

# Event-to-sink Spectrum-Aware Clustering in Mobile Cognitive Radio Sensor Networks

Mustafa Ozger, *Student Member, IEEE*, Etimad Fadel, and Ozgur B. Akan, *Senior Member, IEEE*

**Abstract**—Cognitive radio sensor networks (CRSNs) are event-based systems such that sensor nodes detect events and the event readings of the sensors are collaboratively conveyed in a multi-hop manner through vacant channels from event regions to a sink. Hence, the event-to-sink communication and the dynamic radio environment require a coordination scheme in CRSNs. In this paper, we propose a spectrum-aware clustering protocol to address the event-to-sink communication coordination issue in mobile CRSNs. Our clustering scheme consists of two phases. The first phase is the determination of nodes eligible for clustering, and the second phase is to form clusters among those nodes according to vacant spectrum bands. Clusters are temporary and they are not preserved after the end of events. Furthermore, we find average re-clustering probability, expected cluster coverage area, and find maximum event generation frequency for energy-efficient operation of our protocol. We study performance of our protocol in terms of control and data packet exchange, time steps required for clustering, connectivity of clusters, energy consumed for clustering, and re-clustering ratio due to the mobility. Performance comparison simulations show that our algorithm has better performance in terms of connectivity and energy consumption.

**Index Terms**—Cognitive radio, mobile sensor networks, event-driven communication, spectrum-aware clustering

## 1 INTRODUCTION

COGNITIVE radio technology has been proposed as a solution for the electromagnetic spectrum scarcity problem. Wireless nodes with cognitive radio (CR) capability can utilize the idle licensed bands opportunistically without any interference on primary users (PUs) [1], [2]. Wireless sensor networks (WSNs) can benefit from this advantage of CR technology to overcome the spectrum scarcity challenge. To this end, a new wireless networking paradigm, cognitive radio sensor networks (CRSNs), has been proposed by enabling dynamic spectrum access (DSA) schemes in WSNs [3].

Different from the fixed frequency allocation approach in WSNs, CRSN nodes opportunistically utilize the idle licensed bands. The opportunistic usage of the spectrum is provided by the cognitive cycle operations. CRSN nodes sense the spectrum to determine vacant bands by the spectrum sensing, choose their operating frequencies by the spectrum decision, and change the frequency by the spectrum hand-off according to the licensed user activities [3]. CRSNs are distributed networks such that they do not depend on any central entity to regulate the communication. Packets generated by the event detecting nodes are transmitted in a multi-hop manner from the event

region to the sink in a dynamic radio environment. Technological advances and research efforts make the deployment of CRSNs realizable. To this end, the requirements of cognitive radio and sensor networks and their unique challenges are addressed by a growing number of studies on CRSN. Some of recent works focus on channel assignment according to residual energy [4], energy-efficient channel management scheme [5], packet size optimization [6], power and rate adaptation for maximization of information theoretical capacity [7], and performance analysis in terms of delay [8].

Although CR capabilities offer the flexibility on wireless communication by using different spectrum bands, they also pose some important challenges. First of all, the routing in a dynamic radio environment is performed with an additional constraint of the spectrum availability. Also, the transmission between the sensor nodes can be interrupted by the cognitive cycle operations. The fluctuating nature of available channels due to the PU activity deteriorates the success of the wireless communication. The flat structure of CRSN also causes scalability issues due to an increase in the network size [9]. Furthermore, these challenges are amplified if the sensor nodes are mobile. The mobility of the nodes causes frequent topology changes. Spectrum management and dynamic topology change must be taken into consideration together to satisfy the requirements of the spectrum-aware communication and the management of the mobility. Furthermore, the event-driven communication nature injects bursty traffic which causes channel contentions and packet drops. Hence, these energy consuming operations must be handled.

Mustafa Ozger and Ozgur B. Akan are with the Next-generation and Wireless Communications Laboratory (NWCL), Department of Electrical and Electronics Engineering, Koc University, Istanbul, 34450, Turkey (e-mail: {mozger, akan}@ku.edu.tr).

Etimad A. Fadel is with the Computer Science Department, Department of Electrical and Electronics Engineering, King Abdulaziz University, Jeddah, Saudi Arabia (e-mail: eafadel@kau.edu.sa).

An earlier version of this paper has been presented at IEEE INFOCOM 2013 [17].

Event-to-sink communication may be handled by a clustering approach due to following reasons. First of all, the event routing problem necessitates a cooperation scheme since the network is decentralized, and the channel availabilities of nodes and the topology of the network change dynamically. The event packets must be conveyed from the event region to the sink in such a dynamic environment. Clusters group nodes to cooperate with each other for the event packet routing, the channel assignment and sharing the channel sensing results reliably. Data channel can be determined by the cluster-head and cluster members for intra-cluster communication [10]. Clustering also provides network performance guarantees by organizing the nodes with respect to the dynamic channel availabilities. Cluster-heads and gateway nodes form a backbone for distributed CRSNs such that routing, cognitive cycle operations and mobility management become more manageable.

In a mobile and dynamic radio environment, keeping cluster structure before any event-to-sink communication may lead to excessive energy consumption. Hence, we propose forming clusters after event occurrences in CRSNs. The formed clusters are spectrum-aware such that the clustering criteria do not only depend on physical proximity but also available channels of the nodes and their positions in the network relative to the sink. In this paper, we present mobility-aware Event-to-sink Spectrum-Aware Clustering (mESAC) protocol for mobile CRSNs. The contributions of our paper are itemized as follows.

- Re-clustering probability and expected cluster coverage area are calculated.
- Maximum event generation frequency achieving energy-efficiency with respect to the whole network clustering is studied.
- The effects of mobility and the weight of the mobility in the cluster-head selection process are investigated.

## 2 PROBLEM MOTIVATION

We first describe the network model for CRSN. Afterwards, we give the motivation for the event-driven clustering.

### 2.1 Network Model

PUs utilize the licensed bands without any restrictions. CRSN nodes are the secondary users (SU) which can opportunistically access licensed channels if there is no primary user activity [2]. CRSN nodes can move according to random waypoint mobility model [11]. The nodes form a graph denoted by  $G = (V, E)$  where  $V$  is the set of nodes and  $E$  is the set of edges connecting the nodes. Two neighboring CRSN nodes are connected if they have a common channel.

There are  $M$  non-overlapping orthogonal channels having unique IDs. CRSN nodes can detect vacant

TABLE 1: Notations For Clustering Protocol

Notation	Explanation
$C_i$	Available channel list of node $i$
$E_i$	Remaining energy of node $i$
$d_{i,j}$	Euclidean distance between nodes $i$ and $j$
$d_i^e$	Eligible node degree of node $i$
$v_i$	Speed of node $i$
$\mu$	Mean of the distribution of node speed
$P_i$	Weight for cluster-head selection
$R$	Event radius in meters
$D_{es}$	Distance between event and sink in meters
$r$	CRSN node transmission range in meters

spectrum bands by wideband spectrum sensing. The transmission ranges of PUs and SUs are comparable such that the sensing results of SUs are correlated for the same neighborhood.

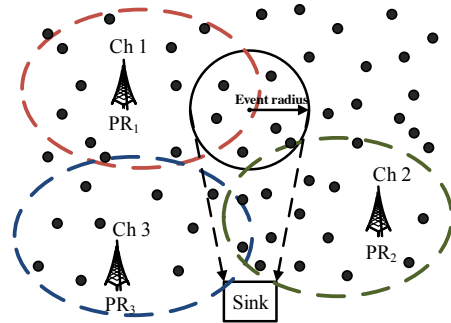


Fig. 1: A typical CRSN topology with an event and the sink.

The communication is event-driven, and CRSN nodes located within the event radius detect the event. These nodes generate packets to be delivered to the sink. Many-to-one traffic pattern from the event detecting nodes toward the sink node is generated in a CRSN as illustrated in Fig. 1. There are 3 primary radios (PRs) and the coverage areas of them are shown by the circles with dashed lines and their operating channels are shown. Inside these circles, SUs can not use those channels if PUs are active.

Clustering needs the exchanges of control messages. The exchanges of these messages are provided via a common control channel with a very small bandwidth. It is available for SUs at any time [12]. Time is slotted and CRSN nodes can transmit their control information in these time synchronized slots. The time synchronization is not in the scope of the paper. Overlay spectrum sharing is used as an interference avoidance model, i.e., SUs use portions of the spectrum not used by the PUs [2]. SUs know their location by utilizing localization algorithms. Neighbor discovery is provided by the periodic common control channel signaling. All nodes know one-hop neighbors and their vacant channels by this signaling [13]. Table 1 shows the notations and their explanations which are useful for explaining the clustering protocol.

