Seong *et al. EURASIP Journal on Wireless Communications and Networking* (2016) 2016:162
DOI 10.1186/s13638-016-0644-y

EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

Reliable low-energy group formation for infrastructure-less public safety networks

Seonik Seong[1], Illsoo Sohn[2], Sunghyun Choi[1] and Kwang Bok Lee[1*]

Abstract

In this paper, we study an infrastructure-less public safety network (IPSN) where energy efficiency and reliability are critical requirements in the absence of cellular infrastructure, i.e., base stations and wired backbone lines. We formulate the IPSN group formation as a clustering problem. A subset of user equipments (UEs), called group owners (GOs), are chosen to serve as virtual base stations, and each non-GO UE, referred to as group member, is associated with a GO as its member. We propose a novel clustering algorithm in the framework of affinity propagation, which is a state-of-the-art message-passing technique with a graphical model approach developed in the machine learning field. Unlike conventional clustering approaches, the proposed clustering algorithm minimizes the total energy consumption while guaranteeing link reliability by adjusting the number of GOs. Simulation results verify that the IPSN optimized by the proposed clustering algorithm reduces the total energy consumption of the network by up to 31 % compared to the conventional clustering approaches.

Keywords: Infrastructure-less public safety networks, Energy efficiency, Reliability, Group formation, Affinity propagation

1 Introduction

A public safety network (PSN) has been developed as a special class of wireless communication network that aims to save lives and prevent property damage. PSNs have evolved separately from commercial wireless networks satisfying various requirements and regulatory issues associated with them [1, 2]. With growing needs for the transmission of multimedia data, existing voice-centric PSN technologies are facing hurdles in fulfilling the demand for high capacity and different types of services. Mission-critical requirements for PSNs include the guaranteed dissemination of emergency information such as alarm texts, images, and videos of disasters even in the absence (or destruction) of cellular infrastructure [3].

Many research projects have been launched to meet the mission-critical requirement of PSN, e.g., Aerial Base Station with Opportunistic Links for Unexpected & TEmporary events (ABSOLUTE), Alert for All (Alert4All), Mobile Alert InformAtion system using satellites (MAIA) [4]. The research projects include the emergency communications using satellite communications, aerial eNodeBs, and terrestrial radio access technologies. The approaches take advantages of inherent broadcasting and resilience with respect to Earth damages for disseminations of alert messages. In this paper, we limit our interests to terrestrial radio access technologies, e.g., Long-Term Evolution (LTE), TErrestrial Trunked RAdio (TETRA), TETRAPOL, and Digital Mobile Radio (DMR), because PSNs should be operational even in the low-class user equipments (UEs) that are lacking in satellite communication functionalities.

The current technology standards for PSNs, such as TETRA, Tetrapol, and DMR, which provide terrestrial radio access technologies using direct communications between devices when base stations and wired backbone lines are not operational in emergency situations as shown in Fig. 1 [3, 5, 6]. To utilize unprecedentedly rapid advances in commercial wireless networks, the 3rd Generation Partnership Project Long-Term Evolution (3GPP LTE) has recently been adopted as a baseline platform for the next-generation PSNs in the USA, the UK, and South Korea. 3GPP standardization body has been evolving LTE specifications to support direct communication between

*Correspondence: klee@snu.ac.kr
[1]Department of Electrical and Computer Engineering, INMC, Seoul National University, Gwanak-ro, Seoul 08826, South Korea
Full list of author information is available at the end of the article

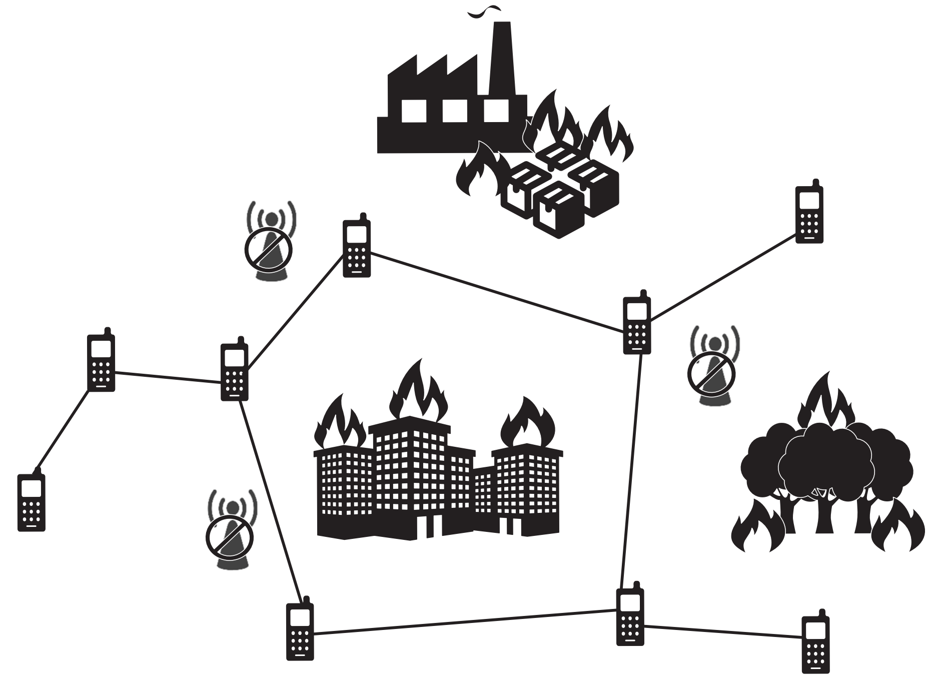

Fig. 1 Overview of an infrastructure-less public safety network

UEs in out of coverage [7–9]. To this end, recent evolution of 3GPP LTE has introduced group owner mode to enable public safety networks [10]. The group owner mode is a key feature for direct communications between UEs especially in infrastructure-less PSN (IPSN) where data transfer via infrastructure is unavailable. Some UEs are designated as group owners (GOs) in the group owner mode and are responsible for group formation, communication with their group members (GMs), and inter-group routing as a wireless backbone. Each non-GO UE becomes a GM of its proximate group by performing authentication with the GO of the group. The benefits from the group-based hierarchy in IPSN include better scalability, bandwidth reuse, and simple routing [11]. However, efficient group formation methods have not been fully studied.

In this paper, we focus on two critical goals to organize groups for PSNs: energy efficiency and reliability. IPSNs should minimize energy consumption of UEs to maximize the survival time of the networks. In addition, IPSNs should guarantee reliable communications. In order to guarantee reliable communications, any pair of UEs should be connected through qualified links only considering wireless coverage. These goals are extremely important in IPSNs, as failing to meet the goals could directly lead to losses of lives or property. To resolve the challenges, we formulate a constrained clustering problem for group formation of IPSNs considering both energy efficiency and reliability. Our formulation becomes a mixed-integer programming which requires a combinatorial optimization. We propose a low-complexity clustering algorithm to solve this problem. As will be discussed in detail in the next section, existing techniques are not efficient enough to provide a satisfactory solution to the problem, thus motivating us to introduce affinity propagation (AP) [12], a state-of-the-art message-passing technique. AP has been proven to be a very efficient tool for various types of optimization problems in communication networks [13–15]. The proposed clustering algorithm based on the AP framework efficiently minimizes the total energy consumption of IPSNs while guaranteeing reliable communications.

Our main contributions are summarized as follows:

- *IPSN organization:* We formulate the IPSN group formation as a *constrained* clustering problem. Unlike the conventional clustering problems studied in the literature, our clustering formulation minimizes energy consumption while guaranteeing reliable communications. This requires a computationally demanding optimization.
- *Low-complexity algorithm:* We propose a low-complexity clustering algorithm for the constrained clustering problem in the AP framework. The proposed clustering algorithm iteratively performs AP to determine the number of clusters adaptively. In

addition, we present how to determine the initial parameters to reduce the number of required iterations and improve the convergence speed.
- *Energy-saving performance:* Our simulation results verify that the proposed clustering algorithm reduces the total energy consumption in the network by up to 31 % compared to the conventional clustering algorithms. Note that such energy saving is achieved while guaranteeing reliable communication in the network.

The remainder of the paper is organized as follows. Section 2 describes the related work in comparison with our work. The system model and the problem formulation are described in Section 3. In Section 4, we propose a novel clustering algorithm in the AP framework. In Section 5, how to determine the initial parameter $p^{(1)}$ is presented to reduce the number of iterations required by the proposed clustering algorithm. We show the simulation results of the proposed clustering algorithm in Section 6. Finally, Section 7 concludes the paper.

Notations: $\max(x, y)$ denotes the maximum of x and y. $\lceil x \rceil$ is used to refer to the smallest integer not less than x. $|x|$ is the absolute value of a real number x. $|\mathcal{A}|$ denotes the cardinality of a set $\mathcal{A}$. The same notation is used to refer to the absolute value and the cardinality, and the meaning depends on the input of the operator. $\mathbb{E}[\cdot]$ denotes the expected value of a random variable. A summary of the notations frequently used in the paper is listed in Table 1.

2 Related work

2.1 Energy-minimizing network formation

The previous studies have investigated minimizing the total energy of the network in wireless sensor networks (WSNs) [16–21]. In WSNs, low-energy adaptive clustering hierarchy-centralized (LEACH-C) [18] proposed a pioneering idea, which minimizes the total energy consumption of the network using a clustered hierarchy. In LEACH-C, simulated annealing-based optimization algorithm [22] is adopted for cluster formation. Several improvements have been proposed to enhance LEACH-C by considering residual energy [19], a re-clustering frequency [20], and solar cells [21]. In [23–25], LEACH-C has been improved by replacing simulated annealing with K-means clustering [26], which has been proved to be efficient for most clustering tasks in machine learning fields.

