

Research Article

Fault-Tolerant Relay Node Placement in Wireless Sensor Networks for Surveillance of Overhead Transmission Lines

Kaixuan Wang  and **Jian Yang** 

Shanxi University of Finance and Economics, Taiyuan, China

Correspondence should be addressed to Kaixuan Wang; wangkx@sxufe.edu.cn

Received 29 September 2021; Revised 24 December 2021; Accepted 31 January 2022; Published 20 March 2022

Academic Editor: Mukhtaj khan

Copyright © 2022 Kaixuan Wang and Jian Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Wireless sensor networks provide a promising technology to surveil overhead transmission lines in smart grid communication. However, a great challenge for researchers is proposed by adverse outdoor environments and essential requirements of strong as well as flexible smart grid communication. Specifically, nearly linear topology, limited energy, and tolerance for faults are entangled to make the surveillance system complex. Because of the need for deep understanding of system, significant efforts have been made in the past few years. We have proposed a fault-tolerance framework for surveillance of overhead transmission lines. In this paper, we follow the framework and further explore deployment problems systematically. Firstly, we present a fault-tolerance placement model. After analyzing the model, we identify the optimal placement for fault tolerance and then propose a placement algorithm with optimal deployment. An extensive experiment highlights some useful observations for design of the fault-tolerance system and demonstrates efficient fault tolerance for relay nodes. Numerical results show that the fault-tolerance ratio is improved by at least 6% compared to that of previous algorithms.

1. Introduction

Smart grid, integrated with widespread green energy (e.g., solar, wind, and thermal energy), represents a revolution for traditional power utilities, because it can exploit advanced information and communication technologies to facilitate the infrastructure to improve the efficiency, sustainability, and flexibility of power grids [1]. Smart grid communication, responsible for two-way, cybersecurity communication, and cutting-edge information technologies, has substantially enhanced the intelligence of current energy systems and has attracted huge attention from many researchers [2]. Human inspection, currently the main maintenance of transmission lines in a distributed power network, suffers from time-consuming, exhausting, and inefficient inspection and is further likely to subject workers involved in power line repairs to serious hazards [3]. Wireless sensor networks (WSNs) have been widely considered in the efficient communication on monitoring application because of their improving performance, relatively low cost, and flexible

nature [4]. Consequently, they are instinctively applied in smart grid communications [5]. For labor-replacing purposes, the authors have initiated a scheme applying WSNs to monitor overhead transmission lines in the power distribution [6]. In this scheme, surveillance involves collecting sensing data from a few sensors scattered around a pole, transmitting data to a relay node installed in the pole, and relaying data from different poles to a control center, which is usually situated in a town or a city.

The scheme focuses on relay nodes and ignores ordinary sensors around poles since these sensors are relatively simple in terms of function, but the relay nodes are responsible for a large amount of work, such as collecting data, transmitting data, and communication with a control center [4]. In WSNs [7], some prerequisites (e.g., time delay, energy consumption) are possible to be problems [8]. But how to deal with these problems in the open air? As the poles supporting power transmission lines are generally extend through tens of kilometers in rural areas, relay nodes installed in these poles have a linear topology. The linear topology represents a

challenge for a network, because it is vulnerable in light of topology tolerance [9], and makes the monitoring application more complex.

The smart grid is a self-healing, strong, and flexible system, in which fault tolerance is indispensable. Most of the existing works, however, emphasize the external characteristics of networks, such as reliability, survivability, and vulnerability [10, 11]. These characteristics are constrained to statistical properties over a long time period and cannot respond to the network in real time. Many works neglect the significance of internal faults so that communication networks are difficult to be understood from the fault perspective. Additionally, N-1 contingency is of critical importance in a power system, because it can provide high reliability and make a system functional in the event of one random component failure [12]. Therefore, this N-1 contingency is an essential principle as well as design requirements in the practical power network.

Green energy has aroused wide concern of researchers since it can dramatically reduce their carbon footprint and hopefully facilitate the smart grid [13]. For the scheme, although the overhead transmission lines carry electricity, it is too powerful to supply for relay nodes directly [14]. As faults are inevitable [7], we aim to provide a fault-tolerance system for surveillance of transmission lines by WSNs considering green energy envisioned for the smart grid [15].

In our previous work [16], a fault-tolerance framework has been proposed to address the fault tolerance of relay nodes in WSNs for surveillance. In the framework, three sections were implemented to conduct fault tolerance in different stages: planning, deployment, and operation. In this study, we follow our fault-tolerance framework, explore the deployment of relay nodes, and concentrate on the backup placement of relay nodes to determine the optimal placement for fault tolerance. Specifically, in view of linear topology as well as N-1 contingency, we emphasize the communication of relay nodes for the monitoring application by WSNs. From the fault tolerance perspective, we propose a fault-tolerance placement model (FaTPM) and clarify the deployment question to determine the optimal fault-tolerance placement. Based on the FaTPM, we present an initial backup deployment problem (IBDP) by energy harvesting out of doors. Then, we analyze the tolerance requirements for compliance with N-1 contingency. Finally, we propose a dynamic optimal deployment placement (DODP) algorithm to solve the fault-tolerance placement problem.

2. Contributions

The main contributions of this paper are as follows:

- (1) On the basis of our fault-tolerance framework, we propose a fault-tolerance placement model for relay nodes to supervise overhead transmission lines.
- (2) Based on the fault-tolerance placement model, we highlight the placement of relay nodes and propose an initial backup deployment problem in consideration of linear topology.

- (3) Complying with the significant N-1 contingency in a power system, we analyze the requirements of fault tolerance and develop a dynamic optimal deployment placement algorithm for fault tolerance.
- (4) We perform experiments to verify the model and algorithm. Numerical results show that the tolerance ratio improves at least 6% compared to that of previous algorithms.

To the best of our knowledge, this study is the first to address fault tolerance for relay nodes by considering N-1 requirements to build a strong, flexible communication system.

3. Organization and Notations

The remainder of this research is organized as follows: Section 2 reviews related works on surveillance by WSNs and fault tolerance. Section 3 introduces the FaTPM and proposes the IBDP. Section 4 presents N-1 implications, analyzes tolerance requirements, and gives the DODP algorithm. The numerical results are provided in Section 5. Finally, Section 6 provides concluding remarks and directions for future research.

The frequently used notations are summarized in Table 1. We also use bold to represent vectors.

4. Related Works

The surveillance of network had been studied in the literature since it was first presented in [6]. A hierarchical network was designed to meet the requirements of smart grid applications by combining wired and wireless technologies [4]. In order to reduce the time delay, P.W.T. Pong et al. introduced a novel configurable network using cellular network on account of the linear network topology [9]. The authors in [17] investigated the transmission delay, selected a “representative node” scientifically, and presented a grouping mechanism to minimize transmission delay for efficient communications. In view of linear topology, these two studies used the cellular network to directly transmit data [9, 17]. In addition, for the gateway of the optical fiber composite overhead ground wire alongside the transmission line, Yi Qian et al. studied the power allocation under the assumption that sensors are powered by green energy (e.g., solar energy). Then, they proposed centralized and distributed power schemes [18]. For the energy consumption of relay nodes, the authors utilized a magnetoelectric composite to supply sensor nodes and built a self-powered sensor network by scavenging energy from an alternating current power line [19]. Despite the remarkable research works, a network without fault tolerance could hardly work well for practical surveillance systems.

A few previous works also studied fault tolerance. The authors in [20] surveyed technologies for monitoring power lines and proposed future research directions. Broadly speaking, the methods of fault tolerance in communication networks could be categorized into two classes: topology reconfiguration and backup. The topology reconfiguration

TABLE 1: Frequently used notations.

