Abstract—Physical attacks, which aim to render one or more sensor nodes non-operational by physically capturing and destroying them, are among the most serious security threats in Wireless Sensor Networks (WSNs). In case of dense deployment (a desirable property in the design of WSNs) multiple sensor nodes acquire redundant (highly correlated) data. As a result, even if some nodes are dead, the remaining nodes can successfully complete the sensing task. However, as the number of nodes in the network decreases, the remaining nodes are burdened with the extra load (energy dissipation). In this study, we investigate the energy cost of survivability in the presence of physical attacks through a novel Linear Programming framework. We explore the energy dissipation characteristics of the network for different physical attack scenarios.

Index Terms—wireless sensor networks, linear programming, energy efficiency, network lifetime, physical attacks

I. INTRODUCTION

Self organization is one of the key concepts in wireless sensor networks (WSNs) which enables deployment to areas where human intervention is too risky/costly. Self organization property together with infrastructure-less characteristics of WSNs make them favorable solutions for many surveillance and control applications [1].

Communication and computation are the two main energy dissipation categories for a typical WSN and communication energy usually dominates the energy expenditure. Therefore, reducing the amount of data conveyed to the base station is critical for improving energy efficiency. Since sensor nodes are generally deployed densely, highly correlated measurements performed by a multitude of sensor nodes lead to data redundancy. By exploiting the redundancy, it is possible to decrease the amount of data relayed towards the base station. One such redundancy elimination approach is the aggregation of the data coming from a group of sensor nodes. Another related technique is the avoidance of redundant data collection. By assigning sensing tasks to nodes intelligently, no redundant data is generated, thus, energy dissipation for communication is minimized. These energy-efficient strategies not only prolong the network lifetime but can also help to satisfy survivability requirements in case of a security attack.

Physical attacks are a class of security attacks on WSNs in which sensor nodes are physically destroyed and rendered permanently non-operational [2], [3]. Physical attacks are easy to conduct, yet, their effects can be catastrophic (i.e., these attacks can destroy vital nodes in the network and lead to inefficient energy dissipation trends in the whole network). Self organization is an effective way of countering against these attacks assuming that the number of nodes adversaries can destroy is not limitless. If the network is designed for survivability then, in case of a security attack remaining operational nodes can reorganize by taking over the functions (e.g., relaying function) of the attacked nodes. Redundancy is another key factor for survivability; even if one or more nodes are dead, remaining nodes that share the same observation window (i.e., if two nodes sense highly correlated data then their observation windows are overlapping) can perform the sensing functionality. However, such additional functions increase the energy dissipation of the remaining nodes.

Linear Programming (LP) is a technique to solve the problem of maximizing or minimizing a linear function whose variables are required to satisfy certain constraints that are expressed either as linear equalities or linear inequalities. There are many optimization problems that can be solved using LP such as network flow problems. In the context of the WSNs, LP approach was previously used to analyze various important subjects (e.g., [4], [5]).

In this study, we consider a WSN that is tasked with monitoring an area for a predetermined period of time. The area is divided into regions and sensor nodes are deployed uniformly throughout the network (i.e., initially there are equal number of nodes in each region). Through this representative application, we investigate the energy cost of combating against physical attacks by using a novel LP framework.

The rest of the paper is organized as follows. Related work is reviewed in Section II. The system model and LP framework is described in Section III. The results of our analysis are presented in Section IV. Conclusions are drawn in Section V.

II. RELATED WORK

In this section, we present a review of the studies on physical attacks in WSNs. We use the term physical attack to refer to the attacks aiming at physically destroying the sensor nodes. We note that in other studies (e.g., [6]) the same term is also used for the concept of tampering with nodes (i.e., capturing a sensor node and gaining direct access to its cryptographic material). By capturing a node, attackers can bypass cryptographic protection and conduct effective denial-of-service attacks [7]. In our study, we are not interested in these more advanced attack techniques and their countermeasures.