Recent studies have developed network formation techniques based on the game theory [27–32]. Among various studies exploiting coalition game for communications networks, the study in [27] proposed a hybrid homogeneous LEACH protocol (HHO-LEACH) to minimize energy consumption of the network. As described in its name, HHO-LEACH is based on LEACH in minimizing energy consumption. However, the connectivity constraints are not considered in [27] like other studies that have enhanced LEACH.

Table 1 Frequently used notations

Notation	Definition
N	Number of UEs
$\mathcal{V}$	Set of GOs
$\mathcal{N}_j$	Set of GMs associated with GO j
d_{ij}	Distance between UE i and UE j
r_1	Maximum range for reliable intra-group link
r_2	Maximum range for reliable inter-group link
$w(d_{ij})$	Power consumption for the transmission between GO j and UE i
$\bar{w}$	Power consumption for the management of each group
$s(i,j)$	Similarity of UE i to UE j

2.2 Reliable network formation using CDS

Network formation under connectivity constraint has been focusing on connected dominating set (CDS) [33–36]. In graph theory, CDS is defined as a connected subgraph of a graph to which every vertex not belonging to the CDS is adjacent. For the last two decades, various algorithms such as Guha and Khuller's algorithm and Ruan's algorithm have been developed for various types of CDS constructions [37, 38]. In ad hoc networks, a CDS is used to select a wireless backbone network guaranteeing reliable connection for routing and control [33–36]. Vertices and edges in a graph represent the set of nodes and the set of links in ad hoc network, respectively. Whether any pair of nodes are adjacent to each other or not (that is, whether two vertices are connected by an edge in a graph) depends on the distance between them.

2.3 Limitations of previous approaches

The aforementioned studies cannot be directly applied to IPSNs because they focus on either minimizing the total energy consumption or guaranteeing reliable connection. Clustering algorithms that minimize total energy consumption do not consider connectivity constraints for reliable communication between UEs. On the other hand, minimization of total energy consumption of a network is not considered in the CDS-based network formation techniques. Thus, we propose a new network formation scheme for IPSN considering both minimizing total energy consumption and guaranteeing reliable communications.

3 System model and problem formulation

3.1 Channel model and network structure

An IPSN composed of N UEs is considered. We consider a hierarchical network topology in IPSN composed of GOs and GMs as shown in Fig. 2. Black circles, white circles, solid lines, and dashed lines are GOs, GMs, intra-group links between GO and GM, and inter-group links between neighbor GOs, respectively. GOs serve as virtual base stations. GOs coordinate their GMs for synchronizations, resource allocations, and initial attachments. They also relay data to support inter-group communications. Each GM is associated with the nearest GO. A set of GOs is denoted as $\mathcal{V}$, and a set of GMs associated with GO j ($\in \mathcal{V}$) is denoted as $\mathcal{N}_j$.

We consider path loss channel model defined in 3GPP specification for direct communication between UEs [9, 39]. The received signal power at UE j from UE i is given by

$$P_{ij}^{\mathrm{rx}} = P_i L_0 \left(\frac{d_{ij}}{d_0}\right)^{-\alpha}, \tag{1}$$

where P_i, d_{ij}, α, d_0, and L_0 denote the transmit power of UE i, the distance between UE i and UE j, path loss exponent, the reference distance, and the transmission loss at the reference, respectively. Then, the signal-to-noise ratio (SNR) at UE j from UE i is $\mathrm{SNR}_{ij} = P_{ij}^{\mathrm{rx}}/\sigma^2$, where σ^2 represents the noise power. A link from UE i to UE j is regarded as a reliable link when SNR_{ij} is larger than the SNR threshold γ, i.e., $\mathrm{SNR}_{ij} \geq \gamma$. The value of γ depends on modulation, code rate, and acceptable bit error rate. The maximum range of a reliable link with transmit power P and SNR threshold γ can be determined as

$$f(P,\gamma) = d_0 \left(\frac{PL_0}{\gamma\sigma^2}\right)^{1/\alpha}, \text{ for } P \leq \frac{\gamma\sigma^2}{L_0}. \tag{2}$$

To organize a reliable clustered hierarchy for an IPSN, two different transmission ranges for reliable links are considered depending on the link types. The transmission range for intra-group links is $r_1 = f(P_1, \gamma_1)$, where P_1 and γ_1 are maximum transmit power for inter-group links and SNR threshold for intra-group links, respectively. The transmission range for inter-group links is $r_2 = f(P_2, \gamma_2)$, where P_2 and γ_2 are the maximum transmit power for inter-group links and SNR threshold for inter-group links, respectively. Normally, higher transmit power, lower-order modulation, and lower-code rate are required for inter-group links of a wireless backbone, implying that $r_1 \leq r_2$ [11].

3.2 Problem formulation

We focus on minimizing total energy consumption in IPSN. Two types of energy consumption are considered, i.e., the energy consumption for transmissions between GO-GM pairs and the energy consumption for the management of groups. The energy consumption for the transmissions between GO j and GM i can be expressed using

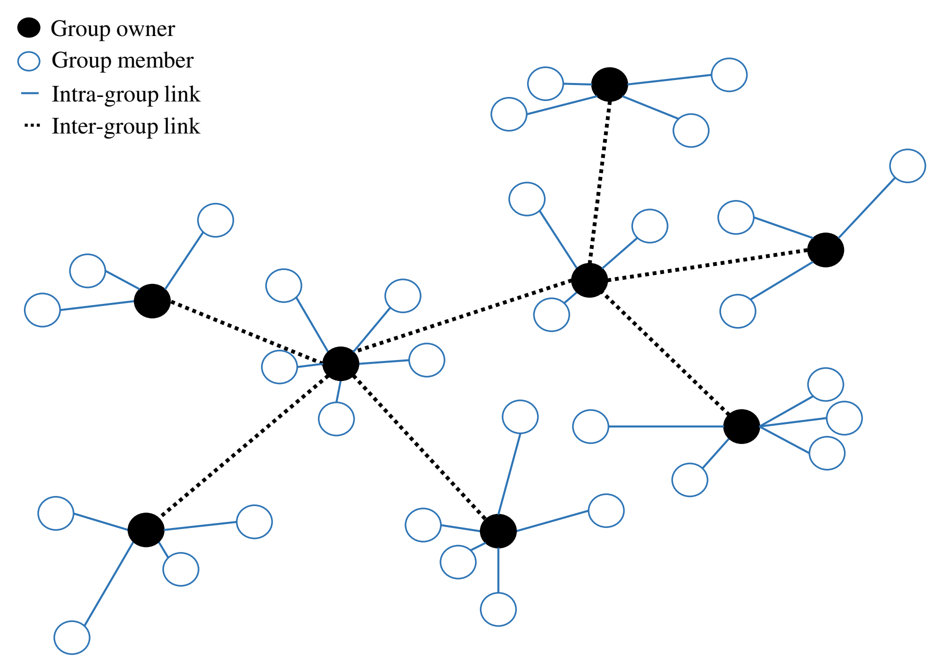

Fig. 2 Structure of an infrastructure-less public safety network

the SNR threshold γ_1. The minimum power consumption for transmission between GO j and GM i to be successfully detected at distance d_{ij} is given as

$$w(d_{ij}) = \frac{\gamma_1 \sigma^2}{L_0} \left(\frac{d_{ij}}{d_0} \right)^{\alpha}, \tag{3}$$

where σ^2, d_0, and L_0 are noise power, reference distance, and transmission loss at the reference, respectively. The power consumption for the management of each group is a system design parameter related to synchronization, resource allocation, inter-group transmissions, and other system factors. The constant power consumption per cluster is denoted as $\bar{w}_1$ and the power consumption affected by the number of the group members is assumed as $\bar{w}_2$ per intra-cluster link. The power consumption for the management of GO j's cluster can be calculated as $\bar{w}_1 + \sum_{i \in \mathcal{N}_j} \bar{w}_2$. The total power consumption for the management of all clusters in an IPSN can be derived as $\sum_{j \in \mathcal{V}} \bar{w}_1 + \sum_{j \in \mathcal{V}} \sum_{i \in \mathcal{N}_j} \bar{w}_2 = N\bar{w}_2 + \sum_{j \in \mathcal{V}} (\bar{w}_1 - \bar{w}_2)$, where N is the number of UEs in the IPSN. The constant term $N\bar{w}_2$ does not affect the optimization decision of clustering, and $\sum_{j \in \mathcal{V}} (\bar{w}_1 - \bar{w}_2)$ can be converted into $\sum_{j \in \mathcal{V}} \bar{w}$, where $\bar{w} = \bar{w}_1 - \bar{w}_2$. Then, the minimization of the total power is formulated as

$$\text{minimize} \sum_{j \in \mathcal{V}} \sum_{i \in \mathcal{N}_j} w(d_{ij}) + \sum_{j \in \mathcal{V}} \bar{w}. \tag{4}$$

In a group-based hierarchy in IPSN, reliable communication between any pair of UEs can be guaranteed by two connectivity conditions, namely, intra-group connectivity and inter-group connectivity. Intra-group connectivity condition implies that each GO should be in the communication range of all GMs associated with it. This condition can be represented as $d_{ij} \leq r_1,\ \forall j \in \mathcal{V},\ \forall i \in \mathcal{N}_j$. Inter-group connectivity condition implies that any GO should not be isolated from the network of GOs. For the mathematical representation of inter-group connectivity, we define the inter-group adjacency matrix $\mathcal{C}$ and the inter-group connectivity matrix $\mathcal{A}$. Matrix $\mathcal{C}$ is an $L \times L$ matrix whose (l, l') entry is one when the distance between the lth GO and the l'th GO is less than r_2, and is zero otherwise, where $L = |\mathcal{V}|$. Matrix $\mathcal{A}$ is an $L \times L$ matrix which is defined as $\mathcal{A} = \sum_{k=1}^{L-1} \mathcal{C}^k$. The (l, l') entry of $\mathcal{C}^k$ is greater than zero if the lth GO and the l'th GO are connected via k reliable inter-group links, and is zero otherwise. Thus, $a_{ll'} > 0$ means that the lth GO and the l'th GO are connected via one or multiple inter-group reliable links, where $a_{ll'}$ is the (l, l') entry of $\mathcal{A}$. Then, inter-group connectivity condition can be represented as $a_{ll'} > 0,\ \forall\, l, l' \in \{1, 2, \ldots, L\}$ such that $l \neq l'$.

