

# A New Green Clustering Algorithm for Energy Efficiency in High-Density WLANs

**Yang Lu, Xuezhi Tan, Yun Mo and Lin Ma**

Communication Research Center, Harbin Institute of Technology  
Harbin, China

[e-mail: tanxz1957@hit.edu.cn]

\*Corresponding author: Xuezhi Tan

*Received August 22, 2013; revised December 20, 2013; accepted February 9, 2014; published February 28, 2014*

---

## **Abstract**

In this paper, a new green clustering algorithm is proposed to be as a first approach in the framework of an energy efficient strategy for centralized enterprise high-density WLANs. Traditionally, in order to maintain the network coverage, all the APs within the WLAN have to be powered-on. Nevertheless, the new algorithm can power-off a large proportion of APs while the coverage is maintained as its always-on counterpart. The two main components of the new approach are the faster procedure based on K-means and the more accurate procedure based on Evolutionary Algorithm (EA), respectively. The two procedures are processes in parallel for different designed requirements and there is information interaction in between. In order to implement the new algorithm, EA is applied to handle the optimization of multiple objectives. Moreover, we adapt the method for selection and recombination, and then introduce a new operator for mutation. This paper also presents simulations in scenarios modeled with ray-tracing method and FDTD technique, and the results show that about 67% to 90% of energy consumption can be saved while it is able to maintain the original network coverage during periods when few users are online or the traffic load is low.

---

**Keywords:** Green clustering, energy efficiency, EA, optimization of multiple objectives

---

This research is supported by National Science Foundation (Grant No. 61101122), National Science and Technology Major Project (Grant No. 2012ZX03004-003), National High Technology Research and Development Program of China (863 Program) (Grant No. 2012AA120802), Postdoctoral Science-Research Development Foundation of Heilongjiang Province (Grant No. LBH-Q12080) and Specialized Research Fund for the Doctoral Program of Higher Education (Grant No. 20122301120004).

<http://dx.doi.org/10.3837/tiis.2014.02.001>

## 1. Introduction

More and more enterprise offices [1], university campuses [2] and municipal downtowns [3] deploy WLANs for flexible Internet connectivity. In many offices and campuses WLANs are deployed consisting of hundreds even thousands of APs (access point) which make the WLAN high-density to meet the increasing demand of network accesses. Many of those WLANs, so called traditional WLANs, are designed to improve the usage of users during times of peak demands, but the peak demand rarely occurs [4]. All APs in traditional WLANs are always-on to ensure the coverage of network, and many of the APs are considered idle or redundant when no users are associated with them. The utilization rate will be even lower during weekends or school holidays, and this is a kind of energy wastage which can be evitable.

SEAR (Survey, Evaluate, Adapt and Repeat) is an energy efficient strategy in high-density WLANs proposed by Jardosh et al. to manage the network resource [5]. In [5], a green clustering algorithm is introduced to form AP clusters and select the cluster-head APs for SEAR. Green clustering is the first algorithm for energy efficient strategies such as SEAR and the green clustering selects cluster-head APs to form a topology which can maintain the same (almost the same) coverage of network as the traditional WLANs.

In this paper, a new green clustering algorithm based on Evolutionary Algorithm (EA) [6] is proposed, and this method will henceforth be referred to as NGCA. NGCA consists of two different procedures which are the faster procedure implemented by  $K$ -means algorithm and the more accurate procedure implemented by EA, respectively. The new approach forms AP clusters and selects cluster-head AP in a completely new perspective. Actually, either of the two procedures may achieve the designed requirements of the new algorithm, but reasonably combining the two approaches and integrating as the proposed NGCA can lead to a further improvement in terms of efficiency and accuracy.

The remainder of this paper is arranged as follow. An overview of related work is given in Section II. The high-density WLAN is introduced firstly, followed by expression about the mechanism of SEAR, and then we detail the differences with respect to the green clustering algorithm of SEAR in Section III. While in Section IV we propose NGCA based on  $K$ -means algorithm and EA. Section V provides simulation results and performance analysis. Section VI is conclusion and discussion of the paper.

## 2. Related Work

Concerning of energy efficiency in WLANs, D. Alessandro et al. introduce a multi-carrier infrastructure to implement energy conservation in WLANs [7]. According to the analysis in [8], L. Miliotis et al. study about the relations between communication efficiency and energy consumption which are important to wireless network management in wireless communication. In [9], J. Rabaey considers that most energy consumption of wireless communication device is in standby mode and reducing the energy consumption of components is an important way to implement green communication. Marsan et al. introduce an analytical model for traffic in WLANs and present the threshold determined by the number of clients and traffic of an AP switch algorithm [10]. Y. Chen [11] and A. Kushki et al. [12] use information theory and probability theory to implement AP selection and A. Kushki et al. quantify the resolution of location on the basis of Fisher Rule for AP selection [13]. In [14], S. H. Fang et al. use a prediction-based method to cluster APs in a decorrelated space In WSNs

(wireless sensor networks) Y. Kang et al. study about the low-power asynchronous broadcast MAC protocol which employs the strobe preamble, compute the energy consumption via rigorous mathematical analysis, and the results show that the proposed scheme outperforms B-CAST, a simple broadcast extension of the well-known B-MAC [15]. Lim S. et al. consider the state-of-the-art ACK-based low power listening (LPL) scheme in WSNs suffers from collision problems due to the protocol incompleteness, and propose  $\tau$ -duration CCA and Short Preamble Counter to mitigate the collision problem in ACK-based LPL MAC protocol and conserve more energy by reducing unnecessary overhearing, respectively, and they demonstrate the improvement of their scheme via a mathematical analysis and experiments, and their results show that up to 36% of energy is saved [16]. TDMA (Time Division Multiple Access) is a widely used MAC technique, and in TDMA energy saving is achieved by switching radio of nodes which are not engaged off, but the switching may increase end-to-end delay. In [17], Y. Zuo et al. propose hybrid algorithm which combines genetic algorithm and simulated annealing algorithm to solve a multi-objective TDMA scheduling problem. Kim D. W. and Park T. study about energy efficient MAC protocol for WSNs, and their scheme use the superframe structure which is bounded by the transmission of a beacon frame and have active and inactive portions, to save power consumption and guarantee QoS of real-time traffic [18]. S. Kang et al. consider that the clustering method is one of the preferred ways to produce a topology for reduced electrical energy consumption in wide area sensor networks, and propose cluster topology method based on sociological structures and concepts [19]. In WSN, the most important and common requirements regardless of application types are to provide a long network lifetime and small end-to-end delay. J. Oak et al. propose Early Preamble MAC with improved energy conservation and low latency, and the proposed scheme is based on CMAC and adopts an early preamble type [20]. Y. Zhang et al. propose the unit QoE per Watt, which is termed QoE efficiency, as a user-oriented metric to evaluate energy efficiency for wireless networks, investigate the tradeoffs between QoE efficiency and spectral efficiency, and their results show the proposed unit are helpful for network design and optimization [21]. J. Zhu studies about energy efficient cellular transmission which can improve the energy efficiency of wireless communication, he considers a cellular network consisting of one base station (BS) and multiple user terminals and explore the network coding for enhancing the energy efficiency of cellular downlink transmission from BS to users, and he proposes a network coded cellular transmission scheme which significantly outperforms the traditional cellular transmission in terms of energy efficiency [22]. Due to the uncertain of connections in Delay Tolerant Networks (DTNs), most routing algorithms in DTNs need nodes to forward the message to others based on the opportunistic contract which is related with the beaconing rate, and bigger beaconing rate uses more energy. In [23], Y. Wu et al. try to exploit the optimal beaconing rate and forwarding rate when total energy is constraint. Employing mobile sinks for data gathering become a trend for energy conservation in WSN, and in [24] M. Eslaminejad et al. study about an energy efficient dual-sink algorithm with role switching mechanism which utilizes both static and mobile sinks. Their algorithm could be employed in event-driven and multi-hop scenarios with improvements on the lifetime, the load and the end-to-end delay. However, few of these works above involve the concept of energy efficiency in high-density WLANs or green clustering.

SEAR (Survey, Evaluate, Adapt and Repeat) is an energy efficient strategy in high-density WLANs proposed by Jardosh et al. to manage the network resource [5]. An energy efficient algorithm in high-density WLANs based on SEAR is studied in this paper, and green clustering algorithm is proposed and implemented.

Green clustering algorithm is an algorithm which generates the initial topology of APs for the energy efficient strategy. An AP selection is done by green clustering algorithm, and the coverage can be ensured by the selected APs which are a small part of the total ones in a high-density WLAN. The selection process is dynamic. APs are powered-off for energy conservation in an energy efficient WLAN, and the transmission power of cluster-head APs is optimized for energy conservation further more. In some particular scenarios which are not strict with coverage of network and allow spectrum holes to emerge only in the edge region APs can continue declining the transmission power and even be powered-off to reduce the number of idle APs.

### 3. Analysis of Energy Efficient Strategy and Green Clustering Algorithm

#### 3.1 Centralized High-Density WLANs

In traditional WLANs APs are independent of each other and access Internet through wired switches. As the enterprise WLAN generally consists of hundreds even thousands of APs, it is obviously inconvenient for management. For the purpose to make WLAN easier and friendlier to manage, many WLAN vendors such as Aruba Networks, Meru Networks, Symbol Technologies and Trapeze Networks have adopted the centralized approach to WLANs, as shown in Fig. 1.

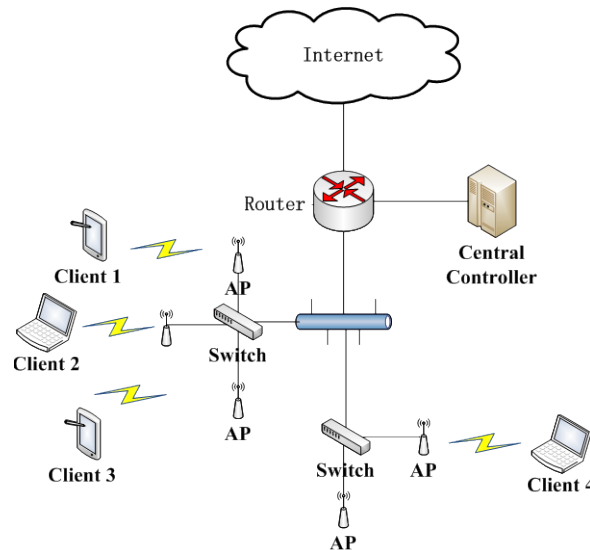


Fig. 1. A centralized WLAN infrastructure.

The deployment of centralized WLANs, which are convenient for network resource management, contributes to the growth and proliferation of enterprise WLANs. The number of APs in an enterprise WLAN increases exponentially every year [1][2][25]. So the energy efficient strategy studied in this paper is mainly concentrated on centralized high-density WLANs. To determine that either a WLAN is high-density or not, a parameter is introduced and defined as

$$\mu_{AP} = \frac{n_{AP}}{S} \quad (1)$$

where  $n_{AP}$  is the number of APs in the WLAN, and  $S$  is the coverage of the WLAN relatively. So  $\mu_{AP}$  denotes the number of APs per unit area in the coverage region. Apparently, a bigger value means the density of APs within the WLAN is comparatively higher.

We present case studies from two different large-scale enterprise WLANs. One example of such an enterprise WLAN is installed at Intel Corporation's building in Portland, OR, U.S., where 125 APs have been deployed at a distance of 5m from each other. There are four floors in the building and each floor is 80m × 38m. Another example is in the Microsoft campus at Redmond, WA, U.S., where 5000 APs have been deployed and the area is about 32,000 square meters [4]. According to (1), the values of  $\mu_{AP}$  are 0.01028 in Intel Corporation's building and 0.01563 in the Microsoft campus, respectively. From the two instances mentioned above, in the case that the value of  $\mu_{AP}$  within a concerned WLAN is more than 0.01, the corresponding WLAN is considered high-density.