Physical attacks can be classified into two groups: blind physical attacks and search-based physical attacks [8]. In [2], blind physical attacks are studied. These attacks are performed
after detecting the deployment area without considering where each sensor node is located. Sensor nodes may be destroyed blindly using a brute force approach (e.g., by bombing the area). In such a situation, the research problem is to determine the minimum number of sensor nodes together with their location information to achieve the desired lifetime. In [3], search-based physical attacks, in which sensor nodes are targeted and destroyed individually, are investigated. To defend against these attacks, a protocol is proposed based on the assumption that sensor nodes are able to detect attackers. In order to evaluate the performance of the proposed protocol, a metric called Accumulative Coverage is defined considering that primary success criteria of an attacker is the amount of coverage reduction in the network.

LP is the subclass of mathematical programming models which uses linear functions both for objective and constraints. There are a wide variety of problems where LP is used for modeling (e.g., transportation, production scheduling, and network flow problems) [9]. Analyzing WSNs through LP based models is an approach employed in many previous studies [4], [5], [10], [11]. But, up to our best knowledge, our work is the first study that investigates the problem of physical attacks in WSNs within an LP framework.

III. System Model

In this section, we first define the research problem informally. Then, the formalization is carried out with LP modeling.

A. Motivation and Problem Definition

During their operational lifetime, sensor nodes can take at least two different roles: sensing role and relaying role. In the sensing role, a sensor node gathers data from the environment and either transmits the data directly to the base station or relays it to another sensor node acting as a relay. It is possible that some of the data can be transmitted directly to the base station and some of the data is conveyed via relay nodes. In the relay role, a sensor node forwards the data it received from either the source node or from another relay node to the base station or yet to another relay node. Data can be relayed successively until final destination (i.e., the base station) is reached. Using wireless communication, sensor nodes form a network in self-organized fashion. This means nodes could collaborate and exchange the roles when needed.

To accomplish the task of monitoring a region for a predetermined amount of time, the network must possess certain resources and capabilities. Energy is usually the most critical resource. A sensor node can satisfy its assigned tasks as long as it has sufficient energy resources. Sensor nodes deployed in hostile areas are vulnerable to physical attacks. If an attacker destroys a sensor node, other nodes have to bear the extra burden to fulfill the responsibility of the destroyed node. If the possibility of such a physical attack has not been considered, in case of an attack even though the remaining nodes are capable of taking over the roles of the destroyed nodes, their energy limitations can prevent them to accomplish the task. In other words, the self organizing nature of the network enable the network to operate with the remaining sensor nodes but these sensor nodes cannot continue to monitor the operation area as planned since their batteries do not last long enough.

We consider a WSN in which sensor nodes are deployed to monitor an area (a battlefield for instance) which is composed of a certain number of non-overlapping regions. The network must be able to collect data from the area for a predetermined period of time. Before the deployment, each sensor node should be charged with the energy that is sufficient to accomplish the task. Initial energy of the nodes would not be enough to function long enough if the possibility of physical attacks was not considered because node failures may lead to inefficient energy dissipation trends. Informally speaking, the research problem is finding the minimum amount of additional energy to collect information collaboratively for a period of time from a designated area if certain number nodes fail to operate due to a physical attack.

In our threat model, we assume the base station is physically well-protected. We further assume that the attacker is not capable of destroying all of the nodes in a given region (i.e., sensor nodes are camouflaged, hence, it is not possible to locate all of them in a reasonable amount of time). We consider two attack models: (i) uniform attack and (ii) non-uniform attack. In uniform attack, we consider a more powerful adversary which picks a certain number of nodes in each region and renders the targeted nodes inoperable. In non-uniform attack, nodes only in a single region are destroyed by the attacker.

