IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON NEW TECHNOLOGIES FOR SMART FARMING 4.0: RESEARCH CHALLENGES AND OPPORTUNITIES

Received July 9, 2019, accepted August 19, 2019, date of publication August 27, 2019, date of current version September 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2937881

Affordable Broad Agile Farming System for Rural and Remote Area

MOHAMMAD AMMAD UDDIN[1], MUHAMMAD AYAZ[1], EL-HADI M. AGGOUNE[1], ALI MANSOUR[2], AND DENIS LE JEUNE[2]

[1]Sensor Networks and Cellular Systems Research Center, University of Tabuk, Tabuk 71491, Saudi Arabia
[2]Lab STICC, ENSTA Bretagne, 29806 Brest, France

Corresponding author: Mohammad Ammad Uddin (mashfaq@ut.edu.sa)

This work was supported in part by the SNCS Research Center and in part by the Deanship of Scientific Research at the University of Tabuk, Tabuk, Saudi Arabia.

ABSTRACT We develop a fast-deployed crop health monitoring system using state-of-the-art technologies to collect data from crop fields in order to take appropriate and timely actions. For the proposed resource optimized system, Saudi Arabian agriculture is taken as a case study. To achieve the desire goals, we harness IoT (Internet of Things) and drones in agriculture to establish rapid system deployment. This paper focuses on the data collection from crop field by organizing heterogeneous IoT devices in clusters and localise them for data harvesting. Clusters are formed by considering the path of UAVs, sensors heterogeneity, weather conditions, fluctuation of sensor nodes, and the communication cost of IoT devices. For localisation, carrying larger or heavier arrays of antennas and receivers with a small size UAV is also a major issue considered in this paper. Hence, we introduce a dynamic clustering and virtual antenna array to develop a complete data collection scheme supported by simulations and experimental tests with proof-of-concept devices. The results are analysed and found promising in terms of energy efficiency, throughput, ease of use, and deployment time. Whole the system is developed with the concept that it can install in rural and remote area with minimum deployment time and agile enough that can collect data in worst conditions (bad weather, hostile environment, fluctuating nodes, poor infrastructure, with or without an established network). In broader sense it can map easily in many similar applications where data is needed to be harvested from a wide range of heterogeneous sensors without existing any infrastructure and ground topology.

INDEX TERMS Smart farming, Internet of Things (IoT), clustering, localization, virtual array antenna, dynamic data collection.

I. INTRODUCTION

The Kingdom of Saudi Arabia (KSA) is the 12th largest country in the world at 2,149,690 km^2 but remains not highly populated. KSA spends $24 billion annually to import food to fulfil the demand of 32 million residents [1]. While the country has the potential to grow most of the crops needed by the population, the major hurdle is a scarcity of water, especially when the underground water supply is not replenished with rare rainfalls. To save on the resource, crops are grown in large circular farms irrigated through central pivot systems. Managing hundreds of thousands of these dispersed crop circles in remote areas is an additional challenge. Considering these issues, this research is motivated to develop a complete solution for smart agriculture. The main objective is to modernise the agriculture sector with the state-of-the-art technologies to manage cop circles with a minimal human interaction. Further to increase the quantity and quality of food by utilizing precise amounts of resources (e.g., pesticides, fertilisers, and water).

The agricultural environment in KSA is considered as a case study as it faces more challenges than most of the other countries, including scarcity of water, mostly deserted lands, lack or limited existence of infrastructure, and adverse weather conditions characterized by high temperatures, frequent sand or dust storms, and dry air. Emerging tools and technologies, such as small and smart sensors [2], wireless communication [3], and Unmanned Aerial Vehicles (UAVs) [4], are introduced in traditional agriculture to overcome these challenges. Additionally, better food quality, higher yields, optimization of resources and economy can be achieved. In this research, smart technology is coupled

The associate editor coordinating the review of this article and approving it for publication was Kun Mean Hou.

with agriculture goals through the following independent steps:

1. A wide range of abundant heterogeneous sensor nodes is deployed across crop field. These sensors are heterogeneous in nature, size, purpose, and energy, and they monitor the environment, soil, and crops.
2. UAVs are utilised to construct an inexpensive, convenient, and swiftly-deployed communication infrastructure among the sensors and end-users.
3. Dynamic clusters of sensor nodes are formed for efficient data collection. Heterogeneous sensing devices are grouped to form a physical or logical cluster according to specified criteria, such as the type of required data, geographical location, communication characteristics, the path of the UAV, and node similarities. The selection of Cluster Head (CH) is an important task in our clustering scheme, which selects the node in good health, the best overall specification, and most suitable for UAV connectivity.
4. A localisation technique explores the ground sensing devices by an aerial UAV and connects them for collecting data.

An important issue addressed in this research is data gathering from many sensors unable to establish an ad-hoc communication due to widespread deployment, geographical constraints, weather conditions or power constraints. An Affordable, broaD and Agile farMing (ADAM) system especially for rural and remote area is proposed in this research paper. This ADAM is composed of UAVs and field sensor nodes and is not limited to prerequisites, such as communication infrastructure, network updates to maintain routes and clusters, installation of special CH nodes for data gathering, and knowledge of UAV paths. All sensor nodes are considered heterogeneous, location unaware, inexpensive, and remain unattended. The function of the field sensor nodes is limited to only sense and sleep during daily routines to preserve the node's energy and maximally prolong the network life. Node energy is further optimised by introducing a dual frequency support in each sensor. A high-power radio transmitter unit in each node remains off until a low power radio receiver picks up an UAV activation call. Figure 1 illustrates the lifecycle of our proposed system (ADAM) consisting of seven steps:

1. UAV is the primary component performing as a data sink and the means of communication between the sensor nodes and the end-user. The lifecycle begins with UAV beaconing, an activation call intended for sensor nodes in close vicinity. Typically, we require specific data (such as a fruit size identification or soil moisture measurements) from within a target area (e.g., the area of interest) to monitor, whole the data from the entire field is rarely needed.
2. Only selected sensor nodes targeted by the beacon become activated in Step 2.
3. The UAV dynamically identifies clusters as it continues along its scheduled path in Step 3. Next, the CH is selected, and two scenarios are important: either none of the field sensors are in good health and capable of communicating with the UAV, or the UAV becomes saturated with numerous responses from ground nodes and is unable to locate all simultaneously.
4. The shunting in Step 4 is introduced to handle both conditions to choose an acceptable number of field sensors for further processing. In this step, some nodes can be pushed or pulled from the process to allow them to be within an acceptable range. The UAV coordinates with the sensor nodes to assist in the selection of the CH.
5. At Step 5, the CH is selected and announced to all nodes.
6. The next step estimates the location of selected CH node to visit for data collection in Step 7. A specially-designed, energy efficient, lightweight, and small antenna is proposed to accomplish this task.
7. The data collection is performed in the final Step 7 during which the CH transmits the collected data to the UAV through a high-bandwidth dedicated link.

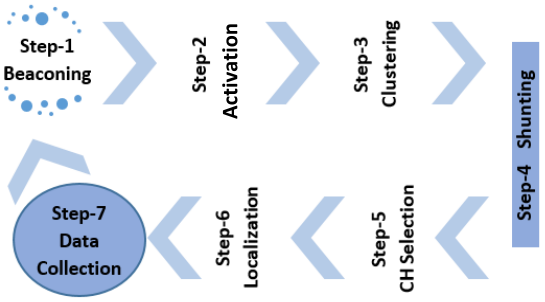

FIGURE 1. Proposed system (ADAM) lifecycle.

This lifecycle continues until all data is harvested from the target area. The system spans in three distinct dimensions including dynamic runtime clustering, real-time CH selection, and localisation of ground sensors by the UAV.

The remainder of the paper is organised as follows. The next section reviews the dynamic cluster formation and cluster head selection. In Section III, we elaborate on the architecture of the ADAM system in terms of layers and phases. Section IV discusses the localisation process for which we propose a lightweight virtual phase array antenna carried by the UAV to locate the position of the field sensors. In Section V, a simulation study of the clustering technique is performed with results discussed, and the localisation system is simulated in Section VI along with the results. A proof of concept development and performance testing of these devices are provided in Section VII followed by our conclusions about the proposed system in Section VIII.

II. PROPOSED UAV-ASSISTED DYNAMIC CLUSTERING ARCHITECTURE

Clustering is a prominent feature for enhancing the reliability and lifetime of a network, which is a key issue for any

Wireless Sensor Network (WSN). Existing clustering techniques are categorized into four broad classes which are static sink immobile nodes (classical HEED [5], LEACH [6]), static sink mobile nodes (cellular Network [7]), mobile sink immobile nodes (Rendez-vous base routing [8]–[11]), and mobile sink mobile nodes (ad-hoc routing [12]). This research relates to the third category where all agriculture sensors are considered static, and the UAV (i.e., sink node) is used to traverse and collect data from the crop field. In the case of static sensor nodes, a predefined distance vector (by periodic updates) for the routing and clustering schemes fits well into the data collection process. This type of clustering is not feasible in our case. First, bad weather and a harsh environment cause many sensor nodes to become unavailable. Second, the type of target sensors and data are not known in advance (e.g., temperature sensors, fruit size sensors or soil sensors). Third, the path of the UAV is not known to the sensor nodes as only the sensors residing along the path will be activated for data collection.

In our scenario, it is rare for a pre-determined CH node to reside along the path of an UAV while being simultaneously available. The Rendez-vous base and network-defined clustering are proposed in the literature, where the Rendez-vous Point (RVP) and CH along with the ground topology are always maintained by sending periodic updates (PUs). The UAV must collect data from these RVPs by following the predefined ground network topology. Here, PUs are overhead, and the UAV must search for and visit the network assigned to the RVPs to collect data. This will deflect the UAV from its path affecting the throughput of the system. As to the best of our knowledge, all these factors (UAV path, node residual energy, node energy consumption, type of data, and area of interest) have not been previously considered collectively as clustering criterion.

With our runtime dynamic clustering scheme proposed, all field sensors are considered initially indistinguishable, and anyone may be selected as the CH based on the factors discussed above.

A. DYNAMIC CLUSTER HEAD SELECTION

All active nodes have an equal opportunity to be selected as the CH per a Bayesian probability [13].

Assume n sensor nodes $S = (s_1, s_2, \ldots, s_n)$ are installed in a field, and each node s_i has z independent attributes denoted by the vector $\mathcal{A} = (a_1, a_2, \ldots, a_z)$. If a node s_i can either be the Cluster Head (CH) or a Cluster Member (CM), then it has two representative states denoted as $State = (CH, CM)$. Some of the prior and posterior probabilities can be defined as:

$P\left(s_i = CH|a_{ij}\right)$. Posterior probability of any node s_i to be selected as CH based on its attribute a_{ij} (jth attribute of ith node).

$P(s_i = CH)$. Prior probability of any node to become the CH.

$P(s_i = CM)$. Prior probability of a s_i sensor node to be a CM.

$P(a_{ij}|s_i = CH)$. Likelihood probability that the selected CH is based on a_{ij}.

$P(a_{ij}|s_i = CM)$. Probability that a CM node (not selected as the CH) has a good value of attribute a_{ij}.

