International Journal of Computer Application (2250-1797)
Volume 7– No.1, January– February 2017

DESIGN AND ANALYSIS OF A P2P FILE ACCESS CONVENIENCE IN MOBILE UNEXPECTED NETWORKS ALTHOUGH REPLICATION FOR ECONOMICAL FILE SHARING

ADIREDDI RATNAM [#1] , K.SATYANARAYANA MURTHY [#2]

[#1] M.Tech Scholar,Department of Computer Science and Technology,
Baba Institute of Technology and Sciences,Visakhapatnam,AP,India.

[#2] Assistant Professor,Department of Computer Science and Engineering,
Baba Institute of Technology and Sciences,Visakhapatnam,AP,India.

ABSTRACT

A wireless sensor network (WSN) is defined as a collection of autonomous sensors, which are mainly used to monitor physical or environmental conditions, such as climate, terrestrial hazards, temperature, sound, pressure, etc.Now a day’s these WSN or MANETs have attracted more and more attention of users in communicating between each other for data transfer but fails in some properties. As we all know that WSN or MANETs don’t require the fixed topology for communication and due to its node mobility and limited communication range, it is always very difficult for a node to create a replica in the network.However, despite the efforts on file replication, there was no research done or focused on the concept like global optimal replica creation with a minimum average querying delay. In the current various replication protocols that was available in MANETs, there was some or other limitations that takes place during the replica generation. The first limitation what the primitive replication protocols lacks is they were unable to allocate limited resources to different files that are available in the network in order to reduce the average querying delay within the network. Next they always neglect the fact that file holders frequency to meet the other nodes plays a very important role rather than simply considering the available storage resources. In this paper we for the first time have implemented a novel concept of resource for file replication and also derive the advantage of node storage and also the meeting frequency for resource allocation in a distributed environment. By conducting various experiments on our proposed model, our simulation results clearly tells that this proposed mechanism with real traces can achieve shorter average querying delay at a lower cost than various baseline replication protocols.

Key Words: Distributed Networks, File Replication, Resource Allocation, Autonomous Sensors.

I. INTRODUCTION

In current days as there was a lot of development in electronic technology, which laid the way for the development of a new generation of wireless sensor networks (WSNs) consisting of a large number of low-power, low-cost sensor nodes that communicate remotely [1]. Such sensor networks can be used in a wide range of applications, such as for information sensing in military and tracking the information over wireless sensors in a secure manner, health monitoring system, data acquisition in times of hazardous environments, and habitat monitoring [2]. All the sensed data which is extracted from various sources should be sent back to the base station for analysis of that extracted data. However, when the sensing field is too far from the base station, transmitting the data over long distances using multihop may weaken the security strength (e.g., some intermediate may modify the data passing by, capturing sensor nodes, launching a wormhole attack [3], a Sybil attack [4], selective forwarding [5], [6], sinkhole [7]), and increasing the energy consumption at nodes near the base station, reducing the lifetime of the network. Therefore, mobile sinks (MSs) (or mobile soldiers, mobile sensor nodes) are essential components in the operation of many sensor network applications, including data collection in hazardous environments , localized reprogramming, oceanographic data collection, and military navigation [8]- [11].