B. Linear Programming Framework

In our system model, energy consumption of sensor nodes is dominated by communication energy dissipation rather than sensing and processing energy dissipation. We adopt the energy model from [5]. In this model, the amount of energy to transmit one bit of data is \( P_{tx,i,j} = \rho + \varepsilon l_{ij} \) and to receive one bit of data is \( P_{rx} = \rho \), where \( \rho \) represents the energy dissipated in the electronic circuitry, \( \varepsilon \) denotes the transmitters efficiency, \( \alpha \) represents the path loss exponent and \( l_{ij} \) is the distance between node-\( i \) and node-\( j \).

In our framework, we assume that there are \( N \) sensor nodes and a single base station in the network. Data generated at each node, transferred either directly (single-hop) or through other sensors acting as relays (multi-hop), terminates at the base station. The network topology is represented as a directed graph \( G = (V,A) \). \( V \) is the set of all nodes, including the base station as node-1. We also define set \( W \), which includes all the nodes except node-1. \( A = \{ (i,j) : i \in W, j \in V - i \} \) is the set of arcs (links). The amount of data sent on the directed link \( (i,j) \) is denoted as \( f_{ij} \).

We consider a uniform random deployment scenario in which sensor nodes are deployed over a rectangular area that includes \( Z \) number of regions to be monitored (\( Z \) denotes the set of regions and the members of set \( Z \) are denoted by \( Z_k \)). The set of all sensor nodes located within region-\( Z_k \) are denoted with \( W_k \). Number of days that sensor node-\( i \) monitors region-\( Z_k \) is denoted as \( d_{i,k} \) and the total time region-\( Z_k \) monitored by all nodes located in it is denoted by \( D_k \). In

\[1\] For example, communication energy dissipation constitutes 91% of the total energy dissipation in Telos sensor nodes [12].
each region, $s_i$ unit of raw data per day is to be conveyed to the base station. We also define a set $F_k$, which consists of failed (destroyed) nodes in region-$Z_k$ and the number of elements in set $F_k$ is $M_k$ (the union of all $F_i$’s constitute the set $F$).

The optimization problem is formulated as an LP problem. Figure 1 presents the basis of our formulation, which is similar to the models in earlier work [5] (we modify this basic model to suit for our needs and expand it with additional constraints). Since the objective is to minimize battery, the problem is the minimization of the maximum battery requirement of the nodes in the network by finding the $f_{ij}$’s (flows) that satisfy the constraints.

\begin{equation}
\text{Minimize battery}
\end{equation}

Subject to:

\begin{align}
& f_{ij} \geq 0 \forall (i, j) \in A & (1) \\
& f_{ij} = 0 \text{ if } i = j \forall (i, j) \in A & (2) \\
& \sum_{j \in V} f_{ij} - \sum_{j \in W} f_{ji} = d_i s_i = 0 \forall i \in W & (3) \\
& P_{rx} \sum_{j \in W} f_{ji} + \sum_{j \in V} P_{tx,ij} f_{ij} - e_i \leq 0 \forall i \in W & (4) \\
& e_i = \text{battery} \forall i \in W & (5)
\end{align}

Fig. 1: LP model as the basis for investigating the energy cost of mitigating physical attacks in WSNs.

Equation 1 states that all flows are non-negative. Equation 2 is used to eliminate infinite loops - there cannot be a flow from the base station to other nodes or from a node to itself. Equation 3 states that the difference between the data flowing out of node-$i$ and the data flowing into node-$i$ is the data generated at node-$i$. Equation 4 states that for all nodes except the base station the energy consumed for transmission and receipt of data is equal to or less than the energy stored in batteries. Equation 5 is used to assign equal energy to each sensor node. As noted earlier, the model presented in Figure 1 and explained so far is the basic model for flow balancing (i.e., all data generated at the sensor nodes eventually terminate at the base station) and energy minimization (i.e., to minimize the maximum energy dissipation of nodes, all sensor nodes are forced to dissipate their energies in a balanced fashion).