$P(a_{ij})$. Prior probability of the ith node, jth attribute a_{ij} to have the best or highest value. The chance of a sensor node s_i to be selected as the CH by knowing one parameter a_{ij} is computed using the Bayesian theorem as follows.

$$\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|a_{\mathrm{ij}}\right) = \frac{\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}\right)\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i = \mathrm{CH}\right)}{\mathrm{P}(a_{\mathrm{ij}})}$$

Using the total probability we can write

$$\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|a_{\mathrm{ij}}\right) = \frac{\mathrm{P}\left(\mathrm{s}_i{=}\mathrm{CH}\right)\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i{=}\mathrm{CH}\right)}{\mathrm{P}\left(\mathrm{s}_i{=}\mathrm{CH}\right)\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i{=}\mathrm{CH}\right){+}\mathrm{P}\left(\mathrm{s}_i{=}\mathrm{CM}\right)\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i{=}\mathrm{CM}\right)} \quad (1)$$

If $\mathcal{A}_i = (a_1, a_2, \ldots, a_z)$ is the set of all parameters of a sensor node s_i (where all set members are independent in nature and known), then the conditional probability of all parameters knowing s_i as a CH is calculated as

$$\begin{aligned}&\mathrm{P}\left(a_{\mathrm{i1}}, \mathrm{a_{i2}}, \ldots, a_{\mathrm{iz}}|\mathrm{s}_i = \mathrm{CH}\right)\\ &= \prod_{j=1}^{z}\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i = \mathrm{CH}\right) = \prod_{j=1}^{z}\frac{\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|a_{\mathrm{ij}}\right)\mathrm{P}\left(a_{\mathrm{ij}}\right)}{P\left(\mathrm{s}_i = \mathrm{CH}\right)}\\ &= \frac{\prod_{j=1}^{z}\mathrm{P}_{ij}\mathrm{P}\left(a_{\mathrm{ij}}\right)}{[P\left(\mathrm{s}_i = \mathrm{CH}\right)]^z}\end{aligned} \quad (2)$$

$$\begin{aligned}&(a_{\mathrm{i1}}, \mathrm{a_{i2}}, \ldots, a_{\mathrm{iz}}|\mathrm{s}_i = \mathrm{CM})\\ &= \prod_{j=1}^{z}\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i = \mathrm{CM}\right) = \prod_{j=1}^{z}\frac{\mathrm{P}\left(\mathrm{s}_i = \mathrm{CM}|a_{\mathrm{ij}}\right)\mathrm{P}(a_{\mathrm{ij}})}{P(\mathrm{s}_i = \mathrm{CM})}\\ &= \frac{\prod_{j=1}^{z}(1 - \mathrm{P}_{ij})\mathrm{P}(a_{\mathrm{ij}})}{[1 - P\left(\mathrm{s}_i = \mathrm{CH}\right)]^z}\end{aligned} \quad (3)$$

If: The set of parameters $\mathcal{A}_i$ (parameter set of node s_i) are known, then P_i will be the probability $P\left(s_i = CH|\mathcal{A}_i\right)$ of sensor node s_i to be selected as the CH. P_{ij} is the probability written as $\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|a_{ij}\right)$, and is the chance of the ith sensor node s_i to be the CH by seeing its jth attribute a_{ij} from the parameters set $\mathcal{A}_i$.

Then:

$$\begin{aligned}&\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|a_{\mathrm{ij}}\right)\\ &= \frac{\mathrm{P}\left(\mathrm{s}_i{=}\ \mathrm{CH}\right)\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i = \mathrm{CH}\right)}{\mathrm{P}(a_{\mathrm{ij}})}\\ &= \frac{\mathrm{P}\left(\mathrm{s}_i{=}\mathrm{CH}\right)\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i{=}\mathrm{CH}\right)}{\mathrm{P}\left(\mathrm{s}_i{=}\mathrm{CH}\right)\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i{=}\mathrm{CH}\right){+}\mathrm{P}\left(\mathrm{s}_i{=}\mathrm{CM}\right)\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i{=}\mathrm{CM}\right)}\\ &\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|\mathcal{A}_i\right)\\ &= \frac{\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}\right)\mathrm{P}\left(\mathcal{A}i|\mathrm{s}_i = \mathrm{CH}\right)}{\mathrm{P}\left(\mathcal{A}_i\right)}\\ &= \frac{\mathrm{P}\left(\mathrm{s}_i{=}\mathrm{CH}\right)\mathrm{P}\left(\mathcal{A}_i|\mathrm{s}_i{=}\mathrm{CH}\right)}{\mathrm{P}\left(\mathrm{s}_i{=}\mathrm{CH}\right)\mathrm{P}\left(\mathcal{A}_i|\mathrm{s}_i{=}\mathrm{CH}\right){+}\mathrm{P}\left(\mathrm{s}_i{=}\mathrm{CM}\right)\mathrm{P}\left(\mathcal{A}_i|\mathrm{s}_i{=}\mathrm{CM}\right)}\end{aligned} \quad (4)$$

Substituting the values from (2) and (3) into (4), we obtain the equation can be derived, as shown at the bottom of this page

Next, consider that $P(a_{ij})$ is constant, then the previous equation may be simplified as

$$\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|\mathcal{A}_i\right) = \frac{\prod_{j=1}^{z} \mathrm{P}_{ij}}{\prod_{j=1}^{z} \mathrm{P}_{ij} + \left(\frac{\mathrm{P}(\mathrm{s}_i=\mathrm{CH})}{1-P(\mathrm{s}_i=\mathrm{CH})}\right)^{z-1} \prod_{j=1}^{z}(1-\mathrm{P}_{ij})} \tag{5}$$

If any node in the network has an equal prior chance to become a CH or CM, referred to as the ''not biased'' condition, then (1) may be written

$$\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|a_{\mathrm{ij}}\right) = \frac{\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i = \mathrm{CH}\right)}{\mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i = \mathrm{CH}\right) + \mathrm{P}\left(a_{\mathrm{ij}}|\mathrm{s}_i = \mathrm{CM}\right)} \tag{6}$$

Considering (6), the equation (5) can then be rewritten as

$$\begin{aligned}\mathrm{P}_i &= \mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|\mathcal{A}_i\right) \\ &= \frac{\mathrm{P}_{i1}.\mathrm{P}_{i2}\ldots\ldots\mathrm{P}_{iz}}{\mathrm{P}_{i1}.\mathrm{P}_{i2}\ldots\ldots\mathrm{P}_{iz}+(1-\mathrm{P}_{i1}).(1-\mathrm{P}_{i2})\ldots\ldots(1-\mathrm{P}_{iz})}\end{aligned} \tag{7}$$

Equation (7) is expressed in the log domain because it may result in a floating-point underflow problem.

$$\begin{aligned}\frac{1}{\mathrm{P}_i} - 1 &= \frac{(1-\mathrm{P}_{i1}).(1-\mathrm{P}_{i2})\ldots\ldots(1-\mathrm{P}_{iz})}{\mathrm{P}_{i1}.\mathrm{P}_{i2}\ldots\ldots\mathrm{P}_{iz}} \\ \Longrightarrow \ln\left(\frac{1}{\mathrm{P}_i} - 1\right) &= \sum_{j=1}^{z}\left(\ln(1-\mathrm{P}_{ij}) - \ln \mathrm{P}_{ij}\right)\end{aligned} \tag{8}$$

In the case where P_{ij} is close to zero,$\ln(P_{ij})$ can produce an issue, but the results remain correct asymptotically. If $P_{ij} \cong 0 \Rightarrow$ we don not have the right to pass from (7) to (8).

According to (7), if $P_{ij} \to 0^+$, then $P_i \to 0^+$. In this case, the log in the right-hand side of (8) will asymptotically be

$$P_{ij} \to 0^+ \Rightarrow ln\left(P_{ij}\right) \to -\infty$$

Then, the right-hand side of (8) will tend to

$$\sum_{j=1}^{z}\left(\ln(1-\mathrm{P}_{ij}) - \ln \mathrm{P}_{ij}\right) \underset{P_{ij\to 0^+}}{\longrightarrow} +\infty$$

Therefore, we concluded that

$$\ln\left(\frac{1}{P_i} - 1\right) \underset{P_{ij\to 0^+}}{\longrightarrow} +\infty$$

$$\Rightarrow \frac{1}{P_i} - 1 \underset{P_{ij\to 0^+}}{\longrightarrow} +\infty$$

$$\Rightarrow \frac{1}{P_i} \underset{P_{ij\to 0^+}}{\longrightarrow} +\infty \Rightarrow P_i \underset{P_{ij\to 0^+}}{\longrightarrow} o^+$$

Each node will calculate its probability to become CH and inform the UAV if it is above a threshold, as the reply to beacon message. The developed system consists of two components of ground sensing nodes and a flying UAV, the functionality of each is expressed as under.

B. UAV DESIGN

The design of UAV plays an important role throughout the entire process. The UAV beacon triggers the system initially, and later it performs the major tasks of cluster formation, CH selection, localisation, communication, and data gathering. The UAV beacon (e.g., the activation call for the sensors) has four parameters:

1. A sensor type needed to activate a group of sensors.
2. A data collection height that is initially proposed by the UAV and can be negotiated later. The sensor nodes with enough energy and range to communicate over at least this distance can be a candidate for the CH.
3. A threshold value set to limit the number of sensors competing for the CH selection.
4. The trailer containing additional details used for the Error Detection Code (EDC) or any other necessary information.

Type	Height	Threshold	Others
6 byte	2 byte	1byte	2 bytes
Header		**Payload**	**Trailer**

The UAV is the most important component in our system, since it maintains the majority of the computing, storing, and communication burden to allow the sensor nodes to be light and simple. The working cycle of the UAV is shown as a flow chart in Figure 2. The variables and notations of the algorithm are listed in Table 1.

C. SENSOR NODE DESIGN

Different sensors are installed in the crop field to monitor various parameters about the environment, crop, and soil. These sensor nodes are assumed to have the following properties:

- Inexpensive, GPS-free, and function unattended.
- Enabled for processing with on-board memory.

$$\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|\mathcal{A}_i\right) = \frac{\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}\right)\frac{\prod_{j=1}^{z}\mathrm{P}_{ij}\mathrm{P}(a_{\mathrm{ij}})}{[P(\mathrm{s}_i=\mathrm{CH})]^z}}{\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}\right)\frac{\prod_{j=1}^{z}\mathrm{P}_{ij}\mathrm{P}(a_{\mathrm{ij}})}{[P(\mathrm{s}_i=\mathrm{CH})]^z} + [1-P\left(\mathrm{s}_i = \mathrm{CH}\right)]\frac{\prod_{j=1}^{z}(1-\mathrm{P}_{ij})\mathrm{P}(a_{\mathrm{ij}})}{[1-P(\mathrm{s}_i=\mathrm{CH})]^z}}$$

$$\mathrm{P}\left(\mathrm{s}_i = \mathrm{CH}|\mathcal{A}_i\right) = \frac{\frac{1}{\mathrm{P}(\mathrm{s}_i=\mathrm{CH})^{z-1}}\prod_{j=1}^{z}\mathrm{P}_{ij}\mathrm{P}(a_{\mathrm{ij}})}{\frac{1}{\mathrm{P}(\mathrm{s}_i=\mathrm{CH})^{z-1}}\prod_{j=1}^{z}\mathrm{P}_{ij}\mathrm{P}(a_{\mathrm{ij}}) + \frac{1}{[1-P(\mathrm{s}_i=\mathrm{CH})]^{z-1}}\prod_{j=1}^{z}(1-\mathrm{P}_{ij})\mathrm{P}(a_{\mathrm{ij}})}$$

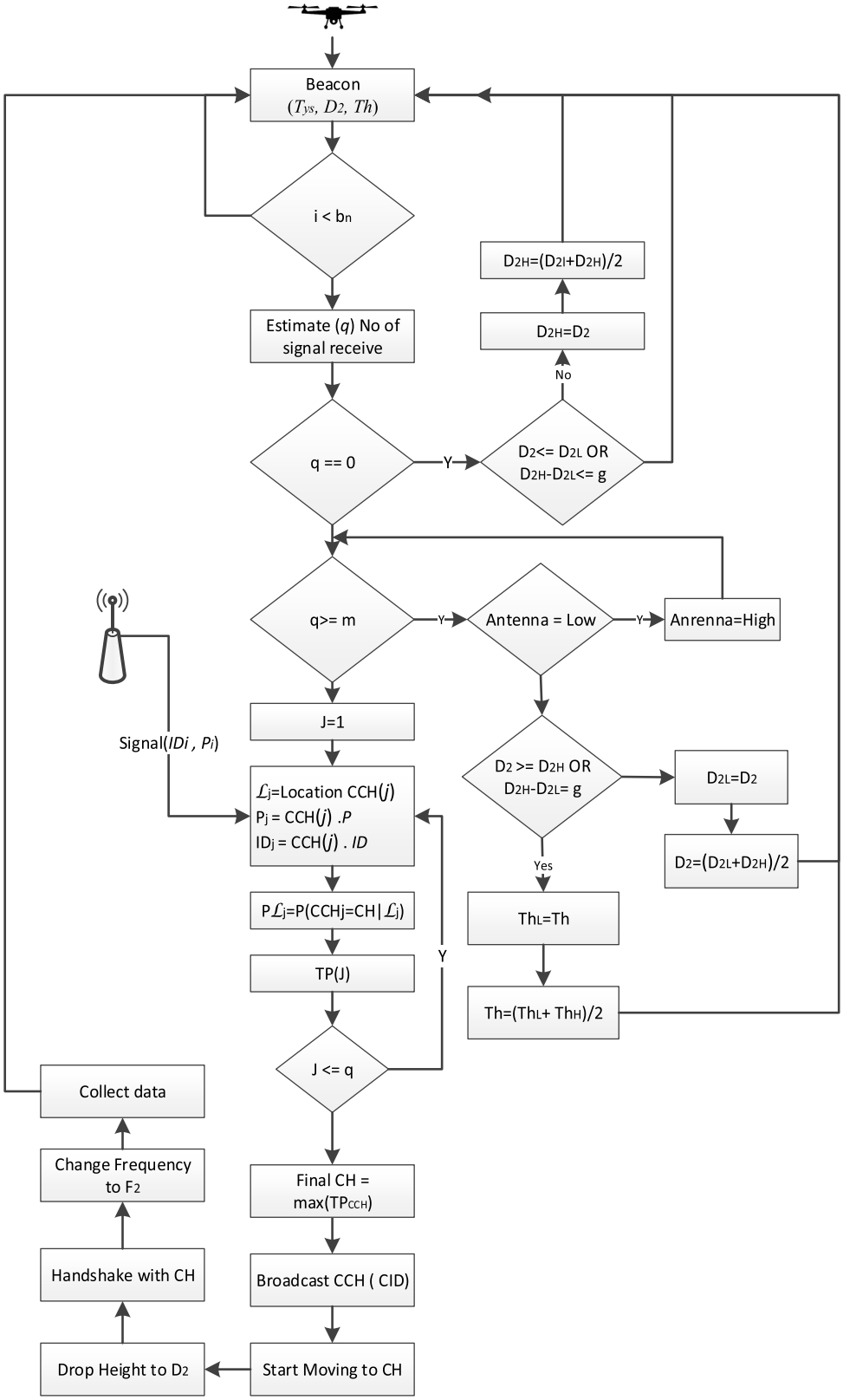

FIGURE 2. A flow chart representing the working of UAV.

