ISSN 2349-4506
Impact Factor: 2.265

Global Journal of Engineering Science and Research Management

PUBLISHING AND QUERYING DATA SEARCH SYSTEM FOR LARGE-SCALE MOBILE WIRELESS NETWORKS USING EFFICIENT HYBRID PROTOCOL

Sulu George*[1] Arifa Azeez[2]
*[1]Department of Computer Science and Engineering, KMEA Engineering College, India.
[2]Department of Information Technology, KMEA Engineering College, India.

KEYWORDS: Scalability, Data Publishing, Querying, Flexibility, Reliability.

ABSTRACT
This paper mainly concentrates on a data search problem where there is a huge usage of the internet now a days. 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, 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 efficiency, we use Efficient Hybrid Routing protocol which reduces the complexity to find angle of arrival information which is needed for the existing method region based geographic routing protocol. Also piggybacking algorithm is used instead of backtracking algorithm 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 zone and retrieves the data to the user.

INTRODUCTION
Mobile wireless sensor networks can simply be defined as a wireless sensor network (WSN) in which the sensor nodes are mobile. They are a smaller, emerging field of research in contrast to their well-established predecessor. They are much more adaptable than static sensor networks as they can be deployed in any situation or state and cope with rapid topology changes. A mobile ad hoc network (MANET) is a continuously self-configuring, infrastructure-less network of mobile devices attached without wires. Each device in it is free to move independently in any direction, and will therefore change its links to other devices frequently. 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. Also, considering the theatrical growth of mobile devices (e.g., laptops, smart phones, and communication-enabled vehicles) 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 (e.g., WiFi enabled cabs in the Manhattan Area). 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 (e.g., mobile phones, pagers, etc.) 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.

MATERIALS AND METHODS
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

ISSN 2349-4506
Impact Factor: 2.265
Global Journal of Engineering Science and Research Management

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.

Metadata Publishing and Querying
Locality sensitive hash function (LSH) hashes two similar keyword groups to close values with high probability. LORD uses LSH to hash a file to 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 using the Hybrid Routing protocol. 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 Hybrid Routing protocol to send requests with as the destinations. The requests 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.

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.

Piggybacking 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 based on the Hybrid Routing protocol. 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. 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.

Coloring-Based Partial Replication
In Coloring-Based Partial Replication, storing a metafile in every node in a region enables mobility-resilient and fast file retrieval but produces a high overhead for node storage, data mapping updates, and location updates. To handle this problem, a coloring-based partial replication algorithm is proposed. The coloring policy in graph theory aims to prohibit two neighboring nodes in a graph from having the same color. Stimulated by this idea, the coloring-based partial replication algorithm aims to guarantee that a node has at least one neighbor holding a metafile while avoiding having the metafile in neighboring nodes.

In coloring-based partial replication algorithm, when a node in a region receives the first metadata of the region, it stores the metadata and broadcasts it along with a flip-flop key with a preliminary value of zero and a TTL (Time to Live). Because of high node mobility, the neighboring relationships between the nodes in a region always vary. To manage the coloring status, each node in a region needs to periodically verify its neighbors to ensure that it can retrieve the metafile within one hop. Specifically, each node appends a flag bit in the periodic hello message

ISSN 2349-4506
Impact Factor: 2.265
Global Journal of Engineering Science and Research Management

to specify whether it is a replica node. Through the hello messages, each node periodically checks whether one of its neighbors is a replica node and records the neighbor's ID. When a non-replica node observes that none of its neighbors has a replica, it sends a metafile request with a TTL to a arbitrarily chosen neighbor. The request is forwarded until meeting a replica node, which sends a metafile to the requester along the original path. If the requester does not receive a response during TTL, it assumes that there is no metafile in its region and its subsequent received metadata is the first metadata in its region.

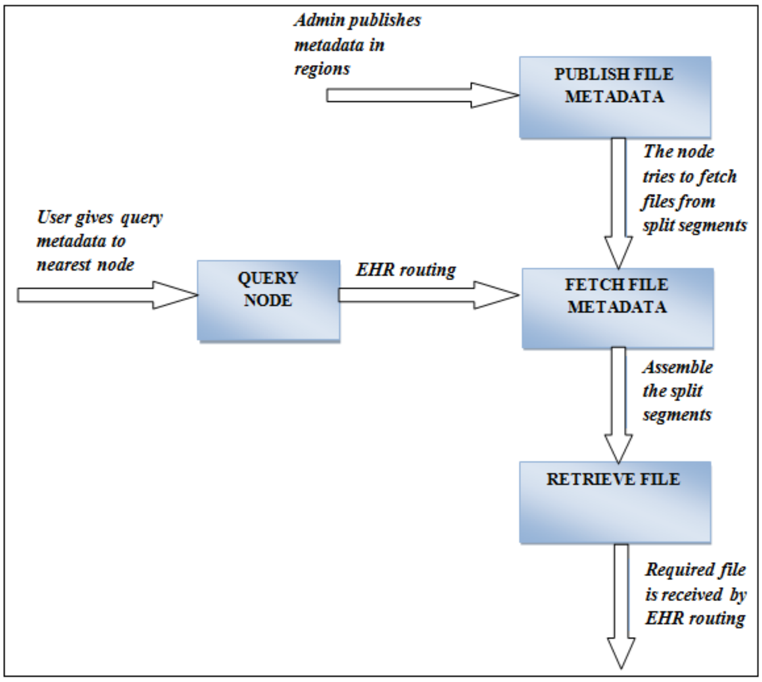
When a node retrieves a data for storing, deleting, or updating a file's metadata, if it is a replica node, it conducts the operation accordingly and forwards the message to its neighbors. If it is a non-replica node, it directly forwards the message to its neighbors. In this manner, all replicas in the region are updated. When a node receives a metadata query, if it is a replica node and is lightly loaded, it responds with the queried metadata. Otherwise, it forwards the query to the neighbor that has the replica. The coloring-based partial replication algorithm also minimizes the overhead due to intra-region mobility. When a node shifts into a new region, without the algorithm, it necessitates getting hold of metadata and drops its old metadata. With the algorithm, only when the node is a replica node, it needs to shift its metadata to its old neighbor; if the node has a neighbor with metadata in the new region, it does not need to acquire metadata.