An optimization problem for energy efficiency and reliability in an IPSN is formulated as

$$\text{minimize} \sum_{j \in \mathcal{V}} \sum_{i \in \mathcal{N}_j} w(d_{ij}) + \sum_{j \in \mathcal{V}} \bar{w} \tag{5a}$$

$$\text{subject to } \mathbf{C}_1 : d_{ij} \leq r_1,\ \forall j \in \mathcal{V},\ \forall i \in \mathcal{N}_j, \tag{5b}$$

$$\mathbf{C}_2 : a_{ll'} > 0,\ \forall\, l, l' \in \{1, 2, \ldots, L\} \text{ such that } l \neq l', \tag{5c}$$

where (5a) represents the minimization of the total energy consumption through the determination of $\mathcal{V}$, and $\mathcal{N}_j$'s, (5b) denotes the connectivity constraints $\mathbf{C}_1$ for intra-group connectivity, and (5c) denotes the connectivity constraints $\mathbf{C}_2$ for inter-group connectivity. Note that the clustering problem for minimizing the sum of the energy consumption without any constraints is known to be an NP-hard problem [40]. In addition to this problem, our formulation includes complicated integer constraints to guarantee reliability. This is very challenging, and none of the previous study investigates an efficient solution. To this end, we develop a novel low-complexity algorithm by introducing AP framework.

4 Constrained clustering algorithm for IPSN

First, we present some background with regard to the AP framework. Secondly, we embed the optimization problem (5) into the AP framework through similarity modeling. Finally, we develop a low-complexity clustering algorithm for IPSNs.

4.1 Preliminaries: affinity propagation

AP is a state-of-the-art clustering algorithm developed in computer science based on a message-passing algorithm [12, 41]. It was originally proposed to find a set of representative data points (or exemplars) to partition a set of data points into subsets of data points. AP has various advantages. First, it outperforms existing clustering algorithms, i.e., K-means clustering and simulated annealing. Secondly, AP provides a deterministic result, unlike other clustering algorithms whose performance depends on the choice of the initial point. Thirdly, AP can be adapted to various complicated problems owing to its flexibility.

The inputs of AP are the real-valued similarities, and the objective of AP is to select GOs which maximize the sum of similarities as formally described by

$$\text{maximize} \sum_{j \in \mathcal{V}} \sum_{i \in \mathcal{N}_j} s(i, j), \tag{6}$$

where $s(i, j)$ denotes the similarity of UE i to UE j. The similarity $s(i, j)$ of UE i indicates the suitability of UE j to be the GO of UE i. The preference (or self-similarity) $s(j, j)$ of UE j indicates the suitability of UE j to be a GO. AP considers all UEs as potential GOs and selects GOs by passing messages iteratively between the UEs. Two types of messages are defined in earlier work [12] as *responsibility* and *availability*. After each iteration of passing messages, the messages are updated by the update rule derived in

[12, 41] based on the max-sum algorithm in a factor graph. The message of responsibility $r(i,j)$ from UE i to potential GO j reflects the accumulated evidence about the suitability of UE j serving as the GO for UE i, considering other potential GOs for UE i. The initial value of responsibility is set to $r(i,j) = s(i,j) - \max_{j' \text{s.t.} j' \neq j} s(i,j')$. The update rule of the responsibility $r(i,j)$ from UE i to potential GO j is

$$r(i,j) = s(i,j) - \max_{j' \text{s.t.} j' \neq j} \left\{ a(i,j') + s(i,j') \right\}. \tag{7}$$

The availability $a(i,j)$ from the potential GO j to UE i reflects the accumulated evidence about the suitability choosing UE j as the GO of UE i, considering the support from other UEs that UE j should be a GO. The initial value of availability is set to $a(i,j) = 0$. The update rule of the availability $a(i,j)$ from potential GO j to UE i is

$$a(i,j) = \min \left(0\,,\; r(j,j) + \sum_{i' \text{s.t.} i' \notin \{i,j\}} \max\left(0, r(i',j)\right) \right). \tag{8}$$

The self-availability of UE j reflects the accumulated evidence about UE j being a GO, considering the positive responsibilities sent to candidate GO j from other UEs. The initial value of availability is also set to $a(j,j) = 0$. The self-availability of UE j is updated differently with the availability $a(i,j)$ as

$$a(j,j) = \sum_{i' \text{s.t.} i' \neq j} \max\left\{0, r(i',j)\right\}. \tag{9}$$

Messages are exchanged iteratively until the fixed number of iterations is reached or the decision of the GOs remains unchanged for the fixed number of iterations. After termination of message exchanges, the GO $j^\star$ of UE i is determined as

$$j^\star = \arg\max_j \left\{ a(i,j) + r(i,j) \right\}, \tag{10}$$

if $j^\star \neq i$. UE i itself becomes a GO if $j^\star = i$.

The overall computational complexity of affinity propagation increases as $\mathcal{O}(N\bar{N}t_{\max})$, where $\bar{N}$ denotes the average number of UEs adjacent to a UE and $t_{\max}$ denotes the maximum number of iterations. Decades of iterations are enough to converge in AP [15, 42] and the convergence properties and proofs are discussed in depth in [15, 42]. Another advantage of this approach is that the computational load of affinity propagation can be implemented in a distributed fashion using multi-core processors [43]. Although the renewal periods of clusters depend on the mobility patterns of UEs, clusters organized by the proposed algorithm are expected to be maintained for a few minutes without a renewal of the clusters. Considering that the proposed clustering algorithm is performed on a long-term time scale, the computation cost for the proposed clustering algorithm is manageable.

4.2 Similarity modeling

We embed the optimization problem (5) into the AP framework with the following similarity modeling. We define the similarity of UE i to UE j as

$$s(i,j) = \begin{cases} -w(d_{ij}), & \text{if } d_{ij} \leq r_1, \\ -\infty, & \text{otherwise.} \end{cases} \tag{11}$$

Then, minimizing the total energy consumption in IPSNs is rendered into maximizing the sum of similarities in AP. Our similarity modeling prevents cases in which the distance between a GM and its GO exceeds r_1. In such cases, the sum of the similarities goes to $-\infty$. Therefore, the intra-group connectivity condition is always satisfied. In addition, we define the preference (or self-similarity) of UE j as

$$s(j,j) = \begin{cases} p, & \text{if UE } j \text{ is eligible for a GO,} \\ -\infty, & \text{otherwise,} \end{cases} \tag{12}$$

where p takes a value that is less than zero. If UE j is not eligible to become a potential GO due to its high mobility, low residual energy, or UE type, it is precluded by setting $s(j,j)$ to $-\infty$.

4.3 Proposed clustering algorithm

When an UE detects the disruption of infrastructure, the UE tries to find new clusters of an IPSN to join. If the UE fails to find any available cluster, the UE initiates the procedures to form a clustered hierarchy for an IPSN. To notify the initiation of forming clusters, a flooding-based route discovery can be used, which is well known in distributed network management such as the generation and maintenance of a routing table in ad hoc network [44–46]. The flooding-based route discovery is useful in the disruption scenarios because it is assuming an ad hoc network without infrastructure. A UE initiates the procedures to form a clustered hierarchy by broadcasting an information request (IREQ) packet. Some of UEs receiving the IREQ packet become relay nodes and rebroadcast the IREQ packet. The IREQ packet is rebroadcast for a fixed number of hops. The UEs that received IREQ send an information response (IREP) packet to the source UE of the IREQ packet along the reverse path of the IREQ packet. The IREP contains the identification and location information. The source UE performs the proposed clustering algorithm with the collected UE information. If multiple UEs initiate clustering procedures simultaneously, the UE responsible for performing computation can be decided by the pre-determined priority, e.g., the ordering of user identification number. In case of the periodic renewals of clusters, GOs collect UE information of associated GMs and one of GOs collects all the information from GOs to perform the proposed clustering algorithm. After performing the proposed clustering algorithm, the result is delivered to each UE by relaying the result through the selected GOs. According to [44–46], it is known that the signaling overhead of the flooding-based signaling is manageable if it is performed on a long-term time scale as in the proposed clustering algorithm.

The outputs resulting from AP are a set of GOs, $\mathcal{V}$, and sets of GMs associated with each GO, $\mathcal{N}_j$'s. The intra-group connectivity is always satisfied for all outputs resulting from AP by our similarity modeling.

If the inter-group constraint is satisfied, the total power consumption $E(p)$ can be calculated using the power consumption model described in Section 3.2 as

$$E(p) = \begin{cases} \sum_{j\in\mathcal{V}} \sum_{i\in\mathcal{N}_j} w(d_{ij}) + \sum_{j\in\mathcal{V}} \bar{w}, & \text{if } \mathbf{C}_2 \text{ in (5c) is satisfied,} \\ \infty, & \text{otherwise,} \end{cases} \tag{13}$$

where $\mathcal{V}$ and $\mathcal{N}_j$ denote the set of GOs, and the sets of GMs associated with GO j, resulting from AP with preference p. If the result of AP does not satisfy the inter-group connectivity constraint, the outputs of AP with p are not valid. In such a case, we consider $E(p)$ as ∞. The proposed clustering algorithm finds the value of p minimizing $E(p)$ as described below. Therefore, the final outputs of the proposed clustering algorithm always satisfy the inter-cluster constraint.