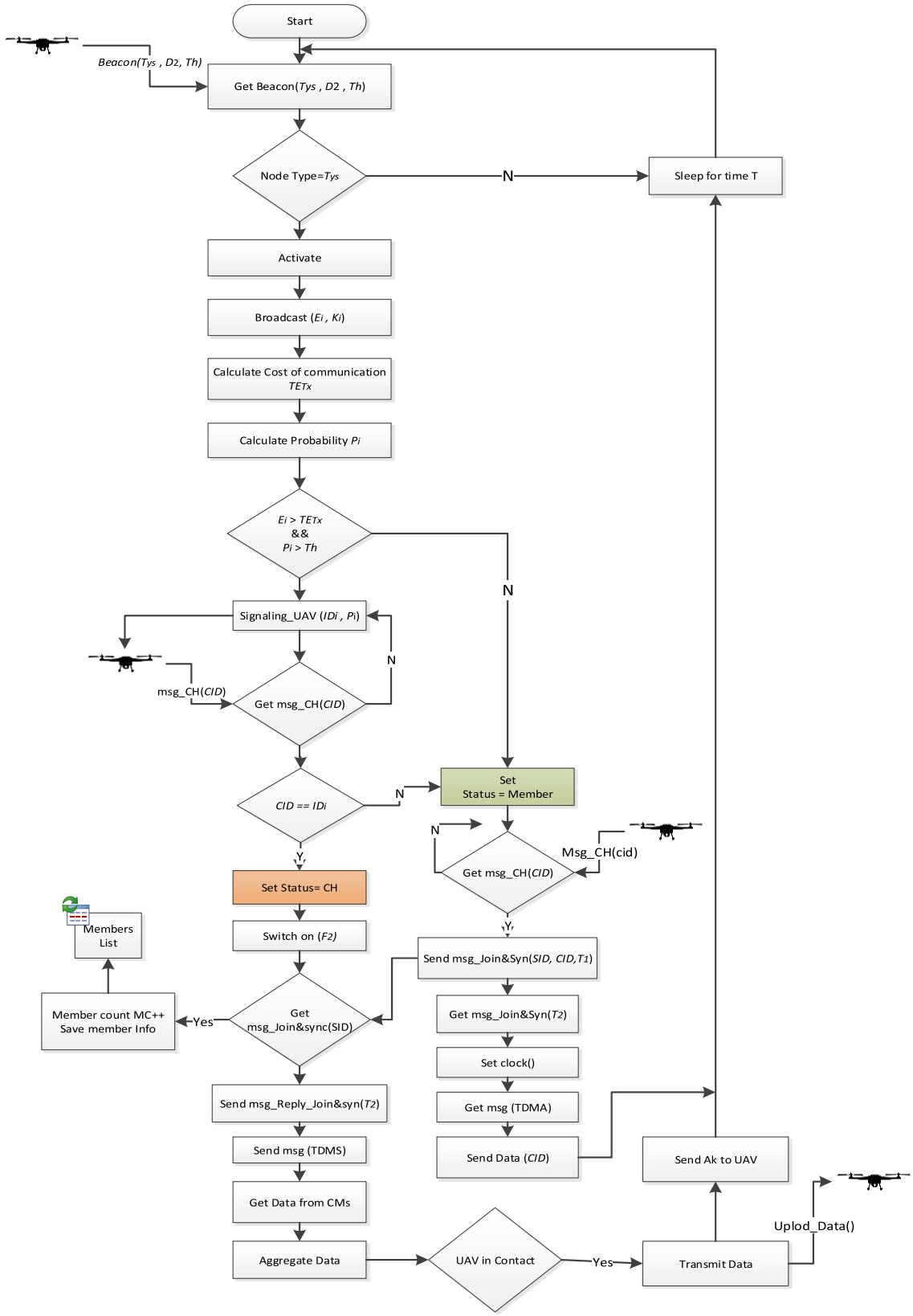

FIGURE 3. A flow chart representing the sensor node functionality.

TABLE 1. Variable list for the state diagram.

Var.	Meaning	Var.	Meaning
CCH	Candidate cluster heads.	$\mathcal{L}_i$	Location of node i.
CID	CH ID broadcast by the UAV.	m	Number of antenna elements.
D_1	Distance for localization must satisfy the far field condition of the frequency used.	n	Number of Sensor nodes $(S1, S2, \ldots, Sn)$.
D_2	Data collection height.	$P\mathcal{L}_i$	Probability of node i by knowing its location.
D_{2L}	Minimum possible D_2.	q	CCH count.
D_{2H}	Maximum possible D_2.	$T\mathscr{b}$	Total data as per equation (9).
E_i	Energy of node i.	Th	Threshold value.
TE_{Tx}	Total energy as per equation (10).	F_1	1st frequency, 433 MHz
T_{ys}	Type of sensor node targeted by the UAV.	F_2	2nd Frequency, 4.2 GHz.
P_i	Probability of node i.	g	GPS accuracy.
ID_i	i^{th} sensor node ID.	$\mathscr{b}_i$	i^{th} node data.
bn	Beaconing period	T	Sleep Time

- Operate at different frequency levels, supporting at least two frequencies, 2.1 GHz and 433 MHz, for data transmission and localisation. The default operative frequency is 433 MHz to listen for the UAV beacon as well as for synchronisation and localisation. The 2.1 GHz transducer is activated on demand for high bandwidth data transmission only.
- All nodes can listen to the UAV when in range but may not reply because of internal parameters constraints.
- Maximum transmission range is 500 m.

A unique ID is assigned to each sensor node containing the crop circle number in the field, type, subtype, and purpose of the sensor node as well as a unique number to identify the same types of sensors. The length of the node ID is 6 bytes, and the UAV selects any type or combination in its beacon message. The format of this ID is as defined as:

Circle No	Type	Sub Type	Purpose	Unique No	Total size
1byte	1byte	1byte	1byte	2 byte	6 byte
Prefix					

There are no predefined or special CH nodes as, initially, all sensor nodes are undisguisable and obtain roles (e.g., members, candidates or CH) based on the context and its health.

All sensor nodes continue monitoring five parameter values about their health, including residual energy, power consumption rate, availability of renewable energy, and data volume.

Activated sensor nodes become the CH as per probability value based on the nodes health indicators along with adding UAV factors, such as its path and node location.

The sensor node sends 9 bytes reply message to the UAV including node ID in the packet header, CH probability in the payload, and Cyclic Redundancy Check (CRC) trailer. The node reply is defined as:

ID	Probability	Trailer
6 byte	2 byte	1 byte

The working flow of the sensor node is outlined in Figure 3.

III. PROPOSED MULTI-LAYERED AND MULTI-PHASED ARCHITECTURE FOR ADAM

This ADAM system is developed with a layered architecture to simplify the design. Each layer is responsible for handling a system component (e.g., a member node, cluster head or the UAV). The first, bottom-most layer is the sensing layer consisting of the installed sensors and deals with the heterogeneity of the sensor nodes, sensing and logging activities, listening to the UAV activation call, CH selection, and sending data to the CH. The identified CH will form a middle layer to bridge the lower layer sensors with the UAV and transmit aggregated data to the UAV. The top-most layer performs as an umbrella to manage all UAV-related tasks, including path planning, navigation, and speed and height management.

The layers are further subdivided into three phases of discovery, navigation, and communication, as shown in Figure 4. The assumptions for this system include the following:

1. All sensor nodes are indistinguishable with no predefined CH. The advantage of this assumption is that data will be collected only from active and available sensors. A node that is unavailable for any reason (e.g., covered by sand, water, plant follicle or drained battery) does not affect the data gathering process.
2. No predefined data routes and clusters are considered so that periodic updates for route discovery and maintenance are not required, which saves energy of the sensor nodes.
3. The UAV and sensor nodes are equipped with dual frequency support, UHF (433 MHz) for localisation and Wi-Fi (4.2 GHz) for data collection, to optimise the energy utilisation.
4. Independent UAV paths enable the UAV to move freely while scanning a target area and instruct the selected ground sensors to activate and aggregate required data at a suitable point.

As mentioned, each layer is further divided into the three phases the detail about each one is provided in this subsection.

A. DISCOVERY PHASE

The UAV initiates the process with the discovery phase composed of four tasks of beaconing, shunting, localising, and nominating the CH, as shown in Figure 4.

Assume the UAV operates at a height of D_1 and sends a beacon $B(T_{ys}, D_2, Th)$ to initiate the data gathering process where T_{ys} is the type of the targeted nodes needing activation, D_2 is the suggested data collection height,

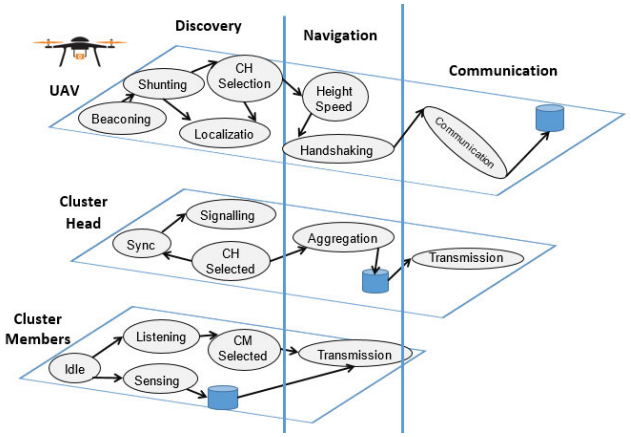

FIGURE 4. The layered architecture for the proposed ADAM system.

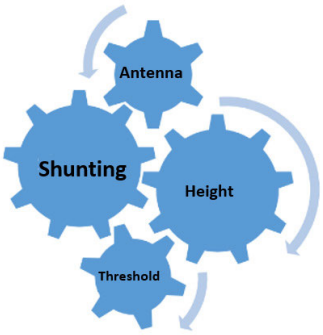

FIGURE 5. The adjustable three parameters of shunting process.

and Th is the threshold value to regulate the participating nodes competing for the cluster head. If there exist n nodes $S = \{s_1, s_2, \ldots, s_n$ deployed in a targeted area and $\aleph$ nodes are activated in response of the beacon message, such that $1 \leq \aleph \leq n$, then the set of nodes $SS = \{SS \subseteq S | SS \in T_{ys}\} = \{ss_1, ss_2, \ldots, ss_\aleph$ represents a cluster.

