables assessing the variability of a simple zoning scheme. The most notable feature of DZS+RW is that it yields an energy consumption profile that is almost similar across all levels with values that do not exceed the peak for the baseline case. Hence, the energy overhead is imposed disproportionately based on the node's role in data routing and the network lifespan is not negatively impacted by the anonymity countermeasure. DZS+RW is not only energy efficient but also highly effective in boosting the BS anonymity. As seen in Table 7, DZS+RW achieves a 34% reduction in the adversary's success rate. That is 22% more than RW.

TABLE 6. ZPF calculation for the considered example topology.

Zone	Level	LPF	ZPF
1	1	00995	First zone does not participate. $ZPF_1 = 0$
	2	0.0392	
	3	0.0860	
2	4	0.1478	$ZPF_2 = 0.22$
	5	0.2211	
	6	0.3023	
3	7	0.3873	$ZPF_3 = 0.43$
	8	0.4727	
4	9	0.5551	$ZPF_4 = 0.59$
	10	0.6321	

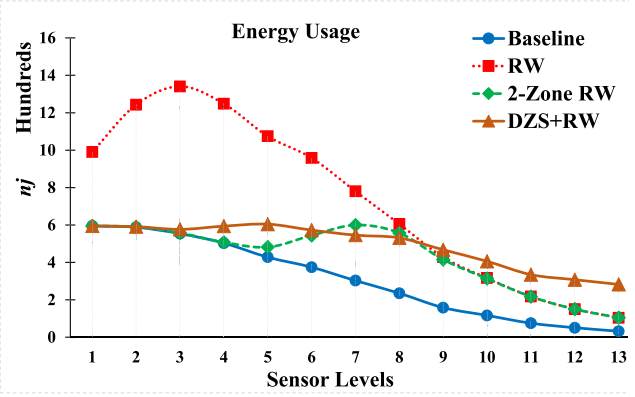


FIGURE 7. Reporting the total energy consumed by nodes in each level when DZS is applied in comparison with the baseline case where no countermeasure is applied and the case where RW is aligned across the network and only on 50% of the levels (two-zones).

VI. CROSS LAYER DESIGN

As discussed in the previous section, the key factors for determining the zone boundaries in DZS are $E_{i,i-1}$ and $E_{i,i+1}$, which reflect the connectivity of a level L_i to the previous and next level, respectively. The goal is to create zones such that more outward ET evidence, away from the BS, is collected than inward evidence, which degrades the effectiveness of the traffic analysis. In this section, we promote a cross-layer scheme to further strengthen DZS, by factoring in link-layer features. Specifically, we incorporate the Transmission Range Increase (TRI) traffic analysis countermeasure [42] to further boost the effectiveness of DZS. Before presenting

such a cross-layer design, we provide a brief overview of TRI.

TABLE 7. The impact of DZS on the BS anonymity in comparison to RW, two-zone and baseline.

Countermeasure	Baseline	RW	2-Zone RW	DZS+RW
Success Rate	60%	48%	38%	26%

A. TRI OVERVIEW

TRI instructs the WSN nodes to use higher transmission power to reach farther than the location of the next hop. TRI strives to influence anonymity in two ways: (i) lowering the adversary's confidence in ET analysis, and (ii) increasing ET's complexity, as explained below.

1) ADVERSARY's CONFIDENCE

By growing the transmission range beyond the next hop, TRI aims to increase the number of potential destinations within the reach of a node. Fig. 8 illustrates the idea. In Fig. 8-a, a transmitter – green dot in the central cell – reaches neighboring cells in each direction, meaning that there are 9 potential destinations (including the source cell). On the other hand, in Fig. 8-b, the transmission range covers two neighboring cells in each direction, which elevates the number of reachable cells to 25. Such an increase in the number of potential destinations directly influences the *Total Evidence* and *Total Belief* in Eq. (2) and Eq.(4), respectively. Furthermore, the *Belief* metric, defined in Eq. (3), depends on the collective evidence of links that are part of a path. In other words, Eq. (1) highlights the importance of a link, while the path is factored in the *Belief* metric. Therefore, by applying TRI some links stand out and some lose significance. A short transmission range localizes the impact on the collected evidence to the vicinity of the transmission source. Meanwhile, increasing the range broadens the scope by introducing evidence to more cells, which bridges the evidence gap among cells and diminishes the adversary's confidence about the inferred data paths. The extreme example would be the case where all nodes have a transmission range that covers the whole network area; to an observer, such a network would have equal evidence for all cells, and hence the traffic analysis would fail.

2) ATTACK COMPLEXITY

TRI has an exponential impact on the complexity of the adversary's traffic analysis. Here, we just highlight such an impact; interested readers can refer to [42] for details.

In a grid of $N \times N$, the adversary would form a graph G consisting of N^2 nodes, each corresponds to a cell. To calculate the value of Eq. (1), the adversary must traverse G multiple times, in each the start point is set to a different cell. Therefore, the complexity of such computation would be $N^2 \times O(b^d)$, in which b is the branching factor and d is the depth of search. Because data is routed to reach the BS, the depth of such search can be capped to N, reflecting the length of the grid diagonal. Therefore, the traverse/complexity can be re-written as $N^2 \times O(b^N)$. The branching factor of b is the key factor on which TRI has a direct impact. For a transmission range that covers up to k cells in each direction, b would be:

$$b = \sum_{i=1}^{k} 2 \times (2i+1) + 2 \times (2\,(i-1)+1) = \sum_{i=1}^{k} 8i$$

Therefore, the complexity of traffic analysis can be represented with:

$$\text{Complexity} = N^2 * O\left(\left(\sum_{i=1}^{k} 8 * i\right)^N\right) \tag{7}$$

In our example shown in Fig. 8, by assuming a 8×8 grid, complexity increases over 6561 times, from over 1 million to over 7 billion, when nodes increase their transmission range by one-fold.

