Efficient Data Search System for Large-Scale Mobile Wireless Networks using Secured and Centralized Routing Protocol

Sulu George
Department of Computer Science and Engineering
KMEA Engineering College
Kerala, India

Arifa Azeez
Department of Information Technology
KMEA Engineering College
Kerala, India

Abstract— ***This paper particularly addresses the data search problem in large-scale wireless networks with high mobility and density. Due to this, in a large scale network it is difficult to fetch the data from the nodes into a wireless network. Because of high overhead and difficulty of data discovery, the nodes will fail to fetch the data flexibly from the nodes and they cannot give the query service as we expect. Locality based distributed Data search system (LORD) is been used in the paper in order to overcome this problem. To increase the security and efficiency, we use Efficient Hybrid Routing protocol which reduces the complexity by minimizing the transmission time. Also backtracking algorithm is used for updating when one node moves into another region. The data will be published to a particular zone in a large network. When a user receives a metadata query from the user node, the file will be fetched directly from the particular centralized server and retrieves the data to the user.***

Index Terms— ***Scalability, metadata publishing, file querying, Flexibility, Reliability, transmission time.***

1 INTRODUCTION

A mobile ad hoc network is an infrastructure-less network with wireless nodes moving from one location to another. Instead of using a central base station to which all the computer must be connected and communicate, it identifies a particular network where every mobile node acts as base station and can move around. Due to the mobile nature of nodes, frequent updates are required to maintain connectivity information among mobile nodes in the network. In other words, it is a self-configuring network. Typical applications for these networks include: information exchange in disaster situations, establishing connectivity in situations with networks on the go. To simulate a mobile ad hoc network, several properties need to be modeled, e.g., mobility of mobile nodes. Recent technologies have enabled the development of a large-scale wireless network (wireless sensor network) and mobile ad hoc network consisting of a huge number of mobile nodes isolated over extensive area. An important issue in such wireless networks is data search. This paper particularly addresses the data search problem in large-scale wireless networks with high mobility and density. WSN are used in variety of applications such as military sensing and tracking, habitat monitoring, health monitoring, environmental pollutant discovery, and wildfire tracking. In a WSN, sensors coordinate to perform distributed sensing of environmental phenomena, and collect and share widely-scattered distributed data in a supportive mode, which makes data search serious to WSNs.

Considering the theatrical growth of mobile devices and the restrictions of wired communication, mobile data search applications that enable ubiquitous data access wherever will proliferate in the near future. It is envisioned that there will be omnipresent wireless devices, and some urban areas will be densely covered by ubiquitous mobile nodes. In a cellular network, each cell uses a different set of frequencies from neighboring cells, to avoid interference and provide guaranteed bandwidth within each cell. When joined together these cells provide radio coverage over a wide geographic area. This enables a large number of portable transceivers to communicate with each other and with fixed transceivers and telephones anywhere in the network, via base stations, even if some of the transceivers are moving through more than one cell during transmission. Therefore, an efficient data search system for a highly mobile and dense wireless network is needed. However, current wireless network data search systems are not suitable for such an environment. It increases the burden to the network.

2 RELATED WORKS

The existing systems are not efficient due to many disadvantages. The flooding mechanism and local-broadcasting methods in WSNs and in MANETs are not energy efficient because of transmitting a high amount of messages. Local-broadcasting also cannot guarantee data discovery. In the topological routing based method in MANETs nodes advertise their obtainable data, build content tables for received advertisements, and forward data requests to the nodes with high probability of holding the data. However, this method causes high overhead for advertising and table maintenance. Also, it cannot guarantee data discovery because of possible expired routes in the content tables owing to node mobility. Geographic routing based data

