IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 8, 2021, accepted January 13, 2021, date of publication January 18, 2021, date of current version February 8, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3052207

Data Collection Mechanism Based on Wavelet Multi-Resolution for Opportunistic Social Networks

WEIMIN CHEN[1,2], FANG CUI[2], AND KELVIN KIAN LOONG WONG[1,3]

[1]School of Information and Electronic Engineering, Hunan City University, Yiyang 413000, China
[2]School of Computer Science and Engineering, Central South University, Changsha 410075, China
[3]School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA 5800, Australia

Corresponding author: Kelvin Kian Loong Wong (csu_cwm@csu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61672540, Grant 61379057, and Grant 61309001; in part by the Natural Science Foundation of Hunan Province under Grant 2017JJ2057; and in part by the Project of Education Department of Hunan Province under Grant 18A397.

ABSTRACT With the shaping of universal computing concept and the development of microelectronics technology, the mobile terminal devices have the strong functions of computing, storage and communication. The opportunistic social networks composed of a large number of terminal devices can be widely used in various scenarios by deploying them anytime and anywhere. One of its most important tasks is to collect data in order to communicate between people and things. The existing researches mainly focus on routing strategies that aim to improve the performance of data collection by optimizing the routing algorithm. However, the inherent characteristics of the opportunistic social networks such as the intermittent communication opportunities make it difficult to improve routing performance because nodes can not get global network topology information. In order to solve this problem, we should establish an efficient data collection mechanism based on wavelet multi-resolution to improve the efficiency of data collection from the source, which mainly studies the multi-resolution compression storage method of node data, the spatial multi-resolution data hierarchical storage framework, and the multi-resolution data management mechanism of mobile node. The experimental results show that the multi-resolution communication mode based on integer wavelet transform can greatly reduce the amount of data in the network and the energy consumption of nodes, **and it is** beneficial to the data collection of opportunistic social networks.

INDEX TERMS Data collection, energy consumption, integer wavelet transform, multi-resolution communication, opportunistic social networks.

I. INTRODUCTION

In recent years, the progress of wireless communication and microelectronics technology has promoted the rapid popularization of mobile intelligent terminals, such as mobile phones, tablets and on-board devices. Using the embedded sensors such as temperature sensors, acceleration sensors, ultrasound sensors, magnetic force sensors, direction sensors, and so on, the mobile intelligent terminal devices take advantage of the encounter opportunity formed by the movement of people, cars or other carriers, and can realize the sense, collection and transmission of all kinds of information in the surrounding environment. This kind of network, which is composed of a large number of mobile intelligent nodes and relies on the encounter opportunity between nodes to achieve data exchange and transmission, is called Opportunistic social Networks (OppNet) [1]. Different from the traditional wireless sensor network, OppNet can communicate under the harsh conditions of network isolation, and break through the limitations of time and space of data collection. Therefore, it has a very good application prospect in various scenarios, such as absence of basic communication facilities, temporary use in emergency situations, and so on. The emergence of OppNet expands the application scope of wireless communication, and greatly promotes the communication between people and things. It is an indispensable part of ubiquitous communication in the future [2], [3].

In OppNet applications, there are many events, such as the movement of nodes, the failure and repair of sensors, which make the network is often split, the communication

The associate editor coordinating the review of this manuscript and approving it for publication was Deyu Zhang.

link is disconnected, the encounter between nodes is unpredictable and short. Therefore, OppNet needs to capture the encounter opportunities to collect and transmit the data in the way of relay forwarding [4]. However, from time to time, fleeting communication opportunities and the storage space constraints are the two main factors that restrict the efficiency of data collection, which makes the research of data collection methods in OppNet is full of challenges and opportunities [5], [6].

At present, the research of data collection in OppNet mainly focuses on routing strategies which aim to improve the efficiency of data collection by optimizing routing algorithm, such as Direct Transmission Algorithm [7], Epidemic Routing Algorithm [8] and other algorithms between the above two extremes [9], [10]. Unfortunately, in OppNet, it is difficult for nodes to obtain reliable information about the global network topology, and to predict the encounter opportunities between nodes. Therefore, the mechanism in the traditional network first establishes end-to-end connection, and then collects data, which is difficult to run in OppNet. In order to further improve the efficiency of data collection in OppNet, it is necessary to explore new methods and models. In this paper, we study the data collection mechanism in OppNet, our main work includes the following three aspects: (1) In view of the inherent characteristics of OppNet, we propose a data compression method based on integer wavelet transform to reduce the amount of transmission data in the network and improve the utilization rate of encounter opportunities. (2) In order to improve the efficiency of data collection and reduce the energy consumption of nodes, we have established an efficient data collection mechanism based on wavelet multi-resolution by studying the multi-resolution compression storage method of node data, spatial multi-resolution data hierarchical storage framework and mobile node multi-resolution data management method. (3) Using a large amount of real data, we evaluate the data collection mechanism based on wavelet multi-resolution.

II. RELATED WORK

In the OppNet, limited resources and network fragmentation make data collection very difficult. At present, the existing researches mainly focus on the routing optimization to collect the data of network [11]–[13]. Based on the routing strategy of data transmission, the researchers deeply studied the data collection method. Shah *et al.* [14] discussed the data collection method in Data-Mule system, and designed a model with three-layer routing. Among them, the lower layer was the wireless sensor that perceived the surrounding environment, the middle layer was the agent called Mule, and the upper layer was a batch of Internet access points (AP). After the underlying node collected data from the environment, the system looked for an opportunity to transmit it to the Mule, and the Mule forwarded the data to the AP. There were many types of environment-aware data, and different data generally had inconsistent emergency and security requirements, so different transmission modes need to be adopted in system. In order to solve this problem, Sun Limin *et al.* [15] proposed a quality adaptive data collection model. This mode set the data forwarding probability index for the nodes in the network, allocated the important factor to the generated message, so that the data transmission mode and the collection quality could be selected dynamically according to the importance of the message. Boldrini and Conti [16] proposed that other nodes in the networks should undertake the collection and transmission of the data if conditions permit, because the source node can not communicate directly with the destination node without a connected link, and it did not have the relevant situation of the target. In contact with each other, the node first understood each ''hobby'' of other nodes, that is, what type of data the other nodes expected to receive, and evaluated the importance of this interest. Then, the node determined a ''utility'' value according to the degree of importance as a condition for judging whether or not to collect and forward these interest data.