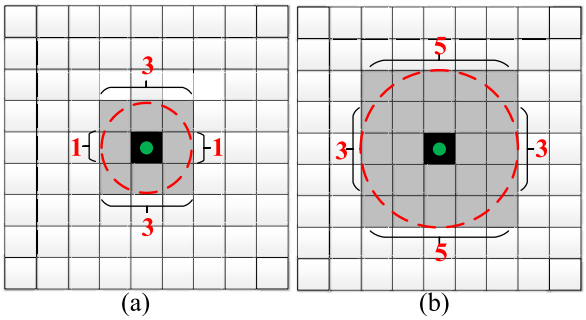


FIGURE 8. Illustrating the transmission range impact on ET. Since the receiver could be in the same cell as the transmitter, there are: (a) 9 potential destinations, and (b) 25 potential destinations. A higher number of destinations implies decreased adversary's confidence.

B. CROSS-LAYER DZS (CL-DZS)

DZS improves anonymity of BS while being conscious of the energy profile of the nodes in the various levels. Zones are formed to diminish the viability of ET as a traffic analysis methodology. As was presented earlier, nodes in a level L_i are included into a zone if it has greater number of edges to higher level ($E_{i,i+1}$) than to lower level ($E_{i,i-1}$); $E_{i,i-1} < E_{i,i+1}$. Since TRI increases the transmission range, it directly influences the number of edges. By doubling the transmission power, a node in L_i could reach beyond nodes in the next levels in each direction, e.g., $E_{i,i-2}$ and $E_{i,i+2}$. We exploit the effect of TRI on the number of edges to further strengthen DZS by boosting directional connectivity of a zone to its outward neighboring zone, and hence introducing evidence for more links that point away from the BS. Another potential benefit of TRI is the increased inter-zone connectivity which creates more inconclusive evidence and grows the complexity of the traffic analysis. Finally, we iterate that TRI exponentially increases the computational complexity of the ET attack, see Eq. (7).

1) EXAMPLE AND NOTATION

To illustrate, let us consider an example network that has 9 levels, for which DZS is assumed to have formed the following zones {(1, 2, 3), (4, 5, 6), (7, 8, 9)}. According to Algorithm 1, L_6 is not part of the third zone alongside L_7 since $E_{6,5} > E_{6,7}$. Doubling the transmission range for nodes in L_6 enables them to reach nodes in L_4 and L_8 and more nodes in L_5 and L_7. Such increase in connectivity for L_6 implies that any transmission with the new range would yield more evidence in both outward and inward directions. The key is that if connectivity of L_6 to L_7 and L_8 becomes considerably greater than its connectivity to L_4 and L_5, i.e., $E_{6,4} + E_{6,5} < E_{6,7} + E_{6,8}$, the ET analysis will be steered away from the BS position. In addition, when a node in L_5 applies TRI and boosts its transmission range, it can increase the number of its neighbors within the second zone, i.e., nodes in L_4 to L_5 and L_6. Such an increase causes links between a node in L_5 to others in both to L_5 and L_6 to outnumber those to nodes in L_4 and in essence diminishes the relevance of the evidence pointing to inner levels.

For simplicity, in the previous example we assumed that applying TRI only results in new edges, represented by $E_{6,4}$ and $E_{6,8}$. But in practice, increasing the transmission range in L_6, could also boost $E_{6,5}$, $E_{6,7}$, and even $E_{6,6}$. Furthermore, depending on TRI's new range, L_6 may even reach beyond L_4, L_8, or not even reach them. To holistically capture the impact of TRI's range on connectivity of L_6, let us refer to all the new edges between L_6 and higher-level nodes as C_6^+, while C_6^- represents all the new edges to lower levels as the result of TRI. C_6° represents new edges between L_6 nodes. With the new notation, TRI could steer away ET analysis if L_6's new range results in

$$C_6^- + E_{6,5} < E_{6,6} + E_{6,7} + C_6^\circ + C_6^+$$

and if L_5's new range leads to having

$$C_5^- + E_{5,4} < E_{5,5} + E_{5,6} + C_5^\circ + C_5^+$$

2) CL-DZS ALGORITHM

Fundamentally, there are two options for integrating TRI with DZS. The first option is to consider the potential range increase during the zone formation process. We deem such an option to be impractical given the major growth in complexity. Basically, all possible transmission ranges for nodes in every level should be factored in and all possible grouping of levels into zones should be considered. Recall that the zone formation problem is NP-hard. Therefore, we adopt the second option where the zone formation in CL-DZS follows the steps in Algorithm 1. In other words, DZS is applied to form zones and then TRI is incorporated to further boost the BS anonymity without reducing the network lifetime. Hence, CL-DZS needs to determine the following: (i) the maximum increase in the transmission range within a zone, which reflects the concern about energy overhead, and (ii) the transmission range for each level to improve anonymity.

To address the first issue, CL-DZS utilizes the ZPF values calculated by DZS. Then, the maximum transmission range of a node, R_{max}, is considered as the highest acceptable range for each zone in proportion to its ZPF. In other words, the outermost zone, Z_{out}, is allowed to increase transmission range up to R_{max}, while in zone Z_j the range is bound to:

$$UBR\,(j) = \frac{ZPF_j}{ZPF_{out}} \times R_{max} \tag{8}$$

That means, the transmission range of the nodes in Z_j can be set to any value between R_{base} and $UBR(j)$. Z_1 does not apply TRI given its peak energy consumption.

