RESEARCH ON DISTRIBUTED DATA STORAGE BASED ON INFORMATION PROCESSING CLOUD

LUO Qing-Yun, ZHU Ling-Zhi [*], CHAGN Yun-Jie[*], ZHAO Jin-Guo, LIAO Wei-Sheng, HE Rui

Department of Computer and Information Science, Hunan Institute of Technology, Hengyang 421002, Hunan, China

Emails: 121471998@qq.com , lingzhi0825@163.com *, changyunjie@aliyun.com *

Submitted: May 3, 2016 ***Accepted: July 30, 2016*** ***Published: Sep. 1, 2016***

Abstract- Sensor network is a data-centric network, which provides data collection, storage and query services. Data storage and query is one of the hot spot in the research of sensor networks. In order to solve the problem of low efficiency of storage and query，high energy consumption in sensor networks, we put forward a scheme that storing distributed data of wireless sensor network based on information processing cloud. Information processing cloud is made up of a group of sensor nodes around the network center, which have the ability to absorb and process data from other nodes of sensor network which do not belong to the information processing cloud. The group of sensor nodes around the network center respond data query requests from anywhere of the network, and sensor nodes can be adjusted dynamically according to real situation, the cloud of nodes and non-cloud of nodes can be dynamically transformed as well. When non-cloud of sensor nodes store data or send query request, they only need to do centripetal movement, centrifugal movement or circumfusing movement. The analysis shows that the proposed scheme can simplify route algorithm of data storage and data query, and it also has less computation cost and storage cost than the existing schemes.

Key words: sensor networks; information processing cloud; data storage; data query

I. INTRODUCTION

The distributed data storage scheme of wireless sensor network generally choose some nodes of the network for storing data, other nodes which do not store data are called as common node. Usually, the data generating rate of every node are fixed, we hope nodes for storing data and high data generating rate nodes are as close as possible under the prerequisite that data query is convenient so that energy cost of delivering data can be decreased, which means sensor network should be able to adjust the location of nodes for storing data dynamically when querying data is convenient. Energy balance means the high energy nodes in the network shoulder data storage and query request task together to reduce nodes' energy cost outside the cloud. Relatively less data storage nodes will be set in areas with small data size in order to reduce the consumption of data query and delay. More data storage nodes will be set in areas with larger data size to share the load and avoid data overflow. Information processing cloud is an assemblage of nodes around the network center which consists of data storage nodes, link storage nodes, intermediate and their neighbor nodes.

II. RELATED WORK

At present, the data storage of wireless sensor network is classified into four categories like external storage, local storage, data-centered storage and position-centered storage. External storage is that all nodes transmit their sensitive data to the base station which put the data into external database; this way of storage is convenient for data query. When the frequency of querying is much high than that of generating data, choosing this way of querying data is more suitable. Otherwise, this way will consume so much energy when transmitting data. Local storage is that nodes store detected data locally and response data query requests by themselves, this method will be relatively effective only when the frequency of generating data is much higher than that of querying data, otherwise nodes will consume massive energy due to the flooding query requests. Data-centered storage is that executing storage by nominating data and use geographic hash table to map a certain type of data to a certain position in the network, in this way we need only send query requests to associated storage

nodes of attribution data to execute query.

Fortunately, the research has made some development on the data storage and data query in WSNs. Technically, these methods are mainly based on the idea. We focus on the following classic schemes of the data storage and data query in WSNs:

a. DCS and GHT

Ratnasamy puts forward the concept of data-centric storage (DCS), and designs the data storage algorithm based on geographic information mapping table (Geographic Hash Table, GHT). The basic principle of the GHT algorithm is based on data attribute to store data, it uses geographic hash table map to a certain location in wireless sensor network, and chooses the location nearest node as the data storage nodes. Therefore，the same type of data is stored in the same node, and query only need to be sent to the including event nodes without flooding query, so the way can greatly save the energy dissipations of the query. But all of the same event and query in the network to be sent to the master node, the energy consumption of master nodes and non-master nodes is most serious, and then energy balance are very poor. Although this way of storing data is convenient for querying, still there are several problems in this scheme:

1) The cost of storing data is very high.

2) Not sensitive to distance. Occasionally when the position where generates data is quite close to the place where data query requests comes from, the data still must be send to associated nodes which may much far from the position where generated data to store.

3) Low reliability. Data storage nodes are easier to lose effectiveness for consuming too much energy.

4) Unable to support compound data type convergence effectively like query “searching information about detected animals and conveyances”.

b. C-DCS

C-DCS [1] is a distributed data storage and retrieval scheme. In C-DCS, the sensor network space is divided into equal-sized clusters, and the number of these clusters is c. C-DCS provides that all the nodes are divided into two categories: the mobile node and stable node,

and only stable node can participate in data storage and routing. Therefore, each node needs to store a stable neighbor node list, when the data forwarding, select only stable nodes in the neighbor list as the next hop node. Similar to the data storage process, data retrieval nodes sends a query to the query data mapping location of nodes, and the query data mapping location of the covering cluster of head nodes returns the query results to the data retrieval node. The specific process is shown in Figure 1.