In addition, the energy consumption of nodes in OppNet has the characteristic of data center [17]. The nodes in the data receiving area and in the communication key area take on more tasks of data receiving, storing and forwarding, which inevitably causes the energy of these nodes to be rapidly reduced, exhausted and even unable to work. Therefore, how to use node energy in a balanced and efficient way is of great practical significance [18]. According to the temporal and spatial correlation of data, the references [19]–[21] studied the process of data propagation, network load balancing and transmission bottleneck in the classic Epidemic routing and Spray forwarding algorithm. They confirmed that the cooperative transmission mechanism can reduce the energy consumption of nodes.

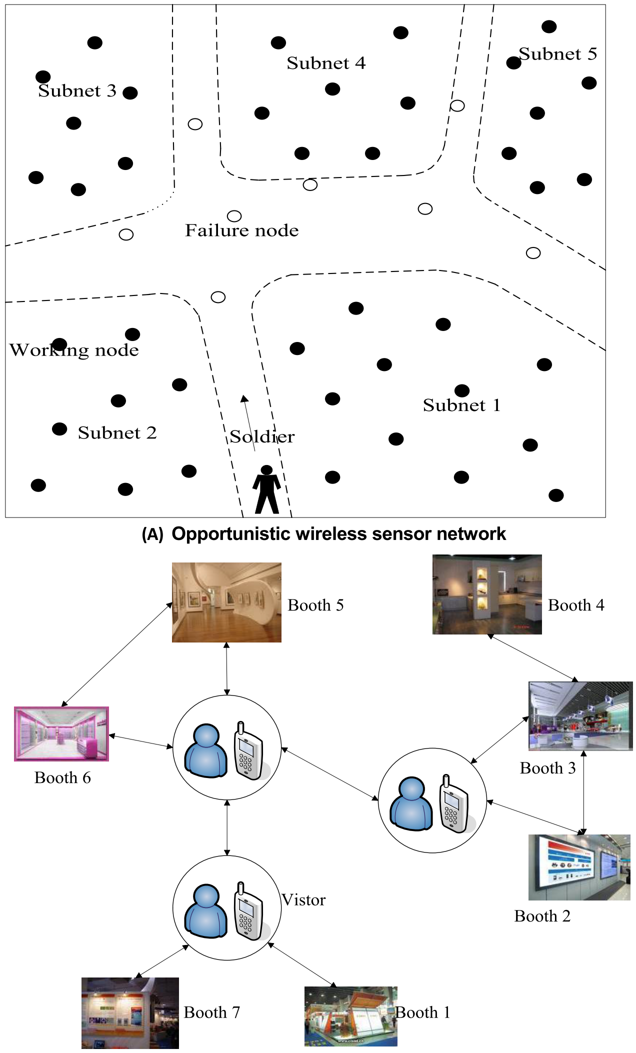
In recent years, the idea of wavelet multi-resolution analysis had appeared in the data collection method of sensor networks [22]–[24]. In clustering-based sensor networks, the underlying nodes sent the data collected from the surrounding environment to the cluster head. According to a certain algorithm, the cluster head integrated the data together and transmitted it to the upper cluster head for further processing. Different application environments often had different requirements for the accuracy and efficiency of data fusion. Some application scenarios may need cluster heads to provide high precision data fusion results, while others only need low precision data to meet the application requirements, so there was no need to transmit high precision data to save the network bandwidth, energy and overhead. References [25], [26] solved the contradiction between occasion and data accuracy by using multi-resolution data analysis. Ganesan *et al.* [27], [28] proposed the method of wavelet multi-resolution analysis to collect and store data in wireless sensor networks. The sensor used wavelet transform to compress the collected data and sent the compressed fusion data to the cluster head. After completely collecting the data of the next cluster head, the upper cluster head used wavelet transform to compress the data, then transmits it to the upper cluster head, and then repeated the process to the final Sink node. In this way, the data stored in the network formed a multi-resolution layer. Among them, the bottom node stored the high precision raw data, with the improvement of the network level, the resolution of each layer cluster head storage data gradually decreased. In the data query, the opposite "drill down" method was used to query the low precision data at the top level, and if you were interested in the data, we further drill down to the lower level to obtain higher precision data. In reference [29]–[31], the hierarchical processing and application of wavelet transform in wireless sensor networks were discussed. Each layer of data first used wavelet transform to remove the correlation and then sent it to the cluster head, which could reduce the amount of data, improved the transmission efficiency and reduced the network overhead.

In a word, the existing data collection method based on routing strategy has some limitations in transmission performance and energy consumption, and this multi-resolution data processing method widely used in sensor networks provides an effective way to solve the resource constraints on data transmission for the limited networks. Therefore, in this paper, we use multi-resolution methods to improve the data collection performance of OppNet.

III. APPLICATION SCENARIOS

OppNet is widely used in various scenarios, such as lack of infrastructure, temporary emergency using, and so on. Figure 1 are two typical application scenarios of OppNet. Due to the particularity of deployment location, the malevolence of natural environment, and the failure of sensors and communication interference, wireless sensor networks are divided into many subnets. In Figure 1 (A), the dotted line denotes the boundary of the connected subnet, the black dot denotes the working node, and the blank circle denotes the failure node. Soldiers with intelligent terminal devices scout the surrounding environment, and the arrow in the picture denotes the way forward. The arrival of soldiers brought opportunities for isolated subnets to connect and communicate with each other, then soldiers use the data collected from the network to judge the security of the surrounding environment in order to determine their direction. In addition, an example of the OppNet of a digital exhibition hall is shown in Figure 1 (B). The rectangular picture represents a booth, and the circle picture represents a visitor. The specific layout of the exhibition hall enables some booths to communicate with each other, but does not form a fully connected network. Visitors with smartphones and other communication terminal devices visit the exhibition hall, and their random movement brings communication opportunities for Internet connectivity. They view the content of the booth and assess the importance of the booth by collecting online data, and choose the booth they are interested in to browse.

In Figure 1 (A), soldiers collect data from the wireless opportunistic networks to provide a basis for their next action decisions. In Figure 1 (B), visitors determine their visit priorities and routes by collecting and analyzing the data of each booth in the exhibition hall.



(A) Opportunistic wireless sensor network

(B) Digital exhibition hall opportunistic networks

FIGURE 1. Two application scenarios of the opportunistic social networks.

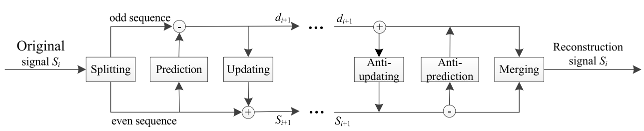


FIGURE 2. Integral wavelet transform framework.