Meanwhile, the second issue is handled by tracking the impact on inter- and intra- level connectivity. For each level L_i in Z_j, the transmission range, R_i, is iteratively reduced starting from $UBR(j)$, and C_i^-, C_i^+, and C_i° are counted. We define $G = \left(E_{i,i+1} + E_{i,i} + C_i^+ + C_i^\circ\right) - (E_{i,i-1} + C_i^-)$, as the gain in useful links. We reduce R_i by a *step* and repeat to calculate G, until reaching R_{base}. The range corresponding to the largest G becomes the preferred setting for the nodes of L_i. Obviously, if the increase in transmission range is not useful, the calculated G will have negative values and R_{base} is kept since it corresponds to $G = 0$. It should be mentioned that the *step* controls how fine-grained the designer prefers the G value calculation to be. The *step* can be the smallest unit of distance in a given WSN application.

As there are no nodes/levels beyond Z_{out}, the last levels in the network, L_d, L_{d-1}, would have $C_d^+ = 0$. Similarly, C_{d-1}^+ might suffer if the TRI's range is long enough to reach beyond L_d. Therefore, levels in Z_{out} are exempted and would adopt $UBR(d)$ to signify the impact on C_d°. Even though such an approach could also increase C_d^-, the far proximity of L_d to the BS, makes the imposed complexity from Eq. (7) to outweigh the effect of C_d^-. The pseudo code summary of CL-DZS is provided in Algorithm 2.

3) IMPOSED COMPLEXITY

As was mentioned earlier in Section VI-II, TRI imposes significant computational complexity increase on adversary, Eq. (7). CL-DZS applies TRI with energy consumption in mind and uses ZPF to control the max range in each zone. Therefore, imposed complexity is tied to zone formation and ZPF value of each zone. Using Eq. (8), the reachable cells in each direction per zone can be defined as

$$RC_z = \frac{UBR(z)}{cell\_size} \tag{9}$$

Therefore, combining Eq. (7) and Eq. (9), with assumption of equal number of cells in each zone, the overall complexity imposed by CL-DZS can be defined as

$$\sum_{z=1}^{n} \left(\frac{N^2}{n} * O\left(\left(\sum_{i=1}^{RC_z} 8 * i \right)^N \right) \right) \tag{10}$$

C. DETAILED CL-DZS EXAMPLE

To illustrate the application and effectiveness of CL-DZS, we use the same topology of the earlier DZS example (Section V.C). Since CL-DZS uses the same procedure of DZS to form zones and calculate ZPF, the zone boundaries defined in Table 6 still hold. As shown in Table 2, the base and max transmission ranges for this example are 120m and 240m, respectively. Therefore, using $R_{max} = 240$, and ZPF values shown in Table 6, CL-DZS calculates UBR for each zone as shown in Table 8. The UBR value of each zone can vary between R_{base} and $UBR(j)$; because $UBR(2) < R_{base}$, it is set back to R_{base}.

Algorithm 2 Pseudo Code Summary of the CL-DZS Steps

```
CL DZS()
{
DZS() //Forms Zones; Algorithm 1
ZPF() //Sets ZPF of each zone
UBR() //Sets Upper Bound range per zone; Eq. 8
foreach L in node_Levels descending order
    if L in zone_1
    //leave range_1 = base_range
    End; //No need to process rest of levels
    if L in zone_out
     set range[ L] =UBR(L) //to upper bound range
     continue;
    max_G =0
    max_range =base_range
    for (r = UBR(L) to base_range; r=r-step)
        //connectivity and added edges to same levels
        E_o = ConnectivityToSameLevel(L)
        C_o = NumberOfNewEdgesTo(L)
        //connectivity and added edges to higher levels
        E_plus =ConnectivityToHigherLevel (L)
        C_plus =NumberOfNewEdgesToHigher (L)
        //connectivity and added edges to lower levels
        E_minus = ConnectivityToLowerLevel(L)
        C_minus = NumberOfNewEdgesToLower( L)
        //Gain
       G = (E_plus E_o +C_plus +C_o) -
(E_minus+C_minus)
        if G >max_G
        max_range =r
        range[L] =max_range
}
```

TABLE 8. Range calculation based on UBR for each zone in example 2.

Zone	ZPF_i	UBR_i	$range_i$
1	0.0	0	Zone 1 does not participate; → 120
2	0.22	89	$89 < R_{base}$ → 120
3	0.43	175	175
4	0.59	240	240

Table 8 reveals that Z_1 and Z_2 are not participating in CL-DZS and only Z_3 and Z_4 would increase the transmission range. All levels in Z_4 fully participate in CL-DZS and increase their transmission range to 240. On the other hand, for the levels in Z_3, namely L_7 and L_8, CL-DZS must decide on an effective range that would result in the highest G, by following Algorithm 2. Table 9 and Table 10 show the results for L_8 and L_7, respectively. For sake of simplicity and a good balance of fine-grained G calculation without imposing unnecessary overhead, *Step* is set to 20. *G=50* corresponds to *range=175* for L_8, while *G=78* is the highest gain with *range=155* for L_7.

TABLE 9. 175 results in best G for L_8 in example 2.

Rang	$E_{8,9}$	$E_{8,8}$	$E_{8,7}$	C_8^+	C_8°	C_8^-	G
175	36	54	50	60	28	78	50
155	36	54	50	26	20	49	37
135	36	54	50	20	16	30	46
$115 < R_{base}$			Not considered.				

TABLE 10. 155 results in best G for L_7 in example 2.