Efficient Hybrid Routing Algorithm

Hybrid Routing Protocol (HRP) is a network routing protocol that combines Distance Vector Routing Protocol (DVRP) and Link State Routing Protocol (LSRP) features. HRP is used to determine optimal network destination routes and report network topology data modifications. Hybrid Routing, commonly referred to as balanced-hybrid routing, is a combination of *distance-vector routing*, which works by sharing its knowledge of the entire network with its neighbors and *link-state routing* which works by having the routers tell every router on the network about its closest neighbors. Hybrid Routing is a third classification of routing algorithm. Hybrid routing protocols use distance-vectors for more accurate metrics to determine the best paths to destination networks, and report routing information only when there is a change in the topology of the network. Hybrid routing allows for rapid convergence but requires less processing power and memory as compared to link-state routing. HRP features are: requires less memory and processing power than LSRP, integrates reactive and proactive routing advantages.

Hybrid routing algorithms aim to use advantages of table driven and on demand algorithms and minimize their disadvantages. Position based routing algorithms that are classified in the hybrid routing algorithms category include the properties of table driven and on demand protocols and are usually interested in localized nodes. Localization is realized by GPS that is used to determine geographical positions of nodes. Position changes which occur because of nodes mobility in MANET cause changes in routing tables of nodes. The GPSs, which are embedded in nodes, are used to update information in tables in position-based algorithms. That makes position-based algorithms different from the table driven and on demand algorithms.

Global Journal of Engineering Science and Research Management

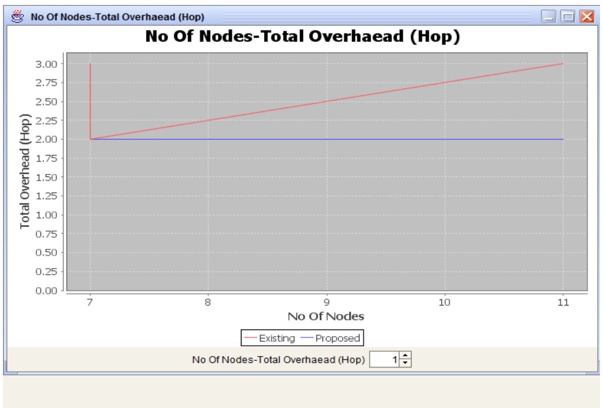


LORD Architecture

RESULTS AND DISCUSSION

Simulation results shows the superior performance of LORD compared to representative data search systems in terms of scalability, overhead, and mobility resilience in a highly dense and mobile network. The results also show the high scalability and mobility-resilience of LORD in an unbalanced wireless network with sparse regions, and the effectiveness of its coloring-based partial replication algorithm. This section includes the comparison of the performance of the existing and the proposed scheme. In the simulation, we divided the area into regions by using landmarks and set the signal range of each landmark is to. All nodes move according to the movement pattern in with 0 pause time. That is, each node randomly selects a region as its destination and moves to the region, and then it randomly selects another region as the destination and moves towards that region. We randomly selected the nodes every second to query for randomly selected files. In the experiments, every message was transmitted once without retransmission. It is reported that the results are within 95% confidence interval.

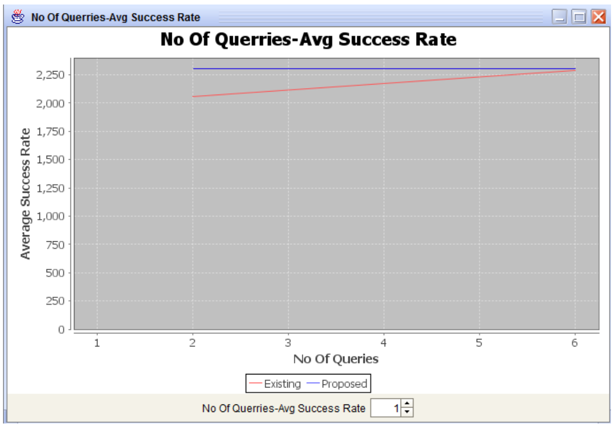
The parameters for evaluation are: Overhead and success rate. The graph shows that the overhead for the proposed system is less than the existing system. The success rate graph shows that the success rate is higher for the proposed system more than the existing system.



Total overhead comparison graph

Global Journal of Engineering Science and Research Management

Success rate and overhead is evaluated and comparisons are shown by the graphs. Success rate is comparatively higher in proposed system than in existing system. Red line indicates existing system and blue line indicates proposed system. Additional nodes can be added after the first stage of file query and repeat the process. Continue the transmission and comparisons can be made.



Average success rate comparison graph

CONCLUSION

There is a huge growth in the usage of internet nowadays. 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. 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.

REFERENCES

1. Haiying Shen, Ze Li and Kang Chen, "A Scalable and Mobility-Resilient Data Search System for Large-Sc ale 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.