If $\mathscr{b}_i$ bits are required to transmit by each member node ss_i, then $T\mathscr{b}$ is the total data expected to be aggregated and transferred by the cluster SS through the CH as

$$T\mathscr{b} = T\mathscr{b}(SS) = \sum_{i=0}^{\aleph} \mathscr{b}_i \tag{9}$$

The energy consumed by the CH to send $T\mathscr{b}$ bits to the UAV over a distance D_2 is

$$TE_{Tx} = (E_{elec} T\mathscr{b}) + \left(E_{amp} T\mathscr{b} (D_2)^2\right) \tag{10}$$

The set of ground sensors SS will begin in the setup phase and remain in this phase as long as the UAV remains in the discovery phase, as shown in Figure 4. Each member of the set SS computes a probability P_i by knowing its attribute set $\mathcal{A} = (a_1, a_2, \ldots, a_z)$. Based on this probability P_i and the energy E_i of the node ss_i, a subset of SS, called the Candidate Cluster Heads, $CCH \subseteq SS \Rightarrow CCH = \{cch_1, cch_2, \ldots, cch_q\}$, is formed where $1 \leq q \leq O$.

$$CCH = \left\{CCH \subseteq SS | \forall E_{i_{current}} T\mathscr{b} \& P_i > Th\right\} \tag{11}$$

All CCH nodes send narrow band signals as a beacon to reply to the UAV. At this stage, if the UAV realizes that $q = 0$ or $q \geq m$, where q is the number of replying nodes and m is the on board antenna element, then it initiates the shunting process to limit the number of CCH to a reasonable range of $0 < q < m$. A special virtual antenna is proposed in this study to be used by the UAV to locate the CCH nodes. Based on the P_i values and CCH locations, the final CH is nominated by the UAV and broadcasted as an information message to all members in the set SS.

Shunting is important because the performance of the entire system depends on the accuracy of the shunting process. It defines the CH selection criteria. Very loose criteria allowing many nodes to be contesting for CH, but very tight criteria results in no CH. In both cases, the system may degrade badly or fails. The process handles the situation when $q = 0$ *or* $q \geq m$ such that if $q = 0$, then $CCH = \emptyset$ and no CCH member has the ability to transmit cluster data to the UAV. In this scenario, shunting instructs the UAV to reduce D_2 iterations as it can go up to a minimum possible height (i.e. a safe flight distance). On the other hand, if $q \geq m$, then there are many members of the set CCH contacting the UAV, but it does not have the capacity to locate all simultaneously. Here, shunting performs three steps:

1. Upgrade the antenna capabilities from m to $2m$.
2. Increase the height D_2 iteratively up to the communication range of F_2.
3. Increase the value of Th.

In short, shunting is a trade-off between data collection height antenna adjustment and threshold value as shown in Figure 5.

Once the CH is selected, then the discovery phase is complete, and the next phase of navigation begins.

B. NAVIGATION PHASE

When the UAV begins the navigation phase, the CMs and CH switch to communication and aggregation mode respectively. The role of each sensor node (CH, CM) was assigned in the previous phase. All CM nodes change their transmitting frequencies from F_1 to F_2, and only the CH operates at both frequencies as it uses F_1 for UAV synchronization and navigation and F_2 to collect data from the CMs. The CH aggregates all neighbouring data and waits to establish a link with the UAV. When the UAV approaches and attains an agreed height D_2, then it begins a handshake procedure with CH and switches on its F_2. This phase concludes with the establishment of a point-to-point high bandwidth backbone link between the CH and UAV.

C. COMMUNICATION PHASE

The final phase is communication, and only the top two layers (CH and UAV) participate. The CH transmits the aggregated data to the UAV through the high bandwidth link of frequency F_2. After successful transmission, the CH will enter a sleep

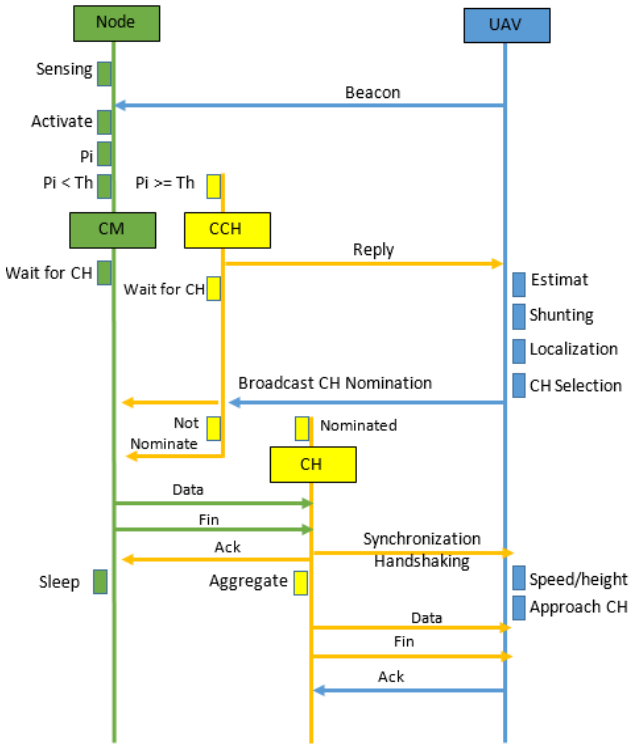

FIGURE 6. The state diagram of the ADAM system.

mode for a specific time T and phase over along with the system lifecycle.

The algorithm of the proposed ADAM architecture is provided in next section.

D. PROPOSED SYSTEM ALGORITHM

The functionality of the ADAM system is elaborated by a homological sequence of steps represented by the state diagram in Figure 6. Each vertical colour code represents a layer as green, yellow, and blue, which represent the bottom, middle, and top layers, respectively.

UAV sends beacon message to the sensor nodes to start the process. Sensor nodes activate and calculate their probabilities and decide either they are candidate of cluster head or not and decide their role (CM or CCH) accordingly. CCHs starts negotiation with UAV with at a very energy economical frequency F_1. UAV decides the final CH and broadcasts a message to all to let them know. All the CCHs change their role again to CM accept one which has chosen as a CH. All the CMs send their data to CH, meanwhile UAV navigates to CH and adjusts its height and speed. Finally, CH will transmit all the aggregated data to UAV. All participated nodes will turned to sleep mode for a specific time while UAV will start making the next cluster.

E. LINK BUDGET

Energy is the critical parameter in WSN. The relationship between the frequency, distance travelled, and energy consumed is expressed as [14].

$$Pathloss = 32.45\ dB + 20\ Log_{10}\ (frequency\ in\ MHz) + 20\ Log_{10}\ (distance\ in\ Km) \quad (12)$$

All system components are equipped with a dual frequency module. The lower frequency is used for large distances and low bandwidth communication is required for identification, navigation, and synchronisation. The higher frequency is used for short range and high bandwidth data exchange. In this system, two types of communication exist between node-to-node and node-to-UAV. For node-to-node communication, energy is conserved by making clusters and limiting the exchange of control packets. While in node-to-UAV, energy is preserved by managing the height of the UAV since high bandwidth transmission is more energy demanding and escalates exponentially with increasing distance. The link budget for both frequencies is described in the following.

1) F_1 UHF 433 MHZ FOR WIDE RANGE COMMUNICATION

If a small radiated power of $100\mu W$ is taken for the transmission, then

$$PW = 100\ \mu W = 10\ log_{10}\ (PW\ mW)\ dBm = 10\ log_{10}\ (0.1)\ dBm = -10dBm \quad (13)$$

In the case of UHF, long-distance communication in a green vegetative area, the masking factor is also very small at about MF =5 dB. If the receiver sensitivity of the UAV is $\xi = -130\ dBm$ and the demodulation signal-to-noise ratio is $SNR = 15\ dB$, then the maximal path loss constraint is calculated as

$$M_{PL} = PW - \xi - SNR - MF = -10 + 130 - 15 - 5 = 100\ dB$$

The maximum distance covered by the UHF communication is based on (13) and calculated as

$$20log_{10}\ (D_1) = 100 - 32.45 - 20\log\ (F_1) = 100 - 32.45 - 52.7 = 14.85D_1 = 10^{0.7425} \cong 5.527\ Km$$

$D_1 = 5.527\ km$ Is the maximum communication range between the CH and UAV for this frequency. In the proposed system, if the UAV operates at a height of 300 m and it activates nodes in a 500 m^2 area, then only 0.825 μW of power is required to transmit the signal at F_1 as well as to conduct the initial UAV-to-sensor node communication for localization, synchronization, and handshaking.

2) F_2 WI-FI 4.2 GHZ FOR SHORT RANGE, HIGH BANDWIDTH BACK BONE COMMUNICATION

Similarly, the maximum distance covered by the high bandwidth Wi-Fi signal is calculated by considering the same parameters as above except for a higher masking factor of

$MF = 20\ dB$. If $\xi = -130\ dBm$ is the receiver sensitivity and $SNR = 15dB$, then the maximum path loss constraint is

$$M_{PL} = PW - \xi - SNR - MF = 85\ dB$$

And the maximum distance covered by the Wi-Fi signal is

$$\begin{aligned} 20\ log_{10}(D_2) &= 85\ dB - 32\ dB - 20\ log_{10}(F_2) \\ &= (85 - 32.45 - 67.6)\ dB = -15.05dB \\ Log_{10}(D_2) &= -\frac{15.05}{20} = -0.7525 \Longrightarrow D_2 \\ &= 10^{-0.7525} \cong 0.176\ km = 176\ m \end{aligned}$$

IV. LOCALIZATION USING VIRTUAL ANTENNA ARRAY

Data collection from remote agricultural areas, as in our case study, includes many heterogeneous sensor nodes dispersed across a large geographical area. Efficient localisation of ground sensors is vital for this data gathering scheme and can follow three options. First, each sensor nodes can be furnished with a GPS module to inform the UAV with its position as in [15]. For this configuration, the end nodes require more processing and energy, but small sensors such as leaf or stem sensors could not accommodate this capability. Second, the UAV may include a localisation array, but such a heavy, large, and energy-demanding antenna system mounted on a small UAV limits what it can carry for its payload and on-board battery. Third, a special Virtual Phase Array (VPA) antenna system, we developed to handle these challenges, replaces the physical ULA antenna with a single antenna [16]. The replacement VPA is so light that the UAV can carry it without disturbing its flight specifications.

Our proposed VPA antenna overcomes the following existing limitations of the classic systems:

1. Size and weight: Conventionally, the location of a signal-emitting node is estimated using an array of multiple antennas. Installing these onto a small UAV increases its weight and results in decreased flight time, speed, payload capacity, and agility. In our proposal, the consolidated VPA replaces the physical antenna.
2. Capabilities: A physical antenna cannot change its steering capability while, in a virtual antenna, the steering precision level is a trade-off between speed and accuracy.
3. Throughput: Compared to a fixed infrastructure with very low throughput and a high round trip time (RTT), the UAV with a VPA reduces these issues by managing the localization precision level, speed, and height of the UAV.
4. The total detectable targets are limited to the number of antenna elements, and this quantity is virtually adjustable with the VPA.
5. Multiple frequencies are not easy to implement in a physical on-board antenna as inter-element spacing is not easily modified. However, this is possible with the VPA.
6. Typical challenges with a physical antenna, such as current gains, mutual coupling, and inter-channel phases between array elements, do not exist in our proposed VPA.

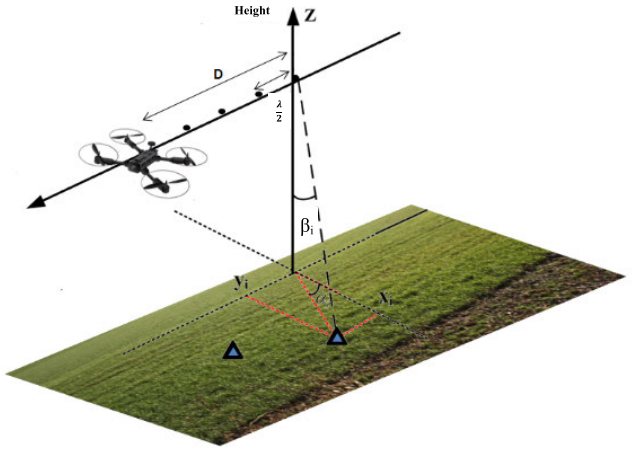

FIGURE 7. A conceptual model for the VPA antenna.

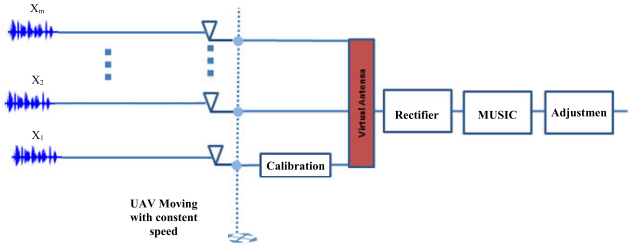

FIGURE 8. The component design of our virtual phase array antenna.

