An Incentive Mechanism-Based Stackelberg Game for Scheduling of LoRa Spreading Factors

Preti Kumari, Hari Prabhat Gupta, *Member, IEEE*, and Tanima Dutta

***Abstract*—Wireless Local Area Networks (WLANs) are one of the most popular networks for the Internet-of-Things (IoT) applications. Among various WLAN technologies, the Long-Range WAN (LoRaWAN) has gained a high demand in recent years because of its low power consumption and long-range communication. However, the Long-Range (LoRa) network suffers from interference problem among LoRa Devices (LDs) that are connected to the LoRa gateway by using the same Spreading Factors (SFs). In this article, we propose a game theory-based approach for estimating the time duration of transmission of data on suitable SFs such that interference problem is reduced and network devices maximize their utilities. We next propose a scheduling algorithm that schedules the allocated time duration on the SFs such that the waiting time of the network can be minimized. We finally use the network simulator-3 for validating the propose work. Various experiments are performed which demonstrate the improvement in the network performance.**

***Index Terms*—Long-range, Nash equilibrium, scheduler.**

I. Introduction

INTERNET-OF-THINGS (IoT) services traditionally employ long-range communication protocols to ensure extended coverage and high transmission rates [1]. However, such protocols suffer from high power consumption [2]. As IoT devices are battery-powered, realizing the demand of services. The Long-Range (LoRa) standard is a common Low Power Wide Area Network (LPWAN) specification that has gained wide interest in IoT applications [3]. LoRaWAN uses LoRa communication protocol that is based on the Chirp Spread Spectrum technique [4]. LoRa supports different Spreading Factors (SFs) for low-power and long-range transmission in parallel.

Fig. 1 illustrates a scenario of smart home in urban area using a LoRaWAN network with different IoT sensors to collect sensory data. It provides home automation to the users for convenience and increases the resident satisfaction. Along with this, it also ensures efficient resource utilization and detect issues before they become catastrophic. For example, sensory data of people's thermostats help to detect the cause of power spikes, automatically switch off the lights whenever a user is not at home to save the electricity *etc.* This scenario consists of multiple apartments in the building. Multiple LoRa Devices (LDs) are installed in each apartment for collecting data from the sensors. These LDs broadcast data to the deployed LoRa Gateways (LGs) within the given range. However, multiple LGs in the network increase the possible combination of SFs for each LD and therefore increase the computational time complexities of the network. Such high complexities also increase the energy consumption of the network. To reduce such complexities, our proposed work considered one LG is bound to the LDs that fall within its range. If the number of LDs within the range of the LG increases, then there is a need to increase the LGs also to mitigate the problem of overload. In that case, each LG is bound to a different group of LDs and receive packets from those bounded LDs only. The other LGs are separately scheduled the other unbound LDs. The LG forwards the collected data from the bounded LDs to the Network Server (NS) and finally to the applications.

Applicants, such as an electricity grid wants thermostats sensory data to know the power spikes and to forecast the electricity demands, use applications running on the collected data. The collection of personal data, in the home, exposes residents to the privacy and security risks. For example, smart light and smart door lock devices can be used as a pathway to enter into the home network to steal the data and find out more about the home user such as when the home is vacant and safe to rob. Due to such of these reasons, apartment owner does not want to share the data collected from the sensors, i.e., End Users (EUs) willingly. In order to motivate EUs and to compensate their cost or risk, EUs usually be provided a reward as a monetary incentive.

LoRaWAN suffers from the interference problem when large number of LDs try to transmit the sensory data to the LG using SF. Interference problem can be reduced if each LD is allocated a time duration for its transmission and leaves the SF free for the remaining time during which it can be used by other LDs. Due to the interference problem, a LD needs to wait for accessing the allocated SF. This article addresses the problem: *how long can a LD be scheduled on a given SF so that the utility of the LD can be maximized and minimizes the waiting time of the network?* We propose an approach that allocates suitable SFs to the LDs and computes the optimal transmission time duration of sensory data on those SFs such that load among SFs is balanced. This time duration depends on the various factors such as the size of data transmission,

Manuscript received August 20, 2019; revised March 30, 2020, August 16, 2020, and September 1, 2020; accepted September 3, 2020. Date of publication September 29, 2020; date of current version December 9, 2020. The work of Hari Prabhat Gupta and Tanima Dutta are partially supported by SERB grants under award numbers ECR/2016/000406 and ECR/2017/002419, respectively. The associate editor coordinating the review of this article and approving it for publication was M. Tortonesi. *(Corresponding author: Preti Kumari.)*

The authors are with the Department of Computer Science and Engineering, Indian Institute of Technology (BHU) Varanasi, Varanasi 221005, India (e-mail: pretikri.rs.cse17@iitbhu.ac.in; hariprabhat.cse@iitbhu.ac.in; tanima.cse@iitbhu.ac.in).

Digital Object Identifier 10.1109/TNSM.2020.3027730

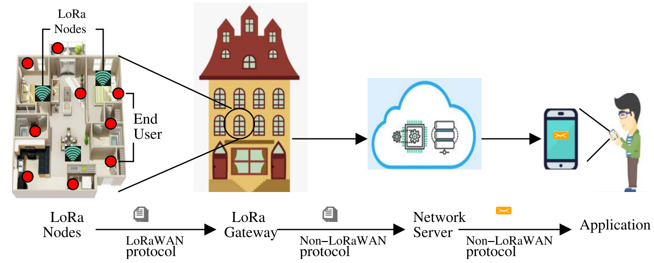


Fig. 1. A smart home scenario using LoRaWAN network.

processing cost, dissatisfaction cost (due to SF interference), and the price received from the applicants connected to the NS. LG gains the price from those applicants via NS for providing data. Further the LG pays some partial amount from the price gain to the LDs for forwarding the data. This is because, LDs consume their own resources for processing and forwarding data and they need to pay incentives to the EUs for collecting data. Such factors are modelled as monetary transfers in the proposed game. We minimise the waiting time of the LDs by scheduling them on the given SFs.

A. Related Work and Motivation

This section summarizes the state-of-the-art research and their limitations motivating the proposed work.

1) Interaction Among LDs: Authors in [5] designed a full MAC protocol enabling collision resolution. Authors in [6] proposed interference model for the LoRa modulation for different coding rates and SFs. Authors in [7] and [8], analyzed inter-SF and co-SF interference in the LoRa network. Based on the simulation results they also state that, for higher SFs, the effects of inter-SF interference is more noticeable. The main limitation in the existing work is that they do not consider interactions among LDs. Multiple access in wireless network using game theoretic approach is surveyed in papers [9], [10]. These survey shown that the game models are useful for designing distributed channel access mechanisms in wireless networks to achieve stable and efficient solutions.

2) Load Balancing: Though there exist several work on the performance analysis of LoRa network [8], [11]–[14], most of them do not consider load balancing while allocating the SFs for improving the network performance. Authors in [8] and [11] proposed SF allocation scheme in the presence of co-SF interference and inter-SF interference, respectively that can improve network performance. Authors in [12] formulated an optimization problem for maximizing the average packet transmission to propose a sub-optimal SF allocation. Optimal transmission policies and channel access control protocol are proposed in [13] and [14], respectively. Authors in paper [15]–[17], focuses on the game theoretic approach for resource allocation in wireless network. The main drawback of such schemes is that they did not consider load imbalance on SFs.

3) Scheduling of LDs: LoRa network faces the problem of scalability when they connect large number of LDs. Authors in [18] proposed a MAC layer protocol in which LG schedules LDs by dynamically specifying the allowed transmission

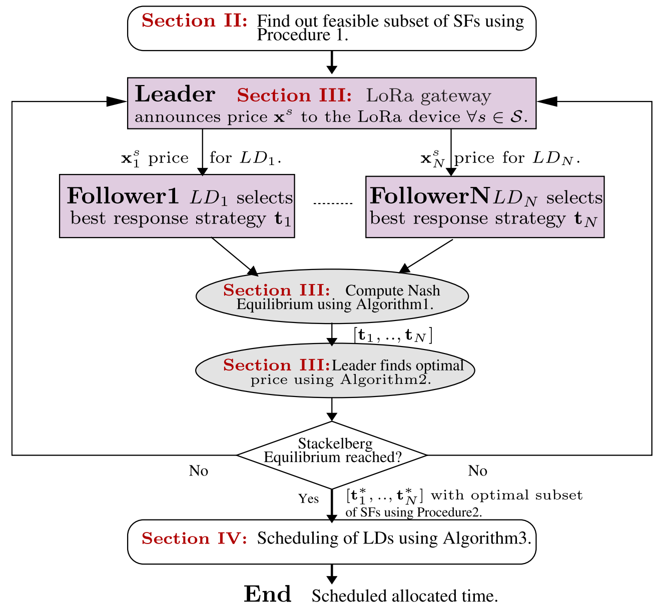


Fig. 2. Block diagram of the proposed approach.

powers and SFs. In paper [19] authors used matching game approach for allocating channels to the LDs such that they can maximize their utility. A nonconvex optimization problem for maximizing the network energy efficiency by considering user scheduling, SF assignment, and power allocation is presented in [20]. Authors in [21] proposed a synchronization and scheduling mechanism for LoRaWAN networks whereas a new communication-planning mechanism and a collision-free time-slotted scheduling approach developed in paper [22] and [23], respectively. The main limitation in existing work is that they did not consider the scheduling of allocated time duration on SFs to minimize the waiting time of LDs.