As in [14], the traffic of primary channel  $m$  is modeled as semi-Markov ON-OFF process where channel

$m$ 's state changes with probabilities  $p_{on,m}$  and  $p_{off,m}$ . OFF state of channel  $m$  is considered as a spectrum opportunity for SUs in the communication range of PU using channel  $m$ . SUs can use this channel for the communication and the cluster formation. This is the temporal variation of the spectrum availabilities. Also, PUs can be assumed to be distributed uniformly over the network area [15]. This may be interpreted as the spatial variation of the spectrum availabilities.

The free space path loss model is adopted such that  $\epsilon_{fs} = 10 \text{ pJ/bit/m}^2$  and  $E_{elec} = 50 \text{ nJ/bit}$  [16]. We also assume that the control packets for clustering are  $200 \text{ bits}$  [17]. We do not take into account the energy for spectrum sensing since it is a common and periodic process for each CRSN node. Our focus is the extra energy consumption due to the channel switching during the event packet transport. The energy consumption for the channel switching is assumed to be  $J_{sw} = 40 \text{ mJ}$  [18].

## 2.2 Problem Motivation

Multi-channel sensor networks and CRSNs utilize multiple channels for the communication. The most advantageous aspect of utilizing multiple channels in CRSNs and multi-channel WSNs is empowering parallel transmissions [19]. This aspect of the multi-channel communication alleviates the contentions and the collisions in these type of dense sensor networks. Multi-channel sensor networks are closely akin to CRSNs in this aspect but there is a fundamental difference. In case of the channel degradation and the interference, sensor nodes in multi-channel WSNs change the operating frequency band without any limitation by a licensed user. However, the channels CRSN nodes opportunistically utilize are the licensed channels and they can be used only when these channels are idle in CRSNs. The CR capabilities of sensor nodes provide the adaptability to the dynamic radio environment. Hence, the solutions for CRSNs must take into account the PU activities.

Although multi-channel WSNs and CRSNs have similar challenges such as channel switching and channel coordination, the realization of CRSNs poses some extra challenges due to the specific properties of CR and sensor networks, and the challenges are amplified by their unique union. The limited battery capability requires energy-efficient and low cost solutions as in the case of multi-channel WSNs. On the other hand, the spectrum management is a result of CR capability of CRSN nodes. The dynamic radio environment due to CR capabilities requires spectrum-awareness. Furthermore, the operation over multiple channels also poses challenges which are also experienced by multi-channel sensor networks. One of these challenges is channel switching which causes energy consumption. Channel coordination is another challenge faced by CRSNs and multi-channel WSNs.

Two neighbor nodes must be on the same channel if they intend to communicate. However, solutions proposed for multi-channel WSNs are not spectrum-aware since they do not consider licensed user activities [20], [21]. On the other hand, cognitive radio network (CRN) solutions discard the hardware and the energy limitations. The solutions for multi-channel WSN and CRN do not address energy-efficiency and spectrum-awareness simultaneously.

The event-driven communication nature of CRSN causes a burst packet injection towards the sink. This yields a communication corridor from the event region to the sink [22]. There exist communication activities only when events occur and there are unique challenges of spectrum-aware communications and the mobility of the nodes. Hence, we propose an event-driven spectrum-aware clustering scheme in the corridor between event and sink to overcome these challenges. The main motivations for proposing clustering technique to this ephemeral event routing problem can be itemized as follows.

- Data aggregation is one of the most important features of the clustering technique. The nodes closer to each other make correlated observations [23]. This decreases the number of data packet transmission which is the most energy consuming part of event reporting. Furthermore, these nodes make correlated spectrum sensing observations [15]. Correlated spectrum sensing logically partitions the nodes in the network into groups with similar spectrum availabilities. Hence, the clustering fulfills the need for the data aggregation, and also the coordinate the spectrum-aware communication.
- The spectrum mobility and the node mobility pose a significant challenge on communication. The channel on which the sensor nodes decide to communicate may become occupied and a new communication channel must be decided. A node may lose its contact with its communicating neighbor due to the mobility of nodes and a new neighbor node must be decided for the event data routing. However, the clusters can deal with the spectrum mobility and the node mobility since the cluster may switch to another data channel among its cluster channels and the communicating node find easily another node in the cluster to route its event packets. Once, the clustering proposes redundancy against spectrum mobility and node mobility.
- A node must exchange control packets with all other neighbors to establish a communication link if the clustering is not utilized [10]. The constraint of having at least a common channel and mobility may be very energy-consuming for the communication in CRSN without clustering since the nodes exchange control packets at each

hop. However, clustering forms a communication structure keeping some neighbor nodes tuned to the same channel. The cluster structure may reduce the channel switching during transmissions of event packets.

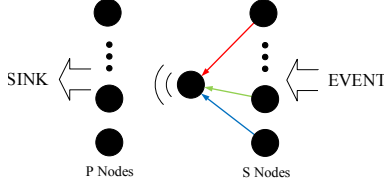


Fig. 2: A communication scenario in a one hop neighborhood during an event to sink communication.

These advantages can be shown with an example for an event-to-sink communication scenario shown in Fig. 2. For clustering case,  $S$  nodes and  $P$  nodes are the cluster members, and the bottleneck node is the cluster-head.  $S$  transmissions on the same channel which causes  $SE_{ept}$  energy consumption where  $E_{ept}$  is the energy for one-hop transmission. The cluster-head aggregates these packets into one packets and transmits that packet to the nodes closer to the sink. The total energy consumption for the clustering case is  $(S + 1)E_{ept}$ . If the clustering is not used for this scenario,  $S$  nodes send their packets to the bottleneck node and  $W_1$  transmissions require channel switching where  $0 \leq W_1 \leq S$ . The bottleneck node sends all the received packets without aggregation which requires  $S$  transmissions and  $W_2$  channel switching where  $0 \leq W_2 \leq S$ . The total energy consumption is  $2SE_{ept} + (W_1 + W_2)J_{sw}$ . This simple scenario shows the advantage of the clustering scheme in spectrum-aware communication.

### 3 RELATED WORK

There is a substantial amount of work on the clustering for cognitive radio networks (CRNs) and WSNs in literature. We classify some of the important works as those which have spectrum-awareness and not. For instance, in one of the spectrum-aware clustering approach in [24], the authors propose a distributed coordination architecture to overcome the spectrum heterogeneity challenge and the scalability issues. The proposed scheme groups nodes according to the selection of coordination channels adaptively and forms clusters in CRN. The scheme tries to maintain the groups during the network operation. Users are clustered in the same group if they have at least one common channel and two nodes in the same cluster can communicate via multiple hops. The aim is to minimize the number of clusters in the network. However, it decreases the number of common channels in clusters. Consequently, this feature of the protocol increases re-clustering probability since it causes less common channels among the cluster members. In [25],

the network is partitioned into clusters by grouping neighbor nodes sharing local common channels. The network is formed by interconnecting the clusters gradually. Furthermore, it provides mechanisms for neighbor discovery, cluster formation, network formation, and network topology management. [26] uses graph theory for spectrum opportunistic clustering (SOC) and assigns control channel to each cluster so that each node in a cluster can communicate within the cluster by local common channels. The recent work [15] presents network topology and spectrum availability as bipartite graphs. Every node constructs bipartite graphs with its one-hop neighbors and available channels. Biclique graphs are obtained from these bipartite ones according to metrics such as maximum edge, maximum node and maximum edge one sided cardinality. These approaches do not use common channel over the entire network, however, a control channel is assigned to each cluster in the network by the clustering technique. The approaches in [26], [15] are not applicable to CRSN since there are three steps to form clustering which needs excessive energy consumption. The authors in [27] construct minimal number of clusters in cognitive radio networks using affinity propagation (AP) message-passing technique. ROSS-DGA and ROSS-DFA are two distributed clustering approaches offered in [28]. They form robust clusters by providing inter-cluster and intra-cluster connectivity using the game theory. These approaches are not compatible with the resource-constrained nature of CRSN nodes. They do not consider the event-driven communication paradigm which is essential for CRSNs.

