

Article

# SF-Partition-Based Clustering and Relaying Scheme for Resolving Near–Far Unfairness in IoT Multihop LoRa Networks

Dick Mugerwa <sup>1</sup>, Youngju Nam <sup>1</sup>, Hyunseok Choi <sup>1</sup>, Yongje Shin <sup>2</sup> and Euisin Lee <sup>1,\*</sup>

<sup>1</sup> School of Information and Communication Engineering, Chungbuk National University, Cheongju 28644, Republic of Korea

<sup>2</sup> Research Institute for Computer and Information Communication, Chungbuk National University, Cheongju 28644, Republic of Korea

\* Correspondence: eslee@cbnu.ac.kr

**Abstract:** Long range (LoRa) is one of the most successful low-power wide-area networking technologies because it is ideally suited for long-distance, low-bit rate, and low-power communications in the unlicensed sub-GHz spectrum utilized for Internet of things (IoT) networks. The effectiveness of LoRa depends on the link budget (i.e., spreading factor (SF), bandwidth (BW), and transmission power (TX)). Due to the near–far effect, the allocation of a link budget to LoRa devices (LDs) in large coverage regions is unfair between them depending on their distance to the GW. Thus, more transmission opportunities are given to some LDs to the detriment of other LD's opportunities. Numerous studies have been conducted to address the prevalent near–far fairness problem. Due to the absence of a tractable analytical model for fairness in the LoRa network, however, it is still difficult to solve this problem completely. Thus, we propose an SF-partition-based clustering and relaying (SFPCR) scheme to achieve enormous LD connectivity with fairness in IoT multihop LoRa networks. For the SF partition, the SFPCR scheme determines the suitable partitioning threshold point for bridging packet delivery success probability gaps between SF regions, namely, the lower SF zone (LSFZ) and the higher SF zone (HSFZ). To avoid long-distance transmissions to the GW, the HSFZ constructs a density-based subspace clustering that generates clusters of arbitrary shape for adjacent LDs and selects cluster headers by using a binary score representation. To support reliable data transmissions to the GW by multihop communications, the LSFZ offers a relay LD selection that ideally chooses the best relay LD to extend uplink transmissions from LDs in the HSFZ. Through simulations, we show that the proposed SFPCR scheme exhibits the highest success probability of 65.7%, followed by the FSRC scheme at 44.6%, the mesh scheme at 34.2%, and lastly the cluster-based scheme at 29.4%, and it conserves the energy of LDs compared with the existing schemes.

**Keywords:** long range (LoRa); spreading factor (SF); lower spreading factor zone (LSFZ); higher spreading factor zone (HSFZ)



**Citation:** Mugerwa, D.; Nam, Y.; Choi, H.; Shin, Y.; Lee, E. SF-Partition-Based Clustering and Relaying Scheme for Resolving Near–Far Unfairness in IoT Multihop LoRa Networks. *Sensors* **2022**, *22*, 9332. <https://doi.org/10.3390/s22239332>

Academic Editors: Giacomo Peruzzi and Alessandro Pozzebon

Received: 25 October 2022

Accepted: 25 November 2022

Published: 30 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

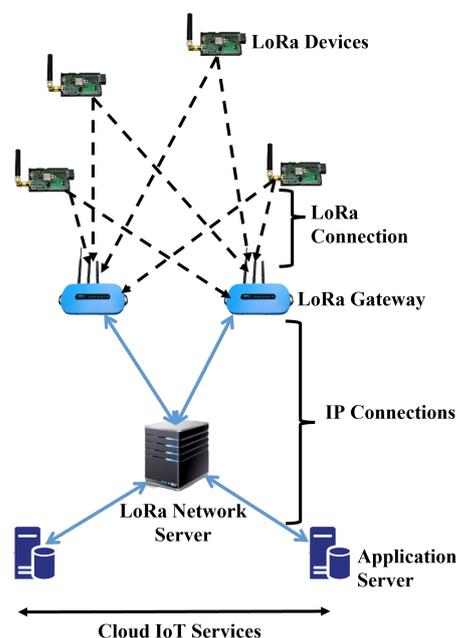
In the dawn of the new era of the Internet of things (IoT), the fast development of wireless communications and device technologies enables the interconnection and exchange of data among sensors. In the near future, the IoT will have an extraordinary impact on humans as a result of an increase in the ubiquity of the Internet by integrating every object with a motive of interaction via embedded systems in many fields such as asset tracking, agriculture, smart homes, and smart cities [1].

There are numerous predictions being made about the growth of connected devices.

According to recent studies [2], every person will be expected to have at least six IoT devices, with a monthly global mobile traffic of 77 exabytes. Usually, the devices are expected to cover a large geographical area for a longer span of five to ten years and send interesting data of small size, such as humidity, temperature, and other variables around them, over a longer distance using wireless communications.

For reflecting the above requirements, low-power wide area networking (LPWAN) technologies with a low cost, low bandwidth, and a low per-unit consumption have been developed for ubiquitous IoT connectivity of a large geographical area [3], different from conventional IoT networking technologies such as ZigBee and Bluetooth based on a shorter communication range. As representative protocols of LPWAN, LoRa, Sigfox, RPMA (Random Phase Multiple Access), and the Weightless protocol have been developed for LPWAN-based IoT implementation within the unlicensed ISM bands. Among them, LoRa is one of the most successful LPWAN technologies because it is ideally suited for long-distance, low-bit rate, and low power communications. LoRa communication technology was first proposed by SemTech and is now being developed by LoRa™ Alliance [4]. LoRa is a physical layer technology that uses a proprietary spread spectrum technique to modulate signals in unlicensed sub-GHz ISM bands (868 MHz in Europe, 915 MHz in North America, and 433 MHz in Asia).

LoRa provides bidirectional communication via the chirp spread spectrum (CSS) modulation, which spreads a narrow-band signal over a wider channel bandwidth, making it more interference resistant [5]. LoRa primarily employs six spreading factors (SF)  $\in \{7, 8 \dots, 12\}$  to adapt to the trade-off between the data rate and the communication range. In fact, the greater the SF, the greater the range with a lower data rate, and vice versa [6]. The data rates vary with the SF and channel bandwidth, ranging from 300 bps to 50 kbps. According to [7], LoRa technology is based on a star of stars topology, in which GWs relay messages between end devices and a central network server. This model is distinguished by the provision of a long-distance and reliable link via a special modulation technique, in which a LoRa GW collects raw data directly from end devices by single-hop communications and forwards it to a network server (NS) connected by a high-speed backhaul network as shown in Figure 1 [8].



**Figure 1.** General network architecture for IoT LoRa networks.

LoRaWAN is a LoRa MAC layer that employs the ALOHA protocol [9], which is managed by an NS [10]. Furthermore, LoRaWAN defines three types of devices with capabilities (classes A, B, and C). Class A devices use pure ALOHA access for the uplink, are battery limited, and can only receive an ACK in two receiving windows during downlink transmission. Class B devices, in contrast to class A, are beacon-based and can open an additional receiving window at predetermined times. Except when transmitting, class C devices are always listening to a channel and consume a lot of energy [11]. The three classes can coexist on the same network, and devices can switch between them. Because of the

orthogonality of the sub-bands and the quasi-orthogonality of different SFs, a LoRa GW can receive packets from multiple LoRa devices at the same time. The LoRa Alliance has defined the higher layers and network architecture known as LoRaWAN, with the medium access control protocol (MAC) layer essentially being an ALOHA variant of random access.

Generally, a massive number of LoRa devices may be connected to a GW because of the vast coverage area factor, which creates a near–far fairness problem due to the heavy route loss as the bottleneck in overall system performance in LoRa networks. Moreover, LoRa networks also suffer from collisions of concurrent transmissions on the same channel and SF in closely packed deployment situations. In this regard, the conventional protocol for pure ALOHA in IoT LoRa networks is proven to be oversimplified because it fails to account for channel fading, power control, and aggregate interference [12]. To overcome this issue in LoRa, previous studies of [13–15] proposed multihop communication schemes using relay devices and a programmed e-node that acted as a range extender to increase the reliability of data forwarding from LoRa devices with duty-cycle constraints to a GW in IoT LoRa networks. In this case, class C devices acted as relay devices, intercepting data transmissions from LoRa devices by overhearing and relaying them to the GW. Furthermore, each relay device alternated between overhearing the end devices' transmissions during the receive window and forwarding the contents of the overheard packets to the GW on a regular basis. However, these studies focused on only increasing coverage extension probability while ignoring fairness in terms of network area success probability.

### 1.1. Background and Motivation

To solve the shortcomings of the existing studies, we propose an SF-partition-based clustering and relaying scheme to address the near–far unfairness problem caused by the low success probability, long channel occupancy time (also known as time on air (ToA)), redundancy, low coverage, capture effect challenge, and collisions in IoT LoRa networks. For this, the proposed SFPCR scheme has three methods: SF-based network partitioning, density-based subspace clustering, and relay LD selecting. The SF-based network partitioning method is used to determine the appropriate partitioning threshold point for bridging the packet delivery success probability gap between different SF regions in an IoT LoRa network. Then, the packet delivery success probability is evaluated based on the hop distance of each hop in multihop LoRa communications. By the partitioning threshold point, the IoT LoRa network is separated into two SF zones: the lower SF zone (LSFZ) and the higher SF zone (HSFZ). To avoid long-distance transmissions with low reliability from individual LDs in the HSFZ of the GW, the density-based subspace clustering method creates arbitrary-shaped clusters by grouping adjacent LDs in the HSFZ and selects the CH of each cluster for gathering and transmitting data packets of cluster members. For clustering the arbitrary-shaped clusters and choosing CHs based on the CH conditional probability, a combination of the DBSCAN algorithm and the naive Bayes classifier is used. To support reliable data transmissions from each CH in the HSFZ to the GW by multihop communications, the relay LD selecting method chooses the best relay LD in the LSFZ to extend uplink transmissions from the CH to the GW. Then, the best relay LD is determined by the harmonic mean of the packet success probability and the remaining energy selection strategy of all candidate LDs in the LSFZ.

### 1.2. Contribution

In order to achieve massive scalability and connectivity, we aim to maximize the minimum success probability of all distant LDs and optimize SF allocation in all SF zones of the IoT LoRa network within a predetermined duty-cycle limitation.

The detailed contributions of this paper can be summarized as follows.

1. We provide an SF-based network partitioning method to determine the best partitioning threshold point for dividing SF zones of an IoT LoRa network into the LSFZ and the HSFZ. To bridge the performance gap among the SF zones, an optimal SF division method based on a heuristic algorithm is proposed.

2. We provide a density-based subspace clustering method for grouping adjacent LDs in the HSFZ. To construct clusters of arbitrary shape and to select CHs by a binary score representation, a combination of the DBSCAN algorithm and naive Bayes classifier is proposed.
3. We provide the relay LD selecting method to ideally choose the best relay LD only in the LSFZ for extending transmissions of CHs in the HSFZ to the GW by multihop communications. The harmonic mean based on the packet success probability and the remaining energy is proposed to choose the best one among all candidate relay LDs.
4. We maximize the performance of packet success probability, scalability, and fairness in the SF zones of an IoT LoRa network. To ensure the optimal zone allocation with connection fairness, the packet success probability of each LD in a certain SF zone is examined.
5. We conduct simulations in various environments for evaluating the performance of the proposed SFPCR scheme. Simulation results demonstrate that the proposed SFPCR scheme raises the packet success probability by bridging the packet delivery success probability gap between SF zones while conserving more energy among LDs by reducing the number of packets compared with existing schemes.

### 1.3. Organisation

The remainder of this paper is organized as follows. First, we provide the related works on the proposed scheme in Section 2. The system model of the proposed scheme is presented with the network, energy, and success probability models in Section 3. Our SF-partition-based clustering and relaying scheme is described in Section 4. The performance evaluation and numerical results are presented and discussed in Section 5. Finally, the paper is concluded in Section 6.

## 2. Related Works

Recently, numerous overviews on LoRa technology and LoRaWAN networks have been published [16,17]. They deal with various research fields such as network lifetime, spreading factor allocation, energy conservation, coverage, connectivity, and performance improvement. LoRa technology has grown in popularity in a variety of industries, including smart metering [18], smart cities [19], fleet tracking [20], and energy management. IoT applications necessitate devices transmitting over long distances, which exposes signals to interference from various sources. Similarly, LoRaWAN coverage and scalability have been studied both indoors and outdoors to achieve a coverage range of 5 to 15 km in an open area with limited interference [16]. However, due to the changing environment, these studies on interference have not fully mitigated the problem to improve the transmission success probability. Unlike the Sigfox communication model, LoRa uses chip spread spectrum (CSS) modulation, which allows for a higher receiver sensitivity and thus a greater resistance to interference. However, there is still a lot of research being conducted on how to reduce cospreading factor interference and collisions caused by two or more LoRa devices transmitting their packets at the same SF [6]. Many other works assume that the SFs are static and are assigned based on the distance between the LoRa devices and their nearest GW, forming a kind of tiered ring structure around each GW. To solve the CSS modulation problems of LoRa communications, there are three research areas: spread factor allocation, clustering, and multihop communication. Thus, we examine the three research areas in the next three subsections in detail, respectively.

