ISSN 2348 – 7968

Optimizing Social Networks For P2P Content Based File Sharing In Detached MANET'S

Akalya K S, Pavithra B , Mr.Vimal S,
B.Tech.,Information Technology, Jeppiaar Engineering College
Chennai,India

Abstract:

The P2P file sharing model makes large-scale networks a blessing instead of a curse, in which nodes share files directly with each other without a centralized server. The successful deployment of P2P file sharing systems and the aforementioned impediments to file sharing in MANETs make the P2P file sharing over MANETs (P2P MANETs in short) a promising complement to current infrastructure model to realize pervasive file sharing for mobile users. As the mobile digital devices are carried by people that usually belong to certain social relationships, we focus on the P2P file sharing in a disconnected MANET community consisting of mobile users with social network properties. In such a file sharing system, nodes meet and exchange requests and files in the format of text, short videos and voice clips in different interest categories. Peer-to-peer file sharing methods in mobile ad hoc networks (MANETs) can be classified into three groups: flooding-based, advertisement-based and social contact-based. The first two groups of methods can easily have high overhead and low scalability. They are mainly developed for connected MANETs, in which end-to-end connectivity among nodes is ensured. The third group of methods adapt to the opportunistic nature of disconnected MANETs but fail to consider the social interests of mobile nodes, which can be exploited to improve the file searching efficiency. Peer-to-peer (P2P) file sharing methods in mobile ad hoc networks (MANETs) can be classified into three groups: flooding-based, advertisement-based and social contact-based. The first two groups of methods can easily have high overhead and low scalability. They are mainly developed for connected MANETs, in which end-to-end connectivity among nodes is ensured. The third group of methods adapt to the opportunistic nature of disconnected MANETs but fail to consider the social interests of mobile nodes, which can be exploited to improve the file searching efficiency.

Keywords:

Distributed computing, peer-to-peer computing, distributed information sharing, peer communities.

I.Introduction:

During past years personal mobile devices like laptops, mobile phones were more popular. Smartphone usage increased around 118 million around the world in 2007 which then became 300 million by 2013.This leads to a promising future, where sharing the files have become more easier than before. Mobile user communicate using a infrastructure formed by graphically distributed stations. But, nowadays they can find themselves using wireless networks.Peer to Peer file sharing method makes large-scale networks where nodes share files directly with each other without a centralized server. This method has become popular and successful among millions of people in file sharing. Deployment of files sharing via peer to peer networking ,has become a complement to current infrastructure model to realize pervasive file sharing for mobile users. People carrying mobile devices were usually have link to certain social relationship. Here, we make a discussion on peer to peer file sharing in disconnected MANETs community which has users with social network properties. In this type of sharing, nodes meet and exchange requests, files which are in the text format, small videos and audios in various interest categories. We can consider a scenario, sharing system in a school campus. In the scenario, nodes that shares the same interests or related topics or the same. In the MANETs, nodes keep constantly moving, which forms disconnected MANETs with positive node encountering. This network connections made a challenge for development of peer to peer MANETs. Peer to peer MANETs supports by traditional methods are flooding based and advertisement based. The main aim is to search a file. This file searching lead to a complexity in broadcasting. In forthcoming methods, nodes will build content tables, and forward files with those tables. As the content tables has expired routes caused by transient network connection, they have low efficiency. Many research has been made on this disconnected MANETs. Searching process in these methods may lead to higher delay when compared to interest based

ISSN 2348 – 7968

searching. Below methods increase the efficiency of message forwarding

1.Nodes usually follow certain movements patterns (i.e) local gathering, device centralities and skewed visiting preferences.
2.Mobile users usually visit or search on certain file interests and they were based on power-law distribution.
3.Users of common interests meet often rather than others.
By using these properties of social networks, we propose sharing via disconnected MANETs using SPOON.Those four components are,

1.Interest Extraction Algorithm.
2.Community Construction Algorithm.
3.Node Role Assignment Algorithm.
4.Interest-Oriented Routing Algorithm.