IV. INTEGRAL WAVELET TRANSFORM

Wavelet transform is an ideal tool for signal time-frequency analysis and processing, which can provide a time-frequency window. The integral wavelet transform is the wavelet construction mode advocated by Sweldens scholars [32]. It consists of two processes: transformation process and reconstruction process. The transformation process involves three operations: splitting of raw data, prediction of odd numbers according to even numbers, and updating of scale coefficients. The reconstruction process also includes three operations: anti-updating of restoring even-number, anti-prediction of restoring odd-number, and merging of raw data. Its overall framework is shown in figure 2 as described below.

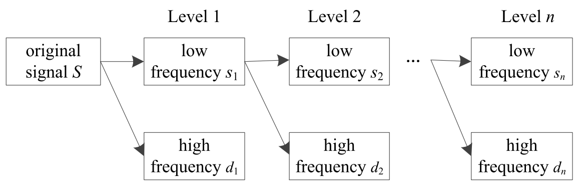


FIGURE 3. *n* **decomposition of integral wavelet.**

A. TRANSFORMATION PROCESS

1) SPLITTING

A set of original sequence signals S is divided into two subsets s_0 and d_0 according to a certain method. The general practice is to divide it into even sequence s_0 and odd sequence d_0, that is:

$$split(S) = (s_0, d_0) \tag{1}$$

2) PREDICTION

According to the correlation between adjacent signals, the value of odd signals is estimated from even signals by using prediction operations p that are not related to signals S. That is: $d_0 = p(s_0)$. The data d_1 is used to represent the difference between the real value d_0 and the estimated value $p(s_0)$ to describe the detailed information of the signal S, that is:

$$d_1 = d_0 - \lfloor p(s_0) \rfloor \tag{2}$$

where, $\lfloor \cdot \rfloor$ means rounding down.

3) UPDATING

The sequence s_0 generated after splitting and the original signal S are not consistent in terms of mean, vanishing moment, etc., so it is necessary to introduce an operator u to update the s_0 in order to obtain the scale coefficient. that is:

$$s_1 = s_0 + \lfloor u(d_1) \rfloor \tag{3}$$

the resulting subset s_1 continues the above-mentioned operation, and is divided into s_2 and d_2. After n decomposing, the original signal S is transformed into $\{s_n, d_n, d_{n-1}, d_{n-2}, \ldots, d_1\}$. The process is shown in figure 3. Among them, s_n represents the low frequency information of the signal S, and $\{d_n, d_{n-1}, d_{n-2}, \ldots, d_1\}$ represents the high frequency information of the signal S.

B. RECONSTRUCTION PROCESS

1) ANTI-UPDATING

The even sequences s_0 is restored by using s_1 and d_1, that is:

$$s_0 = s_1 - \lfloor u(d_1) \rfloor \tag{4}$$

2) ANTI-PREDICTION

The odd sequences d_0 is restored by using s_0 and d_1, that is:

$$d_0 = d_1 + \lfloor p(s_0) \rfloor \tag{5}$$

3) MERGING

The original signal S is restored by merging s_0 and d_0, that is:

$$S = (s_0, d_0) \tag{6}$$

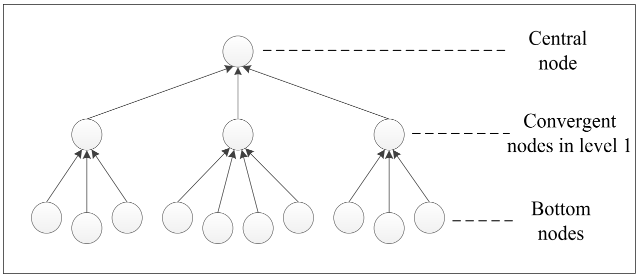


FIGURE 4. **Hierarchical structure diagram of OppNet.**

The original signal S can be restored after n reconstruction operation for the decomposition signal $\{s_n, d_n, d_{n-1}, d_{n-2}, \ldots, d_1\}$.

In the scheme of integral wavelet transform, the downward rounding operation is added in order to calculate easily. However, this rounding operation does not affect the reconstruction of original signal S, because it does not change the final calculation results. That is to say, the rounding operation will only change the value of the d_1 into an integer, and the value of s_0 and d_0 will not change as a result. After the inverse transformation operation, the data can be restored and the original signal S can not be changed. In addition, integer wavelet transform does not need convolution operation, the whole calculation process only integer shift and addition and subtraction operations. Therefore, the hardware requirement is low and easy to implement, for example, the design of 5/3 filter requires 4 multipliers and 6 adders in traditional wavelet transform, but only need respectively 2 and 4 in integer wavelet transform, which is beneficial to the implementation in OppNet with limited resources.

V. DATA COLLECTION MECHANISM BASED ON WAVELET MUTI-RESOLUTION

A. BUILDING A HIERARCHICAL NETWORK STRUCTURE

In OppNet, the data transmission between nodes can not be carried out directly because of the network splitting, and it can only be relaid by using the encounter opportunities. Therefore, it is impossible to form a hierarchical structure in generally in OppNet. However, In a certain cases, the locally connected region can establish a hierarchical structure, such as, an OppNet formed by a wireless sensor network for the failure of nodes, a locally region is connected at a given time interval. Therefore, it can be built to a hierarchical tree structure, **as shown** in figure 4. multiple bottom nodes can send data to a convergent node, and then the convergent node can gradually send it up to the upper level, thus a hierarchical structure is formed. This structure can reduce the participation of irrelevant nodes and improve communication efficiency. At the same time, it is beneficial to aggregate fusion and reduce correlation of data.

The construction process of the hierarchical structure in OppNet are as follows:

1) FORMING OF PARTITIONED NETWORK STRUCTURE

For any locally connected regional network, for example, a community, it is assumed that the network area is a circle which can be an arbitrary figure in practice, it is divided into $(2^*r_c/a)^2$ sub-networks. Where, r_c is the maximum radius of the network, a is the size of the side length of the subarea network. It is assumed that nodes within each sub-regional network are directly connected, then $2^*a^2 \leq r^2$. Where, r is the maximum communication distance of the node.

Figure 5 shows this partitioned network structure, a locally connected network is divided into 16 sub-regional networks, and each subnet is corresponded to a label. For example, s_i represents the i-th subnet. Meanwhile, each node in the network can calculate which sub-regional network does it belong to according to its position and the number of hops with the central node.