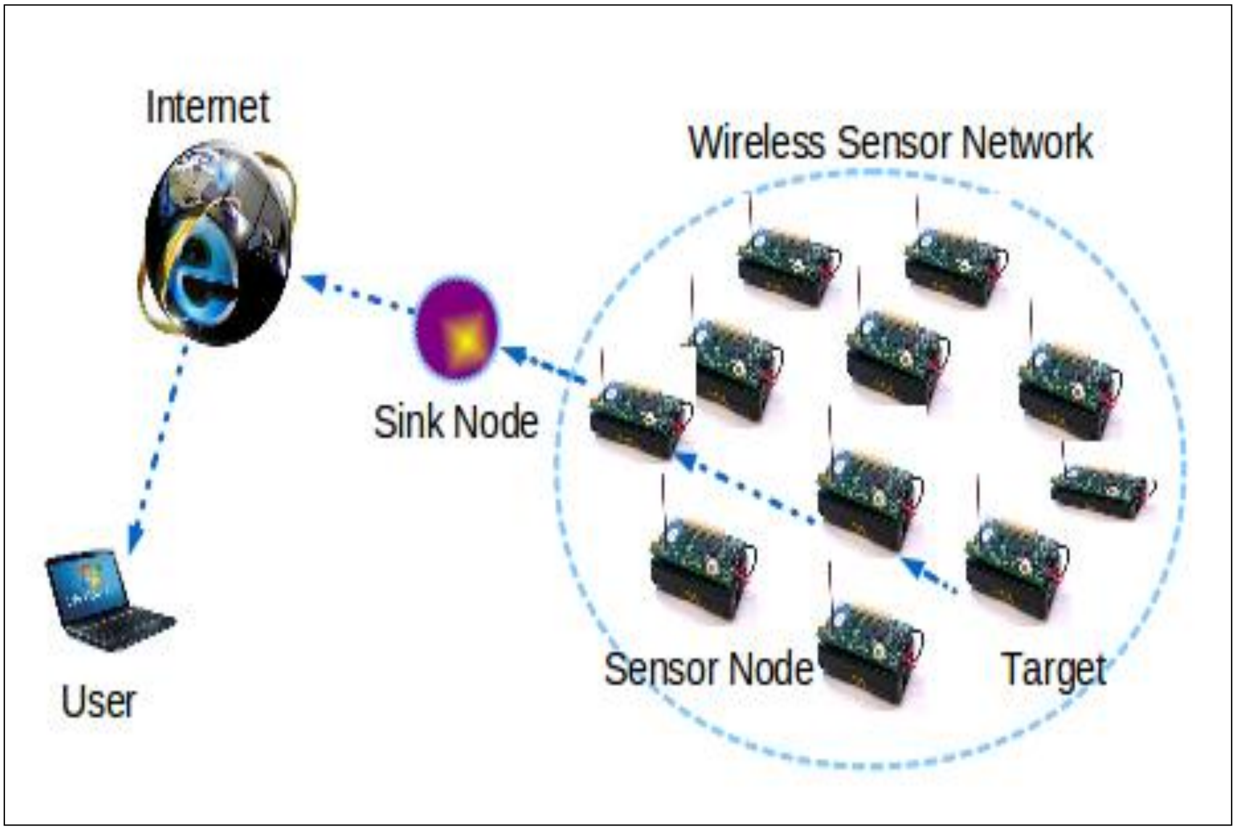


Figure 1. Represents the Architecture of a Typical Wireless Sensor Network or Mobile Adhoc Networks

From the above figure 1, we can clearly get an idea about the typical wireless sensor network or mobile adhoc networks mainly consists of a set of sensor nodes with in a region and a sink node which is connected in order to carry the information from the user node to the target node with the internet. Generally in the WSN applications almost sensor nodes transmit critical information over the network; therefore, security services, such as, authentication and pair wise key establishment between sensor nodes and mobile sinks, are important. However, the resource constraints of the sensors and their nature of communication over a wireless medium make data confidentiality and integrity a nontrivial task. Traditional schemes in ad hoc networks using asymmetric keys are expensive due of their storage and computation cost. These limitations make key pre-distribution schemes the tools of choice to provide low cost, secure communication between sensor nodes and mobile sinks.

Replication is an effective method introduced in order to reduce the file transfer time and bandwidth consumption in Data Grids—placing most accessed data at the right locations can greatly improve the performance of data access from a user's perspective. Now a days we all know that a lot of data intensive applications are evolved, which mainly aim to answer some of the most fundamental questions which are faced by the network users and which is very prevalent topics in the present engineering domains. Examples of such a data replication mainly includes a topic like human genome mapping system [12], and another major concept like astronomy [10], [13], and climate change modeling [14]. In such applications, large amounts of data sets are generated, accessed, and analyzed by scientists worldwide. The Data Grid is an enabling technology for data intensive applications. It is composed of hundreds of geographically distributed computation, storage, and networking resources to facilitate data sharing and management in data intensive applications. One distinct feature of Data Grids is that they produce and manage very large amount of data sets, in the order of terabytes and petabytes.