The number of clusters is the key parameter for both total energy consumption and connectivity constraints. As the number of clusters increases, the first term of $E(p)$ decreases and the second term of $E(p)$ increases. If there are too few clusters, the connectivity constraints cannot be satisfied. The proposed clustering algorithm finds the proper number of clusters for the problem (5) by controlling the value of p considering total energy consumption and the connectivity constraints. Finding the optimal value of p is a complicated problem, and we propose an efficient method to find the optimal value of p based on what is known as golden section search [47]. Before beginning the golden section search, the value of preference changes with the moving rate of ρ to obtain the initial search interval of preference value to find the optimal value of p. The search interval of preference value is represented by the preference triplet $(\phi_{\min}, \phi_{\mathrm{c}}, \phi_{\max})$. The preference triplet $(\phi_{\min}, \phi_{\mathrm{c}}, \phi_{\max})$ consists of the minimum of the search interval $\phi_{\min}$, the maximum of the search interval $\phi_{\max}$, and the internal point of the search interval ϕ_{c}. The triplet should satisfy $\phi_{\min} < \phi_{\mathrm{c}} < \phi_{\max}$, $E(\phi_{\min}) > E(\phi_{\mathrm{c}})$, and $E(\phi_{\max}) > E(\phi_{\mathrm{c}})$. After finding the initial preference triplet, the preference triplet is iteratively updated based on the golden section search to narrow the search interval of preference value by evaluating the total power consumption at a new value in the search interval, namely ϕ. The overall procedure of the proposed clustering algorithm is shown in Algorithm 1, where $\mathbb{AP}(p)$ represents the optimization via AP with the preference p, and $\mathbb{GSS}(\phi_{\min}, \phi_{\mathrm{c}}, \phi, \phi_{\max})$ denotes the updates of the preference triplet and ϕ based on the golden section search.

Algorithm 1 Proposed clustering algorithm

```
Initialization: m = 2, p^(2) = p^(1) ρ
E(p^(1)) ← AP(p^(1)), E(p^(2)) ← AP(p^(2)), Δ = sign(E(p^(1)) − E(p^(2)))
Phase I. Determination of initial triplet
repeat
    m = m + 1
    p^(m) = p^(1) ρ^(Δ(m−1+(Δ−1)/2))
    E(p^(m)) ← AP(p^(m))
    (φ_min, φ_c, φ_max) = (p^(m−1−Δ), p^(m−1), p^(m−1+Δ))
until E(φ_min) > E(φ_c) and E(φ_max) > E(φ_c)
Phase II. Golden section search
repeat
    (φ_min, φ_c, φ, φ_max) ← GSS(φ_min, φ_c, φ, φ_max)
    E(φ) ← AP(φ)
until |φ_max − φ_min| < ε(|φ_c| + |φ|)
if E(φ) > E(φ_c) then p_r = φ_c
else p_r = φ
end if
return E(p_r) ← AP(p_r)
```

To find the initial triplet of the preferences, we consider the initial value of p, denoted by $p^{(1)}$, and the moving rate of p, denoted as ρ. Initially, AP is performed with the preference of $p^{(1)}$ and $p^{(2)}$, where $p^{(2)} = p^{(1)}\rho > p^{(1)}$. If $E(p^{(1)}) > E(p^{(2)})$ or $E(p^{(1)}) = E(p^{(2)}) = \infty$, AP is iteratively performed by increasing $p^{(m)} = p^{(1)}\rho^{m-1}$ $(m \geq 3)$ until finding M such that $E(p^{(M-1)}) \leq E(p^{(M)})$. The initial triplet of the preferences in this case is $(\phi_{\min}, \phi_{\mathrm{c}}, \phi_{\max}) = (p^{(M-2)}, p^{(M-1)}, p^{(M)})$. If $E(p^{(1)}) < E(p^{(2)})$, AP is iteratively performed by decreasing $p^{(m)} = p^{(1)}\rho^{-(m-2)}$ $(m \geq 3)$ until finding M such that $E(p^{(M-1)}) \leq E(p^{(M)})$. The initial triplet of the preferences in this case is $(\phi_{\min}, \phi_{\mathrm{c}}, \phi_{\max}) = (p^{(M)}, p^{(M-1)}, p^{(M-2)})$.

After determining the initial triplet of preferences, the preference value is updated based on the golden section search to find the final preference value. The golden section search narrows successively the search interval of preference value by updating the triplet of preferences. Using $(\phi_{\min}, \phi_{\mathrm{c}}, \phi_{\max})$, a new preference of ϕ is calculated as

$$\phi = \begin{cases} \phi_{\mathrm{c}} + \left(2 - (1+\sqrt{5})/2\right)(\phi_{\max} - \phi_{\mathrm{c}}), & \text{if } \phi_{\mathrm{c}} - \phi_{\min} < \phi_{\max} - \phi_{\mathrm{c}}, \\ \phi_{\mathrm{c}} - \left(2 - (1+\sqrt{5})/2\right)(\phi_{\mathrm{c}} - \phi_{\min}), & \text{otherwise.} \end{cases} \tag{14}$$

Among $\phi_{\min}$, ϕ_{c}, $\phi_{\max}$, and ϕ, the golden section search determines a new triplet of preferences based on the results of AP with the preference of $\phi_{\min}$, ϕ_{c}, ϕ, and $\phi_{\max}$ as

$$(\phi_{\min},\phi_{\mathrm{c}},\phi_{\max}) \leftarrow \begin{cases} (\phi_{\min},\phi,\phi_{\mathrm{c}}), & \text{if } E(\phi) \le E(\phi_{\mathrm{c}}) \text{ and } \phi < \phi_{\mathrm{c}}, \\ (\phi,\phi_{\mathrm{c}},\phi_{\max}), & \text{if } E(\phi) > E(\phi_{\mathrm{c}}) \text{ and } \phi < \phi_{\mathrm{c}}, \\ (\phi_{\min},\phi_{\mathrm{c}},\phi), & \text{if } E(\phi) > E(\phi_{\mathrm{c}}) \text{ and } \phi > \phi_{\mathrm{c}}, \\ (\phi_{\mathrm{c}},\phi,\phi_{\max}), & \text{if } E(\phi) \le E(\phi_{\mathrm{c}}) \text{ and } \phi > \phi_{\mathrm{c}}. \end{cases} \tag{15}$$

The update of the preference triplets terminates if $\phi_{\max} - \phi_{\min} < \epsilon|\phi_{\mathrm{c}} + \phi|$, where ϵ denotes the termination threshold. The final preference value is $p_{\mathrm{r}} = \phi_{\mathrm{c}}$ if $E(\phi) > E(\phi_{\mathrm{c}})$, and the final preference value is $p_{\mathrm{r}} = \phi$ if $E(\phi) < E(\phi_{\mathrm{c}})$. The final clustering outputs are the $\mathcal{V}$ and $\mathcal{N}_j$s resulting from AP with the preference p_{r}.

5 Determination of the initial point

The proposed clustering algorithm is based on iterative updates of the AP solution that requires long processing time and high computational complexity. The required number of iterations of the proposed clustering algorithm depends on the choice of the initial value of preference $p^{(1)}$. If the iteration begins with an arbitrary value of $p^{(1)}$, the proposed clustering algorithm finds inefficiently the final value of p. To reduce the number of iterations of the proposed clustering algorithm, we determine the $p^{(1)}$ by estimating the minimum number of clusters, denoted by κ, that satisfies the connectivity constraints.

The estimated minimum number of clusters, κ, depends on r_1, r_2, and S because the entire area of S should be covered by the clusters, and r_1 and r_2 limit the maximum radii of clusters. When $r_1 < r_2/2$, intra-group connectivity is a major factor that limits the maximum radii of clusters and determines κ. When $r_1 > r_2/2$, inter-group connectivity is a major factor that limits the maximum radii of clusters and determines κ. Therefore, we conclude that κ can be calculated as

$$\kappa = \left\lceil \max\left(\frac{S}{\pi\,(r_1)^2}, \frac{S}{\pi\,(r_2/2)^2}\right)\right\rceil. \tag{16}$$

Then, we determine $p^{(1)}$ which corresponds to κ via their mathematical relation. We determine the relation between the number of clusters and the value of p in AP by averaging the distribution of the UEs.

Lemma 1 *Let N and S be the number of UEs and the area, respectively. Then, the relation between the number of clusters, K, and the value of preference p in AP is*

$$p = -\frac{\gamma_1\sigma^2(2K + \alpha(N-K))}{KL_0(\alpha+2)}\left(\frac{S}{\pi K d_0^2}\right)^{\alpha/2}. \tag{17}$$

Proof See Appendix. □

Using Lemma 1, $p^{(1)}$ can be set to

$$p^{(1)} = -\frac{\gamma_1\sigma^2(2\kappa + \alpha(N-\kappa))}{\kappa L_0(\alpha+2)}\left(\frac{S}{\pi \kappa d_0^2}\right)^{\alpha/2}. \tag{18}$$

The derived relation between the number of clusters and the value of p in AP is evaluated in Fig. 3 with $N = 400$ and 1000, and the other parameters are same with the parameters described in Table 4. Figure 3 shows that the relation is well estimated such that it can be utilized in the determination of $p^{(1)}$.

6 Simulation results

We present the simulation results of the proposed clustering algorithm. The simulation parameters used in this paper is listed in Table 2. A rectangular area of 2×2 km is considered. A WINNER+ B1 path loss model with $\alpha = 4.37$, $d_0 = 1$ m, and $L_0 = 0.068$ is set with a height of 1.5 m and a frequency of 700 MHz, which is the most widely used bandwidth [9, 39]. For intra-group links, SNR threshold γ_1 is set to 9 dB with 16 QAM, a 616/1024 code rate, and 1 % bit error rate based on the performance evaluation in an LTE system [48]. Similarly, γ_2 is set to 3 dB with QPSK, a 602/1024 code rate, and 1 % bit error rate for inter-group links. Maximum transmit powers of $P_1 = 23$ dBm and $P_2 = 30$ dBm are considered for intra-group link and inter-group link, respectively. A single set of affinity propagation terminates when the number of iterations is 1000 or the decisions of the GOs are identical for 10 iterations. A moving rate of preference is set to $\rho = 0.3$. A termination threshold of $\epsilon = 0.01$ is considered in the golden section search. The proposed clustering algorithm is compared with K-means clustering [23], LEACH-C [19], and CDS-based formation [36]. The performances are averaged over 100 independent realizations of user distributions.