Notation	Description	Notation	Description
RN	Relay node of wireless sensor network	CM	Cellular module for transmission
SD_i	Sensor data in pole i	En	Energy consumption for (.)
r	Compare ratio for data	$Dist[a, b]$	Distance between a and b
td	Time delay for (.)	Gp_i	Set of relay nodes in i group
tr	Transmission rate for (.)	D_t	Time delay
R_j	Coverage range for cellular module j	fre_i	Frequency of occurred fault
ME_i	Magnetolectric transducer in pole i	SP_i	Solar panel in pole i
$me(t)$	Magnetolectric harvesting rate	$ch(t)$	Solar charging rate
Ex	Expense of equipment (.)	T	Observation time
f_i	Fault in pole i	Thr	Threshold for delay
TTR_i	Time to repair in pole i	TBF_i	Time between failures in pole i
X	$X_i \in \{0, 1\}$ Vector for RN backup	Y	$Y_j \in \{0, 1\}$ Vector for CM backup

was extensively studied in general WSNs by estimating cost of reconfiguration [21]. It was an NP-hard problem to determine the minimum numbers of relay nodes to achieve fault tolerance in a fully connected network [22]. To solve the problem, the authors developed heuristic algorithms, such as moth flame optimization algorithm and bat algorithm. Because of the linear network architecture, the fault tolerance by reconfiguring topology was not fit for the surveillance scheme.

Some studies used the backup method to obtain the fault tolerance. A reverse transmission approach was presented by [23]. This approach aimed to ensure intact data and to transmit information in the reverse direction when a fault occurred. In [24], the authors employed several power lines in parallel and backed up data in a neighboring power line to tolerate faults. To enhance the robustness [10], the authors provided multiple node-disjoint paths between a communication node and a control center to achieve fault tolerance.

Although previous studies have produced important results, the fault-tolerance surveillance system is still to be fully understood. The fault-tolerance framework has been proposed [16]. In the framework, we use backup method to obtain the fault tolerance considering the linear topology. The numbers of backup, placement of backup, and adaptive switching algorithm are included in this framework. Although the placement of backup nodes is already referred to, in this study, we follow this framework and discuss the fault-tolerance placement of relay nodes comprehensively.

4.1. Network Model. In this section, we take a systematic look at the network model, introduce the FaTPM, and elaborate factors. Then, we propose the initial backup deployment problem.

In the power distribution network, the monitoring system by WSNs can be modeled as shown in Figure 1. Between two adjacent power substations (i.e., Sta_1 and Sta_2), the relay nodes (RNs) are installed in each pole to relay data to a control center (CC). The CC is responsible for aggregation of data and operation of algorithm. Data are transmitted in a hop-by-hop style. Owing to the constraint of time delay in linear topology, a cellular-enabled module (CM) has been introduced [10]; then a group is formed by CM, in which several relay nodes share a common CM.

The short-distance wireless communication technology (e.g., Zigbee) is used for relay nodes as well as sensors scattered around the pole. The long-distance wireless communication (e.g., 4G/GSM) is adapted to the CM. Generally, substations are connected with CC by optical fiber. Thus, we model the monitoring system as an undirected graph $G = (V, E)$ due to two-way communication. V represents vertices, which stand for n poles, two adjacent power substations (sometimes we use RN_0 and RN_{n+1} refer to Sta_1 and Sta_2), and the control center (CC). That is, $V = \{RN_1 \dots RN_n \cup Sta_1 \cup Sta_2 \cup CC\}$. E represents edges, which include short-distance wireless communication, long-distance wireless communication, and optical connections.

For the FaTPM, we reinforce this model with fault tolerance. There are n relay nodes between substations with linear topology, $RN_1 \dots RN_n$. These relay nodes have real positions, and there is different spacing between poles according to geographical locations. The CMs (i.e., $CM_1 \dots CM_m$) are responsible for directly transmitting data. Thus, relay nodes are grouped by CMs to $Gp_1 \dots Gp_m$. In each group Gp_k , $k \in [1..m]$, several relay nodes are included, for instance, $RN_{k1} \dots RN_{kq} \in Gp_k$. As the fault is unavoidable, backup for the RN and CM is indispensable to achieve fault tolerance since a network with linear topology can hardly tolerate any failures via reconfiguration. In addition, backup relay nodes should have the same functions with the original ones from the fault-tolerance view. In contrast to [9], we determine that the last relay node is a sink in each group for fault-tolerance purposes. Therefore, $G = (V, E)$ is the fault-tolerance placement model, as shown in Figure 1.

4.2. Data Aggregation. In the FaTPM, data are attached in a hop-by-hop manner in each group. All data are finally transmitted to CC by CM at the last relay node in the same group. For instance, there are four relay nodes in a group as demonstrated in Figure 2. Data from RN_1 to RN_4 are transmitted to CC in the hop-by-hop manner. The last relay node, RN_4 , is a sink that collects all data in this group and triggers the cellular network by CM.

For every relay node, there are three necessary and successive actions: sensing, compressing, and transmitting data. A relay node collects data from around sensors by

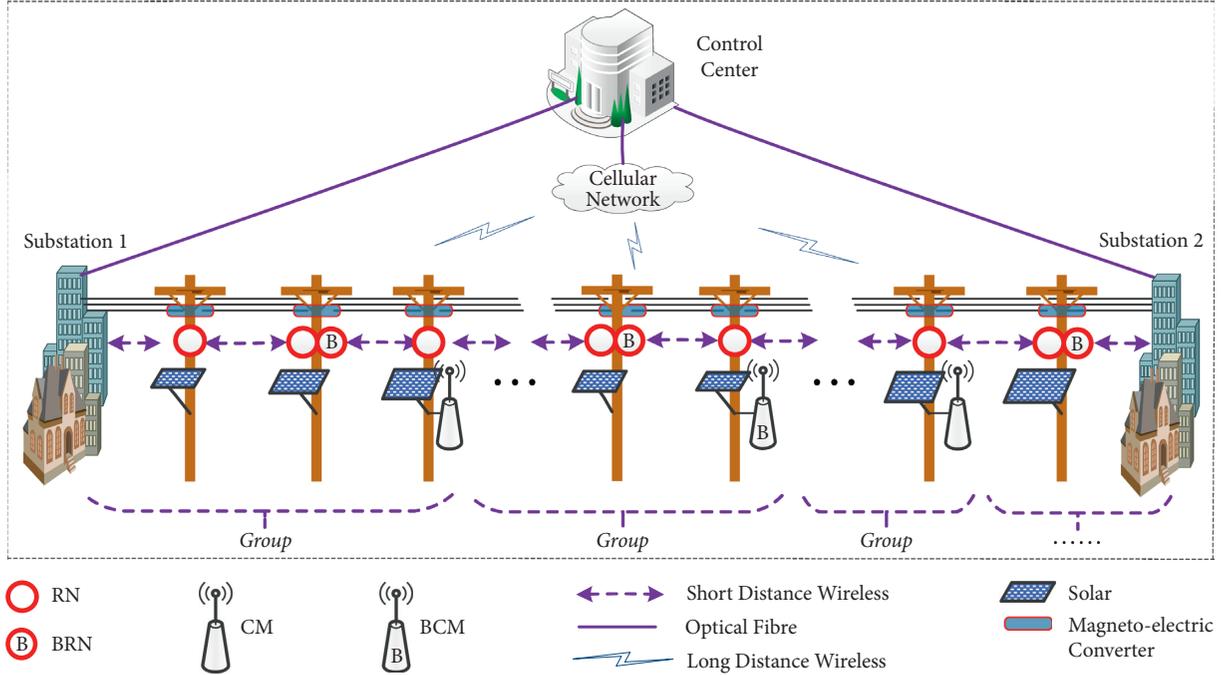


FIGURE 1: Fault-tolerance placement model.