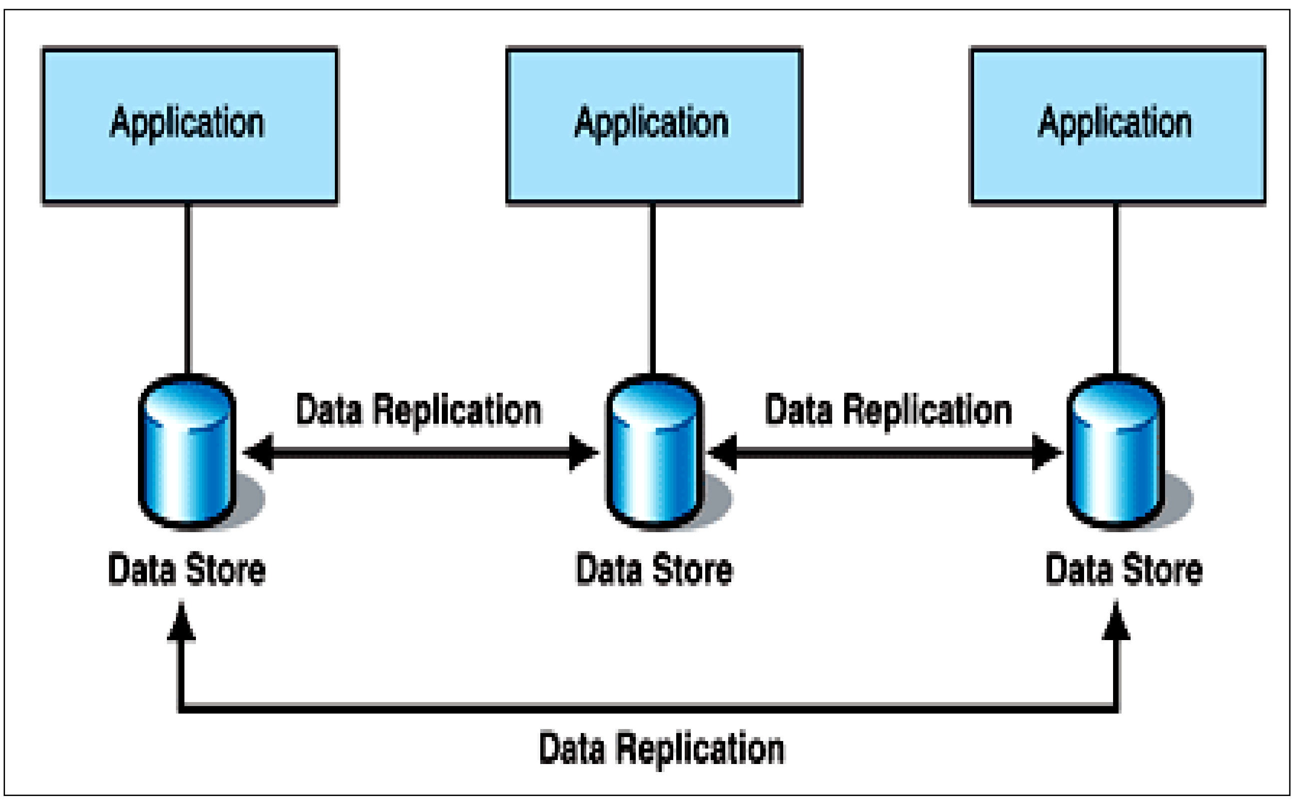


Figure 2. Represents the Process of Data Replication in Maintaining the Storage of Multiple Copies of Same Data

From the above figure 2, we can clearly find out that instead of sharing a single instance of a database between various applications, we can make multiple copies of the database so that each application has its own dedicated store. In order to keep these copies synchronized, we need copy data from one data store to the other data store. This type of approach is most common with packaged applications because it is not intrusive. Here from the above figure we can clearly find out that for each and every application there will be a single data store that will store all the information of that application into it and if the same data store need to be access for other applications also within the same server ,we will try to replicate the data store of one application with replication of other application.

II. RELATED WORK

In this section we will find the background /related work that was analyzed and studied in order to implement this current paper. This section will describe the work that is related to manet's and replication.

MAIN MOTIVATION

As we all know that there was a huge increase in popularity of mobile devices e.g. Smart phones and other hand held devices including Palmtops, IPods and Laptops. As these were increasing its usage and sale their demand also gradually increases in and around our location. One among the best mobile devices is MANETs, where the term MANETs indicates Mobile Adhoc NETworks.This is again classified into two types like normal MANETs which is connection oriented and disconnected MANETs which are having problems during continuous connection also termed as delay tolerant networks (DTNs).