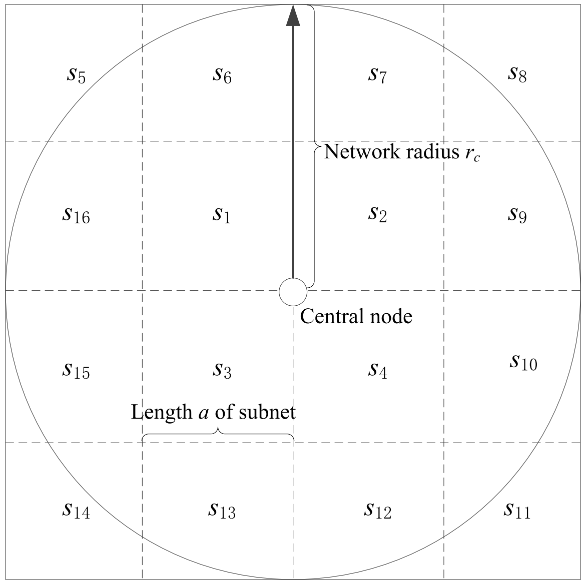


FIGURE 5. Partitioned network structure diagram of OppNet.

2) SELECTION OF CONVERGENT NODES FOR EACH SUB-NETWORK

For each sub-regional network, the convergent nodes are selected according to the residual energy, storage space and relational metric weights of nodes, that is:

$$c_i = \alpha_1 \frac{E_i}{E_{\max}} + \alpha_2 \frac{S_i}{S_{\max}} + \alpha_3 \frac{M_i}{M_{\max}} \tag{7}$$

where, the E_i is the current energy value of the node v_i, and the $E_{\max}$ is the maximum energy of the node v_i. The S_i is the current remaining storage space of v_i, and the S_{max} is the maximum remaining storage space of v_i. The M_i is the relational metric weight which can be calculated according to a certain rules such as the relationship metrics. α_1, α_2 and α_3 are adjusted parameters and meet $\alpha_{1+}\alpha_{2+}\alpha_3 = 1$. In this paper, the parameter α_1, α_2 and α_3 are equal, that is $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$, and the main reason is that the residual energy, storage space and relational metric weights of nodes are equally important in general. Then, according to the calculated value c_i, the convergent node with the largest value is selected. If the maximum value is juxtaposed, one of them is randomly selected as the convergent node. The main reason for this selection is to balance the energy consumption of nodes and improve the transmission capacity of data.

3) COMMUNICATION MODE OF HIERARCHICAL STRUCTURE

In the hierarchical network structure, the data collected by the bottom node, it is firstly sent to the convergent node of the subnet in the region, then sent to the next convergent node, and finally sent to the central node of the subnet.

B. MULTI-RESOLUTION PROCESSING OF DATA

In this section, the data collection mechanism based on the integral wavelet multi-resolution is constructed, as shown in figure 6. Firstly, the multi-resolution data compression storage method of single node is constructed, that is, the time correlation of the node data is removed. Then, the multi-resolution data storage method of network data is constructed, that is, the spatial correlation of the network data is removed.

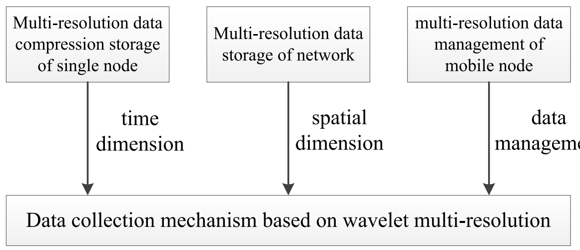


FIGURE 6. Diagram of Data collection mechanism based on wavelet multi-resolution in OppNet.

Finally, the multi-resolution data management method of mobile nodes is determined, that is, the mobile nodes collect the data removed temporal and spatial correlations in the network.

1) MULTI-RESOLUTION DATA COMPRESSION STORAGE OF SINGLE NODE

Various signals collected continuously in the network are a time series data flow, and it is represented with X_i, where i is the serial number of the node v_i. The data flow X_i consists of a sequence of data, which can be expressed as $<x_{ij}, t_{ij}>$, where x_{ij} represents the j-th data of the node v_i, and t_{ij} is the time-stamp of the x_{ij}. This data has strong correlation and periodicity, so it can be multi-resolution compressed and stored in local memory by integer wavelet transform.

Figure 7 shows the basic process of data compression using integer wavelet transform, and it is described in detail as follows:

Step 1: Data preprocessing. The task of preprocessing is to realize integer transformation of data by some means, such as unit conversion, displacement, and so on. If it is lossy compression, it is also necessary to make appropriate choices according to the requirements of data accuracy. In addition,

TABLE 1. Common integral wavelet transform.

Wavelets	Transformation Formula	
Haar Wavelet	$d_{1,j}=s_{0,2j+1}-s_{0,2j}$	
	$s_{1,j}=\left\lfloor \frac{1}{2}(s_{0,2j}+s_{0,2j+1})\right\rfloor$	
5/3 Wavelet	$d_{1,j}=s_{0,2j+1}-\left\lfloor \frac{1}{2}(s_{0,2j}+s_{0,2j+2})\right\rfloor$	
	$s_{1,j}=s_{0,2j}+\left\lfloor \frac{1}{4}(d_{1,j-1}+d_{1,j})+\frac{1}{2}\right\rfloor$	
9/7 Wavelet	$d_{1,j}^{1}=s_{0,2j+1}-\left\lfloor \alpha(s_{0,2j}+s_{0,2j+2})+\frac{1}{2}\right\rfloor$	$\alpha\approx-1.586134342$
	$s_{1,j}^{1}=s_{0,2j}+\left\lfloor \beta(d_{1,j}^{1}+d_{1,j-1}^{1})+\frac{1}{2}\right\rfloor$	$\beta\approx-0.05298011854$
	$d_{1,j}=d_{1,j}^{1}+\left\lfloor \gamma(s_{1,j}^{1}+s_{1,j+1}^{1})+\frac{1}{2}\right\rfloor$	$\gamma\approx 0.8829110762$
	$s_{1,j}=s_{1,j}^{(1)}+\left\lfloor \delta(d_{1,j}+d_{1,j-1})+\frac{1}{2}\right\rfloor$	$\delta\approx 0.4435068522$

the data can be translated by subtracting a value, for example, the median value of the data, so that the data have smaller absolute values, and the coding overhead is reduced and the compression ratio is improved.

Step 2: Integer wavelet transform. The key work of integer wavelet transform is to select appropriate filter. This selection needs to be based on the actual situation and specific needs, compromise between compression effect and calculation speed. In common integer wavelet filters, there are three wavelets: Haar wavelet, 5/3 Wavelet and 9/7 Wavelet, as shown in Table 1. Among them, Haar wavelet is the simplest, **but** the redundancy removal ability is weak, the data compression performance is limited; 9/7 Wavelet is lossy compression based on floating point operations, and has better performance in image compression, However, the computation is the largest; 5/3 Wavelet have only integer addition and shift operations, can be carried out both lossless and lossy compression, and has the advantages of less computation and better compression performance. Therefore, in this paper, 5/3 wavelet is used to non-loss or loss transform for the collected data in OppNet.