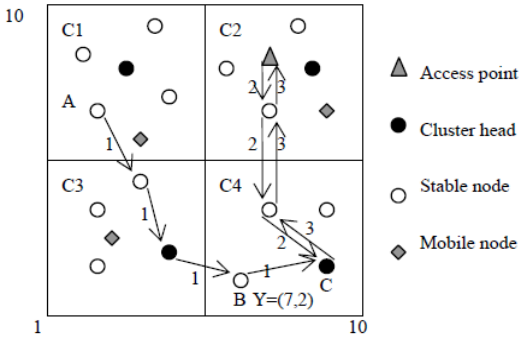

Figure1. Data storage and retrieval process

When the data storage and retrieval of the mobile nodes needs to broadcast the help message of REQ to the neighbors, after the neighbor nodes receives the help message of REQ, if the neighbor nodes is willing to help the mobile nodes forward the data node, the mobile node returns to the agreement help message of APV. After the mobile node receives the agreement help message of APV, and select the most close to the target location of neighbor nodes as the next-hop node. This process is shown in Figure 2.

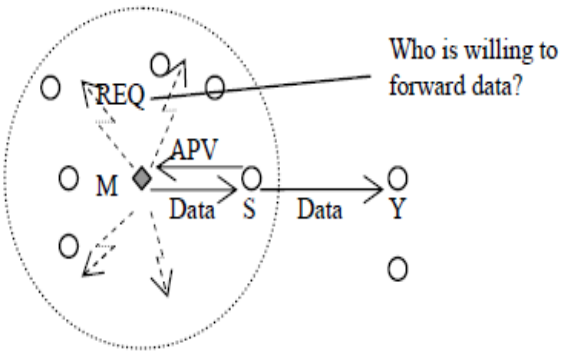

Figure2. Mobile node

In order to reduce the energy consumed of data transmission, event data generated by all

nodes of the cluster is first sent to the cluster head nodes aggregation, and then the local cluster head node converge the results to the covering data storage locations mapped clusters of cluster head nodes. Cluster head nodes receive a query request and converge toward the stored data, and will gather the results back to the inquirers. Improved data storage and retrieval process is shown in Figure 3.

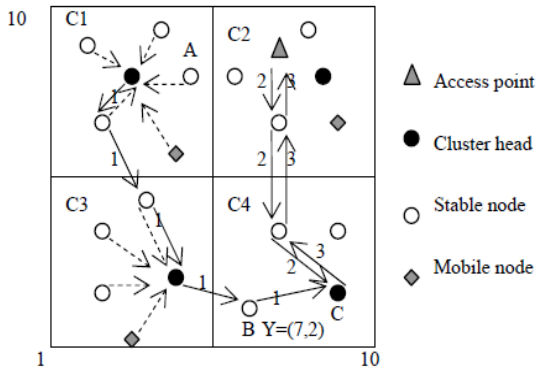

Figure3. Two levels of aggregation

c. DRIB

DRIB [2] is a kind of distributed data storage scheme based on data transmission path. The scheme builds a virtual boundary of four anchor nodes and four segments constitute the axis. In the virtual boundary, *X, Y, Z and Z' are* four anchor nodes, *ZX, XZ', Z'Y*, and *YZ* axis are the four segments. Double-ruling based on the spherical is shown in Figure 4. For data storage, if the minimum number of hops of the data producer to the virtual boundary of the nodes is greater than 1, the data producer copy data to the first direction closest of nodes. On the contrary, the data producer copy data to the side edge *YZ* and *XZ'*. From Figure 5, when the network boundary shapes rules, the transmission path of query request and the replication path of event data coverage inevitably. There are same problems in DRIB, when the network boundary irregular shape, the transmission path of query request and the replication path of event data don't coverage.

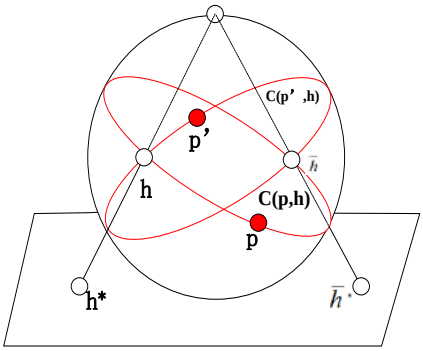

Figure4. Double-ruling based on the spherical

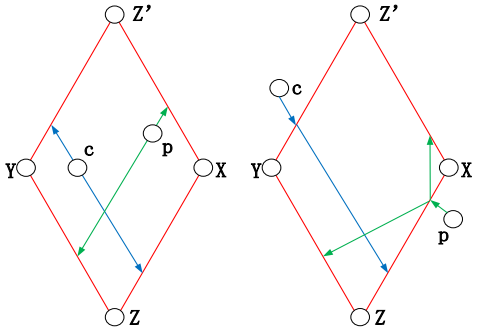

Figure5. DRIB

d. ADS

The main idea of ADS is to divide the sensor network area into grids. Every grid and the same grid at different time periods use different data storage and retrieval strategies. ADS in each mesh data storage solutions can convert between the two data storage solutions: local storage and data-centric storage.