File searching scheme has two phases.They are Intra community searching and Inter community searching.SPOON uses social network based on both interest and movement pattern. Initially, it classifies common-interest and often encountered nodes as one social community. Second, it takes a frequency at which nodes meet various interests rather than various nodes in file searching. Third, it assigns stable nodes in a community as co-ordination and mobile nodes as ambassadors which travel often to foreign communities. SPOON also contain extra functions such as files pre-fetching, completion of queries and preventing formation of loops and churning of nodes for further enhancement in file searching.

II.Related Works:

P2P File Sharing In MANETs:

Flooding-Based Methods:

This is the foremost approach in p2p MANETs file sharing. It deals with moving nodes in a geographical area to transfer the files between neighbours. Passive Distributed Indexing(PDI) is a general file searching algorithm. Local broadcasting is used for content searching and those content indexes are setup at the nodes along with reply path. Application layer has been a part in file searching and transfer algorithm proposed by Klemn. This Algorithm combines the query results from p2p to remove the routing paths. Hayes extended Gnutella system moving environments and explained the use of keywords that represents use of sharing file.

Social-Network Based Methods:

Long-term neighbouring relationships were utilized by services in MOPS. Nodes that has frequent contents are grouped together and the nodes with different searches are grouped for intercommunity communication. Files and contents are disseminated through brokers to reach inter-communities. Node movement is the only criteria in MOPS and whereas SPOON enhances the use of intercommunity Appointing one ambassador for each and every foreign community, which helps in forwarding a query to the point directly.The messages from the ambassadors are utilized by the stable nodes.Hence, the above methods are used to transfer information to the matched subscriber. So, these methods cannot be applied to search a file directly.

interests. Thus, this flood-based methods results in high overhead because of broadcasting.
Tchakarov and Vaidya proposed a method on content encountering on location based ad-hoc network. It transfers content and request in different direction to ensure this encountering. P2P gathers both advertisement and discovery process. Each file holder regularly transmits on advertisement based information to neighbouring nodes about it files. The discovery process encounters the desired file and leaves the clue to other subsequent search request. Repanties and Kalogerabi proposed a system by the use of Bloom Filter which forum the content page and transmit them to all others nodes who makes this query requests. This method has no guarantee of successful file search and share because of high mobility of nodes.

P2P File Sharing in Detached MANETs:

Cache/Replication-Based Methods:

Huanget proposed a system which considers many criteria like node mobility, file popularity and file server in file replica creation for file presence in PDI. Gao et al proposed at a co-ordinated accumulation in tolerable networks. It copies every search file into central locations so

III.SPOON Design:

Usually p2p file sharing in MANETs has these methods,

ISSN 2348 – 7968

1. Content representation.

2. Structure of Node Management.

3. File sharing from above methods.

Trace Data Analysis:

Based on the correlation between nodes interest search and these frequency of contact, validation is performed. We say here, the total meeting time of nodes is the sum of time length of trace. For each encountering. The nodes which has meeting time length larger than that of four times were grouped together as a community that has half of all nodes in the community. Average number of shared interests tracks for each node is calculated in its own community and also with other community nodes. Atleast, the node's average values in each community is calculated.

Community	Avg no shared interests with nodes in community	Avg no shared interests with nodes not in community
1	1.50	0.99
2	0.83	0.69
3	1.17	0.79
4	1	0.39
5	1.19	0.94
6	0.33	0.21
7	1.1	0.71

Interest Extraction:

Node Information can be classified into different interests categories. It is identified that nodes usually have few files that they share or which they frequently ask queries on those interested files. These are just a 20 percent of total file categories. In other type of sharing system, nodes that share file with common interest can form a separate community. These are the methods we follow under Interest Extraction technique. Thus, SPOON extracts the nodes interests from its files.