Failed nodes cannot participate in data gathering or relaying, thus, any flows originating or flowing through such a node should be set to zero. Equation 6 incorporates this restriction into our model:

\begin{equation}
f_{ij} = 0 \text{ if } (i \in F \text{ or } j \in F) \forall (i, j) \in A & (6)
\end{equation}

To model the optimal case of monitoring the area in which sensor nodes in the same region operate in a coordinated fashion and redundancy is totally eliminated, we introduce the following constraint:

\begin{equation}
\sum_{i \in W_k} d_i = D_k \forall Z_k \in Z & (7)
\end{equation}

Equation 7 formulates the case of optimal cooperation between sensor nodes in each region for monitoring the region (i.e., no redundancy exists in the transmitted data). Total network area is divided into $N_Z$ non-overlapping regions and region-$Z_k$ must be monitored $D_k$ days by the sensor nodes that are located in region-$Z_k$. Note that in each region-$Z_k$ the set of nodes located in that region ($W_k$) are essentially observing the same phenomena (i.e., they would acquire redundant data if they transmitted data simultaneously). The model consisting of constraints presented in Equations 1, 2, 3, 4, 5, 6, and 7 is called Optimal Role Assignment (ORA) Model. In ORA model only a single node performs the sensing operation at a time in each region (i.e., nodes take turns for sensing the data).

Equation 8 formulates the operation of WSN without any cooperation in data acquisition (i.e., the redundant data acquisition case). Each sensor node in the same region-$Z_k$ collects the same amount of data and all redundant data is conveyed to the base station.

\begin{equation}
d_i = D_k \forall i \in W_k, Z_k \in Z & (8)
\end{equation}

The set of constraints presented in Equations 1, 2, 3, 4, 5, 6, and 8 are used to model the network operation mode where nodes do not perform sensing collaboratively (i.e., each sensor node performs sensing and sends the data to the base station without exploiting the data redundancy). We call this model Redundant Data Sensing (RDS) model. Note that in RDS model energy minimization is achieved by optimizing only $f_{ij}$’s (i.e., $d_i$’s are fixed), however, in ORA model both flows ($f_{ij}$’s) and monitoring times ($d_i$) are jointly optimized. Note that Equation 7 and Equation 8 are not used simultaneously. All system variables with their acronyms and descriptions are presented in Table I.

\section*{IV. Analysis}

In our analysis, we solve the LP problems introduced in previous section for a simple exemplifying topology with a square area of size 200 m x 200 m. As shown in Figure 2, the area consists of 32 regions (i.e., $N_Z = 32$). The task of the sensor nodes in each region is to monitor it for 50 days (i.e., $D_k = 50$ days). Note that we opt to adopt a uniform monitoring time for all regions. In each day, same amount of data ($s_i = M_b$) is collected for each region. Each region has the same number of randomly deployed sensor nodes. We use 192 nodes ($N = 192$) in total and the base station is at the center (i.e., there are six sensor nodes in each region; $N_k = 6$). All nodes perform the relaying operation collaboratively and there are no restrictions on their transmission ranges (e.g., any node-$i1$ in any region-$Z_k$1 can send data to any other node-$i2$ in any other region-$Z_k$2). We use GAMS [13] to solve the LP models. All data points are the averages of the results of 100 random topologies. The parameters used in the analysis are presented in Table II.