If $Event_{num}^{t+1}$ and $Query_{num}^{t+1}$ satisfies a formula (1), it switches to local storage mode:

$$Event_{num}^{t+1} > (Query_{num}^{t+1} + 1) * \sqrt{n} + T \tag{1}$$

If $Event_{num}^{t+1}$ and $Query_{num}^{t+1}$ satisfies a formula (2), it switches to a data-centric data storage mode:

$$Event_{num}^{t+1} < (Query_{num}^{t+1} - 1) * \sqrt{n} - T \tag{2}$$

e. EDFC

EDFC uses coding techniques for distributed data storage. In EDFC, according to Robust Soliton distribution [3], each node selects its own encoding degree d, and each data source node selects b paths randomly and then forwards the data packet by random walk method. The b calculation formula is:

$$b = \frac{N \sum_{d=1}^{k} x_d d \mu(d)}{k} \tag{3}$$

and then the coding agree of nodes is *d and* the number of the data source packets of these nodes is $x_d d$. $\mu(d)$ shows the specific gravity of these nodes.

III. DEFINITION AND THEOREM

a. The node classification

Provided nodes were random distributed and fully distributed network model was adopted as system model. Sensor nodes in the information processing cloud:

1) The main link nodes: connection nodes and data storage nodes

2) The sign nodes;

Nodes outside in the information processing cloud: The centripetal node and the centrifugal nodes (there is a mark variable, rout_tag, inside each node, when the value of rout_tag is 0, the node is centripetal node, if the value is 1, the node is centrifugal node).

b. Definition

1) The transmitting radius of the node: The maximum distance the signal launched by the node can reach with proper power.

2) The turning radius of the node: In the process of expanding data storage nodes, when the newly participated node is seeking another one to join in, the node will start whirling clockwise or anticlockwise around the line which connect the node and the node to be expanded with a certain length, in this process the turning length is named as the turning radius of the node.

3) The extended circle of data storage node: make the data storage node that need to expand as the center of the circle, set the transmission radius of the sensor nodes as the radius of the circle.

4) Launch node of the extended circle: the node which is the nearest to the associated extended circle of the data storage which is executing expansion at both sides of the node on the main link of the information processing cloud.

c. Theorem

Theorem: Information processing cloud primary link is closed, it is the flat area where the sensor network is divided into two parts inside and outside. In the process of information processing dynamic expansion and contraction of the cloud, the network center position has always been that part of the area included in the information processing cloud surrounded the main link.

When the data storage node on the primary link from the network to a central location is less than twice the radius of the transmission node, if the data acceptance rate of the data storage node is greater than the specified upper threshold data acceptance rate, and the received data center network location closer to the center position, at which point the data storage node is not performing dynamic expansion algorithm, but to broadcast the message to change its neighbor nodes type, the message contains the new node types and the message sender node ID and node types . The node receiving the message checks the message sender node type of the node, if the node type for the data storage node and is a Class A index node, the node compared with the sender of the message from the network to a central location, while the node to the network center position is less than that when the message sender to the location of the center of the network, the node change node type to data storage nodes, which storage they generated and received data. Due to node type change message the new data storage node received and does not need to set up the establishment of a new message to its neighbor nodes broadcast data storage node. If the data storage node on the primary link caused by the outward expansion or energy is below a given threshold or other reasons, it is no longer being a data storage node, if the node is also a Class A index node, the node broadcasts news to its neighbors index revocation.

When the data storage node of main link data found that in the incidence of a node is greater than the data rate of the received upper limit, then the data storage node on the primary link increases in the index node class node type -B, saving the node pointer which data generation rate is greater than the data rate of the receiving node pointer upper limit, the node sends to a node type change message, it will notify its node type change into data storage node.

IV. TDISTRIBUTED DATA STORAGE BASED ON INFORMATION PROCESSING CLOUD

a. Construction of the initial information processing cloud

Assume that the sensor network center position is known. To find the center position of network nodes nearest distance in the network, the method is: after sensor nodes arranged in, by the network at any node using GPSR routing protocol to send the center location of the network for the distance to the center of the network recently location request message. The message to the destination node is the distance to the center of the network location of the closest node. In order to facilitate the description, we call near the center position of the nearest node to node C.