### 2.1. Spread Factor Allocation in LoRa Networks

The performance of LoRa networks has been aided by appropriate resource allocation strategies, one of which is SF, which defines the number of bits that can be encoded by a symbol. SF assignment schemes include, but are not limited to, random, area-proportional, and distance-based distribution schemes from the GW [10,21]. Furthermore, SF assignments in the LoRa network are more attributable to the work in [22], which indicated that SF allocation was primarily based on the power level of the signal that the GW receives from

the end devices and the GW's sensitivity without taking end device location into account. Similarly, in [23], an efficient interference-aware SF allocation strategy was introduced. In reality, SF can be assigned based on the signal-to-noise ratio (SNR) of the received packets, and as a result, devices in a certain range can use the same SF [24]. However, the challenge to optimize the SF assignment over each channel realization and high-density building obstacles was ignored. In [25], the authors considered a fair SF clustering of end devices in LoRa networks using k-means clustering. Their algorithm determined the minimum cost function, which was defined as the distance between each point in the data set and its nearest centroid. Furthermore, the k-means clustering method achieved robust clustering with ease. However, this method of SF allocation did not take into account the SF overlap of edge nodes at different annulus boundaries as well as the distance of the end devices from the GW. Studies in [26] proposed an SF allocation algorithm based on matching theory to maximize the minimum achievable average rate in LoRa networks to improve network throughput and fairness. This theory classifies resource allocation as a many-to-one matching problem with conventional externalities and peer effects. In this case, a device prefers to be paired with the SF that provides the most utility. That algorithm, however, performed better with only heterogeneous LoRa devices. Reynders et al. [27] proposed a heuristic SF allocation scheme in which users with similar path losses were simply assigned to the same channel, then to each SF based on their distance from the GW. The work in [28] proposed a suboptimal SF allocation strategy to maximize packet success probability. Inter-SF interferences, on the other hand, were ignored in that work. In particular, the proposed SFPCR scheme allocates SF based on a combination of the device's distance and the SNR of the received packets. The SF allocation technique utilized in the SFPCR scheme is modified by assuming that all LoRa devices inside the same contour lines have the same SF.

## 2.2. Clustering in LoRa Networks

Clustering is one of the useful methods for enhancing network connectivity and prolonging the network lifetime in wireless networks by managing nodes efficiently. In clustering for IoT LoRa networks, LoRa devices are grouped into clusters, and the cluster head (CH) in each cluster acts as the mediator for data transmissions between LoRa devices and a GW. Clustering not only reduces packet redundancy but also manages efficient data routing, resulting in improved efficiency, scalability, and the avoidance of redundant message exchange. Clustering protocols enable end devices to reduce data packets on networks through data aggregation, which improves network lifetime and energy consumption. As a result, several studies based on clustering algorithms have been proposed for IoT LoRa networks. In [29], the authors proposed a two-stage energy-efficient cluster-based solution for data collection by a mobile sink in a robot network by considering the first stage as a travelling salesman problem and the second stage as the removal of some CH robots from the path to reduce the UAV's energy consumption. In [30], the cluster-based layering approach for uplink multihop communication was analysed. In that paper, the authors presented a multihop uplink communication scheme that took advantage of LoRa's ability to customize its transmission parameters and combines them with a novel routing protocol to solve network coverage and high energy consumption problems. Clusters were used in that scheme to associate an end device with exactly one GW, and once an end device was associated with a GW, the end device communicated through the GW. This type of scheme, however, limited the ability of end devices closer to the GW to transmit directly to the main GW. In [21], an SF-based clustering scheme was proposed to improve the multihop capacity of the LoRa network. Based on network clustering, each subnet rooted at a sink node was assigned a specific SF. The procedure focused on balancing airtime between subnets while maintaining connectivity.

However, end device extraction from a higher SF subnet has an impact on network maintenance and appears to be an unrealistic scenario. Based on this clustering, Ref. [31] described a method for logically partitioning end devices into several clusters. End devices in nearby clusters served as relay devices for end devices in the next more distant zone.

As a result, clustering with relaying provided an appealing solution to this problem by focusing on distant end devices one at a time. Clusters formed in that case were either the same or different widths. The clustering style in that approach considered clusters of the entire network but ignored the transmission capability of end devices closer to the GW.

As examined above, the existing studies concentrate on clustering techniques that partition the entire network in order to reduce energy consumption and improve connectivity between LoRa devices. However, these approaches deprive LoRa devices with a high link quality of directly connecting to the GW and acting as relay LoRa devices of distant devices. As a result of only forming subspace clusters in the designated HSFZ, among the benefits of the proposed density-based subspace clustering are a reduced energy consumption, communication overhead, and increased connectivity.

### 2.3. Multihop Communication in LoRa Networks

Multihop communications are used to extend the connection distance over the communication range of single-hop communications. To enable coverage extension in IoT LoRa Networks based on single-hop communications, many studies have been published in IoT LoRa networks. The first multihop LoRa network was investigated in [32]. As a result, when a data packet was broadcast, all neighbouring end devices checked its origin in terms of hop distance and relayed the messages if they were closer to the sink. The algorithm allowed the use of beacons for time synchronization as well as for the communication distance between each end device and the GW. However, it caused a lot of packet redundancy and delay, which slowed down network performance. Aslam et al. [19] investigated the use of multihop LoRa topologies to enable energy-efficient connectivity in smart city applications. In this study, packet reception rates for various source-to-destination distances, SFs, and transmission powers were used to evaluate single-hop and multihop LoRa topologies. Then, the results of the study demonstrated that two-hop networks could significantly outperform single-hop networks in terms of range. However, that study did not go into detail about the best relay placement to influence range extension for distant end devices and the powerful impact. In [33], the authors proposed a multihop communication scheme called FSRC based on a selective relay operation. By controlling communication parameters, the relay control strategy of the FSRC maximized both coverage probability and minimum success probability for all SF regions. Like other similar schemes [34,35], they introduced a programmed e-node and an implicit relay node to act as a transparent range extender with an overhearing operation, respectively. With the primary goal of replicating the same packet sent by distant end devices, they received it and relayed it to the GW by overhearing. However, these implementations generated a large amount of network traffic from broadcasts, resulting in performance degradation.

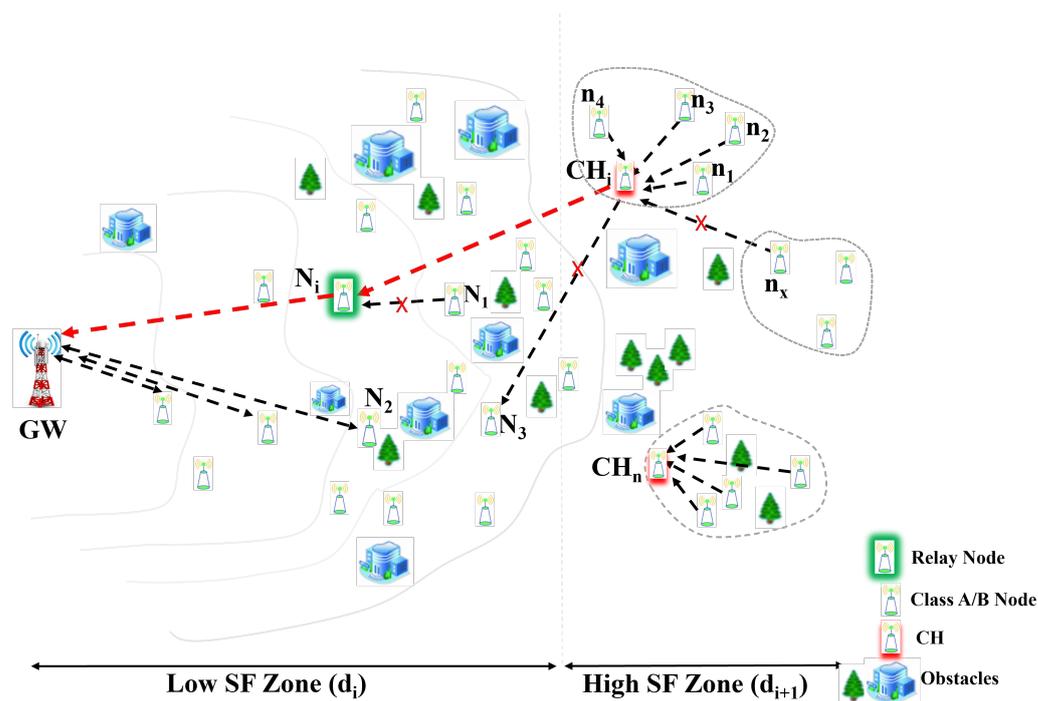
In [36,37], an energy-efficient multihop communication scheme called e2MCH and a LoRa mesh networking system for large-area monitoring of IoT applications were studied. The research presented in these papers introduced multiple relay devices for collecting data from IoT sensors spread across a large geographical area and forwarding them to a GW in close proximity. The research, however, was vulnerable to packet collisions, which reduced the reliability and energy efficiency of LoRa networks. The proposed SFPCR scheme employs a centralized optimum relay assignment scheme that iteratively chooses the best relay LoRa device from the candidate set of devices in the LSFZ based on both packet success probability and residual energy, to transmit a data packet from the subspace clusters in the HSFZ. Consequently, the proposed multihop approach offers a practical framework for reducing the transmission distance and transmission power required by distant nodes with large spreading factor values.

## 3. System Model

In this section, we provide a detailed description of the system model in the proposed scheme. We first present the network model for deploying an IoT LoRa network. Next, we derive the energy model and the success probability model for LoRa devices in various spreading factor zones based on the network model, respectively.

### 3.1. Network Model

As the IoT LoRa network shown in Figure 2, we considered a classic uplink LoRa system with a single GW at the centre of a network field of radius  $R = 6$  km and a set of a lot of LoRa devices (LDs), which was defined by  $LD = n_i | I \in \{1, \dots, N\}, N = |LD|$ . LoRa devices were allocated in Gaussian distribution within the network field. For the proposed SFPCR scheme, we divided the IoT LoRa network into two regions: a low spreading factor zone (LSFZ) and a high spreading factor zone (HSFZ). In our network model, a LoRa device located in the HSFZ joins the CH in a cluster through a single hop communication and sends its data packet to the CH. Then, the CH forwards data packets from LoRa devices in its cluster to a selected relay LoRa device in the LSFZ, which transmits these data packets in a single-hop communication to the GW, as shown in Figure 2.



**Figure 2.** Model of the relay and cluster process in the proposed SFPCR scheme.

The performance of LoRaWAN is restricted by a duty cycle in a sub-band for the amount of time spent transmitting packets, which is referred to as the time on air (ToA). As a consequence of this, we took into consideration a duty cycle ( $\delta_i$ ) of 1% [4], operating at a certain frequency ( $F_c$ ) in the 868 MHz band with a code rate (CR) of 4/5 and a variable bandwidth (BW) ranging from 125 kHz to 500 kHz. We explicitly deployed class A and B LoRa devices, with LoRa devices always initiating transmission. During a time slot, each LoRa device  $n_i$  with matching  $SF \in \{7, 8, \dots, 12\}$  and values for signal-to-noise ratio (SNR), receiver sensitivity for a specific bandwidth, and bit rate, as shown in Table 1, transmits a packet. The LoRa devices use different SFs for transmission to ensure orthogonality and multiple detections at the receiver.

In [38], the authors alluded that different SFs caused significantly different times on air for symbol transmission. A symbol can encode  $k$  information bits into a chirp with  $SF = k$ , and the bit rate  $R_b$ , so the symbol period is evaluated by  $T_{symbol} = \frac{2^k}{BW}$ , where BW is the bandwidth. When  $SF = k + 1$ , one symbol period is  $\frac{2^{k+1}}{BW}$ , which doubles the transmission time by sending only one more bit. However, a higher SF indicates a greater resistance to interference and noise, resulting in a greater communication range.

According to recent research, we assumed that an IoT LoRa network was divided into multiple annular rings of the same width. Thus, the width of the annulus was defined

here as the Euclidean distance between the two closest points on the inner boundary  $C_{in}$  and the outer boundary  $C_{out}$ , respectively. Recent work has shown a spreading factor distribution scheme based on a certain range of distances to the GW. According to [22,39], the spreading allocation is primarily based on the power level of the signal that the GW receives from the devices and the GW sensitivity without taking the device location into account. Consequently, it is not trivial to consider a general coverage contour metric to determine the degree of coverage at any point within the IoT LoRa network while avoiding interference from obstacles. In contrast, the proposed SFPCR scheme determines and allocates the appropriate SF based on the transmission distance and SNR of the received packets, as described in [24]. In addition, the SF allocation technique utilized in the SFPCR scheme was modified by assuming that all LoRa devices inside the same contour lines had the same SF. Due to obstructions and interference, LoRa devices may be the same distance apart but have different SNR values. Therefore, these devices cannot share an SF.