A clustering scheme is proposed for multi-channel WSNs to achieve power-efficiency [20]. It proposes a backbone by clustering and decreases the amount of packet forwarded to the sink by data aggregation by cluster-heads. However, clustering in [20] does not take into account the PU activities. The recent work in [29] suggests a spectrum-aware clustering protocol to enable energy-efficient communication in CRSN by intra-cluster aggregation and inter-cluster relaying. It also finds the optimal number of nodes in a cluster. However, the authors do not consider the connectivity between cluster-heads for inter-cluster communication. On the other hand, the clustering approaches proposed for WSNs consider energy-efficiency but lack spectrum-awareness. The clusters are formed by first selecting cluster-heads in WSN clustering schemes. Optimal cluster-head selection is an NP hard problem [30]. A novel cluster-head selection algorithm, Weighted Clustering Algorithm (WCA) [31], is proposed based on the combined weight metric. This metric is calculated by taking into account several system parameters. Despite vast amount of clustering approaches in WSN, the idea of clustering between event and sink, namely Event-to-Sink Directed Clustering (ESDC), is proposed in [22]. It suggests forming

clusters in the corridor between the event and the sink in WSNs. However, this clustering scheme is not applicable to CRSN since WSN nodes are not aware of CR functionalities and ESDC does not address spectrum heterogeneity.

Although there exists significant amount of clustering approaches on CRNs, they do not address the challenge of limited energy resource and hardware capacity of sensor nodes. Thus, energy-efficient clustering solution is required to address the limited energy resource challenge as well as opportunistic spectrum access challenges. In this paper, we provide energy-efficiency by exploiting the event-driven communication nature, and perform the spectrum-aware clustering for the coordination in a dynamic radio environment.

## 4 MOBILITY- AWARE EVENT - TO - SINK SPECTRUM - AWARE CLUSTERING (MESAC)

Our clustering protocol, mESAC, provides an on-the-fly coordination scheme for the communication in CRSN. We propose a two phase clustering protocol to establish the coordination between the event and the sink. The first phase is for the determination of the intermediate eligible nodes between the event and the sink, i.e., the formation of an eligibility routing corridor for clustering. In the second phase, the nodes in this corridor form the spectrum-aware clusters with their one-hop neighbors. Clusters are maintained until the end of the event. These phases are explained individually in the following subsections.

### 4.1 Eligibility Corridor Determination

The communication in CRSN starts with event occurrences, i. e., resources are used in the network for event occurrences. Hence, we designed our protocol according to communication needs of CRSN. The first step of our protocol is to determine the corridor between event and sink. This step is performed in a distributed fashion since the network is decentralized. We specify this corridor by changing the status of the nodes in this region. The first step of determining the eligible nodes is the event detection in our distributed approach. The event detecting nodes become eligible for clustering directly. Afterwards, Eligibility For Clustering REquest (EFC\_REQ) messages are sent by these nodes to their one-hop neighbors through a common control channel by piggybacking it to the spectrum sensing packets through cooperative diversity mechanism [15], [32]. Each node receiving this message becomes an eligible node if it is located closer to the sink than the EFC\_REQ sender. The new eligible nodes send EFC\_REQ to their non-eligible one-hop neighbors and the process continues until EFC\_REQ reaches the sink. Non-eligible nodes do not send EFC\_REQ messages, hence, the corridor can

not be expanded further in the directions other than the direction towards the sink. Finally, the eligibility corridor is determined by these messages. Algorithm 1 outlines the eligibility process for the clustering. mESAC determines the eligible nodes in a distributed manner by Algorithm 1. In this phase, we assume that the sensor node density is sufficient to not terminate the formation of eligibility corridor due to routing towards the sink.

---

#### Algorithm 1 Determination of the Eligible Nodes for Clustering

---

```

1: Consider node  $i \in V$  where  $V$  is set of nodes in
   the network
2: if Node  $i$  detects event then
3:   It is eligible for clustering
4:   State( $i$ ) = "Eligible Ordinary Node"
5:   Start sending EFC_REQ to one-hop neighbors
   other than event detecting neighbors
6: else
7:   if It receives EFC_REQ from its neighbor then
8:     Compare the distance
9:     if  $d_{i,sink} \leq d_{EFC\_REQ-sender,sink}$ 
        $\wedge d_{i,event} \geq d_{EFC\_REQ-sender,event}$  then
10:      Node  $i$  is eligible for clustering
11:      State( $i$ ) = "Eligible Ordinary Node"
12:      Send EFC_REQ to one-hop neighbors
13:     else
14:       Not eligible
15:     end if
16:   else
17:     Node  $i$  is not eligible
18:   end if
19: end if

```

---

If a node receives an EFC\_REQ message by one of its neighbors, the sender already knows if the request receiver is eligible due to the neighbor discovery. However, the nodes that are eligible or non-eligible for clustering inform their neighborhood by Eligibility For Clustering REply (EFC\_REP) message upon receiving an EFC\_REQ message. Due to the mobility, some nodes leave or enter the eligibility region, the topology changes in the eligibility region are tracked by the neighbor discovery process which is not in the scope of this paper. These nodes become eligible or non-eligible according to the eligibility condition by the requests from their eligible neighbors. After this process, a new subnetwork is generated which is denoted by  $G' = (V', E')$  where  $V'$  is the set of eligible nodes and  $E'$  is the set of edges connecting eligible nodes.  $V'$  is a subset of  $V$ .

The steps of the first phase of the protocol is shown in Fig. 3. The mobility management in this corridor is done in a cooperative manner. If an eligible node moves out of the eligibility corridor which is a result of having no eligible neighbors, it becomes non-

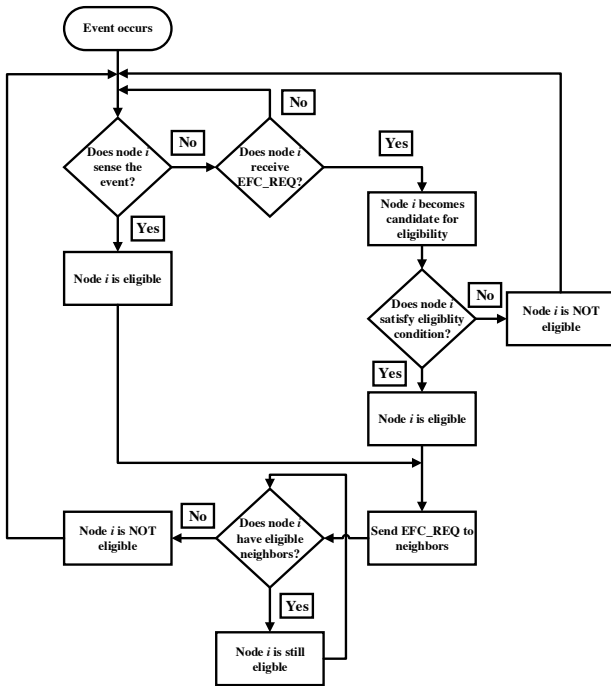


Fig. 3: Steps for managing eligibility and mobility of the nodes during an event.

eligible and waits for an EFC\_REQ. If a non-eligible node encounters an eligible node and it satisfies the eligibility condition, it becomes eligible.

## 4.2 Cluster-head Selection

The motivation of clustering in sensor networks is to create organizational structures and to manage distributed wireless networks in a structured way. The dynamic radio environment intensifies the need for a clustering scheme to coordinate the communication in a typical CRSN. In this work, we group the eligible nodes in the event-to-sink corridor according to their spectrum availabilities while maintaining intra-cluster and inter-cluster communication. Only the nodes that are appointed as eligible form clusters. The clusters are not maintained in the network if the event detecting nodes do not further sample the event.

Cluster-heads are the local coordinators in the cluster. They perform data aggregation, manage the change of the idle spectrum bands in the clusters and manage the mobility of the nodes. They also receive the event reports from the cluster members that are closer to the event. Afterwards, they perform data aggregation to reduce the event reports and route these reports to the nodes closer to the sink.

Optimal cluster-head selection is an NP-hard problem [30]. Hence, heuristic algorithms are proposed for this problem in literature. In our case, the solution to the cluster-head selection problem also requires spectrum-awareness in the dynamic radio environment. Hence, the cluster-head selection must take into account the vacant spectrum bands of the eligible

nodes. In this paper, we propose an algorithm to determine cluster-heads among the eligible nodes after the first phase. The formed clusters must be stable against the fluctuations of the spectrum availabilities, the energy outage, and the mobility of the nodes. To this end, we take into account several system parameters as in [31]. Every eligible node is assigned a weight according to a combination of different parameters. This weight shows the capability of a nodes to serve as a cluster-head. In our heuristic selection method, we maximize the weight of the selected cluster-heads. Hence, we select the nodes in the network having maximum capability to serve as cluster-heads. The parameters for selecting cluster-heads can be itemized as follows.