B. Major Contributions and Overview of the Solution

In this article, we propose a single-leader and multiple-followers Stackelberg game based approach to estimate the transmission time duration of LDs on feasible SFs as shown in Fig. 2. This computation is performed at LG because the LDs have limited processing and storage capacity. Once the LG received the request from LD, they computes the transmission time duration and schedule the LDs on the SFs. After computation, LG sends the control messages to the LDs for SF allocation along with the time slot of the LDs for data transmission. This article has following major contributions:

- We first propose a Hasse diagram based procedure for finding the subsets of the SFs for each LD, called as feasible subsets of SFs, on which they can successfully transfer the data to the LG.
- We next propose a Stackelberg game that models the interactions among LG (leader) and LDs (followers) which are connected by using feasible subsets of SFs. We propose an algorithm that chooses a feasible subset of SFs, called as optimal subset of SFs, which gives the maximum utility of the network and minimizes the network interference.

TABLE I
SYMBOLS USED IN THIS ARTICLE

Symbol	Description
$\mathcal{N}$	Set of LDs
N	Total number of LDs
n	Index of LDs
$\mathcal{S}$	Set of SFs
s	Index of SFs in $\mathcal{S}$
$\mathcal{C}_n$	Set of SFs of LD n
C_n	Max. number of SFs in $\mathcal{C}_n$
c	Index of SF in $\mathcal{C}_n$
t_n^c	Time duration of LD n
x_n^s	Price of LD n

- To reduce waiting time of the network, we finally present an algorithm for scheduling the transmission of LDs on the SF. Such LDs are allocated the same SF. We also used Network Simulator-3 [6] to demonstrate the impact of the propose approach on the performance of the network.

The remainder of this article is organized as follows: next section presents a procedure for finding the feasible subsets of SFs. Section III presents the game analysis and finds the optimal subsets of SFs. Section IV describes the scheduling of LDs. The simulation results are presented in Section V and this article is concluded in Section VI.

II. FINDING THE FEASIBLE SUBSETS OF SFS

We consider a network scenario which consists of multiple LDs, one LG, one NS, and an application, as shown in Fig. 1. We use Calligraphic letters such as $\mathcal{N}$, Italic capital letters such as *N*, and Italic small letters such as *n* to denote sets, numbers, and index, respectively. **Bold** notations are used for vector quantity throughout this article. Denote the set of LDs by $\mathcal{N} = \{1, 2, \dots, N\}$ and set of SFs by $\mathcal{S} = \{7, 8, \dots, 12\}$. Let a LD $n \in \mathcal{N}$ consists C_n feasible SFs. All possible feasible SFs of *n* is denoted by $\mathcal{C}_n = \{1, 2, \dots, C_n\}$. We assume that uniformly and randomly deployed LDs have constant stream of information to transmit at the NS via LG. Table I illustrates the symbols used in this article.

Definition 1 (Suitability of SF): The notation d_s is the maximum distance at which a LD can transmits the data to the LG by using the SF $s \in \mathcal{S}$. A SF *s* is suitable for a LD *n* for transmitting its data if *n* lies within the distance d_s.

Definition 2 (Partially Ordered Set): Partially ordered set (Poset) is an ordered pair of binary relation defined over a set $\mathcal{Z}$, (i.e., $\leq$, $\mathcal{Z}$) that satisfies following properties $\forall u, v, w \in \mathcal{Z}$:

- *Reflexivity:* i.e., $u \leq u$.
- *Antisymmetry:* i.e., if $u \leq v$ and $v \leq u \Rightarrow u = v$.
- *Transitivity:* i.e., if $u \leq v$ and $v \leq w \Rightarrow u \leq w$.

A Poset of all possible subsets can be visualized using Hasse diagram where each node corresponds to a subset of SFs for a given LD. In a Hasse diagram, we evaluate each node for its feasibility and assign a color accordingly, i.e., 'red' for feasible and 'white' otherwise. A feasible node in Hasse diagram for a LD $n \in \mathcal{N}$ consists not more than C_n suitable SFs.

Procedure 1 illustrates the steps for finding the feasible subsets of SFs for $\mathcal{N}$. The inputs of the procedure are the set of LDs, LG, and maximum number of SFs that can be used by each LD, i.e., C_n where $n \in \mathcal{N}$. Steps (2)-(4) draw the Hasse diagram for *n* where nodes consist the subsets of SFs. Steps (6)-(8) check the feasibility of each node *n*, i.e., all SFs lie in the node are suitable and total number of SFs are less than C_n. The algorithm changes the color of the node from white to red if it is feasible. Let z_n^i and SF_{nj}^i denote i^{th} feasible subset of LD *n* and j^{th} suitable SF of z_n^i subset, respectively. The outputs of Procedure 1 are the feasible subsets of SFs, denoted by $\mathbf{Z}$, where $\mathbf{Z} = [\mathbf{Z_1}, \dots, \mathbf{Z_n}, \dots, \mathbf{Z_N}]$, $\mathbf{Z_n} = [z_n^1, \dots, z_n^i, \dots, z_n^z]$, z_n^i consists $\{SF_{n1}^i, \dots, SF_{nk}^i\}$ SFs, and $k \leq C_n$.

Procedure 1: Finding the Feasible Subsets of SFs

Input: Set *N*, LG, and C_n for $n \in \mathcal{N}$.
Output: Feasible subsets $\mathbf{Z}$ of SFs for $\mathcal{N}$.

```
1 for each LD n ∈ 𝒩 do
2     Create Hasse diagram of Poset for n;
3     Assign 'white' color to all nodes of Hasse diagram;
4     Compute the possible SFs for n;
5     for each node do
6         if it is white color node then
7             if it consists not more than C_n suitable SFs
                then Assign 'red' color to node;
```

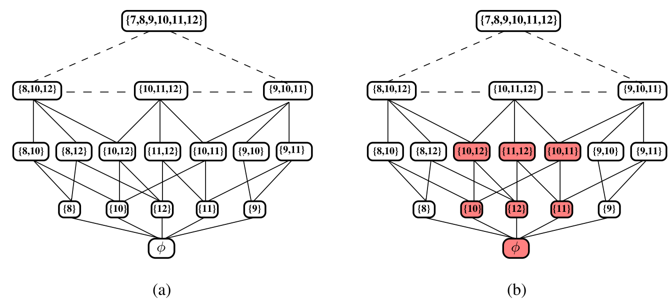


Fig. 3. Illustration of Hasse diagram of feasible subsets of SFs.

Example 1: Let consider a scenario where the suitable SFs for a LD $n \in \mathcal{N}$ are SF_{10}, SF_{11}, and SF_{12} because it lies outside the transmission range of SF_7, SF_8, and SF_9. In this scenario, if *n* wants to use maximum two out of three possible SFs ($C_n = 2$) then it can use any of the following feasible combination: SF_{10}, SF_{11}, SF_{12}, (SF_{10}, SF_{11}), (SF_{10}, SF_{12}), or (SF_{11}, SF_{12}). Such combinations are the power order set (Poset) of the possible three SFs. The Hasse diagram for all possible subsets of SFs is shown in part (a) of Fig. 3. Initially, all nodes are of 'white' color. Part (b) of Fig. 3 shows the red color nodes which consist feasible subsets of SFs for *n*.

III. IDENTIFY OPTIMAL SUBSET OF SFS FOR EACH LD

In this section, we propose a follower and leader model and analyze the Nash Equilibrium (NE) among LDs and Stackelberg Equilibrium (SE) between LG and LDs, respectively. We then, propose a procedure to identify the optimal subset of SFs for each LD out of the estimated feasible subsets of SFs by using the follower and leader model.

A. Analysis of Follower (LoRa Device) Game

This section derives the expression of the follower model for estimating the utility and formulates the optimization problem of followers by using the follower model. At the beginning of the game, LG allocates suitable SFs and price per data for each LD. Such prices are given based on the SFs and significance of the data. A LD uses the strategy of LG for selecting an optimal strategy. Such strategy maximizes the utility of the LD. Let t_n^c be the time duration of *n* on $c \in \mathcal{C}_n$. Similarly, the time vector $\mathbf{t}^c = [t_1^c, \dots, t_n^c, \dots, t_N^c]^T$. Let x_n^s be the price per unit time paid by *n* on $s \in \mathcal{S}$ to the LG. The price vector on *s* is given by $\mathbf{x}^s = [x_1^s, \dots, x_n^s, \dots, x_N^s]^T$. The utility function of a LD $n \in \mathcal{N}$ consists the following terms:

The first term is the cost paid to the end users. Let t_n^c denotes the time duration of data forwarding from the connected EUs to the LG via LD $n \in \mathcal{N}$ on SF $c \in \mathcal{C}_n$. The LD *n* needs to pay cost to the EUs for that data. The cost function due to the data transmission for $\mathbf{t}_n$ time duration on all available SFs in the combination $\mathcal{C}_n$ from EUs to LD *n* is denoted by $L_e(\mathbf{t}_n)$, is given as

$$L_e(\mathbf{t}_n) = \sum_{c=1}^{C_n} \left[-e_n^c \log\left(1 - \frac{t_n^c}{\delta}\right)\right], \tag{1}$$

where e_n^c is the pricing coefficient determined by LD *n* on SF *c* and δ is a parameter that is introduced to give cost values very close to the values obtained by a quadratic one. By using the Taylor expansion,

$$L_e(\mathbf{t}_n) = \sum_{c=1}^{C_n} \left[\frac{e_n^c t_n^c}{\delta} + \frac{e_n^c {t_n^c}^2}{2\delta^2}\right]. \tag{2}$$

Next, let $L_p(\mathbf{t}_n)$ be the processing cost for computing the received data from the EUs to LD *n* during $\mathbf{t}_n$ time duration on all available SFs in the combination $\mathcal{C}_n$. The processing cost depends on the hardware configuration of the LD *n* where the magnitude of the processing cost is negative. Let a_n and b_n are the hardware constants of LD *n*. Similar to the cost function modeled in paper [24], the processing cost function is defined as,

$$L_p(\mathbf{t}_n) = \sum_{c=1}^{C_n} \left[a_n \times (t_n^c)^2 + b_n \times t_n^c \right]. \tag{3}$$

Next term is the dissatisfaction cost of a LD. It occurs due to interference problem during data transmission from LDs to LG. Let h_n^c, p_n^c, σ^2, and *W* are the channel gain, transmission power, bandwidth, and power of white Gaussian noise of LD $n \in \mathcal{N}$ on SF $c \in \mathcal{C}_n$, respectively. Using Shannon channel capacity [25], the transmission rate (bits/second) of LD *n* on SF *c* is given by

$$R_n^c = W \log\left(1 + \frac{p_n^c h_n^c}{\sigma^2}\right). \tag{4}$$

Let the number of LDs which uses same or lower SF from LD *n* is denoted by d_n^c. Such d_n^c LDs create the interference problem to LD *n* because they also use the equal or higher transmission power during transmission of data to the LG. Such interference problem creates dissatisfaction to LD *n* during data transmission. The channel gain and transmission power of other d_n^c LDs are denoted by h_j and p_j, where $0 \le p_n^c \le p_j$ and $1 \le j \le d_n^c$, respectively. The transmission rate of LD *n* in the presence of other d_n^c LDs is given by

$$r_n^c = W \log\left(1 + \frac{p_n^c h_n^c}{\sum_{j=1}^{d_n^c} p_j h_j + \sigma^2}\right). \tag{5}$$

Dissatisfaction cost of LD *n* during t_n^c time duration is the difference of the transferred data with and without interference problem. This cost increases based on the deviance from the transmission rate without interference. Therefore, using Taguchi loss function, the dissatisfaction cost during $\mathbf{t}_n$ time duration can be defined as

$$L_s(\mathbf{t}_n) = \sum_{c=1}^{C_n} \left[\theta_n^c (t_n^c R_n^c - t_n^c r_n^c)^2 \right], \tag{6}$$

where θ_n^c is the loss function coefficient of LD *n* on SF *c*.

Last term is the price gained from LG. Price function of a LD *n* in Stackelberg game also depends on the strategy of other LDs. The Stackelberg Equilibrium (SE) estimates the total data from all the LDs which is known as forecasted data denoted by D. Each LD tries to forward more data to increase its incentives. We use the price function which reduces the incentive if a LD uses SF more than its allocated time, i.e.,

$$p(\mathbf{t}) = \lambda \left(1 + \frac{D - \sum_{i=1}^{N} t_i^c r_i^c}{\alpha}\right) + x_n^c, \tag{7}$$

where $\mathbf{t} = [t_1^c, t_2^c, \dots, t_N^c]^T$, $\forall c \in \mathcal{C}_n$, λ is the pricing parameter to give the elasticity in pricing, α is normalization parameter, and x_n^c is the base price assigned by LG to LD *n* on available SF *c*. The base price depends on the importance of transmitted data from LD *n* and the load on SF *c*. The price paid by LG to LD *n* for the data that is generated in t_n^c time duration is the product of the data transmitted by LD *n* and p(**t**), i.e.,

$$L_g(\mathbf{t}_n) = \sum_{c=1}^{C_n} [p(\mathbf{t}) t_n^c r_n^c], \; n \in \mathcal{N}. \tag{8}$$

1) Follower Game: Each LD works as a player and the follower game consists of *N* players. Let LD *n* forwards EUs data for t_n^c time duration by using SF *c*. Such time duration is the strategy of the LD *n*. The payoff of player *n* is the net utility, which is the price gained from LG minus the cost paid by the LD, (i.e., EUs cost, processing cost, and dissatisfaction cost). Let $\lambda' = \lambda + \frac{\lambda}{\alpha} D$, ${b'_n}^c = \frac{e_n^c}{\delta} + b_n$, and ${a'_n}^c = \frac{e_n^c}{2\delta^2} + a_n$. The payoff of player *n* from Eqs. (2), (3), (6), and (8) is given as

$$U_n^F(\mathbf{t}_n, \mathbf{t}_{-n}) = L_g(\mathbf{t}_n) - \big(L_e(\mathbf{t}_n) + L_p(\mathbf{t}_n) + L_s(\mathbf{t}_n)\big). \tag{9}$$

The game at the follower level is given as

Problem 1 (Follower Level Game):

$$\begin{aligned} \max_{\mathbf{t}_n} \quad & U_n^F(\mathbf{t}_n, \mathbf{t}_{-n}) \\ s.t. \quad & \sum_{c=1}^{C_n} t_n^c \le {t_n}^{max}, t_n^c \ge 0, \; n \in \mathcal{N}, c \in \mathcal{C}_n. \end{aligned} \tag{10}$$

Algorithm 1: Estimation of Best Response Strategy

Input: Precision threshold η, feasible subset of SFs;
Output: Best transmission time duration t_n^c of n;
Initialization: $\tau = 0$, $t_n^c[0] = \eta$;
do
  Increase τ by one;
  /* Use Eq. (11) to calculate $t_n^c[\tau+1]$. */
  $t_n^c[\tau+1] = Q_n^c[\tau] - \frac{1}{A_{nn}\sum_{c=1}^{C_n} A_{nn}}(\sum_{c=1}^{C_n} Q_n^c[\tau] - t_n^{max})$;
while $(\|t_n^c[\tau+1] - t_n^c[\tau]\| > \eta)$;

Problem 1 expresses that the LD n optimizes the strategy to maximize its utility and constraint imposes that the duration of forwarding the data from EUs to the LG must not be greater than its duty cycle, where $n \in \mathcal{N}$.

2) Nash Equilibrium (NE) Analysis: NE is a stable outcome of the follower game where all LDs interact through self-optimization and reach a point by calculating best response strategy where no LD has any incentive to deviate.

Theorem 1: Let $t_n^{c\star}$ denotes the best response strategy of a LD n on SF $c \in \mathcal{C}$. It is given as

$$t_n^{c\star} = Q_n^c - \frac{1}{A_{nn}\sum_{c=1}^{C_n}\left(\frac{1}{A_{nn}}\right)}\left(\sum_{c=1}^{C_n} Q_n^c - t_n^{max}\right),$$

where

$$Q_n^c = \frac{\lambda' r_n^c - {b'_n}^c + x_n^c r_n^c - \frac{\lambda}{\alpha} r_n^c \sum_{i=1, i\neq n}^{N} t_i^c r_i^c}{A_{nn}},$$
$$A_{nn} = 2{a'_n}^c + 2\frac{\lambda}{\alpha}{r_n^c}^2 + 2\theta_n^c(R_n^c - r_n^c)^2. \tag{11}$$

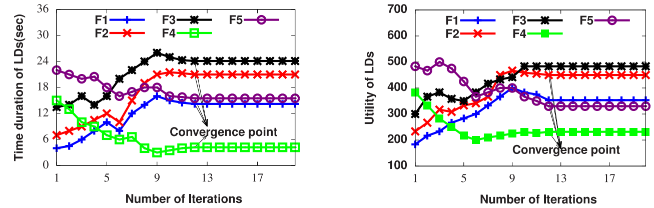
Proof: Appendix A gives the proof of the theorem. ■

Proposition 1: The best response strategy of a LD n, i.e., $t_n^{c\star}$, is unique.

Proof: Appendix B gives the proof of the proposition. ■

The estimated best response strategy of LDs is used to find out the NE among LDs as shown in Algorithm 1. Each LD selects its best response strategy and announces its transmission time duration on SF $c \in \mathcal{C}_n$ to the LG. This procedure repeats until the Algorithm 1 converge to a fixed point, i.e., a NE point similarly as in [26].

Example 2: Let we consider 10 LDs to obtain the convergence rate of the follower game which will increase as the number of followers increases. Initially, all followers set their time duration randomly but after each iteration, according to the strategies of other LDs, they re-calculate their time duration at which utility of all LDs maximized. Parts (a) and (b) of Fig. 4 illustrate that the followers (denoted as F1$\cdots$F5 among 10 LDs) find the stable value at iteration 13. Note that time duration of F1 is less than F5 in the stable state in part (a) of Fig. 4 whereas utility is high of F1 in part (b) of Fig. 4. This is because of the different pricing coefficients.