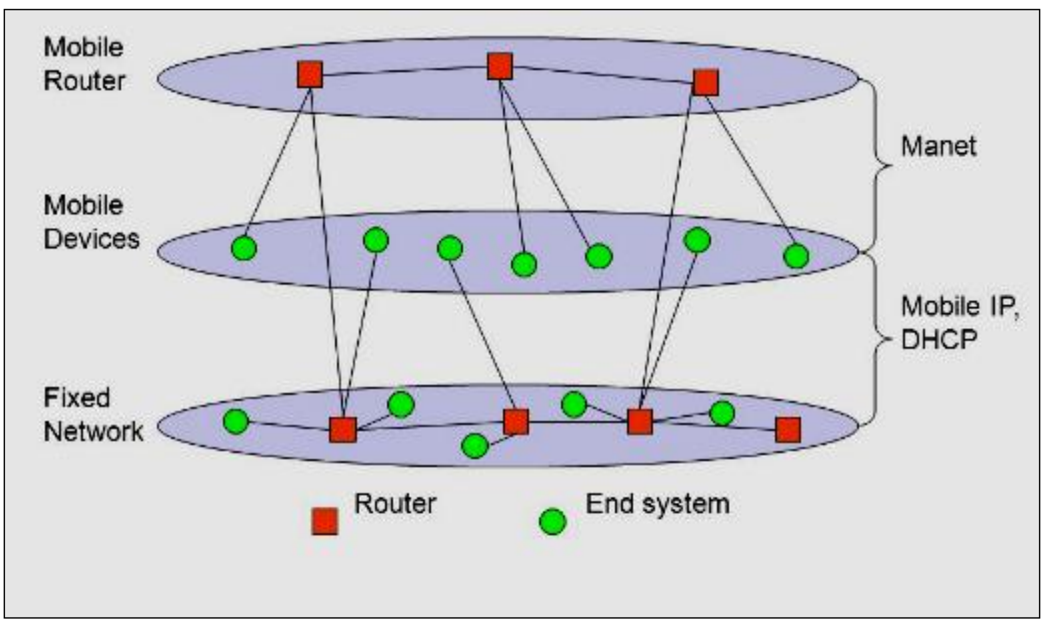


Figure 3. Represents the Architecture of a Mobile Adhoc Network with a Set of Mobile Devices and Router

If we go with the former MANETs architecture it has relatively dense node distribution in its area and while the later is designed as a sparsely distributed nodes with opportunistic in nature.As this mobile adhoc networks were emerged with a new file sharing features ,this gave rise to development of a new framework like peer-to-peer to be designed over MANETs. In the creation of peer –to-peer, the local P2P file sharing model is initially designed and it has the following 3 advantages compared with other type of communication networks, now let us look at those advantages in detail as follows: First the proposed local P2P enables the file sharing facility without having the base station readily available. (Eg: In rural Areas).From the above figure 3,we can clearly get an idea that MANETs contain a set of mobile devices all connected with a set of single or multiple routers that are available within the network, so that all the end systems are connected with the centralized router for communication.

ABOUT DATA REPLICATION

Replication is one of the best active research topic for many years in WWW[10], peer-to-peer networks [11] and mesh networks [14].Through MANETs we can able to communicate with one another for the enormous scientific data and complex scientific applications call for new replication algorithms, which have attracted much research recently. The authors study data replication on Data Grids as a static optimization problem. They show that this problem is NP-hard and non approximable, which means that there is no polynomial algorithm that provides an approximation solution if $P \neq NP$. The authors discuss two solutions: Integer programming and simplifications. They only consider static data replication for the purpose of formal analysis. The limitation of the static approach is that the replication cannot adjust to the dynamically changing user access pattern. Furthermore, their centralized integer programming technique cannot be easily implemented in a distributed Data Grid. Moreover, Baev et al. [16] show that if all the data have uniform size, then this problem is indeed approximable. And they find 20.5-approximation and 10-approximation algorithms. However, their approach, which is based on rounding an optimal solution to the linear programming relaxation of the problem, cannot be easily implemented in a distributed way. In this work, we follow the same direction (i.e., uniform data size), but design a polynomial time approximation algorithm, which can also be easily implemented in a distributed environment like Data Grids.

A Well known author like Raicu et al. [17], [18] studied and analyzed both theoretically and empirically the resource allocation in data intensive applications. They propose a "data diffusion" approach that acquires computing and storage resources dynamically; replicate data in response to demand, and schedules computations close to the data. They give a O (N M) competitive ratio online algorithm, where N is the number of stores, each of which can store M objects of uniform size. However, their model does not allow for keeping multiple copies of an object simultaneously in different stores. In our model, we assume each object can have multiple copies, each on a different site.