- **Eligible node degree,  $d_i^e$ :** The node with the highest eligible one-hop neighbor must be more likely to be a cluster-head since it has more opportunity to form a cluster and to maintain the formed cluster. Hence, selecting the node with the highest node degree as cluster-head will have higher chance and combination alternatives to form robust cluster with its eligible one-hop neighbors.
- **Number of available channels,  $C_i$ :** Spectrum-aware clustering is performed according to the available channels of cluster-heads and their one-hop neighbors. Hence, the cluster-head must be the one having the highest number of available channels to increase the number of available channels in a cluster. If a node with the highest available number becomes a cluster-head, the cluster becomes more stable since there will be higher chance to have greater number of vacant bands in the formed cluster. If the node with less number of vacant bands is selected as a cluster-head, it will confine the cluster structure to less common channels in the formed cluster and smaller number of cluster member nodes. The selection of cluster-head node with higher vacant channels also increases the probability of inter-cluster communication since the nodes with the highest number of vacant channels in its neighborhood have higher possibility to possess a common channel with the neighboring cluster-heads.
- **Remaining energy,  $E_i$ :** Cluster-heads are more active than ordinary node due to the intra-cluster communication. Therefore, the remaining energy is an important parameter while selecting the cluster-heads.
- **Distance to the sink,  $d_{i,sink}$ :** The data collected by the sensor nodes is aggregated in the cluster-head. The collected data is transferred by the cluster-heads in the event routing corridor. Hence, the selection of cluster-heads closer to the sink decreases the energy consumption in CRSN.

- **Node speed,  $v_i$ :** It is a good choice to select the nodes whose speeds are low as cluster-heads for stable clusters. In the weight calculation, the node speed is inversely proportional to the weight. The cluster-head is the main entity for a cluster, and if it leaves the cluster region, all the member nodes become non-clustered.

The above parameters are selected to maintain the stability of the formed clusters and to provide the energy-efficiency. These parameters are used separately in literature to select the cluster-heads. For example, the node degree is used for WSN clustering [33], the number of available channels is used CRN clustering [24], [25], [26], [27], [28], [29], the remaining energy and the distance to the sink are used for WSN clustering [22]. However, to the best of our knowledge the combination of these parameters for the cluster-head selection in a dynamic radio environment has not been considered. Hence, we propose a cluster-head election process based on the combination of these parameters. According to these parameters, every eligible node has a weight for the cluster-head election which is defined as

$$P_i = w_1 |d_i^e| + w_2 |C_i| + w_3 |E_i| + w_4 |d_{i,sink}| + w_5 \left| \frac{1}{v_i} \right| \quad (1)$$

for an arbitrary eligible node  $i$ . The sum of the weight of the parameters is 1.  $|d_i^e|$  is obtained by the EFC\_REQ messages of the neighbor nodes. The other parameters are contained in the node  $i$ . After the first phase, the eligible nodes determine their weights for the cluster-head selection. The units of the parameters are different, however, we are interested in their magnitudes.

The weight of each parameter depends on the system requirements. Although there is not an exact procedure to select the weights of the parameters, the following steps can be carried out to tune the weights of these parameters. Firstly, if one or more parameters are not important to the reader, then the weight of this/these parameters should be zero. The reader should divide the total weight, which is 1, between the remaining parameters. If the remaining parameters are equally important, the weights of the remaining parameters stay the same. However, if there is one parameter whose importance is greater than the remaining parameters, the weight of this parameter should be increased incrementally by decreasing the weights of the other parameters until the desired system performance is achieved. If there are more than one parameter whose importance is greater than the remaining parameters, the reader should also increase the weights of these important parameters by decreasing the weights of the less important parameters. Hence, the reader/researcher should tune the weight of the parameters to achieve the desired system performance according to the importance of the parameters. The more important parameter should have the greater weight and the less

important parameter should have the less weight. For example, if the nodes are not mobile,  $w_5$  must be zero. For an other case where  $w_1 = 1$  and  $w_2 = w_3 = w_4 = w_5 = 0$ , the cluster-head selection depends only on the node degree, and cluster-head selection becomes the highest node degree algorithm in [33]. We are offering a general weighing scheme and optimal values of the parameter weights can be found for different network setups. This optimization is beyond our scope.

In this procedure, the nodes with the highest weight in their neighborhood become cluster-heads. The new cluster-heads and their one-hop neighbors are deprived from the cluster-head election process. The cluster-head election process continues with the nodes that are not deprived from the election procedure. These remaining nodes again control whether they are the nodes with the highest weight among their non-clustered neighbor nodes. This process will go on until all nodes are clustered.

#### 4.2.1 Time Complexity of Cluster-Head Selection

The cluster-head selection is a completely distributed process. The node which does not have the highest weight among its eligible neighbors waits for its eligible neighbor nodes to become cluster-heads or members of cluster-heads. Thus, this node can become a node having the highest weighted node among its non-clustered eligible neighbors or can be deprived from the cluster-head selection process by a newly selected neighboring cluster-head. Hence, the time required for cluster-head election procedure depends on the maximum number of hops in the eligibility corridor. It is directly related to the hop diameter of the eligibility corridor which is generally the number of hops between the event and the sink. Thus, the time complexity of the cluster-head selection process depends on the distance between the event and the sink.

### 4.3 Determining Cluster Members

Firstly, each cluster-head investigates whether common channels exist among all its eligible one-hop neighbors. If the number of the calculated common channel is nonzero, then the cluster-head sends C\_REQ to its eligible one-hop neighbors. If there is not even a common channel among the eligible one-hop neighbors, the node calculates the weights of its vacant channels. The weight of a channel is calculated by counting the number of eligible one-hop neighbors having that channel in their vacant channel set. The highest weighted channel becomes the desired cluster channel, and the other common channels are the backup channels against PU appearance. The eligible one-hop neighbors having that channel become the desired cluster members. A tie is resolved in favor of the channel having smaller channel ID number. The



cluster-head node sends C\_REQ to these desired cluster members. However, a node may be a neighbor of two or more cluster-heads and it becomes a member of only one cluster-head. Hence, the desired calculated structure may not be achieved since a node will reply to only one of C\_REQs.

The node receiving C\_REQs from multiple neighboring cluster-heads selects the node with the lowest weight. The lowest weight selection is due to the load balancing between cluster-heads. The cluster-head with the lowest weight has less member, and the cluster-heads with higher weight has greater number of members. In order to decrease the burden of the cluster-heads with higher weights, the cluster-head with the lower weight is selected for the cluster membership.

#### 4.3.1 Dominating set

The cluster-head selection is based on the iterative selection of cluster-heads according to the weights of one-hop neighborhood. Hence, the cluster-head selection procedure forms a dominating set among eligible nodes such that every eligible node is either a cluster-head or a neighbor of a cluster-head. Every non-clustered eligible node is covered by cluster-head set. However, an eligible node may not receive any C\_REQ from any of its neighbor cluster-head nodes. In this case, these nodes become cluster-heads. Each eligible node is either a cluster-head or a cluster member. Hence, the set with eligible nodes forms a dominating set.

#### 4.3.2 Convergence of mESAC

Nodes in a formed eligibility corridor are connected due to the assumption of connectivity of the network such that all event messages from the event region reach to the sink ultimately. Hence, the convergence of the eligibility corridor formation, which is the first phase of the algorithm, is guaranteed due to the connectivity.

In the second phase, the cluster-head selection procedure is performed iteratively in a distributed fashion among the formed eligible network,  $G'$ . Hence, the procedure finishes until all the nodes in  $V'$  becomes either a cluster-head or a neighbor of a cluster-head in which the time required for this process depends on the diameter of  $G'$ . Afterwards, the eligible nodes that are not cluster-head respond to the clustering requests from the neighboring cluster-heads having the lowest weight. They become members of the corresponding cluster-heads. If a non-clustered eligible node does not receive any request from its neighboring cluster-heads, it becomes a cluster-head. Thus, each node in  $V'$  becomes either a cluster-head or a member of a cluster-head. This finalizes our algorithm which guarantees the overall convergence.

### 4.4 Topology Change due to Mobility

A non-eligible node can leave or enter the eligibility corridor during the event transport due to the mobility of nodes. The node leaving the eligibility region may be a cluster-head or a member node. If it is a member node, cluster structure continues without that node, and the members and the channels of the cluster change accordingly. On the other hand, if the leaving node is a cluster-head, the cluster-head selection procedure is implemented as in Section 4.2 and Section 4.3.