(a) (b)

Fig. 4. Avg. time duration (part(a)) and utility (part(b)) of LDs at convergence of Algorithm 1.

B. Analysis of Leader (LoRa Gateway) Game

This section derives the expression of the leader model for estimating the utility. The section also formulates and analyzes the optimization problem of leaders by using the leader model.

1) Leader Model: The leader makes the pricing strategy in such a way that the LDs use all SFs voluntary for forwarding the data and maximizes their own net utility. The utility function of the leader consists the price paid to the LDs and the revenue gain from applications. The LG pays prices to the LDs for the received data as given in Eq. (8). The LoRa network consists N LDs where LD n forwards data for $\mathbf{t}_n$ time duration with $\mathbf{r}_n$ transmission rate. The price paid to the LDs by the LG is given as

$$G_p(\mathbf{t}) = \sum_{n=1}^{N}\sum_{c=1}^{C_n}\left(\lambda\left(1 + \frac{D - \sum_{i=1}^{N} t_i^c r_i^c}{\alpha}\right) + x_n^c\right) t_n^c r_n^c. \tag{12}$$

The LG gains price from the applications which are connected to the NS for providing the data. The revenue generated from the data received by the LDs is given by

$$G_r(\mathbf{t}) = \sum_{n=1}^{N}\sum_{c=1}^{C_n}\left(g_{1n} t_n^c r_n^c - g_{2n}(t_n^c r_n^c)^2\right), \tag{13}$$

where g_{1n} and g_{2n} are the coefficients such that $g_{1n} >> g_{2n}$ which are used to measure the gain received from data on LG. The quadratic form of the utility function allows for tractable analysis and also serves as a good second-order approximation for the broader class of concave utility.

2) Leader Game: The LG works as the player in leader game. The strategy of the game is the price x_n^c paid to the LD n for forwarding the data using SF c. The payoff for a leader is the net utility which is the generated revenue minus price paid to LDs.

$$U^L(\mathbf{x}_n, \mathbf{x}_{-n}) = G_r(\mathbf{t}) - G_p(\mathbf{t}). \tag{14}$$

In SG, the leader aims to maximize its revenue. The game at leader level is given by

Problem 2:

$$\max_{\mathbf{x}_n} \ U^L(\mathbf{x}_n, \mathbf{x}_{-n})$$
$$s.t. \ \ 0 \le x_n^c \le X^{max}, \ n \in \mathcal{N}, c \in \mathcal{C}_n, \tag{15}$$

where X^{max} is the maximum price limit given by LG to the LD. LG optimizes strategy to maximize its utility and constraint imposes that price paid to the LD to be greater or equal to zero and not exceed its limit. Problem 2 can be rewritten as

Problem 3:

$$\max_{\mathbf{x}_n} \ U^L(x_n^c, t_n^c) = \sum_{n=1}^{N}\sum_{c=1}^{C_n}\Bigg(t_n^{c\star}\Bigg(g_{1n} - \lambda' + \frac{\lambda}{\alpha}\sum_{i=1}^{N} t_i^{c\star} r_i^c - x_n^c\Bigg) r_n^c - g_{2n}\big(t_n^{c\star} r_n^c\big)^2\Bigg).$$

$$s.t. \ \ 0 \le x_n^c \le X^{max}, \ \ n \in \mathcal{N}. \tag{16}$$

3) Stackelberg Equilibrium (SE) Analysis: The optimal strategy of the leader and the best response strategies of the followers, constitutes SE of the game. To find out the optimal strategy of LG, Proposition 2 proves the optimality of the best response strategies of LDs.

Proposition 2: The best response strategy of a LD $n \in \mathcal{N}$ is an optimal solution.

Proof: Appendix C gives the proof of proposition. ■

We propose the following theorem to prove that the obtained unique and optimal best response strategies of LDs is admitted by LG.

Theorem 2: The LG finds the optimal strategy by using the optimal strategies of LDs, i.e., estimated transmission time duration at NE state.

Proof: Appendix D gives the proof of theorem. ■

We use Karush-Kuhn-Tucker conditions [27] for finding the optimal solution of Problem 3. Therefore, the Lagrangian of Problem 3 is given as

$$\mathcal{L}^L(x_n^c, t_n^c) = \sum_{c=1}^{C_n} t_n^{c\star}\Bigg(g_{1n} - \lambda' + \frac{\lambda}{\alpha}\sum_{i=1}^{N} t_i^{c\star} r_i^c - x_n^c\Bigg) r_n^c - g_{2n}\big(t_n^{c\star} r_n^c\big)^2 + \Lambda_1 x_n^c - \Lambda_2(x_n^c - X^{max}), \tag{17}$$

where Λ_1 and Λ_2 are the Lagrangian multipliers of x_n^c. On equating the derivative of $\mathcal{L}^L(x_n^c, t_n^c)$ with respect to x_n^c with zero, the value of x_n^c is given as

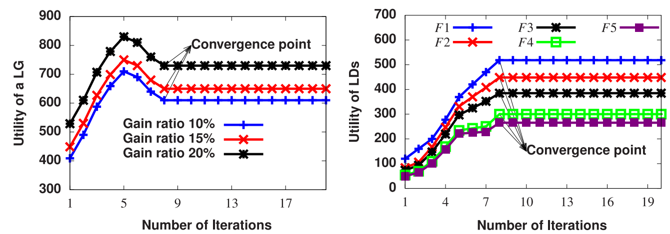
$$x_n^{c\star} = \frac{t_n^{c\star}}{\Delta t_n^c}\left(\frac{\lambda}{\alpha}\Delta t_n^c r_n^c - 1\right) - 2 g_{2n} t_n^c r_n^c + \left(g_{1n} - \lambda' + \frac{\lambda}{\alpha}\sum_{i=1}^{N} t_i^{c\star} r_i^c\right). \tag{18}$$

The LoRa network consists a SE when a LG estimates the optimal price for maximizing their utility while the LDs select their time duration for data forwarding to maximize their benefit. We propose an algorithm as explained in Algorithm 2, for finding the SE in LoRa network.

Example 3: The convergence rate of Algorithm 2 is shown in Fig. 5. Similar to Example 1, we consider 10 LDs to obtain the convergence rate of Algorithm 2. Gain ratio defined as the difference between g_{1n} and g_{2n} which increases as we decrease g_{2n}. Parts (a) and (b) of Fig. 5 illustrate that the leader and each follower iteratively update their utility based on the price of the leader, respectively and converges to a stable value at iteration 8, confirming the convergence and stability of the Algorithm 2. It can be seen from part (a) of Fig. 5 that utility of the LG increases with the increase in gain ratio and this increment increase exponentially due to the quadratic nature of second term in Eq. (14).

Algorithm 2: Finding the Stackelberg Equilibrium

/* Run at LoRa gateway for all connected $n \in \mathcal{N}$. */
Input: ϵ, ω, $\tau \leftarrow 0$, $x_n^c[0]$;
Output: $t_n^{c\star}$, $x_n^{c\star}$;
do
  $\tau \leftarrow \tau + 1$;
  /* Follower game: Each n maximizes its net utility */
  Estimate t_n^c by using Algorithm 1;
  /* Leader game: Use Eq. (18) for estimating $x_n^c[\tau+1]$ */
  $x_n^c[\tau+1] = x_n^c[\tau] + \epsilon \bigtriangledown_{U^L}(x_n^c[\tau])$;
while $(\|x_n^c[\tau+1] - x_n^c[\tau]\| < \omega x_n^c[\tau])$;



(a) (b)

Fig. 5. Utility of LG (part (a)) and LDs (part (b)) at the convergence of Algorithm 2.

C. Finding Optimal Subset of SFs

In this section, we propose a procedure to identify the optimal subset of SFs for each LD out of the estimated feasible subsets of SFs. Such optimal subset satisfies service requirement of LD and maximizes the revenue of the network. Procedure 2 uses the output of Procedure 1, i.e., feasible subsets of SFs $\mathbf{Z_n}$, where $n \in \mathcal{N}$. It illustrates the following steps to identify the optimal subset of SFs for each n from $\mathbf{Z_n}$. Steps (1) and (2) of Procedure 2 repeat for each LD n and each feasible subset for n, respectively, to estimate the best response strategy of n and optimal strategy of LG. To do this, Step (3) estimates t_n^i and x_n^i for z_n^i using Algorithm 2, where $z_n^i \in \mathbf{Z_n}$ subset of SFs of n consists $\{SF_{n1}^i, \ldots, SF_{nk}^i\}$ SFs. The procedure repeats this step for all feasible subsets of n and selects an optimal subsets of SFs for n which maximizes the utilities of n and LG. The outcomes of Procedure 2 can be arranged in a two-dimensional matrix $a[u][n]$ where a LD n allocates $a[u][n]$ time duration for using SF $u \in \mathcal{S}$ and LD $n \in \mathcal{N}$.