Rang	$E_{7,8}$	$E_{7,7}$	$E_{7,6}$	C_7^+	C_7°	C_7^-	G
175	50	63	70	67	32	72	70
155	50	63	70	42	22	29	78
135	50	63	70	33	9	29	56
$115 < R_{base}$			Not considered.				

We ran experiments to evaluate the effectiveness of CL-DZS configuration using the simulation parameters of Section IV.A. In Table 11 the adversary's success rate is shown for CL-DZS in comparison to DZS+RW, RW, and baseline. We note 18% reduction compared to DZS+RW, which is a significant achievement and confirms the effectiveness of the cross-layer design. In Fig. 9, the energy consumption per level for CL-DZS is shown in comparison to baseline, RW, and DZS+RW. The trend of CL-DZS curve in Fig. 9 is considerably lower than DZS+RW, which implies that not only CL-DZS is achieving higher anonymity but also consuming less energy compared to DZS+RW.

TABLE 11. The impact of CL-DZS on the BS anonymity in comparison to baseline, RW, and DZS+RW.

Countermeasure	Baseline	RW	DZS+RW	CL-DZS
Success Rate	60%	48%	26%	8%

VII. SIMULATION RESULTS

Validation of DZS and CL-DZS is done through extensive simulations. We have used the same setup discussed earlier in Section IV-A. The validation is geared for reporting the average performance and comparing DZS and CL-DZS to competing countermeasures in the literature.

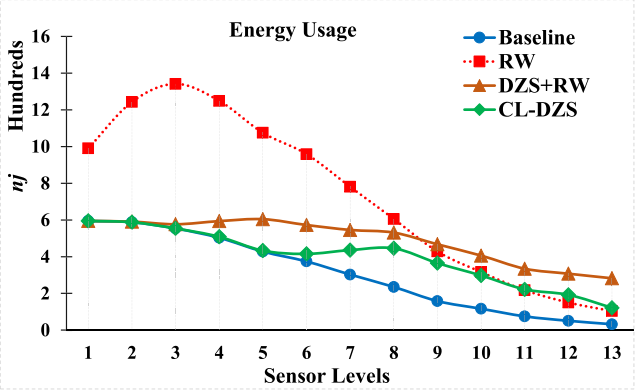


FIGURE 9. Reporting the total energy consumed by nodes in each level when CL-DZS is applied in comparison with baseline, RW, and DZS+RW.

A. AVERAGE PERFORMANCE

The viability of DZS and CL-DZS have been demonstrated through an example in Sections V and VI. Here we study the average performance over numerous topologies. We note that both schemes allow the incorporation of contemporary traffic analysis countermeasures, such as route alteration and redundant traffic generation. For consistency with previous sections, we use RW to demonstrate the benefits of our design in sustaining (and increasing) the effectiveness of such a countermeasure while avoiding the negative energy implication that it inflicts on the network nodes. We report the results where CL-DZS is used alone, and each DZS and CL-DZS is applied in combination with RW.

RW categorizes neighbors of a node S into "*Parents*" and "*Not-Parents*". Neighbors at a level less than that of S (i.e., closer to the BS) are labeled as *Parents* while the rest are labeled as *Not-Parents*. The variable *pr* is used to determine the probability that a data path is altered. If RW decides to alter a route, it would randomly pick the next hop from *Not-Parents*; otherwise, it pursues the preferred-next-hop according to shortest-path routing tree. Integrating RW with DZS is discussed in Section V and is referred to as "DZS+RW". The same process applies to incorporate RW with CL-DZS (CL-DZS+RW). In CL-DZS+RW, we set RW to use a higher transmission range only when forwarding packets to *NotParents*.

Fig. 10 compares the energy consumption of the baseline case to that of DZS+RW and CL-DZS with and without RW. For all DZS variants, the increase in energy consumption is capped to the maximum in the baseline case, i.e., all are below 600*nj*. The difference between the CL-DZS+RW and DZS+RW curves reflects the impact of TRI, where the leap in energy consumption grows with the increase in the level count, i.e., as we move away from the BS. We also like to point out the impact of RW on CL-DZS. When CL-DZS is applied alone, the transmission range increases only from level 6 and up, while incorporating RW pumps the energy depletion rate from level 4 and up. Overall Fig. 10 confirms that our energy-aware design prevents energy consumption at all levels from exceeding the baseline's peak and hence does not reduce the WSN lifespan.

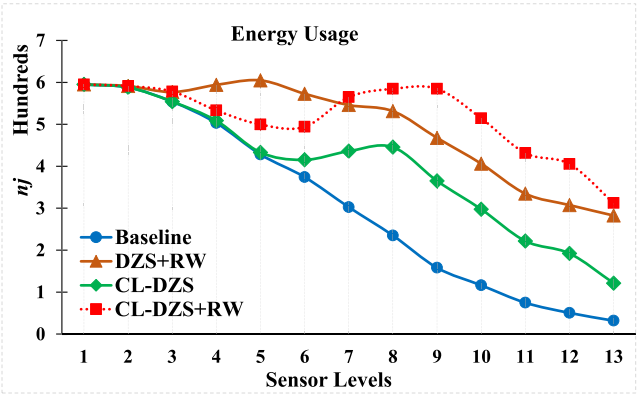


FIGURE 10. Energy consumption in each level for CL-DZS+RW in comparison to baseline, DZS+RW, and CL-DZS.

The BS anonymity is measured using the *Success Rate* metric and is reported in Table 12. Generally, CL-DZS achieves remarkable reduction in the attack effectiveness; such performance even gets improved when combined with RW. The 4% gain (reduction in success rate) contributed by RW comes at the cost of a 14% increase in transmission rate over CL-DZS. It should be noted that in our design of CL-DZS+RW we set ZPF to half the value used in DZS+RW, hence the number of transmissions is decreased in CL-DZS+RW. As CL-DZS is only utilizing TRI, no altered route is applied and therefore its transmission rate is equal to baseline. Considering the results of Fig. 10 and Table 12 collectively confirms the great advantages of CL-DZS.