Step 3: Wavelet coefficients quantification. The quantification of wavelet coefficients affects the performance of data compression. The more accurate the quantification of wavelet coefficients is, The smaller the compression ratio of the data is. There are two common methods of coefficient quantification: vector quantification and scalar quantification. Among them, Scalar quantification is especially suitable for resource-constrained opportunistic networks because of its simple algorithm and low requirements. The process of scalar quantification is as follows: Let the maximum of data is $c_{\max}$, the minimum is $c_{\min}$, the quantification number is n, Then the step size is $(c_{\max}\text{-}c_{\min})/2^{n}$. The bigger the quantification number n is, the shorter the step size is, the higher the accuracy is, and the more the computation is. The value of n is decided by the the requirement of the quantification accuracy, and it

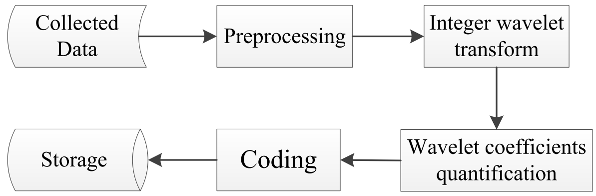


FIGURE 7. Diagram of Multi-resolution data compression storage.

should be as small as possible when the compression quality is satisfied in the practical applications.

Step 4: Coding. The deflate algorithm is used to encode wavelet coefficients. As a perfect combination of dictionary coding lz77 and Hoffman coding, the process of compression and decompression is short and the compression performance is good.

2) MULTI-RESOLUTION HIERARCHICAL STORAGE OF NETWORK DATA

The data flow generated by a single node has a strong time correlation because of the short interval, while the data set generated by the whole network has a certain spatial correlation due to the close distance. For spatial correlation between data, integer wavelet transform can also be used to remove redundant parts.

If a single node is regarded as a component of a network data sequence signal, then multiple nodes can be used to cooperate to complete the integral wavelet transform, and remove the spatial correlation between nodes. The nodes of the same level in OppNet are divided into the even node $s_{\mathrm{n}}^{(0)}$ and the odd node $d_{\mathrm{n}}^{(0)}$. Then, the node $s_{1}^{(0)}$ transmits collected data to the adjacent node $d_{1}^{(0)}$, and wavelet coefficients $d_{1}^{(1)}$ and $s_{1}^{(1)}$ can be calculated by $d_{1}^{(0)}$. In the same way, The node $s_{2}^{(0)}$ sends the data to the node $d_{2}^{(0)}$ and the $d_{2}^{(1)}$ and $s_{2}^{(1)}$ are calculated. This progress goes on, we can get all coefficients of 1-level wavelet decomposition $d_{0}^{(1)},\ldots,d_{\mathrm{n}}^{(1)}$ and $s_{0}^{(1)},\ldots,s_{\mathrm{n}}^{(1)}$. After that, A low frequency coefficient $s_{0}^{(1)},\ldots,s_{\mathrm{n}}^{(1)}$ is sent to the upper layer of convergence nodes and a new data sequence is formed, The 2-level wavelet transform coefficients $d_{0}^{(2)},\ldots,d_{\mathrm{n}}^{(2)}$ and $s_{0}^{(2)},\ldots,s_{\mathrm{n}}^{(2)}$ can be obtained by repeated above operation. Finally, the central node completes n-level wavelet transform. As shown in Figure 8, Each level of convergence node receives a transformed wavelet low frequency coefficients, and it eliminates redundancy between node data and can effectively reduce the amount of data, storage space and energy consumption.

3) MULTIRESOLUTION DATA MANAGEMENT FOR MOBILE NODES

The main task of mobile nodes access OppNet is to collect data. In the above research, the multi-resolution data compression storage of single node and the multi-resolution hierarchical data storage of multiple nodes are constructed. Therefore, mobile nodes can collect data with different resolutions according to actual needs in the process of data

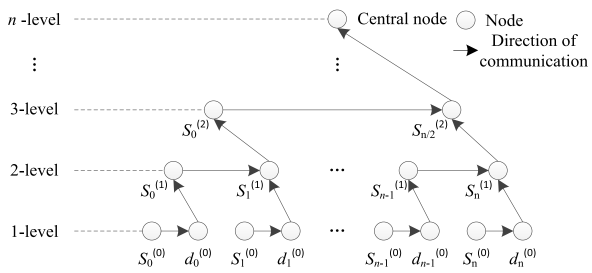


FIGURE 8. Diagram of Multi-resolution hierarchical storage of network data.

collection, and not require each sensor node to provide high precision data. In this way, for the area far away from the mobile node, the removal of data redundancy can be gradually increased, and the amount of transmission data can be reduced based on the reduction of resolution. Therefore, it can save the storage space of the mobile node, and seize the opportunity of intermittent communication.

The implementation of multi-resolution data management mechanism of mobile nodes can be based on hierarchical sub-regional network model. First, OppNet is divided into sub-regional networks to form multiple virtual grids in different positions. Then, a representative node is elected in each grid as the convergent node, and the rings of different radii are constructed with the spatial position of the mobile node as the center. Finally, the adjacent nodes of mobile nodes store the high-precision data according to the multi-resolution data hierarchical storage mechanism, and long-distance nodes store on the corresponding ring away from the central node.

VI. SIMULATIONS

To evaluate the performance of data collection mechanism based on wavelet multi-resolution in OppNets, a few simulation experiments are done with several real data.

A. EFFECT OF INTEGRAL WAVELET DECOMPOSITION

In the experiment, the time data stream collected by nodes are used to analyze the effect of integral wavelet decomposition.

The data used in the experiment were derived from the Tropical Atmospheric Ocean Project (TAO) [33]. From 1984, the project deployed about 100 sensors at different depths in 71 moorings of the tropical Pacific Ocean to capture the temperature of seawater in the relevant regions. In the experiment, we use the temperature data collected by the sensor (T0N125W) at 12:00 every day in 2011, and these data generated by the sensor is intact and without loss, omissions or errors.