Figure 4 shows the average number of clusters that minimizes the total energy consumption as a function of $\bar{w}$ for different numbers of UEs, N. The proposed clustering algorithm autonomously finds the number of clusters which minimizes the total energy consumption while satisfying the connectivity constraints. In K-means clustering and LEACH-C, the average number of clusters that minimizes the total energy consumption is attained by performing the algorithms for $1 \le K \le N$ clusters. The value of K that satisfies connectivity constraints in less than 90 % of the realizations is excluded from the results for a fair comparison. As the value of $\bar{w}$ increases, the number of clusters decreases in the low $\bar{w}$ region of $\bar{w} \le 25$ dBm because the total energy consumption can

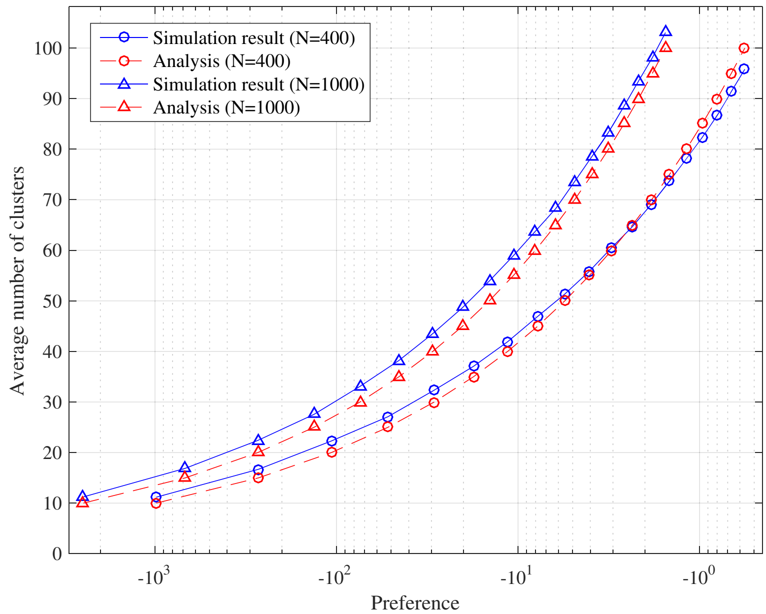

Fig. 3 Relation between the preference and the average number of clusters

be minimized with less number of clusters. In this region, total energy consumption determines the resulting number of clusters because the connectivity constraints are always satisfied given this number of clusters. On the other hand, the numbers of clusters are nearly identical in the high $\bar{w}$ region of $\bar{w} > 25$ dBm. Despite the fact that fewer clusters is preferred in the perspective of total energy consumption, the number of clusters cannot be less than the minimum number of clusters in the region to satisfy connectivity constraints. The proposed clustering algorithm can satisfy the connectivity constraints with fewer clusters because the proposed clustering algorithm is designed considering the connectivity constraint by the similarity modeling. CDS-based formation gives the same results regardless of $\bar{w}$ because CDS-based formation does not consider power consumption.

Table 2 Simulation parameters

Parameter	Value
Area, S	2×2 km
Path loss exponent, α	4.37
Transmission loss at $d_0 = 1$ m, L_0	0.068 dB
Noise power, σ^2	-104 dBm
Maximum Tx power for intra-group link, P_1	23 dBm
Maximum Tx power for inter-group link, P_2	30 dBm
SNR threshold for intra-group link, γ_1	9 dB
SNR threshold for inter-group link, γ_2	3 dB

CDS-based formation does not differentiate the maximum ranges according to the link type (intra-group link or inter-group link) [36]. We consider the maximum range for intra-group r_1 as a maximum range in CDS-based formation to guarantee that the connectivity constraints are satisfied.

Figure 5 shows how the average total energy consumption in the proposed clustering algorithm varies with $\bar{w}$ when the number of UEs is $N = 400$. The proposed clustering algorithm is compared with K-means clustering with $K = 35$, 55, and 74 clusters and LEACH-C with $K = 41$, 61, and 80 clusters. The selected values of K are the value of K minimizing the total power consumption at $\bar{w} = 35$ dBm in Fig. 4, the value of K minimizing the total power consumption at $\bar{w} = 15$ dBm in Fig. 4, and the mean of these two values. CDS-based formation creates the same number of clusters regardless of $\bar{w}$ as shown in Fig. 4. The proposed clustering algorithm shows the lowest energy consumption compared to K-means clustering, LEACH-C, and CDS-based formation, irrespective of the value of $\bar{w}$. Specifically, the proposed clustering algorithm provides less energy consumption than K-means clustering, LEACH-C, and CDS-based formation by up to 31 % in the high $\bar{w}$ region of $\bar{w} > 25$ dBm where connectivity constraints dominate the resulting number of clusters. The performance gain of the proposed clustering algorithm in the high $\bar{w}$ region comes from the relatively small number of clusters satisfying the connectivity constraints in the proposed clustering algorithm. In the region of $\bar{w} < 25$ dBm, the proposed clustering algorithm

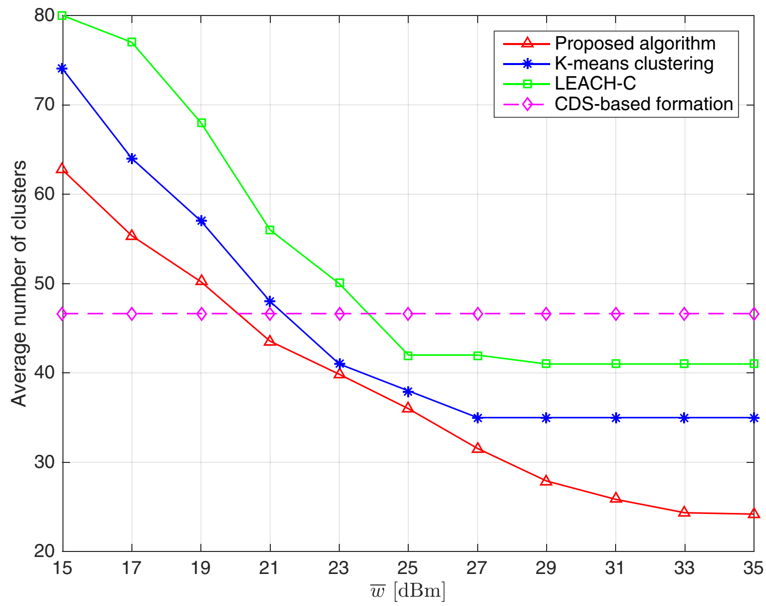

Fig. 4 Average number of clusters versus the value of $\bar{w}$, where $N = 400$

provides better power minimization than K-means clustering by about 12 %. The performance gain of the proposed clustering algorithm in the low $\bar{w}$ region comes from the outstanding clustering performance of the AP framework.

Figure 6 compares the average total energy consumption in the proposed clustering algorithm with those of K-means clustering and LEACH-C as a function of the number of UEs when $\bar{w} = 30$ dBm. The selected values of $K = 35$ for K-means and $K = 41$ for LEACH-C are the minimum number of clusters satisfying the connectivity constraints in more than 90 % of the realizations because $\bar{w} = 30$ dBm is in the high $\bar{w}$ region. The average total energy consumption gradually increases with the number of UEs. The proposed clustering algorithm shows more energy saving than the existing clustering algorithms by

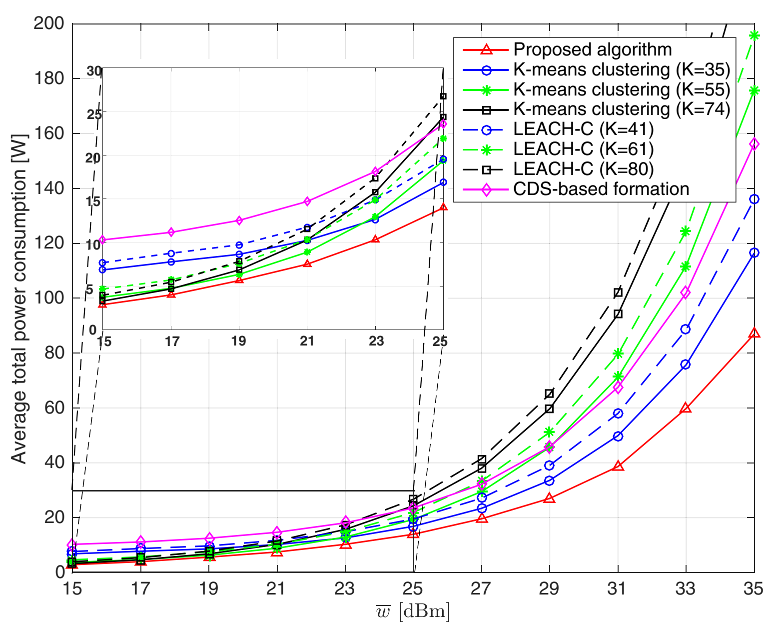

Fig. 5 Average total power consumption versus the value of $\bar{w}$, where $N = 400$