To evaluate the benefits of eliminating data redundancy we first perform an analysis without any node failures ($M_k = 0$, $\forall Z_k \in Z$) by using ORA and RDS models. The required battery energy for monitoring the network is found to be 4.37 J
TABLE I: Terminology for LP Formulations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Number of nodes</td>
</tr>
<tr>
<td>Z</td>
<td>Set of the regions in the network</td>
</tr>
<tr>
<td>f_{ij}</td>
<td>Flow from node-i to node-j</td>
</tr>
<tr>
<td>s_i</td>
<td>Data generated at node-i in one day</td>
</tr>
<tr>
<td>d_i</td>
<td>Number of days that node-i monitors a region</td>
</tr>
<tr>
<td>P_{rx}</td>
<td>Energy consumption for receiving one bit of data</td>
</tr>
<tr>
<td>P_{tx,ij}</td>
<td>Energy consumption for transmitting one bit of data from node-i to node-j</td>
</tr>
<tr>
<td>l_{ij}</td>
<td>Distance between node-i and node-j</td>
</tr>
<tr>
<td>ρ</td>
<td>Energy dissipated in the electronic circuitry</td>
</tr>
<tr>
<td>ε</td>
<td>Transmitters efficiency</td>
</tr>
<tr>
<td>α</td>
<td>Path loss exponent</td>
</tr>
<tr>
<td>e_i</td>
<td>Energy requirement for sensor node-i</td>
</tr>
<tr>
<td>G</td>
<td>Directed graph that represents network topology</td>
</tr>
<tr>
<td>V</td>
<td>Set of nodes, including the base station as node-1</td>
</tr>
<tr>
<td>W</td>
<td>Set of nodes, except the base station (node-1)</td>
</tr>
<tr>
<td>A</td>
<td>Set of edges (links)</td>
</tr>
<tr>
<td>D_k</td>
<td>Total number of days that region-Z_k must be monitored</td>
</tr>
<tr>
<td>Z_k</td>
<td>A member of set Z</td>
</tr>
<tr>
<td>N_Z_k</td>
<td>Total number of regions</td>
</tr>
<tr>
<td>W_k</td>
<td>Set of sensor nodes in region-Z_k</td>
</tr>
<tr>
<td>N_k</td>
<td>Total number of sensor nodes in region-Z_k</td>
</tr>
<tr>
<td>F</td>
<td>Set of failed nodes in the network</td>
</tr>
<tr>
<td>M</td>
<td>Total number of failed nodes in the network</td>
</tr>
<tr>
<td>F_k</td>
<td>Set of failed nodes in region-Z_k</td>
</tr>
<tr>
<td>M_k</td>
<td>Total number of failed nodes in region-Z_k</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>20</td>
<td>19</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>28</td>
<td>27</td>
<td>26</td>
<td>25</td>
</tr>
<tr>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
</tr>
</tbody>
</table>

Fig. 2: Region map in our analysis

TABLE II: Parameters used in the analysis

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network area</td>
<td>200 m X 200 m</td>
</tr>
<tr>
<td>ρ</td>
<td>50 nJ/bit</td>
</tr>
<tr>
<td>ε</td>
<td>100 pJ/bit/m²</td>
</tr>
<tr>
<td>α</td>
<td>2</td>
</tr>
<tr>
<td>N</td>
<td>192</td>
</tr>
<tr>
<td>N_Z_k</td>
<td>32</td>
</tr>
<tr>
<td>D_k</td>
<td>50 day</td>
</tr>
<tr>
<td>s_i</td>
<td>1 Mb/day</td>
</tr>
</tbody>
</table>

and 28.58 J for ORA and RDS models, respectively. In RDS model, amount of data collected is six times more than ORA model but corresponding energy requirement increases slightly more than six times. In the ORA model, sensor nodes monitor the area by taking turns, however, they always participate in data relaying. The solution of the LP model with the chosen parameter set provides the optimum amount of time assigned to each sensor node to monitor their regions. For example, sensor nodes in region-1 has the following monitoring times; node-1: 2.52 days, node-2: 2.48 days, node-3: 15.09 days, node-4: 2.83 days, node-5: 23.78 days, and node-6: 3.30 days. Depending on their locations, some nodes use most of their energies for relaying data, thus, they take less role in sensing and generating data and dissipate less energy for it. In the remaining analysis, we concentrate on the ORA model and do not consider RDS model.