The experimental results are shown in figure 9. The abscissa is the time and the ordinate is the temperature of the sea water. Figure 9 (A) shows the curves is drawn from the original data, Figure 9 (B) is drawn from the low frequency coefficients after an integral wavelet transform, Figure 9 (C) is drawn from the low frequency coefficients after three integral wavelet transform, and Figure 9 (D) is drawn from the high frequency coefficients after an integral wavelet transform. It can be seen that after integral wavelet transform decomposition, the low frequency coefficients retain the overall contour characteristics of the original data, and can be used as an approximation of the original data. The high frequency coefficients contains the main feature details of the original data. It can also be seen that the data collected by the sensor still have clear original data characteristics after 3-5 times wavelet transform. That is to say, in the actual OppNet application, the data still has a certain reference significance after 3-5 times convergence.

B. ANALYSIS OF DATA FLOW

There are two key steps in the transmission of multi-resolution data. First, the real-time transmission of low-frequency data. After the bottom nodes collect the original data, the low frequency coefficients generated by the integral wavelet transform decomposition are transmitted to the upper convergent node, and the low frequency and high frequency coefficients are stored in the memory of the bottom nodes at the same time. Then, the demand transmission of high frequency data. When the user queries and collects the data, he/she first browses the outline data of the central node which is the low frequency data formed by the integer wavelet multilevel transformation. If it meets the needs of the user, the high frequency data is transmitted to the user by the next convergent node, and the original data can be restored by inverse transformation of integer wavelet. Figure 10 shows the process, where the solid line represents the real-time transfer of the low frequency coefficients, while the dashed line represents the demand transfer of the high frequency coefficients.

For further analysis, the temperature data in TAO are simulated and analyzed. We take 360 data from the original temperature data, and carry out the following process. First, taking unit conversion. The real value of temperature is converted into integer value. Then, taking three-level integral wavelet transform decomposition and its boundary extension. After transformation decomposition, the number of high or low frequency coefficients becomes $(n + 8)/2$. Therefore, The 1-level wavelet coefficients s_1 and d_1 formed by integral wavelet transform have 184 data, the 2-level wavelet coefficients s_2 and d_2 have 96 data, and the 3-level wavelet coefficient s_3 and d_3 have 52 data. In the data transmission process, The bottom node v_i first transmit the low frequency coefficient s_1 to 2-level convergent node cv_2, and cv_2 also transmits the s_2 generated by decomposition to 3-level convergent node cv_3. At this point, the data in the memory of each level node are: v_i has s_1 and d_1, cv_2 has s_2 and d_2, cv_3 has s_3 and d_3. Then, When users query and browse these data, he/she is first received the data s_3 from the 3-level convergent node cv_3, and carries on the data summary display. If the user needs to get the raw data, he/she can receive the detail data d_3, d_2, and d_1 from the node cv_3, cv_2 and v_i, and use the wavelet transform to get the data s_2, s_1 and s_0. In the process,

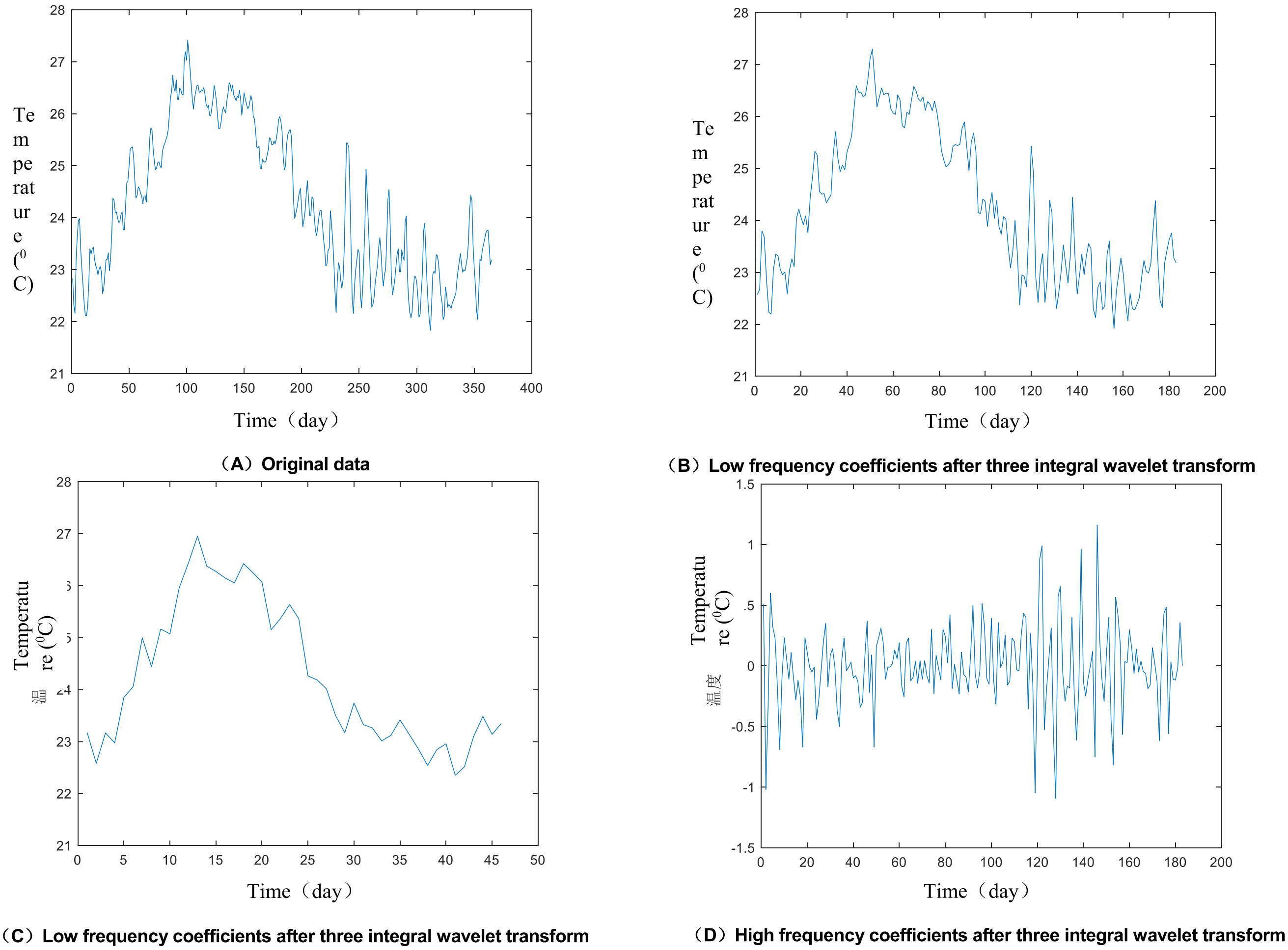


FIGURE 9. Diagram of Integral wavelet decomposition.