TABLE 12. Impact of RW, DZS+RW, CL-DZS, and CL-DZS+RW on the adversary attack success rate and the transmission rate in the network.

Countermeasure	Success Rate	Percentage of Transmission Rate Relative to the Baseline
Baseline	60%	
RW	48%	234%
DZS+RW	26%	148%
CL-DZS	8%	100%
CL-DZS+RW	4%	114%

B. COMPARISON WITH COMPETING APPROACHES

We note that the most successful countermeasures in the literature are the ones that heavily rely on generating bogus packets. Here we are comparing the performance with four of these countermeasures, even though CL-DZS+RW does not generate any bogus packets and only relies on route alteration. The considered countermeasures are:

- *Differential Fractal Propagation (DFP)*: DFP combines RW with fake packet generation using probabilistic measures driven from the node's forwarding rates [13].
  - *Deceptive Packets, Multiple Destinations —— Single Packet (DP-MS):* Nodes generate deceptive

packets and route them to pre-define dummy sinks. The approach aims to flatten the *Belief* metric in order to boost the BS anonymity [22].
  - *Assisted Deception (AD)*: Nodes coordinate locally to generate deceptive packets to prevent temporal correlation of consecutive transmissions [44].
- *Anti-Traffic Analysis (ATA)*: It aims to achieve a transmission rate uniformity among all nodes by generating fake packets [23].

Table 13 shows the relative transmission rate increase of countermeasures relative to the baseline. ATA's brute force approach grows the transmission rate by 3379 times. DFP, DP-MS, and AD are more conscious of their impact on nodes, yet they impose significantly high overhead, 348%, 449%, and 710%, respectively. On the other hand, CL-DZS+RW is the absolute winner with only 14% increase over the baseline. Table 13 also shows the *Success Rate* metric when each of the countermeasures is applied. ATA unsurprisingly results in 0% *Success Rate*, given the massive fake packet transmissions that are nearly 30 and 5 times more than CL-DZS+RW and AD, respectively. Meanwhile AD and CL-DZS+RW both result in 4% *Success Rate*; yet AD achieves that using over 7 times the number of transmissions in the network.

TABLE 13. Impact of DFP, DP-MS, ATA, AD, and CL-DZS+RW on anonymity and transmission rate.

Countermeasure	Success Rate	Percentage of Transmission Rate Relative to the Baseline
Baseline	60%	
DFP	20%	348%
DP-MS	12%	449%
ATA	0%	3379%
AD	4%	710%
CL-DZS+RW	4%	114%

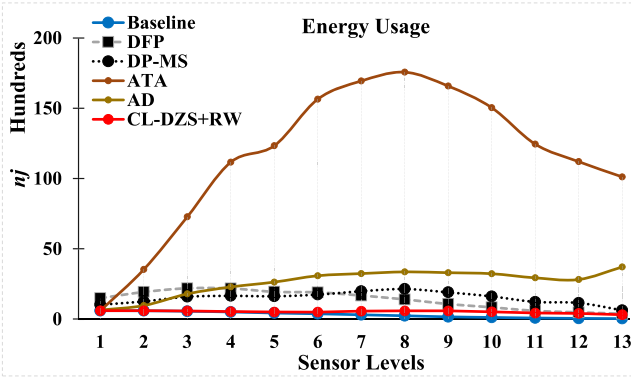


FIGURE 11. Energy consumption in each level.

Energy consumption in each level is shown in Fig. 11. The impact of ATA's 3379% increase in transmission rate (Table 13) on energy is clearly visible in Fig. 11, where the overhead is too high that ATA's curve dominates the chart. Comparing the peak value at L_8 with the baseline's peak (L_1), ATA is shortening the network lifespan by 28 times. AD comes second in terms of the significance of energy overhead, where its peak value at L_{13} is over 6 times of that of the baseline. For better visibility, Fig. 12 shows the results while excluding ATA and AD. Both DFP and MP-MS increase energy consumption across all levels. DFP's and DP-MS's peak values are at L_3, and at L_8 respectively. Both experience 3.6 times growth in maximum energy consumption compared to the baseline peak, implying 3.6 times shorter WSN lifespan. Unlike all other countermeasures, CL-DZS+RW keeps its energy consumption less than the peak of baseline and has zero impact on the time for the first node to die.

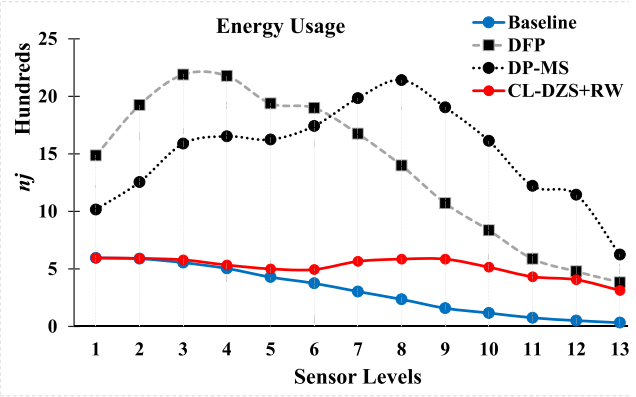


FIGURE 12. Reporting the energy consumption in each level; this is a version of Fig. 11 that does not show the curves of ATA and AD.