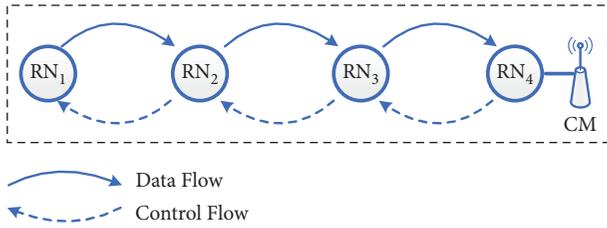


FIGURE 2: Data relay scheme.

sensing action and then compresses data to reduce redundancy due to duplicated data. After these two actions, the relay node transmits compressed data to the next relay node. Therefore, for the most of relay nodes, received data come from themselves and from the previous relay node. Moreover, there is an optional action before transmitting, backing up data, if the relay node has a backup node.

Several types of sensors are considered (e.g., tension sensors and insulator degradation sensors). RN_{i+1} senses and collects sensor data $SD_{i+1,j}$ from sensor type j around pole $i+1$. Consequently, the data from itself are $\sum_j SD_{i+1,j}$. After sensing and collecting data, the RN_{i+1} then compresses these data. If there is the backup relay node for RN_{i+1} , RN_{i+1} backs up the compressed data before these data are transmitted [9]. If r is a ratio of compression, the received data in RN_{i+1} can be calculated as $\sum_j SD_{i+1,j} + r \cdot SD_i$.

Suppose that sensing data (SD) are the same size in each RN, and the received data in last relay node SD_q are $((q-1) \cdot r + 1) \cdot SD$; if q nodes are in group Gp_k , $k \in [1..m]$, where $|Gp_k| = q$, therefore, the total data to be transmitted by the CM are $r \cdot q \cdot SD$.

Energy consumption is composed of at least three parts for RN_i according to actions: energy for sensing En_{sen} ,

compressing En_{comp} , and transmitting En_{tx} . The En_{backup} stands for optional backup energy due to fault tolerance. The energy of sensing, compressing, and backup is only relevant to data from itself, while the energy of transmitting is related to the previous relay node. Therefore, for RN_i , we have $En_{sen} = p_s \cdot SD_i$, $En_{comp} = p_c \cdot SD_i$, and $En_{backup} = p_b \cdot r \cdot X_i \cdot SD_i$. p_s , p_c and p_b are energy parameters of sensing, compressing, and backing up one bit of data, respectively. X_i functions as an indicator if RN_i has the backup node. For simplicity, we replace these parameters with a compound p_0 and use En_{scb} in place of En_{sen} , En_{comp} and En_{backup} . In general, under the assumption of a $(1/d^n)$ path loss, we obtain

$$\begin{aligned} En_{scb} &= (p_0 + r \cdot X_i) \cdot SD_i, \\ En_{tx} &= (p_1 + p_2 \cdot d^n) \cdot SD_i, \end{aligned} \quad (1)$$

where p_1 is energy/bit of the transmitter and receiver electronics, and p_2 is that for the transmit op-amp. Parameter d is transmission distance between poles. Thus, energy consumption of RN_i is $En_i = En_{scb} + En_{tx}$. From the FaTPM, we know that energy consumption is unequal or uneven for each relay node because there are hop-by-hop relay and differences for transmitting data.

4.3. Time Delay. In a surveillance application, time delay is extremely important compared to other factors [4]. The relay can be carried out in a parallel manner in groups. We concentrate on time delay in a group. The time delay can be divided into two sections, relay time delay within a group td_{rt} and directed link delay td_{dl} ; the first represents delay between relay nodes in the same group; the other represents transmission delay by the directed link CM.

The td_{rt} is related to the specific protocol. In [17], Zigbee is used since it has low cost and energy requirements. For wireless technology, we must consider collision avoidance owing to the hidden station problem in WSNs. The most commonly available technology is CSMA/CA. For simplicity, we refer to td_{at} as an average channel access time with CSMA/CA. Notably, there is a unique characteristic of the CM: in contrast to a base station, CM has wider bandwidth for uplink than downlink. Consequently, for the specific group Gp_j , time delay td_{rt} within the group is the sum of $td_{at} \cdot (|Gp_j| - 1)$ and $(r/tr_{zigbee}) \sum_{k=1}^{|Gp_j|} SD_k$, $k \in Gp_j$. Thus, we have two components of delay: the time delay within a group and the delay of direct transmission:

$$\left(td_{at} (|Gp_j| - 1) + \frac{r \sum_{k=1}^{|Gp_j|} SD_k}{tr_{zigbee}} \right) + \frac{r \cdot SD_{|Gp_j|}}{tr_{cm}} \leq D_t, \quad (2)$$

where tr_{zigbee} and tr_{cm} signify a transmission rate of Zigbee and CM, respectively. D_t is the threshold of time delay.

4.4. Energy Harvesting. Energy is one of fundamental factors in WSNs [25]. Since overhead power lines in distribution power networks are always in open areas, relatively ample sunlight is available. However, solar energy is prone to weather changes, for example, clouds and rain. Typically, it results in instability. Accordingly, we add a battery Bt to overcome instability. Previous research has proposed a wireless charging approach to solve this problem [26], but linear topology limits the effectiveness of a charging algorithm. Alternatively, magnetic-field energy harvesting can be a beneficial supplement in the absence of sunshine [27], because power lines are always surrounded by electromagnetic fields. In practice, solar panels and magnetoelectric transducers can be firmly fixed to the pole, as shown in Figure 1.

Assuming that the charging rate of solar energy is symmetric during day time, we can achieve

$$ch(t) = \frac{w}{\sqrt{2\pi\sigma^2}} e^{-((t-u)^2/2\sigma^2)}, \quad t_r < t < t_s, \quad (3)$$

where w is determined by the size of the solar panel, and t_r and t_s (hours) are the sunrise time and sunset time, respectively. Because of the unbalanced energy consumption, solar panels are naturally not all the same size, as shown in Figure 1. Furthermore, we simplify magnetoelectric energy harvesting into a time-related function; hence, it can be defined as $me(t) = \phi t / (t_b - t_a)$, $t_a < t < t_b$, where ϕ is proportional to the number of coil turns and the flux gradient [19]. t_a , t_b are parameters with the dimension of hours. Similar to solar energy harvesting, magnetoelectric energy harvesting varies adaptively due to the uneven energy consumption of relay nodes. Therefore, for the specific relay node RN_i , as the maximal time of data reporting is D_t , we have

$$En_i \leq Bt + \int_{t_r}^{t_s} ch_i(t) dt + \int_{t_a}^{t_b} me_i(t) dt \quad (4)$$

to show that the energy of consumption should not exceed the harvesting energy.

4.5. Coverage. The CM is relevant to group information. Researchers have previously resolved CM numbers in detail [17]. Fault tolerance, however, is neglected from previous studies. To achieve fault tolerance, we relax boundaries of groups and double the number of CM because the strict boundaries leave no space for fault tolerance. Thus, different from [17], we use the last relay node as the representative node of a group to assemble data and transmit data directly, as shown in Figure 2.

As there is different spacing between poles, for CM_j , if R_j is a coverage range, the number of poles covered by CM_j is at least

$$\frac{R_j}{\max_{i, (i+1) \in Gp_j} \text{Dist}[RN_i, RN_{i+1}]} + 1. \quad (5)$$

The part of denominator indicates the maximum distance between adjacent poles in group Gp_j .