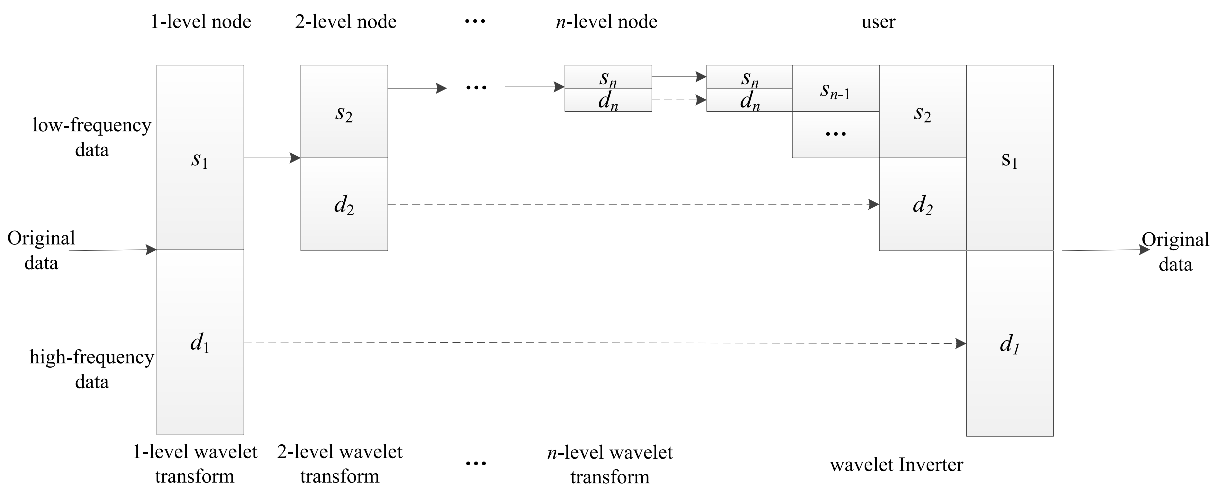


FIGURE 10. Diagram of Multiresolution data collection process.

the data received by users is 52≤data≤384. Among them, its minimum amount of data is 52 when user is not interested in the data or the accuracy of the data can meet the needs of the user, and its maximum amount of data is $52 + 52 + 96 + 184 = 384$ when user need to receive data s_3, d_3, d_2 and d_1 in turn to get the original data, he/she receives 24 more

TABLE 2. Traffic of multi-resolution data.

Communication mode	Node v_i	2-level convergent node cv_2	3-level convergent node cv_3	User
Traditional mode	360 (s)	360 (s)	360 (s)	360 (s)
Multi-resolution mode	368 (s_1、d_1)	192 (s_2、d_2)	104 (s_3、d_3)	52≤ data ≤384

data than the original data because of the need for boundary processing, The average amount of data is only 192 when it is assumed 50% probability of abandoning or continuing data query. Compared with the traditional communication mode, the results are shown in Table 1, Wavelet transform data traffic undoubtedly has certain advantages.

In the above data transmission process, the lossless multi-resolution storage of wavelet transform data is only considered, and the lossy compression of data after removing time and space correlation is not involved. The lossless data compression based on integral wavelet transform can achieve compression ratio of 4-19 times, and lossy compression can achieve higher compression ratio. Therefore, the multi-resolution communication mechanism based on integer wavelet transform can greatly reduce the amount of data in the network and the energy consumption of nodes, which is beneficial to the data collection of OppNet.

VII. CONCLUSION

Data collection is the main task and core problem of OppNet. Compared with routing algorithm, the research on the processing method of data is the fundamental motive force to improve the efficiency of data collection. In this paper, we propose the multi-resolution data collection mechanism based on integral wavelet transform, which studies the multi-resolution compression storage of the time series data flow, the multi-resolution hierarchical storage of network data, and multi-resolution data management of mobile node. It can effectively remove data redundancy, reduce the amount of data involved in communication in the network energy consumption of nodes, and is an efficient data collection mechanism in OppNet.

REFERENCES

[1] Y. P. Xiong, L. M. Sun, J. W. Niu, and Y. Liu, "Opportunistic networks," *J. Softw.*, vol. 20, no. 1, pp. 124–137, 2009.

[2] H. H. Wu, "Research on the key technologies of video delivery in opportunistic networks," Beijing Univ. Posts Telecommun., Beijing, China, Tech. Rep., 2015.

[3] Y. Wang, S. F. Wang, S. S. Xu, C. X. Zhao, and L. Chang, "Node cooperative forwarding mechanism of opportunity social network based on Merkle hash tree," *J. Chin. Comput. Syst.*, vol. 40, no. 7, pp. 104–109, 2019.

[4] J. Wu, Z. Chen, and M. Zhao, "Information cache management and data transmission algorithm in opportunistic social networks," *Wireless Netw.*, vol. 25, no. 6, pp. 2977–2988, Aug. 2019.

[5] M. Guo and M. Xiao, "MSSN: An attribute-aware transmission algorithm exploiting node similarity for opportunistic social networks," *Information*, vol. 10, no. 10, pp. 299–320, Sep. 2019.

[6] R. Wang, X. Wang, F. Hao, L. Zhang, S. Liu, L. Wang, and Y. Lin, "Social identity-aware opportunistic routing in mobile social networks," *Trans. Emerg. Telecommun. Technol.*, vol. 29, no. 5, p. e3297, May 2018.

[7] V. Kuppusamy, "Performance analysis of epidemic routing in destinationless OppNets," in *Proc. WoWMoM*, 2018, pp. 1–3.

[8] D. Dash and D. K. Sanyal, "Steiner system-based topology-transparent priority scheduling for wireless ad hoc networks," *Internet Technol. Lett.*, vol. 2, no. 3, p. e102, May 2019.

[9] D. Lynch, M. Fenton, D. Fagan, S. Kucera, H. Claussen, and M. O'Neill, "Automated self-optimization in heterogeneous wireless communications networks," *IEEE/ACM Trans. Netw.*, vol. 27, no. 1, pp. 419–432, Feb. 2019.

[10] L. Zhang, C. Yu, and H. Jin, "Dynamic spray and wait routing protocol for delay tolerant networks," in *Proc. Netw. Parallel Comput. (NPC)*, Gwangju, South Korea, Sep. 2012.

[11] D. Zhang, Z. Chen, M. K. Awad, N. Zhang, H. Zhou, and X. S. Shen, "Utility-optimal resource management and allocation algorithm for energy harvesting cognitive radio sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3552–3565, Dec. 2016.