VIII. CONCLUSION

The vital role of the base-station in a WSN has made it a preferable target of attacks. Even if packet encryption is pursued, traffic analysis can be conducted using intercepted transmissions to infer the presence of communication links among nodes and uncover the network topology. For that, Evidence Theory proved to be an effective attack model. The high and uneven energy overhead that contemporary anti-traffic analysis techniques impose on the individual sensor nodes tend to exhaust the limited energy resources of sensor nodes faster than normal, and hence shorten the practical lifespan of the network. In this paper, after pointing out the problem, a novel dynamic multi-zone design has been proposed to determine how a countermeasure can be applied to address energy concerns while achieving the BS anonymity goal. Such a multi-zone design has been further extended through the incorporation of cross-layer feature that increase effectiveness in countering the traffic analysis threat. We have demonstrated the effectiveness of our approach and how it can be integrated with other countermeasures. The simulation results have shown that our approach has zero impact on the WSN lifespan while significantly improves the BS anonymity. As a future extension, we plan to utilize machine learning techniques to predict energy consumption patterns and engage the countermeasures adaptively to impose minimum energy overhead while maximizing anonymity.

APPENDIX

See Tables 14 and 15.

TABLE 14. Acronym Reference.

AD	Assisted Deception
ATA	Anti-Traffic Analysis
BS	Base-Station
CL-DZS	Cross-Layer Dynamic Zoning Strategy
DFP	Differential Fractal Propagation
DP	Deceptive Packets
DP-MS	DP - Multiple Destinations Single Packet
DZS	Dynamic Zoning Strategy
ET	Evidence Theory
LPF	Level Participation Factor
RW	Random Walk
TRI	transmission range increase
UBR	Upper Bound Range
WSN	Wireless Sensor Network
ZFP	Zone Participation Factor

TABLE 15. Main Variables.

C_i^-	Number of new edges between node in level i and levels smaller than i
C_i^+	Number of new edges between node in level i and levels greater than i
C_i°	Number of new edges between node in level i
d	Maximum level value in network
$E_{i,j}$	Number of edges between nodes in level i and j
i	Level of a node
R_{base}	Based Transmission Range in network
R_{max}	Maximum transmission range in network

REFERENCES

[1] K. Chopra, K. Gupta, and A. Lambora, "Future internet: The Internet of Things—A literature review," in *Proc. Int. Conf. Mach. Learn., Big Data, Cloud Parallel Comput. (COMITCon)*, Feb. 2019, pp. 135–139.

[2] M. Abdelhafidh, M. Fourati, L. C. Fourati, and A. Chouaya, "Wireless sensor network monitoring system: Architecture, applications and future directions," *Int. J. Commun. Netw. Distrib. Syst.*, vol. 23, no. 4, pp. 413–451, Jan. 2019.

[3] S. Yinbiao and K. Lee, "Internet of Things: Wireless sensor," Int. Electrotech. Commission, Geneva, Switzerland, White Paper 78, Sep. 2014.

[4] J. Jiang, G. Han, H. Wang, and M. Guizani, "A survey on location privacy protection in wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 125, pp. 93–114, Jan. 2019.

[5] Y. A. Yahya, S. Raed, A. M. H. Darghaoth, and S. A. Majeed, "Secure routing protocol for wireless sensor networks: Survey," in *Proc. 8th Int. Eng. Conf. Sustain. Technol. Develop. (IEC)*, Feb. 2022, pp. 155–160.

[6] J. Wang, F. Wang, Z. Cao, F. Lin, and J. Wu, "Sink location privacy protection under direction attack in wireless sensor networks," *Wireless Netw.*, vol. 23, no. 2, pp. 579–591, Feb. 2017.

[7] J. R. Jiang, J. P. Sheu, C. Tu, and J. W. Wu, "An anonymous path routing (APR) protocol for wireless sensor networks," *J. Inf. Sci. Eng.*, vol. 27, no. 2, pp. 657–680, 2011.

[8] J. Kong, X. Hong, and M. Gerla, "An identity-free and on-demand routing scheme against anonymity threats in mobile ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 8, pp. 888–902, Aug. 2007.

[9] N. Baroutis and M. Younis, "Location privacy in wireless sensor networks," in *Mission-Oriented Sensor Networks and Systems: Art and Science: Advances*, vol. 1, H. M. Ammari, Ed. Cham, Switzerland: Springer, 2019, pp. 669–714.

[10] Y. Qin, D. Huang, and B. Li, "STARS: A statistical traffic pattern discovery system for MANETs," *IEEE Trans. Depend. Sec. Comput.*, vol. 11, no. 2, pp. 181–192, Mar. 2014.

[11] P. Venkitasubramaniam, T. He, L. Tong, and S. Wicker, "Toward an analytical approach to anonymous wireless networking," *IEEE Commun. Mag.*, vol. 46, no. 2, pp. 140–146, Feb. 2008.

[12] J. Deng, R. Han, and S. Mishra, "Decorrelating wireless sensor network traffic to inhibit traffic analysis attacks," *Pervasive Mobile Comput.*, vol. 2, no. 2, pp. 159–186, Apr. 2006.

[13] J. Deng, R. Han, and S. Mishra, "Countermeasures against traffic analysis attacks in wireless sensor networks," in *Proc. 1st Int. Conf. Secur. Privacy Emerg. Areas Commun. Netw. (SECURECOMM)*, Sep. 2005, pp. 113–126.

[14] M. Boulaiche and M. Younis, "Increasing base-station anonymity through illusive void formation," *Int. J. Commun. Netw. Distrib. Syst.*, vol. 25, no. 4, pp. 433–460, 2020.