A Data Grid can be represented as an undirected graph G (V, E), where a set of vertices $V = \{1,2, . . . n\}$ represents the sites in the Grid, and E is a set of weighted edges in the graph. The edge weight may represent a link metric such as loss rate, distance, delay, or transmission bandwidth. In this paper, the edge weight represents the bandwidth and we assume all edges have the same bandwidth B (in Section 6, we study heterogeneous environment where different edges have different bandwidths). There are p data files D = {$D_{1,}$ $D_{2,}$…………….. D_p } in the Data Grid, Dj is originally produced and stored in the source site Sj ε V . Note that a site can be the original source site of multiple data files. The size of data file Dj is sj. Each site i has a storage capacity of mi (for a source site i , mi is the available storage space after storing its original data). We begin this section by considering an illustrative example which serves as the basis of our problem statement and will be used throughout the paper to demonstrate the main features of our system.

III. FILE SHARING IN MOBILE ADHOC NETWORKS

In this section we will find out the file sharing in mobile adhoc networks that was used in our current application. Here we can discuss about both file sharing in normal MANETs as well as file sharing in Delay Tolerant Network (DTNs).Now we can discuss about that in detail as follows:

FILE SHARING IN NORMAL MANETs

In this section we mainly discuss about the topic of file replication for efficient file sharing applications in MANETs which was been under study. If we go with some of the previous papers [10], [11], [12], an individual node or a group of nodes decide the list of files to replicate according to file querying frequency.

Two well known authors like Hara and Madria [10] proposed three file replication protocols:

a. Static Access Frequency (SAF),
b. Dynamic Access Frequency and Neighborhood (DAFN), And
c. Dynamic Connectivity Based Grouping (DCG).

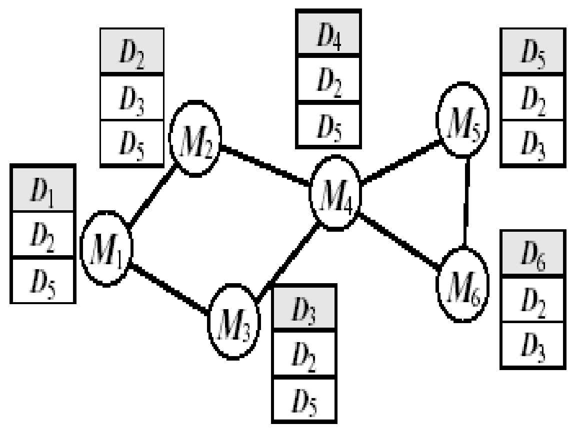
Now let us discuss about the each and every file replication protocols in detail as follows: Out of all the three replication protocols we used in the paper, SAF is the first and primitive protocol that was launched in which each node replicates its frequently queried files until its all the available storage space is used up. This protocol may lead to many duplicate replicas among neighboring nodes when they have the same interested files. Now if we go with the second type of protocol like DAFN , this was launched in order to eliminate the duplicate replicas among neighbors' nodes that are available within the network. Now we can further go with the last protocol like DCG in which all the duplicate replicas that are available in a group of nodes with frequent connections by creating replicas for files in the descending order of their group based querying frequencies.Though DAFN and DCG enable replicas to be shared among neighbors, neighboring nodes may separate from each other due to node mobility. Also, they incur high traffic load in identifying duplicates or managing groups.

A well known writer like Gianuzzi [21] investigated the probability of acquiring a file, which has n replicas in the network, from the potentially partitioned network. He also studied the file retrieval performance when erasure coding [22] is employed. Chen [23] discussed how to decide the minimal number of mobile servers needed to ensure that every data item can be obtained within at most k (k >= 1) hops by any node in the system.

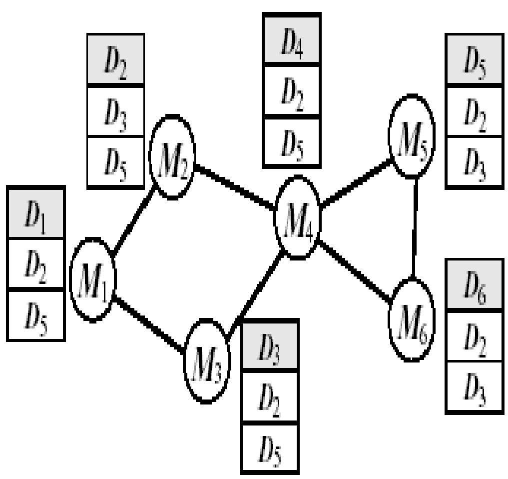
Now let us look about each and every protocol in detail with following below examples.