### 3.2 Study of Green Clustering Algorithm in SEAR

SEAR is proposed by Jardosh as a demand-driven strategy to manage APs in high-density WLANs efficiently. And it is a policy-based method which can be tailored to achieve the performance desired by WLAN administrators. Moreover, based on some certain policies used, SEAR can conserve energy while the same performance is maintained as clients receive in its always-on counterpart.

SEAR resides on the central controller of a centralized WLAN through which it can control all APs. SEAR is assumed to have complete knowledge of the physics position and status of all APs managing the WLAN to achieve the designed requirements efficiently and switch the APs on or off in accordance with the policy used. SEAR consists of four components which are green clustering algorithm, user demand estimation, topology management, and user association, as shown in Fig. 2 [5].

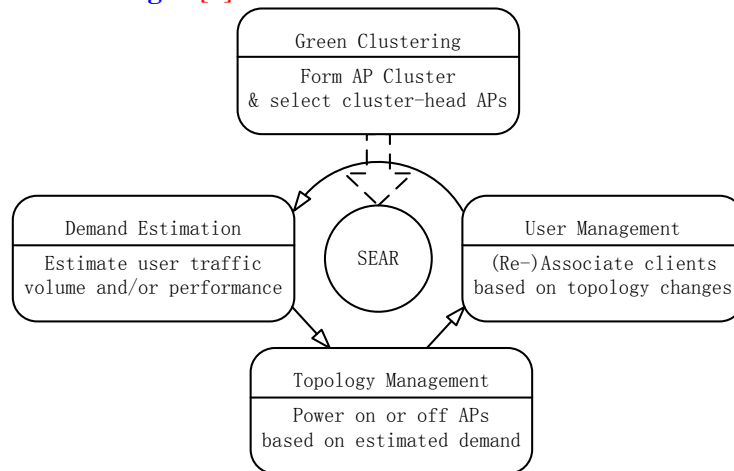


Fig. 2. Components of SEAR.

A green clustering algorithm forms AP clusters and selects a cluster-head AP for per cluster. SEAR uses green clustering algorithm as the initialization to provide necessary information for the further processing. The premise of green clustering is that a single AP from each of those clusters is sufficient to provide basic coverage to users in the vicinity of any AP within that cluster. As shown in Fig. 3, consider five APs (1 to 5) are placed within close proximity. Before applying green clustering all the five APs work in smaller transmitting powers.

However, in case of green clustering deployed, AP with label 1 is able to provide coverage to the region while others are shut down. Moreover, the transmitting power of AP1 has been increased for extended coverage. Therefore, the five APs form a new cluster and AP1 is the cluster-head AP as shown in Fig. 3.

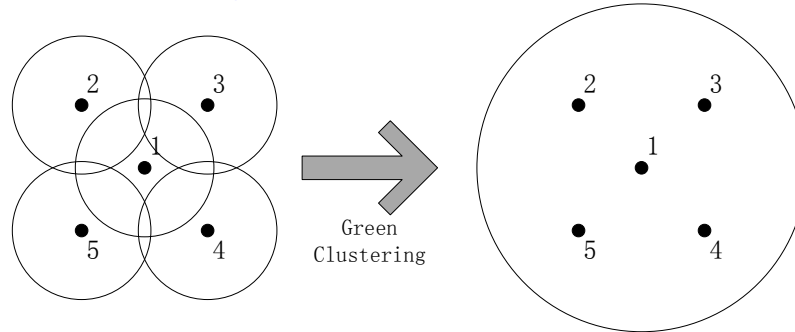


Fig. 3. Illustration of cluster formation.

Besides, it is worth mentioning that the transmitting power of a net device is only a tiny part of the total power consumption. Taking Cisco WAP4410N Wireless-N AP of S series for example, its transmitting power with single antenna is 13dBm to 17dBm (12.6 to 50.1 mW), while this AP's total power consumption is about 10W during normal work. Therefore, even the transmitting power increases for satisfying the algorithm requirement, the impact on total power consumption could be ignored because the energy saving by powering-off APs is much more significant.

In SEAR, green clustering is implemented in two steps which are neighborhood discovery and cluster formation, respectively.

In the first step of SEAR's green clustering, the APs need to work as a sniffing monitor and the central controller need to configure all the APs within the WLAN to the same channel for more than 1 min interval, even while the AP may be providing connectivity to clients. During that interval, each AP records the number and the signal strength of beacon frames from every other AP. If and only if the number and the signal strength are both no smaller than the pre-determined thresholds, the two AP are assumed to be in very close proximity. By deploying this method, the number of neighbors within each AP's neighborhood is obtained and then we let  $n_i$  denote the number mentioned for AP  $i$ .

In the second step, similar to the algorithm suggested by Bejerano [26], [5] utilizes a fast greedy clustering approach. In this step,  $n_i$  occurs as attribute value of AP  $i$ . To implement cluster formation, the maximum  $n_i$  is chosen first and the corresponding AP forms the first cluster. After that, a set as a cluster is formed,  $i$  is added into the set  $R_i$  as the first component of the cluster and at the meantime  $i$  is removed from all neighborhood sets of other APs. After  $i$  is added, all the APs  $j$  in the neighborhood of  $i$  are stepped through and added to the set  $R_i$  if and only if the new  $j$  is in the neighborhood set of every other AP already added to  $R_i$ . Once all the APs that satisfy the condition that being added to  $R_i$ , AP  $i$  is made the cluster-head of the cluster. Then the algorithm of cluster formation is iterated until all the APs in the WLAN have been added to some sets.

Once SEAR forms green clusters, the cluster-head APs remain powered-on by default, while the other ones in each cluster, which are called secondary APs, are supposed to be off for

energy conservation. Moreover, the transmitting power of cluster-head APs needs to be increased to maintain the network coverage.

Green clustering algorithm of SEAR is measurement-based to form AP clusters which are in close proximity of each other. The neighborhoods can be varied based on the thresholds of signal strength and number of received beacon messages, and the thresholds directly affect the results of green clustering. In other words, if low thresholds are chosen, SEAR's green clustering algorithm forms smaller size of clusters which means smaller power saving. Conversely, high thresholds may results in the lack of the network coverage. The choice of thresholds is critical for SEAR's green clustering.

#### 4. A New Green Clustering Algorithm

In Section III we illustrate about SEAR and its green clustering. To implement the initialization of the energy efficient strategy in high-density WLANs, we propose a new green clustering algorithm (NGCA) with a completely different perspective and approach. In order to implement our method, we analyze and study the propagation model of a typical indoor environment to indicate the coverage of an AP, and then on the basis of the propagation models a new green clustering consisting of the faster procedure and the more accurate procedure, is proposed to form green clusters of APs.

##### 4.1 Estimation of APs' Coverage

As mentioned in Section III, a central controller has the complete knowledge of physics positions of all APs in the WLAN, and the coverage of AP is the evaluation index of our algorithm for future processing. The coverage region of an AP can be described as an area in which the Rx RSS (received signal strength) from the current AP is over a pre-determined criterion, and the value determining the boundary of access service can refer to Rx sensitivity of some protocol. In order to specify network parameters, the WLAN we investigate is assumed as an IEEE 802.11b WIFI network, and typical Rx sensitivities for IEEE 802.11b are -86dBm at 1Mbps, -83dBm at 2Mbps, -79dBm at 5.5Mbps and -76dBm at 11Mbps, respectively. Therefore, for better network performance the area in which Rx RSS from an AP exceeds the biggest Rx sensitivity of 802.11b is considered to be in the coverage region.

NGCA usually works in indoor environments, and consists of two procedures which are for clustering fast and accurately, respectively. When few users are online or the traffic load is low, in order to maintain the basic network coverage which is measured by coverage of each AP, the fast procedure could supply a coarse result of AP clusters in a relatively short time, then the algorithm carries on to achieve a more accurate AP deployment for further power conservation. For different purposes of the algorithm, the coverage of an AP should be measured in different methods. Therefore, we introduce a simple and fast method based on the Friis free space equation for the fast procedure, and then employ a complex but accurate method based on ray-tracing and FDTD (finite-different time-domain) for the accurate one. First, the simple method (named simple modeling) would be presented as follow.

In the free space, the power received by a receiver antenna which is separated from a radiating transmitter antenna by a distance  $d$  is given by the Friis free space equation

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 l} \quad (2)$$

where  $P_t$  is the transmitting power,  $P_r$  is the received power which is a function of the distance  $d$ ,  $G_t$  is the transmitter antenna gain,  $G_r$  is the receiver antenna gain,  $l$  is the system loss factor not related to propagation ( $l \neq 1$ ), and  $\lambda$  is the wavelength in meters.

The environment algorithm is generally concentrated in an indoor environment. Due to the complex propagation environment, the attenuation factor model is introduced and analyzed to estimate the value of coverage radius of APs. According to (2), considering of the propagation loss of EM wave, the path loss of EM wave in the band around 2.4 G is given by approximately

$$L = 46 + 10n \cdot \log d \quad (3)$$

where  $L$  denotes the path loss and  $n$  is the attenuation factor which differs in different environments. Generally, in the relatively isolated environment,  $n$  is set 3.0 to 3.5. So the coverage radius  $Rad$  can be given by

$$Rad = 10^{\left(\frac{L-46}{10n}\right)} \quad (4)$$

And if the value of  $P_t$  is set to 15dBm,  $Rad$  can be considered as 30m.

For the fast procedure, the coverage of an AP is considered as a circle of which radius can be estimated according to (4). However, for future energy efficiency, a more accurate method to estimate the coverage should be employed. The coverage region can be measured with devices, and we have made a radio map consisting of RSS values from APs deployed in the building where our office resides (detailed in Subsection 5.2), but there are several serious problems to seek for the boundary of AP in this way, and some of the problems are that if the location or Tx power of some AP changes, or a new AP is deployed, the radio map might have to be rebuilt which makes the timeliness poor, so building a radio map, particular in a large area, is a hard work but with limited robustness and poor generalization. Therefore, for better exactness and robustness, a method for indoor propagation modeling (named complicated modeling) based on ray-tracing and FDTD is introduced.

Ray-tracing method is presented to evaluate the indoor wave propagation and penetration [27]. According to ray-tracing, waves from a Tx antenna could be modeled as ray tubes, and from geometrical optics, the E-field of the ray tube at Rx can be given by

$$\vec{E} = \vec{E}_0 \cdot \left\{ \prod \bar{\bar{R}}_i \right\} \cdot \left\{ \prod \bar{\bar{T}}_i \right\} \cdot \left\{ \prod \exp(-\gamma_i l_i) \right\} \cdot SF \quad (5)$$

where  $\vec{E}_0$  denotes the E-field at a reference point  $RP_0$ ,  $\left\{ \prod \bar{\bar{R}}_i \right\}$  and  $\left\{ \prod \bar{\bar{T}}_i \right\}$  represent the reflection and transmission coefficient dyads along the whole path, respectively.  $\left\{ \prod \exp(-\gamma_i l_i) \right\}$  denotes the propagation phase variations and exponential losses for the ray contribution from  $RP_0$ . From the conservation of energy flux in a ray tube,  $SF$  can be obtained as

$$SF = \sqrt{A_0/A} \quad (6)$$

where  $A_0$  and  $A$  are the cross-sectional areas at  $RP_0$  and the filed point, respectively.