search systems have been proposed for WSNs for high scalability. Specifically, a file is mapped to a geographic location based on the distributed hash table (DHT) data mapping policy (a file is mapped to a location whose ID is closest to the file's ID), and stored in a node nearest to the geographic location using geographic routing. To search a file, a requester computes the mapped geographic location and uses geographic routing to transmit the query to the location. However, in a highly mobile network, a file needs to be frequently transferred to its new mapped file holder, which produces high overhead. Also, a delayed data mapping update may lead to a querying failure. In addition, geographic routing needs exact geographic node localization (e.g., x, y) using GPS or virtual coordinates which exacerbates overhead burden and increases energy consumption. GPS consumes nodes precious energy resources and may not provide location information in some situations (e.g. indoors). The virtual coordinate methods need periodic coordinator updates, which produces high overhead in a highly mobile network. All these leads to find an efficient data search in terms of scalability, overhead, and mobility resilience in a highly dense and mobile network.

3 PROPOSED LOCALITY-BASED DATA SEARCH SYSTEM DESIGN

To build a scalable and reliable data search system for large-scale highly mobile and dense wireless networks, we propose a LOcality-based distRibuted Data search system (LORD). LORD divides the entire wireless network area into a number of geographic regions. The metadata1 of a file is mapped to a region using the DHT data mapping policy, and it is stored in all or a subset of the nodes in the region, thus improving mobility resilience. A node needs to update data mapping only when it moves across regions, thus minimizing maintenance overhead. LORD has a novel centralized routing protocol for data publishing and querying. After a node retrieves its queried metadata, it requests different file segments from physically closest file holders. It calculates the sizes of segments requested from different file holders based on their distances to minimize the file fetching latency. After retrieving the file, the requester publishes the metadata of the file to its mapped regions. To increase the efficiency, we use Centralized Routing protocol which reduces the complexity to find angle of arrival information which is needed for the existing method region based geographic routing protocol. The central servers are storing the information such as received metadata by the subset of nodes. Thus, it can reduce overhead and traffic in the network. Also backtracking algorithm is used for updating when one node moves into another region. That is, to avoid packet dropping.

A parallel file fetching algorithm is used to determine numerous physically close file servers to send various file segments to reduce file retrieval latency. Extensive experimental results demonstrate the superior performance of LORD in comparison to previous data search systems and the efficiency of LORD components.

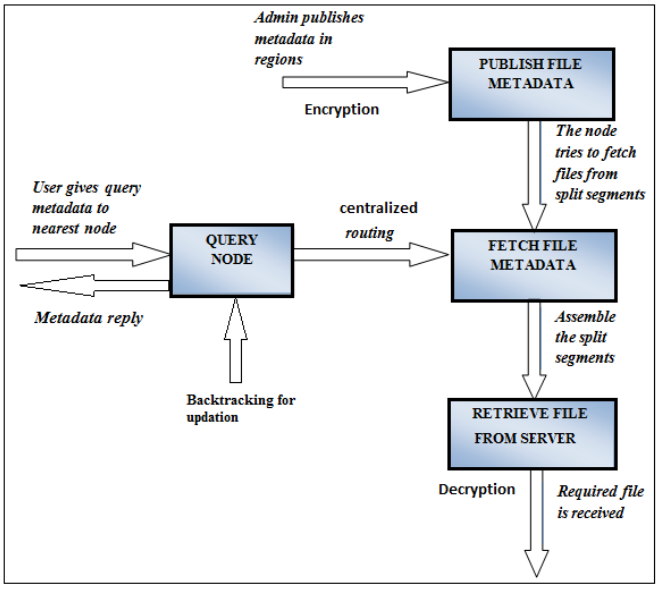

Fig: LORD Architecture

Fig 3.1 shows the architecture of LORD. LORD is also characterized by metadata storage instead of data storage. Most current wireless data search systems use data storage. Admittedly, data storage methods avoid an additional file querying step after locating the file hosts in the metadata storage methods. However, in a highly mobile and dense wireless network, metadata storage has the advantage in terms of overhead for data mapping updates determined by message size. High node mobility leads to frequent mapped data (metadata or files) transfer. Metadata storage only produces additional operations of small-size metadata querying and replying.