SAF EXAMPLE

Data	Mobile host					
	M_1	M_2	M_3	M_4	M_5	M_6
D_1	0.65	0.25	0.17	0.22	0.31	0.24
D_2	0.44	0.62	0.41	0.40	0.42	0.46
D_3	0.35	0.44	0.50	0.25	0.45	0.37
D_4	0.31	0.15	0.10	0.60	0.09	0.10
D_5	0.51	0.41	0.43	0.38	0.71	0.20
D_6	0.08	0.07	0.05	0.15	0.20	0.62
D_7	0.38	0.32	0.37	0.33	0.40	0.32
D_8	0.22	0.33	0.21	0.23	0.24	0.17
D_9	0.18	0.16	0.19	0.17	0.24	0.21
D_{10}	0.09	0.08	0.06	0.11	0.12	0.09



DAFN EXAMPLE



M_1-M_2 : $D_2 \rightarrow D_7\,(M_1)$, $D_5 \rightarrow D_8\,(M_2)$

M_1-M_3 : $D_5 \rightarrow D_8\,(M_3)$

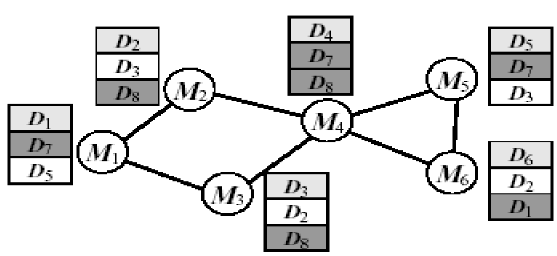
M_2-M_4 : $D_2 \rightarrow D_7\,(M_4)$

M_3-M_4 : No duplication

M_4-M_5 : $D_5 \rightarrow D_8\,(M_4)$

M_4-M_6 : No duplication

M_5-M_6 : $D_2 \rightarrow D_7\,(M_5)$, $D_3 \rightarrow D_1\,(M_6)$



DCG EXAMPLE

Data	Mobile host						Group	
	M_1	M_2	M_3	M_4	M_5	M_6	G_1	G_2
D_1	0.65	0.25	0.17	0.22	0.31	0.24	1.29	0.55
D_2	0.44	0.62	0.41	0.40	0.42	0.46	1.87	0.88
D_3	0.35	0.44	0.50	0.25	0.45	0.37	1.54	0.82
D_4	0.31	0.15	0.10	0.60	0.09	0.10	1.16	0.19
D_5	0.51	0.41	0.43	0.38	0.71	0.20	1.73	0.91
D_6	0.08	0.07	0.05	0.15	0.20	0.62	0.35	0.82
D_7	0.38	0.32	0.37	0.33	0.40	0.32	1.40	0.72
D_8	0.22	0.33	0.21	0.23	0.24	0.17	0.99	0.41
D_9	0.18	0.16	0.19	0.17	0.24	0.21	0.70	0.45
D_{10}	0.09	0.08	0.06	0.11	0.12	0.09	0.34	0.21