FDTD is presented by Kane Yee for numerical solution of Maxwell's Equations [28]. In FDTD, a space grid is introduced in which E-field components and H-field components at different directions are discretized, so each E-field component is encircled by four H-field components, and also each H-field component is encircled by four E-field components. Moreover, in time domain, E- and H-field components are also discretized arranging at regular intervals. According to FDTD, in a 3-D scenario, for instance, E-field component at direction x ( $E_x$ ) can be given by

$$E_x \Big|_{i,j+1/2,k+1/2}^{n+1/2} = \left( \frac{1 - \frac{\sigma_{i,j+1/2,k+1/2} \Delta t}{2\epsilon_{i,j+1/2,k+1/2}}}{1 + \frac{\sigma_{i,j+1/2,k+1/2} \Delta t}{2\epsilon_{i,j+1/2,k+1/2}}} \right) E_x \Big|_{i,j+1/2,k+1/2}^{n-1/2} + \left( \frac{\frac{\Delta t}{\epsilon_{i,j+1/2,k+1/2}}}{1 + \frac{\sigma_{i,j+1/2,k+1/2} \Delta t}{2\epsilon_{i,j+1/2,k+1/2}}} \right) \bullet \left( \frac{H_z \Big|_{i,j+1,k+1/2}^n - H_z \Big|_{i,j,k+1}^n}{\Delta y} - \frac{H_y \Big|_{i,j+1/2,k+1}^n - H_y \Big|_{i,j+1/2,k}^n}{\Delta z} - J_{SOURCE_x} \Big|_{i,j+1/2,k+1/2}^n \right) \quad (7)$$

where  $n$  denotes time,  $\epsilon$  and  $\sigma$  are material constants,  $E$  and  $H$  represent values of E- and H-field, respectively, and propagation directions of EM waves are denoted by x, y and z in a rectangular coordinate system.  $\Delta x$ ,  $\Delta y$  and  $\Delta z$  are space increments in Yee grid,  $i$ ,  $j$  and  $k$  are parameters to describe distances at direction x, y, and z, respectively.  $\Delta t$  is the time increment. Analogously, components  $E_y$ ,  $E_z$ ,  $H_x$ ,  $H_y$  and  $H_z$  can be obtained with Yee grid through FDTD technology. Therefore, with previous time field, another field at adjacent nodes, electric and magnetic current sources, values of E- and H-field can be achieved at arbitrary point in Yee grid.

However, in FDTD, the space grid size must be a fraction of the wavelength for ensuring that over one increment the EM field does not change significantly. That means modeling indoor wave propagation in a larger area could be very difficult using existing equipment and the computation amount would be huge. Balancing between exactness and feasibility, a combination of ray-tracing and FDTD is introduced.

In a scenario of indoor environment illustrated by Fig. 4, the wave propagation between AP and Location D is investigated. The combination of ray-tracing and FDTD is that the propagation between AP and B, C and D is analyzed based on ray-tracing method, and the propagation between B and C is calculated on the basis of FDTD technique, similarly with the algorithm proposed in [29]. Every wall of the building is discretized into a unit cell (brick), Tx and Rx are added. The interaction between Tx and the bricks can be calculated iteratively according to (8), (9) and (10), and to simplify the explanation, vertical polarization with the three field components  $E_z$ ,  $H_x$  and  $H_y$  is assumed. The derivation for horizontal polarization is analogous.

$$\vec{E} = \frac{\omega\mu I_0}{4} H_0^{(2)}(k|\vec{r} - \vec{r}'|) \bullet \hat{z} \quad (8)$$

where  $H_0^{(2)}$  is the Hankel function of second kind,  $\vec{r}$  and  $\vec{r}'$  denote the vectors to Tx and to the brick point, respectively.  $k$  is the wavenumber.

$$\vec{M}_s = 2\vec{E}_s \times \vec{n} \quad (9)$$

where  $\vec{E}_s = E_s \hat{z}$  are the electric surfaces fields in the z-direction, and  $\vec{n}$  is the surface normal vector of the brick.

$$\vec{E} = \frac{k}{4j} \bullet \int 2E_s \bullet H_0^{(2)\gamma}(k|\vec{r} - \vec{r}'|) \bullet \frac{\vec{n} \bullet |\vec{r} - \vec{r}'|}{|\vec{r} - \vec{r}'|} d\vec{l} \bullet \hat{z} \quad (10)$$

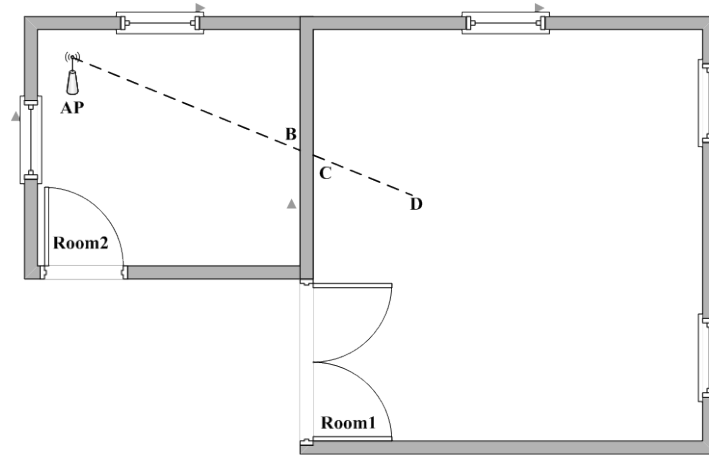


Fig. 4. Illustration of modeling indoor wave propagation.

And the analogous combinations of ray-tracing and FDTD have already been proved in [29-31]. Thus, for obtaining APs' coverage, it is unnecessary to configure all the APs within the WLAN to the same channel as in SEAR, and the experience of clients in the WLAN is guaranteed. In our algorithm, with the indoor propagation model, the central controller would update the information about the radio map in the WLAN which can be estimated and AP clusters will be formed on the basis of this data set.

#### 4.2 The Faster Procedure Based on K-means Algorithm

As mentioned above, NGCA could form green clusters based on access coverage of APs which is needed to be evaluated in an indoor environment. One simple but quick method for supplying coverage is based on the Friis free space equation, the other one which is a hybrid method consisting of ray-tracing and FDTD is more complex, time-consuming but accurate. Therefore, in our algorithm, two parallel procedures are introduced. One, named the fast procedure, is for forming coarse results about AP clusters in a relatively short time. And the other one, named the accurate procedure, is for achieving a more accurate AP deployment. Moreover, both procedures would work on the radio map built by modeling methods discussed in Subsection 4.1. In order to obtain a solution fast,  $K$ -means technique is employed.

$K$ -means algorithm is a kind of simple and efficient clustering approach. With  $K$ -means a  $D$ -dimension data set consisting of  $N$  samples can be divided into  $K$  categories by minimizing a criterion function  $J$  which be expressed by

$$J = \sum_{n=1}^N \sum_{k=1}^K r_{nk} \|\mathbf{x}_n - \boldsymbol{\mu}_k\| \quad (11)$$

where  $r_{nk}$  is a clustering index,  $r_{nk} = 1$  and  $r_{nj} = 0$  for  $j \neq k$ .  $\boldsymbol{\mu}_k$  is the cluster center of the  $k$ th cluster and  $\|\bullet\|$  is an operator that calculates Euclid distance between  $\mathbf{x}_n$  and  $\boldsymbol{\mu}_k$ .

In order to apply  $K$ -means the number of categories  $K$  should be given, and the first step of  $K$ -means is to initialize a cluster center for each category at random.  $K$ -means consists of E-step and M-step. In E-step, on the basis of  $r_{nk}$ ,  $\boldsymbol{\mu}_k$  is generated following optimization criterion expressed by  $\boldsymbol{\mu}_k = \frac{\sum_n r_{nk} \mathbf{x}_n}{\sum_n r_{nk}}$ . In M-step, according to  $\boldsymbol{\mu}_k$ ,  $r_{nk}$  is generated following optimization criterion.  $K$ -means iterates E-step and M-step until the value of the criterion function  $J$  is smaller than a very small number  $\xi$ , and  $\boldsymbol{\mu}_k$  is the ultimate cluster center of the  $k$ th category.

The faster procedure of our algorithm is based on  $K$ -means, each AP in the WLAN is regarded as the sample to be classified while the trade-off of this approach is the number of categories  $K$  which is problem-specific.

To begin the procedure we choose a bigger  $K$  and the method is discussed in the next section where we present a specific scenario (in Subsection 5.1).

The faster procedure of NGCA is the following.

- (1) Initialization:
  - (a) Obtain the modified positions of all APs in the WLAN.
  - (b) Calculate and determine the number of categories  $K$ .
- (2) Generate cluster centers for  $K$  categories at random.
- (3) Apply  $K$ -means algorithm, and obtain the results of clustering expressed as  $R_K$ .
- (4) **WHILE** the condition of network coverage is met:
  - (a) Step through all the samples in  $R_K$  and attempt to remove a single AP while keeping others APs reside in  $R_K$ , and check the condition of network coverage. The condition of network coverage is the constrain condition of our new algorithm. In order to fulfill the constrain condition, we mesh the region, calculate the area covered by cluster-head APs, and compare with the total area. If the difference between the two areas is a very small number, we consider the condition of network coverage is met.
  - (b) **IF** any of the conditions calculated is met, **THEN** remove the corresponding AP from  $R_K$  and increase the loop counter.

**IF** all of the conditions calculated are not met, **THEN GOTO** Step 5.
- (5) **END WHILE**
- (6) After removing some APs,  $R_K$  is the result obtained with the faster procedure.

### 4.3 The More Accurate Procedure Based on EA

In this procedure, we form AP clusters in another different perspective from both the approaches in our faster algorithm and in SEAR. In the faster procedure, we obtain results of green clustering with a bigger number first, and we express the results consisting of a set of cluster-head APs as  $R_K$ . Then we remove APs from  $R_K$ . In term of the number of trail cluster-head APs, the more accurate step is an opposite procedure during which we begin the approach with a smaller number of clusters, and then step up the number, while in the faster procedure a larger number is pre-determined and the number is stepped down.

#### 1) Problem description

In the more accurate procedure, we regard the area covered by the WLAN as the boundary for feasible solutions, and a set of trial cluster-head APs as a feasible solution  $R$ . And corresponding APs' physics positions express as  $Loc$ . The problem that green clustering attempts to handle can be addressed and define as

$$\begin{aligned} & \arg \min \left\{ \sum_{i \neq j} (S_i \cap S_j) \right\} \\ & \text{s. t. } \bigcup_{i=1}^n S_i \geq S_{\text{basic}} \end{aligned} \quad (12)$$

where  $S_i$  and  $S_j$  are the cover area of the  $i$ th AP and the  $j$ th AP, respectively, which are both in the set of a feasible solution, and  $i \neq j$ .  $\sum_{i \neq j} (S_i \cap S_j)$  represents overlap area, and  $S_{\text{basic}}$  represents the condition of network coverage.

We adapt the basic problem and transform it into a matter of multiple objective optimization which can be expressed as

$$\begin{aligned} & \text{Minimum } F = [\text{objfunc1}(Loc), \text{objfunc2}(Loc)] \\ & \text{subject to } R \in \Omega \end{aligned} \quad (13)$$

where  $R$  is a feasible solution consisting of a set of trial cluster-head APs,  $\text{objfunc1}(Loc)$  denotes the function that calculates the coverage with  $Loc$ . As this is a minimizing process the results  $Scale$  is handled by  $\text{abs}(1 - Scale)$  which means the area not covered by  $R$ . While  $\text{objfunc2}(Loc)$  represents the function that calculates the overlap area,  $\Omega$  denotes the region of the WLAN, and  $R \in \Omega$  means that the location of each AP in  $R$  is reasonable.