It is also assumed that all ground sensor nodes emit a narrow band signal that is monitored by a single antenna mounted over the UAV. This solitary antenna acts as a linear array of M antennas by exploiting the motion of the UAV. As the UAV acquires snapshots of the data after a fixed time and distance intervals, M snapshots are considered as the same outcome of a physical antenna with M elements, as shown in Figure 7.

A. VIRTUAL ANTENNA ARRAY

The UAV carrying the VPA, referred to as the Synthetic Aperture UAV (SA-UAV), summarises the received data in response to incoming signals of different phases at different times. The Angle of Arrival (AoA) of a signal is estimated if the channel is assumed to be quasi-stationary during data collection at least for one snapshot (less than a second), which is practical.

We assume the SA-UAV acquires snapshots at M different places to estimate the direction of N scattered target nodes emitting narrowband signals where $N = M - 1$. The snapshots are taken at a fix period of Δt. The SA-UAV collects R_{over} samples of data for each snapshot. If the SA-UAV is operating at a height satisfies the far field condition of the frequency used, then all received signals can be considered a plane wave. The VPA is more complex than a physical ULA, and special consideration is required to collect and process

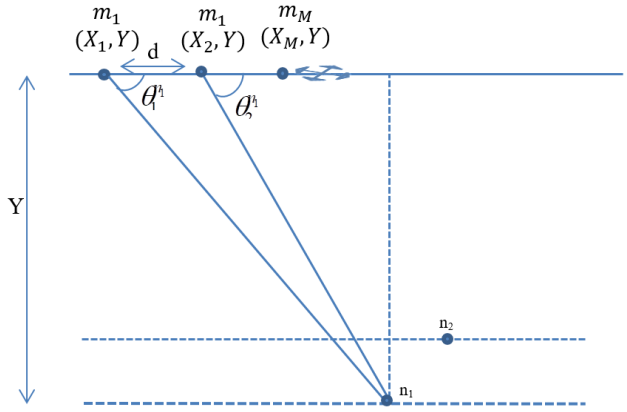

FIGURE 9. The geometric angle variation for the virtual phase array antenna.

samples due to changing characteristics of the incoming signals. To deal with these issues, new modules are incorporated into conventional localization systems (e.g., MUSIC), such as calibration, rectification, and adjustment. Our designed VPA is illustrated in Figure 8.

If $x_i(t)$ is x independent signals collected at time t at the i^{th} snapshot then factors such as signal-to-noise ratio (SNR), geometrical variation of the antenna, and desynchronization of the UAV and ground sensor clock cause an extra deviation in the phase of a received signal for each snapshot. The phase difference F_{offset} caused by the desynchronized clock between the devices for two consecutive samples is estimated by introducing a calibration module during the first snapshot. Once F_{offset} is estimated, it is considered the same for all proceeding samples. Data from all snapshots are given to the Virtual Antenna Module (VAM) for further processing. A matrix is created by storing the response of each signal resulting in a steering vector, which is handed over to the Rectifier Module (RM). The estimated F_{offset} is applied next in the RM to rectify the entire data set and to construct a covariance matrix R_{AS}. This rectified covariance matrix is input to the classical MUSIC algorithm with its output (angle θ) being corrected by the adjustment module. The details of each proposed module are described in the following sections.

B. GEOMETRICAL VARIATION

The classical MUSIC algorithm cannot be applied directly because the snapshots are not all taken simultaneously (as with a physical array antenna). The continuously changing position of the antenna causes deviation in the signal characteristics. First, the geometrical variation of the antenna must be defined. If N sensor nodes $(n_1, n_2, \ldots, n_N)$ are deployed, then a signal transmitted by node n_i makes the angle $(\theta_1^{n_1}, \theta_2^{n_1}, \ldots, \theta_M^{n_1})$ with M virtual antenna elements comprising the UAV antenna (see Figure 9). For each snapshot, the UAV motion is taken as a straight path and the X axis is considered linearly corresponding to the UAV trajectory.

As shown in Figure 9, N target nodes $(n_1, n_2..n_N)$ are installed in a field where $1 \leq i \leq (N = M - 1)$. The ordinate of UAV, Y, is the same for all nodes while the abscissa is X_k at time t_k where $1 < k \leq M$.

The SA-UAV acquires M snapshots of the data after a fixed distance of

$$d = X_{k+1} - X_k = c_u \Delta t \tag{14}$$

If the UAV covers a distance d in time Δt with a speed c_u while taking two successive snapshots, then the angle of the first antenna element X_1 with the sensor node n_i is determined by

$$\tan \theta_1^{n_1} = \frac{X_1}{Y} \tag{15}$$

If the first snapshot is taken as a reference, then the angle for the second snapshot is calculated as

$$\tan \theta_2^{n_1} = \frac{X_1 - d}{Y} \tag{16}$$

Similarly, for the linear flight of UAV, we write

$$\theta_k^{n_i} = \theta^{n_i}(t_k) = \tan^{-1}\left(\frac{X_k - C_u t_k}{Y}\right) \tag{17}$$

The SA-UAV takes in the received signals after an interval of $\Delta t = \frac{\lambda}{2c_u}$ seconds (where $d = \lambda/2$ is the antenna element separation). Then, it constructs a $M \times 1$ steering vector, or an array manifold, for the node i of angle θ. We construct the observation matrix as

$$A(\theta) = [a_1(t_1), a_2(t_2), \ldots a_M(t_N)]^T \tag{18}$$

If t_i and the observations $A(\theta)$ are known, then the VPA can model the physical array of the antenna. As the data sampling frequency is higher and the frequency of snapshotting is related to the speed of the UAV, it can be assumed that the UAV is operating at a constant speed of c_u for one antenna length making a uniform VPA antenna of M elements. The array's first element is taken as a reference, and the steering vector model is written as

$$\begin{aligned} A &= \left[e^{j\beta d_i \sin(\theta_i)}\right] = \left[e^{j\beta c_u t_i \sin(\theta_i)}\right] \\ &= \left[1, e^{j\beta(c_u \times t_1)\sin(\theta)} \ldots e^{j\beta(c_u \times t_N)\sin(\theta + \Delta\theta_M)}\right] \end{aligned} \tag{19}$$

where, $i = 1, 2, \ldots, M$, the inter-snapshot spacing is $d = \lambda/2$, $\beta = \frac{2\pi}{\lambda}$, λ is the wavelength, and θ is the AoA of the sensor. The UAV begins acquiring snapshots at time $t_1 = 0$ and it approaches the M^{th} snapshot at time t_m. The change of the incident angle of the signal is $\Delta\theta$ for snapshot N due to movement of the UAV.

A key benefit of our proposed system is that it can operate at different frequencies and can adapt inter-element spacing and multiple samplings as required.

C. SAMPLING

If M snapshots are required, then R_{over} samples are needed between every two snapshots m_k and m_{k+1}. The R_{over} also depends on the sampling frequency of the Analog to Digital Converter (ADC) used. If F_{ADC} (in Hz) is the measured ADC sampling frequency, then oversampling for one virtual array element is

$$R_{over} = \lfloor F_{ADC}\,\Delta t \rfloor \quad where\ \lfloor\ \rfloor\ stand\ for\ integer\ part \tag{20}$$

If a standard ADC converter (e.g., Voltage Controlled Crystal Oscillator) has an accuracy of 10 ppm $<$ Frequencyaccuracy $<$ 50 ppm (ppm = part per million).

And if $F_{radio} = 433\ MHz$ is the frequency considered for communication, then it generates a frequency offset of

$$4.33\ KHz < \Delta F < 21.65\ KHz$$

By sampling the period $\frac{1}{F_{ADC}}$, the frequency offset generates a rotation of phase according to $\Delta\psi_{offset} = \frac{\Delta F}{F_{ADC}}$

$$0.02165 \times 2\pi < \left|\Delta\psi_{offset}\right| < 0.10825 \times 2\pi$$

Therefore, the poorest accuracy results in a $\frac{2\pi}{10}$ phase rotation between two consecutive samples. This means that, following the above conditions, every consecutive data sample may have a $\frac{2\pi}{10}$ phase difference instead of what is expected. In contrast to the physical antenna, the VPA samples are independent (taken one by one), and each sample receives an additional $\Delta\psi_{offset}$, which must be corrected.

D. CALIBRATION

Due to desynchronization of the clocks in the sensor nodes and the UAV, the precedent ΔF applies to each clock and the resultant inaccuracy (relative to a theoretical) clock frequency is the sum of both, which may be in opposite directions. The ADC oversampling causes an additional phase difference between two successive data samples. Since the phase rotation between two consecutive samples is $< \frac{2\pi}{10}$, the calibration can estimate this rotation for compensation before the array processing procedure.

To fix our idea without a loss of generality, suppose the scenario with a singular source, and let $s(k)$ denote the kth sample of that source received by the antenna at k discrete time.

$$s\,(k) = e^{-j2\pi\,\Delta F_{total}\,\frac{k}{F_{ADC}}} \tag{21}$$

where, $k = 1, 2, \ldots, R_{over}$. ΔF_{total} is the frequency offset difference of the two clocks, UAV, and the node expressed as $\Delta F_{total} = F_{UAV} - F_{node}$ the angle of rotation between two samples due to ΔF_{total} is

$$\begin{aligned}\Delta\Phi &= \Phi_k - \Phi_{k-1} \\ &= \frac{2\pi\,\Delta F_{total}\,k}{F_{ADC}} - \frac{2\pi\,\Delta F_{total}\,(k-1)}{F_{ADC}}\end{aligned} \tag{22}$$

From (20) and (22), the frequency offset difference ΔF_{total} can be written as

$$\Delta F_{total} \approx \frac{\Delta\Phi R_{over}}{2\pi\,\Delta t} \tag{23}$$

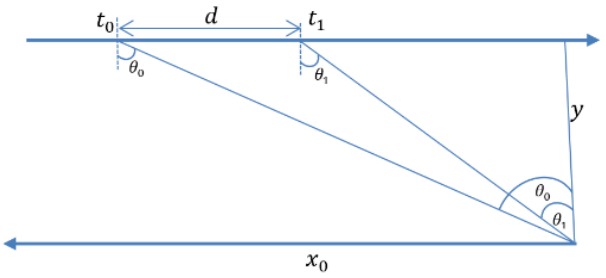

FIGURE 10. The angle adjustment for the virtual phase array antenna.

ΔF_{total} may be considered constant during an observation time. For a unitary observation time of Δt (between two samples of the VPA), we approximate as

$$\Delta F_{total} \approx \widetilde{\Delta F}_{\text{offset}} \cong \frac{\sum_{k=2}^{R_{over}} (\Phi_k - \Phi_{k-1})}{R_{over} - 1} \tag{24}$$

E. RECTIFICATION

After the estimation of $\widetilde{\Delta F}_{\text{offset}}$, a rectification is applied to compensate for the phase offset as

$$S_r\,(k) = e^{-j2\pi\,\widetilde{\Delta F}_{offset}\,\frac{k}{F_{ADC}}}\,s\,(k) \tag{25}$$

F. ADJUSTMENT OF THE INCIDENT ANGLE

In our scenario, the incident θ changes for each sample by the ratio of the UAV covered distance and its speed. The final output θ value must be adjusted as following Figure 10.

As in Figure 10 let θ_i be the incident angle at the ith snapshot, x_i is the horizontal distance between the UAV and node at the ith snapshot, and y is the height of UAV. In this case, we conclude that $x_i = x_{i+1} + d$ and $\theta_i = tan^{-1}\left(\frac{x_i}{y}\right)$. Where, d is the distance between two successive snapshots as $D = c_u\Delta t$. Here, c_u is the speed of the UAV and Δt is the time between two snapshots. Using a first-order approximation, the adjustment is initially written as

$$\begin{aligned}\Delta\theta &= \theta_i - \theta_{i+1} = \tan^{-1}\left(\frac{x_i}{y}\right) - \tan^{-1}\left(\frac{x_i+1}{y}\right) \\ &\approx \frac{x_i}{y} - \frac{x_i - d}{y} = \frac{c_u\Delta t}{y}\end{aligned} \tag{26}$$

Therefore, the total adjustment $\Delta\theta$ should be

$$\Delta\theta_T = M\,\Delta\theta(27) \tag{27}$$

V. PERFORMANCE ANALYSIS ON THE CLUSTERING TECHNIQUE

To evaluate the performance of our designed system, simulation models are developed using three different software platforms:

1. OMNeT++ (Objective Modular Network Testbed in C++ [17], [18]) is shown in Figure 11 with video results available online [19].
   In this simulation 2D analysis of cluster, formation is conducted. Blue color doted and the diagonal line

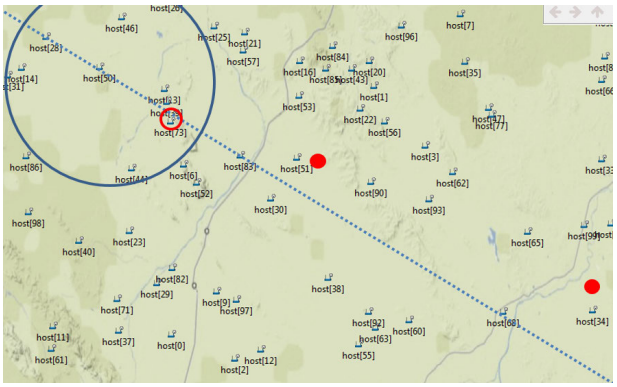

FIGURE 11. A snapshot of the OMNeT++ simulation.