**Table 1.** LoRa parameters based on SF.

SF (i)	SNR (dB)	Receiver Sensitivity qs (dBm)	Bit Rates (kbps)
7	−6	−123	5.47
8	−9	−126	3.13
9	−12	−129	1.76
10	−15	−132	0.98
11	−17.5	−135.5	0.54
12	−20	−137	0.29

### 3.2. Energy Model

For the energy model of LoRa devices in the proposed scheme, we employed the same energy model as [38,40]. Unlike class C LoRa devices, which are usually mains-powered, the energy consumption of each LoRa device must be precisely stated. The energy consumption  $E_{tx,i}$  for transmitting a  $\beta$ -bit LoRa packet from a LoRa device LD(i) to the GW can be categorized as follows: the LoRa device waking up, low-power listening, radio preparation, signal transmission, radio off, and processing. Furthermore, the consumed energy for signal transmission is affected by the transmission power as well as the varying transmission time caused by the spreading factor. It is precisely defined as (1) [41],

$$E_{tx,i} = P_{tx} \times T_{SF(i)} \quad (1)$$

where  $P_{tx}$  denotes the transmit power with time on air for a LoRa device with  $SF(i)$ , and  $T_{SF(i)}$  denotes the time on air for LD(i) expressed as (2) and (5) [38].

$$T_{SF(i)} = T_{symbol} \times (N_{pr} + N_{pl}) \quad (2)$$

$$N_{pr} = (n_{preamble} + 4.25) \times \frac{2^{SF_i}}{BW} \quad (3)$$

$$N_{pl} = 8 + \max\left(\left\lceil \frac{8L - 4SF_i + 44}{4(SF_i - 2DE)} \right\rceil CR, 0\right) \times \frac{2^{SF_i}}{BW} \quad (4)$$

$$T_{SF(i)} = (20.25 + \max\left(\left\lceil \frac{8L - 4SF_i + 44}{4(SF_i - 2DE)} \right\rceil CR, 0\right)) \times \frac{2^{SF_i}}{BW} \quad (5)$$

In Equation (2),  $N_{pr}$  and  $N_{pl}$  represent the number of symbols in the packet preamble and payload, respectively. Substituting for  $T_{symbol}$ ,  $N_{pr}$ , and  $N_{pl}$  in (2), we get (5), if low data rate optimization is enabled,  $DE = 1$ , otherwise  $DE = 0$ ,  $CR = 4/5$  represents the code rate, and  $n_{preamble} = 8$  is the preamble length. It is worth noting that the energy consumption of a LoRa device in the annulus ( $j$ ) is proportional to airtime ( $\tau_j$ ), with large values for a high SF.

Subsequently, we evaluated the energy consumption  $E_{n-hops}(SF, P_{tx})$  using the multi-hop communication from a LoRa source device(i) to the GW through n hops. Then, the energy consumption for sending a packet was computed as follows:

$$E_{n-hops}(SF, P_{tx}) = \sum_{i=1}^n (E_{tx,i} + E_{rx,i}) \quad (6)$$

where  $P_{tx}$  is the consumed power according to the supply current and supply voltage for the transmitter,  $E_{tx,i}$  and  $E_{rx,i}$  are the energy consumption of the transmitter and receiver in the  $i$ th hop according to its current configuration on the transmitted power, spreading factor, and bandwidth, respectively.

Class A and B LoRa devices on the LoRa network typically spend the bulk of their time in light-sleep mode due to their low energy consumption and battery power. Meanwhile, in order to calculate the total dissipated energy consumption  $E_{total\_N_i}(t)$  of a LoRa device  $N_i$  at a specific time (t), we considered two states, active and inactive modes, given as (7) and (8) and thus derived  $E_{total\_N_i}(t)$  as follows.

$$E_{total\_N_i}(t) = E_{inactive\_N_i}(t) + E_{active\_N_i}(t) \quad (7)$$

Here,  $E_{inactive\_N_i}(t)$  is the dissipated energies in the state where devices are in sleep mode ( $E_{sleep\_N_i}(t)$ ) relative to 1.8  $\mu$ A [42] and  $E_{active\_N_i}$  is the dissipated energies by the LoRa device during the transmission mode, reception mode, data measurement, processing, and wake up. Thus,  $E_{active\_N_i}$  can be calculated as follows:

$$E_{active\_N_i}(t) = E_{Wu} + E_{proc} + E_{Wut} + \sum_{j=1}^n E_s(i) \quad (8)$$

where  $E_{Wu}$  is the energy consumption when a device wakes up,  $E_{proc}$  is the energy consumption of data processing, and  $E_{Wut}$  is the amount of energy to move the transceiver from sleep mode to active mode, respectively. In particular,  $E_s(i)$  is the energy consumption of a LoRa device in terms of data rate to transmit  $L_{packet}$  with  $n$ -hop communication evaluated as (9) [43] and is calculated as follows:

$$E_s(i) = \frac{L_{packet}}{R_b} \times (P_{tx} + P_{rx}) \quad (9)$$

where  $L_{packet}$  is the size of the transmitted packet in bits,  $P_{tx}$  and  $P_{rx}$  are the consumed power according to the supply current and supply voltage for transmitter and receiver, respectively, and  $R_b$  is the bit rate defined in Equation (10).

$$R_b = \frac{CR \times SF}{\frac{2^{SF}}{BW}} \quad (10)$$

Here,  $SF$  is the spreading factor,  $BW$  is the bandwidth, and  $CR$  is the code rate defined as  $\frac{4}{(\iota+4)}$  with  $\iota \in \{1, 2, 3, 4\}$ , respectively.

Ultimately, to understand the LoRa device's energy consumption, it is essential to efficiently utilize the LoRa devices' residual energy  $E_{re\_N_i}(t)$  at a given time (t) for the effective selection of CHs and also prevent the untimely demise of devices to ensure that the energy of each LoRa device in the network is consumed in a relatively balanced manner. Thus,  $E_{re\_N_i}(t)$  can be computed as (11),

$$E_{re\_N_i}(t) = E_0 - E_{total\_N_i}(t) \quad (11)$$

where  $E_0$  is the initial energy of LoRa devices.

### 3.3. Success Probability Model

In this section, we examine the success probability of a signal in LoRa's respective spreading factor regions under intra- and inter-SF interference. In this paper, we investigated the average success probability of LoRa devices in random locations in the field of an IoT LoRa network in order to determine the optimal number and positions of relay LoRa devices (LDs) in relation to the distance between a source LD and the GW. The proposed SFPCR scheme considers the evaluation of a source LD's packet success probability using a varying SF to be dependent on the distance from the receiving LoRa device (i.e., the source LD (S) to the CH, the CH to the relay LD (R), and R to the GW). The collision of packets in a LoRa network is classified into two types: intra-SF collision and inter-SF collision [44]. If a transmission collides, the LoRa device becomes backlogged and slows down, resulting in excessive delay and packet loss. As a result, it is critical for the LoRa link to have a high data reliability.

Transmission from LDs located within a distance  $\leq r$  to the GW is considered successful in a LoRa network with a uniform distribution of  $N$  nodes over the disk of density  $\lambda = \frac{N}{\pi r^2}$  if and only if a source LD (S) transmits a packet at distance  $x$  from the GW that is not affected by either intra-SF or inter-SF collisions. Furthermore, we assumed that LoRa devices released packets within predetermined time slots within a deployment radius of  $r$ , and that packet transmission intensity was  $\theta = \frac{pkt}{\tau}$  (bits per second), where  $pkt$  is the number of packets and  $\tau$  is the time interval in seconds.

Intra-SF interference occurs when LDs in the same spreading factor region and BW collide when the distance from the GW is less than  $xR$ . Here,  $x$  is a location of a LoRa node and  $R = e^{\frac{6}{10\gamma}} > 1$  [45], and the density of the LDs within the same spreading factor region  $SF(i)$  is represented by  $\rho_i = \frac{\alpha_i N}{\pi r^2}$  [44]. Subsequently, the number of potential interferers  $I_{SF(i)}$  within the same spreading factor region (intra-SF) is represented by Equation (12).

$$I_{SF(i)} = \alpha_i N \frac{(\min(xR, r))^2}{r^2} \quad (12)$$

Here,  $N$  is the number of LoRa devices in the disk,  $\alpha_i$  is the percentage of LDs in a particular SF with  $\sum_7^{12} \alpha_i = 1$ . Assuming that none of the potential interfering LDs initiate a transmission during the vulnerable period of  $2T_i$  duration,  $P_{suc}^i(loc(x))$  is the probability of successful transmission of a LoRa device at location  $x$  in an IoT network area and is represented by the Equation (13),

$$P_{suc}^i(loc(x)) := e^{-2T_{SF(i)}\theta I_{SF(i)}} \quad (13)$$

where  $T_{SF(i)}$  is the time on air evaluated in Equation (2).

Consequently, the combined success probability  $P_{suc}(S, GW)$  from the source LD (S) to the GW in the proposed SFPCR scheme takes advantage of the varying SF as defined in (14):

$$P_{suc}(S, GW) = P_{suc}^{SF(i)}(S, CH) \times P_{suc}^{SF(j)}(CH, R) \times P_{suc}^{SF(k)}(R, GW) \quad (14)$$

Here, S represents the source LoRa device  $n_i$ , CH is the cluster head represented as  $CH_i$ , R is the relay LoRa device also represented as  $N_i^*$ , and GW is the gateway, respectively. Then,  $P_{suc}^{SF(i)}(S, CH)$  is the success probability from S to CH,  $P_{suc}^{SF(j)}(CH, R)$  from CH to R, and  $P_{suc}^{SF(k)}(R, GW)$  from R to GW.  $SF(i)$ ,  $SF(j)$ , and  $SF(k)$  are the spreading factors assigned to each of them.

## 4. SF-Partition-Based Clustering and Relaying (SFPCR) Scheme

In this section, we explain the proposed SF-partition-based clustering and relaying (SFPCR) scheme in detail, which has three methods. First, we provide a detailed explanation of SF partitioning, followed by the density-based subspace clustering in the HSFZ and relay LoRa device selection mechanism, respectively.

#### 4.1. SF-Based IoT Network Partitioning

In the SFPCR scheme, we divided an IoT LoRa network field into two zones, the low spreading factor zone (LSFZ) and the higher spreading factor zone (HSFZ), to resolve the near-far fairness problem in the packet success ratio. In the HSFZ, the LoRa nodes were clustered to enhance data collection ability and transmission reliability, and the cluster head (CH) in each cluster used a LoRa node in the LSFZ as a relay node to forward data to the GW. Algorithm 1 shows an SF partitioning algorithm to find an optimal partitioning threshold point ( $P_{opt}$ ) for dividing an IoT LoRa network field into an LSFZ and an HSFZ. Then, a virtual partition threshold point ( $\Theta_{th}$ ) is used for making the LSFZ and the HSFZ. The algorithm employs a heuristic process with ( $0 \leq \Theta_{th} \leq 100$ ) until no packet success probability gaps exist. At each iteration, the algorithm evaluates the modified packet success probability of all the LoRa devices in the respective zones until  $P_{opt} = \Theta_{th}$ . Moreover, if  $P_{suc}^i(loc(x)) < \Theta_{th}$ , the packet success probability of LDs in the HSFZ is being improved and the reverse is true for the packet success probability in the LSFZ.