3.1 AREA PARTITION

A highly mobile and dense wireless network with nodes spreading over an area and are independently and identically distributed. LORD is proposed for a wireless network with a number of landmarks. Considering the promising ubiquitous computing environment in the future, such static landmarks (e.g., base stations, WIFI access points) will not be difficult to find. Once the landmarks are determined, LORD divides the entire area into a number of regions. A region is the neighboring zone in the transmission range of a landmark and centered by the landmark. Each region is identified by an assigned integer ID. To make LORD adaptive to general case, the regions can be any shape. The number of landmarks (regions) can be determined based on the transmission range of the nodes and the size of the entire area. Here, it is focused on a certain area, such as a campus, a habitat monitoring area or a wildfire tracking area.

3.2 METADATA PUBLISHING AND FILE QUERYING

Locality sensitive hash function hashes two similar keyword groups to close values with high probability. LORD store the metadata of similar files into the same region for similarity search. A file's keywords can be its file name or the keywords

retrieved using information retrieval algorithms. The number of LSH hash values of a file can be one or more than one based on the settings of LSH.A file's keywords can be its file name or the keywords retrieved using information retrieval algorithms. A file host publishes the metadata to the mapped regions. The node in a destination region that firstly receives the metadata broadcasts it to all other nodes in the region. When a mobile node wishes to query a file, it calculates the coordinate, of the file's metadata and uses centralized routing protocol to send metadata and file. Those are forwarded to the destination regions, which are exactly the regions that hold the metadata of the queried file. If the first query receiver in the destination region is lightly loaded, it responds to the requester. Otherwise, the query is forwarded to a randomly selected neighbor continuously until reaching a lightly loaded node in the region, which will respond to the requester. The requester can specify a similarity threshold. The similarity between the keywords of a file and the queried keywords is calculated. The query receiver responds to the requester with the metadata that has a similarity to the queried keywords greater than the required threshold. All receiving data is transferred and stored in the corresponding server and makes the node free.

3.3 PARALLEL FILE FETCHING

In Parallel File Fetching Algorithm, after receiving the metadata of its queried file, a requester can retrieve the region IDs of the file's holders. It then places the file holders in the region map initially configured to itself. To reduce file fetching latency, LORD uses a parallel transmission algorithm, in which different file segments are simultaneously transmitted from different file holders to the file requester. Since each segment has a shorter data stream than the whole file, the total time period for transmitting all segments to the file requester is shorter than transmitting the whole file from one file holder. Specifically, the file requester chooses geographically close file holders among the located ones, and asks each file holder to transmit a segment of the file. Different segments destined to the same destination may arrive at the same node in routing. Then, this node can merge these segments before forwarding them out to save energy for forwarding.

3.4 BACKTRACKING ALGORITHM

Back-Tracking Algorithm, a data requester integrates the ID of its region (i.e. source region) into its request when querying for metadata or data. The required metadata or data will be sent back to the requester. In a highly mobile wireless network, the requester may move out of its region or even pass through a number of regions before the reply arrives at the source region. LORD has a backtracking algorithm to keep track of the requester's movement. In the algorithm, if a requester moves out of its current region before receiving the response, it sends a back-tracking message (including its current region) to the source region. The message is piggybacked on the hello messages between neighbor nodes. After reaching another region, node will drop all its old data and obtain those data from its new neighbor. Thus, each node in the source region keeps a back-tracking message of the requester. Using this message, the response can be forwarded to the requester that moves out of the source region.

3.5 SECURED AND CENTRALIZED ROUTING PROTOCOL

Security is the one of the biggest concern in different type of networks. Due to diversity nature of network, security breaching became a common issue in different form of networks. Solutions for network security comes with concepts like cryptography in which distribution of keys have been done. If you want to send data to some other persons through network then if you truly want to keep the information secret, you need to agree on some sort of key that you and he can use to encode/decode messages. But no need of keep using the same key, it makes easier and easier for others to crack your cipher. Moreover the complete data will be encrypted before the transmission takes place. The RSA algorithm is used for the encryption and the decryption purpose. The encrypted data will then piggybacked by which the data is hided and then moved to the destinations. Through this the congestion on the network can be controlled. RSA is one of the first practical public-key cryptosystems and is widely used for secure data transmission. In such a cryptosystem, the encryption key is public and differs from the decryption key which is kept secret.