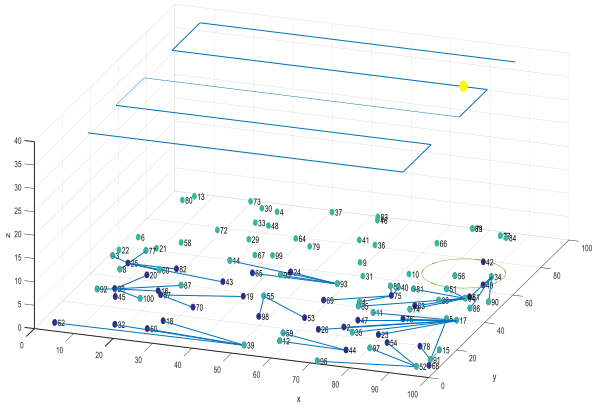

FIGURE 12. The MATLAB simulation snapshot.

represents the path of UAV while blue circle shows its communication range. The small red colored circles are the node which was selected as CH in the previous iteration.

2. The snapshot of our MATLAB simulation is provided in Figure 12 with a complete video available online [20]. The 3D analysis is conducted in this simulation. A 100 × 100 meter plane bed is taken for sensor nodes deployment. A small yellow dot representing the UAV, flying at the height of 40 meters and at the same height blue zigzag lines represent its path. Connected nodes with blue lines at the ground level represent a cluster where central node is a CH.
3. The AGI-STK is presented in Figure 13 with videos available at [21]–[23]. A real-time simulation study is conducted in AGI-ST. STK is taking real-time images from google earth with exact land geography of the area including land characteristics (height, depth, roads, trees, etc.) [24]. The big green circles are representing the original farm field located in the Tabuk region. The white color doom shapes, we created to analyse the ranges of different sensor nodes. A real-time flying drone can also be seen in the picture, operating for data collection.

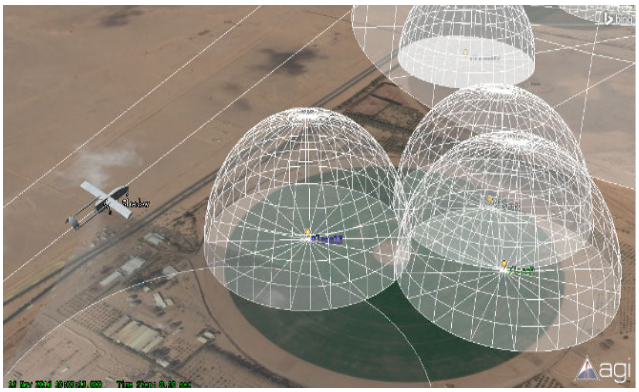

FIGURE 13. An illustration of the STK simulation.

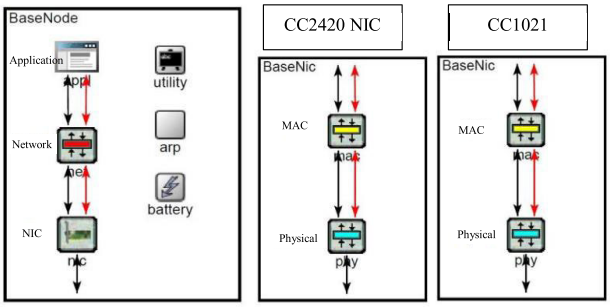

FIGURE 14. Illustration of the sensor node and installed network interface card.

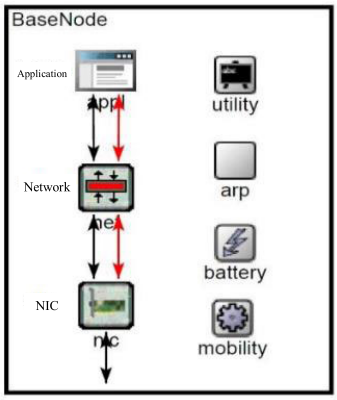

FIGURE 15. The UAV simulation model with mobility module.

As per protocol, wireless sensor nodes are developed with dual frequencies support. We developed a simulation model for the sensor node using two NICs, CC1021 [25] and CC2420 [26]. Figure 14 and Figure 15 shows the simulation model of the sensor node and UAV sink, respectively. The CC1021 is a low power RF transceiver for transmitting narrowband signals of frequency 433 MHz. On the other hand, the CC2420 operates at 2.4 GHz with compliance to the 802.15.4-enabled Network Interface Card (NIC) and has a built-in CSMA/CA base MAC layer.

The simulation models for the sensor nodes are developed in such a way that they can negotiate the communication height with the UAV by considering the remaining energy

TABLE 2. Simulation setup and list of parameters.

Simulation Setup Parameters	Values
Beacon sending period	Every 2 seconds
Battery capacity of each sensor	1 joule
Speed of UAV	20 m/h
Height of UAV during mission	400m to 10m
cc2420	802.15.4 compliant network interface card operating frequency 2.4 GHz built-in CSMA/CA
cc1021	Narrowband communication NIC of 433 MHz
Number of sensor	100
Simulation area	$2km^2$

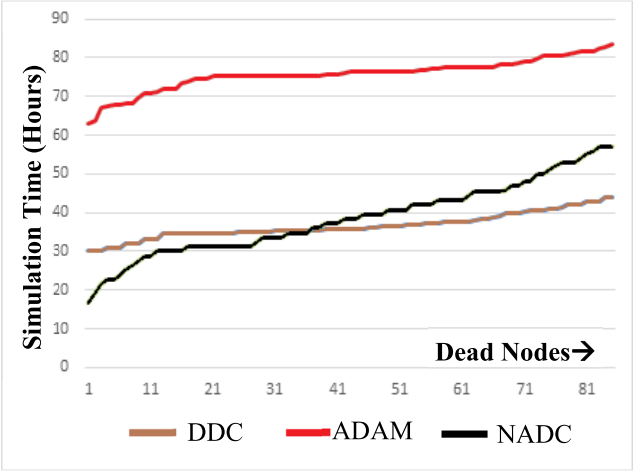

FIGURE 16. Simulation study of energy utilization in different system.

level and amount of data needed to transmit. Localisation requires only 0.825 μW by operating with the CC1021 module. The data transmission by the CH to the UAV uses a higher frequency (Wi-Fi) and causes the largest energy depletion.

The developed sink traverses the network to collect data for the static sensor nodes. In this scenario, three types of data collection are possible:

1. DDC (Direct Data Collection) by UAV [27]
2. NADC (Network Assisted Data Collection) [28]
3. ADAM (our designed UAV controlled routing protocol)

In DDC, the UAV visits each sensor node individually. For NADC, the sensor nodes make a cluster to which the UAV visits only the CH to collect the data, but the path of the UAV is defined and controlled by the ground network topology. In our proposed system, ADAM, the UAV moves freely, guides the ground sensor nodes in forming clusters, and help in the selection of the CH. To evaluate the performance of the system, we created a simulation model for each type with different simulation runs to study the variation of the parameters. Initial simulation setup and different parameters used are listed in table 2.

Some of the simulation results are explained below.

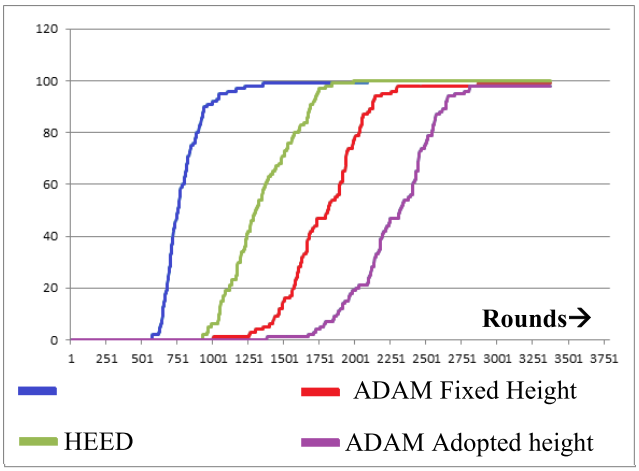

FIGURE 17. The pattern of dead nodes in each round in our proposed system vs LEACH and HEED.

A. NUMBER OF DEAD NODES VS SIMULATION TIME

In this simulation, we test the energy utilisation of nodes and estimate how quickly the sensor nodes become exhausted completely. We simulate each case (NADC, DDC, and ADAM) separately against the same parameters as shown in Figure 16.

Our design can enhance the network lifetime by 20% as well as improve the cost of adding more intelligence in sensor nodes, dual frequency support, and the virtual antenna in UAV.

B. DEAD NODES INVESTIGATION IN MATLAB

The results of our simulations on dead nodes in each UAV round with respect to different routing schemes are shown in Figure 17. Each scheme is represented as a different color curve with the higher performance being observed in the lower curve. The red and purple lines show the performance of our proposed ADAM protocol.

We test the ADAM system against two cases with a fixed height and an adoptive height. In the first case, although the UAV flies at a constant height (the UAV is not reducing the height to collect data) even then the proposed system performs well due to duel frequency use, no flooding of information, no periodic updates, and improved CH selection. The red curve shows the results of this scenario. In the second case, the performance of the proposed system is tested with the adaptive height scenario. The results denoted as the purple curve. In this case, before transmitting the data, a suitable height is negotiated by the UAV and CH. Energy depletion accrues exponentially if the distance increases during high bandwidth communication. Adjusting the height positively impacts the overall system lifetime.

C. INTER-CLUSTER COMMUNICATION ASSESSMENT

Figure 18 reflects the amount of overhead required to build a network with different routing schemes. All the control packets are exchanged to make clusters, choose the CH, and build a network. The different color lines represent protocols

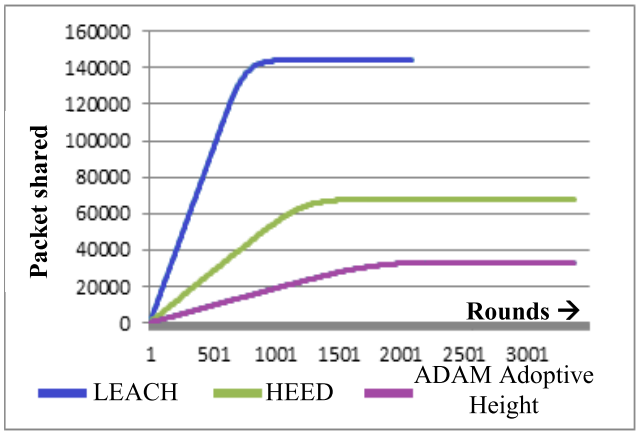

FIGURE 18. The inter-cluster packet exchange vs the number of rounds.

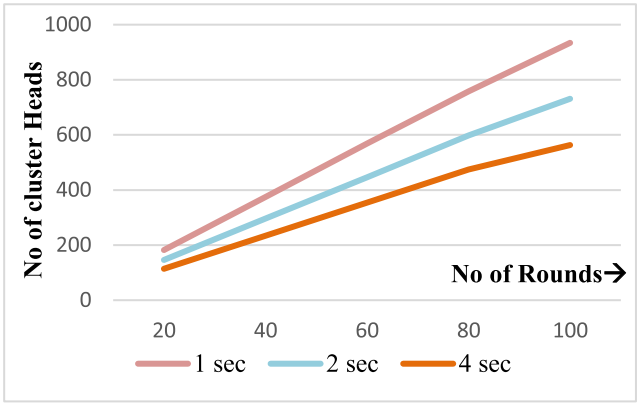

FIGURE 19. The number of clusters vs the rounds with varying beaconing periods.

and the lowest is the best in the system. In a conventional system, 140k control packets (i.e., periodic updates and CH notifications) are exchanged throughout the simulation run of 1000 rounds among the sensor nodes, which only keep the network alive and updated. In the proposed system, these control messages are reduced to 20k, which indicates clear improved performance.