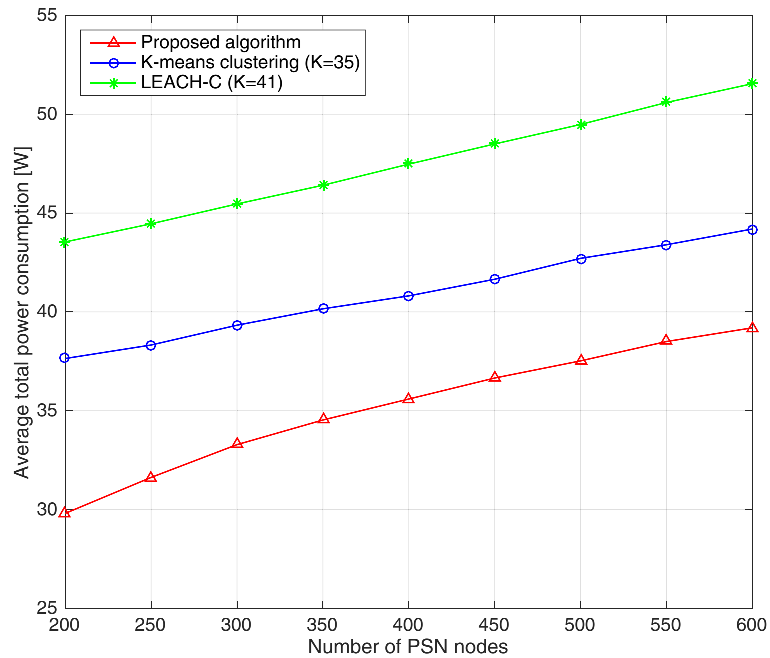

Fig. 6 Average total power consumption versus the number of UEs, where $\bar{w} = 30$ dBm

26 % when $N = 200$ and by 12 % when $N = 600$. As noted above, the performance gain of the proposed clustering algorithm comes from the relatively small number of clusters in the high $\bar{w}$ region. As the number of UEs increases, the performance difference decreases because the portion of $\sum_{j\in\mathcal{V}} \bar{w}$ out of the total power consumption decreases.

Similarly, Fig. 7 compares the average total energy consumption in the proposed clustering algorithm with that of K-means clustering and LEACH-C as a function of the number of UEs when $\bar{w} = 20$ dBm in the low $\bar{w}$ region. $K = 41$, 53, and 65 are selected for K-means clustering and $K = 45$, 65, and 84 are selected for LEACH-C. The selected values of K are the value of K minimizing the total power consumption at $N = 200$, the value of K minimizing the total power consumption at $N = 800$, and the mean of these two values. As the number of UEs in IPSN increases, more number of clusters are created to minimize the total power consumption. The proposed clustering algorithm reduces the total power consumption by about 12 % compared to K-means clustering and by about 30 % compared to LEACH-C. The performance gain which comes from the outstanding clustering performance of the AP framework remains steady regardless of the number of UEs.

Table 3 shows how the average total power consumption varies with the number of iterations, denoted by $t_{\max}$, in the proposed clustering algorithm. Case A considers $N = 400$ and $\bar{w} = 20$ dB, case B considers $N = 400$ and $\bar{w} = 30$ dB, and case C considers $N = 800$ and $\bar{w} = 20$ dB. In general, larger number of iterations improves the performance of the proposed algorithm. However, the algorithm complexity can be considerably reduced by sacrificing the performance slightly. Note that the proposed clustering algorithm with $t_{\max} = 70$ still notably outperforms the conventional clustering algorithm. Considering that the proposed clustering algorithm is performed on a long-term time scale as described in Fig. 10, the computation cost for the proposed clustering algorithm is manageable.

Figure 8 shows the empirical cumulative distribution function (CDF) of the number of hops between UEs, where $N = 400$. The average latency in packet delivery between the UEs is proportional to the number of hops between them. More than 90 % UE pairs can deliver packets to each other within four hops when the simulation area is 2×2 km. The structures of clustered hierarchies are different at $\bar{w} = 20$ and 30 dB due to the different number of clusters. However, the average number of hops between any two pairs is similar.

To provide the simulation results in more practical scenarios, option 1, which defines the layout for the simulation of IPSNs in 3GPP LTE [9], is used in Figs. 9 and 10. The parameters defined in option 1 are described in Table 4. eNodeBs are used only for describing the UE distributions, and it is assumed that no eNodeB is enabled for IPSNs as defined in [9]. Shadowing is assumed to follow log-normal distribution, and the shadowing

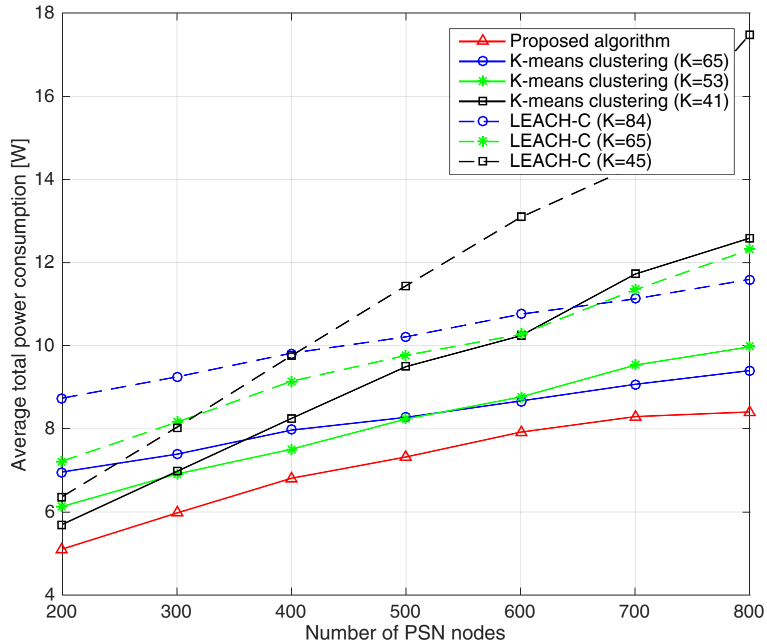

Fig. 7 Average total power consumption versus the number of UEs, where $\bar{w} = 20$ dBm

standard deviation, denoted by σ_{sh}, is varied from $\sigma_{\text{sh}} = 0$ to 10 dB. In the practical model, the connectivity constraints cannot be always guaranteed in the proposed clustering algorithm due to the channel fluctuation. We define the outage probability as the probability that the received SNR at a UE from its GO is less than the required threshold γ_1. The outage probability shows how reliable the links between GMs and GOs are in a fading model and a mobility model.

Figure 9 shows the average outage probability in the proposed clustering algorithm as a function of the shadowing standard deviation. The proposed clustering algorithm always satisfies the connectivity constraints, and the outage probability equals zero when shadow fading is not considered. As the shadowing standard deviation increases, the outage probability increases. The proposed clustering algorithm can reduce the outage probability by conservative setting in the connectivity constraints. In other words, the outage probability can be reduced if r_1 and r_2 are set to smaller values. The conservative setting for reliability results in higher power consumption in the proposed clustering algorithm. The tradeoff between the reliability and the power consumption should be considered in the practical implementations.

Figure 10 shows the average outage probability in the proposed clustering algorithm as a function of time and UE speed, denoted by v. A random walk mobility model [49], which is one of the most employed models used for ad hoc network [50] and sensor network [51], is used in the mobility pattern of UEs. To focus on the impact of mobility in the proposed clustering algorithm, the impact of shadow fading is excluded here. With no mobility of UEs, the outage probability equals zero by the clustering result ensuring the connectivity constraints from the clustering algorithm. As UEs move faster, the outage probability increases fast and the renewal of clusters should be performed more frequently to maintain a reliable hierarchy. However, the renewal of clusters every 10 min is enough to maintain a reliable hierarchy even when $v = 9$ km/h. The outage probability can be reduced by setting r_1 and r_2 to smaller values, and the similar tradeoff between the reliability and the power consumption with the fading model exists in the mobility model.

Table 3 Average total power consumption vs the number of iterations

	Proposed algorithm (W)				*K*-means clustering (W)
Number of iterations	100	90	80	70	
Case A	6.8	6.8	6.9	7.1	7.5
Case B	35.5	36.2	37.7	38.5	41.1
Case C	8.3	8.4	8.5	8.7	9.3

7 Conclusions

A novel clustering algorithm based on affinity propagation is proposed to provision both energy efficiency

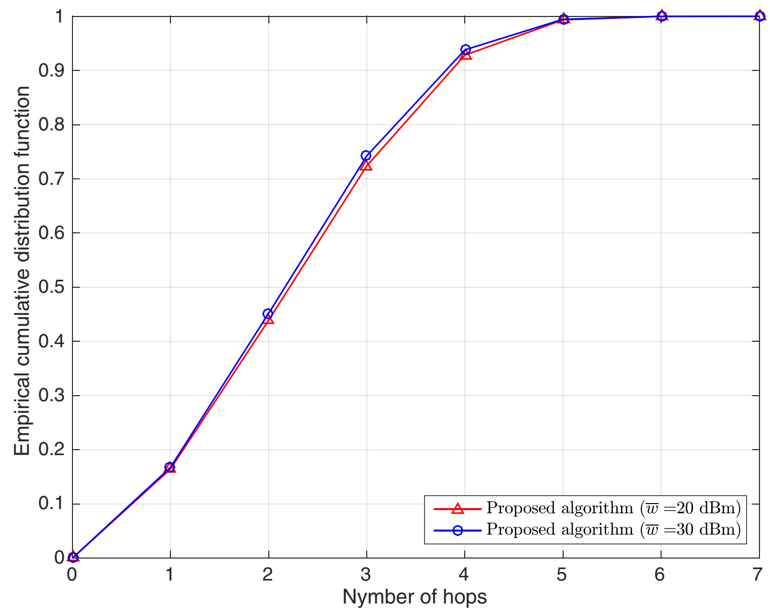

Fig. 8 Empirical CDF of the number of hops, where $N = 400$

and reliability in IPSNs. As a key novelty, we embed the optimization problem of IPSNs into the AP framework by means of similarity modeling and proposing an efficient method to find the number of GOs based on the golden section search. The proposed clustering algorithm adaptively determines the number of clusters that minimizes the total energy consumption while satisfying the connectivity constraints. The simulation results have shown that the proposed clustering algorithm considerably outperforms K-means clustering, LEACH-C, and CDS-based formation in various environments. Future research directions include studies of the impact of UE mobility, heterogeneous traffic patterns, and different device types.