For efficient routing protocol, centralized servers are used in each region so that it can store all kind of information about its subset of nodes. If a node from one region transmit the metadata and file to the requestor in another region, the node which receiving information stores in the centralized server for that particular region. So that whenever any neighbor node needs the data, it can obtain data from its neighbor node. Even if it is not in the neighbor node it can get it from the central server. Thereby, traffic in the regions can be reduced. Each node can maximize its storage space so that it can act as lightly loaded buffers. Thus nodes can receive more number of files at any time. This will be more accurate when compared to the geographic region based algorithm. Transmission time for receiving the information can also be minimized. It simplify the design, increases its reliability, and gain flexibility. It makes much easier to implement and replicas can be reduced.

4 CONCLUSION

There is a great development of smart phones, laptops and huge growth in the usage of internet nowadays. It consist large number of nodes which are spread over a wide area. It has been a very hectic for the nodes to store the files. Due to the heavy storage, the data is not able to be fetched so easily and reliably. Data discovery is not possible so data searching is becoming a crucial problem to the wireless network. So there is a requirement in an advancement of technology to overcome this problem. This paper introduces LORD system, which helps in publishing the file easily among the nodes and retrieves the file exactly from that node. This merely reduces

the data overhead and increases the flexibility, makes the network more reliable. This works well in a large scale network which increases the scalability by retrieving the files in an easy manner. It provides accurate data and less transmission time.

REFERENCES

[1] Haiying Shen, Ze Li and Kang Chen, "A Scalable and Mobility-Resilient Data Search System for Large-Scale Mobile Wireless Networks", IEEE transactions on parallel and distributed systems, vol. 25, no. 5, may 2014.

[2] S. Guo, Y. Gu, B. Jiang, and T. He, ''Opportunistic Flooding in Low-Duty-Cycle Wireless Sensor Networks With Unreliable Links,'' in Proc. MobiCom, 2009, pp. 133-144.

[3] K.-W. Fan, S. Liu, and P. Sinha, ''Dynamic Forwarding Over Tree-on-Dag for Scalable Data Aggregation in Sensor Networks,'' IEEE Trans. Mobile Comput., vol. 7, no. 10, pp. 1271-1284, Oct. 2008.

[4] S. Schnitzer, H. Miranda, and B. Koldehofe, ''Content Routing Algorithms to Support Publish/Subscribe in Mobile Ad Hoc Networks,'' in Proc. IEEE LCN Workshops, 2012, pp. 1053-1060.

[5] N. Shah and D. Qian, ''An Efficient Unstructured P2P Overlay Over MANET Using Underlying Proactive Routing,'' in Proc. MSN, 2011, pp. 248-255.

[6] P. Sharma, D. Souza, E. Fiore, J. Gottschalk, and D. Marquis, ''A Case for MANET-Aware Content Centric Networking of Smartphones,'' in Proc. IEEE WOWMOM, 2012, pp. 1-6.

[7] C. Tien Ee and S. Ratnasamy, ''Practical Data-Centric Storage,'' in Proc. NSDI, 2006, pp. 325-338.

[8] J. Xu, X. Tang, and W.-C. Lee, ''A New Storage Scheme for Approximate Location Queries in Object Tracking Sensor Networks,'' IEEE Trans. Parallel Distrib. Syst., vol. 19, no. 2, pp. 262-275, Feb. 2008.

[9] M. Demirbas, X. Lu, and P. Singla, "An in-Network Querying Framework for Wireless Sensor Networks,'' IEEE Trans. Parallel Distrib. Syst., vol. 20, no. 8, pp. 1202-1215, Aug. 2009.