In Figure 3, relative energy overhead (i.e., percentage energy increase when compared to the case where none of the nodes fail – $M_k = 0 \forall Z_k \in Z$) is presented as a function of number of remaining nodes in the network for the case of uniform attack. In uniform attacks, equal number of nodes are dead in each region (e.g., for 128 remaining nodes in the network a total of 64 nodes, two from each region, are dead). Percentage energy overhead grows from 20% (when only one sensor node dies in each region – 160 sensor nodes remain in the network) to 574% (when five sensor nodes die in each region – 32 sensor nodes remain in the network).
In Figure 4, relative energy overhead for each region is plotted for different numbers of nodes failed due to a nonuniform attack. In this case, indicated number of nodes are dead only in one region – none of the nodes in other regions are dead. After nodes are failed due to a physical attack, the network reorganizes itself and remaining sensor nodes update their sensing and relaying patterns. We assume that at most five sensor nodes fail to operate in each region after the attack and one node is enough to cover and monitor the region. In Figure 4, regions are ordered in groups of 4 according to their distance from the base station.

Figure 4 reveals an interesting energy dissipation trend, as described as follows. Sensor nodes close to the base station directly send most of their data to the base station but sensor nodes far away from the base station need to transmit via a multi-hop route (i.e., directly sending collected data to the base station is not energy efficient). As a result, sensor nodes close to the base station become more heavily used than the others. For this reason, attacks against regions close to the base station cause higher energy overhead per sensor node when compared to the attacks performed to other regions of the network. There is an exception to this mechanism when there are five failed nodes in each region (i.e., there is only one surviving node in each region; $M_k = 5$). In this situation, the minimum energy requirement of the network is the highest when attacks are directed against the farthest regions from the base station. The reason for such behavior is that network cannot share the burden of monitoring a region when there is only one node left alive in the region (the nodes in other regions can cooperate only for relaying the data out of the region and the remaining operational node in the region has to get the data out of the region on its own). The burden of getting the data out of the farthest regions is especially heavy due to the extended distances to the base station.

V. CONCLUSION

In this study, through a novel LP framework we investigate the energy overhead arising due to physical attacks in WSNs. We model optimal network behavior to balance the energy dissipation throughout the network so that incapacitation of any node (due to a physical attack) does not result in an overwhelming energy cost for the whole network. We consider two attack scenarios: (i) uniform attack, where all network regions are equally affected by the attack (i.e., the same number of nodes are dead in each region) and (ii) non-uniform attack, where only the nodes in a single region are affected. Our results show that the energy cost of uniform attacks can efficiently be shared by the network (e.g., if one sixth of the nodes are incapacitated by an attack the energy overhead is one fifth of the energy required to accomplish the task without any node failure). However, for non-uniform attacks the energy overhead depends on the targeted region. Physical attacks targeting the regions closer to the base station lead to more energy overhead when compared to the attacks targeting other regions provided that there are multiple nodes left operational in the targeted region after the attack. On the other hand, if there is only one node left in the targeted region then the attacks on the furthest regions from the base station result in the highest energy overhead due to the lack of available redundancy, which limits the network wide cooperation to only the relaying operation for the attacked region.

Note that in this study we characterize the energy overhead due to node failures induced by physical attacks. However, similar energy overhead characteristics are observed if the nodes fail with the same spatial pattern due to natural factors. Nevertheless, nature cannot target the nodes which lead to the most damage in terms of the energy overhead for the remaining nodes (unlike the case of intelligent attackers).

In our models, nodes are either active or failed for the entire network operation period (i.e., we do not investigate the case where nodes operate some time and then fail). Since it is not possible to know the exact timing of an attack, a reasonable strategy is to allocate enough resources to mitigate the effects of the worst-case scenario in which the attack is conducted just after network starts operating and damages nodes permanently.

The LP framework we present in our study can easily be tailored to accommodate other aspects of physical attacks in WSNs. For instance, a natural extension of our analysis would be to examine energy dissipation characteristics when the base station is not located at the center of the monitored region.

REFERENCES