The construction of the message processing cloud is initiated by the C node. Specific method is: C nodes choose a distance to point C as the center of a circle with R as the radius of the circle of circumference of the nearest node B and is transmitted to the node information cloud processing building message. After receiving the message, node B along with C as the center of a circle, with R as the radius of the circle as a circular motion, on his way through the point as data storage nodes, each a

Data storage nodes to the hop neighbor nodes broadcast new data storage node setup message is received the news of the neighbor node to the node type of landmark nodes and also record new data storage node ID.

b. Dynamic Adjustment for Information Processing Cloud

Cloud information processing of dynamic adjustment includes information processing cloud expansion, shrinkage, the replaced node, cloud internal node type conversion, and conversion of cloud node types etc.

b.i Cloud decided to dynamically adjust the information processing factors

In most of the time, information processing cloud is in a stable state, only under certain conditions, information processing cloud will be dynamically adjusted.

The first factor is the data acceptance rate of the information processing cloud main chain. Information processing of the main backbone of the road nodes is divided into two kinds, one

is the data storage node, and the other is the connection node. For the data storage nodes, one of which is in steady state conditions is the data storage node data receiving rate to the door under the constraints and threshold value. The door limit is the maximum data refers to the data storage nodes which can accept rate under the threshold is the smallest index according to the data storage nodes can be allowed to accept rate. When the data storage node data receiving rate is greater than the threshold value, or will the information processing cloud neighbor nodes in the information processing of cloud road connections to the backbone data storage node (node on the road belongs to the neighbor nodes of the backbone node of the connection node), or dynamic expansion in the node nearby, along the extension node corresponding circle track to find the data storage nodes to share the new, by increasing the number of the old data storage node near the new data storage node Old data storage node load through to the data to mobile data storage node position to reduce the energy consumption required for data storage. When the data storage node data receiving rate is less than the threshold value, data storage nodes to of node type conversion, the own node types connected node, while notifying the shift of its neighbor nodes.

The second factor is that the residual energy information processing nodes in the road. Below the cloud backbone nodes at a given threshold energy (the original data storage node) can choose distance data from their neighbor nodes have higher rate of regional center position is near, the residual energy of a larger data storage node as a new node, and you jump up and down the sending node to the node information. If the original data storage node in the backbone of the road next hop node is also a new data storage node's neighbors, the new data storage nodes can be directly and the original data storage node on the next hop node to connect. If not, the new data storage nodes need the help of another a bridge node indirectly connect bridge node cannot be the backbone node way and also to become the data storage nodes join the backbone path. Then, the original data storage Node to the neighbor node broadcast data storage node to delete information, and the new data storage node to the neighbor node broadcast the new data node set up information.

b.ii Cloud decided to dynamically adjust the information processing factors

The information processing in the cloud data storage node data receiving rate is not too high, too high that the data storage node overload; at the same time, the data storage node, data receiving rate cannot be too low, too low will reduce the efficiency of data query. We set the upper and lower data rate for receiving data storage section when the data storage node, data receiving rate is lower than the lower limit, the data storage node the node type is changed into a connecting node; receiving rate when the data storage node data is higher than the upper limit, the data storage node first checks its neighbors in the information processing of cloud backbone node has no road connection, if there is. The node will notice the node type into a data storage node, and then by the node to its one hop neighbor node sends the new data storage node establishment message; if not, the Information Office of Li cloud Expansion. Expansion method is higher than the data receiving rate of nodes on the threshold of the first to the two hop neighbor broadcast message expansion and expansion that is contained in a message to be dynamically extended data storage node locations and ID information, assume that nodes from start sending extended message to destination node completely to receive extended news consumption time t, expansion message sending nodes waiting for the 3t after sending the message expansion in time in the neighborhood of that point, if in this time adduction to neighbor nodes cancel extended information is the nodes wait for a data collection period T and send the message expansion. A node receives from two different nodes of the extended message, to compare the two nodes of the ID size, and then unextend sends a message to the node ID smaller. When extended message sent to the neighbor node has no data received in waiting for 3t time after unextend news, the calculation of a cycle received in the average position of lc source node data, if lc covered in the data storage node, the data storage node selects data from the neighbor node yields several nodes larger as the data storage node of the new place of its own into the information processing link. Otherwise the comparison of cloud distance between the lc and the data storage nodes to the Network Center, if the lc to the network center distance is greater than the data storage node to the network center distance, the information processing in the cloud data Storage node outward expansion adjustment, otherwise inward expansion adjustment.

The extension method of information processing cloud on the data storage node is： First, the data storage node sends an extension start message to the start node of the expansion circle

corresponding to the data storage node. The initial rotation radius of the node ID and node of the extended start message is included in the extended start message. The initial radius of rotation of the joint is calculated by the data storage node which needs to be expanded according to the local information. The calculation formula is as follows:

$$R_t = \min\left\{ \frac{\pi R_\tau}{(r_T * (N_c - 2)/r_{normal})}, R_c \right\} \quad (1)$$

R_t node represents the rotation radius, R_c represents the node transmission radius, r_T denotes the need extended data storage nodes in a time period of data receiving rate, N_c represents the expansion of the number of nodes on the backbone road in round, r_{normal} denote the maximum value allowed for data storage node data receiving rate.