4.6. Fault. In open areas, the frequency at which faults occur is generally not equal in different places owing to the geographical locations. Moreover, some types of faults occur with high-frequency (e.g., partial discharge faults at the peaks of hills or mountains), while other types of faults occur with low frequency. For the purpose of fault tolerance, we cannot treat different position equally. We use weight vector wei to identify the different placement. We denote fre_{\max} and fre_{\min} as the maximum frequency and minimum frequency, respectively. fre_{avg} represents the average frequency. The frequency is determined based on historical records. We attempt to assign a greater weight to high-frequency faults, whereas low-frequency faults are weighted based on the harmonic mean. Thus, the effects are less extreme for fault events that occur with low frequency. Therefore, for RN_i , the weight wei_i is

$$wei_i = \begin{cases} \frac{(fre_i - fre_{\text{avg}})^2}{(fre_{\max} - fre_{\min})^2}, & fre_i \geq fre_{\text{avg}}, \\ \frac{i}{\sum_i 1/fre_{\text{avg}} - fre_i}, & fre_i < fre_{\text{avg}}. \end{cases} \quad (6)$$

When a fault occurs, it should be tolerated by the system. Specifically, a backup relay node should take control and ensure continued operation. Given the total number of backup relay nodes (TN_{BRN}) as well as the total number of backup cellular modules (TN_{BCM}), we should carefully determine placement of backup nodes to maximize tolerance capability. Because the CM is also likely to fail, we should also back up the cellular module to keep the monitoring system functional.

$$\sum_i X_i \cdot wei_i \leq TN_{BRN}, \sum_j Y_j \leq TN_{BCM}. \quad (7)$$

The backup numbers of relay nodes, TN_{BRN} and TN_{BCM} , have been studied in previous literature [16, 17].

4.7. Cost. Cost is one of crucial factors in realizing the smart grid [28]. The cost of fault-tolerance system consists of four parts: cost for relay nodes and backup; $Ex_{RN} + X_i \cdot Ex_{RN}$; cost for the cellular module and backup; cost for solar panels and cost for the magnetoelectric transducer. The last two parts are $l \cdot \{Ex_{SP} + Ex_{ME}\}$, where l is a position coefficient in a group. As CM and backups for CM need a license and subscription fee sub from other operation companies, the cost is $(1 + \text{sub})(Ex_{CM} + Y_i \cdot Ex_{CM})\epsilon$.

Functioning as data sensing, compressing, communication, and control, relay nodes have high cost as the

equipment of energy harvesting and cellular module is closely related to them. Therefore, backups for all relay nodes are impractical. But we still need to consider the fault tolerance. Therefore, we determine the appropriate placement of backups to maximize capability of fault tolerance.

Since backup nodes should work exactly the same as original relay nodes in the event of a fault, the backups are also constrained by factors. As the number of backup nodes is resolved in [16], we target backup cost in this placement problem. Thus, the IBDP can be stated as follows.

Given the definitive backup numbers, TN_{BRN} and TN_{BCM} , we determine the optimal deployment placement of backup relay nodes to maximize fault tolerance. The IBDP can be formulated mathematically as follows: given system model $G = (V, E)$, RN , CM , TN_{BRN} , TN_{BCM} and group information Gp , we obtain the exact solution for the following minimal optimization problem:

$$\begin{aligned} f(X_i, Y_j) &= \sum_{i=1}^{|RN|} Ex_{RN} \cdot X_i + (1 + \text{sub}) \sum_{j=1}^{|CM|} Ex_{CM} \cdot Y_j \\ &+ \sum_{j=1}^{|CM|} \sum_{k=1, k \in Gp_j}^{|Gp_j|} l \cdot (Ex_{SP_k} + Ex_{ME_k}), \\ \text{s.t.} &\left(td_{at}(|Gp_j| - 1) + r \cdot \sum_{k=1}^{|Gp_j|} \frac{SD_k}{tr_{zigbee}} \right) \cdot X_i + Y_j \\ &\cdot r \cdot \left(\frac{SD_{|Gp_j|}}{tr_{cm}} \right) \leq D_t, k \in Gp_j, \\ En_i &\leq \left(Bt + \int_{t_r}^{t_s} ch_i(t)dt + \int_{t_a}^{t_b} me_i(t)dt \right) \\ Y_j &\cdot (|Gp_j| - 1) \cdot \max_{i, (i+1) \in Gp_j} \text{Dist}[RN_i, RN_{i+1}] < Y_j \cdot R_j, \end{aligned} \quad (8)$$

$$\sum X_i \cdot wei_i \leq TN_{BRN},$$

$$\sum Y_j \leq TN_{BCM},$$

$$\left| \bigcup_{j=1}^{|CM|} Gp_j \right| = |RN|,$$

$$X_i, Y_j \in [0, 1], i \in [1..|RN|], j \in [1..|CM|].$$

Equation (9) constrains time delay. Energy is limited in equation (10). Equations (11) and (12) account for the coverage and fault weight, respectively.

5. Fault-Tolerance System

In this section, we first introduce requirements of the fault-tolerance system. Then, we provide an analysis of tolerance

requirements. Next, the IBDP is reformulated in compliance with N-1. Finally, we propose the dynamic optimal deployment placement (DODP) algorithm to solve the fault-tolerance placement problem.

5.1. Tolerance Requirements. For a reliable power system, N-1 plays a significant role in fault tolerance [29]. N-1

implies that if any independent component fails or is removed, it cannot cause power system failure among N components of a power system [12]. In this subsection, we discuss N-1 requirements for communication system.

5.2. N-1 Implications for Communication Networks. For communication networks, we focus on node faults since many link faults are caused by relay nodes in wireless technologies. Further, a fault involves time of occurrence and duration time, besides the frequency and type, which are already discussed. More accurately, fault tolerance means that the backup really works for the time in which a fault has occurred and continued. Therefore, to fully understand faults, we introduce two variables, TTR and TBF, which are related to the fault time t [30]. TTR_i stands for the time required to repair faults on i pole, while TBF_i signifies the time between adjacent failures on i pole. The state of each RN_i is either TTR_i or TBF_i . Thus, the N-1 can be transcribed as follows: $\forall j$ in n poles,

$$P\left(\sum_{i,i \neq j}^{n-1} TBF_i(t), TTR_j(t), TBF_j(t)\right) \approx \sum_i^n TBF_i(t). \quad (9)$$

Equation (9) suggests that the probability of TTR_j be as small as possible to be neglected.

Definition 1. (N-1 requirements for communication networks). In the monitoring system, for n relay nodes, $\forall j, j \in n, \exists \varepsilon > 0$ is small positive number, we have

$$\lim_{t \rightarrow +\infty} P\{TTR_j(t) < \varepsilon\} = 1. \quad (10)$$

Equation (10) focuses on a specific fault j , with the same meaning as (9).

Proposition 1. (N-1 back up completely). For compliance with the N-1 requirements, for a linear monitoring system, N-1 requires that all relay nodes be completely backed up at all times.

Proof. Suppose that there is a linear network $\{RN_1, RN_2, \dots, RN_n\}$ composed of n poles. f_j indicates that a fault has occurred on the j th pole, $j \in [1..n]$. According to Definition 1, $\forall t$, there is always a sufficiently small positive number ε that makes the expression $\lim_{t \rightarrow +\infty} P\{TTR_i(t) < \varepsilon\} = 1$ hold. This expression indicates that the event $TTR_i(t) < \varepsilon$ is an inevitable event as time passes. The inevitable event implies that when the fault f_j has occurred on the j th pole, fortunately, there is always a backup relay node if we neglect the switching time from the fault node to backup node. As j is arbitrary, we can conclude that each pole requires a backup relay node to comply with the requirements of N-1. \square

5.3. Relaxation. A relay node is more expensive than an ordinary sensor node. It is impractical to back up all relay nodes as the above Proposition 1 claims. Therefore, N-1 requirements are too strong to apply for an actual fault-

tolerance system. Naturally, we attempt to relax N-1 for practical use. Here, we perform the relaxation based on two assumptions: (1) a threshold Thr is allowed for a monitoring application in a power system, and (2) TTR is limited and not long time. Thus, we can translate the N-1 as follows: in observation time T ,

$$P(TTR_i(t) < Thr) = 1, t \leq T. \quad (11)$$

Equation (11) makes compliance of the communication network with the N-1 requirements feasible. For the tolerance system, N-1 requires the probability $P(TTR_i(t) \geq Thr)$ to be as close to 0 as possible, even if a fault occurs in practice.