A node can enter the eligibility corridor. It becomes eligible by EFC\_REQ message from its eligible one-hop neighbors. The incoming node firstly becomes eligible, then it associates itself to the nearest cluster-head if it has the channels of that cluster. Otherwise, it becomes a cluster-head.

### 4.5 Inter-cluster Communication

Inter-cluster communication is performed through the gateway nodes. Due to the fluctuations of the spectrum availability, a common channel may not exist between neighboring cluster-heads. Hence, the inter-cluster communication is performed via gateway nodes and it is coordinated by the cluster-heads. The inter-cluster communication depends on the common channels between the nodes in the neighbor clusters.

The parameter which defines the number of vacant channels while selecting cluster-heads increases the possibility of common channels between the neighboring cluster-heads. If there is a common channel between any pair of neighboring cluster-heads, there is no need to change the operation channel between the cluster-heads and the gateway nodes. This provides energy-efficiency since the channel switching is energy consuming operation.

## 5 EFFICIENCY REQUIREMENT OF EVENT-TO-SINK APPROACH FOR CLUSTERING

### 5.1 Average Re-clustering Probability

The spectrum opportunity occurs for a node, if it senses the channel as idle. This event occurs if the node is not around any PU and the channel is successfully sensed as idle, or if it is in the transmission range of the PU and the PU is not active and the node senses the channel as idle, or PU is active and the node falsely senses the channel idle. To this end, we first find the probability that a node is not in the transmission range of a PU. This probability is obtained as

$$p_w = 1 - \frac{\pi\tau^2}{A} \quad (2)$$

where  $\tau$  is the transmission range of a PU and  $A$  is the area of the network.



The imperfections on the channel sensing cause errors. The probabilities of detection and mis-detection can be defined respectively as

$$p_d = Pr[\text{Ch. } i \text{ sensed idle} \mid \text{Ch. } i \text{ is idle}], \quad (3)$$

$$p_m = Pr[\text{Ch. } i \text{ sensed idle} \mid \text{Ch. } i \text{ is occupied}]. \quad (4)$$

The channel usage of a PU depends on its ON-OFF probabilities. These probabilities can be denoted as  $p_{on}$  and  $p_{off}$ . A CRSN node can have the channel  $i$  in its vacant channel list in 3 cases which can be explained as follows.

- If the node is outside of the transmission range of the PU and the state of the channel  $i$  is detected successfully, the channel  $i$  is in the vacant channel list. This probability can be written as  $p_w p_d$ .
- If the node is inside the transmission range of the PU which is OFF, and the state of the channel  $i$  is detected successfully, the channel  $i$  is in the vacant channel list. This probability can be written as  $(1 - p_w) p_{off, i} p_d$ .
- If the node is inside the transmission range of the PU which is ON, and the state of the channel  $i$  is mis-detected as idle, the channel  $i$  is in the vacant channel list. This probability can be written as  $(1 - p_w) p_{on, i} p_m$ .

The probability that channel  $i$  is in the vacant channel list of a node can be obtained by the union of these 3 independent cases. Hence,  $p_{l, i}$  is

$$p_{l, i} = p_w p_d + (1 - p_w) p_{off, i} p_d + (1 - p_w) p_{on, i} p_m. \quad (5)$$

Since ON probabilities of the channels are the same, we can omit  $i$  in the subscripts in Eq. (5) and can state the probability that an arbitrary channel in the system is in the list of a node as  $p_l$ . This probability defines the change of the channel availability in our dynamic radio environment under the imperfect channel sensing and the spatial variation of PUs.

Under random waypoint mobility model in a wireless network, the distribution of the distance between any pair of nodes,  $D$ , can be found for a square network whose one side length is  $L$ . It is defined for the square of the distance, hence, for  $0 < D \leq L^2$ , the distribution of the distance between any two nodes smaller than  $L^2$  can be written as in Eq. (6).

Two nodes become neighbors if they are in the transmission ranges of each other. Hence, the probability that the two nodes are neighbors, which can be denoted by  $p_n$ , can be given as

$$p_n = Pr[D \leq r^2] = \int_0^{r^2} \phi(z) dz. \quad (7)$$

The re-clustering condition for a cluster is to have no common channel among its members. Hence, the re-clustering probability of the cluster can be found by considering the overall probability that there is no

common channel among the possible cluster members. Firstly, we have a cluster-head for a cluster structure. There can be  $N - 1$  member of a cluster at most since there are  $N$  nodes in the network. Furthermore, any set of channels from the set of channels in the network,  $C$ , can be the common channel set of the cluster. There must be an event occurrence for the cluster formation since our clustering algorithm is event-driven. We denote event occurrence as  $E = 1$ , and the no event case as  $E = 0$ . To this end, the probability of the re-clustering of a cluster given that an event occurred,  $p_{rc|E=1}$  can be expressed as

$$p_{rc|E=1} = \sum_{k=1}^{N-1} (p_n)^k \sum_{m=1}^M \binom{M}{m} ((1 - p_l)^m (1 - p_l^{(M-m)}))^k \quad (8)$$

where  $M$  is the size of the channel set in the system. Hence, the probability of re-clustering can be found as

$$p_{rc} = p_e \sum_{k=1}^{N-1} (p_n)^k \sum_{m=1}^M \binom{M}{m} ((1 - p_l)^m (1 - p_l^{(M-m)}))^k \quad (9)$$

where  $p_e = Pr[E = 1]$  is the probability that an event has occurred.

## 5.2 Expected Cluster Coverage Area

Cluster coverage area is the area covered by the cluster-head plus the extra coverage regions contributed by possible cluster members. Applying the procedure in [34], the coverage area of a cluster can be found as follows. In the first step, the coverage area contributed only by the cluster-head is  $\pi r^2$ . Maximum possible total area of a cluster is  $\pi(2r)^2$  since possible one-hop cluster members can be located at a distance  $r$  to the cluster-head so that the maximum radius from the cluster-head becomes  $2r$ . We ignore the edge effects in our computations.

We start with the cluster-head in the coverage area computation of a cluster structure. The cluster-head coverage, which is the minimum cluster coverage area, can be expressed as  $\gamma_1 = \gamma_{min} = \pi r^2$  and the maximum cluster coverage area can be expressed as  $\gamma_{max} = \pi(2r)^2$ . In the next steps, we assume that the cluster-head gains member nodes and we investigate the contribution of the incoming cluster member to the cluster coverage area. For instance, if there are  $k - 1$  nodes in the cluster and one more node joins the cluster, the cluster coverage area of  $k$  nodes becomes as

$$\gamma_k = \gamma_{k-1} + \alpha_k \quad (10)$$

where  $\alpha_k$  is the contribution from the  $k^{\text{th}}$  member node. If we take expectation on both sides of Eq. (10), it becomes  $E[\gamma_k] = E[\gamma_{k-1}] + E[\alpha_k]$ .

The fraction of the contribution by the  $k^{\text{th}}$  node with respect to the maximum cluster coverage area

$$\phi(D) = \frac{6D^3}{L^8} - \frac{6D^4}{5L^{10}} + \frac{6D^5}{125L^{12}} + \frac{96\sqrt{D^3}}{5L^5} - \frac{1584\sqrt{D^5}}{125L^7} + \frac{36\pi}{25L^2} + \frac{192D^2\sqrt{D^3}}{175L^9} - \frac{36\pi D}{5L^4} - \frac{48D\sqrt{D^3}}{5L^7} + \frac{9\pi D^2}{2L^6} \quad (6)$$

can be written as

$$f_k = \frac{\gamma_{max} - \gamma_{k-1}}{\gamma_{max}}. \quad (11)$$

The contribution of the  $k^{\text{th}}$  node is obtained as

$$\alpha_k = f_k \gamma_1 = \left( \frac{\gamma_{max} - \gamma_{k-1}}{\gamma_{max}} \right) \gamma_1. \quad (12)$$

The expected cluster coverage area can be written as

$$E[\gamma_k] = E[\gamma_{k-1}] + E\left[\frac{\gamma_{max} - \gamma_{k-1}}{\gamma_{max}}\right] \gamma_1 \quad (13)$$

where  $\gamma_1$  is a constant.