When the start node of the expanded circle of the data storage node receives the expanded round starting message from the data storage node, the start node of the expansion circle takes the R_t as the radius from the starting point to point to the direction of the data storage node corresponding to the expansion circle and rotates clockwise or counterclockwise(If the starting point of the extended circle points to the data storage node in the expansion circle, the starting point of the extended circle can be rotated clockwise to reach the starting point of the expansion circle and the direction of the vector of the direction of the data center, rotating the start node of the extended circle clockwise, otherwise rotate clockwise.), to the reverse direction of the direction when it arrives at the beginning of the rotation. This process we call a scan, in a region that is scanned in a scan, we call the scanning area of the node that is initiated. After a scan of the start node of the extended circle, select the range of the scanning area to extend the circle recently and to extend the circle distance less than R_{c/K_1} $(k_1 > 1)$ of the node as the new next hop data storage node, and forward the extension to the node to start the message.

If there is no node that satisfies the condition in the scanning area of the node, the node adjusts its scan radius to re scan.

Scanning radius of the first $k_2(k_2 = 1,2,3,\ldots.)$ scan is:

$$k_2 R_\tau = \min\left\{ k_2 \bullet \pi R_\tau \Big/ \left(r_T * (N_c - 2) \Big/ r_{normal} \right) \right\} \quad (2)$$

Receive the next hop node data storage node of the extended boot messages in the same way to find their next hop new data storage node, the process continues until a new data storage node finds another start node contained in the message in the starting scan area. At this time, found included in the message extension start another extended start node new data storage node expansion success message to the start node when the expansion, expansion started to receive this extended success message node, the data storage node sends it to the corresponding extended circle, then the data storage node corresponding to the extended circle two hop neighbor broadcast message expansion success.

When the expansion node circle (link two extended start node except) received the expansion of successful message, delete information to its neighbor nodes broadcast link node. Link node deletion information contains the message publisher information and an extended EXPEND_TAG. EXPEND_TAG sign value of 0 information processing cloud backbone road at the node expansion, and when the EXPEND_TAG value is 1 node information processing cloud node to delete the information received extended inwards. The main link node link node information from the data storage node of the node set or link node delete. Centralized data storage nodes when a node set and connection set for a null value, the node checks that the two set by at least one non null to null value for link node delete the message. If the value of EXPEND_TAG in EXPEND_TA G value is 0, the node will set its own node type to the centrifugal node, if the value of EXPEND_TAG is 1, the node will set its own node type to the central node.

When the establishment of new data storage nodes receives the extending success message, they broadcast the message of data storage node established to the new neighbor node. When they receive the message of data storage nodes established, if the neighbor nodes are cloud nodes, the neighbor nodes save the ID of data storage nodes established to themselves data storage node set. If the neighbor nodes aren’t cloud nodes, these nodes will vary the type of own node to sign nodes, and then generate an empty data storage nodes set, the establishment of new data storage nodes added to this data storage nodes set.

b.iii The Dynamic Shrinking of the Information Processing Cloud

If the information processing cloud to perform only extended operation, information processing cloud backbone road will become more and more. In this way, when the network aware data generation was reduced backbone road will appear a large number of connections and queries on these nodes not to need the data but will also must go through these nodes, so in the data query will waste a lot of energy. In order to reduce such a waste of energy, information processing cloud need dynamic contraction to reduce the information processing cloud backbone road connecting node number.

(1) Dynamic contraction information processing cloudy conditions

There are no other data storage nodes on the main link between the two data storage nodes on the main link， and the number of connection nodes between the two data storage nodes is larger than a predetermined threshold value. In order to save the energy consumption of data query, information processing cloud may need to be carried out in these two data storage nodes to adjust the shrinkage. Whether the two data storage nodes to adjust the conditions of contraction, in addition to the above mentioned one, but also need another. That is, the number of connection nodes between the two data storage nodes is larger than the distance between the two data storage nodes and the ratio of the transmission radius of the sensor nodes is () times. If representation between the two data storage nodes, the distance between these two nodes, the nodes represent transmission radius, the information processing in the cloud node A, B between the conditions required for shrinkage adjustment can be summarized as:

1) On the main link between A and B only connect the nodes；

2） $N_{reallink} \rangle \max\left\{k_{\Im} * {}^{D_{A\_B}}\!/_{R_C}, n_{link}\right\})$