If we look closely at it, TTR_i can be an accumulation of multiple faults in i pole for several types of sensors. That is, $TTR_i = \sum_j TTR_{i,j}$. Generally, the different fault types on a pole are not independent, whereas the faults in different poles are independent. An occurring fault is always relevant to the $TTR_{i,j}$. Suppose that k is a fault number (FN); the event $TTR_i \geq Thr$ can be reformulated according to the total probability theorem,

$$P(TTR_i(t) \geq Thr) = \sum_k P(X = k) \cdot P\left(\sum_j TTR_{i,j} \geq Thr | X = k\right). \quad (12)$$

$\sum_j TTR_{i,j}$ is the cumulative time of the fault, and k is the specific fault number. Then, (12) can be rearranged as follows:

$$P(TTR_i(t) \geq Thr) = \sum_k P(X = k) \cdot P\left(\sum_{j=1}^k TTR_{i,j} \geq Thr\right). \quad (13)$$

We assume that FN follows a Poisson distribution with λ and $TTR_{i,j}$ follows a normal distribution [30]. Thus, (13) is the probability of sum of nonindependent random variables, essentially, that is, convolution of computing. In fact, (13) is not helpful in making decisions since it involves complex integration. We resort to a generating function of the random variables to achieve approximate results. Here, we directly provide a proposition. More details can be found in the Appendix.

Proposition 2. (binomial approximation). For a fault that has occurred, N-1 can be approximately computed by the binomial distribution $B(n, (\lambda_i/n))$ with the constraints that λ_i is very small and $\mu < 0$. That is,

$$P(TTR_i \geq Thr) \approx P\left(B\left(n, \frac{\lambda_i}{n}\right) \geq Thr\right) \longrightarrow 0 \text{ with } \mu < 0, \\ \lambda_i \longrightarrow 0, \\ n \longrightarrow \infty. \quad (14)$$

The proposition signifies that if the occurring fault is a rare event and TTR_i is very small, we can approximately

replace the fault time with the fault number, which is subject to a binomial distribution with constraints.

5.4. Problem Reformulation. The fault-tolerance system integrates optimal placement and fault tolerance. We can reformulate the problem based on the IBDP. Given system model $G = (V, E)$ and related parameters, we obtain a fault-tolerance system that complies with the N-1 requirements. Therefore, when a fault has occurred, the fault-tolerance system can be reformulated as follows:

$$f(X_i, Y_j) \text{ s.t. (9)(10)(11)(12)(13)(14)(20).} \quad (15)$$

5.5. Dynamic Algorithm. In operation of networks, we compute the optimal backup placement to provide the best locations to obtain fault tolerance within observation time T . In the next observation time, we can repeat this process and achieve the new optimal placement. We need to record the TTR of faults. Thus, from sequential time series, we can compute the optimal placement dynamically at multiple observation times. In addition, the TTR time of faults can further accumulate data of prior probability and facilitate the IBDP, as demonstrated in Algorithm 1.

The algorithm can be run in a server of the control center (CC). CC sends control commands to relay nodes by control flow, as shown in Figure 2.

5.6. Complexity and Confidence Interval. We provide complexity of above algorithm with respect to network scale n . The complexity of Step 1 is $O(n)$. For steps 3 to 7, complexity heavily depends on the 0–1 integer linear program (ILP) owing to Step 5. Since each binary variable is either 0 or 1, the complexity is $O(k * 2^m)$ theoretically in the worst case, where m is the number of binary variables, and k is the number of constraints. According to (8), we have $m \leq n$ since the number of variables m is definitely not more than the network scale n . Therefore, complexity of the algorithm is $O(k * 2^n)$ in the worst case.

For the confidence interval of (14), we use the normal approximation [31], $p \pm z_{1-\alpha/2} \sqrt{p(1-p)/n}$, $z_{1-\alpha/2} = 1.96$ for 95% confidence since it is easy to calculate.

6. Evaluations and Results

6.1. Experimental Configuration. We use Matlab to perform experiments to validate the model and algorithm. We ran codes on a CPU with a 2.5 GHz Intel i5 processor, 4 GB memory, and Linux Mint 19.1 operating system. The configuration data are mainly from the testbed and are slightly modified in terms of fault tolerance [9]. We consider a maximum of 100 poles/towers spanning approximately 40km, with an average span length of 1300ft. Additionally, t_{at} is 41ms, and D_t is 2s. The parameters p_1 and p_2 are (50nj/bit) and (100pj/(bit/m²)), respectively. t_r and t_s are defined as 9:00 and 16:00 of each day. The cost ratio of Ex_{SP} , Ex_{EM} and Ex_{CM} is 5: 10: 20. The distance between poles follows a normal distribution. We run experiments 50

times and calculate the average value. We experimentally verify the proposed DODP algorithm and compare it with random deployment and the BBPP algorithm from [16].

6.2. Metric. We use a tolerance ratio (Tr) to evaluate the fault tolerance capability, as defined in (16). A correct placement, $cp_i \in \{0, 1\}$, means that a fault has occurred on a pole i that has a backup node. A false placement, $fp_i \in \{0, 1\}$, denotes that a fault has occurred on a pole i that has no backup in (16). The higher Tr indicates greater tolerance. Meanwhile, we also consider overheads and time delay.

$$Tr = \frac{\sum_i cp_i}{\sum_i cp_i + \sum_i fp_i}. \quad (16)$$

6.3. Results. First, we show how sensor data (SD) affect time delay with respect to different numbers of poles. SD varies from 2K, 4K to 8K. As shown in Figure 3, the time delay strictly increases with the number of poles, which makes sense because a greater number of poles represent a longer distance and more time consequently. The parameter r varies from 0.125, 0.25 to 0.5, as shown in Figures 3(a)–3(c), respectively. The time delay is 6.2049s where SD is 8K and r is 0.5. A large amount of data undoubtedly consumes plenty of time in the sensing, compressing, and transmitting processes. However, time always has a deadline in monitoring applications; therefore, a large amount of data requires careful design. On the basis of data in Figure 3, we conclude that a reasonable data volume and a high compression ratio r are two goals of the system. However, a high compression ratio r will lead to greater energy consumption for the relay nodes.

Figure 4 evaluates the ratio of energy consumed in the sensing, compressing, backing up, and transmitting process. According to experimental data, the energy of backing up is nearly 0, which indicates that we can ignore the backing up process when considering energy consumption. In Figure 4, we compare energy consumption of sensing and compressing (En_{scb}) with that of transmitting (En_{tr}). Furthermore, we display the parameter n in (1), $n = 2$ in Figure 4(a), $n = 2.5$ in Figure 4(b), and $n = 3$ in Figure 4(c). The comparison reveals that transmitting energy is the main source of energy consumption and En_{scb} energy is a minor source in Figure 4(a). However, in Figure 4(c), these roles are reversed. The results confirm that the index n in (1) is strongly related to energy consumption. Thus, in design of the communication system, we should focus on the outdoor transmission path.