---

#### Algorithm 1: Pseudocode for boundary cut-off point of LSFZ and HSFZ.

---

**Input:**  $P_{suc}^i(loc(x))$ : LoRa device success probability  
**Input:**  $N$ : number of IoT LoRa devices  
**Input:**  $x$ : LoRa device's location  
**Input:** BW: bandwidth  
**Input:** SF: spreading factor of a LoRa device  
**Output:** Boundary point LSFZ and HSFZ,  $P_{th}$

- 1: Initialization
- 2:  $Sum \leftarrow 0$
- 3:  $P_{max} \leftarrow 0$
- 4:  $P_{opt} \leftarrow 0$
- 5: **for**  $\Theta_{th} = 0 : 100$  **do**
- 6:   **for**  $i = 1 : N$  **do**
- 7:     Calculate  $P_{suc}^i(loc(x))$  in Equation (13)
- 8:     Calculate  $P_{\Delta suc}^i(P_{suc}^i(loc(x)), \Theta_{th})$
- 9:      $Sum + = P_{\Delta suc}^i(P_{suc}^i(loc(x)), \Theta_{th})$
- 10:   **end for**
- 11: Calculate  $P_{avg}(\Theta_{th}) = \frac{Sum}{N}$
- 12: **if**  $P_{max} < P_{avg}(\Theta_{th})$  **then**
- 13:    $P_{max} = P_{avg}(\Theta_{th})$
- 14:    $P_{opt} = \Theta_{th}$
- 15: **end if**
- 16: **end for**
- 17: **return**  $P_{opt}$

---

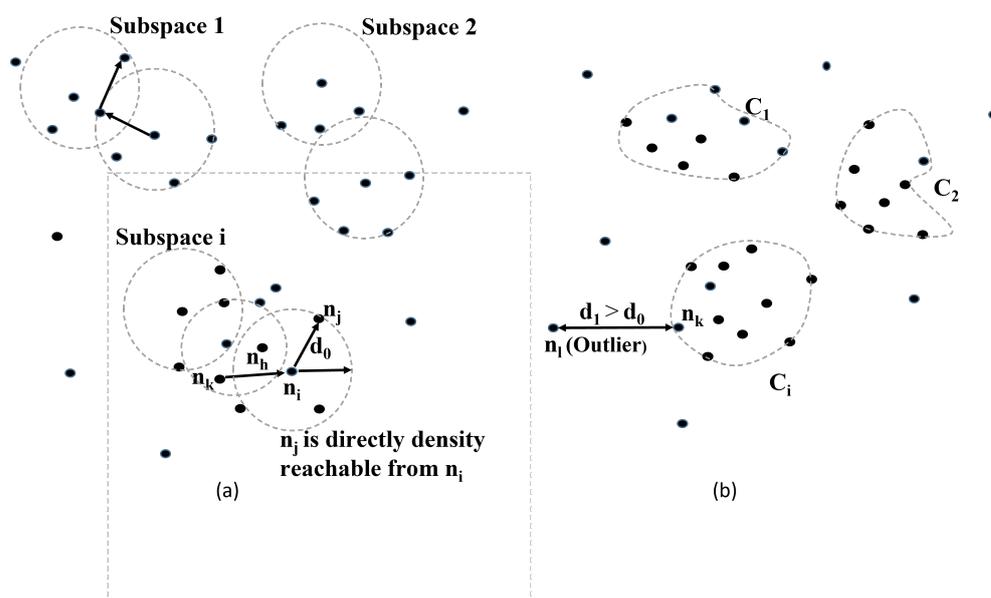
Subsequently, the GW was placed at the network's midpoint, and its location was used as a reference point to define the network's zones. It should be noted that the optimal coverage range of each zone determined the LoRa devices density deployed in each zone. The initialization and beaconing processes provided the GW with global knowledge of the entire network. Thence, we deduced the average success probability  $P_{avg}(\Theta_{th})$  for all LoRa devices, necessitating the determination of the partition threshold point  $\Theta_{th}$  for the SF zones. Specifically, distinct virtual zones could be formed based on the network reliability deduced from the average success probability  $P_{avg}(\Theta_{th})$  for all the LoRa devices at a particular distance from the GW as shown in Equation (15) for respective SF among all nodes.

$$P_{avg}(loc(x), \Theta_{th}) = \frac{1}{N} \sum_{i=1}^N (P_{suc}^i(loc(x)), \Theta_{th}) \quad (15)$$

In Equation (15),  $N$  is the number of LoRa devices whose packet success probability has already been registered at the GW. Under the condition that  $P_{suc}^i(loc(x)) > P_{opt}$ , LDs in the LSFZ have a success probability greater than the calculated threshold value. Here,  $P_{opt}$  is based on the average success probability ( $P_{avg}$ ) of all the LDs that take up the SF regions with faster chirps. Any LoRa device in the LSFZ, in particular, is a potential relay candidate for CHs. Otherwise, the remaining SF zone is referred to as the HSFZ. LoRa devices in the HSFZ have a lower success probability than the computed threshold if  $P_{suc}^i(loc(x)) < P_{opt}$ , which includes all LDs far from the GW and within the transmission radio range ( $T_r$ ) that suffer from near-far effect due to the fading channel and path loss caused by obstacles.

#### 4.2. Density-Based Subspace Clustering in the HSFZ

Clustering is one of the main methods used to divide a network into a number of groups (clusters), with one node designated as the cluster head (CH) for each cluster [46]. Intuitively, a cluster in the SFPCR scheme is precisely defined as a set of density-connected LoRa devices in a particular region. Usually, the expected type of cluster is determined by the clustering criterion. The SFPCR scheme constructs LoRa nodes in the HSFZ as clusters. For clustering of the HSFZ, we used a combination of the DBSCAN algorithm [47] and naive Bayes classifier, a density-based subspace clustering algorithm that created arbitrary-shaped clusters and eliminated outliers, as shown in Figure 3. The naive Bayes classifier [48] was used for the CH selection based on the CH's conditional probability. The DBSCAN clustering algorithm was suitable for the discovery of subspace clusters of arbitrary shape with a high efficiency in the network. The SFPCR scheme employs a centralized clustering strategy in which the network exerts complete control over the clustering procedure via the GW. In this method, the GW always selects which LD will function as CHs in the HSFZ, requiring periodic information on all LoRa devices to select the most suitable nodes. Moreover, there are eight standardized message types for LoRaWAN. We used the proprietary message type to encode a collection of messages necessary for network layering, cluster formation, and multihop message relaying. To facilitate proper cluster formation, we also defined a control message called the "HELLO" message. The GW periodically broadcasted the control message to all LDs within the cluster. The centralized clustering algorithm was dissected into two main phases: an initial setup phase and a steady-state phase for the subspace cluster formation and the CH's selection criteria, respectively. We explain each of them in detail.



**Figure 3.** Arbitrary subspace clusters in the HSFZ: (a) Node density-reachability, (b) Node density-connected arbitrary subspace clusters.

- **Initial setup phase:** During this phase, the GW applies global knowledge from all the LoRa devices to build density-based subspace clusters based on the LoRa devices' distribution and information. The SFPCR scheme generates density-based subspace clusters of LoRa devices whose neighbourhood contains a minimum number ( $LD_{min}$ ) of other devices within a given radius ( $d_0$ ). In this case,  $d_0$  is calculated using the  $k$ -nearest neighbour to its immediate neighbours [49]. Meanwhile, given the value  $k$  is represented as  $LD_{min}$ , we chose 4 as the  $k$  value, and the 4-nearest distance from  $LD(i)$  was the Euclidean distance from  $LD(i)$  to its 4-nearest neighbours, independent of the transmission range.

Without loss of generality, three types of LoRa devices were defined: core LoRa devices, which contained at least  $LD_{min} \geq 4$  LoRa devices [50] in their  $d_0$  neighbourhood. Border LoRa devices did not have enough devices in their neighbourhood, but they were close to some core LoRa devices. Finally, other devices were considered outliers. Assuming a set of LoRa devices  $N = \{N_1, N_2, \dots, N_n\}$  in the HSFZ made up of  $\psi$  subspaces  $\{C_i\}_{i=1}^{\psi}$  ( $i = 1, \dots, \psi$ ), let  $N_i$  be the set of  $S_i$  LoRa devices belonging to subspace  $C_i$  and  $n = \sum_{i=1}^{\psi} S_i$ . Starting with an arbitrary LoRa device  $n_i$ , the method returned all LoRa devices that were density-reachable from  $n_i$  using  $d_0$  and  $LD_{min}$ . As illustrated in Figure 3a, LoRa device  $n_j$  is directly density-reachable from LoRa device  $n_i$  if  $n_i$  is a core LoRa device and  $n_j$  is in its  $d_0$  neighbourhood. A LoRa device  $n_j$  is defined as density-reachable from a core LoRa device  $n_i$  if there exists a chain of devices from  $n_i$  to  $n_j$ , with each device being directly density-reachable from the prior LD. However, if  $n_i$  is a border device, no devices are density-reachable from  $n_i$  and then the algorithm visits the next LoRa device.

To deduce the parameter  $d_0$ , we first computed the Euclidean distances with Equation (16) of all the  $k(n_i)$ -nearest neighbours from Equation (17), and selected the maximum 4th-nearest neighbour distance  $d_0$  as shown in the Equation (18).

$$dist(n_i, n_j) = \sqrt{(X_{n_i} - X_{n_j})^2 + (Y_{n_i} - Y_{n_j})^2} \quad (16)$$

where  $dist(n_i, n_j)$  is the Euclidean distance from the core LD  $n_i$  to device  $n_j$ .

$$k = \min(LD_{max}, LD_{min}) \quad (17)$$

In this case,  $LD_{max}$  is the number of neighbours within an arbitrary subspace cluster distance ( $d_{max}$ ), and  $LD_{min}$  is the minimum number of LoRa devices to make a local subcluster, respectively.

$$d_0 = \underset{i}{\operatorname{argmax}} (dist(n_i, n_j)) \quad (18)$$

for  $n_j \in J_i$ , where  $J_i$  represents  $n_i$ 's nearest neighbours such that the number of LoRa devices in ( $J_i$ ) is equivalent to  $k$ . As a result, the algorithm merges all the local subclusters to form the arbitrary density-based subspace cluster based on ( $k(n_i) = LD_{min}$ ) as shown in Figure 3b. However, if ( $k(n_i) < LD_{min}$ ), no cluster is formed. A density connection can also refer to the relationship that exists between border LDs that are part of the same cluster but do not have a core LD in common with which they share any density reachability. A local subcluster is formed by a core LD and all of its retrieved nearby LDs within a predefined  $d_0$  distance, and it expands using the two fundamental concepts of density-reachable and density-connected, as long as the conditions are met [51]:

- $n_j$  is a member of a cluster  $C_i$  if  $n_i$  is a member of  $C_i$  and  $n_j$  is density-reachable from  $n_i$ .
- $n_i$  and  $n_j$  are density-connected if both of them are members of a cluster  $C_i$ .

Assume the distance in the HSFZ between two sets  $S_1$  and  $S_2$  of LDs is defined as  $dist(S_1, S_2) = \min\{dist(n_i, n_j) \mid n_i \in S_1, n_j \in S_2\}$ . Two sets with at least the density of the thinnest cluster will be separated only if the distance between them is greater

than  $d_0$ . The LoRa device density connection determines the maximum number of LDs ( $LD_{max}$ ) in an arbitrary cluster. Furthermore, in the steady-state phase, the GW is critical in computing the CHs of clusters by using network information from the IoT LoRa devices during network initialization.

- **Steady-state phase:** In this phase, the cluster head selection is performed using a naive Bayes classifier [48]. The GW plays a crucial role in generating the score value needed to categorize normal LoRa devices as CHs or cluster members. Based on the Bayes classifier, the GW determines the  $\kappa$  best LoRa devices to become CHs based on their binary score. The value of  $\kappa$  indicates the appropriate number of CHs. In [52], the authors explained what clustering algorithms were and how they worked. Moreover, there were three types of algorithms for choosing CHs: predetermined, random, and attribute-based. In the SFPCR scheme, we adopted attribute-based algorithms to select CHs. The GW constitutes the vector metric attributes  $A^{(i)}$  for a LoRa device LD(i) to become the CH.  $A^{(i)} = \{A_1^{(i)}, A_2^{(i)}, \dots, A_j^{(i)}\}$ , where  $j$  is the vector's dimension. The attributes of an LD(i) include, but are not limited to, the residual energy ( $E_{re\_N_i}(t)$ ), the link quality ( $L^{(i)}$ ) between LD(i) and its neighbour, and the distance ratio ( $D_{ratio}(n_i, N_i)$ ) evaluated as the minimum distance of a source node  $n_i$  to the relay LoRa device  $N_i$  in the LSFZ. As a result, a vector of attributes can be expressed as Equation (19).

$$A^{(i)} = \{E_{re\_N_i}(t), L^{(i)}, D_{ratio}(n_i, N_i)\} \quad (19)$$

Here,  $E_{re\_N_i}(t)$  (24) and  $L^{(i)}$  (13) are considered as the primary parameters, and  $D_{ratio}(n_i, N_i)$  is considered as the secondary parameter and is calculated using Equation (20).

$$D_{ratio}(n_i, N_i) = \left[ \frac{d_i}{\sum_{i=1}^N \left[ \frac{d_i}{N} \right]} \right]^{min} \quad (20)$$

In Equation (20),  $n_i$  is the source LD in the  $i$ th cluster,  $N_i$  is the selected relay LD at a time ( $t$ ), and  $d_i$  is the distance from the source LD to the relay LD, respectively. Then, the goal of the CH selection is to find an appropriate mapping relationship between  $C$  and  $A^{(i)}$ , where  $C$  is the binary score of attributes to which the LD belongs, either a cluster member (CM) or a cluster head (CH), and is expressed as  $C = \{b_0, b_1\}$ .

$$P(C = b_1 | A^{(i)}) = \frac{P(A^{(i)} | C = b_1)P(C = b_1)}{P(A^{(i)})} \quad (21)$$

$P(C = b_1 | A^{(i)})$  expresses the probability of LD(i) becoming the CH with the classification  $A^{(i)}$  using the Bayesian theorem expressed in Equation (21). Finally, the GW selects the LoRa device with the highest probability of becoming a CH within the clustered region based on Equation (22), which is also represented in Algorithm 2.

$$C \leftarrow \underset{b_1}{\operatorname{argmax}} (P(A^{(i)} | C = b_1)) \quad (22)$$

In order to create connectivity between LoRa devices LD(i) in the HSFZ, each CH communicates with cluster members within a transmission radio range ( $T_r$ ) shown as  $CH_i | d(LD(i), CH_i) < T_r$ . In addition, after a certain number of rounds at a time ( $t$ ), the network carries out the clustering procedure again for reclustering. Once reclustering is complete, the GW continuously monitors the residual energy and link quality of LoRa devices in the HSFZ to choose new CHs based on the updated parameter values to ensure network continuity. The energy threshold ( $E_{th}$ ) is established to prevent the early death of some LoRa devices owing to their uneven energy decrease. In the SFPCR scheme, the energy threshold was assessed using the average energy of all LDs.