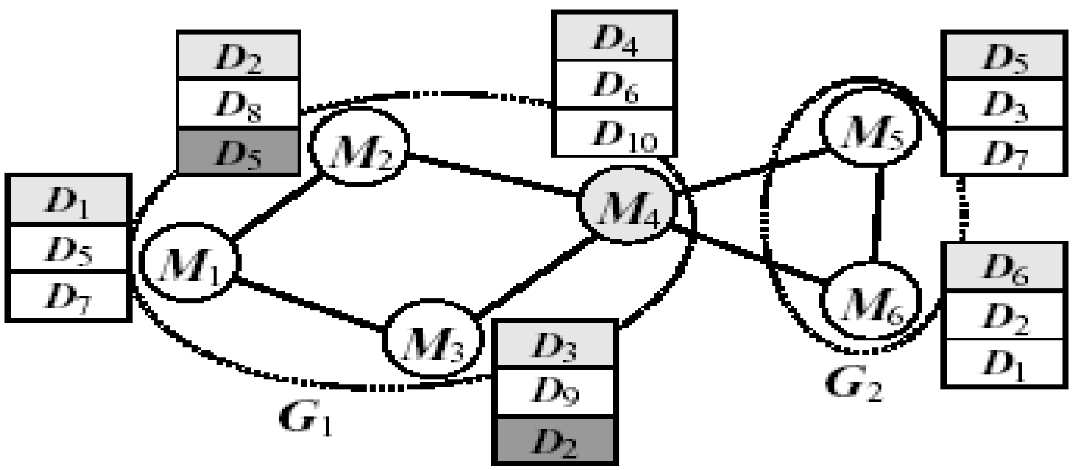


Figure 4. Represents the Examples of Each and Every File Replication Protocol

From the above all figure 4, we can able to find the exact similarities what each and every replication protocol has one with other compared with a set of examples in a mobile adhoc networks.

IV. DESIGN OF FILE REPLICATION PROTOCOL

In this section we will try to find out and examine the design and implementation of a file replication protocol with all its components and this is very important in our proposed application.The two solutions to handle the challenges in achieving the proposed OFRR described above represent a maximal approximation to realize the OFRR in a distributed manner. Based on the solutions, we propose the priority competition and split file replication protocol

(PCS). We first introduce how a node retrieves the parameters needed in PCS and then present the detail of PCS.

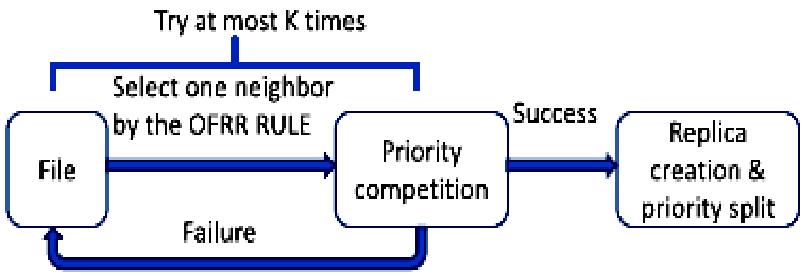


Figure 5. Represents the Architecture of a Replica Distribution Process that was Designed in this Current Application

From the above figure 5,we can clearly show that it keeps trying to replicate file j on nodes it encounters until one replica is created or K attempts have been made. If file j is replicated, its P is split. Next, the node fetches the file from the top of the list and repeats the process. If file j fails to be replicated after K attempts, the node stops launching competition until the next period

V. CONCLUSION

In this paper, we for the first time have investigated a practical problem of how to allocate limited resources for file replication for the purpose of global optimal file searching efficiency in MANETs. Till now there were a lot of primitive protocols that are available in network regarding the same concept but they didn't achieved the highest level of security in achieving all advantages .So in this paper for the first time we try to analyze theoretically the influence of replica distribution on the average querying delay under constrained available resources with two mobility models, and then derived an optimal replication rule that can allocate resources to file replicas with minimal average querying delay. After a deep analysis of our current model we finally designed the priority competition and split replication protocol (PCS) that realizes the optimal replication rule in a fully distributed manner. By conducting various experiments on our proposed model, we finally came to a conclusion that this proposed mechanism gives high level of security in terms of data replication on a shared storage area over a Peer to Peer Network.

VI. REFERENCES

[1] H. Chan, A. Perrig, and D. Song, "Key Distribution Techniques for Sensor Networks," Wireless Sensor Networks, pp. 277-303, Kluwer Academic, 2004.

[2] A. Rasheed and R. Mahapatra, "An Efficient Key Distribution Scheme for Establishing Pairwise Keys with a Mobile Sink in Distributed Sensor Networks," Proc. IEEE 27th Int'l Performance Computing and Comm. Conf. (IPCCC '08), pp. 264-270, Dec. 2008.

[3] T. Gao, D. Greenspan, M. Welesh, R.R. Juang, and A. Alm, "Vital Signs Monitoring and Patient Tracking over a Wireless Network," Proc. IEEE 27th Ann. Int'l Conf. Eng. Medicine and Biology Soc.(EMBS), Sept. 2005.

[4] A. Rasheed and R. Mahapatra, "An Energy-Efficient Hybrid Data Collection Scheme in Wireless Sensor Networks," Proc. Third Int'l Conf. Intelligent Sensors, Sensor Networks and Information Processing,2007.

[5] L. Hu and D. Evans, "Using Directional Antenna to Prevent Wormhole Attacks," Proc. Network and Distributed System Security Symp., 2004.