[12] D. Zhang, Y. Qiao, L. She, R. Shen, J. Ren, and Y. Zhang, "Two time-scale resource management for green Internet of Things networks," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 545–556, Feb. 2019, doi: 10.1109/JIOT.2018.2842766.

[13] F. Tang, Y. Kawamoto, N. Kato, K. Yano, and Y. Suzuki, "Probe delay based adaptive port scanning for IoT devices with private IP address behind NAT," *IEEE Netw.*, vol. 34, no. 2, pp. 195–201, Mar. 2020.

[14] R. C. Shah, S. Roy, S. Jain, and W. Brunette, "Data MULEs: Modeling a three-tier architecture for sparse sensor networks," in *Proc. 1st IEEE Int. Workshop Sensor Netw. Protocols Appl.*, May 2003, pp. 30–41.

[15] L. M. Sun, Y. P. Xiong, and J. Ma, "Adaptive data collection mechanism in mobile sensor networks," *J. Commun.*, vol. 29, no. 11, pp. 186–193, 2008.

[16] M. Conti, M. Mordacchini, and A. Passarella, "Data dissemination in opportunistic networks using cognitive heuristics," in *Proc. IEEE World Wireless, Mobile Multimedia Netw.*, Jun. 2008, pp. 1–6.

[17] D. Zhang, Z. Chen, J. Ren, N. Zhang, M. K. Awad, H. Zhou, and X. S. Shen, "Energy-harvesting-aided spectrum sensing and data transmission in heterogeneous cognitive radio sensor network," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 831–843, Jan. 2017.

[18] J. Zhang, D. Zhang, X. Xu, F. Jia, Y. Liu, X. Liu, J. Ren, and Y. Zhang, "MobiPose: Real-time multi-person pose estimation on mobile devices," in *Proc. 18th ACM Conf. Embedded Netw. Sensor Syst. (SenSys)*, New York, NY, USA, Nov. 2020, pp. 136–149.

[19] L. Buttyán, L. Dóra, M. Félegyházi, and I. Vajda, "Barter trade improves message delivery in opportunistic networks," *Ad Hoc Netw.*, vol. 8, no. 1, pp. 1–14, Jan. 2010.

[20] G. Dini and A. Lo Duca, "A reputation-based approach to tolerate misbehaving carriers in delay tolerant networks," in *Proc. IEEE Symp. Comput. Commun.*, Jun. 2010, pp. 772–777.

[21] D. Zhang, L. Tan, J. Ren, M. K. Awad, S. Zhang, Y. Zhang, and P.-J. Wan, "Near-optimal and truthful online auction for computation offloading in green edge-computing systems," *IEEE Trans. Mobile Comput.*, vol. 19, no. 4, pp. 880–893, Apr. 2020.

[22] L. J. Wang, Z. Y. Gao, and P. S. Yao, "Data compression algorithm based on space-time correlation for sensor networks," *J. Jilin Univ.*, vol. 242, no. 2, pp. 155–160, Jun. 2020.

[23] H. Dong, J. G. Lu, and Y. X. Sun, "Distributed wavelet compression in wireless sensor networks," *Chin. J. Sens. Actuators*, vol. 20, no. 11, pp. 2481–2486, 2007.

[24] J. L. Wang, *A Research on Wavelet Data Compression Algorithms for Wireless Sensor Networks*. Changsha, China: Hunan Univ., 2007.

[25] H. Qi, S. Iyengar, and K. Chakrabarty, "Multiresolution data integration using mobile agents in distributed sensor networks," *IEEE Trans. Syst., Man Cybern. C, Appl. Rev.*, vol. 31, no. 3, pp. 383–391, Aug. 2001.

[26] H. Qi, Y. Xu, and X. Wang, "Mobile-agent-based collaborative signal and information processing in sensor networks," *Proc. IEEE*, vol. 91, no. 8, pp. 1172–1183, Aug. 2003.

[27] D. Ganesan, D. Estrin, and J. Heidemann, "Dimensions: Why do we need a new data handling architecture for sensor networks?" *ACM SIGCOMM Comput. Commun. Rev.*, vol. 33, no. 1, pp. 143–148, Jan. 2003.

[28] D. Ganesan, B. Greenstein, D. Estrin, J. Heidemann, and R. Govindan, "Multiresolution storage and search in sensor networks," *ACM Trans. Storage*, vol. 1, no. 3, pp. 277–315, Aug. 2005.

[29] A. Ciancio, S. Pattem, A. Ortega, and B. Krishnamachari, "Energy-efficient data representation and routing for wireless sensor networks based on a distributed wavelet compression algorithm," in *Proc. 5th Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Nashville, TN, USA, 2006, pp. 309–316.

[30] R. S. Wagner, R. G. Baraniuk, S. Du, D. B. Johnson, and A. Cohen, "An architecture for distributed wavelet analysis and processing in sensor networks," in *Proc. 5th Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, 2006, pp. 243–250.

[31] F. Tang, Z. M. Fadlullah, B. Mao, N. Kato, F. Ono, and R. Miura, "On a novel adaptive UAV-mounted cloudlet-aided recommendation system for LBSNs," *IEEE Trans. Emerg. Topics Comput.*, vol. 7, no. 4, pp. 565–577, Oct. 2019.

[32] M. D. Adam and F. Kossentni, "Reversible integer-to-integer wavelet transforms for image compression: Performance evaluation and analysis," *IEEE Trans. Image Process.*, vol. 9, no. 6, pp. 1010–1024, Jun. 2000.

[33] (2018). *Nation Data Buoy Center*. [Online]. Available: https://tao.ndbc.noaa.gov/proj_overview/proj_overview_ndbc.shtml

WEIMIN CHEN received the B.Sc. degree in physics from Hunan Normal University, China, in 1995, the M.Sc. degree in software engineering from Hunan University, China, in 2011, and the Ph.D. degree in computer science and engineering from Central South University. His research interests include opportunistic networks, wireless networks, software engineering, and bio-computer.

FANG CUI received the B.Sc. degree from Hunan City University, China, in 2008, and the M.Sc. degree in computer science and engineering from Central South University, China, in 2020. Her research interests include opportunistic networks and wireless networks.

KELVIN KIAN LOONG WONG received the B.Sc. degree from Nanyang Technological University, Singapore, in 2001, the M.Sc. degree from The University of Sydney, Australia, in 2003, and the Ph.D. degree from The University of Adelaide, Australia. His research interests include artificial intelligence and bioinformatics.