The procedure for finding the expected cluster coverage area can be iteratively found from the first node, i.e. the cluster-head, to the  $n^{\text{th}}$  node. It can be obtained as

$$E[\gamma_k | k = n] = \left[ 1 - \left( 1 - \frac{\gamma_1}{\gamma_{max}} \right)^n \right] \gamma_{max}. \quad (14)$$

Some of the nodes in the neighborhood of the cluster-head have the common channels with the cluster-head and they contribute to the expected cluster coverage area. If we denote the probability that there are  $n$  nodes satisfying the clustering criteria as  $Pr[k = n]$ , the expected cluster coverage area can be written as

$$E[\gamma] = \sum_{n=0}^{N-1} E[\gamma_k | k = n] Pr[k = n]. \quad (15)$$

Eq. (15) can be expanded so as to write the explicit form for  $Pr[k = n]$  as

$$E[\gamma] = \sum_{h=0}^{N-1} \sum_{j=h}^N \sum_{m=1}^M \binom{M}{m} p_n^j (1 - p_n)^{N-j} p_l^{hm} (1 - p_l)^{(j-h)m} E[\gamma_k | k = h]. \quad (16)$$

### 5.3 Event Characteristic

The energy-efficiency of our protocol depends on the event characteristics. In the whole network clustering approach, all nodes form and maintain the clusters without any event occurrences. Event-to-sink approach forces the cluster formation with the event occurrences in the region between the event and the sink. Hence, the event occurrence frequency highly affects the energy-efficiency of our approach. We find the average event occurrence frequency,  $f_{ef}$ , for the energy-efficient operation of our approach and the steps to find it can be found in Appendix. The upper limit of event occurrence frequency for different event occurrence frequency is given in the following table. The below table shows that as the event radius

increases, the event frequency satisfying the energy-efficiency decreases due to the resulting larger event routing corridor.

	$R = 0.1L$	$R = 0.2L$	$R = 0.3L$
$f_{ef}$	15.57	6.80	4.12

## 6 PERFORMANCE EVALUATION

We establish a simulation environment in MATLAB to study the performance of our clustering protocol. We perform simulations to test our protocol in terms of the number of data and control packet transmissions, the time steps for clustering, the connectivity, the energy consumed for the clustering, and changes in the distribution of speed of the nodes and coefficient of the speed in node weight calculation. Nodes are distributed randomly in 100 m x 100 m area. The common parameter that we change is the distance between event and sink,  $D_{es}$ . The sink is located at (0, 50) and events are generated on the horizontal line at 50 m level. Simulations are performed for 100 different topologies and the results are averaged to draw a conclusion about our algorithm. Our algorithm is compared in the simulations with DSAC [29], SOC [26] and Maximum Node Biclique (MNB) Clustering in [15].

The other parameters, which are the CR transmission range,  $r$ , and the event radius,  $R$ , are changed during the simulations. These parameters are used since their changes reflect the performance of our protocol, and the eligibility corridor and spectrum-aware clustering are highly related to these parameters. Furthermore, there are 5 licensed channels in the system. Each PU utilizes a randomly selected channel from these licensed channels.

In our simulations, the coefficients in the weight calculation are selected as  $w_1 = 0.4$ ,  $w_2 = 0.3$ ,  $w_3 = 0.1$ ,  $w_4 = 0.1$ , and  $w_5 = 0.1$ . The sum of the coefficients is 1 and these coefficients can be changed according to the importance of the parameters in the system.

### 6.1 Data and Control Packet Exchange

We investigate the benefits of clustering on the reduction of data packets at the cost of control packets. We explore the performance of mESAC and no clustering (NC) approaches with and without mobility. In NC case, the packets are routed to the neighbor having at least one common channel and located nearest to the sink. The nearest neighbor to the sink is known by the neighbor discovery [13] as in mESAC. The control packet exchange is performed by the receiver and the transmitter to determine the communication channel. Each event detecting node generates one event packet,

this is the worst possible case since the number of data packet is the minimum possible. As the number of event packets increases, the number of data packets prevails the number of the control packets and the effect of data packets becomes predominant. Furthermore, PUs are always ON. Hence, the time variation of the spectrum availability does not change but the spatial spectrum availability changes due to the random placement of the primary users. The always ON PU is the worst case possible for the spectrum opportunity for CRSN nodes since there would be no spectrum opportunity in the time domain. In the first study, we observe the change in the number of packets in terms of  $R$  and  $D_{es}$  without mobility. We study 4 cases and the specifications of these cases can be seen in the following table.

Case 1	Case 2	Case 3	Case 4
$R = 20$	$R = 25$	$R = 20$	$R = 25$
$D_{es} = 50$	$D_{es} = 50$	$D_{es} = 75$	$D_{es} = 75$

As seen in Fig. 4, in all cases, the number of transmitted data packets is significantly reduced as compared to the NC case due to the aggregation. The number of control packets increases due to the formation of clustering. On the other hand, there is substantial number of control packet exchange for NC approach to coordinate the communication. It does not benefit from the data aggregation which is one of the main motivation to use clustering. The efficiency of the clustering may be less in Case 3 and Case 4 which mainly depends on the size of the data and control packets. This is the result of the eligibility corridor formation. If the size of the data packet is 2 times the size of the control packet, mESAC becomes inefficient as seen in Fig. 4 in Case 3. If the size of the data packet is 4 times greater than the control packet, mESAC becomes more efficient in all cases.

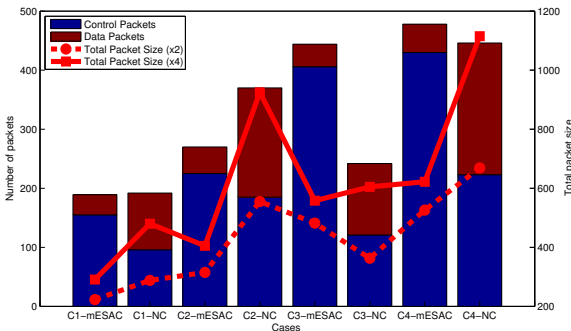


Fig. 4: Data and control packet exchange for different cases without mobility.

As the event radius increases, there would be more event detecting nodes. It results in more data and control packet transmissions. Furthermore, the increase in  $D_{es}$  causes more control packet exchange since the event-to-sink corridor is expanded. For NC approach, the increase in  $D_{es}$  causes more data and control packet transmission due to the increase in the number

of hops between the event and the sink.

The effect of the mobility is also investigated. In this set of study, the random waypoint mobility model is used and the mean of the nodes speed is changed. CRSN nodes move in random directions in the magnitude of their speed at each time step. We have studied 4 cases where  $R = 20 m$  which can be tabulated as follows.

Case 1	Case 2	Case 3	Case 4
$\mu = 0$	$\mu = 1$	$\mu = 0$	$\mu = 1$
$D_{es} = 50$	$D_{es} = 50$	$D_{es} = 75$	$D_{es} = 75$

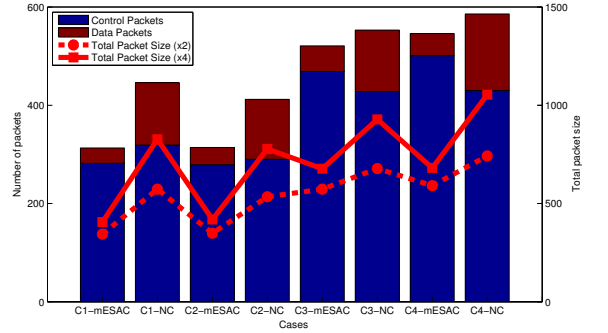


Fig. 5: Data and control packet exchange for different cases with mobility.

Mobility increases the control packet exchange since the neighborhood of the nodes changes and the routes between the event and the sink changes accordingly. There may be a need of control packet exchanges to establish the routing path from the event to the sink without the clustering. However, the clustering has a resistance to the mobility since it groups the nodes to use other members in case of the changes in the network structure due to the mobility. As seen in Fig. 5, as we increase the speed of the nodes, the number of control packet exchange increases in all cases for the two approaches. In mESAC, the cluster nodes are reserved for the possible relay node selection in the clustering structure. However, in NC approach, the transmitter may lose its contact with the receiver node that is located the nearest to the sink which may cause extra control packet exchange. In this set of study, since the size of the data packet is greater than the size of the control packets, mESAC is more efficient in all the cases as seen Fig. 5.

## 6.2 Time steps for clustering

Establishing clusters requires some time steps since they are formed from scratch. In Section 4.2.1, we stated that the time complexity of selecting cluster-heads is related to the diameter of the eligibility corridor. In addition to the selection of cluster-heads, forming the eligibility corridor also requires time.