Community Construction:

In Social Network group, nodes with similar interest tend to meet frequently. SPOON classifies them according to common interest and frequent contact for interest-based file searching. Nodes with multiple fascinates belong to multiple communities. The community construction can facilely be conducted in a centralized manner by amassing node fascinates and contact frequencies from all nodes to a central node. However, considering that the proposed system is for distributed disconnected MANETs, in which timely information accumulation and distribution is nontrivial, we further propose a decentralized method to ascertain the adaptive of SPOON in authentic environment. When two nodes, verbally express N1 and N2, meet, they consider two cases for community engenderment:

1) they do not belong to any communities.

2) at least one of them is already a member of a community.

The purpose of taking into account the weight of each interest group is to eliminate the noise of interest groups with a minuscule number of files and achieve more preponderant interest clustering. If N1 and N2 have at least one pair of matched interest group, and their contact frequency is higher than the top h1 percent highest encountering frequencies in either node, the two nodes form an incipient community. The keywords in their matched interest groups and corresponding weights constitute the community vector of the community. In the second case, suppose N2 is already a member of community C, N1 calculates for each of its interest groups, verbalize , to decide whether it should join in community C. If the homogeneous attribute value for one interest group is more immensely colossal than TG, and N1's contact frequency with community C is higher than the top h2 percent of N1's contact frequencies with all nodes it has met, N1 is granted the membership to community C. The contact frequency with community C refers to the accumulated contact frequency with nodes in C. It is updated upon each encountering with a node in C. This denotes that Ni contacts members in community C frequently enough to assure connections. N1 then copies the community vector and other community information from N2. Additionally, when a node meets the community coordinator, it reports its files to the coordinator to update its file index and community vector. The coordinator then forwards the updated community information to community members when meets them. With above community construction method, nodes with prevalent fascinates and frequent contacts gradually form a community. However, nodes that appear later have more stringent community acceptance requisite. Its contact frequency to the community needs to be higher than that of more nodes, and its interest vector is compared with a longer community vector. Withal, nodes in a group admit incipient members distributive. As a result, nodes in a group may not have very kindred intrigues or high contact frequencies. We propose two solutions to alleviate this quandary. First, we set an initial period for incipiently joined nodes in which they accumulate contact frequencies with others. Then, when a node commences to join in communities, its meeting frequencies with others are relatively stable, which provides more precise quantification for determining the communities to join in. Second, we utilize group member pruning. Subsisting community members can have a second round voting to corroborate the eligibility of incipient community members. Categorically, if N2 in community C finds a node, verbally express N1, slakes the requisites of C, it awards N1 a potential membership for C. Then, other community members in C further checks N1's eligibility to join in C. That is, every time when N1 meets a subsisting member of C, verbally express N3, N3 checks whether they have at least one pair of matched interest group and whether their contact frequency is higher than the top h1 percent of N1's highest contact frequencies. If affirmative, N3 approves the membership of N1.

Node Role Assignment:

A precedent study has shown that in a social network consisting of mobile users, only a component of nodes have high degrees [20]. We can often find a consequential or popular person who coordinates members in a community in our daily life. For example, the college dean coordinates different departments in the college, and the department head connects to faculty members in the department. Thus, we capitalize on variants of node mobility for file sharing. We define community coordinator and ambassador nodes in the view of a social network. A community coordinator is a consequential and popular node in the community. It keeps indexes of all files in its community. Each community has one ambassador for each kenned peregrine community, which accommodates as the bridge to the community. The coordinator in a community maintains the vector of communities and corresponding ambassadors to map queries to ambassadors for intercommunity.

Community Co-ordinator Node Selection:

A stable node that has tight frequency contact with other community is defined to be community co-ordinator. The importance of centrality is used to estimate the efficiency of vertex within the network.