(2) Dynamic Shrinkage implementations of information processing cloud

Each data storage node saves a T_{shrink} contraction cycle and a timer. In order to avoid frequent contraction of energy waste, the data storage node information processing on the cloud backbone only in the query request arrival, and the timer value is greater than the

systolic contraction cycle check node values stored in the reserve work. The data storage node on the backbone when receiving the query, the query first checks whether there have variable $n_{reallink}$, ID_{last} and LC_{last}, if these three variables, $n_{reallink}$ check whether the value of the number of nodes connected is greater than the threshold value n_{link}. if less than the threshold value, then check the node timer the time value is greater than the stored T_{shrink} value, if the value is less than, T_{shrink} the node will be variable $n_{reallink}$, delete the ID_{last} and LC_{last} from the query, and then execute the query, nodes in the forwarding operation; if the timer. Moment values greater than the value of T_{shrink} for its preservation, the query request variables in $n_{reallink}$ value is set to 0, will the variables ID_{last} and LC_{last} values were set for the data storage node ID value and the coordinate position. If the $n_{reallink}$ value is larger than the connection node number of threshold value of n_{link}, then the node calculated e value and a value. If the value of $k_2 * D_{A\_B} / R_C$ is greater than $n_{reallink}$, the operation is repeated when the value of $n_{reallink}$ is less than n_{link}; if the value of $k_2 * D_{A\_B} / R_C$ is less than $n_{reallink}$, the node performs information processing of the cloud shrink operation.

If the data storage node on the backbone receive request does not contain variables $n_{reallink}$, ID_{last} and LC_{last} are the data storage nodes to check the timer time value is greater than the stored T_{shrink} value, if the value is less than T_{shrink}, the node only executes query forwarding operations; if the timer value is greater than T_{shrink} the value of the node in the query which added three variables $n_{reallink}$, ID_{last} and LC_{last}. And set the initial value of $n_{reallink}$ is 0, the variable ID_{last} and LC_{last} values respectively for the data storage node ID and location coordinates, then execute the query forwarding operations. When a node receives a connection the query, the query first checks whether there have variable $n_{reallink}$.

if there is, it will be the value of $n_{reallink}$ plus 1, and then forwards the query request; when a connection node also received two contains these three variables in the query, the node two The size of the variable ID_{last} in the query request, the ID_{last} value of the smaller query request in the two variables ID_{last} and $n_{reallink}$ removed, and then forwarded the two query requests.

(3) And changing the node type of judgment

When the outermost node receives the data, the data sending node is informed that the node type is changed to the center node，and re transmitted the previous data to the heart. The innermost node is the nearest neighbor node which is closest to the center of the network, when the data packet is received, the source node of the data packet is notified to change the node type to the centrifugal node. The source node of the data packet changes its node type and then re sends the data packet.

c. Information processing cloud of self-healing

When the information processing in the cloud information processing cloud link node was destroyed or information processing in the cloud contains a region node data storage node or nodes are destroyed when information processing cloud starting the self-repair function. Detection information processing the main link node of the cloud can be monitored by the method to determine the main chains on the way up and down nodes are damaged.

c.i. Information processing cloud self-healing theme ideas

Information processing cloud recovery functions to achieve the main idea is that first found by opening information processing cloud link node generates a message handling cloud message recovery, the message before along the circuit breaker in the area of information processing cloud extending along the direction of propagation until they encounter information by another breakpoint of backbone road, cloud processing. Information processing the main backbone of the main chain of a message to receive information processing to generate a link after the recovery of the message to establish a successful

message, and the message along the original road to return to send.

The following is a detailed introduction of information processing cloud recovery process. That circuit breaker information processing cloud link node generates a message handling cloud message recovery, the message contains a containing node ID and location information cannot be successfully established the connection information processing cloud link node set (obviously, at the start of this collection only one node information).Then the message generation node selects a suitable node to send the message. This suitable node is satisfied:

Its remaining energy is greater than a threshold. Received information cloud processing nodes of the message recovery first check your own node types, see is not the roadmap node, if the node is not of landmark nodes, or the node is of landmark nodes, but corresponding to the centralized data storage nodes, nodes are included in information processing cloud recover the message with a set of nodes, the nodes along to the node and not last successfully established connection information processing cloud link node around the node inverse clockwise to find an energy value is greater than the cloud link node minimum energy threshold of node information processing and information at the cloud will send a message to the node. If the node is the landmark nodes, and the data storage nodes of the corresponding node contains a centralized information processing cloud with message recovery nodes not included (as a set of these nodes is denoted as s), the node checks whether there exists a set s is a node of breakpoint information processing chain cloud the road, if not, the node to the end node along with a collection of s in left and right ends of node distance node a node farthest to the node for the link to another endpoint node around the rotating counterclockwise, find a residual energy greater than the node information processing cloud the lowest threshold link node as its next hop, and will send out a message to restore cloud information processing, If it exists, the node and the node connection is established, and its node type variable node for the data storage and notify its one hop neighbor nodes, then the original road return link message is successfully established, the node received the news of the variable nodes of its type node for the data storage and its one hop neighbor nodes.

Figure6 Damaged area

c.ii The measures to avoid repetition of establishing primary link

When the information is allowed to deal with the existence of multiple main chain paths between any two nodes of the main backbone of the cloud，On the one hand, repeat the establishment of backbone road connection will cause energy waste, another face to face, the data storage nodes in between these two points in data receiving rate is too high to expand the algorithm more complex, increasing the node computation, reduces the efficiency of the nodes, at the same time, also for data query trouble, queries must be to layers of depth to each sub link query, the link results in a link aggregation, increase the query of the energy consumption, increase the query delay. Therefore, we have to ensure that the information processing of the main chain of the main road between any two nodes and there is only one responsible for the data storage and processing of the main link of the main road.