#### 2) EA in the new algorithm.

We adapt the basic problem and transform it into a matter of multiple objective optimization [32]. Similar with the faster procedure, the more accurate procedure forms green clusters also based on the radio map. The methods to build a radio map here are measuring and simulating, but in both methods RSS values obtained at reference points are pre-determined. According to (13), the working mechanism to solve the transformed problem would be optimizing processes over locations, so the selection of step-size would be critical for optimization and convergence time. Therefore, an optimization algorithm which could support multi-step sizes should be introduced. In general, different step sizes should be applied to different data sets, and most methods implement multi-step size mechanisms through independent calculation processes

rather than one execution procedure, so there would be no relations amongst results obtained with different step sizes during calculating. However, interactions among results are also critical for optimization and convergence time. In order to handle the new problem, we adapt and apply Evolutionary Algorithm (EA) in which multi-step sizes can be involved to subpopulations, and interactions among results can be implemented through Reinsertion, Migration and Competition amongst subpopulations [33].

EA is a stochastic search method that mimics the metaphor of natural biological evolution and models natural processes, such as selection, recombination, mutation, migration, locality and neighborhood [34]. To handle the new problem we adapt some processes in EA, and introduce a multi-population model. The structure of multi-population EA is shown in Fig. 5.

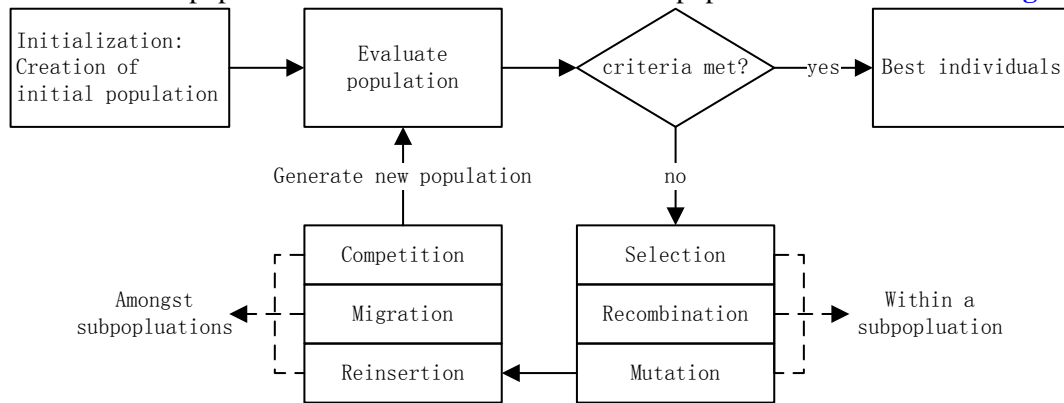


Fig. 5. The structure of multi-population EA.

In single-objective problems, rank-based fitness assignment is involved, but in our problem there are two criteria to be considered in order to evaluate the quality of an individual. The superiority of one feasible solution over the other can be decided by comparing the two solutions. It can be carried out following Pareto Dominance.

**Definition 1** (Pareto Dominance): A vector  $\mu = (\mu_1, \dots, \mu_k)$  is said to dominate  $\mathbf{v} = (v_1, \dots, v_k)$  (denoted by  $\mu \circ \mathbf{v}$ ) if and only if  $\mu$  is partially less than  $\mathbf{v}$ , i.e.,  $\forall i \in \{1, \dots, k\}, \mu_i \leq v_i \wedge \exists i \in \{1, \dots, k\} \mu_i < v_i$ .

If in a set of feasible solutions which can be regard as Pareto Optimal Set, a vector is not dominated by any other vector, it is non-dominated. All the non-dominated vectors in the Pareto Optimal Set can form a front, and in the general case, it can be considered that these non-dominated points produce the Pareto Front. The purpose of applying multi-objective algorithm is to achieve the Pareto Front approximately.

In order to apply EA in NGCA for better performance, we adapt some processes which are shown as follow.

(1) The structure of EA. The population is divided into five subpopulations and in each subpopulation some parameters and methods of EA, such as mutation rate, mutation range and the approach of recombination, are accordingly different.

(2) Selection and recombination. In terms of the new problem the coverage is considered as a more important property than the overlap. However, it is difficult to assign them weights accurately. To deal with that, we introduce bias as an assistant approach to our method. We use a repository *REP* to store non-dominated vectors and the corresponding sets of APs. In most EAs, individuals to produce offspring are chosen from population, and offspring are produced

according to the rule expressed as [35]

$$Val_i = Val_i^{P1} \cdot a_i + Val_i^{P2} \cdot (1 - a_i) \quad i \in \{1, 2, \dots, n\} \quad (14)$$

where  $a_i$  is a scaling factor and chosen uniformly at random over an interval  $[-d, 1+d]$  for the  $i$ th dimension, and a value of  $d = 0.25$  ensures that the variable area of the offspring is the same as the variable area spanned by the variables of the parents statistically.  $Val_i$  is the variable of the  $i$ th dimension from offspring,  $Val_i^{P1}$  and  $Val_i^{P2}$  are the variables from Parent1 and Parent2, respectively. In the new algorithm, a scheme is introduced to apply bias. If the event of bias is triggered, offspring are produced according to the rule expressed as

$$Val_i = Val_i^P \cdot a_i + REP[h]_i \cdot (1 - a_i) \quad i \in \{1, 2, \dots, n\} \quad (15)$$

where  $Val_i^P$  is the variable of the parent chosen from population, and  $REP[h]$  is a value taken from the repository. As non-dominated vectors are stored in  $REP$ , we repeatedly bisect the range in each objective to obtain several hypercubes. Such hypercubes have as many components as objective functions. Those hypercubes containing more than one individual are assigned a fitness equal to  $x/n$ , where  $n$  is the number of individuals within the corresponding hypercube and  $x$  is a number (we use  $x=1$  here). We decrease the fitness of those hypercubes containing more individuals and it can be seen as a form of fitness sharing [36]. Then, we apply roulette wheel selection [37] and tournament selection [38] with the fitness values to select the hypercubes within which  $REP[h]$  is selected at random.

(3) Mutation. In each subpopulation different mutation rate and range are applied to maintain genetic diversity. At the beginning of the search, the mutation operator attempts to explore all the new individuals. Then, the mutation rate might be decreased in accordance with the number of generations. And, as a consequence, the mutation range is affected by the rate. The pseudo code of the mutation operator applied in the new algorithm is the following.

```

% individual = individual to be mutated
% dims = number of dimensions
% currentgen = current generation
% totalgen = total number of generations
% mutrate = mutation rate
% mutrange = mutation range
function mutation_operator( individual , dims , currentgen , totalgen , mutrate )
IF flip( (1 - currentgen / totalgen)1/mutrate2 ), THEN
    whichdim = random(0, dims - 1)
    mutrange = (upperbound[whichdim] - lowerbound[whichdim]) • (1 - currentgen / totalgen)1/mutrate2
    ub = individual[whichdim] + mutrange
    lb = individual[whichdim] - mutrange
    IF ub > upperbound[whichdim], THEN ub = upperbound[whichdim]
    IF lb < lowerbound[whichdim], THEN lb = lowerbound[whichdim]

```

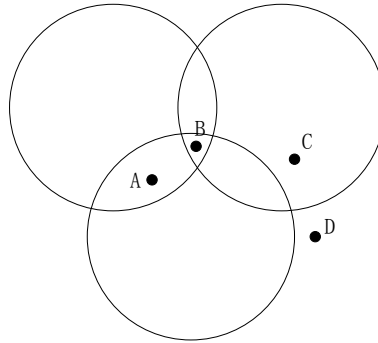
$individual^{NEW}[whichdim] = \text{RealRandom}(lb, ub)$

**END IF**

**END function**

3) The more accurate procedure based on EA.

We mesh the WLAN region to calculate the coverage and overlap, and more specifically, we represent the coverage and overlap of a grid area with the status of a point within it. As shown in Fig. 6, a dot circle represents the coverage area of an AP, and each point amongst A, B, C and D denotes a grid area, respectively. For example, the grid area denoted by A is covered by two APs all together, and the property of coverage and overlap within point A are 1 and 1, respectively. Similarly, the values are 1 and 2 for point B, 1 and 0 for point C, 0 and 0 for point D. The trade-off for meshing is the choice of the grid number which affects the operation speed directly, so we introduce the concept of resolution for meshing. With low and high resolutions, coarse and precise results can be obtained.



**Fig. 6.** Illustration of calculating coverage and overlap.

The more accurate procedure of NGCA is the following.

- (1) Initialization:
  - (a) Obtain the physical positions of all APs in the WLAN.
  - (b) Calculate and determine the original number of green clusters  $N$ , and set an upper bound  $Nub$  for  $N$ .
  - (c) Set a smaller number of total generations for EA expressed as  $totalgen1$ , a low resolution and a smaller ratio  $\xi_1$  which represents the uncovered area.
- (2) **WHILE**  $totalgen1$  has not been reached.
  - (a) Apply the modified EA and the objective values are calculated with the low resolution.
  - (b) Increase the counter of generations.
- (3) **END WHILE**
- (4) Evaluate the results with  $\xi_1$ .
  - (a) **IF** any of feasible solutions' coverage is no bigger than  $\xi_1$ , **THEN GOTO** step 5.
  - (b) **IF** all of feasible solutions' coverage is bigger than  $\xi_1$ , and  $N$  has not reached  $Nub$ , **THEN**  $N++$  and **GOTO** Step 2.
  - (c) **IF** all of feasible solutions' coverage is bigger than  $\xi_1$ , and  $N$  has reached  $Nub$ , **THEN** the algorithm fails. To avoid this situation, we recommend a bigger  $Nub$ , and in the new green clustering we handle this by combining the faster procedure, and we will discuss the approach in the next subsection.
- (5) Initialization for another EA processing:
  - Set a high resolution and a bigger number of total generations for the current EA

expressed as  $totalgen2$ , and  $totalgen2 \geq 2 \times totalgen1$  is recommended.

(6) **WHILE**  $totalgen2$  has not been reached.

(a) Apply the modified EA and the objective values are calculated with the high resolution.

(b) Increase the counter of generations.

(7) **END WHILE**

(8) Set a bigger ratio  $\xi_2$ , the value of  $\xi_2$  is problem-specific and  $\xi_2 \leq \xi_1 / 5$  is recommended.

With  $\xi_2$  the feasible solutions in  $REP$  whose first objective value (calculated by  $objfunc1(Loc)$ ) is no bigger than  $\xi_2$ , are acceptable. Once acceptable solutions are obtained, we choose the solution from the acceptable solutions which is most close to the acceptable hyperplane  $ObjVal1 = \xi_2$  as the ultimate feasible solution

#### 4.4 A New Green Clustering Algorithm

The new green clustering algorithm (NGCA) consists of modeling the environment, analysis of APs' coverage, the faster procedure and the more accurate procedure. For an illustrative purpose, the structure and operation mechanism of the proposed method are shown in Fig. 7.