D. EFFECT OF VARYING THE BEACONING PERIOD ON NODE ENERGY UTILIZATION

By prolonging the beaconing period, larger and fewer clusters are formed, as shown in Figure 19. However, by increasing the number of clusters, there is a chance to leave some nodes unattended because they may not be assigned to any cluster. As shown in Figure 20, many nodes are left unattended (the higher peaks) at the end of the simulation but are not the case in Figure 21.

VI. PERFORMANCE ANALYSIS ON THE LOCALIZATION SYSTEM

The proposed localisation system reviewed by simulation in MATLAB. A 433 MHz single antenna is mounted over the UAV that moves with a constant speed of 20 m/s. Two target nodes emitting narrowband signals $a(\theta)$ are placed at ground

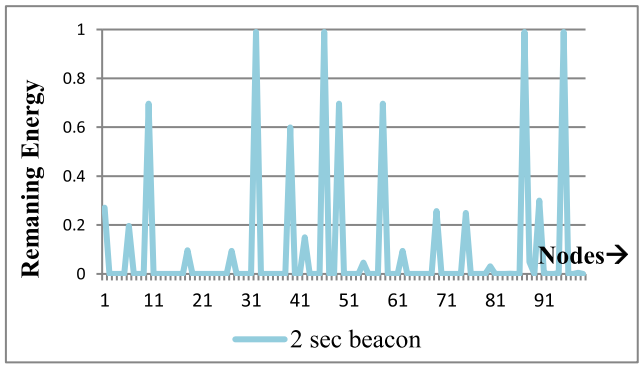

FIGURE 20. The energy utilisation of each node in 2-second beacon periods.

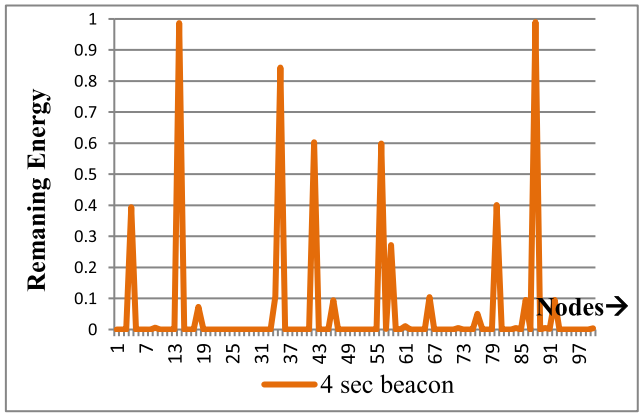

FIGURE 21. The energy utilisation of each node in 4-seconds beacons.

TABLE 3. List of variables used in localization simulation.

Variable	Value	Discerption
F	4.3 x 10^8	Hz radio frequency for transmission
F_{ADC}	2 x 10^5	Hz analogy to digital converter frequency
F_{offset}	2 x 10^4	Hz frequency offset between transmitter and receiver
V	20	m/s Speed of UAV
dt	λ /(2*v)	Time difference between 2 snapshots (2 virtual antennas)
R_{over}	F_{ADC}*dt	Oversampling of the virtual array= number of samples %between 2 snapshots of the virtual array
M	20	Number of snapshots for 6 degree accuracy
y	100	Height of UAV
dt_{over}	1/fADC	Time between 2 ADC samples.

level with the azimuth angles of 20^0 and -60^0. The elevation angle is taken to be 0 for all nodes. The SA-UAV locates the target nodes. The parameters used in simulation are listed in table 3.

As shown in Figure 22, the AoA of these two target nodes are measured accurately (20^0 and -60^0) by the UAV using our proposed system. The directions of both targets are measured with and without our proposed system. The black curve shows the estimation by MUSIC without our modification,

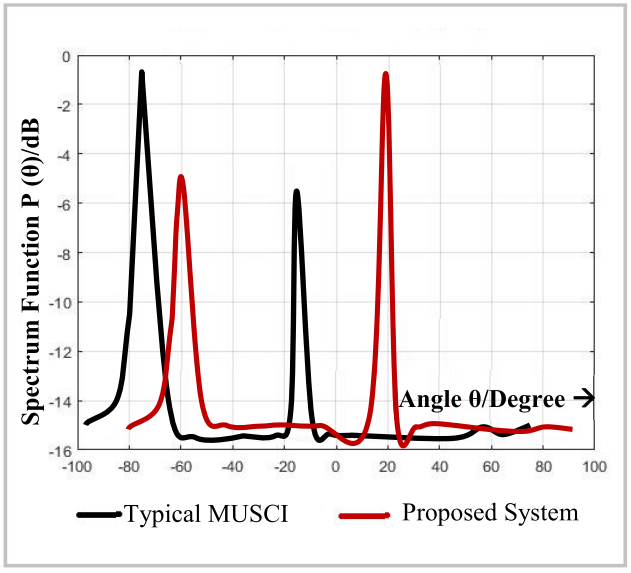

FIGURE 22. Our proposed system vs classical MUSIC.

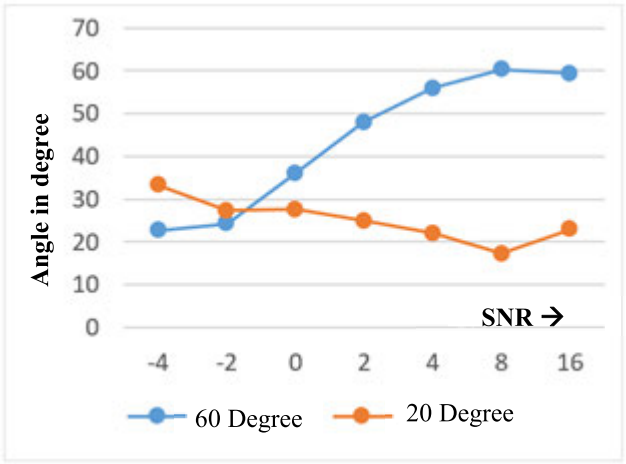

FIGURE 23. The effect of increasing the SNR on the proposed AoA estimation.

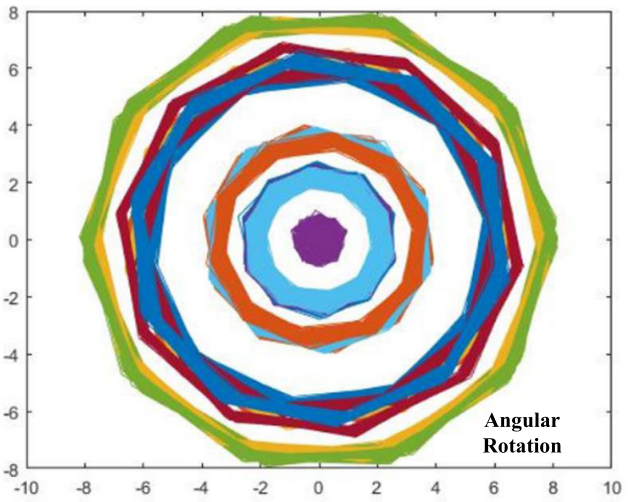

FIGURE 24. The received signal without rectification.

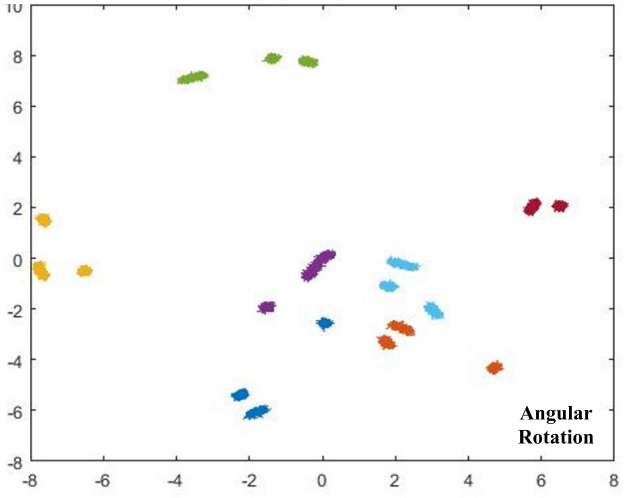

FIGURE 25. The received signal after rectification.

which does not work well. The red curve represents a successful MUSIC process after incorporating our developed modules, including the calibration and rectification. So, estimation provides accurate results.

The system performance may degrade with an increasing SNR, and its relation with the angle detection accuracy is shown in Figure 23. The proposed system works well with a low value of SNR up to 8 dB.

Figure 24 shows the original received signal without a rectification, and each signal is observed to receive 10 rotations per sample while causing significant deviations in the AoA estimation. A special treatment called rectification is required on the received signal to deal with this issue.

The received signals after correcting with our proposed rectification are shown in Figure 25 where the rotation of phases due to ADC oversampling is treated with the proposed rectification module and fixes them successfully.

All simulation results provided above demonstrate that the VPA can model the physical antenna successfully without degradation of results. The designed system makes it possible to develop a SA-UAV where a single antenna acts as a virtual phase array of M antennas to provide the benefits of ease of use, cost-effectiveness, light-weightiness, energy efficiency, flexibility, and adaptability.

VII. PROOF OF CONCEPT

An Arduino microcontroller is used to build our concept devices because it offers many advantages and conveniences including:

1. An open-source microcontroller that is easy to program and build.
2. Includes many compatible modules and applications.
3. Offers multiple sizes (e.g., Uno, Micro control, and Mini) with the same interface, so that if a component is built on an Uno, then it is easy to change it to a Micro.

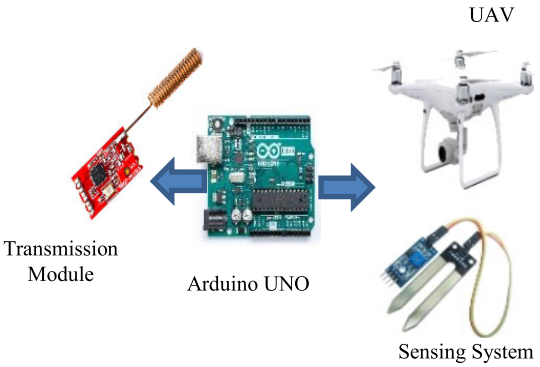

FIGURE 26. The proof of concept devices for our proposed system.

FIGURE 27. Dr. Ammad Uddin is deploying the prototype sensor node in the crop field.

FIGURE 28. Our UAV flying during a sensing mission.

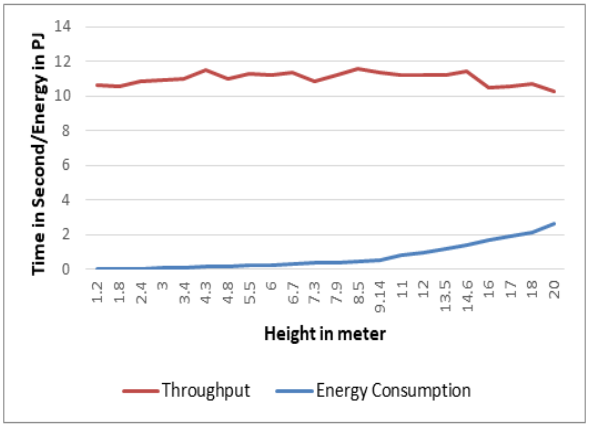

FIGURE 29. The UAV data collection from varying heights.

4. Provides different shields (e.g., GSM, Motor, MP3, Zigbee, and a joystick) that are readily available to design a component in layers.

The UAV sink node and ground sensor nodes are two components of this system made by the Arduino microcontroller. A customised UAV is also built to carry this developed sink node that operates by following our proposed algorithm. The UAV sink and sensor nodes are shown in Figure 26. Both components are developed using Arduino Uno and an nRF24L01 pair used for processing and communication, respectively. The Arduino Uno connects with a sensing device (e.g., temperature sensor, humidity sensor, leaf sensor or soil sensor) to make the ground sensor node (see Figure 27), and, if connected with the UAV, will act as a sink. We develop two software separately for the sensor node and the sink node, which were installed in Arduino's ROM.

The sensor-sink prototype was tested near the Tabuk region's ASTRA farms where alfalfa crops are cultivated. Figure 27 was taken when we were deploying the sensor nodes in crop field and Figure 28 was taken while the UAV was on a mission to collect data from the crop filed, and the complete video of the field experiment is available online [29].