Then we define the centrality node that assigns weight to each contact frequency, for the selection of co-ordinator which exposes the tighter among the co-ordinator discovery. Consider a node(N) in a community gets all the information from the neighbours in the same community. The centrality degree is calculated by,

$$D(p)=\sum_{i=1,j\neq i}^{N} w$$

Where w is the link between different nodes. In order to reflect the property, the co-ordinator has link between all other nodes in the same and all community members, w equals 1, if frequency contact is between two nodes which is larger than threshold and 0 otherwise. Though, this method cannot have connection with all community members, it has overall tightest connection with all members in the community.

Every nodes checks for its centrality and broadcast information to all members in the community. If a node does not receive any high centrality score than itself, it fixes itself as a potential co-ordinator. It requests all information old co-ordinator. If it becomes the potential co-ordinator, it request all information from old co-ordinator. Now, when new co-ordinator meets members in community, they disseminate information for group vector update and request routing.

Community Ambassador Node Selection:

An ambassador is the one who bridges the co-ordinator in its own community as well as the foreign community. In order to select The ambassador, we consider the node frequency contact with the co-ordinator and the foreign community. Each node(i) calculates the value of utility for foreign community k by,

$$U=F(N,C)*F(N,N)$$

Where c represents foreign community N is the co-ordinator in its own community and F(.) denotes meeting frequency. Now, community co-ordinator choose an ambassador for every foreign community. Default ambassador is the one, which has weighter frequency contact with all foreign community. If the request fails to find ambassador, the default ambassador carries and perform the request. If the ambassador losses the connection with the co-ordinator, the default ambassador (or) new ambassador satisfies the requests. This method facilitates the interest-oriented file sharing and searching in MANETs.

In the design, ambassador is the heart that connects different communities efficiently, Though this method has the limited cost because,

1. This happens among the community members.

2.As nodes have stable centrality, longer broadcast time can be set.

ISSN 2348 – 7968

Ambassador selection involves the nodes to inform its utility to the co-ordinator using beacon messages. Therefore, it does not cost much.

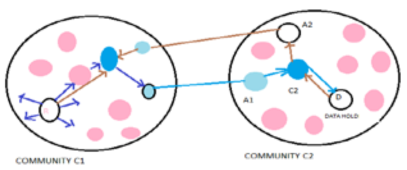
Interest-Oriented File Searching and Retrieval:

In social network, people normally have same file interests and visible patterns in certain distribution. Hence, people with same interests tend to have contact with each other's frequency. Thus, it is forming a good guidance for file searches and shares. Considering the relation among nodes moving pattern and interests, we can route requests the file holders based on frequency meeting between different nodes. This interests based file searching is of two types,

1.Intracommunity search
2.Intercommunity search

A node initially makes searches inside its own community. If the co-ordinator requests is not satisfied in its own community, it then moves to intercommunity for the satisfaction of requests (ie)it forwards the requests to the foreign community. The requests will be deleted if the time expires. During this search, nodes sends messages to another node using Interest-Oriented Routing Algorithm, where messages were forwarded to the node that has requests alike. The retrieved file is routed to the search path.

Fig 1.



INFORMATION EXCHANGE AMONG NODES:

The information exchanged among nodes can be summarized as follows.i.e.,there are two kinds of phases-community construction phase and node assignment phase.in the construction phase,two node which have come across exchange their interest and community vectors.For the assignment of role,the nodes transmit their degree centrality for coordinator selection within their communities.Once the coordinator has been selected,its ID is also transmitted to all the other nodes inside the community and then each node would report its contact frequencies with foreign communities for ambassador selection.In addition to these when a node comes across a coordinator of its own community,it send its updated node vector to the coordinator for update and retrieval.Whereas when an ambassador meets the coordinator it reports about the community vector of foreign community.the two nodes which have came across exchange their node and history vectors for packet routing.Each node has to check the packets in it to decide which packet can be forwarded to the other node based on the file searching algorithm.If the network is stable the frequency of information exchanged for these two phases can be reduced to save money.