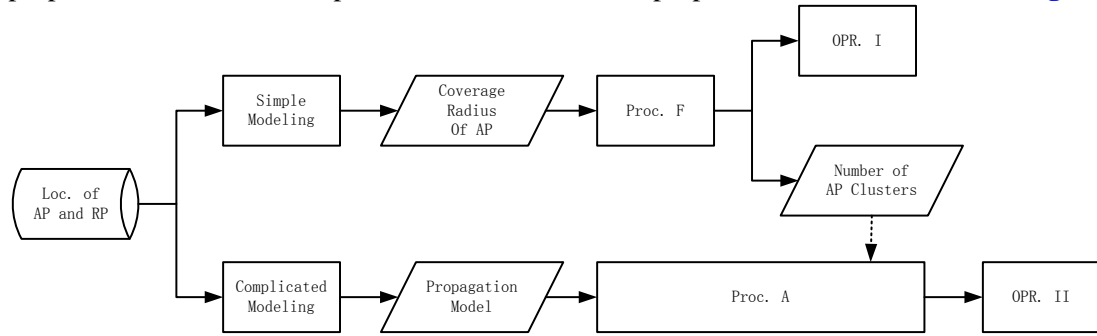


Fig. 7. The structure and operation mechanism of NGCA.

In Fig. 7, Loc. of AP and RP denotes the data base which stores physical positions of APs and RPs (reference point), Proc. F is short for the faster procedure, and Proc. A is short for the more accurate procedure. Once Proc. F has been executed, results  $R_F$  consisting of feasible APs can be obtained. Similarly, the results of Proc. A expressed as  $R_A$  can also be obtained, and the APs in  $R_A$  are the ultimate green cluster-head APs. In Fig. 7, OPR. I denote operations of increasing the transmitting power of APs in the set of  $R_F$  and then powering-off APs in  $R_F^\delta$ , and OPR. II denote operations of powering-on APs with maximum transmitting power in  $R_F^\delta \cap R_A$  and then powering-off APs in  $R_F \cap R_A^\delta$ . After OPR. II, cluster-head APs are formed, and the results of NGCA are obtained.

Once Proc. F obtains results, information interaction occurs between Proc. F and A. Proc. F can determine the number of clusters, this number is set as the upper bound of green cluster number used in Proc. A, and as shown in Fig. 7 the processing of interaction expresses as the vertical dot line from the results of Proc. F towards Proc. A.

Going through procedures and operations in NGCA, the green clusters can be obtained in a different but new perspective.



## 5. Simulation Results and Analysis of New Green Clustering Algorithm

In this section we carry on simulations to verify the new green clustering algorithm (NGCA) proposed in this paper. Firstly, an ideal model is introduced and simulation is conducted in this model, then we model our office as an actual scenario and run NGCA on it. Besides, the green clustering algorithm in SEAR is presented as a comparison in the last subsection.

### 5.1 Simulation and Analysis in the Ideal Model

The deployment of APs in an ideal enterprise (where deployed high-density WLAN) is modeled for our simulation experiment first. The size of experimental region is  $100\text{ m} \times 100\text{ m}$ , and the deployment of APs is shown in Fig. 8 where an empty circle denotes an AP and the number nearby is the label of the corresponding AP. According to (1), the value of  $\mu_{AP}$  in our simulation experiment is 0.0081 and deploying APs (where distance in between is 10m) can be considered as a high-density WLAN. As experimental setup, we need to set an indoor maximum coverage radius of AP expressed by  $Rad$ . Assuming that the Rx sensitivity is  $-76\text{dBm}$ , the maximum indoor propagation distance is about 31.62m according to (3), and in the simulation experiment of this subsection,  $Rad$  is set to 30m.

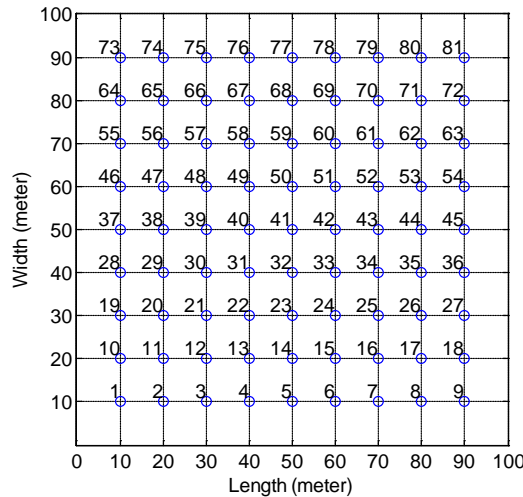


Fig. 8. Deployment of APs in simulation experiment.

In Proc. F,  $K$ -means is applied and the number of original categories  $K$  is determined as following. In ideal conditions, to cover an experiment area which is  $100\text{ m} \times 100\text{ m}$ , only

$$N_0 = \text{ceil} \left( \frac{100^2}{\pi \cdot Rad^2} \right) = 4 \text{ APs are needed, though it is unreachable in practice. Supposing that}$$

the area as a piece of paper, and folding the paper in half horizontally and vertically in turn. Then repeat these processes until the folded area can be fully covered by one single AP with  $Rad$ , and denoting the times of folding is  $n$ ,  $2^n$  is the number we looking for. Therefore,  $K$  is 16 for our simulation.

In Proc. A, as discussed in the above section, the upper bound of green cluster number is derived from the results of Proc. F. The first objective value  $\xi_1$  for low resolution is set to 0.01, and  $\xi_2$  for high resolution is set to 0.002. And the low resolution is set to  $100 \times 100$  which

means the experimental region is divided into  $100 \times 100$  meshes, while the high resolution is set to  $1000 \times 1000$ .

According to Fig. 7, to simulate NGCA, the first step is to decide the coverage area of an AP. In the simulation, the AP deployment is shown as Fig. 8, and it is an experimental scenario in which no walls, doors or furniture exist, and that means the coverage can be estimated through the radius. Then, Proc. F and A would be carried on simultaneously. Proc. F completes faster than Proc. A, and the results of Proc. F are shown in Fig. 9.

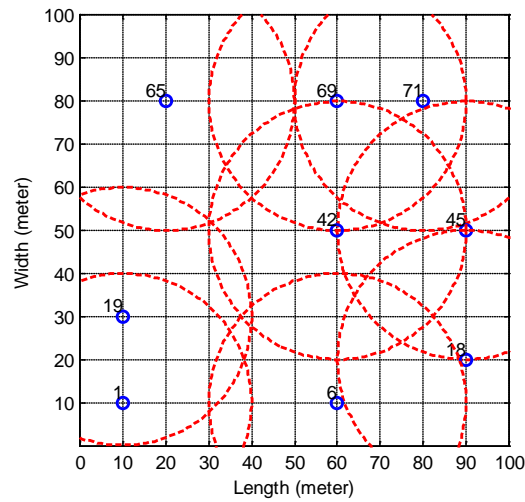


Fig. 9. Results of PROC. F, in which 9 APs are powered-on.

where a circle denotes a cluster-head AP, and the number nearby with the dotted line around it are the AP's label and its coverage border, respectively. Fig. 9 shows that to cover such a region 9 APs are needed according to Proc. F. At this point, the number of clusters obtained is passed to Proc. A as the upper bound of green cluster number. After passing parameters, Proc. F completes, and Proc. A carries on. In EA, feasible solutions are stored in the repository *REP*, after the total generations are reached, the feasible solutions are shown in Fig. 10.

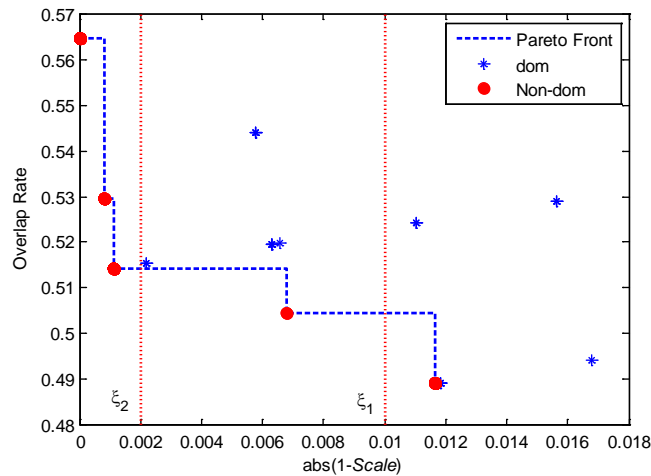
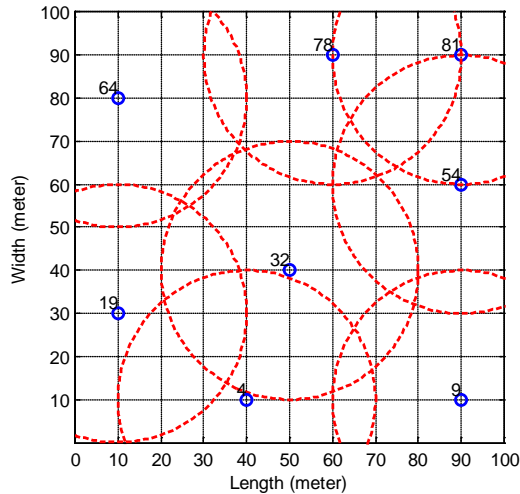


Fig. 10. Illustration of feasible solutions, the objectives are to minimize uncovered and overlap rates.

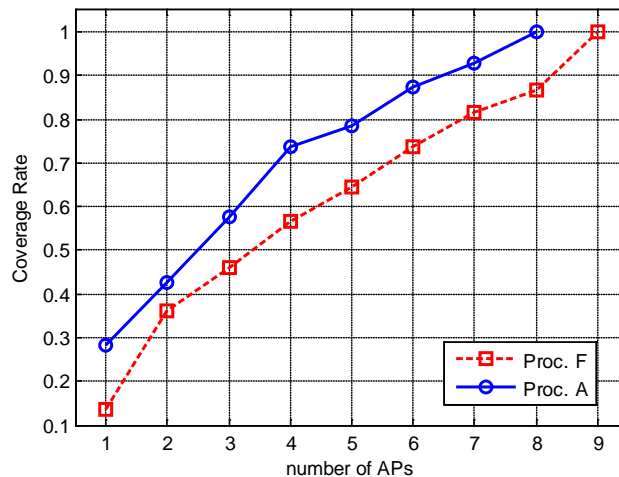
where a solid red circle denotes a non-dominated solution, a blue asterisk denotes a dominated solution and the dotted fold line is the approximate Pareto Front.  $\xi_1$  and  $\xi_2$  are the thresholds for our modified EA,  $x = \xi_1$  and  $x = \xi_2$  are acceptable lines. The non-dominated solutions which are on the left of  $x = \xi_2$  are acceptable solutions, and the one close to  $x = \xi_2$  most is the ultimate feasible solution which is shown in **Fig. 11**.



**Fig. 11.** Results of NGCA,

in which 8 APs are powered-on and 88.89% of energy can be saved.

According to **Fig. 11**, it shows that to cover such a region, 8 APs are needed deriving from Proc. A. Concerning of coverage rate, a comparison is introduced between Proc. F and A, and the results are shown in **Fig. 12**.



**Fig. 12.** Illustration of the comparison between Proc. F and A.

**Fig. 12** clearly illustrates that Proc. A maintains coverage with less APs than Proc. F. Thus far, Proc. A is finished and after some operations the new algorithm is completed.

In traditional WLANs, to cover our experimental area, all the 81 APs are powered-on. With the new approach, 8 APs are needed. As the first step of energy efficient strategy, about 90% energy is saved with NGCA.

## 5.2 Simulation and Analysis in the Actual Model

In the above subsection, simulations based on an ideal model have been conducted, and the validity and efficiency of NGCA have been proved. In this subsection a practical model will be introduced, and for verifying the proposed algorithm, 27 AP are deployed on the floor where our office resides as shown in Fig. 13. The area of our floor is about 1875 m<sup>2</sup>, and according to (1), the value of  $\mu_{AP}$  is 0.0144.