The nRF24L01 is a transceiver operating at 2.4 GHz frequency, and it can transmit data at a rate up to 2 Mbps with very low power consumption. Initially, we installed five sensor nodes in the crop field, each loaded with 20 MB of data. As soon as the UAV approached the field and it started to process, one sensor was selected as the CH dynamically, which then aggregated all data from the cluster nodes and transferred 100MB data to the UAV using a high bandwidth channel.

The GPS coordinates of all sensors were predefined and known to the UAV because the localisation system, including the virtual antenna and UHF transceiver for the sensor nodes, remain in development. The data collected from this proof of concept experiment in the farm fields are plotted in Figure 29 with an evaluation of the effect of varying the UAV height on the system performance. There exists no noticeable effect on the data collection time by increasing the data collection height of the UAV, but it does have a significant impact on the field nodes' energy utilisation.

VIII. CONCLUSION

Our focus in this study is to harness IoT and UAVs in conventional agriculture to produce quality and quantity crops with optimum resource utilisation. The challenge was to develop

an IoT based system that can resist in harsh weather, able to work with no or limited infrastructure, integrates heterogeneous sensors to collect the required data. To cope with these constraints, we developed ADAM system, a UAV controlled routing protocol for data collection.

The formation of dynamic clusters from heterogeneous sensor nodes, a Bayesian-based cluster head selection, and a special localisation system are the key features of our design. An important advantage of this scheme is that the UAV can move independently to scan the desired area and harvest only selected IoT nodes. Another benefit of this protocol is that it preserves the maximum node energy by shifting most of the working load onto the UAV that uses very low power frequency for localisation and collection of data at a reasonable height. The UAV performs many tasks throughout the data gathering process, including beaconing, shunting of connected nodes, localisation of the cluster head, and data collection. Furthermore, it also plays an important role in cluster formation and CH selection. Localisation of the field sensors by the UAV is also a crucial design of the system with a lightweight, energy efficient virtual phase array antenna system to locate the field sensors.

We evaluated the system in a series of simulation models, and the results reflect success in preserving the maximum node energy at the expense of the UAV energy and working load. The objectives of this research include the monitoring of crop-related parameters, processing collected data from a central point, and taking appropriate action to conserve resources (e.g., water). In future, sensors, actuators, artificial intelligence, and the UAV will be connected in a closed loop system to manage the resources and eliminate opportunities for energy waste.

REFERENCES

[1] S. M. Taha, *Kingdom Imports 80% of Food Products*, Arab News, Riyadh, Saudi Arabia, 2019.
[2] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E.-H. M. Aggoune, "Internet-of-Things (IoT) based smart agriculture: Towards making the fields talk," *IEEE Access*, to be published.
[3] X. Feng, F. Yan, and X. Liu, "Study of wireless communication technologies on Internet of Things for precision agriculture," *Wireless Pers. Commun.*, pp. 1–18, May 2019.
[4] P. Tokekar, J. V. Hook, D. Mulla, and V. Isler, "Sensor planning for a symbiotic UAV and UGV system for precision agriculture," *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1498–1511, Dec. 2016.
[5] C. Fu, Z. Jiang, W. Wei, and A. Wei, "An energy balanced algorithm of LEACH Protocol in WSN," *Int. J. Comput. Sci. Issues*, vol. 10, no. 1, pp. 354–359, 2013.
[6] K. Xu and M. Gerla, "A heterogeneous routing Protocol based on a new stable clustering scheme," in *Proc. Military Commun. Conf. (MILCOM)*, vol. 2, Oct. 2010, pp. 838–843.
[7] J.-Y. Chang and J. Pei-Hao, "An efficient cluster-based power saving scheme for wireless sensor networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, no. 1, Dec. 2012, Art. no. 172.
[8] L. Shangguan, L. Mai, J. Du, W. He, and H. Liu, "Energy-efficient heterogeneous data collection in mobile wireless sensor networks," in *Proc. 20th Int. Conf. Comput. Commun. Netw. (ICCCN)*, 2011, pp. 1–6.
[9] A. W. Khan, A. H. Abdullah, M. A. Razzaque, and J. I. Bangash, "VGDRA: A virtual grid-based dynamic routes adjustment scheme for mobile sink-based wireless sensor networks," *IEEE Sensors J.*, vol. 15, no. 1, pp. 526–534, Jan. 2015.
[10] H. Okcu and M. Soyturk, "Distributed clustering approach for UAV integrated wireless sensor networks," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 15, no. 3, pp. 106–115, Jan. 2014.
[11] A. W. Khan, A. H. Abdullah, M. H. Anisi, and J. I. Bangash, "A comprehensive study of data collection schemes using mobile sinks in wireless sensor networks," *Sensors*, vol. 14, no. 2, pp. 2510–2548, Feb. 2014.
[12] B. T. Sharef, R. A. Alsaqour, and M. Ismail, "Vehicular communication ad hoc routing Protocols: A survey," *J. Netw. Comput. Appl.*, vol. 40, pp. 363–396, Apr. 2014.
[13] *Naive Bayes Classifier*, Wikimedia Found., San Francisco, CA, USA. 2019.
[14] WikiPedia. *Free-Space Path Loss*. Accessed: Sep. 21, 2017. [Online]. Available: https://en.wikipedia.org/wiki/Free-space_path_loss
[15] E. Goldoni, L. Prando, A. Vizziello, P. Savazzi, and P. Gamba, "Experimental data set analysis of RSSI-based indoor and outdoor localization in LoRa networks," *Internet Technol. Lett.*, vol. 2, no. 1, Jan./Feb. 2019, Art. no. e75.
[16] M. Ammad-Uddin, D. Le Jeune, A. Mansour, and E.-H. M. Aggoune, "Direction of arrival of narrowband signals based on virtual phased antennas," in *Proc. IEEE 23rd Asia–Pacific Conf. Commun.*, Dec. 2017, pp. 627–632.
[17] A. Varga, "The OMNeT++ discrete event simulation system," in *Proc. Eur. Simulation Multiconf. (ESM)*, Prague, Czech Republic, Jun. 2001, pp. 1–7.
[18] A. Varga, "Using the OMNeT++ discrete event simulation system in education," *IEEE Trans. Educ.*, vol. 42, no. 4, p. 11, Nov. 1999.
[19] M. Ammad-Uddin. (2017). *OMNet++ Simulation of Dynamic Clustering and Crop Health Monitoring*. [Online]. Available: https://www.youtube.com/watch?v=pYx-Z0chWoA
[20] M. Ammad-Uddin. *MatLab 3D Simulation of Dynamic Clustering*. Accessed: Dec. 3, 2017. [Online]. Available: https://www.youtube.com/watch?v=zVM42i9BQFY
[21] M. Ammad-Uddin. *STK Simulation of Dynamic Data Collection by UAV*. Accessed: Dec. 3, 2017. [Online]. Available: https://www.youtube.com/watch?v=OXtNn94l1xA
[22] (2017). *STK Simulation of Dynamic Data Collection by UAV (Study 1)*. [Online]. Available: https://www.youtube.com/watch?v=YFX1EkgB1m4
[23] M. Ammad-Uddin. *STK Simulation of Dynamic Data Collection by UAV (Study 2)*. Accessed: Jun. 13, 2019. [Online]. Available: https://www.youtube.com/watch?v=dYwm6Mv-9fs
[24] *AGI STK Home Page*. Accessed: Jul. 3, 2019. [Online]. Available: https://agi.com/products
[25] Texas Instruments Web. *CC1021 Single Chip Low Power RF Transceiver for Narrowband Systems*. Accessed: Oct. 6, 2017. [Online]. Available: http://www.ti.com/lit/ds/swrs045e/swrs045e.pdf
[26] Texas Instruments Web. *2.4 GHz IEEE 802.15.4/ZigBee-Ready RF Transceiver*. Accessed: May 2, 2017. [Online]. Available: http://www.ti.com/lit/ds/symlink/cc2420.pdf
[27] X. Li, A. Nayak, and I. Stojmenovic, "Exploiting actuator mobility for energy-efficient data collection in delay-tolerant wireless sensor networks," in *Proc. 5th Int. Conf. Netw. Services*, Apr. 2009, pp. 216–221.
[28] J. Rao and S. Biswas, "Joint routing and navigation Protocols for data harvesting in sensor networks," in *Proc. 5th IEEE Int. Conf. Mobile Ad Hoc Sensor Syst.*, Sep. 2008, pp. 143–152.
[29] M. A. Uddin, *Internet of Things (AG IOT)*, YouTube, San Mateo, CA, USA, Nov. 2017.

MOHAMMAD AMMAD UDDIN received the Ph.D. degree in wireless sensors networks from ENSTA Bretagne, France, the M.S. degree in computer networks from the COMSATS Institute of Information Technology Islamabad, Islamabad, Pakistan, and the M.Sc. degree in computer science in software engineering from Bahria University, Islamabad, Pakistan. He is currently a CCNA and a CCAI. He is currently a Senior Researcher with the Sensor Networks and Cellular Systems Research Centre, University of Tabuk, Tabuk, Saudi Arabia, where he taught many graduate and undergraduate computer courses in a number of universities at Kingdome and abroad. He led many research and development projects in the areas of wireless sensor networks, underwater sensor networks, and smart agriculture. He is the author of many research articles published in the IEEE and other journals. He is listed as an Inventor in a patent registered in the USA. His research areas include routing, clustering, and localization of sensor nodes in wireless sensor networks.

MUHAMMAD AYAZ received the Ph.D. degree in information technology from Universiti Teknologi PETRONAS, Malaysia, and the M.S. degree in computer science from SZABIST, Islamabad, Pakistan, in 2007 and 2011, respectively. He is currently an Assistant Professor and a Researcher with Sensor Networks and Cellular Systems (SNCS) Research Center, University of Tabuk, Saudi Arabia. He led various research projects supported by the University of Tabuk and the Ministry of Higher Education, Saudi Arabia, especially related to water quality monitoring. He is the author of many research articles published under the IEEE, Elsevier, Springer, Wiley, and other well-known journals. His research interests include mobile and sensor networks, routing protocols, network security, and underwater acoustic sensor networks.

EL-HADI M. AGGOUNE received the M.S. and Ph.D. degrees in electrical engineering from the University of Washington (UW), Seattle, USA. He is a Professional Engineer registered in the State of Washington. He has taught graduate and undergraduate courses in electrical engineering at many universities in USA and abroad. He served at many academic ranks including an Endowed Chair Professor, the Vice President, and a Provost. One of the laboratories, he directed won the Boeing Supplier Excellence Award Program. He was also the winner of the IEEE Professor of the Year Award, UW Branch. He is listed as an Inventor in a major patent assigned to the Boeing Company. His research work is referred in many patents including patents assigned to ABB, Switzerland, and EPRI, USA. He is currently a Professor and the Director of the Sensor Networks and Cellular Systems (SNCS) Research Center, University of Tabuk, Tabuk, Saudi Arabia. He has authored many articles in the IEEE and other journals and conferences. He is serving on many technical committees. His research interests include power systems, wireless sensor networks, scientific visualization, and neural networks.

ALI MANSOUR received the M.Sc. and Ph.D. degrees from INPG, France, and the Habilitation à Diriger des Recherches degree (Hons.) from UBO, France. He held many positions including Postdoctoral at LTIRF, France, a Researcher with RIKEN, Japan, a Teacher-Researcher with ENSIETA, France, a Senior Lecturer with Curtin University, Australia, an Invited Professor with ULCO, France, a Professor with Tabuk University, KSA, and a Professor with ENSTA Bretagne, France. He published more than 170-refereed publications and h-index of 27. In addition, he is the author of a book and coauthor of three other books and four book chapters. During his career, he had successfully supervised several postdoctoral students, and the M.Sc. and Ph.D. degree students. He was the Vice President of the IEEE Signal Processing Society in Western Australia for two years. His research interests include source separation, high-order statistics, signal processing, robotics, telecommunication, biomedical engineering, electronic warfare, and cognitive radios. He has also been the Lead Guest Editor of the *EURASIP Journal on Advances in Signal Processing*.

DENIS LE JEUNE received the bachelor's, master's, and Ph.D. degrees in signal processing from the University of Brest. He is currently a Lecturer with the University of ENSTA Bretagne, Brest, France. He was the Project Manager of definition, conception, and the development of automated remote controlled radio stations for HF reception. He was the Head of the Signal Processing Team for télécommunication systems. He is an Engineer with French MoD. He was a Research Engineer with the Institute National de la Recherche Scientifique (INRS), Canada, Altran Group Company, SAGEM, Genevilliers France, and the Signal Processing Laboratory, France. His research interests include signal processing, wireless communication, and antenna formation and localization.