The accurate time delay for different groups is shown in Table 2 since groups are indispensable. In contrast to the previous study [17], we double groups in the FaTPM based on fault tolerance. Table 2 reveals that the time delay sharply decreases if we double groups. For clarity, we list group 8 twice. Table 2 also demonstrates that time delay for different transmission rates by Zigbee and cellular. Ranges of Zigbee are from 20K, 31.25K to 250K. An interesting result is that the time delay in group 14 is the same as that in group 16.

Require: G , observation T , threshold time Thr , and other parameters including location vector \mathbf{RN}
 Ensure: optimal backup location vector \mathbf{RN} with $N-1$

- (1) Initial location vector \mathbf{RN} ;
- (2) $t = 0$
- (3) **while** $t \leq T$ **do**
- (4) When a fault f_i occurs, we access parameters including TTR_i , TBF_i .
- (5) Compute the optimal location with (15) and output \mathbf{X}_i , \mathbf{Y}_j
- (6) Update the backup location \mathbf{RN} adaptively and configure the backup nodes
- (7) **end while**
- (8) Output \mathbf{RN} , TTR_i and TBF_i for the step 1.

ALGORITHM 1: Dynamic optimal deployment placement.

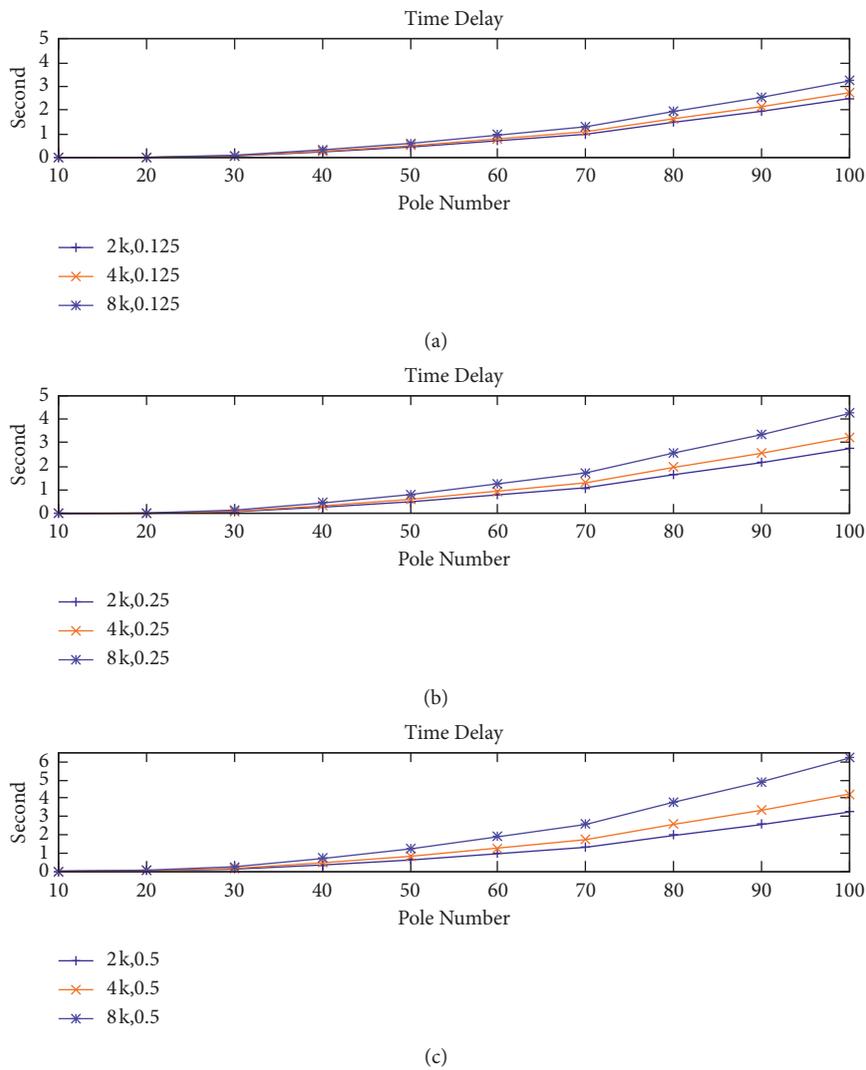


FIGURE 3: Time delay with sensor data SD and compression ratio r .

The same conclusion can be drawn from group 9 and group 10. Therefore, even though time delay can be decreased by increasing the cellular modules, it is not reduced substantially when the group number increases.

We now consider the overheads in Figure 5 with group information Gp . Generally, many groups lead to much heavier overheads. The overheads are displayed in Figure 5(a) with SD 8K and in Figure 5(b) with SD 4K under

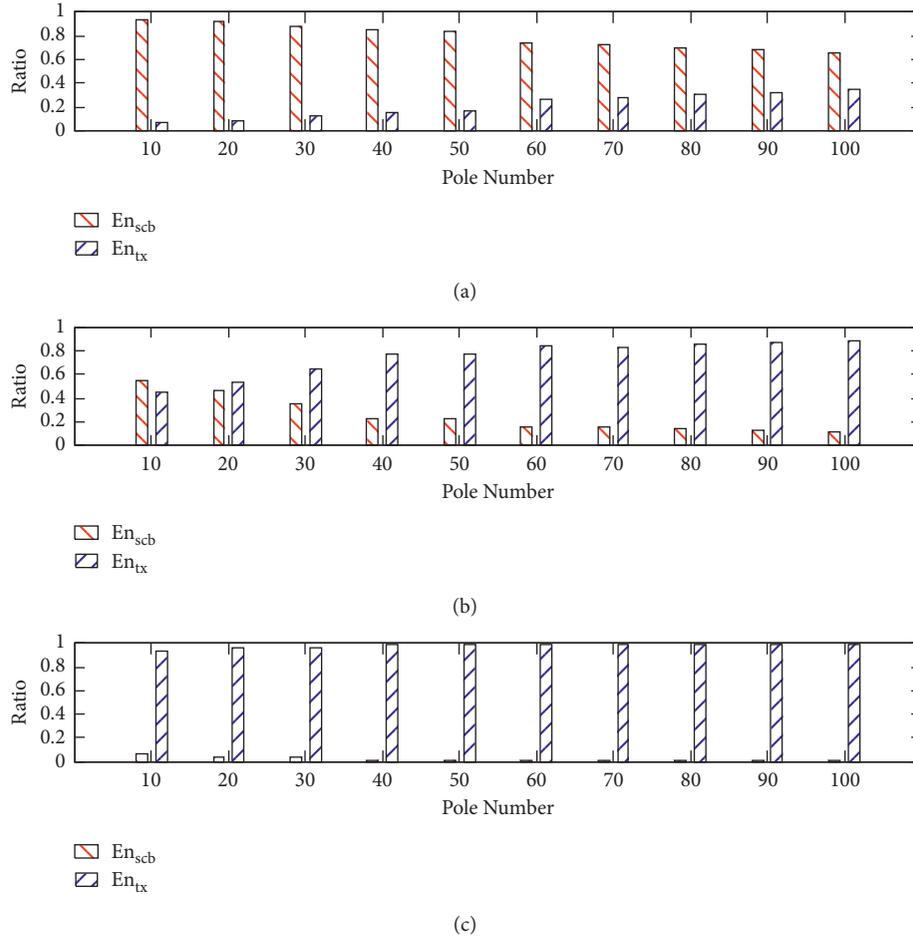


FIGURE 4: Comparison of energy consumption with n . (a) $n=2$. (b) $n=2.5$. (c) $n=3$.

TABLE 2: Time delay (s).