Example 4: Let continue Example 1 where LDs n, m, p, and q want to use maximum two possible SFs for satisfying its service requirement. We assume that all four LDs are lying in third ring and therefore Part (b) of Fig. 3 shows the feasible subsets of SFs for the LD n, m, p, and q. We use Procedure 2 and compute the optimal subset of LDs n, m, p, and q as shown in blue color node of parts (a)-(d), respectively, of Fig. 6. Table II illustrates the time duration of LDs on the

Procedure 2: Identify Optimal Subsets of SFs for LDs

```
Input: Feasible subsets Z = [Z_1, .., Z_n, ..., Z_N].
Output: Optimal subset of SFs for each n ∈ N.
for each LD n ∈ N do
    for each 'red' color node z_n^i ∈ Z_n do
        Estimate t_n^i and x_n^i for z_n^i using Algorithm 2;
    // An optimal subset of SFs for n
    Select t_n^i, x_n^i, and {SF_{n1}^i, ..., SF_{nk}^i} which
      maximize the utilities of n and LG;
```

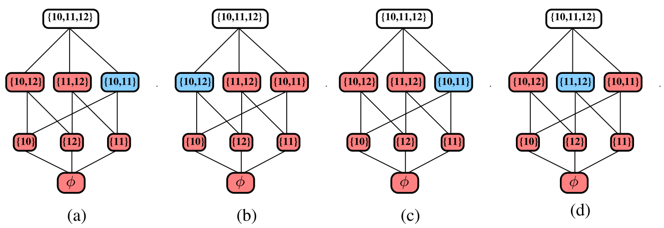


Fig. 6. Illustration of Hasse diagram of optimal subsets of SFs for *n*, *m*, *p*, and *q* shown in parts (a)-(d), respectively.

TABLE II
ILLUSTRATION OF OPTIMAL SUBSETS OF SFs FOR LDs

SF$\backslash\mathcal{N}$	n	m	p	q
SF_7	x	x	x	x
SF_8	x	x	x	x
SF_9	x	x	x	x
SF_{10}	3	2	7	0
SF_{11}	4	0	2	4
SF_{12}	0	2	0	1

optimal subset of SFs, where *0* and **x** denote non-selected and infeasible SFs, respectively. For instance, first column shows that *3* and *4* time durations are allocated to LD *n* for using SF_{10} and SF_{11}, respectively. Such time duration satisfies its service requirement and the network maximizes its revenue.

IV. SCHEDULING OF LDs TO MINIMIZE WAITING TIME

In a LoRa network, a given SF can be used by only one LD for a given time duration while the other LDs need to wait. Part (a) of Fig. 7 illustrates four LDs (*n*, *m*, *p*, *q*) with their time durations for using the allocated SFs. For example, LD *n* uses SF10 and SF11 for 3 and 4 time durations, respectively. Parts (b) and (c) show that the proper scheduling of the allocated time duration reduces the waiting time of the LDs which helps to increase the network performance. The waiting time of the LDs is defined in Definition 3.

Definition 3 (Waiting Time of a LD): Let a set of SFs $\mathcal{C}_n$ is allocated to a LD $n \in \mathcal{N}$. Waiting time is the time *n* has to wait its turn to get allocated SFs free for transmission. It is the difference between the time when *n* is ready for data transmission and the time it starts transmitting the data to the LG by using the allocated SFs.

The propose scheduling algorithm makes the order of using the SFs by the LDs. We use the greedy approach since the scheduling of the LDs is an NP-Hard problem and it is proven in Lemma 1. This greedy approach can be solved in polynomial time. Lemma 2 proves that Algorithm 3 can find out an

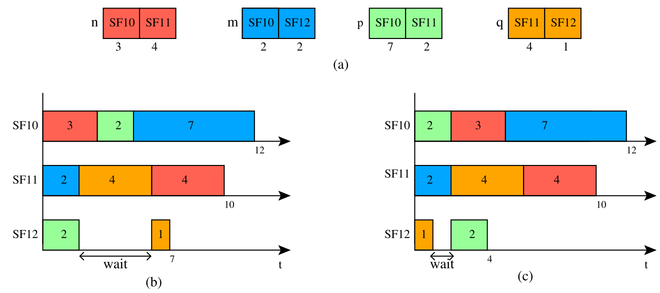


Fig. 7. Illustration of scheduling of SFs. Part (a) shows the time duration for accessing the SFs corresponding to the LDs. Parts (b) and (c) show the without and with scheduled SFs, respectively.

optimal scheduling order of data transmission if the LDs are selected greedily based on ascending allocated time duration. Algorithm 3 uses Procedure 1 and Procedure 2 to schedule the LDs. In other words, we want to minimize $\sum_{i=7}^{12}$ (waiting time on SF_i). Algorithm 3 illustrates the steps of the scheduler. The algorithm uses the outcomes of Procedure 2, i.e., matrix $a[u][v]$ where $v \in \mathcal{N}$ and $u \in \mathcal{S}$. Steps (1) and (2) of Algorithm 3 run for all SFs and LDs in the network. Step (3) searches the LDs with different SFs for parallel scheduling. Step (4) calls *Extract_Min* function for a SF $i \in \mathcal{S}$ that returns the LD which consists minimum allocated time. Step (4) also uses *Difference_Min* function for finding a LD $k \in \mathcal{N}$ which consists minimum difference between the allocated time of *i* and *k*. The output of Algorithm 3 arranged in a matrix $b[u][v][w]$ where a LD *v* allocates $b[u][v][w]$ time schedule for using *u* SF at the w^{th} position.

The time complexity of Algorithm 3 depends on two *for* loops where the loop at Step 2 depends on the number of LDs connected to LG (*i.e., N*), and loop at Step 3 contains two functions. Both functions compute *argmin* operation using min heap which takes $(NlogN)^2$ time. Since total number of LDs in the network, *i.e., M,* is much higher than N, therefore time complexity of Step 3 can be approximated as $O(M(logM)^2)$, where $M \approx N^2$. As SG formulated among *N* LDs, number of iterations required for convergence of Algorithm 1 and Algorithm 2 is in the order of constant because of the small value of *N*. Therefore the time complexity of the proposed scheduling of LoRa spreading factors is $O(NM(logM)^2)$.

Lemma 1: Scheduling of the LDs for accessing the SFs is NP-Hard problem.

Proof: Appendix E gives the proof of lemma. ■

Lemma 2: Algorithm 3 finds an optimal transmission order if the LDs are selected in order of ascending allocated time duration.

Proof: Appendix F gives the proof of lemma. ■

V. RESULTS AND DISCUSSION

In this section, we carry out the experimental evaluation to validate the performance of the proposed approach.

A. Simulation Setup

We use Network Simulator-3 (NS-3) [6] to validate the proposed approach. Each experiment repeats 100 times and

Algorithm 3: Scheduling of LDs on SFs

Input: t, allocated time duration matrix $a[][]$, and ζ.
Output: Scheduled allocated time duration matrix $b[][]$.

// Set $SF \leftarrow 12$

for *int* $i \leftarrow 7$ ***to*** $SF - 1$ **do**
  for *int* $n \leftarrow 1$ ***to*** N **do**
    for *int* $k \leftarrow i + 1$ ***to*** SF **do**
      $Difference\_Min(Extract\_Min(i), k)$;

Function *Extract_Min(i)*
  $u = \underset{j}{\operatorname{argmin}}\{a[i][j]\},\ 1 \leq j \leq N,$
  $b[i][l_i] = u,\ a[i][u] = \zeta,\ l_i = l_i + 1;$
  return

Function *Difference_Min(u, k)*
  $l = \underset{l}{\operatorname{argmin}}\{|u - a[k][l]|\},\ 1 \leq l \leq N,$
  $b[k][l][l_k] = l,\ a[k][l] = \zeta,\ l_k = l_k + 1$
  return

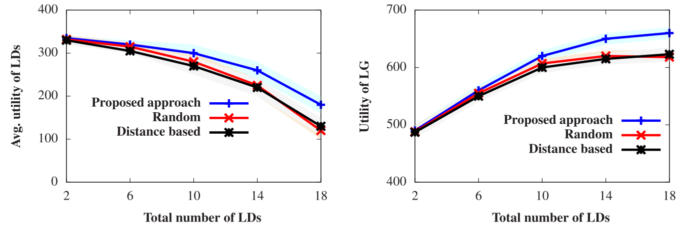
TABLE III
PARAMETER VALUES USED IN THE EXPERIMENTS

Parameter	Symbol	Value
Signal bandwidth	W	125 kHz
Channel gain	h_n^c	0.2
Pricing coefficient of LG	λ	10
Pricing coefficient of EU	e_n^c	[0-1]
Gain coefficient	$\{g_{1n}, g_{2n}\}$	$\{100, 0.1\}$
Loss function coefficient	θ_n^c	[0-1]

present the average. Most of the simulation parameters are considered from [8]. The network is set up by randomly deployed the LDs in a disc-shaped field. The radius of the field is 1.2 Kilometer for a smart home scenario in urban area. It means all the LDs present in the network within the range of LG can be connect to the LG. We considered 18 connected LDs for the simulation which are transmitting data at the same time. However there are more than 18 LDs present in the network. We used noise expression with variance $\sigma^2 = -174 + 10\log_{10} W$ to calculate additive noise present in the communication channel [19]. The other parameters are provided as in Table III. For the performance comparison, we adopted existing schemes such as random, distance based, and equal interval based SF allocation [7], [8], [11], [12].