**Algorithm 2:** Pseudocode to select cluster heads in the HSFZ.

---

**Input:**  $X = X_1, X_2, \dots, X_n$ : set of LoRa Devices (LDs)  
**Input:**  $A^{(i)}$ : LoRa Device metric attributes  
**Input:**  $N$ : number of LoRa Devices  
**Input:**  $\kappa$ : optimal number of CHs  
**Output:** Clusters of all LDs in HSFZ, CH allocation

- 1: Network initialization
- 2: All  $C^{(i)} = b_0$
- 3: **for**  $i \leftarrow 1 : N$  **do**
- 4:   Calculate the  $E_{re\_N_i}(t), L^{(i)}, D_{ratio}(n_i, N_i)$
- 5:   **if**  $L^{(i)} \geq L_{th}$  &&  $E_{re\_N_i}(t) \geq E_{th}$  **then**
- 6:     Add LD(i) to CH candidate set
- 7:   **else**
- 8:     Add LD(i) to cluster member (CM)
- 9:   **end if**
- 10: **end for**
- 11: **for** All LDs  $\in$  CH candidate set **do**
- 12:   Calculate the  $P(C = b_1 | A^{(i)})$
- 13:   **if**  $P(C = b_1 | A^{(i)}) = \max\{P(C = b_1 | A^{(i)})\}$  **then**
- 14:      $C^{(i)} = \{b_1\}$  // LD(i) is set as a CH
- 15:   **else**
- 16:      $C^{(i)} = \{b_0\}$  // LD(i) is set as a CM
- 17:   **end if**
- 18: **end for**
- 19: **return** CH

---

#### 4.3. Relay LoRa Device Selection in the LSFZ

In this paper, we considered exploiting LDs in the LSFZ as relay devices to support reliable data transmissions from source LoRa devices in the HSFZ to the GW. To achieve a reliable relay device selection policy, the SFPCR scheme uses a centralized optimal relay assignment method that iteratively allows a single best relay LoRa device (LD) among the relay candidate set of LoRa devices  $R = \{N_1, N_2, N_3, \dots, N_i\}$  in the LSFZ, to access the channel and assist the bidirectional transmissions of all the distant LDs (represented as CMs) through the CH in the HSFZ.

The goal of our method was to devise a way to randomly select the best relay device to aid bidirectional communication between the source LDs (represented as CHs) and the GW, while taking into account two important performance parameters, namely, the link quality in terms of packet success probability and residual energy  $E_{re\_N_i}(t)$  of the candidate relay LoRa devices. The  $CH_i$  sends the relay a wake-on-radio (WOR) frame. This WOR frame wakes up the selected relay and sends packet information. The system model with  $CH_i$  in the HSFZ acting on behalf of the source LDs is shown schematically in Figure 4. Meanwhile, the set  $R$  of potential candidate relays in the LSFZ includes LoRa devices, that is,  $R = \{N_1, N_2, N_3, \dots, N_i\}$ . Our method compares the metrics to determine the best relay  $N_i^* \in R$ . The packets collected by the  $CH_i$  are transmitted to  $N_i^*$  as the best relay from the candidate set because it meets the packet success probability as the primary evaluation parameter and residual energy threshold conditions for our relay device selection method. However, if any LD in the LSFZ does not meet the predefined threshold conditions, it cannot become a potential relay device.

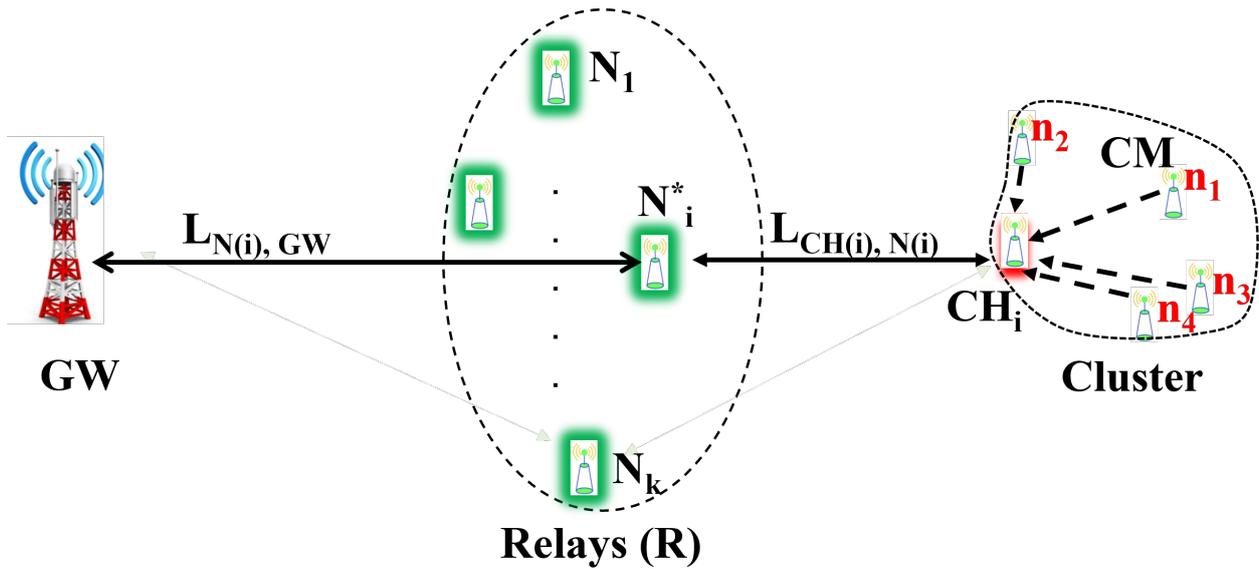


Figure 4. The relay selection model.

We present four main phases to specify our centralized optimal relay assignment method. With the four main phases, our method iteratively selects an optimal relay LoRa device and uses it to achieve cooperative transmission by jointly considering a LoRa device's link quality ( $L_{N_i}$ ) in terms of packet success probability and remaining energy information  $E_{re\_N_i}(t)$ . Initially, the residual energy of each LoRa device is critical in the relay selection process to avert excessive use of a single relay device on the appropriate channel. Thus, the concept of network fairness is not jeopardized in this way. Here, we explain the four phases in detail, respectively.

**Phase 1:** To create a set of potential relay devices, we employed the harmonic mean criterion to assess the link quality [24]. According to two-hop channel information, each LoRa device estimates the integrated link quality  $L(N_i)$  in terms of the packet success probability ( $P_{suc}^i(loc(x))$ ) based on the distance between them, the SNR, and the receiver sensitivity of the device. Then,  $L(N_i)$  can be shown in Equation (23).

$$L(N_i) = 2(P_{suc}^{(S,N_i)}(loc(x)) + P_{suc}^{(N_i,GW)}(loc(x)))^{-1} \quad (23)$$

For all LoRa devices that pass the first criteria, a candidate set of relay devices is constituted for further scrutiny in the succeeding phases in a bid to select the best relay LD.

**Phase 2:** The GW evaluates the residual energy  $E_{re\_N_i}(t)$  of each candidate relay device  $N_i$  at a given time ( $t$ ) since its deployment in an IoT LoRa network. Assuming the total energy consumption  $E_{total\_N_i}(t)$  used by a LoRa device  $N_i$  during inactive and active modes at a particular time ( $t$ ),  $E_0$  is the initial energy of the LD. Then, the residual energy  $E_{re\_N_i}(t)$  of the  $i$ th relay device is represented in Equation (24).

$$E_{re\_N_i}(t) = E_0 - E_{total\_N_i}(t) \quad (24)$$

where  $E_0$  is the initial energy of the LoRa device at the initial time ( $t_0$ ),  $E_{total\_N_i}(t)$  is the energy dissipated by the LoRa device during inactive mode at time ( $t$ ) and active mode.

**Phase 3:** We used a positive weight factor ( $0 \leq \alpha \leq 1$ ) that indicated the preference between the integrated link quality  $\alpha$  and the residual energy weight  $(1 - \alpha)$  for all  $N_i \in N$  LoRa devices. Equation (25) indicates the combined relay parameters  $W(N_i)$ .

$$W(N_i) = \alpha(E_{re\_N_i}(t)) + (1 - \alpha)L(N_i) \quad (25)$$

$L(N_i)$  is the integrated link quality in terms of the LoRa device's packet success probability,  $E_{th}$  is the threshold energy considered to be the average residual energy and  $L_{th}$

represents the threshold link quality. Given, the following conditions,  $E_{th} \leq E_{re\_N_i}(t) \leq 1$  and  $L_{th} \leq L(N_i) \leq 1$  hold true.

**Phase 4:** By the ranking method, when a candidate LoRa device  $N_i$  has the highest rank as  $R_{N_i}^*$  among all the candidate LDs by evaluating with the formula depicted as the Equation (26), it becomes the best relay device.

$$R_{N_i}^* = \arg \max_{i \in N} W(N_i) \quad (26)$$

where  $R_{N_i}^*$  means the candidate  $N_i$  LoRa device with the highest rank that becomes the best relay device to transmit data from the source LD to the GW. Since the relay selection process is iterative, the rest of the devices back off and wait for the next transmission round.

## 5. Performance Evaluation

In this section, we evaluate the performance of the proposed SFPCR scheme through simulations. We compare the SFPCR scheme with three existing schemes: the FSRC scheme [33], the cluster-based scheme [21], and the mesh scheme [37] for performance evaluation. To do this, we first describe our simulation environments and scenarios and next present their performance evaluation metrics. Last, we explain the performance comparison of the FSRC and the three existing schemes by simulation results.

### 5.1. Simulation Environments and Scenario

To validate the SFPCR scheme's performance, an NS-3-based simulation environment was used to model multihop communications in the LoRa network. The performance of the FSRC scheme, the cluster-based scheme, and the mesh scheme were compared to the SFPCR scheme based on the relay installation and selection procedure. The NS-3 simulation tool is suitable for large networks anchored on discrete event simulation models of a system that changes to its state occurrence at discrete points in the simulation time [53].

Except for the GW, all the devices had the same communication needs. Each LoRa device sent out data packets on a regular basis, and transmissions only happened during certain time slots to avoid collisions. LoRa devices always kept their duty cycle below 1%, which is the minimum amount allowed by regulations [4]. Different spreading factors were spread out evenly based on how far away the devices were from the GW.

Table 2 lists the simulation parameters used in our model. We considered a circular network region with a cell radius of  $R = 6$  km and a GW in its centre to handle communication. The LoRa devices' distance from the GW was based on a Gaussian distribution. The parameters bandwidth, spreading factor, transmission power, and code rate were assigned to each device. In addition, each device transmitted an average of 20 bytes every packet. The estimated transmit power could be set from 2 dBm to 14 dBm, the channel bandwidth was  $125 \leq BW \leq 250$  kHz, and the channel code rate was 4/5. It is hard to say how many data the GW can handle, but it is safe to presume that more data provided more often will limit how many devices it can interface with. We compared the performance of two existing schemes to see which one was better. We also investigated the differences between them to see if our SFPCR scheme would be better based on the performance metrics.

**Table 2.** Simulation parameters.

Parameter	Values
Cell radius (R)	6 km
Number of LDs (N)	100–1000
Spreading factor (SF)	7–12
Bandwidth	125,250 (kHz)
Payload length	10–250 (Bytes)
Transmission power	14 dBm
Data rate	0.25–5.47 (kbps)

Table 2. Cont.