Data	Zigbee			Cellular			
	20K	31.25 K	250K	20 M	50M	1G	
Group	4	24.395 3	20.975 2	15.655 2	18.695 2	18.692 0	18.691 3
	5	15.284 4	13.124 5	9.764 4	11.684 4	11.681 9	11.681 3
	6	9.842 6	8.438 6	6.254 6	7.502 6	7.500 5	7.500 1
	7	6.877 2	5.887 2	4.347 2	5.227 2	5.225 4	5.225 0
	8	5.593 2	4.783 3	3.523 3	4.243 3	4.241 6	4.241 3
Double	9	3.422 8	2.918 8	2.134 9	2.582 9	2.581 6	2.581 3
	8	5.593 2	4.783 3	3.523 3	4.243 3	4.241 6	4.241 3
	10	3.422 8	2.918 8	2.134 9	2.582 9	2.581 6	2.581 3
	12	1.782 5	1.512 4	1.092 4	1.332 4	1.331 5	1.331 3
	14	1.161 0	0.981 0	0.701 0	0.861 0	0.860 2	0.860 0
	16	1.161 0	0.981 0	0.701 0	0.861 0	0.860 2	0.860 0
	18	0.672 1	0.564 1	0.396 0	0.492 0	0.491 4	0.491 3

the condition of 2 second time delay. Figure 5(a) shows much higher expense than that of Figure 5(b) since more data need to be transmitted in the restricted time. Accordingly, much more energy is consumed for transmission. The energy constraint becomes sensitive in the optimization, which consequently leads to more cost owing to the solar and magnetoelectric transducer. Figures 5(c) and 5(d) show different fault probabilities from 3% to 5%. In general, a

scenario with high fault probability needs more backup nodes for the relay nodes.

The fault-tolerance rate Tr is listed in Table 3. We compare the DODP algorithm with random deployment and based prior backup placement (BPBP) algorithm [16] in 14 groups. The BPBP algorithm outperforms the random algorithm because it appends prior information of faults. Moreover, the DODP algorithm achieves the highest fault-

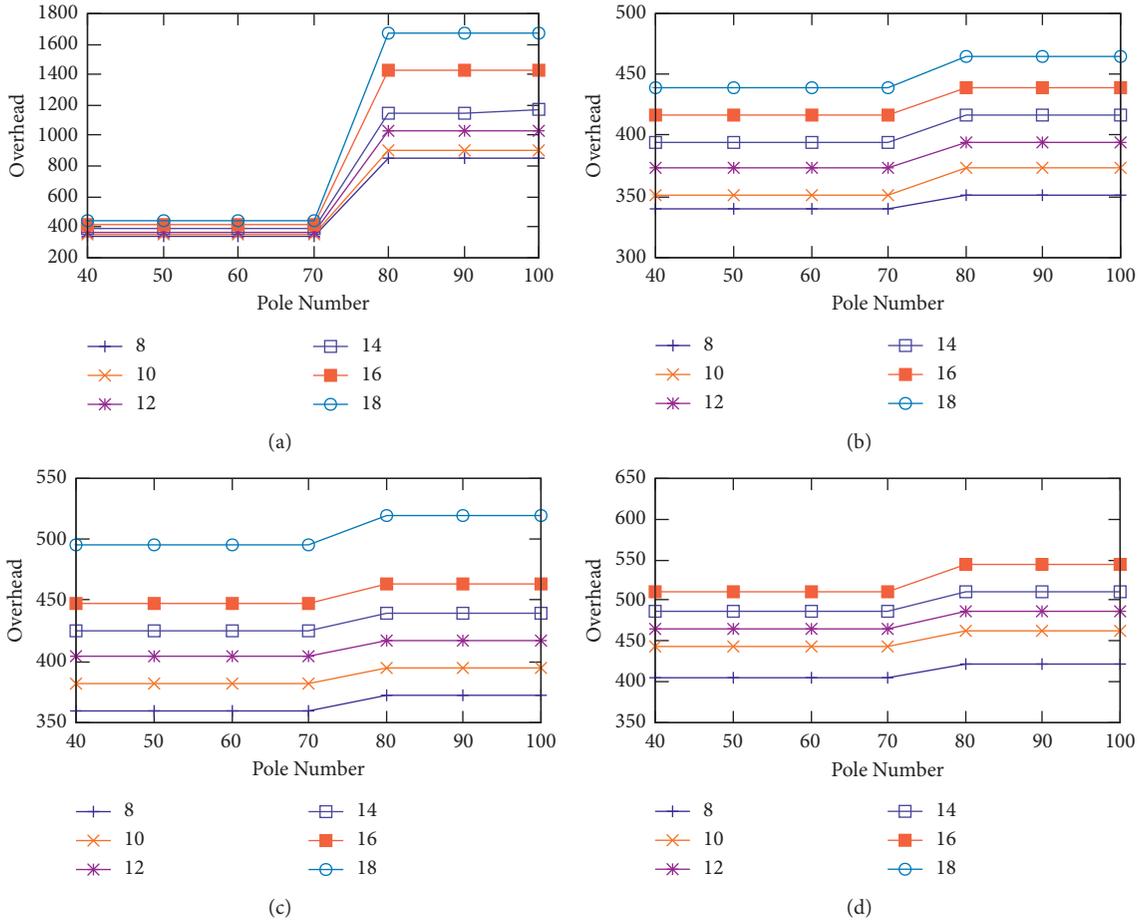


FIGURE 5: Overheads with different groups G_p and sensor data SD .

TABLE 3: Fault-tolerance ratio (Tr).

Data		40	50	60	70	80	90	100
$T = 1$	Random	0.571 4	0.614 9	0.668 8	0.704 7	0.787 6	0.751 3	0.740 9
	BPBP	0.626 4	0.690 4	0.720 6	0.743 5	0.822 8	0.813 9	0.804 7
	DODP	0.705 2	0.724 2	0.782 7	0.826 6	0.881 9	0.868 5	0.884 7
$T = 5$	Random	0.607 5	0.646 6	0.696 4	0.725 6	0.741 3	0.714 4	0.704 4
	BPBP	0.732 5	0.758 9	0.786 1	0.826 3	0.886 1	0.851 3	0.847 9
	DODP	0.766 1	0.810 3	0.858 1	0.898 3	0.926 3	0.901 7	0.896 3
$T = 9$	Random	0.593 1	0.619 6	0.677 6	0.708 9	0.791 0	0.757 6	0.741 2
	BPBP	0.646 0	0.699 0	0.743 7	0.781 6	0.861 3	0.827 7	0.799 2
	DODP	0.694 1	0.736 3	0.784 5	0.820 4	0.895 7	0.875 6	0.886 1

tolerance rate. From data in Table 3, the average Tr is improved by 8.646%, 6.472%, and 6.423% in DODP compared to BPBP for 1-, 5-, and 9-month periods, respectively. Therefore, we conclude that the fault-tolerance ratio in DODP improves at least 6% compared to that of the previous BPBP algorithm.

7. Discussion and Conclusion

Surveillance of overhead transmission lines by wireless sensor networks, which is one of smart grid applications,

faces challenges. In common wireless sensor networks, researchers have developed cutting-edge algorithms and traded off in many factors, such as energy, time delay, and cost. Moreover, with the linear topology and fault tolerance requirements, the surveillance application has posed greater challenges. We have proposed a fault-tolerance framework to deal with this tough questions. In the framework, we have explored backup method to provide fault tolerance. When we examine the backups, numbers of backup, placement of backup, and the adaptive switching algorithm should be determined.

In this paper, we follow the fault-tolerance framework and analyze deployment of relay nodes to monitor transmission power lines via wireless sensor networks. Based on a fault-tolerance placement model, we propose an initial backup deployment problem to determine the optimal backup locations of relay nodes. To build a flexible communication network, we ensure that the fault-tolerance complies with the N-1 of the power system and propose a dynamic optimal deployment placement algorithm. The numerical results present interesting highlights and confirm that the dynamic optimal deployment placement algorithm exceeds the performance of previous algorithms by at least 6%.