B. Experimental Results

1) Utility Gain Perceived From the Proposed Model: We study the comparison of network performance based on the different SF allocation schemes. Utility is compared for SF allocation based on our proposed game model, random, and distance based schemes. Part (a) of Fig. 8 illustrates that average utility of the LDs decreases as the number of devices connected to the LG increases, irrespective of the used SF allocation scheme. It also shows that the average utility of LDs is higher when SF is allocated using our proposed game model as compared with other schemes. The results also show that the utility of LG initially fastly increases with increase in the number of LDs. This is because initially LG received more data with the increase of LDs but as the number of LDs



(a) (b)

Fig. 8. Average utility of LDs (part (a)) and LG (part (b)).

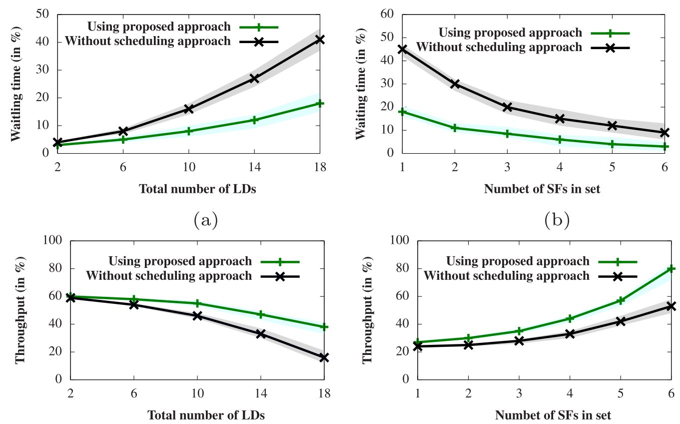
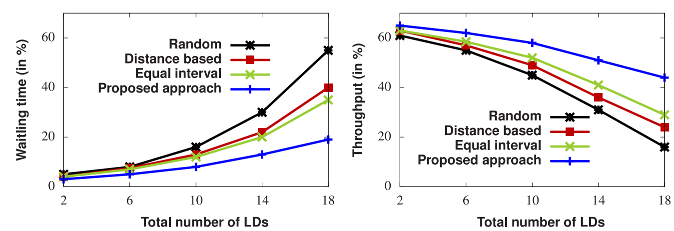


Fig. 9. Impact of SFs and LDs on waiting time and throughput.



(a) (b)

Fig. 10. Comparison of the proposed work with existing work.

become larger, all data can not be reached to the LG due to the interference in the network.

2) Effect of Scheduling of LDs on the Network: Throughput and waiting time are the main performance matrices in the network. Parts of Fig. 9 illustrate that when we use the proposed work, the throughput and waiting time of the network are increased and decreased, respectively. This is because, the proposed work schedules the LDs in a given order which overcomes the interference issue. It is also true that more LDs are also created more interference and therefore throughput goes done. We also compare the proposed scheduler with existing work. We use equal interval, distance based, and random SFs allocation methods [7], [8]. As name suggest, equal interval and random take equal and random time to access the SFs by the LDs, respectively. The SFs in distance based assignment uses the distance between LDs and the LG. Parts of Fig. 10 illustrate the effects of such methods on the performance of the network and the proposed approach outperforms the others.

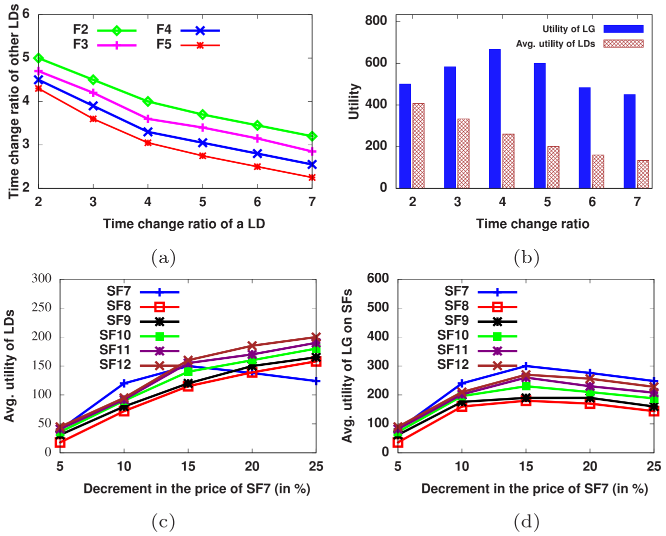


Fig. 11. Impact of time change of a LD on time duration of other LDs (part (a)) and utilities of LDs and LG (part (b)). Impact of the price change of a SF on utility of LD (part (c)) and LG (part (d)).

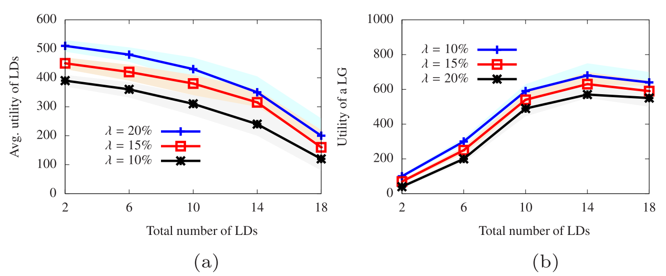


Fig. 12. Impact of number of LDs on utility of LDs (part (a)) and LG (part (b)) at different of pricing parameter λ.

TABLE IV
CONVERGENCE RATE OF ALGORITHM 1 AND ALGORITHM 2

		Algorithm 1			Algorithm 2		
		$\eta = 0.1$	$\eta = 0.4$	$\eta = 0.6$	$\omega = 2$	$\omega = 3$	$\omega = 4$
N	**2**	4	2	2	3	2	2
	6	8	5	3	5	4	3
	10	**13**	**8**	6	**9**	**7**	5
	14	24	17	12	14	12	9
	18	38	27	21	22	18	15

3) Analysis of Impact of Game Parameters: Next, we study the impact of the time duration and pricing strategy selected by each LD and LG, respectively, on the load on SFs. Part (a) and part (b) of Fig. 11 show the result of increases the time duration of a LD versus time duration of other LDs and utility, respectively. It shows that other LDs get less allocated time because all devices are handled by a single LG. The average utility of LDs also decreases as a LD increases time duration by deviating from the equilibrium stage. The utility of LG depends on the total data received and the pricing strategy. A LG starts decreasing its utility after some increment in time duration because with the increased time, interference problem also increases and more possibility to the data not reach at the LG. Part (c) and part (d) of Fig. 11 show the impact of the price of SF7 on the load on other SFs. As we decrease the price of SF7, LDs start to move on different SFs which reduces the interference on SF7 thus increase in the utility on SF7 and other SFs also. But as more LDs shifted to the other SFs, the utility of SF7 start decreasing. With the help of this result, one can find out the point at which decrement in the price of SF7 should stop so that utility on all SFs increases.

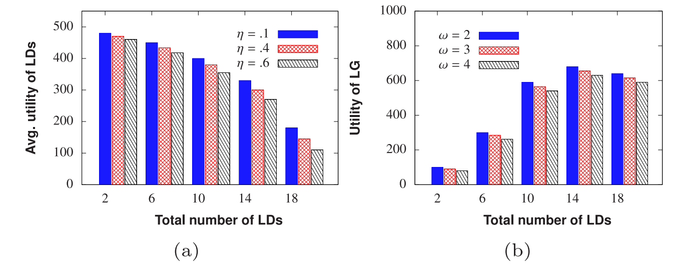


Fig. 13. Impact of the convergence parameters on the utility of LDs.

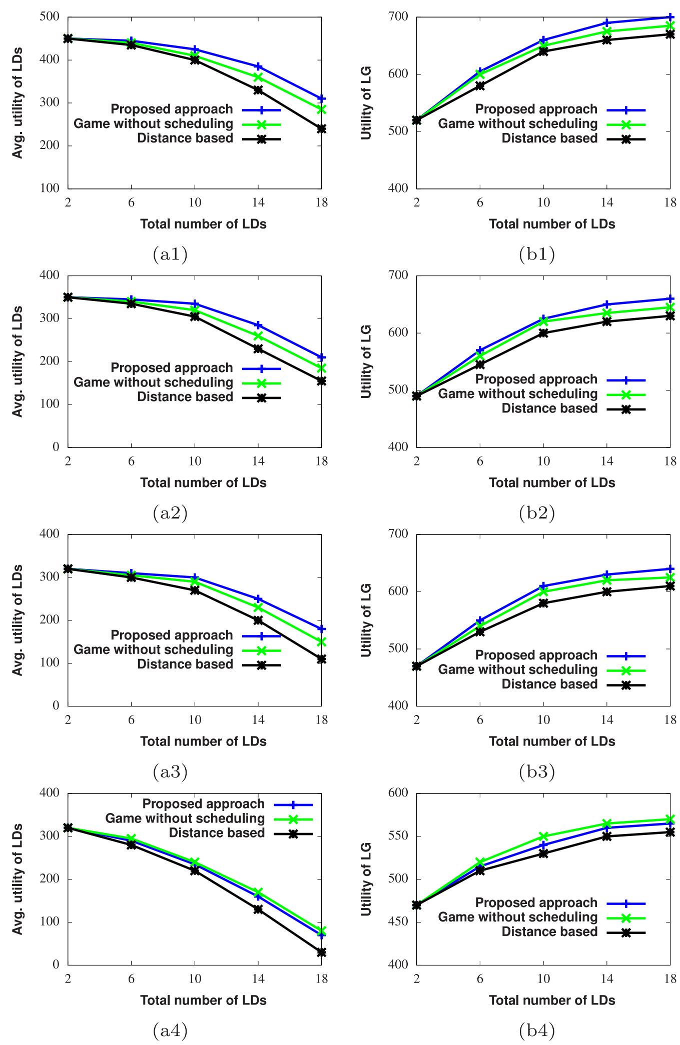


Fig. 14. Impact of the deployment distribution of LDs on the utilities of LDs and LG. Parts (a1)-(a4) and parts (b1)-(b4) show the results of the utilities of LDs and LG using different distributions, respectively.