Parameter	Values
Coding Rate	4/5
Simulation time	3600 s
Payload CRC	ON
Duty cycle regulation	1 %
Channel frequency	868 (MHz)
Header (H)	0
Preamble symbol	8
Packet interval rate	10–15 (min)
Low data rate optimization (DE)	1
Path loss exponent ( $\gamma$ )	4

### 5.2. Performance Evaluation Metrics

The metrics used for the performance evaluation of the proposed SFPCR scheme and the existing schemes [21,33,37] were: packet success probability, energy consumption, and the number of packets. In simple terms:

- **Packet success probability**  
A frame transmission is considered successful when no collisions occur and all the bits in the frame are accurately decoded despite interference [54]. The harsh environment and capture effects are two factors that can have an impact on the success of data transmissions between LDs and GWs. It is critical to consider both the likelihood of the GW receiving an uplink from the source LD and the likelihood of the device successfully receiving a downlink.
- **Energy consumption**  
We considered the energy consumption among LoRa devices, which refers to the number of delivered data bits per unit of energy consumed by a LoRa device [39]. The SFPCR scheme aimed to balance the HSFZ's energy efficiency by shortening the ToA of the LDs.
- **Number of Packets**  
When analysing traffic behaviour in a LoRa network, the number of packets passing through the network to the GW is critical. In the LoRaWAN network, packets are sent infrequently and are affected by a variety of factors. As a result, the primary objective here was to ensure an effective transmission that limited duplicate packets, collisions, and retransmissions, all of which produce large quantities of network congestion and thus impede overall network performance.

### 5.3. Simulation Results

In this section, we present the IoT LoRa performance evaluation using the metrics from Table 2. We compared the proposed SFPCR scheme to the FSRC scheme, also called the relay scheme, the mesh scheme, and the cluster-Based scheme. We focused on three factors: transmission probability, packets sent by a given number of LoRa nodes, and energy consumption. The number of nodes, payload size, and gateway distance were compared. This study validated the suggested technique for multihop LoRa networks by using relay and clustering approaches in corresponding spreading factor zones.

Figures 5–7 illustrate how the packet success probability of LoRa devices varied with the number of nodes, payload size (B), and node distance from the GW (m). All the schemes exhibited an inverse proportionality between the success probability and the variables of number of nodes, payload size, and node distance from the GW. Figure 5 indicates that increasing the number of nodes decreased the chance of packets successfully reaching the GW for the FSRC, cluster-based, and mesh schemes. This was due to the capture effect as nodes competed for the channel and collisions arising from blind transmissions as these nodes transmitted data packets as soon as they were ready to transmit [55], which hindered the successful delivery of packets to the GW. In contrast, the proposed SFPCR scheme

employed collision-free transmission [45]. Due to this regulated access to the medium, each LoRa device transmitted its packets directly to the selected recipient. The SFPCR scheme deployed, among other things, density-based subspace clusters in the HSFZ and relay nodes in the LSFZ in order to boost and improve the performance of remote nodes so that data packets could be successfully sent to the GW with no transmission errors. Consequently, the SFPCR scheme outperformed the other schemes with the highest number of LDs by approximately 47.6% packet success probability, followed by the FSRC scheme at 24.7%, as shown in Table 3. However, as the number of LDs increased, the performance degraded because an increase in the number of nodes caused an automatic increase in the probability of packet error and interference [56].

Figure 6 depicts an inverse proportionality between success probability and payload size (B) for the SFPCR scheme and other comparative schemes. Because the LoRa network is utilized for long-distance transmission, an increase in packet size has a detrimental impact on the network’s performance. Therefore, the smaller the size of a data packet, the greater the possibility of success. The success probability of the SFPCR scheme was roughly 65.7% with a default payload size of 20 bytes, as shown in Table 3. It provided a superior method for reducing the quantity of traffic delivered to the GW by using data aggregation for all nodes clustered in the HSFZ, in contrast to previous techniques.

Table 3. Simulation results.

Scheme	Success Probability (%)			
	FSRC	Cluster Based	Mesh	Proposed SFPCR
Number of nodes (1000)	24.7	20.1	23.3	47.6
Default payload size (20 bytes)	44.6	29.4	34.2	65.7
Distance from the GW (6000 m)	1.12	0.036	0.67	29.1

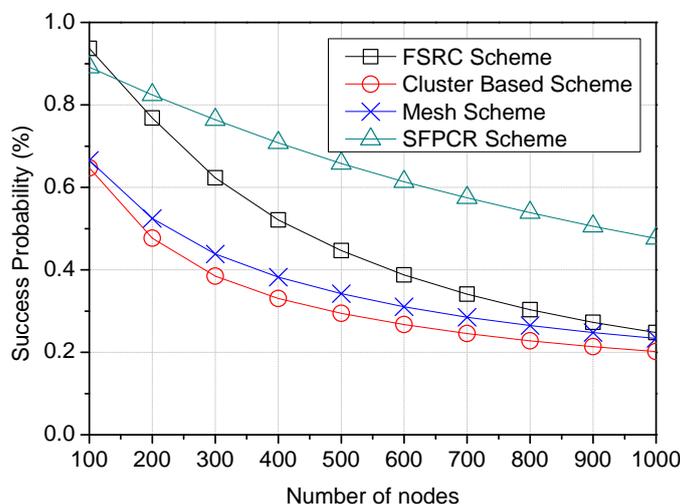
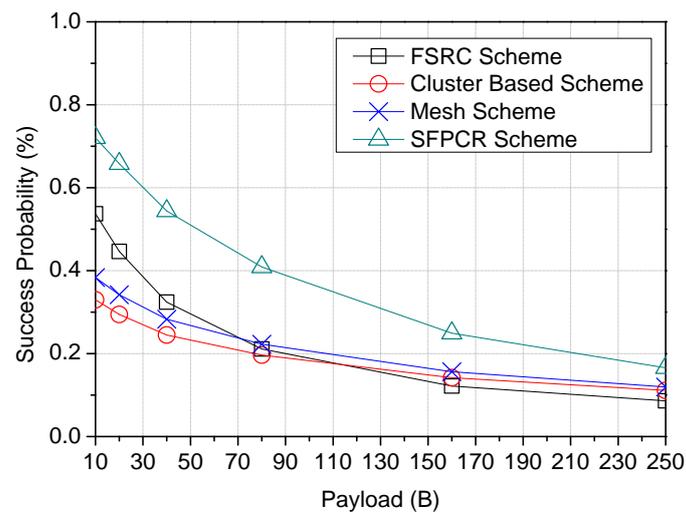
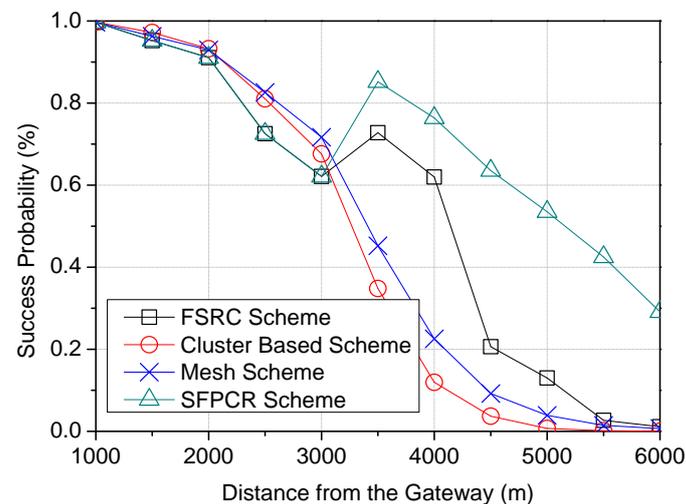


Figure 5. Plots variation in Success Probability according to the number of nodes in the network with a comparison between the SFPCR scheme and other protocols.



**Figure 6.** Plots variation in Success Probability according to the Payload Size (B) with a comparison between the SFPCR scheme and other protocols.

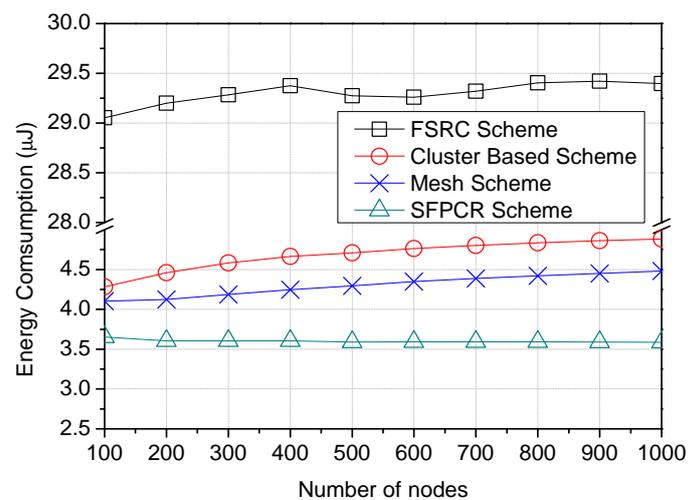


**Figure 7.** Plots variation in Success Probability according to the Distance of LoRa devices from the GW with a comparison between the SFPCR scheme and other protocols.

Figure 7 demonstrates a correlation between the distance of LoRa devices from the GW and their success probability. The success probability of all schemes diminished exponentially with rising average distances. However, at a distance between 3000 m and 3500 m, the SFPCR scheme and the FSRC scheme improved the success probability marginally due to the use of relays to extend packets from far LoRa devices to the GW. Moreover, the SFPCR scheme outperformed the other schemes in terms of success probability by more than 29.1% as the transmission distance from the GW increased. The usage of subspace clusters in the HSFZ and relays with success probabilities above the threshold in the LSFZ to transfer packets from far LDs to the GW mitigated the near–far fairness issue. The other schemes, on the other hand, demonstrated a negative decrease in success probability as the distance between the LDs and the GW increased due to poor link quality, noise, collision among the nodes, and other interference from obstacles such as buildings and vegetation encountered by nodes located far from the GW [55].

Figure 8 illustrates the variation in the number of nodes as a function of the energy consumption for the SFPCR scheme and the three other schemes under consideration. In contrast to the SFPCR scheme, the results for all three comparative schemes demonstrated a slightly direct correlation between energy use and number of nodes. The FSRC scheme utilized the most energy, exceeding 29  $\mu$ J, as a result of using more LDs with a higher SF

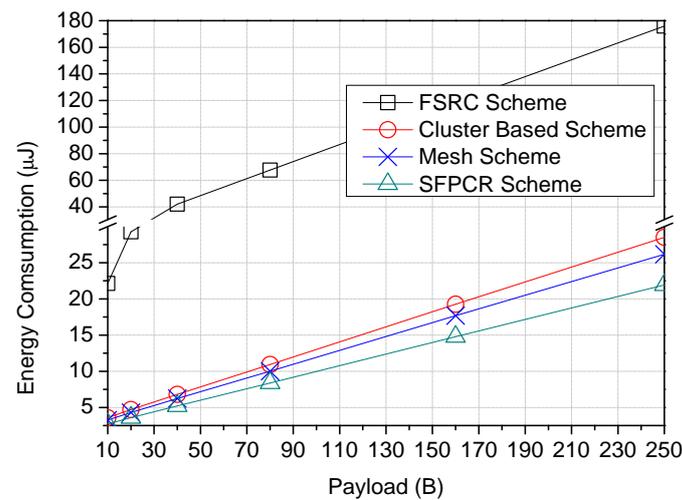
and broadcasting packets sent throughout the network, which introduced considerable routing overhead and used a great deal of energy. Similar to the cluster-based scheme, which had huge cluster sizes, the cluster member nodes had to connect with other nodes located at great distances, resulting in increased energy consumption owing to intracenter distance communication. In addition, the mesh scheme consumed 22.2% more energy than the SFPCR scheme because of the unnecessary multihops to send a signal strong enough to reach the GW. However, the SFPCR scheme consumed approximately 67.4% less energy than the comparative schemes because it primarily utilized short hops with a lower SF, controlled redundant transmission by data fusion from CHs in the HSFZ and also managed the number of transmissions and retransmissions to the GW by employing slotted ALOHA [45]. Consequently, the network's lifespan was increased.



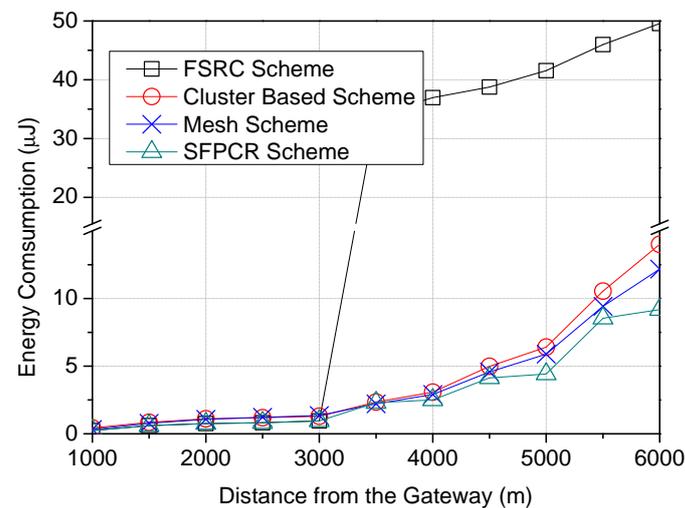
**Figure 8.** Plots variation of energy consumption ( $\mu\text{J}$ ) as a function of number of nodes with a comparison between the proposed SFPCR scheme and other protocols.