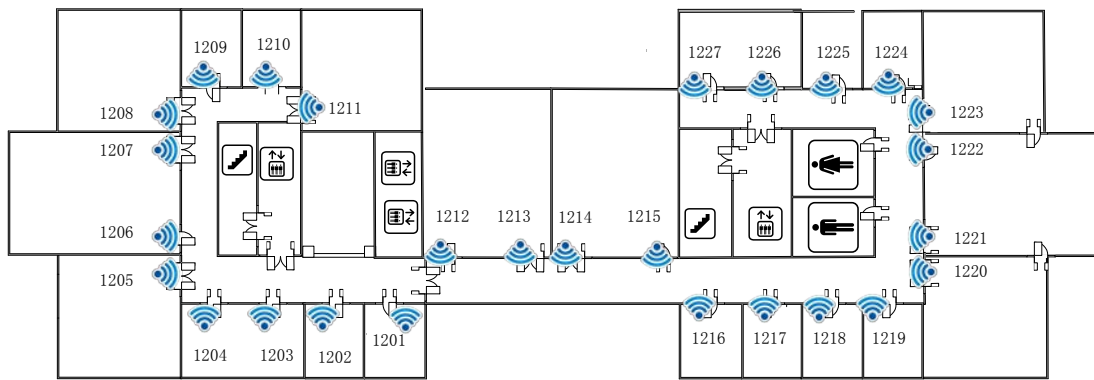


Fig. 13. The floor layout and AP deployment of our office, and the numbers denote doors.

The purpose of the proposed algorithm is that when few users are online or the traffic load is low, for energy efficiency NGCA classifies AP in the WLAN, and remains cluster-head APs on, powers off secondary APs while the basic network coverage of corridors and offices can be maintained. Unlike discussions in the above subsection, the model introduced to decide the coverage of an AP here is more complex and practical, so it is not appropriate to estimate the path loss with (4) and measure the coverage area of AP with a reference circle. Therefore, we select some locations on our floor as reference points (RP), measure and record values of Rx RSS at each RP as shown in Fig. 14, and with these data a radio map can be built. And, with either measured or estimated radio maps, the coverage region of an AP can be evaluated.

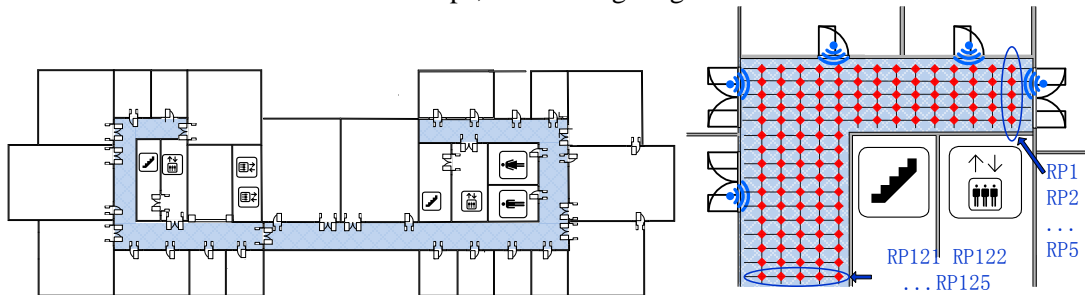
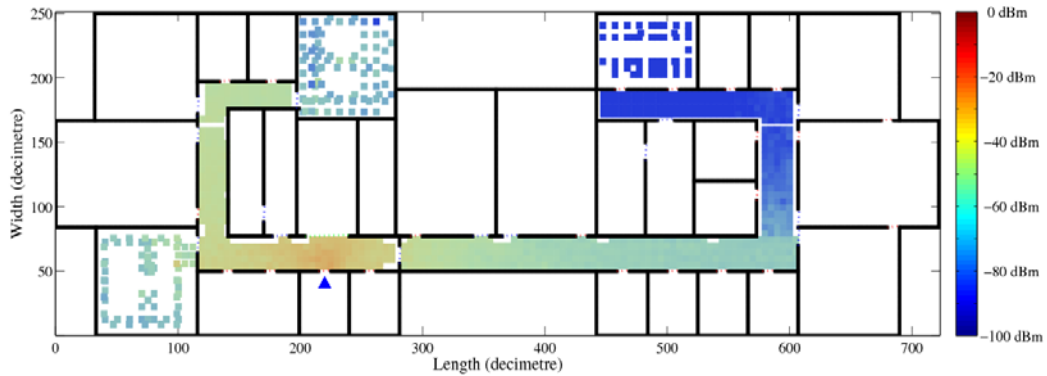
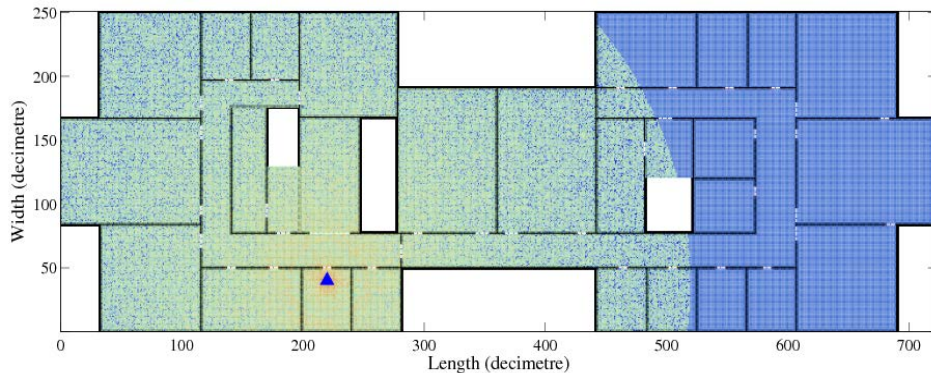


Fig. 14. Illustrations of RPs on corridors.



**Fig. 15.** Measured RSS at RPs from AP deployed in Room1202 with a sapling interval of 0.5 m.

For instance, **Fig. 15** illustrates RSS at RPs received from the AP deployed in Room1202, where values of RSS are represented through different colors, and the AP is denoted by a solid blue triangle. As the constraint of authorization, RSS values are measured and recorded in 3 dedicated offices, and RPs are set unbalanced as all offices are furnished with tables, chairs, computers and other laboratory furniture, which makes only sampling RPs in aisles are available, as shown in **Fig. 15**. But in corridors, RPs are setup homogeneously with a 0.5m sampling interval. Therefore, with the measured radio map, the coverage region of an AP can be evaluated. Moreover, for better accuracy and robustness, the method of complicated modeling is introduced. According to the investigation and discussion in Subsection 4.1, RSS at an arbitrary location on our floor from each AP can be estimated. In order to compare, RSS data of AP in Room1202 is demonstrated as the example shown in **Fig. 16**.

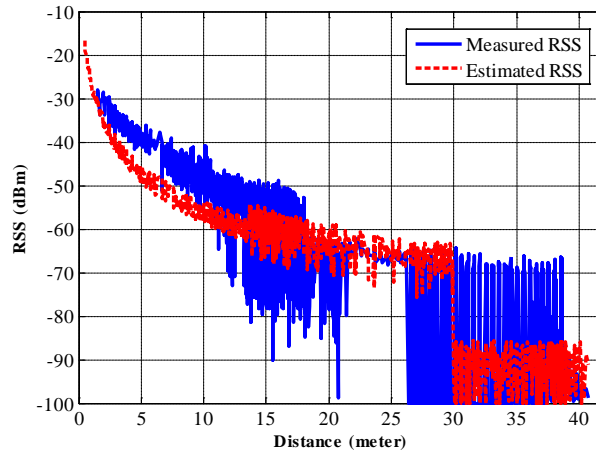


**Fig. 16.** Calculated RSS on our floor from AP deployed in Room1202 with a sampling interval of 0.1 m.

To be specific, RSS has been estimated at almost each location on our floor including all the offices, corridors, the hall, staircases and even W.C., and the 3 blank areas denote elevator shafts where the network coverage is meaningless. It is worth noticing that boundaries of dark blue which denote values of RSS less than -80dBm are arc-shaped. In order to reduce unnecessary calculation, it is assumed that if the distance between an AP and a RP exceeds 30m, the value of RSS at this RP from the AP is concerned as a value between -80 and -100dBm which is similar to the noise power setting in simulations, thus boundaries of RSS values less than -80dBm present arc-shaped.

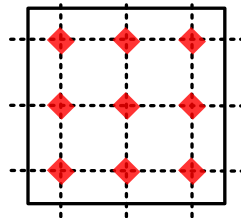
RSS measured and estimated from the AP deployed in Room1210 are shown in **Fig. 17** at different distances. The estimating method could provide more RSS information. Measured RSS changes severely, particularly when RSS values are smaller, but estimated RSS is more

stable. And with the increase of distance, the number of RPs at one same distance grows and wave propagation between the AP to each RP becomes more complex, such as the number of walls and doors to be through may be diverse. Due to the more diverse propagation environment, RSS in a small range could vary a lot, and the increased distance between AP and RP would contribute to this phenomenon which is proved by both blue and red curves of measured and calculated RSS.



**Fig. 17.** Comparison of measured RSS and estimated results, and in order to reduce unnecessary calculation, when the distance between RP and AP exceeds 30m, RSS value is assigned between -90 and -100dBm.

In the method of sampling model which is the preprocessing of Proc. F, estimation of AP's coverage is simply based on coverage radius. However, in the complicated modeling which is the preprocessing of Proc. A, estimation of AP's coverage is based on modeling indoor wave propagation. And in Proc. A, in order to determine if a pre-determined unit area is covered by some AP, the WLAN region is meshed to form several grids. As the value of sampling interval in modeling indoor propagation is smaller, the grid in modified EA can be expressed as relations amongst RPs within it (as shown in Fig. 18). For instances, if RSS from an AP at any RP in the grid is less than the coverage condition this grid would be consider as uncovered by this AP, or conversely, the grid is concerned to be covered if RSS at any RP meets the condition. As the area of a grid is relatively small in term of the WLAN region, the ways to decide a grid's network coverage have little impact on results of NGCA. And in our simulation, if the mean value of RPs' RSS in a grid meets the coverage condition, the gird would be regarded as being covered by the corresponding AP.



**Fig. 18.** Illustration of relations between RPs (denoted by solid red diamonds) built in complicated modeling, and the grid (denoted by the square) in modified EA.

Through the simple modeling and complicated modeling, the coverage of an AP can be

determined, and with the coverage information, NGCA can generate appropriate AP clusters. Similar with the above subsection, results of Proc. F is presented first.

In this actual model, the coverage radius  $Rad$  for Proc. F should decrease as the complex indoor environment, and  $Rad$  is set to 15m. According to the work mechanism of Proc. F, similar with the above subsection, results of coarse clusters are shown in Fig. 19.