As we increase  $r$ , the diameter of the eligibility corridor decreases. It results in a bit less time for clustering in mESAC as seen in Fig. 6. Furthermore, the

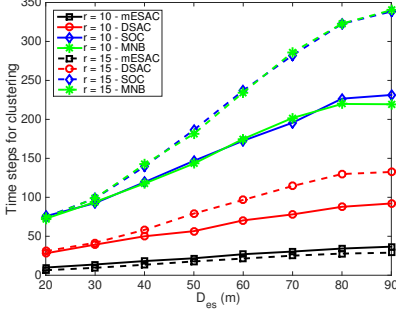


Fig. 6: Timesteps required for clustering vs.  $D_{es}$  with respect to  $r$ .

time steps required for clustering increases with the number of hops to reach the sink. Hence, the increase in the distance between event and sink increases the required time steps. Our algorithm, mESAC, performs the best for event-driven clustering scenarios as seen in Fig. 6. DSAC, SOC and MNB need more control packet exchange for clustering, therefore, they need more time steps to finalize the cluster formation.

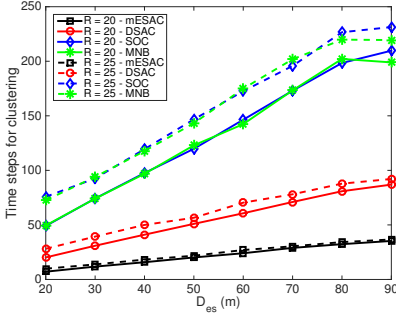


Fig. 7: Timesteps required for clustering vs.  $D_{es}$  with respect to  $R$ .

The time steps required for clustering slightly increase with an increase in the event radius. This is an expected result since the clustering time depends on the size of the eligibility corridor. This slight difference is due to the expansion of eligibility corridor with the increase in the event radius as seen in Fig. 7. This case is the same for the other protocols for an increase in event radius. DSAC performs worse than mESAC and better than SOC and MNB. DSAC needs more message exchange than mESAC since each node agglomerate with its neighbor to form clusters and the formed clusters again combine with neighboring clusters. This increases the message exchange for clustering, and hence the time steps required for clustering.

### 6.3 Connectivity

We define the connectivity as the ratio of average number of common channels of a cluster with its neighbor clusters. In mESAC, we select the nodes with the highest number of vacant channel in their neighborhood to increase the possible number of common channels with neighboring cluster-heads.

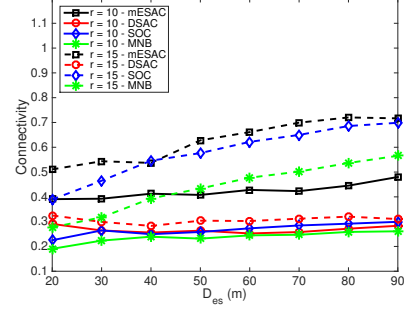


Fig. 8: Connectivity vs.  $D_{es}$  with respect to  $r$ .

In our algorithm, the cluster-heads are selected in one-hop neighborhood of the eligible nodes. The neighborhood depends on the transmission range of the nodes. If  $r$  increases, it will increase the distance between the neighboring cluster-heads. The occupied channel in a neighborhood is more likely to be vacant for more distant nodes. Thus, it will increase the probability of having higher number of vacant channels between clusters. As seen in Fig. 8, the increase in the transmission range increases the connectivity.

The reason for better performance is that the nodes having the highest number of vacant channels in its neighborhood in mESAC are selected as cluster-heads. Having the highest number of vacant channels increases the possibility to be connected with the neighboring cluster-heads. On the other hand, SOC establishes a balance between the number of cluster members and cluster channels. Hence, it does not consider the inter-cluster connectivity. mESAC establishes more connected clusters as seen in Fig 8.

### 6.4 Energy consumption for clustering

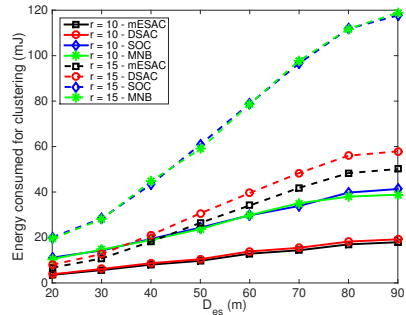


Fig. 9: Energy consumed for clustering vs.  $D_{es}$  with respect to  $r$ .

Energy consumption is a major design concern for our algorithm since CRSN nodes have limited battery power. In this set of study, we observe the effects of  $D_{es}$  and  $r$  on the energy consumption for clustering. As shown in Fig. 9, an increase in the transmission radius increases the energy consumption since more nodes consume energy for cluster formation process.

SOC has three steps to form and to finalize clusters. Every node informs its neighbors in each step and

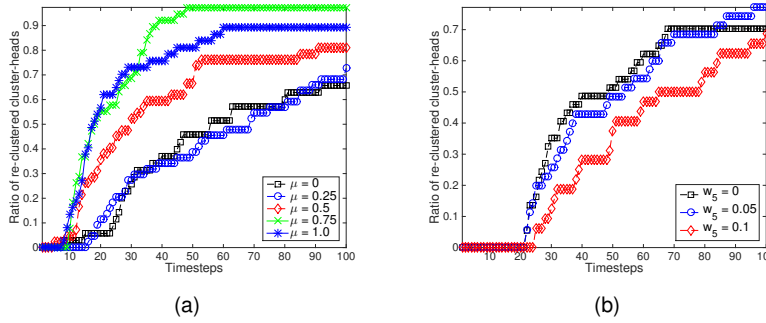


Fig. 10: Ratio of the re-clustered cluster-heads vs. timesteps for (a) different  $\mu$ , (b) different  $w_5$ .

it causes energy consumption. MNB calculates the weight of each node and forms clusters accordingly as in SOC. Hence, mESAC consumes less energy for clustering than SOC which can be seen in Fig. 9 for different  $r$ . DSAC consumes more energy than mESAC due to the iterative agglomeration process.

### 6.5 Re-clustering due to mobility

Mobility of the nodes causes topology changes. It directly affects the cluster stability since the cluster-head may lose its connection with its members due to the changes in the neighborhood and the vacant channel. In this set of experiment, the event is generated at (100,50), and 300 nodes are distributed in the network. The nodes move in the network according to uniform distribution with a density of 1. We observe the effects of the changes in the mean of the uniform distribution. We change the mean of the distribution from  $\mu = 0$  to  $\mu = 1$ . As seen in Fig. 10(a), if we increase the mean of the node speed distribution, the average node speed increases and re-clustering is triggered earlier and more cluster-heads are forced to re-cluster in smaller amount of time. Fig. 10(a) shows that the slope increases due to increasing  $\mu$ . The ratio of the re-clustered cluster-heads reaches 1 in 40 time steps when  $\mu = 1.0$  m/s.

The coefficient of the speed in the node weight calculation is important for the stability of cluster-heads against the mobility. We change the coefficient of the node speed,  $w_5$ , with the coefficient of the vacant channels,  $w_2$ , accordingly to make the sum of all weight equal to one. As seen in Fig. 10(b), if the weight of the speed in the node weight calculation increases, more cluster-heads are forced to re-cluster for decreasing  $w_5$ . Hence, the increase in the coefficient  $w_5$  makes the cluster-heads more immune to the topology changes.

## 7 CONCLUSION

In this paper, we address the need for event-to-sink coordination scheme by presenting a clustering protocol for mobile CRSNs to minimize the energy consumption and to coordinate the communication

between event and sink. We propose a two phase clustering algorithm which determines eligible nodes for clustering and performs clustering based on node weights according to some system parameters in mobile CRSNs. Eligible node degree, number of available channels, remaining energy of nodes, the distance to the sink and the speed of the nodes are selected parameters to increase the energy-efficiency and connectivity of clusters. We further investigate the average re-clustering probability and how our approach decreases this probability. We obtain the maximum average event frequency to achieve energy efficiency for our protocol.

## ACKNOWLEDGMENTS

The work of M. Ozger, and O. B. Akan was supported by The Scientific and Technological Research Council of Turkey (TUBITAK) under grant #110E249.