Figure 9 shows a direct proportionality between payload sizes and energy consumption for all schemes. As packet size increases, so does energy consumption, limiting the lifespan of an IoT network that is energy constrained. This is due to the broadcast storm problem caused by source nodes and relay nodes deployed in the network; in comparison, the cluster-based, mesh, and proposed SFPCR schemes consumed relatively little energy, limiting the use of distant nodes with higher SF to transmit directly to the GW, which consumes a great deal of energy. In addition, the SFPCR scheme transmitted small packets with short hops and a lower spreading factor in order to save transmission power. This implied that the source node in the HSFZ transmitted to the closest CH and relay node, thereby reducing the amount of time spent in the air. Consequently, the amount of energy lost was restricted.

In Figure 10, all schemes demonstrated a low quantity of energy consumption between the GW and approximately 3000 m, where direct transmission occurred; nevertheless, the FSRC scheme lost significantly more energy than the other schemes because of broadcasts made by relays and source nodes. The results demonstrated that the proposed SFPCR scheme consumed the least amount of energy as the distance increased, and the number of transmissions was regulated because all distant nodes in the HSFZ coordinated their transmission through the CHs rather than a direct transmission of data packets to the GW and because of the use of relay nodes in the LSFZ with a low SF to transmit packets of nodes with a success probability below the threshold.



**Figure 9.** Energy consumption as a function of payload size (B) with a comparison between the proposed SFPCR scheme and other protocols.



**Figure 10.** Energy consumption as a function of LoRa devices' distance from the GW (m) with a comparison between the proposed SFPCR scheme and other protocols.

As the number of nodes in the network increased, Figure 11 shows that the variation in the number of packets remained essentially constant across all the different schemes. Following the cluster-based and mesh schemes in terms of the number of packets used for varying numbers of nodes in the IoT sensor area comes the newly proposed SFPCR scheme, which used the fewest packets overall. The FSRC Scheme, on the other hand, had the most packets since it broadcast packets from the source nodes, relay nodes, and normal nodes. This caused a redundancy in the network as a result of the increased number of duplicate packets that were broadcast.

Similarly, Figure 12 demonstrates that the SFPCR scheme forwarded the fewest data packets to the GW. Furthermore, in the SFPCR scheme, the number of sent packets was more important than their size. Consequently, there was always a limit at which increasing the packet size would no longer increase performance and may instead decrease it. Cluster-based and mesh schemes had roughly the same number of packets as the payload size grew because their topologies were nearly identical. In contrast to the SFPCR scheme, the broadcast phase of the FSRC scheme resulted in a large number of duplicate packets in the network.

As depicted in Figure 13, the SFPCR scheme demonstrated a single packet transmission at near ranges from the GW to approximately 3000 m, beyond which the number of packets

increased modestly due to the employment of relay nodes in the LSFZ. In contrast, the CHs in the HSFZ aggregated data packets from all remote source nodes within the HSFZ, hence lowering network traffic flow. Similarly, the FSRC scheme from the GW to a distance of approximately 3000 m demonstrated a single packet transfer because no relay nodes were utilized in that zone. Beyond this stage, relay nodes that duplicated the sent packets were deployed. In the mesh and cluster-based schemes, each node obtained data from one node while passing data to the next node, resulting in an increase in packets. Consequently, as the distance from the GW rose, so did the number of node connections and hops.

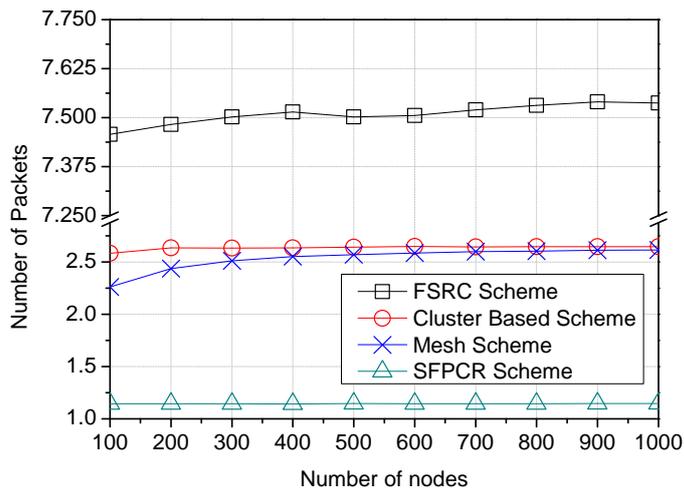


Figure 11. Variation in the number of packets according to the number of nodes in the network with a comparison between the proposed SFPCR scheme and other protocols.

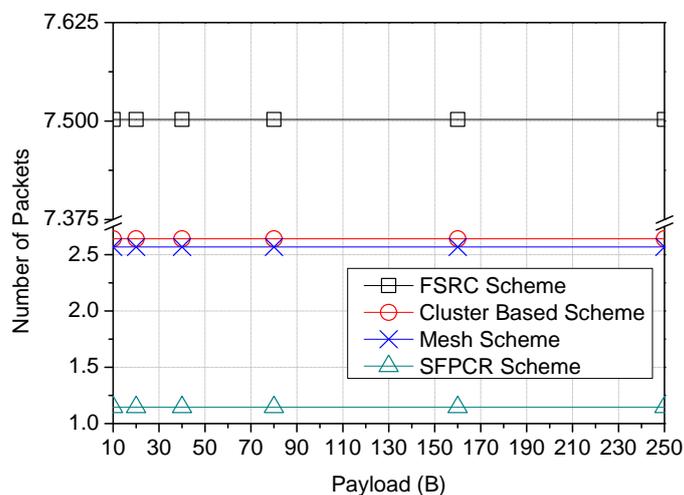
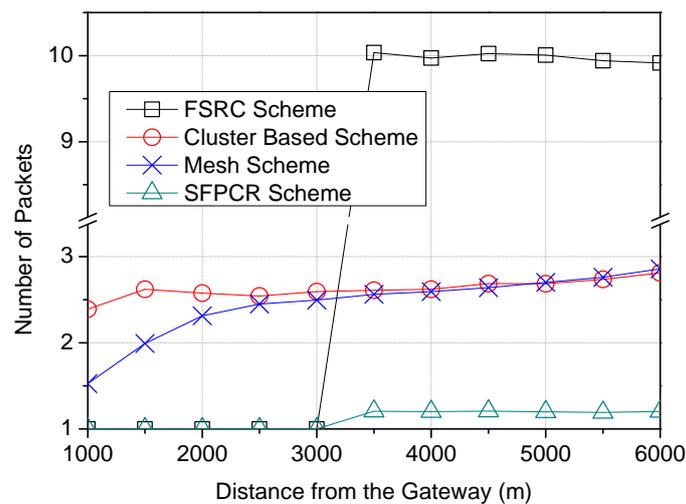


Figure 12. Variation in the number of packets according to the payload size (B) with a comparison between the proposed SFPCR scheme and other protocols.



**Figure 13.** Variation in the number of packets according to the the distance of LoRa devices from the GW with a comparison between the proposed SFPCR scheme and other protocols.

## 6. Conclusions and Discussion

In this paper, a novel SF-partition-based clustering and relaying scheme was proposed to solve the near–far fairness problem in IoT LoRa networks. The proposed SFPCR scheme utilized clustering and a multi-hop relay approach with the aid of low-SF LDs to increase the minimum success probability of distant nodes with a high SF. In the proposed scheme, the SF-based network partitioning method was used to determine the best partitioning threshold point for dividing SF zones into an LSFZ and an HSFZ to bridge the performance gap among SF zones by using an optimal SF division based on a heuristic algorithm. Then, by a combination of the DBSCAN algorithm and naive Bayes classifier, the density-based subspace clustering method constructed clusters of arbitrary shape for adjacent LDs in the HSFZ and selected CHs by a binary score representation. To extend the transmissions of CHs in the HSFZ by multihop communications, the relay LD selecting method chose the best relay LD in the LSFZ by using the harmonic mean based on the packet success probability and the remaining energy. The packet success probability of each LD in a certain SF zone was examined based on the hop distance of each hop to provide the optimal zone allocation with connection fairness. Through comprehensive simulations, we revealed that the SFPCR scheme exhibited the highest packet success probability of 29.1% for distant LDs at 6000 m, followed by the FSRC scheme at 1.12%, the mesh scheme at 0.67%, and lastly the cluster-based scheme at a negligible value of 0.036%. Using the default payload size of 20 bytes, the proposed SFPCR scheme also exhibited the highest success probability of 65.7%, followed by the FSRC scheme at 44.6%, the mesh scheme at 34.2%, and lastly the cluster-based scheme at 29.4%. The SFPCR conserved more energy for the LDs in order to prolong the lifetime of IoT LoRa networks in comparison to previous LoRa multihop schemes. Future work will include the development of intelligent clustering algorithms that can maintain up-to-date information on the link quality between LDs and their neighbours and react to sudden topological changes, as well as the responsiveness to changing network conditions through the use of lightweight reinforcement learning techniques.

**Author Contributions:** Conceptualization, D.M.; methodology, D.M.; software, Y.N.; validation, Y.N., H.C. and Y.S.; writing—original draft preparation, D.M.; writing—review and editing, Y.N. and E.L.; supervision, E.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2022-2020-0-01462) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Nomenclature

Symbol	Description
LD(i)	LoRa device i
CH	Cluster head
$N_{pr}$	Number of symbols in a packet preamble
$N_{pl}$	Number of symbols in a payload
$L_{packet}$	Size of the transmitted packet
$\rho_i$	Density of LDs in a particular SF
$\lambda$	Density of LoRa devices
$\Theta$	Virtual partition threshold point
$R_b$	Bit rate
$A^{(i)}$	LDs metric attribute
$T_{SF(i)}$	Time on air for single transmission of LD(i)
$T_{symbol}$	Time for transmitting a symbol
$I_{SF(i)}$	Interferers in SF region
$P_{opt}$	Optimal partition point
$P_{suc}^i(loc(x))$	Success probability
$P_{avg}(loc(x), \Theta_{th})$	Average success probability
$L_{Ni}$	Link quality
$E_{sleep\_Ni}(t)$	Energy dissipated in light-sleep mode
$E_s(i)$	Energy consumption
$E_{total\_Ni}(t)$	Total amount of consumed energy
$E_{reNi(t)}$	Residual energy of a LoRa node
$LD_{max}$	Maximum number of LoRa devices
$LD_{min}$	Minimum number of LoRa devices
$W(N_i)$	Combined relay parameters with weights
$R_{Ni}^*$	Optimal rank
$dist(n_i, n_j)$	Euclidean distance
$D_{ratio}(n_i, N_i)$	Minimum distance ratio between nodes
$P(C A^{(i)})$	Probability of LD(i) becoming the CH

## References

- Deese, A.S.; Jesson, J.; Brennan, T.; Hollain, S.; Stefanacci, P.; Driscoll, E.; Dick, C.; Garcia, K.; Mosher, R.; Rentsch, B.; et al. Long-Term Monitoring of Smart City Assets via Internet of Things and Low-Power Wide-Area Networks. *IEEE Internet Things J.* **2021**, *8*, 222–231. [CrossRef]
- Barnett, T.; Jain, S.; Andra, U.; Khurana, T. Cisco Visual Networking Index (VNI) Complete Forecast Update, 2017–2022. *Presentation* **2018**, 1–30. Available online: <https://s3.amazonaws.com/media.mediapost.com/uploads/CiscoForecast.pdf> (accessed on 5 September 2022).
- Mekki, K.; Bajic, E.; Chaxel, F.; Meyer, F. A comparative study of LPWAN technologies for large-scale IoT deployment. *ICT Express* **2019**, *5*, 1–7. [CrossRef]
- LoRa Alliance. *LoRaWAN 1.1 Specification*; LoRa Alliance: Beaverton, OR, USA, 2017; Volume 11.
- Reynders, B.; Meert, W.; Pollin, S. Range and coexistence analysis of long range unlicensed communication. In Proceedings of the 23rd International Conference on Telecommunications (ICT), Thessaloniki, Greece, 16–18 May 2016; pp. 1–6. [CrossRef]
- Georgiou, O.; Raza, U. Low Power Wide Area Network Analysis: Can LoRa Scale? *IEEE Wirel. Commun. Lett.* **2017**, *6*, 162–165. [CrossRef]
- Wang, X.; Kong, L.; He, L.; Chen, G. mLoRa: A Multi-Packet Reception Protocol in LoRa networks. In Proceedings of the 2019 IEEE 27th International Conference on Network Protocols (ICNP), Chicago, IL, USA, 8–10 October 2019; pp. 1–11. [CrossRef]
- Centelles, R.P.; Freitag, F.; Meseguer, R.; Navarro, L. Beyond the Star of Stars: An Introduction to Multihop and Mesh for LoRa and LoRaWAN. *IEEE Pervasive Comput.* **2021**, *20*, 63–72. [CrossRef]
- Yuan, L.; Wu, T.; Zhou, A. Performance Research of ALOHA Protocol. In Proceedings of the 7th International Conference on Cyber Security and Information Engineering, Brisbane, Australia, 3–25 September 2022.
- Adelantado, F.; Vilajosana, X.; Tuset-Peiro, P.; Martinez, B.; Melia-Segui, J.; Watteyne, T. Understanding the Limits of LoRaWAN. *IEEE Commun. Mag.* **2017**, *55*, 34–40. [CrossRef]