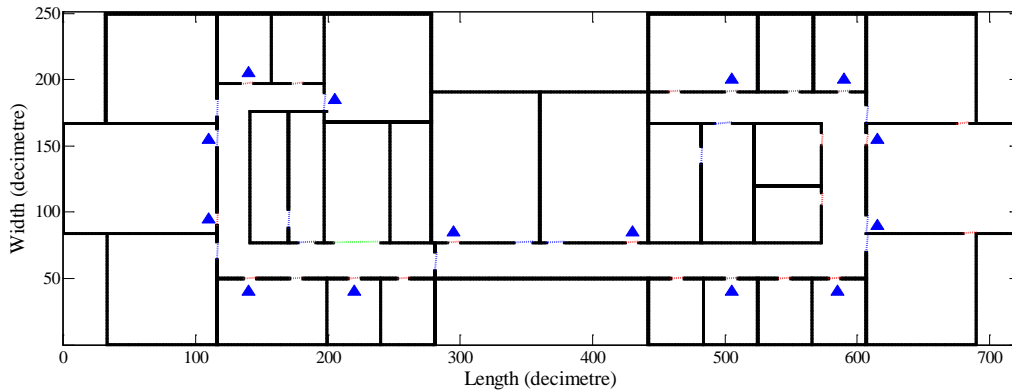


Fig. 19. Results of Proc. F, in which 14 APs are powered-on.

With the results obtained from Proc. F, the upper bound of AP clusters for Proc. A is set to 14, and it is assumed that simulation parameters for Proc. A in the actual model are as same as those in the ideal model. For each feasible solution which represents a set consisting of APs, the network coverage conditions should be calculated of which the results could be concerned as objects. And in this practical model, the network coverage conditions are evaluated through that RSS at all RPs in complicated modeling from each AP in the feasible solution is estimated. According to the RSS values RPs covered by this AP can be determined, and all APs in the feasible solution will be traverse to find the uncovered and overlap rates under the current conditions. Iterating the steps in accordance with the working mechanism of Proc. A, the results, which are also the ultimate solutions of NGCA, can be achieved, as shown in Fig. 20.

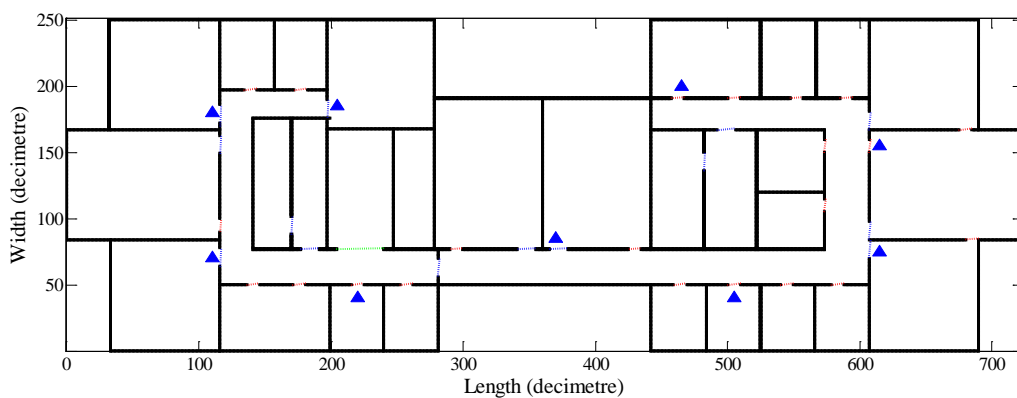
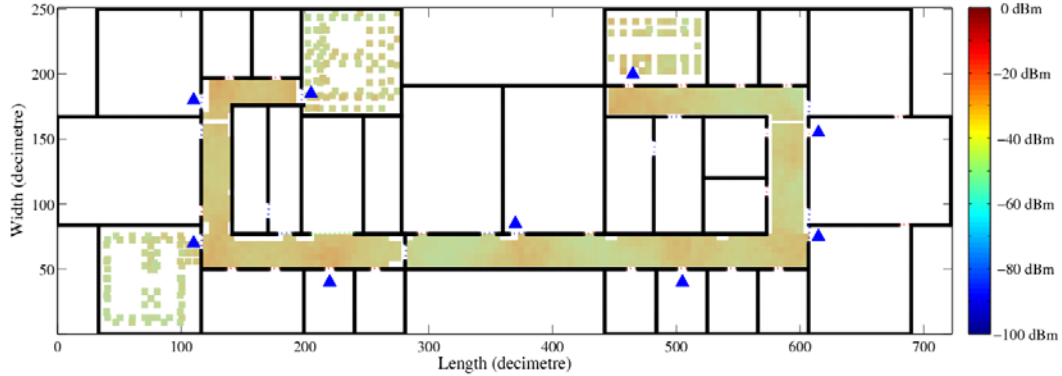


Fig. 20. Results of NGCA obtained in the actual model, in which 9 AP are powered-on and 66.7% of energy can be saved.

Results of NGCA show that 9 APs at locations specified by our algorithm are enough to provide basic network coverage in the scenario modeled as the floor where our office resides, so 18 AP can be powered-off for energy efficiency when few users are online or the traffic load is low, which means about 66.67% of power can be saved. In order to verify the results,

RSS values from these 9 APs are measured and recorded, and the results are shown in [Fig. 21](#).



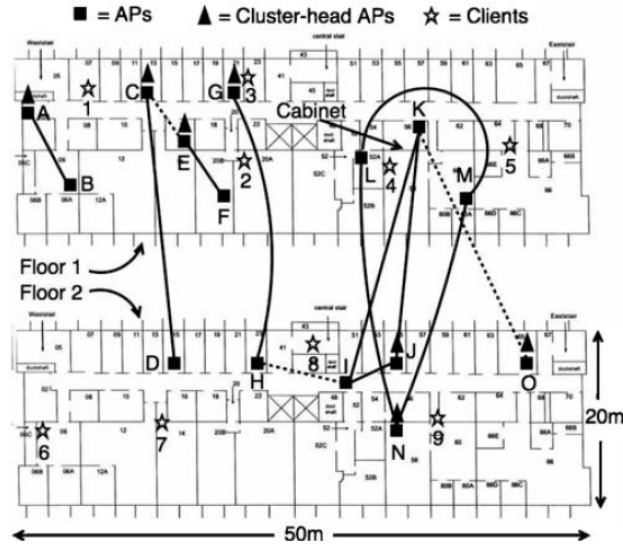
[Fig. 21](#). RSS measured on our floor, and only the maximal RSS at each RP is recorded.

In [Fig. 21](#), only the maximal RSS received among the 9 APs at each RP is recorded. Therefore, according to the measurement, the basic network coverage is maintained by the 9 APs NGCA specified when other APs are powered-off for energy efficiency.

### 5.3 Comparison and Analysis

In Subsection 5.1 and 5.2, simulations have been conducted to verify the validity and efficiency of NGCA in an ideal scenario and an actual scenario modeled as our offices, respectively, in the ideal model 88.89% of energy is saved, and 66.67% is saved in the practical model. Moreover, in this subsection, the green clustering algorithm in SEAR [\[5\]](#) is introduced and compared.

In [\[5\]](#), in order to implement green clustering, Jardosh conducts experiments in their building, and the building layout and AP deployment are shown in [Fig. 22](#).



[Fig. 22](#). The floor layout and AP deployment of the comparison task, where on 2 floors, each of which is 50 m × 20 m, 14 APs are deployed.

In [Fig. 22](#), 14 APs are deployed in their building, and 7 AP clusters are formed according to their green clustering algorithm, which means 50% of energy is saved, while the results of



NGCA are 88.89% in the ideal model and 66.67% in the actual model. Therefore, only in term of the saved energy ratio, our NGCA outperforms their algorithm, especially results in the ideal model. However, the comparison of ratios seems to be unfair and also not convincing due to the different indoor layouts and wave propagation environments. For clarifying issue about the disadvantage of only comparing saved energy ratios, a theoretical example is presented. It is assumed that in an office which is 10 m × 10 m, 100 APs are deployed there, and 99 APs can be powered-off. Therefore, 99% of energy is saved. However, in this assumption the number 99% cannot reflect the better performance as the scenario is too simple.

In order for fairness, a parameter which can describe the gain of algorithms in different indoor environments to a certain extend is introduced, and the parameter  $\theta$  can be defined as

$$\theta = \frac{r_{\text{SAVED}}}{\mu_{\text{AP}}} \bullet n \quad (16)$$

where  $r_{\text{SAVED}}$  denotes the ratio of saved energy,  $\mu_{\text{AP}}$  is the parameter in (1) to describe the density of APs, and  $n$  is the attenuation factor in (3) to demonstrate different environments. So according to (16),  $\theta$  can be calculated in different scenarios, and a bigger value of  $\theta$  denotes the better performance of an algorithm, so  $\theta$  can be introduced as the algorithm gain to measure the performance of different green clustering algorithms. Firstly, we begin with  $\theta$  in the theoretical model.

In the theoretical model,  $\mu_{\text{AP}}$  is 1, and as in a simple wave propagation environment  $n$  is set to 3, so  $\theta$  is 2.97.

In the experiment environment of SEAR, 14 APs are deployed on 2 floors. In order to calculate  $\mu_{\text{AP}}$ , the area of experiment  $S$  should be determined first. It is considered that  $S$  is assigned with only one floor's area, as APs in our simulations can also provide network coverage for upstairs and downstairs, so for uniformity, only the area of one floor is taken into consideration for each simulation experiment. Therefore,  $\mu_{\text{AP}}$  here is 0.014, and  $n$  is set to 3.5, which makes  $\theta$  equal to 125.

And then, referring to our simulation in the ideal model,  $r_{\text{SAVED}}$  is 88.89%,  $\mu_{\text{AP}}$  is 0.0081, and  $n$  is set to 3, so  $\theta$  is 329.22 which is bigger than  $\theta$  of SEAR.

Finally,  $\theta$  of our simulation in the actual model is calculated.  $r_{\text{SAVED}}$  is 66.67%,  $\mu_{\text{AP}}$  is 0.0144 which is similar to the AP density in SEAR, and  $n$  is 3.5 as same as the attenuation factor in SEAR, and the value of  $\theta$  is 162.05 which is also bigger than  $\theta$  of SEAR.

According to these values of  $\theta$  which are environment-specified, the unfairness of comparing only ratios of saved energy can be proved and surmounted, and both in terms of the saved energy ratio and algorithm gain, the performance of NGCA is better provably.

## 6. Conclusion

High-density WLANs are deployed in more and more enterprise offices and college campuses at present. To ensure network coverage, all the APs in the WLANs are powered-on and this causes huge energy wastage. In this paper, the energy efficient strategy is discussed and a new green clustering which is an innovative algorithm of the strategy is proposed.

In this paper, we study the energy efficient strategy and the green clustering algorithm in SEAR. Theoretically, the green clustering in SEAR is a kind of greedy algorithm. And it has to

configure all APs to a same channel compulsively and keeps status for more than 1 minute in order to apply those APs within the WLAN, even while the APs may be providing connectivity to clients meantime, which could severely impact the experience of terminals.

While in the new green clustering algorithm, with a different method to form AP clusters, it is unnecessary for the central controller to configure all APs to a same channel initially, which, to same context, guarantee the performance and facilitate the end users.

Then, we propose a new green clustering algorithm (NGCA) in a completely different perspective. The two main components of our approach are the faster procedure (Proc. F) based on  $K$ -means and the more accurate procedure (Proc. A) based on EA. The two procedures are processes in parallel for different designed requirements, and meantime there is information interaction in between. In Proc. A, for better performance in the new algorithm, we adapt Evolutionary Algorithm. Population is averagely divided into 5 subpopulations for applying different parameter settings, and a repository *REP* is introduced to store feasible solutions obtained by EA. With *REP* additional operations for selection and recombination is introduced, and also a new operator for mutation is applied.

Finally, simulations of NGCA are clearly presented in both ideal model and actual model, besides the comparison with an existing green algorithm is introduced as well. The results of the simulations verify the feasibility of NGCA and show that the new approach can save about 90% energy consumption in the ideal scenario and 67% in the actual scenario meanwhile coverage is maintained during periods when few users are online or the traffic load is low.