Next, we present the impact of the total number of LDs on the proposed game equilibrium. Part (a) of Fig. 12 illustrates that as number of LDs connected to the LG increases, each device gets less time for data transmission on the allocated

SFs. Thus, the utilities of LDs decrease with more number of LDs and with the decrease in λ. Increase in the number of LDs connected to LG leads to the generation of more data and revenue, thus the utility of LG initially increases as shown in Part (b). It starts decreasing after a certain increase in LDs because of the interference problem among LDs.

We analyze the impact of precision threshold on the convergence rate of Algorithm 1 and Algorithm 2 and utilities of LDs and LG, as shown in Table IV and Fig. 13, respectively. For high precision threshold, algorithms converge in less number of iterations which reduces the time complexity of the algorithms. However, it provides less accuracy of the result. Table IV shows the impact of η and ω on the rate of convergence. As value of η increases from 0.1 to 0.4, the required iterations for 10 LDs decrease from 13 to 8. Similarly, iterations required for convergence reduce from 9 to 7 when ω increases from 2 to 3. Parts (a) and (b) of Fig. 13 illustrate the average utility of LDs decreases and utility of LG increases with the increase in the number of LDs.

4) Impact of the Distribution of LDs: To evaluate the performance of the proposed approach, we have performed experiments on uniform, poisson, random, and equal distributions. All LDs deployed using these four distributions and SF is allocated to a LD based on the distance from the LG. Fig. 14 shows that the proposed approach gives better utility than other schemes when deployment of LDs follow uniform, poisson, and random distributions. It also shows that utility gain in the game without scheduling scheme is larger than our proposed approach in an equal distribution. This is because scheduling takes extra time to compute minimum time duration among LDs but there is no need for scheduler in equal distribution because all LDs get same SF.

VI. Conclusion

This work proposed an approach to estimate the time duration of transmission of data on suitable SFs such that interference problem is reduced. We used Hasse diagram for finding the subsets of the feasible SFs for each LD. We next used a game theory-based approach for computing and scheduling the time duration of LDs on suitable SFs. It helps to maximizes the utility of the network and utility of the devices. The network also reached Nash equilibrium and Stackelberg equilibrium. We also validated the work and impact of various parameters on the performance of the network by using the simulator. As future work, we are planning to enhance the scheduling for multiple LGs with mobile LDs.

Appendix A
Proof of Theorem 1

Using Lagrangian multipliers $\gamma_{n,1}$ and $\gamma_{n,2}$ for constraints defined in Eq. (10), the $\mathcal{L}_n^F(\mathbf{t}_n, \mathbf{t}_{-n})$ is given by

$$\begin{aligned}\sum_{c=1}^{C_n}\Bigg[t_n^c\Bigg(\Bigg(\lambda' - \frac{\lambda}{\alpha}\sum_{i=1}^{N} t_i^c r_i^c + x_n^c\Bigg) r_n^c - {b'_n}^c\Bigg) - t_n^{c2} \\ \times \left(\theta_n^c (R_n^c - r_n^c)^2 + {a'_n}^c\right)\Bigg] + \gamma_{n,1} t_n^c \\ - \gamma_{n,2}\left(\sum_{c=1}^{C_n} t_n^c - t_n^{max}\right), \\ s.t.\ \gamma_{n,1} t_n^c, \gamma_{n,2}\left(\sum_{c=1}^{C_n} t_n^c - t_n^{max}\right) = 0, \\ \gamma_{n,1}, t_n^c \geq 0, \gamma_{n,2} > 0.\end{aligned} \tag{19}$$

The second order derivative of $\mathcal{L}_n^F(\mathbf{t}_n, \mathbf{t}_{-n})$ is $-2{a'_n}^c - 2\frac{\lambda}{\alpha} r_n^{c2} - 2\theta_n^c (R_n^c - r_n^c)^2$, which is negative. Therefore, Problem 1 is convex optimization problem and the follower level game has at least one NE [24]. The value of t_n^c can be obtained from the coefficient matrix by setting the first order derivative of $\mathcal{L}_n^F(\mathbf{t}_n, \mathbf{t}_{-n})$ equals to zero, i.e.,

$$t_n^{c\star} = \frac{1}{A_{nn}}\left({\beta_n}^c + x_n^c r_n^c - \frac{\lambda}{\alpha} r_n^c \sum_{i=1, i\neq n}^{N} t_i^c r_i^c\right), \tag{20}$$

where, $A_{nn} = 2{a'_n}^c + 2\frac{\lambda}{\alpha} r_n^{c2} + 2\theta_n^c (R_n^c - r_n^c)^2$, ${\beta_n}^c = \lambda' r_n^c - {b'_n}^c + \gamma_{n,1} - \gamma_{n,2}$. Let $FX = \sum_{c=1}^{C_n} A_{nn}$. By putting value of t_n^c into the constraint of Eq. (19), we get $\gamma_{n,2}$ equals to

$$\frac{1}{FX}\sum_{c=1}^{C_n}\left(\frac{\lambda' r_n^c - {b'_n}^c + x_n^c r_n^c - \frac{\lambda}{\alpha} r_n^c \sum_{i=1, i\neq n}^{N} t_i^c r_i^c}{A_{nn}} - t_n^{max}\right). \tag{21}$$

Using $\gamma_{n,2}$ in Eq. (20) from Eq. (21) and theorem proved.

Appendix B
Proof of Proposition 1

We check positivity, monotonicity, and scalability for the uniqueness of the best response strategy of a follower *n* defined in Eq. (11). Eq. (11) shows that $t_n^c > 0$ because $\sum_{c=1}^{C_n} Q_n^c - t_n^{max}$ will always return negative value and therefore confirms its positivity. Let t_{-n}^c and ${t'}_{-n}^c$ are the data transmission time duration of LD other than *n* on SF *c*. Eq. (11) shows that if $t_{-n}^c \geq {t'}_{-n}^c$ then $f(t_{-n}^c) \geq f({t'}_{-n}^c)$ and therefore t_n^c consists monotonicity. Using Eq. (11), we get

$$\mu f\left(t_{-n}^c\right) - f\left(\mu t_{-n}^c\right) \geq 0. \tag{22}$$

Hence, scalability proved.

Appendix C
Proof of Proposition 2

Since the objective function $U_n^F(\mathbf{t}_n, \mathbf{t}_{-n})$ is quadratic function given x_n^c and constraint is affine, therefore Problem 1 is convex optimization problem. Using Eq. (10), $t_n^{c\star} = \min\{t_n^c, t_n^{max}\}$. Let assume $t_n^c \geq t_n^{max}$ and hence $t_n^{c\star} = t_n^{max}$. The objective function of a LG, i.e., $U^L(\mathbf{t}_n, \mathbf{t}_{-n})$ is to maximize Problem 2 after substituting $t_n^{c\star}$. However, we can observe that $U^L(\mathbf{t}_n, \mathbf{t}_{-n})$ is a decreasing function with respect to x_n^c. For maximizing this function, we must have

$$x_n^c \leq \frac{\left(\tilde{Q}_n^c - \frac{1}{\sum_{c=1}^{C_n} \frac{1}{A_{nn}}}\left(\sum_{c=1}^{C_n} \frac{\tilde{Q}_n^c}{A_{nn}} - t_n^{max}\right) - A_{nn} t_n^{max}\right)}{r_n^c \left(\frac{1}{A_{nn}\sum_{c=1}^{C_n} \frac{1}{A_{nn}}} \sum_{j=1, j\neq c}^{C_n} \frac{x_n^j}{A_{nn}} - 1\right)}, \tag{23}$$

where $\tilde{Q}_n^c = \lambda' r_n^c - {b'}_n^c - \frac{\lambda}{\alpha} r_n^c \sum_{i=1, i\neq n}^{N} t_i^c r_i^c$. Thus $t_n^{c\star} = t_n^c$, which completes the proof.

Appendix D
Proof of Theorem 2

The partial derivative of $U^L(x_n^c, t_n^c)$ with respect to x_n^c and x_l^c, i.e., $d^2 U^L(x_n^c, t_n^c)/dx_n^c dx_l^c$, is given as

$$\begin{cases} 2\Delta t_n^c r_n^c \left(\frac{\lambda}{\alpha}\Delta t_n^c r_n^c - 1\right) - 2 g_{2n} \Delta t_n^{c2} r_n^{c2} & if \ \ n = l \\ 0 & else \end{cases} \tag{24}$$

where $\Delta t_n^c = \frac{r_n^c}{A_{nn}}(1 - \frac{1}{A_{nn}\sum_{c=1}^{C_n} \frac{1}{A_{nn}}})$. Since value of α is almost equal to forcasted demand of data therefore is much larger than λ and $(\frac{\lambda}{\alpha}\Delta t_n^c r_n^c - 1)$ returns negative value. Due to the positivity of Δt_n^c, negativity of $(\frac{\lambda}{\alpha}\Delta t_n^c r_n^c - 1)$, diagonal elements of Hessian matrix will give negative value. Therefore, Problem 3 is a standard convex maximization problem.