11. Phung, K.H.; Tran, H.; Nguyen, Q.; Huong, T.T.; Nguyen, T.L. Analysis and assessment of LoRaWAN. In Proceedings of the 2018 2nd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom), Ho Chi Minh, Vietnam, 29–31 January 2018.
12. Lyu, J.; Yu, D.; Fu, L. Achieving Max-Min Throughput in LoRa Networks. In Proceedings of the 2020 International Conference on Computing, Networking and Communications (ICNC), Big Island, HI, USA, 17–20 February 2020; pp. 471–476. [[CrossRef](#)]
13. Sisinni, E.; Ferrari, P.; Carvalho, D.F.; Rinaldi, S.; Marco, P.; Flammini, A.; Depari, A. LoRaWAN Range Extender for Industrial IoT. *IEEE Trans. Ind. Inform.* **2020**, *16*, 5607–5616. [[CrossRef](#)]
14. Barrachina-Muñoz, S.; Bellalta, B.; Adame, T.; Bel, A. Multi-hop communication in the uplink for LPWANs. *Comput. Netw.* **2017**, *123*, 153–168. [[CrossRef](#)]
15. Borkotoky, S.S.; Schilcher, U.; Bettstetter, C. Cooperative Relaying in LoRa Sensor Networks. In Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), Big Island, HI, USA, 9–13 December 2019; pp. 1–5. [[CrossRef](#)]
16. Augustin, A.; Yi, J.; Clausen, T.; Townsley, W.M. A study of LoRa: Long range & low power networks for the internet of things. *Sensors* **2016**, *16*, 1466.
17. Haxhibeqiri, J.; De Poorter, E.; Moerman, I.; Hoebeke, J. A survey of LoRaWAN for IoT: From technology to application. *Sensors* **2018**, *18*, 3995. [[CrossRef](#)]
18. Enriko, I.K.A.; Abidin, A.Z.; Noor, A.S. Design and Implementation of LoRaWAN-Based Smart Meter System for Rural Electrification. In Proceedings of the 2021 International Conference on Green Energy, Computing and Sustainable Technology (GECOST), Miri, Malaysia, 7–9 July 2021.
19. Aslam, M.S.; Khan, A.; Atif, A.; Hassan, S.A.; Mahmood, A.; Qureshi, H.K.; Gidlund, M. Exploring Multi-Hop LoRa for Green Smart Cities. *IEEE Netw.* **2020**, *34*, 225–231. [[CrossRef](#)]
20. Fargas, B.C.; Petersen, M.N. GPS-free geolocation using LoRa in low-power WANs. In Proceedings of the 2017 Global Internet of Things Summit (Giots), Geneva, Switzerland, 6–9 June 2017.
21. Zhu, G.; Liao, C.-H.; Sakdejayont, T.; Lai, I.-W.; Narusue, Y.; Morikawa, H. Improving the Capacity of a Mesh LoRa Network by Spreading-Factor-Based Network Clustering. *IEEE Access* **2019**, *7*, 21584–21596. [[CrossRef](#)]
22. Farhad, A.; Kim, D.-H.; Pyun, J.-Y. Scalability of LoRaWAN in an Urban Environment: A Simulation Study. In Proceedings of the 2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN), Zagreb, Croatia, 2–5 July 2019; pp. 677–681. [[CrossRef](#)]
23. Farhad, A.; Kim, D.-H.; Sthapit, P.; Pyun, J.-Y. Interference-Aware Spreading Factor Assignment Scheme for the Massive LoRaWAN Network. In Proceedings of the 2019 International Conference on Electronics, Information, and Communication (ICEIC), Auckland, New Zealand, 22–25 January 2019; pp. 1–2. [[CrossRef](#)]
24. Mohammadi, S.; Farahani, G. Scalability Analysis of a LoRa Network Under Inter-SF and Co-SF Interference with Poisson Point Process Model. *J. Comput. Secur.* **2021**, *8*, 43–57.
25. Asad Ullah, M.; Iqbal, J.; Hoeller, A.; Souza, R.D.; Alves, H. K-means spreading factor allocation for large-scale LoRa networks. *Sensors* **2019**, *19*, 4723. [[CrossRef](#)] [[PubMed](#)]
26. Amichi, L.; Kaneko, M.; Rachkidy, N.E.; Guitton, A. Spreading Factor Allocation Strategy for LoRa Networks Under Imperfect Orthogonality. In Proceedings of the ICC 2019–2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–7. [[CrossRef](#)]
27. Reynders, B.; Meert, W.; Pollin, S. Power and spreading factor control in low power wide area networks. In Proceedings of the 2017 IEEE International Conference on Communications (ICC), Paris, France, 21–25 May 2017; pp. 1–6. [[CrossRef](#)]
28. Lim, J.-T.; Han, Y. Spreading Factor Allocation for Massive Connectivity in LoRa Systems. *IEEE Commun. Lett.* **2018**, *22*, 800–803. [[CrossRef](#)]
29. Gul, O.M.; Erkmen, A.M. Energy-efficient cluster-based data collection by a UAV with a limited-capacity battery in robotic wireless sensor networks. *Sensors* **2020**, *20*, 5865. [[CrossRef](#)]
30. Farooq, M.O. Clustering-Based Layering Approach for Uplink Multi-Hop Communication in LoRa Networks. *IEEE Netw. Lett.* **2020**, *2*, 132–135. [[CrossRef](#)]
31. Misic, V.B.; Misic, J.; Khan, M.S.I. Optimum Zoning in RF-Recharged Sensor Networks. In Proceedings of the 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall), Montreal, QC, Canada, 18–21 September 2016; pp. 1–6. [[CrossRef](#)]
32. Bor, M.; Vidler, J.E.; Roedig, U. LoRa for the Internet of Things. In Proceedings of the 2016 International Conference on Embedded Wireless Systems and Networks, Graz, Austria, 15–17 February 2016; pp. 361–366.
33. Lee, S.; Lee, J.; Park, H.-S.; Choi, J.K. A Novel Fair and Scalable Relay Control Scheme for Internet of Things in LoRa-Based Low-Power Wide-Area Networks. *IEEE Internet Things J.* **2021**, *8*, 5985–6001. [[CrossRef](#)]
34. Sisinni, E.; Carvalho, D.F.; Ferrari, P. Emergency Communication in IoT Scenarios by Means of a Transparent LoRaWAN Enhancement. *IEEE Internet Things J.* **2020**, *7*, 10684–10694. [[CrossRef](#)]
35. Mugerwa, D.; Nam, Y.; Cho, H.; Choi, H.; Shin, Y.; Lee, E. Implicit Multi-hop Communication Scheme based on Overhearing in IoT LoRa Networks. In Proceedings of the 2022 Thirteenth International Conference on Ubiquitous and Future Networks (ICUFN), Barcelona, Spain, 5–8 July 2022; pp. 190–195. [[CrossRef](#)]
36. Anedda, M.; Desogus, C.; Murrone, M.; Giusto, D.D.; Muntean, G.-M. An Energy-efficient Solution for Multi-Hop Communications in Low Power Wide Area Networks. In Proceedings of the 2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Valencia, Spain, 6–8 June 2018; pp. 1–5. [[CrossRef](#)]

37. Lee, H.-C.; Ke, K.-H. Monitoring of Large-Area IoT Sensors Using a LoRa Wireless Mesh Network System: Design and Evaluation. *IEEE Trans. Instrum. Meas.* **2018**, *67*, 2177–2187. [[CrossRef](#)]
38. Gao, W.; Du, W.; Zhao, Z.; Min, G.; Singhal, M. Towards Energy-Fairness in LoRa Networks. In Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), Dallas, TX, USA, 7–10 July 2019; pp. 788–798. [[CrossRef](#)]
39. Sheu, J.P.; Chang, G.Y.; Wu, S.H.; Chen, Y.T. Adaptive k-coverage contour evaluation and deployment in wireless sensor networks. *ACM Trans. Sens. Netw. (TOSN)* **2013**, *9*, 1–31. [[CrossRef](#)]
40. Nurgaliyev, M.; Saymbetov, A.; Yashchyshyn, Y.; Kuttybay, N.; Tukymbekov, D. Prediction of energy consumption for LoRa based wireless sensors network. *Wirel. Netw.* **2020**, *26*, 3507–3520. [[CrossRef](#)]
41. Philip, M.S.; Singh, P. Energy Consumption Evaluation of LoRa Sensor Nodes in Wireless Sensor Network. In Proceedings of the 2021 Advanced Communication Technologies and Signal Processing (ACTS), Virtual, 4–6 December 2021; pp. 1–4. [[CrossRef](#)]
42. Cheong, P.S.; Bergs, J.; Hawinkel, C.; Famaey, J. Comparison of LoRaWAN classes and their power consumption. In Proceedings of the 2017 IEEE Symposium on Communications and Vehicular Technology (SCVT), Leuven, Belgium, 14 November 2017; pp. 1–6. [[CrossRef](#)]
43. Ochoa, M.N.; Guizar, A.; Maman, M.; Duda, A. Evaluating LoRa energy efficiency for adaptive networks: From star to mesh topologies. In Proceedings of the 2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Rome, Italy, 9–11 October 2017; pp. 1–8. [[CrossRef](#)]
44. Zorbas, D.; Papadopoulos, G.Z.; Maille, P.; Montavont, N.; Douligeris, C. Improving LoRa Network Capacity Using Multiple Spreading Factor Configurations. In Proceedings of the 2018 25th International Conference on Telecommunications (ICT), Saint Malo, France, 26–28 June 2018; pp. 516–520. [[CrossRef](#)]
45. Zorbas, D.; O’Flynn, B. Autonomous Collision-Free Scheduling for LoRa-Based Industrial Internet of Things. In Proceedings of the 2019 IEEE 20th International Symposium on “A World of Wireless, Mobile and Multimedia Networks” (WoWMoM), Washington, DC, USA, 10–12 June 2019; pp. 1–5. [[CrossRef](#)]
46. Merabtine, N.; Djenouri, D.; Zegour, D.-E. Towards Energy Efficient Clustering in Wireless Sensor Networks: A Comprehensive Review. *IEEE Access* **2021**, *9*, 92688–92705. [[CrossRef](#)]
47. Li, S.-S. An Improved DBSCAN Algorithm Based on the Neighbor Similarity and Fast Nearest Neighbor Query. *IEEE Access* **2020**, *8*, 47468–47476. [[CrossRef](#)]
48. Wang, X.; Wu, H.; Miao, Y.; Zhu, H. A Hybrid Routing Protocol Based on Naïve Bayes and Improved Particle Swarm Optimization Algorithms. *Electronics* **2022**, *11*, 869. [[CrossRef](#)]
49. Hu, L.Y.; Huang, M.W.; Ke, S.W.; Tsai, C.F. The distance function effect on k-nearest neighbor classification for medical datasets. *SpringerPlus* **2016**, *5*, 1–9. [[CrossRef](#)] [[PubMed](#)]
50. Rushikesh, S. Using Densities to Detect Nested Clusters. Master’s Thesis, Itä-Suomen Yliopisto, Kuopio, Finland, 2020.
51. Rezaei, M. Clustering Validation. Ph.D. Thesis, Itä-Suomen Yliopisto, Kuopio, Finland, 2016.
52. Afsar, M.M.; Tayarani-N, M.H. Clustering in sensor networks: A literature survey. *J. Netw. Comput. Appl.* **2014**, *46*, 198–226. [[CrossRef](#)]
53. Available online: <https://www.nsnam.org/> (accessed on 30 June 2022).
54. Kim, Y.K.; Kim, S.Y. Success probability characterization of long-range in low-power wide area networks. *Sensors* **2020**, *20*, 6861. [[CrossRef](#)]
55. Li, J.C.; Gamage, A.; Tengourtius, A.W.; Li, M. Known and unknown facts of LoRa: Experiences from a large-scale measurement study. *ACM Trans. Sens. Netw. (TOSN)* **2019**, *15*, 1–35. [[CrossRef](#)]
56. Hassnawi, L.; Ahmad, R.B.; Yahya, A.; Aljunid, S.; Ali, Z. Performance analysis of workload effects over motorway wireless ad hoc camera networks. *Int. J. Adv. Comput. Technol. (IJACT)* **2013**, *5*, 117–125.