In the future work, we will study adaptive switching following the fault-tolerance framework. Specifically, adaptive switching from a faulty node to a backup node to ensure sustained data acquisition also represents additional challenges for researchers.

Data Availability

Some or all data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Science Foundation of Ministry of Education of China (nos. 21YJCZH197 and 20YJA910005) and was also partially supported by Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (no. 2020L0252).

References

- [1] F. Salvadori, C. S. Gehrke, A. C. de Oliveira, M. de Campos, and P. S. Sausen, "Smart grid infrastructure using a hybrid network architecture," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1630–1639, 2013.
- [2] K. Y. Bendigeri, J. D. Mallapur, and S. B. Kumbalavati, "Direction based node placement in wireless sensor network," in *Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, Coimbatore, India, March 2021.
- [3] Bureau of Labor Statistics, "Line Installers and Repairers," 2018, <https://www.bls.gov/ooh/installation-maintenance-and-repair/line-installers-and-repairers.htm>.
- [4] B. Fateh, M. Govindarasu, and V. Ajjarapu, "Wireless network design for transmission line monitoring in smart grid," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 1076–1086, 2013.
- [5] X. Huan, K. S. Kim, S. Lee, E. G. Lim, and A. Marshall, "A beaconless asymmetric energy-efficient time synchronization scheme for resource-constrained multi-hop wireless sensor networks," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1716–1730, 2020.
- [6] Y. Yang, D. Divan, R. G. Harley, and T. G. Habetler, "Design and Implementation of Power Line Sensor for Overhead Transmission Lines," in *Proceedings of the 2009 IEEE Power Energy Society General Meeting*, pp. 1–8, Calgary, AB, Canada, July 2009.
- [7] M. Z. A. Bhuiyan, G. Wang, J. Wu, J. Cao, X. Liu, and T. Wang, "Dependable structural health monitoring using wireless sensor networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 4, pp. 363–376, 2017.
- [8] L. I. Khudonogova and S. V. Muravyov, "Feasibility estimation of creating fault-tolerant prioritized transmission scheme in wsn," in *Proceedings of the International Siberian Conference on Control and Communications*, Omsk, Russia, May 2015.
- [9] K. S. Hung, W. K. Lee, V. O. K. Li et al., "On wireless sensors communication for overhead transmission line monitoring in power delivery systems," in *Proceedings of the 2010 First IEEE International Conference on Smart Grid Communications*, pp. 309–314, Gaithersburg, MD, USA, Oct 2010.
- [10] P.-Y. Kong, C.-W. Liu, and J.-A. Jiang, "Cost-efficient placement of communication connections for transmission line monitoring," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 4058–4067, 2017.
- [11] C. Vellaithurai, A. Srivastava, S. Zonouz, and R. Berthier, "Cpindex: cyber-physical vulnerability assessment for power-grid infrastructures," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 566–575, 2014.
- [12] C. He, L. Wu, T. Liu, and Z. Bie, "Robust co-optimization planning of interdependent electricity and natural gas systems with a joint n-1 and probabilistic reliability criterion," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 2140–2154, 2017.
- [13] S. Hu, X. Chen, W. Ni, X. Wang, and E. Hossain, "Modeling and analysis of energy harvesting and smart grid-powered wireless communication networks: a contemporary survey," *IEEE Transactions on Green Communications and Networking*, vol. 4, no. 2, pp. 461–496, 2020.
- [14] F. Ye, Y. Liang, H. Zhang, X. Zhang, and Y. Qian, "Design and analysis of a wireless monitoring network for transmission lines in smart grid," *Wireless Communications and Mobile Computing*, vol. 16, no. 10, pp. 1209–1220, 2016.
- [15] X. Wang, Y. Zhang, G. B. Giannakis, and S. Hu, "Robust smart-grid-powered cooperative multipoint systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6188–6199, 2015.
- [16] K. Wang, X. Qiu, S. Guo, and F. Qi, "Fault tolerance oriented sensors relay monitoring mechanism for overhead transmission line in smart grid," *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1982–1991, 2015.
- [17] Y.-C. Wu, L.-F. Cheung, K.-S. Lui, and P. W. T. Pong, "Efficient communication of sensors monitoring overhead transmission lines," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1130–1136, 2012.
- [18] F. Ye, Y. Qian, and R. Q. Hu, "Energy efficient self-sustaining wireless neighborhood area network design for smart grid," *IEEE Transactions on Smart Grid*, vol. 6, no. 1, pp. 220–229, 2015.
- [19] T. Hosseinimehr and A. Tabesh, "Magnetic field energy harvesting from ac lines for powering wireless sensor nodes in smart grids," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 8, pp. 4947–4954, 2016.
- [20] Y.-J. Kim, J. Lee, G. Atkinson, and H. Kim, "Thottan: SeDAX: a scalable, resilient, and secure platform for smart grid communications," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1119–1136, 2012.

- [21] R. N. Moraes, C. Ribeiro, and C. Duhamel, "Optimal solutions for fault-tolerant topology control in wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 5970–5981, 2009.
- [22] S. Sapre and S. Mini, "Moth flame based optimized placement of relay nodes for fault tolerant wireless sensor networks," in *Proceedings of the 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–6, Bengaluru, India, July 2018.
- [23] L. Tang, Y. Rui, Q. Wang, H. Li, and Z. Bu, "A reverse transmission mechanism for surveillance network in smart grid," in *Proceedings of the 2013 Proceedings IEEE INFOCOM*, pp. 3435–3440, Turin, Italy, April 2013.
- [24] J. Xia, K. Yu, C. Yun et al., "A novel mechanism for surveillance transmission in smart grid," in *Proceedings of the 2012 International Conference on Smart Grid (SGE)*, pp. 1–8, Oshawa, ON, Canada, Aug 2012.
- [25] G. Kakamanshadi, S. A. V. I. T. A. Gupta, and S. Singh, "A survey on fault tolerance techniques in wireless sensor networks," in *Proceedings of the IEEE International conference on Green Computing and Internet of Things (ICGCIoT2015)*, Noida, India, October 2015.
- [26] P. Zhou, C. Wang, and Y. Yang, "Self-sustainable sensor networks with multi-source energy harvesting and wireless charging," in *Proceedings of the IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pp. 1828–1836, IEEE, Paris, France, April 2019.
- [27] E. B. Pehlivanoglu, M. Ozger, O. Cetinkaya, and O. B. Akan, "Harvesting-throughput trade-off for wireless-powered smart grid iot applications: an experimental study," in *Proceedings of the 2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, Kansas, MO, USA, May 2018.
- [28] C. Pisani, A. Vaccaro, and D. Villacci, "Conceptualization and experimental deployment of an adaptive synchronized sensing system for power line thermal monitoring," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2158–2165, Dec 2016.
- [29] K. Sundar, H. Nagarajan, L. Roald, S. Misra, R. Bent, and D. Bienstock, "Chance-constrained unit commitment with n-1 security and wind uncertainty," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 1062–1074, 2019.
- [30] M. Xia, M. Tornatore, C. U. Martel, and B. Mukherjee, "Risk-aware provisioning for optical wdm mesh networks," *IEEE/ACM Transactions on Networking*, vol. 19, no. 3, pp. 921–931, 2011.
- [31] D. B. Lawrence, T. T. Cai, and A. DasGupta, "Interval estimation for a binomial proportion," *Statistical Science*, vol. 16, no. 2, pp. 101–117, 2001.