## REFERENCES

- [1] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.
- [2] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation / dynamic spectrum access / cognitive radio wireless networks: A survey," *Computer Networks (Elsevier)*, Sept. 2006.
- [3] O. B. Akan, O. Karli, and O. Ergul, "Cognitive radio sensor networks," *IEEE Network*, vol. 23, no. 4, pp. 34-40, July 2009.
- [4] X. Li, D. Wang, J. McNair, and J. Chen, "Residual energy aware channel assignment in cognitive radio sensor networks," in *Proc. of IEEE WCNC 2011, Quintana-roo, Mexico*, 2011, pp. 398-403.
- [5] J. Han, W. Jeon, and D. Jeong, "Energy efficient channel management scheme for cognitive radio sensor networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1905-1909, May 2011.
- [6] M. C. Oto, and O. B. Akan, "Energy-efficient packet size optimization for cognitive radio sensor networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1544-1553, April 2012.
- [7] B. Gulbahar, and O. B. Akan, "Information Theoretical Optimization Gains in Energy Adaptive Data Gathering and Relaying in Cognitive Radio Sensor Networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1788-1796, May 2012.
- [8] Z. Liang, S. Feng, D. Zhao, and X. S. Shen, "Delay Performance Analysis for Supporting Real-Time Traffic in a Cognitive Radio Sensor Network," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 325-335, Jan. 2011.
- [9] J. Yu and P. Chong, "A Survey of Clustering Schemes for Mobile Ad Hoc Networks," in *IEEE Commun. Surveys Tutorials*, vol. 7, no. 1, pp. 32-48, Jan. 2005.



- [10] K.-L. A. Yau, N. Ramli, W. Hashim, and H. Mohamad, "Clustering algorithms for Cognitive Radio networks: A survey," *Journal of Network and Computer Applications (Elsevier)*, vol. 45, pp. 79-95, Oct. 2014.
- [11] O. Younes, and N. Thomas, "Analysis of the Expected Number of Hops in mobile Ad Hoc Networks with Random Waypoint Mobility," *Journal of Network and Electronic Notes in Theoretical Computer Science (Elsevier)*, vol. 275, pp. 143-158, Sep. 2011.
- [12] J. Jia, Q. Zhang, and X. Shen, "HC-MAC: A hardware-constrained cognitive MAC for efficient spectrum management," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 106-117, Jan. 2008.
- [13] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE DySPAN*, Nov. 2005, pp. 137-143.
- [14] H. Su, and X. Zhang, "Design and analysis of a multi-channel cognitive MAC protocol for dynamic access spectrum networks," in *Proc. of IEEE MILCOM 2008*, San Diego, CA, 2008.
- [15] M. Bradonjic and L. Lazos, "Graph-based criteria for spectrum-aware clustering in cognitive radio networks," *Ad Hoc Networks (Elsevier)*, vol. 10, no. 1, pp. 75-94, Jan. 2012.
- [16] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy Efficient Communication Protocol for Wireless Microsensor Networks," in *Proc. of HICSS 2000*, 2000.
- [17] M. Ozger, O.B. Akan, "Event-driven Spectrum-Aware Clustering in Cognitive Radio Sensor Networks," in *Proc. of IEEE INFOCOM 2013*, Turin, Italy, Apr. 2013, pp. 1483-1491.
- [18] S. Wang, Y. Wang, J. P. Coon, and A. Doufexi, "Energy-Efficient Spectrum Sensing and Access for Cognitive Radio Networks," *IEEE Trans. Veh. Technol.*, vol. 61, no. 2, pp. 906-912, Feb. 2012.
- [19] O. D. Incel, "A survey on multi-channel communication in wireless sensor networks," *Computer Networks*, vol. 55, no. 13, 2011, pp. 3081-3099.
- [20] A. Gupta, C. Gui, and P. Mohapatra, "Exploiting multi-channel clustering for power-efficiency in sensor networks," in *Proc. International Conference on Communication System Software and Middleware*, pp. 1-10, 2006.
- [21] Y. Wu, J. A. Stankovic, T. He, and S. Lin, "Realistic and Efficient Multi-Channel Communications in Wireless Sensor Networks," in *Proc. INFOCOM 2008*, Phoenix, AZ, USA, 2008.
- [22] A. Bereketi, and O. B. Akan, "Event-to-Sink Directed Clustering in Wireless Sensor Networks," in *Proc. of IEEE WCNC 2009*, Budapest, Hungary, 2009, pp. 1-6.
- [23] M. C. Vuran, O. B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: Theory and applications for wireless sensor networks," *Comput. Netw.*, vol. 45, no. 3, pp. 245-259, 2004.
- [24] J. Zhao, H. Zheng, and G.-H. Yang, "Distributed Coordination in Dynamic Spectrum Allocation Networks," in *Proc. of IEEE DySPAN 2005*, Baltimore, MD, 2005.
- [25] T. Chen, H. Zhang, G.M. Maggio, and I. Chlamtac, "CogMesh: A Cluster-Based Cognitive Radio Network," in *Proc. of IEEE DySPAN 2007*, Dublin, Ireland, 2007, pp. 168-178.
- [26] L. Lazos, S. Liu, and M. Krunz, "Spectrum opportunity-based control channel assignment in cognitive radio networks," in *Proc. of IEEE SECON 2009*, Piscataway, NJ, 2009, pp. 135-143.
- [27] K.E. Baddour, O. Ureten, and T. J. Willink, "Efficient Clustering of Cognitive Radio Networks Using Affinity Propagation," in *Proc. of IEEE ICCCN 2009*, San Francisco, CA, 2009.
- [28] D. Li, and J. Gross, "Robust Clustering of Ad-Hoc Cognitive Radio Networks under Opportunistic Spectrum Access," in *Proc. of IEEE ICC 2011*, Kyoto, Japan, 2011.
- [29] H. Zhang, Z. Zhang, H. Dai, R. Yin, and X. Chen, "Distributed spectrum-aware clustering in cognitive radio sensor networks," in *Proc. of IEEE GLOBECOM 2011*, Houston, TX, 2011, pp. 1-6.
- [30] S. Basagni, I. Chlamtac, and A. Farago, "A generalized clustering algorithm for peer-to-peer networks," in *Proc. of Workshop on Algorithmic Aspects of Communication (satellite workshop of ICALP)*, Jul., 1997.
- [31] M. Chatterjee, S. K. Das, and D. Turgut, "WCA: A Weighted Clustering Algorithm for Mobile Ad hoc Networks," *Journal of Cluster Computing (Special Issue on Mobile Ad hoc Networks)*, vol. 5, pp. 193-204, 2002.
- [32] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio networks," in *Proc. of DYSPAN 2005*, Baltimore, MD, USA, 2005, pp. 137-143.
- [33] M. Gerla and J.T.C. Tsai, "Multicenter, mobile, multimedia radio network," *Wireless Networks* 1(3), pp. 255-265, 1995.
- [34] L.-H. Yen, C. Yu, and Y.-M. Cheng, "Expected k-coverage in wireless sensor networks," *Ad Hoc Networks*, vol. 5(4), pp. 636-650, Sept. 2006.



**Mustafa Ozger** [S'12] received his B.Sc. degree in Electrical and Electronics Engineering from Middle East Technical University, Ankara, Turkey, in 2011 and received his M.Sc. degree in Electrical and Electronics Engineering from Koc University, Istanbul, Turkey in 2013. He is currently a research assistant in the Next-generation and Wireless Communication Laboratory and pursuing his Ph.D. degree at the Electrical and Electronics Engineering Department, Koc University, Istanbul, Turkey. His current research interests include cognitive radio networks, and cognitive radio sensor networks.

**Etimad Fadel** received her BS degree in Computer Science from King Abdulaziz University (KAU), Saudi Arabia, in 1994 and Mphil/PhD degrees in computer science from De Montfort University (DMU), United Kingdom, in 2007. Currently, she is an Assistant Professor with the Computer Science Department at KAU, Saudi Arabia. Her research interests are Wireless Networks, Internet of Things, Internet of Nano-Things, Sensor Networks and LTE-A Cellular Systems.



**Ozgur B. Akan** [M'00-SM'07] received his Ph.D. degree in electrical and computer engineering from the Broadband and Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology in 2004. He is currently a full professor with the Department of Electrical and Electronics Engineering, Koc University and the director of the Next-generation and Wireless Communications Laboratory. His current research interests are in wireless communications, nano-scale and molecular communications, and information theory. He is an Associate Editor of the IEEE Transactions on Communications, the IEEE Transactions on Vehicular Technology, the International Journal of Communication Systems (Wiley), the Nano Communication Networks Journal (Elsevier), and the European Transactions on Technology.