However, information processing cloud link node between breakpoints need link repair and in destroyed region on the edge of the two breakpoints, possibly in breakpoint information processing cloud link repair process is another breakpoint also start information processing cloud link repair program. This will cause backbone cloud road, repeat the establishment of information processing.

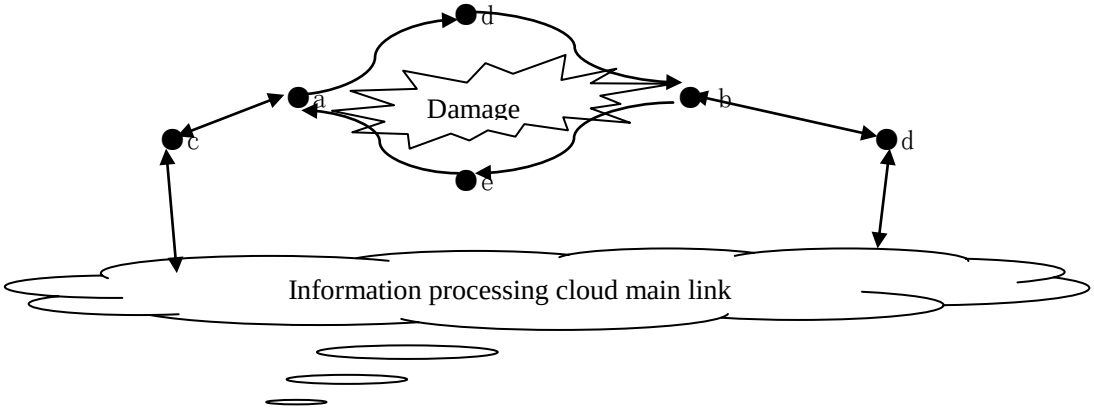

Figure7 the main links of information processing cloud have been restored

As shown in Figure 1,a and b are two breakpoints on both sides of the damage zone, node a first found themselves in the main chain on the way up and down one of the nodes was destroyed, so a boot information processing cloud main link repair program, and when the a will be the main link to the information processing cloud repair message sent to the node D, the node B also find themselves up and down the node in a failure, so b will be the main chain to restore the message sent to the node e, along the other direction of the information processing cloud repair. This will establish the two links between the nodes and the node.

The solution to this problem is to select the A and B two nodes in the node ID the larger the link established by the node, and the removal of the node ID smaller nodes to establish the link(or select a smaller ID link to remove the larger ID link).Assuming that a> b, When the a node receives the main chain path repair message initiated by the b node, the a node generates a link removal message to return to the node of the main chain path repair message initiated by the b node to the a node, the node cancels the main chain repair process initiated by the b node.

V. DATA STORAGE

When the center node is aware of the data generated, the central node uses the GPSR routing protocol to send to the network center location; When the sensor data is generated by the sensor, the node sends the data by the change of the GPSR routing protocol to the center of the network; When the information processing nodes in cloud data, if the node type is a data

storage node, it will store the data locally, if node type is connected to the node, then the node will forward data to the nearby data storage nodes to store, if the node type is of landmark nodes, nodes check their neighbor nodes which are data storage nodes, if there is one or a plurality of data storage node, select a data reception rate smaller data storage node will send data in the past, if not, the data is sent to the connection node.

Old data storage node needs to forward data to a new data storage node, if the data is larger, it can be forwarded to its pointer.

VI. DATA QUERY

Send inquiry initiated radial nodes along the radial direction of the inquiry request; queries centrifugal node sends a query sent by centrifugal direction; query node initiates the cloud can send queries to the primary link node. When a query is sent to the master node link received by the main link in both directions on the first node in the main links of the query request Send request jumping up and down, respectively, in the node on the primary link first, save when receiving a query request originating node location query request and the query, and then went down to hop forwards the query request, and when successful queries forwarded from the node, the node in the local data query, save the query results, corresponding to the query ID and the originating query node location, and delete the corresponding queries. When a node point on the primary link received at the same time from the same query on the main link and down came the hop, the node does not forward queries, but the query ID and its corresponding local query results a direction along the main link forwarding, and forwards the query ID in another direction along the primary link when a node receives a master link containing the query results and query data packet ID, the node queries ID packet corresponding to the query results locally and packet aggregation query results, and will bring together the results sent to the next hop. When a node receives the query results two of the same query on the main link, the node that these two query results and query local query results converge, then the final result of the convergence of the routing protocol to the GPSR initiating node transmits a query.