[6] Y. Tirta, Z. Li, Y. Lu, and S. Bagchi, "Efficient Collection of Sensor Data in Remote Fields Using Mobile Collectors," Proc. 13th Int'l Conf. Computer Comm. and Networks (ICCCN '04), Oct.2004.

[7] J.R. Douceur, "The Sybil Attack," Proc. First Int'l Workshop Peer-to-Peer Systems (IPTPS '02), Mar. 2002.

[8] H. Deng, W. Li, and D.P. Agrawal, "Routing Security in Wireless Ad Hoc Networks," Proc. IEEE Comm. Magazine, pp. 70-75, 2002.

[9] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," Proc. MobiCom, pp. 56-67, 2000.

[10] A. Kansal, A. Somasundara, D. Jea, M. Srivastava, and D. Estrin, "Intelligent Fluid Infrastructure for Embedded Networks," Proc. Second ACM Int'l Conf. Mobile Systems, Applications, and Services (MobiSys '04), June 2004.

[11] B.J. Culpepper and H.C. Tseng, "Sinkhole Intrusion Indicators in DSR MANETs," Proc. First Int'l Conf. Broadband Networks (Broad- Nets '04), pp. 681-688, Oct. 2004.

[12] C. Palazzi and A. Bujari, "A Delay/Disruption Tolerant Solution for Mobile to Mobile File Sharing," Proc. IFIP/IEEE Wireless Days,2010.

[13] Y. Tseng, S. Ni, and E. Shih, "Adaptive Approaches to Relieving Broadcast Storms in a Wireless Multihop Mobile Ad Hoc Network," Proc. 21st Int'l Conf. Distributed Computing Systems (ICDCS), pp. 481-488, 2001.

[14] B. Chiara et al., "HiBOp: A History Based Routing Protocol for Opportunistic Networks," Proc. IEEE Int'l Symp. World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2007.

[15] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic Routing in Intermittently Connected Networks," ACM SIGMOBILE Mobile Computing and Comm. Rev., vol. 7, no. 3, pp. 19-20, 2003.

[16] S. Ioannidis, L. Massoulie, and A. Chaintreau, "Distributed Caching over Heterogeneous Mobile Networks," Proc. ACM Int'l Conf. Measurement and Modeling of Computer Systems (SIGMETRICS),2010.

[17] M.J. Pitkanen and J. Ott, "Redundancy and Distributed Caching in Mobile DTNs," Proc. ACM/IEEE Int'l Workshop Mobility in the Evolving Internet Architecture (MobiArch), 2007.

[18] X. Zhuo, Q. Li, W. Gao, G. Cao, and Y. Dai, "Contact Duration Aware Data Replication in Delay Tolerant Networks," Proc. IEEE 19th Int'l Conf. Network Protocols (ICNP), 2011.

[19] X. Zhuo, Q. Li, G. Cao, Y. Dai, B.K. Szymanski, and T.L. Porta, "Social-Based Cooperative Caching in DTNs: A Contact Duration Aware Approach," Proc. IEEE Eighth Int'l Conf. Mobile Adhoc and Sensor Systems (MASS), 2011.

[20] Z. Li and H. Shen, "SEDUM: Exploiting Social Networks in Utility-Based Distributed Routing for DTNs," IEEE Trans. Computers,vol. 62, no. 1, pp. 83-97, Jan. 2012.

[21] V. Gianuzzi, "Data Replication Effectiveness in Mobile Ad-Hoc Networks," Proc. ACM First Int'l Workshop Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks (PE-WASUN),pp. 17-22, 2004.

[22] S. Chessa and P. Maestrini, "Dependable and Secure Data Storage and Retrieval in Mobile Wireless Networks," Proc. Int'l Conf.Dependable Systems and Networks (DSN), 2003.

[23] X. Chen, "Data Replication Approaches for Ad Hoc Wireless Networks Satisfying Time Constraints," Int'l J. Parallel

VII. ABOUT THE AUTHORS

ADIREDDI RATNAM is currently pursuing her 2 years M.Tech in Department of Computer Science and Technology at Baba Institute of Technology and Sciences, Visakhapatnam, AP, India. Her area of interest includes Networks.

K.SATYANARAYANA MURTHY is currently working as an Assistant Professor in Department of Computer Science and Engineering at Baba Institute of Technology and Sciences, Visakhapatnam, AP, India. He has more than 7 years of teaching experience in various engineering colleges.His research interest includes Data Mining and Networks.