[15] W. Conner, T. Abdelzaher, and K. Nahrstedt, "Using data aggregation to prevent traffic analysis in wireless sensor networks," in *Distributed Computing in Sensor Systems* (Lecture Notes in Computer Science), vol. 4026. Berlin, Germany: Springer, 2006, pp. 202–217.

[16] K. Mehta, D. Liu, and M. Wright, "Location privacy in sensor networks against a global eavesdropper," in *Proc. IEEE Int. Conf. Netw. Protocols*, Oct. 2007, pp. 314–323.

[17] R. El-Badry and M. Younis, "Providing location anonymity in a multi-base station wireless sensor network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 157–161.

[18] I. T. Almalkawi, J. Raed, N. Alghaeb, and M. G. Zapata, "An efficient location privacy scheme for wireless multimedia sensor networks," in *Proc. IEEE Int. Conf. Emerg. Technol. Factory Automation ETFA*, vol. 2019-Septe, pp. 1615–1618, 2019.

[19] Y. Ebrahimi and M. Younis, "Novel assessment metric and countermeasures for traffic attack threats in wireless sensor networks," in *Proc. 37th Annu. IEEE Conf. Local Comput. Netw.*, Oct. 2012, pp. 340–343.

[20] B. Chakraborty, S. Verma, and K. P. Singh, "Temporal differential privacy in wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 155, Apr. 2020, Art. no. 102548.

[21] J. Chen, Z. Lin, Y. Liu, Y. Hu, and X. Du, "Sink location protection protocols based on packet sending rate adjustment," *Int. J. Distrib. Sensor Netw.*, vol. 12, no. 1, Jan. 2016, Art. no. 6354514.

[22] Y. Ebrahimi and M. Younis, "Using deceptive packets to increase base-station anonymity in wireless sensor network," in *Proc. 7th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jul. 2011, pp. 842–847.

[23] B. D. Ying, D. Makrakis, and H. T. Mouftah, "Anti-traffic analysis attack for location privacy in WSNs," *EURASIP J. Wireless Commun. Netw.*, vol. 2014, no. 1, pp. 1–15, Dec. 2014.

[24] Z. W. Hussien, D. S. Qawasmeh, and M. Shurman, "MSCLP: Multi-sinks cluster-based location privacy protection scheme in WSNs for IoT," in *Proc. 32nd Int. Conf. Microelectron. (ICM)*, Dec. 2020, pp. 1–4.

[25] Y. Ebrahimi and M. Younis, "Averting in-situ adversaries in wireless sensor network using deceptive traffic," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2011, pp. 1–5.

[26] U. Acharya and M. Younis, "Increasing base-station anonymity in wireless sensor networks," *Ad Hoc Netw.*, vol. 8, pp. 791–809, Nov. 2010.

[27] S. Alsemairi and M. Younis, "Cross-layer technique for boosting base-station anonymity in wireless sensor networks," *Int. J. Commun. Syst.*, vol. 30, no. 13, p. e3280, Sep. 2017.

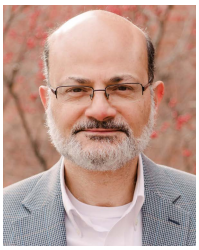
[28] N. Baroutis and M. Younis, "Load-conscious maximization of base-station location privacy in wireless sensor networks," *Comput. Netw.*, vol. 124, pp. 126–139, Sep. 2017.

[29] S. Alsemairi and M. Younis, "Forming a cluster-mesh topology to boost base-station anonymity in wireless sensor networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2016, pp. 1–6.

[30] N. Baroutis and M. Younis, "Boosting base-station anonymity in wireless sensor networks through illusive multiple-sink traffic," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–7.