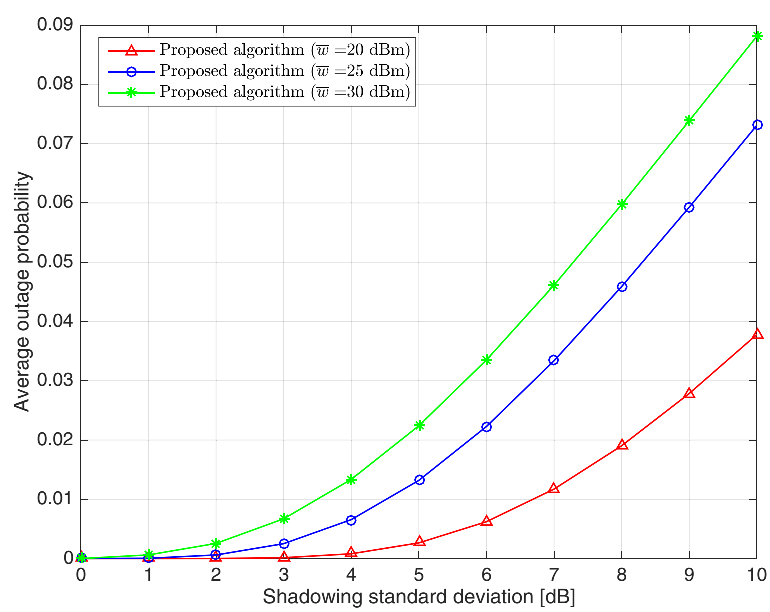

Fig. 9 Average outage probability versus the shadowing standard deviation

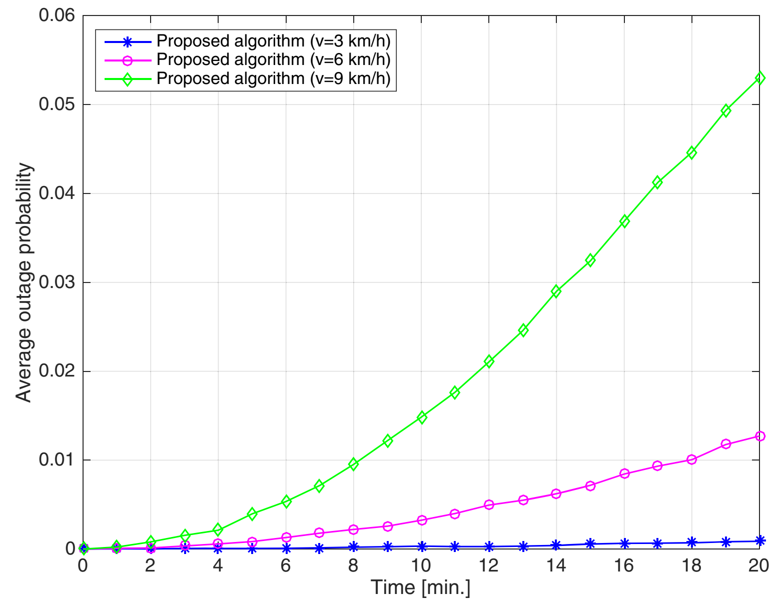

Fig. 10 Average outage probability versus time, where $\bar{w} = 25$ dBm

Appendix

Proof of Lemma 1

We define the sum of similarities with the given p and K as

$$R(\mathcal{V}_K, \mathcal{N}_1, \ldots, \mathcal{N}_K) := \sum_{j \in \mathcal{V}_K} \sum_{i \in \mathcal{N}_j} s(i,j) + \sum_{j \in \mathcal{V}_K} p, \quad (19)$$

where $\mathcal{V}_K$ and $\mathcal{N}_j$ denote the set of K GOs and the set of GMs associated with GO j resulting from AP, respectively. The similarity function is assumed as follows:

$$s(i,j) = -w(d_{ij}) = \frac{\gamma_1 \sigma^2}{L_0} \left(\frac{d_{ij}}{d_0} \right)^{\alpha}. \quad (20)$$

We assume that GMs are uniformly distributed around the GOs within a circle with a radius of $\sqrt{S/\pi K}$. Then, the expectation of $R(\mathcal{V}_K, \mathcal{N}_1, \ldots, \mathcal{N}_K)$ is a function of p and K. It can be derived as

$$\begin{aligned} \bar{R}(p,K) &= \mathbb{E}\left[R\left(\mathcal{V}_K, \mathcal{N}_1, \ldots, \mathcal{N}_K \right) \right] \\ &= \sum_{j \in \mathcal{V}_K} \sum_{i \in \mathcal{N}_j} \bar{s}(K) + \sum_{j \in \mathcal{V}_K} p \\ &= -\frac{2\gamma_1 \sigma^2 (N-K)}{L_0(\alpha+2)} \left(\frac{S}{\pi K d_0^2} \right)^{\alpha/2} + Kp. \end{aligned} \quad (21)$$

Here, $\bar{s}(K)$ is the mean similarity when K clusters exist; it can be calculated as

$$\bar{s}(K) = \int_0^{2\pi} \int_0^{\sqrt{S/\pi K}} \frac{1}{\pi \left(\sqrt{S/\pi K} \right)^2} \cdot -w(r)\, r\, dr\, d\theta \quad (22)$$

$$= -\frac{2\gamma_1 \sigma^2}{L_0(\alpha+2)} \left(\frac{S}{\pi K d_0^2} \right)^{\alpha/2}. \quad (23)$$

As the number of clusters increases, the sum of similarities decreases while the sum of preferences increases proportionally to the number of clusters. AP autonomously finds the number of clusters that maximizes the sum of similarities with the given value of p as

$$\hat{K} = \arg\max_{K} \bar{R}(p,K). \quad (24)$$

Unfortunately, K is restricted to being an integer. Therefore, $\hat{K}$ can be determined by calculating $\bar{R}(p,K)$ for all K such that $1 \leq K \leq N$ which is computationally difficult in the case of a large N. Thus, we propose a suboptimal approach with a relaxation technique. That is, K is relaxed as a positive real number of $K \in [1, N]$. This allows us to

Table 4 Simulation parameters for Figs. 9 and 10

Parameter 2	Value
ISD	500 m
Number of eNodeB	19
Number of sector per eNodeB	3
Number of UE per sector	10 (uniform, outdoor)
eNodeBs enabled	0 %
Shadowing	i.i.d.

determine the partial derivative of $\bar{R}(p,K)$ with respect to K as follows:

$$\frac{\partial \bar{R}(p,K)}{\partial K} = \frac{\gamma_1 \sigma^2 (2K + \alpha(N-K))}{KL_0(\alpha+2)} \left(\frac{S}{\pi K d_0^2}\right)^{\alpha/2} + p. \tag{25}$$

For $2 \leq \alpha$, $\frac{\partial \bar{R}(p,K)}{\partial K}$ is a continuous and monotonically decreasing function of $K \in [1,N]$ and there exists an unique maximum value of $\bar{R}(p,K)$. The relation between p and K which makes the result of (25) equal to zero is a one-to-one function for $K \in [1,N]$. Therefore, by reformulating $\frac{\partial \bar{R}(p,K)}{\partial K} = 0$ in terms of p, we can ascertain the value of p to create K clusters in AP as

$$p = -\frac{\gamma_1 \sigma^2 (2K + \alpha(N-K))}{KL_0(\alpha+2)} \left(\frac{S}{\pi K d_0^2}\right)^{\alpha/2}. \tag{26}$$

This completes the proof.

Competing interests
The authors declare that they have no competing interests.

Acknowledgements
This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIP) (No. 2014R1A2A2A01003637), Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. NRF-2015R1D1A1A01057100), and LG Electronics Co. Ltd.

Author details
[1]Department of Electrical and Computer Engineering, INMC, Seoul National University, Gwanak-ro, Seoul 08826, South Korea. [2]Department of Electronics Engineering, Gachon University, Seongnamdae-ro, 13120 Seongnam, South Korea.