AKNOWLEDGEMENT

This work is supported by the Key Projects of Science and Technology of Hunan Province (No.2013FJ3095),The National College Students' Innovation and Entrepreneurship Training Program ([2015]41-12824), project supported by the Hunan Province College Students' Research Learning and Innovative Experiment Plan Fund(Xiang Jiao Tong[2015]269-562), The Hunan Province College Research Project of the Teaching Reform (Xiang Jiao Tong[2015]291-562),and the scientific research project of Hunan Institute of Technology(HY13001).

REFERENCES

[1] Le H C, Guyennet H, Zerhouni N. "Mobile effect reduction in data-centric storage for wireless sensor networks", 3rd IET International Conference on Intelligent Environments (IE 07). IET Digital Library, 2007, pp: 304-311.

[2] Shen HY, Li T, Schweiger T. "An Efficient Similarity Searching Scheme Based on Locality Sensitive Hashing", In: Proceedings of Int'l Conf. Database Theory (ICDT 2008). Piscataway: IEEE Press, 2008. pp: 123−128.

[3] Lin Y, Liang B, Li B. "Data Persistence in Large-Scale Sensor Networks with Decentralized Fountain Codes", Proceedings - IEEE INFOCOM, 2007, pp: 1658-1666.

[4] Akyildiz I F, Su W, Sankarasubramaniam Y, et al. "A survey on sensor networks", Communications Magazine IEEE, vol.40, No.8, 2002, pp.102-114.

[5] Yu G J. "Adaptive storage policy switching for wireless sensor networks", Wireless Personal Communications, vol.48, No.3, 2009, pp: 327-346.

[6] Li J, GAO H. "Survey on Sensor Network Research", Journal of Computer Research & Development, vol.1, No.45, 2008, pp.1-15.

[7] Zhang Q, Xie Z P, Ling B, et al. "A Maximum Lifetime Data Gathering Algorithm for Wireless Sensor Networks", Journal of Software, vol.11, No.16, 2005, pp.1946-1957.

[8] Gong H, Liu M, Wang X, et al. EADEEG: "An Energy-Aware Data Gathering Protocol for Wireless Sensor Networks", Journal of Software, vol.5, No.18, 2007.pp:1092-1109.

[9] Liang J B, Wang J X, Tao-Shen L I, et al. “Maximum Lifetime Algorithm for Precise Data Gathering Based on Tree in Wireless Sensor Networks”, Journal of Software, vol.9, No.21, 2010. pp: 2289-2303.

[10] X. Li, Y. J. Kim, R. Goninan and W. Hong. “Multi-dimensional Range Queries in Sensor Networks”, International Conference on Embedded Networked Sensor Systems. 2004, pp: 63-75.

[11] L. Xie, L. J. Chen, D. X. Chen. “Ring-Based multi-resolution data storage for sensor networks”, Journal of Software, vol.12, no.20, 2009, pp: 3163-3178.

[12] J. L. Xu, X. Y. Tang, and W. C. Lee, “A New Storage Scheme for Approximate Location Queries in Object-Tracking Sensor Networks”, IEEE Transactions on Parallel & Distributed Systems, vol.2, no.19, 2008, pp: 262-275.

[13] C. Y. Ai, R. Y. Du, M. H. Zhang and Y. S. Li, “In-Network Historical Data Storage and Query Processing Based on Distributed Indexing Techniques in Wireless Sensor Networks”, Wireless Algorithms, Systems, and Applications, International Conference, WASA 2009, Boston, Ma, USA, August 16-18, 2009, pp:264-273.

[14] L. Q. Pan, J. Z. Li, J. Z. Luo. “Approximate Skyline Query Processing Algorithm in wireless sensor networks”. Journal of Software, vol.5, no.21, 2010, pp: 1020-1030.

[15] D. Braginsky and D. Estrin. “Rumor routing algorithm for sensor networks”, First ACM International workshop on wireless sensor networks & applications, 2002, pp: 22-31.

[16] X. Liu, Q. F. Huang and Y. Zhang, “Balancing Push and Pull for Efficient Information Discovery in Large-Scale Sensor Networks”, IEEE Transactions on Mobile Computing, vol.3, no.6, 2007, pp: 241-251.

[17] Z. C. Yu, B. Xiao and S. G. “Zhou Achieving optimal data storage position in wireless sensor networks”, Computer Communications, vol.1, No.33, 2010, pp: 92-102.

[18] Xu J, Qian H, Ying W, et al. “A deployment algorithm for mobile wireless sensor networks based on the electrostatic field theory”, The International Journal on Smart Sensing and Intelligent Systems, Vol.8, No.1, 2015, pp. 516-537.

[19] Bai, Q., & Jin, C, “Image fusion and recognition based on compressed sensing theory”, International Journal on Smart Sensing & Intelligent Systems, Vol.8, No.1, 2015, pp. 159-180.

[20] Qiao J, Liu S, Qi X, et al, “Transmission power control in wireless sensor networks under the minimum connected average node degree constraint”, The International Journal on Smart Sensing and Intelligent Systems, Vol.8, No. 1, 2015, pp.801-821.