[31] N. Baroutis and M. Younis, "Using fake sinks and deceptive relays to boost base-station anonymity in wireless sensor network," in *Proc. IEEE 40th Conf. Local Comput. Netw. (LCN)*, Oct. 2015, pp. 109–116.
[32] V. Kumar, A. Kumar, and M. Singh, "Boosting anonymity in wireless sensor networks," in *Proc. 4th Int. Conf. Signal Process., Comput. Control (ISPCC)*, Sep. 2017, pp. 344–348.
[33] K. Bicakci, I. E. Bagci, and B. Tavli, "Lifetime bounds of wireless sensor networks preserving perfect sink unobservability," *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 205–207, Feb. 2011.
[34] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Netw.*, vol. 38, no. 4, pp. 393–422, 2002.
[35] N. A. Pantazis, S. A. Nikolidakis, and D. D. Vergados, "Energy-efficient routing protocols in wireless sensor networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 551–591, 2nd Quart., 2013.
[36] Y. Chang, X. Yuan, B. Li, D. Niyato, and N. Al-Dhahir, "A joint unsupervised learning and genetic algorithm approach for topology control in energy-efficient ultra-dense wireless sensor networks," *IEEE Commun. Let.*, vol. 22, no. 11, pp. 2370–2373, Nov. 2018.
[37] Y. Chang, H. Tang, Y. Cheng, Q. Zhao, and B. Yuan, "Dynamic hierarchical energy-efficient method based on combinatorial optimization for wireless sensor networks," *Sensors*, vol. 17, no. 7, p. 1665, Jul. 2017.
[38] Z. Fei, B. Li, S. Yang, C. Xing, H. Chen, and L. Hanzo, "A survey of multi-objective optimization in wireless sensor networks: Metrics, algorithms, and open problems," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 550–586, 1st Quart., 2017.
[39] M. Pandey and S. Verma, "Privacy provisioning in wireless sensor networks," *Wireless Pers. Commun.*, vol. 75, no. 2, pp. 1115–1140, Mar. 2014.
[40] N. Li, N. Zhang, S. K. Das, and B. Thuraisingham, "Privacy preservation in wireless sensor networks: A state-of-the-art survey," *Ad Hoc Netw.*, vol. 7, no. 8, pp. 1501–1514, Nov. 2009.
[41] J. R. Ward and M. Younis, "Cross-layer traffic analysis countermeasures against adaptive attackers of wireless sensor networks," *Wireless Netw.*, vol. 25, no. 5, pp. 2869–2887, Jul. 2019.
[42] Y. Ebrahimi and M. Younis, "Increasing transmission power for higher base-station anonymity in wireless sensor network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2011, pp. 1–5.
[43] T. Yan, Y. Bi, L. Sun, and H. Zhu, "Probability based dynamic load-balancing tree algorithm for wireless sensor networks," in *Networking and Mobile Computing* (Lecture Notes in Computer Science), vol. 3619, X. Lu and W. Zhao, Eds. Berlin, Germany: Springer, 2005, pp. 682–691.
[44] Y. Ebrahimi and M. Younis, "Traffic analysis through spatial and temporal correlation: Threat and countermeasure," *IEEE Access*, vol. 9, pp. 54126–54151, 2021.
[45] S. Seys and B. Preneel, "ARM: Anonymous routing protocol for mobile ad hoc networks," *Int. J. Wireless Mobile Comput.*, vol. 3, no. 3, pp. 145–155, Oct. 2009.
[46] L. Zhou, Y. Shan, and X. Chen, "An anonymous routing scheme for preserving location privacy in wireless sensor networks," in *Proc. IEEE 3rd Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC)*, Mar. 2019, pp. 262–265.
[47] M. Saikia, U. K. Das, and M. A. Hussain, "Secure energy aware multi-path routing with key management in wireless sensor network," in *Proc. 4th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Feb. 2017, pp. 310–315.
[48] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler, "SPINS: Security protocols for sensor networks," *Wireless Netw.*, vol. 8, no. 5, pp. 521–534, Sep. 2002.
[49] D. Huang, "On measuring anonymity for wireless mobile ad-hoc networks," in *Proc. 31st IEEE Conf. Local Comput. Netw.*, vol. 1, Nov. 2006, pp. 779–786.
[50] A. Serjantov and G. Danezis, "Towards an information theoretic metric for anonymity," in *Privacy Enhancing Technologies* (Lecture Notes in Computer Science), vol. 2482, R. Dingledine and P. Syverson, Eds. Berlin, Germany: Springer, 2003, pp. 41–53.
[51] A. Liu, X. Liu, Z. Tang, L. T. Yang, and Z. Shao, "Preserving smart sink-location privacy with delay guaranteed routing scheme for WSNs," *ACM Trans. Embedded Comput. Syst.*, vol. 16, no. 3, pp. 1–25, Aug. 2017.
[52] V. Kumar and A. Kumar, "A novel approach for boosting base station anonymity in a WSN," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 9, pp. 114–120, 2017.
[53] Z. Ren and M. Younis, "Effect of mobility and count of base-stations on the anonymity of wireless sensor networks," in *Proc. 7th Int. Wireless Commun. Mobile Comput. Conf.*, Jul. 2011, pp. 436–441.
[54] O. Kulyk, S. Neumann, J. Budurushi, and M. Volkamer, "Nothing comes for free: How much usability can you sacrifice for security?" *IEEE Secur. Privacy*, vol. 15, no. 3, pp. 24–29, Jun. 2017.
[55] R. Ahmad, R. Wazirali, and T. Abu-Ain, "Machine learning for wireless sensor networks security: An overview of challenges and issues," *Sensors*, vol. 22, no. 13, p. 4730, Jun. 2022.
[56] G. Bhatti, Y. Javed, M. Naveed, and S. Asif, "Out-door localization in large-scale wireless sensor networks by using virtual nodes," in *Proc. Adv. Sci. Eng. Technol. Int. Conf. (ASET)*, Feb. 2020, pp. 1–7.
[57] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.
[58] G. Mao, B. Fidan, and B. D. O. Anderson, "Wireless sensor network localization techniques," *Comput. Netw.*, vol. 51, no. 10, pp. 2529–2553, 2007.
[59] P. Rong and M. L. Sichitiu, "Angle of arrival localization for wireless sensor networks," in *Proc. 3rd Annu. IEEE Commun. Soc. Sensor Ad Hoc Commun. Netw.*, Sep. 2006, pp. 374–382.
[60] W. R. Heinzelman, A. Sinha, A. Wang, and A. P. Chandrakasan, "Energy-scalable algorithms and protocols for wireless microsensor networks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 6, Jun. 2000, pp. 3722–3725.

YOUSEF EBRAHIMI received the B.S. degree in computer engineering from the Isfahan University of Technology, Iran, and the M.S. degree in computer engineering from the Sharif University of Technology, Iran. He is currently pursuing the Ph.D. degree in computer engineering with the University of Maryland, Baltimore County. His research interests include wireless sensor networks, network architectures and protocols, communication security, and embedded systems.



MOHAMED YOUNIS (Fellow, IEEE) is currently a Professor with the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County (UMBC). Before joining UMBC, he was with Honeywell International Inc., where he led multiple projects for building integrated fault tolerant avionics and dependable computing infrastructure. He also participated in the development of the redundancy management system, which is a key component of the vehicle and mission computer for NASA's X-33 space launch vehicle. He has published about 350 technical papers in refereed conferences and journals. He has seven granted and three pending patents. His research interests include network architectures and protocols, wireless sensor networks, fault tolerant computing, secure communication, and cyber-physical systems. He is a fellow of the IEEE Communications Society. He serves/served on the editorial board of multiple journals and the organizing and technical program committees of numerous conferences.