Received: 12 November 2015 Accepted: 21 May 2016
Published online: 07 July 2016

References
1. T Doumi, MF Dolan, S Tatesh, A Casati, G Tsirtsis, K Anchan, D Flore, LTE for public safety networks. IEEE Commun Mag. **51**(2), 106–112 (2013). doi:10.1109/MCOM.2013.6461193
2. Technical Advisory Board for First Responder Interoperability, Recommended minimum technical requirements to ensure nationwide interoperability for the nationwide public safety broadband network. Technical report (May 2012)
3. G Baldini, S Karanasios, D Allen, F Vergari, Survey of wireless communication technologies for public safety. IEEE Commun Surv Tutor. **16**(2), 619–641 (2014). doi:10.1109/SURV.2013.082713.00034
4. T Pecorella, LS Ronga, F Chiti, S Jayousi, L Franck, Emergency satellite communications: research and standardization activities. IEEE Commun Mag. **53**(5), 170–177 (2015). doi:10.1109/MCOM.2015.7105657
5. ETSI, *Terrestrial Trunked Radio (TETRA); Voice plus Data (V+D); designer's guide; part 3: direct mode operation (DMO). TR 300-3 V1.3.3*. (European Telecommunications Standards Institute (ETSI), June 2009)
6. D Buchanan, 700 MHz public safety broadband task force report and recommendations. NPSTC (2009)
7. 3GPP, service requirements for the evolved packet system (EPS). TS 22.278 V13.2.0, The 3rd Generation Partnership Project (3GPP) (Dec. 2014)
8. 3GPP, feasibility study for proximity services. TR 22.803 V12.2.0, The 3rd Generation Partnership Project (3GPP) (June 2013)
9. 3GPP, study on LTE device-to-device proximity services. TR 36.843 V12.0.1, The 3rd Generation Partnership Project (3GPP) (Mar. 2014)
10. 3GPP, Study on architecture enhancements to support proximity-based services (ProSe). TR 23.703 V12.0.0, The 3rd Generation Partnership Project (3GPP) (Feb. 2014)
11. X Liu, A survey on clustering routing protocols in wireless sensor networks. Sensors. **12**(8), 11113–11153 (2012)
12. BJ Frey, D Dueck, Clustering by passing messages between data points. Science. **315**(5814), 972–976 (2007). doi:10.1126/science.1136800
13. I Sohn, SH Lee, JG Andrews, Belief propagation for distributed downlink beamforming in cooperative MIMO cellular networks. IEEE Trans Wireless Commun. **10**(12), 4140–4149 (2011). doi:10.1109/TWC.2011.101210.101698
14. M Shamaiah, SH Lee, S Vishwanath, H Vikalo, Distributed algorithms for spectrum access in cognitive radio relay networks. IEEE J Selected Areas Commun. **30**(10), 1947–1957 (2012). doi:10.1109/JSAC.2012.121110
15. SH Lee, I Sohn, Affinity propagation for energy-efficient BS operations in green cellular networks. IEEE Trans Wireless Commun. **14**(8), 4534–4545 (2015). doi:10.1109/TWC.2015.2422701
16. AA Abbasi, M Younis, A survey on clustering algorithms for wireless sensor networks. Comput Commun. **30**(14), 2826–2841 (2007)
17. A Anand, NB Mehta, Quick, decentralized, energy-efficient one-shot max function computation using timer-based selection. IEEE Trans Commun. **63**(3), 927–937 (2015). doi:10.1109/TCOMM.2014.2386330
18. WB Heinzelman, A Chandrakasan, H Balakrishnan, An application-specific protocol architecture for wireless microsensor networks. IEEE Trans Wireless Commun. **1**(4), 660–670 (2002). doi:10.1109/TWC.2002.804190
19. M Tripathi, MS Gaur, V Laxmi, RB Battula, in *Proc International Conference on Computational Intelligence and Information Technology*. Energy efficient LEACH-C protocol for wireless sensor network, (2013), pp. 402–405. doi:10.1049/cp.2013.2620
20. F Zhao, Y Xu, R Li, Improved LEACH routing communication protocol for a wireless sensor network. Int J Distributed Sensor Netw. **2012**, 1–6 (2012)
21. T Voigt, A Dunkels, J Alonso, H Ritter, J Schiller, in *Proc International Symposium on Computers and Communications*. Solar-aware clustering in wireless sensor networks, vol. 1, (2004), pp. 238–243. doi:10.1109/ISCC.2004.1358411
22. T Murata, H Ishibuchi, in *Proc IEEE Conference on Evolutionary Computation*. Performance evaluation of genetic algorithms for flowshop scheduling problems, (1994), pp. 812–8172. doi:10.1109/ICEC.1994.349951
23. D Mechta, S Harous, I Alem, D Khebbab, in *International Conference on Innovations in Information Technology*. LEACH-CKM: Low energy adaptive clustering hierarchy protocol with K-means and MTE, (2014), pp. 99–103. doi:10.1109/INNOVATIONS.2014.6987570
24. GY Park, H Kim, HW Jeong, HY Youn, in *Proc International Conference on Advanced Information Networking and Applications Workshops*. A novel cluster head selection method based on K-means algorithm for energy efficient wireless sensor network, (2013), pp. 910–915. doi:10.1109/WAINA.2013.123
25. P Sasikumar, S Khara, in *Proc International Conference on Computational Intelligence and Communication Networks*. K-means clustering in wireless sensor networks, (2012), pp. 140–144. doi:10.1109/CICN.2012.136
26. J Macqueen, in *Proc. the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. Statistics. Some methods for 2 classification and analysis of multivariate observations, vol. 1, (1967), pp. 281–297. http://projecteuclid.org/euclid.bsmsp/1200512992
27. S Tyagi, S Tanwar, SK Gupta, N Kumar, S Misra, JPC Rodrigues, S Ullah, in *IEEE International Conference on Communications (ICC)*. Bayesian coalition game-based optimized clustering in wireless sensor networks, (2015), pp. 3540–3545. doi:10.1109/ICC.2015.7248873
28. D Wu, Y Cai, L Zhou, J Wang, A cooperative communication scheme based on coalition formation game in clustered wireless sensor networks. IEEE Transactions on Wireless Communications. **11**(3), 1190–1200 (2012). doi:10.1109/TWC.2012.012712.111049
29. W Saad, Q Zhu, T Basar, Z Han, A Hjorungnes, in *IEEE Global Telecommunications Conference*. Hierarchical network formation games in the uplink of multi-hop wireless networks, (2009), pp. 1–6. doi:10.1109/GLOCOM.2009.5426293
30. W Saad, Z Han, M Debbah, A Hjorungnes, T Basar, Coalitional game theory for communication networks. IEEE Signal Process. Mag. **26**(5), 77–97 (2009). doi:10.1109/MSP.2009.000000

31. X Lu, P Wang, D Niyato, in *IEEE Wireless Communications and Networking Conference (WCNC)*. Hierarchical cooperation for operator-controlled device-to-device communications: a layered coalitional game approach, (2015), pp. 2056–2061. doi:10.1109/WCNC.2015.7127784
32. Y Xiao, KC Chen, C Yuen, Z Han, LA DaSilva, A Bayesian overlapping coalition formation game for device-to-device spectrum sharing in cellular networks. IEEE Trans Wireless Commun. **14**(7), 4034–4051 (2015). doi:10.1109/TWC.2015.2416178
33. J Yu, N Wang, G Wang, D Yu, Connected dominating sets in wireless ad hoc and sensor networks—a comprehensive survey. Comput Commun. **36**(2), 121–134 (2013). doi:10.1016/j.comcom.2012.10.005
34. KM Alzoubi, P-J Wan, O Frieder, Distributed heuristics for connected dominating sets in wireless ad hoc networks. Commun Netw J. **4**(1), 22–29 (2002)
35. J Blum, M Ding, A Thaeler, X Cheng, *Connected dominating set in sensor networks and MANETs. Handbook of Combinatorial Optimization: Supplement Volume B*. (Springer, Boston, MA, 2005), pp. 329–369
36. D Kim, Y Wu, Y Li, F Zou, D-Z Du, Constructing minimum connected dominating sets with bounded diameters in wireless networks. IEEE Trans Parallel Distributed Syst. **20**(2), 147–157 (2009). doi:10.1109/TPDS.2008.74
37. S Guha, S Khuller, Approximation algorithms for connected dominating sets. Algorithmica. **20**(4), 374–387 (1998)
38. L Ruan, H Du, X Jia, W Wu, Y Li, K-I Ko, A greedy approximation for minimum connected dominating sets. Theor Comput Sci. **329**(1), 325–330 (2004)
39. J Meinilä, *et al*, D5.3: WINNER+ final channel models. Wireless World Initiative New Radio WINNER (2010)
40. O Kariv, SL Hakimi, An algorithmic approach to network location problems II: The p-medians. SIAM J Appl Math. **37**(3), 539–560 (1979)
41. FR Kschischang, BJ Frey, HA Loeliger, Factor graphs and the sum-product algorithm. IEEE Trans Inform Theory. **47**(2), 498–519 (2001). doi:10.1109/18.910572
42. SH Lee, I Sohn, Distributed relay pairing for bandwidth exchange based cooperative forwarding. IEEE Commun Lett. **19**(3), 459–462 (2015). doi:10.1109/LCOMM.2014.2385064
43. M Wang, W Zhang, W Ding, D Dai, H Zhang, H Xie, L Chen, Y Guo, J Xie, Parallel clustering algorithm for large-scale biological data sets. PloS One. **9**(4), 91315 (2014)
44. CE Perkins, EM Royer, in *IEEE Workshop on Mobile Computing Systems and Applications*. Ad-hoc on-demand distance vector routing, (1999), pp. 90–100. doi:10.1109/MCSA.1999.749281
45. H Dubois-Ferriere, M Grossglauser, M Vetterli, in *Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Networking & Computing*. Age matters: efficient route discovery in mobile ad hoc networks using encounter ages (ACM, 2003), pp. 257–266. doi:10.1145/778415.778446
46. XM Zhang, EB Wang, JJ Xia, DK Sung, A neighbor coverage-based probabilistic rebroadcast for reducing routing overhead in mobile ad hoc networks. IEEE Trans Mobile Comput. **12**(3), 424–433 (2013). doi:10.1109/TMC.2011.277
47. J Kiefer, Sequential minimax search for a maximum. Proc Am Math Soc. **4**(3), 502–506 (1953)
48. J Blumenstein, JC Ikuno, J Prokopec, M Rupp, in *Proc. ELMAR*. Simulating the long-term evolution uplink physical layer, (2011), pp. 141–144
49. T Camp, J Boleng, V Davies, A survey of mobility models for ad hoc network research. Wireless Commun Mobile Comput. **2**(5), 483–502 (2002)
50. J Liu, H Nishiyama, N Kato, T Kumagai, A Takahara, Toward modeling ad hoc networks: current situation and future direction. IEEE Wireless Commun. **20**(6), 51–58 (2013)
51. R Bellazreg, N Boudriga, Dyntunkey: a dynamic distributed group key tunneling management protocol for heterogeneous wireless sensor networks. EURASIP J Wireless Commun Netw. **2014**(1), 1–19 (2014). doi:10.1186/1687-1499-2014-9