Appendix E
Proof of Lemma 1

Scheduling of the allocated time duration for using the SFs is NP-Hard problem which can be proven by reduction to the minimization Knapsack problem. Let w_n^s is the waiting time of LD n using SF s, then scheduling problem can be written as

$$\begin{aligned} \min \ & \sum_{n=1}^{N}\sum_{s=1}^{S} w_n^s y_n^s \\ s.t. \ & \sum_{s=1}^{S} t_n^s \leq {t_n}^{max}, \ t_n^s \geq 0, \ n \in \mathcal{N}, \ s \in \mathcal{S}, \end{aligned} \tag{25}$$

where $y_n^s = 1$ if LD n has allocated SF s, else $y_n^s = 0$. Constraints impose that time duration taken by a LD n on all feasible SFs must not exceed its duty cycle. This is the minimization knapsack problem, which is NP-hard, and thus scheduling problem is also NP-hard and hence proved.

Appendix F
Proof of Lemma 2

For simplicity, let consider one SF onto which multiple LDs want to transmit data. $P = p_1 p_2 \cdots p_y$ is the sequence of LDs who want to send data to the LG in order of increasing time duration. Let t_i is the time duration of LD p_i in the sequence, then total time in the system taken by SF is

$$T(P) = t_1 + (t_1 + t_2) + \cdots = \sum_{k=1}^{y}(y - k + 1)t_k.$$

Let assume that 2 LDs with sequence number $\tilde{a}$ and $\tilde{b}$ change their position where $\tilde{a} < \tilde{b}$. We denote this out of order sequence as P'. Then total time in the system taken by SF when it uses P' sequence, is $T(P') = (y - \tilde{a} + 1)t_{\tilde{b}} + (y - \tilde{b} + 1)t_{\tilde{a}} + \sum_{k=1, k\neq \tilde{a}, \tilde{b}}^{y}(y - k + 1)t_k$. The old sequence is preferable to the new because $T(P) - T(P') = (\tilde{b} - \tilde{a})(t_{\tilde{a}} - t_{\tilde{b}}) < 0$ and lemma proved.

References

[1] M. Amadeo, G. Ruggeri, C. Campolo, and A. Molinaro, "IoT services allocation at the edge via named data networking: From optimal bounds to practical design," *IEEE Trans. Netw. Service Manag.*, vol. 16, no. 2, pp. 661–674, Feb. 2019.

[2] H. Pirayesh, P. K. Sangdeh, and H. Zeng, "EE-IoT: An energy-efficient IoT communication scheme for WLANs," in *Proc. IEEE INFOCOM*, 2019, pp. 361–369.

[3] T. Elshabrawy and J. Robert, "Capacity planning of LoRa networks with joint noise-limited and interference-limited coverage considerations," *IEEE Sensors J.*, vol. 19, no. 11, pp. 4340–4348, Feb. 2019.

[4] R. M. Sandoval, A. Garcia-Sanchez, and J. Garcia-Haro, "Optimizing and updating lora communication parameters: A machine learning approach," *IEEE Trans. Netw. Service Manag.*, vol. 16, no. 3, pp. 884–895, Jul. 2019.

[5] N. El Rachkidy, A. Guitton, and M. Kaneko, "Collision resolution protocol for delay and energy efficient LoRa networks," *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 2, pp. 535–551, May 2019.

[6] F. Van den Abeele, J. Haxhibeqiri, I. Moerman, and J. Hoebeke, "Scalability analysis of large-scale LoRaWAN networks in ns-3," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2186–2198, Oct. 2017.

[7] A. Mahmood, E. Sisinni, L. Guntupalli, R. Rondín, S. A. Hassan, and M. Gidlund, "Scalability analysis of a LoRa network under imperfect orthogonality," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1425–1436, Mar. 2019.

[8] A. Waret, M. Kaneko, A. Guitton, and N. El Rachkidy, "LoRa throughput analysis with imperfect spreading factor orthogonality," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 408–411, Oct. 2019.

[9] K. Akkarajitsakul, E. Hossain, D. Niyato, and D. I. Kim, "Game theoretic approaches for multiple access in wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 372–395, 3rd Quart., 2011.

[10] J. Moura and D. Hutchison, "Game theory for multi-access edge computing: Survey, use cases, and future trends," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 260–288, 1st Quart., 2019.

[11] L. Amichi, M. Kaneko, E. H. Fukuda, N. El Rachkidy, and A. Guitton, "Joint allocation strategies of power and spreading factors with imperfect orthogonality in LoRa networks," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3750–3765, Apr. 2020.

[12] J. Lim and Y. Han, "Spreading factor allocation for massive connectivity in LoRa systems," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 800–803, Jan. 2018.

[13] R. M. Sandoval, A. Garcia-Sanchez, J. Garcia-Haro, and T. M. Chen, "Optimal policy derivation for transmission duty-cycle constrained LPWAN," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 3114–3125, May 2018.

[14] M. O. Farooq and D. Pesch, "A search into a suitable channel access control protocol for LoRa-based networks," in *Proc. IEEE LCN*, 2018, pp. 283–286.

[15] M. Ghazvini, N. Movahedinia, K. Jamshidi, and N. Moghim, "Game theory applications in CSMA methods," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1062–1087, 3rd Quart., 2013.

[16] L. Song, D. Niyato, Z. Han, and E. Hossain, "Game-theoretic resource allocation methods for device-to-device communication," *IEEE Wireless Commun.*, vol. 21, no. 3, pp. 136–144, Jun. 2014.

[17] M. Doudou, J. Barcelo-Ordinas, D. Djenouri, J. García-Vidal, A. Bouabdallah, and N. Badache, "Game theory framework for MAC parameter optimization in energy-delay constrained sensor networks," *ACM Trans. Sensor Netw.*, vol. 12, pp. 1–35, May 2016.

[18] B. Reynders, Q. Wang, P. Tuset-Peiro, X. Vilajosana, and S. Pollin, "Improving reliability and scalability of LoRaWANs through lightweight scheduling," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1830–1842, Mar. 2018.

[19] X. Liu, Z. Qin, Y. Gao, and J. A. McCann, “Resource allocation in wireless powered IoT networks,” *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4935–4945, Jan. 2019.
[20] B. Su, Z. Qin, and Q. Ni, “Energy efficient resource allocation for uplink LoRa networks,” in *Proc. IEEE GLOBECOM*, 2018, pp. 1–7.
[21] J. Haxhibeqiri, I. Moerman, and J. Hoebeke, “Low overhead scheduling of LoRa transmissions for improved scalability,” *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3097–3109, Oct. 2019.
[22] J. Pullmann and D. Macko, “A new planning-based collision-prevention mechanism in long-range IoT networks,” *IEEE Internet Things J.*, vol. 6, no. 6, pp. 9439–9446, Sep. 2019.
[23] D. Zorbas and B. O’Flynn, “Autonomous collision-free scheduling for LoRa-based industrial Internet of Things,” in *Proc. IEEE WOWMOM*, 2019, pp. 1–5.
[24] L. Shi, L. Zhao, G. Zheng, Z. Han, and Y. Ye, “Incentive design for cache-enabled D2D underlaid cellular networks using Stackelberg game,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 765–779, Jan. 2019.
[25] Y. Choi, H. Kim, S. Han, and Y. Han, “Joint resource allocation for parallel multi-radio access in heterogeneous wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3324–3329, Nov. 2010.
[26] H. Zhang, Y. Xiao, L. X. Cai, D. Niyato, L. Song, and Z. Han, “A multi-leader multi-follower Stackelberg game for resource management in LTE unlicensed,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 348–361, Jan. 2017.
[27] M. Li, “Generalized Lagrange multiplier method and KKT conditions with an application to distributed optimization,” *IEEE Trans. Circuits Syst.*, vol. 66, no. 2, pp. 252–256, May 2019.

Preti Kumari received the M.Tech. degree from the Indian Institute of Technology Patna in 2016. She is currently pursuing the Ph.D. degree with the Department of Computer Science and Engineering, Indian Institute of Technology (BHU) Varanasi, India. Her research interests include sensor networks, IoT, and mobile computing.

Hari Prabhat Gupta (Member, IEEE) received the Ph.D. degree in computer science and engineering from the Indian Institute of Technology Guwahati, in 2014. He is currently an Assistant Professor with the Department of Computer Science and Engineering, Indian Institute of Technology (BHU) Varanasi, India. His research interest includes IoT.

Tanima Dutta received the Ph.D. degree in computer science and engineering from the Indian Institute of Technology Guwahati, in 2014. She is currently an Assistant Professor with the Department of Computer Science and Engineering, Indian Institute of Technology (BHU) Varanasi, India. Her research interest includes multimedia.