The green clustering algorithm is proposed as the first method of energy efficient strategy in high-density WLANs in this paper. The algorithm will be the foreshadowing for future study about energy efficient strategy in which the optimization of cluster-head APs' transmitting power will be concerned, and NGCA will be supplemented comprehensively further. After green clustering, the mechanism of estimating user demand based on Markov Chain Model and queuing theory will also be studied, and an access selection based on IEEE 802.11 PCF for energy efficiency in high-density WLANs will be proposed in the near future.

## References

- [1] Aruba selected by Microsoft for next generation wireless LAN (2005). <http://www.arubanetworks.com/news/release/2005/06/13>.
- [2] Dartmouth College WLAN. <http://crawdad.cs.dartmouth.edu>.
- [3] <http://muniwifi.org>.
- [4] A. P. Jardosh, G. Iannaccone, K. Papagiannaki, and B. Vinnakota, "Towards an energy-star WLAN infrastructure," in *Proc. of HotMobile*, vol. 15, pp. 85-90, February 26-27, 2007. [Article \(CrossRef Link\)](#).
- [5] A. P. Jardosh, K. Papagiannaki., E. M. Belding, K. C. Almeroth, G. Iannaccone, and B. Vinnakotan, "Green WLANs: On-demand WLAN infrastructures," *Mobile Networks and Applications*, vol. 14, no. 6, pp. 798-814, December, 2009. [Article \(CrossRef Link\)](#).
- [6] J. H. Holland, "Adaptation in natural and artificial systems", in *The University of Michigan Press*, 1975.
- [7] D. Alessandro, S. Moret, and N. Tonello, "Green hybrid FMT for WLAN applications," in *Proc. of WD*, pp. 1-5, October 20-22, 2010. [Article \(CrossRef Link\)](#).
- [8] L. Miliotis, V. Apostolaras, A. Korakis, T. Tao, and Z. Tassioulas, "New channel allocation techniques for power efficient WiFi networks," in *Proc. of PIMRCW*, pp. 347-351, September 26-30, 2010. [Article \(CrossRef Link\)](#).
- [9] J. Rabaey, "The standby power challenge: Wake-up receivers to the rescue," in *Proc. of VLSI-TSA*, pp. 42, April 27-29, 2009. [Article \(CrossRef Link\)](#).
- [10] M. A. Marsan, L. Chiaraviglio, and D. Ciullo, "A simple analytical model for the energy-efficient

- activation of access points in dense WLANs,” in *Proc. of ICEEC-NE*, pp. 159-168, April 13-15, 2012. [Article \(CrossRef Link\)](#).
- [11] Y. Chen, Q. Yang, J. Yin, and X. Chai, “Power-efficient access-point selection for indoor location estimation,” *IEEE Trans. on Knowledge and Data Engineering*, vol. 18, no. 7, pp. 877-888, 2006. [Article \(CrossRef Link\)](#).
- [12] A. Kushki, K. Plataniotis, and A. Venetsanopoulos, “Kernel-based positioning in wireless local area networks,” *IEEE Trans. on Mobile Computing*, vol. 6, no. 6, pp. 689-705, 2007. [Article \(CrossRef Link\)](#).
- [13] A. Kushki, K. N. Plataniotis, and A. N. Venetsanopoulos, “Intelligent dynamic radio tracking in indoor wireless local area networks,” *IEEE Trans. on Mobile Computing*, vol. 9, no. 3, pp. 405-419, 2010. [Article \(CrossRef Link\)](#).
- [14] S. H. Fang, T. N. Lin, and P. C. Lin, “Location fingerprinting in a decorrelated space,” *IEEE Trans. on Knowledge and Data Engineering*, vol. 20, no.5, pp. 685-691, 2008. [Article \(CrossRef Link\)](#).
- [15] Y. Kang, S. Lim, J. Yoo, and C. Kim, “Design, analysis and implementation of energy-efficient broadcast MAC protocol for wireless sensor networks,” *KSII Trans. on Internet and Information Systems*, vol. 5, no. 6, pp. 1113-1132, June, 2011. [Article \(CrossRef Link\)](#).
- [16] S. Lim, Y. Kang, J. Jeong, and C. Kim, “Design, analysis and evaluation of a new energy conserving MAC protocol of wireless sensor networks,” *KSII Trans. on Internet and Information Systems*, vol. 6, no. 12, pp. 3046-3060, December 31, 2012. [Article \(CrossRef Link\)](#).
- [17] Y. Zuo, Z. Ling, and L. Liu, “Energy-efficient low-delay TDMA scheduling algorithm for industrial wireless mesh networks,” *KSII Trans. on Internet and Information Systems*, vol. 6, no. 10, pp. 2509-2528, 2012. [Article \(CrossRef Link\)](#).
- [18] D. W. Kim ,and T. Park, “An energy efficient MAC protocol providing guaranteed service for wireless sensor network,” *KSII Trans. on Internet and Information Systems*, vol. 5, no. 1, January, 2011. [Article \(CrossRef Link\)](#).
- [19] S. Kang, S. Lee, S. Ahn, and S. An, “Energy efficient topology control based on sociological cluster in wireless sensor networks,” *KSII Trans. on Internet and Information Systems*, vol. 6, no. 1, pp. 361-380, January, 2012. [Article \(CrossRef Link\)](#).
- [20] J. Oak, Y. Choi, W. Park, “EP-MAC: Early preamble MAC to achieve low delay and energy consumption in duty cycle based asynchronous wireless sensor networks,” *KSII Trans. on Internet and Information Systems*, vol. 6, no. 1, pp. 2980-2991, November, 2012. [Article \(CrossRef Link\)](#).
- [21] Y. Zhang, H. Long, Y. Peng, K. Zheng, and W. Wang, “User-oriented energy-and spectral-efficiency tradeoff for wireless networks,” *KSII Trans. on Internet and Information Systems*, vol. 7, no. 2, pp. 216-233, February, 2013. [Article \(CrossRef Link\)](#).
- [22] J. Zhu, “Energy efficiency analysis of cellular downlink transmission with network coding over Rayleigh fading channels,” *KSII Trans. on Internet and Information Systems*, vol. 7, no. 3, pp. 456-468, March, 2013. [Article \(CrossRef Link\)](#).
- [23] Y. Wu, S. Deng, and H. Huang, “Energy-efficiency joint control of epidemic routing in delay tolerant networks,” *KSII Trans. on Internet and Information Systems*, vol. 7, no. 2, pp. 234-252, February, 2013. [Article \(CrossRef Link\)](#).
- [24] M. Eslaminejad, S. A. Razak, and A. S. H. Ismail, “Eedars: An energy-efficient dual-sink algorithm with role switching mechanism for event-driven wireless sensor network,” *KSII Trans. on Internet and Information Systems*, vol. 6, no. 10, pp. 2473-2492, 2012. [Article \(CrossRef Link\)](#).
- [25] Forrester research (2006). <http://www.forrester.com>.
- [26] Y. Bejerano, “Efficient integration of multi-hop wireless and wired network with QoS constrains”, in *IEEE/ACM Trans. on Networking*, vol. 12, no. 6, pp. 1064-1078, 2002. [Article \(CrossRef Link\)](#).
- [27] C. F. Yang, and C. J. Ko, “A ray tracing method for modeling indoor wave propagation and penetration,” *IEEE Trans. on Antennas Propagat.*, vol. 46, pp. 907-919, 1998. [Article \(CrossRef Link\)](#).
- [28] K. S. Yee, “Numerical solution of initial boundary value problems involving Maxwell's Equations in isotropic media,” *IEEE Trans. on Antennas Propagat.*, vol. 14, no. 3, pp. 302-307, 1966. [Article \(CrossRef Link\)](#).

- [29] M. Thiel, and K. Sarabandi, "A hybrid method for indoor wave propagation modeling," *IEEE Trans. on Antennas and Propagat.*, vol. 56, no. 8, pp. 2703-2708, 2008. [Article \(CrossRef Link\)](#).
- [30] Y. Wang, S. Safavi-Naeini, and S. K. Chaudhuri, "A hybrid technique based on combing ray tracing and FDTD methods for site-specific modeling of indoor radio wave propagation," *IEEE Trans. on Antenna Propagat.*, vol. 48, no. 5, pp. 743-754, 2000. [Article \(CrossRef Link\)](#).
- [31] Y. Wang, S. K. Chaudhuri, and S. Safavi-Naeini, "An FDTD/ray-tracing analysis method for wave penetration through inhomogeneous walls," *IEEE Trans. on Antenna Propagat.*, vol. 50, no. 11, pp. 1598-1604, 2002. [Article \(CrossRef Link\)](#).
- [32] K. Deb, "Multi-objective optimization using Evolutionary Algorithm", in *Wiley: Chichester*, 2001.
- [33] D. B. Fogel, "An introduction to simulated evolutionary optimization," *IEEE Trans. on Neural Networks: Special Issue on Evolutionary Computation*, vol. 5, no. 1, pp. 3-14, 1994. [Article \(CrossRef Link\)](#).
- [34] R. L. Haupt, and S. E. Haupt, "Partical genetic algorithm", in *John Wiley & Sons: New York*, 1998.
- [35] H. Muhlenbein, and D. Schlierkamp-Voosen, "Predictive models for breeder genetic algorithm," *Evolutionary Computation*, pp. 25-49, 1993. [Article \(CrossRef Link\)](#).
- [36] K. Deb, and D. E. Goldberg, "An investigation of Niche and Species formation in genetic function optimization," in *Proc. of 3rd Int. Conf. Genetic Algorithm*, pp. 42-50, 1989. [Article \(CrossRef Link\)](#).
- [37] J. E. Baker, "Reducing bias and inefficiency in the selection algorithm," *ICGAI*, pp. 14-21, 1987. [Article \(CrossRef Link\)](#).
- [38] D. E. Goldberg, and K. Deb, "A comparative analysis of selection schemes used in Genetic Algorithm," *FGAI*, pp. 63-93, 1991. [Article \(CrossRef Link\)](#).



**Yang Lu** received his B.S. and M.S. in information and communication engineering from Harbin Institute of Technology, Harbin, China, in 2007 and 2011, respectively. He is currently working toward the Ph.D. degree with School of Electronics and Information Engineering, Harbin Institute of Technology. His research interests include energy efficient strategy in high-density wireless LANs and networking technology in cognitive radio.



**Xuezhong Tan** received his B.S., M.S. and Ph.D. from Harbin Institute of Technology, Harbin, China, in 1982, 1986 and 2005, respectively. From Oct.1988 to Mar. 1990, he was a visiting scholar with Kyoto University, Kyoto, Japan. He is currently a professor with School of Electronics and Information Engineering, Harbin Institute of Technology. He is a Senior Member of Chinese Institute of Electronics and Chinese Institute of Communication. He has published more than 70 papers in international journals. His research interests include wireless communication, digital trunking communication, and cognitive radio.



**Yun Mo** received the B.S. degree from Beijing Technology and Business University, Beijing, China, in 2009, and the M.S. degree from University of Melbourne, Melbourne, Australia in 2011, respectively. He is working currently toward the Ph.D. degree with the School of Electronics and Information Engineering, Harbin Institute of Technology. His current research interests include pattern recognition, machine learning and indoor-positioning.



**Lin Ma** received his B.S., M.S. and Ph.D. in communication engineering from Harbin Institute of Technology, Harbin, China, in 2003, 2005 and 2009, respectively. In 2009, he joined Communication Research Center of Harbin Institute of Technology. His current research activities are in wireless LAN networking, indoor positioning and indoor navigation.