INTELLIGENT FILE PREFETCHING:

The ambassadors has an advantage of meeting nodes holding different since they can travel between different communities. Because of this they can prefetch popular files outside the home community. To forward a file residing in a community through the coordinator of the local community there is a query.The frequency of the local queries are tracked by each coordinator for remote files to provide information to the ambassador in its each community. When a community gets to know that the foreign community neighbors have popular remote files which are requested numerous times by members it stores the files on its memory. These prefetched files serve the requests in the ambassador home community which reduces the file searching delay.

QUERYING-COMPLETION AND LOOP-PREVENTION:

When a file query is given there exists a number of matching files in the system. To specify the number of files that it has to find, a node can associate a parameter S_{max}. When the matching files are discovered we need to ensure that the querying process stops. To solve this we let a query carry the parameter S_{max} once it is generated. When the matching file is found the parameter S_{max} decreases itself by one. When this query is replicated to another node the parameter S_{max} is evenly split to the two nodes. When S_{max} equals to 0 then the query stops searching files and when it wants to find more than one file it forwards the IRA to the same node repeatedly. To avoid this, SPOON incorporates two strategies, they are adding ID to the query before forwarding to the next node and the other one is the node has to record its queries it has received within a certain period of time. The former method avoids sending a packet to nodes it has visited before while the latter method prevents sending replicas of the same query to the same node. When a node N_i forwards a query to a

ISSN 2348 – 7968

newly met node N_j based on IRA, N_i checks the query's record of traversed nodes contains N_j and if yes, it does not forward the query. When a node receives a query, it checks whether the query exists in its record of received queries then the node sends the query back to the sender. These strategies avoid searching loops by simply avoiding the same query forwarding to the nodes which has received the query before.

NODE CHURN CONSIDERATION:

In SPOON, when a node joins a system it checks the community it belongs to know the ID of the community coordinator so that it can report its files and utility values to the community coordinator when meeting it.It helps the coordinators to maintain the updated information of the community members. When users manually stop the SPOON application on their devices the node leaves the system willingly. While leaving, the node informs about the departure through IRA so that when the leaving node is an ambassador then the coordinator have to choose a new coordinator whereas if the leaving node is a coordinator it broadcasts to notify others to select a new coordinator. At the same time a node can abruptly leave the system due to various reasons.Simply depending on the periodical beacon message a node cannot tell whether a neighbor has left or isolated from itself but this is usaual in mobile adhoc networks.To manage this problem the time stamps are recorded while meeting the nodes and it is being send to the coordinator through IRA and after receiving this information the coordinator updates the latest time stamp of each node which was seen by ther nodes.When the timestamp of anode is more than T_x seconds ago then it is considered as the departed node.The timestamps are normally maintained by the nodes to determine whether it is still alive.A node attaches the coordinator departure information on the beacon messages so that the neighbor nodes knows whether the coordinator has left.Note that a node can know the number of community members from the coordinator.When a node came to know that half of the coordinator members are aware about the departure of the coordinator then a coordinator reelection message is broadcasted to select a new coordinator by the same method explained in Section 3.4.1.

4. PERFORMANCE EVALUATION

MOPS is a social-network based content service system. A community is formed by grouping the frequent contact nodes and select the nodes with frequent contacts with other communities as brokers for communication. PDI+DIS is a combination of PDI and an advertisement-based DIS semination method (DIS) . PDI provides distributed search service through local broadcasting (three hops), and builds content tables in nodes along the response paths, while DIS let each node disseminate its contents to its neighbors to create content tables. CacheDTN replicate files to network centers in decreasing order of their overall popularity. In PodNet, nodes cache files interested by them and nodes they have met. We adopted the "Most Solicited" file solicitation strategy in PodNet. We doubled the memory on each node in CacheDTN and PodNet for replicas. In Epidemic, when two nodes meet each other, they exchange the messages the other has not seen.The following experiments are conducted.they are

1. Evaluation of community construction. We first evaluated the proposed community construction algorithm introduced in Section 3.3.
2. GENI experiments. We deployed the systems on the real-world GENI ORBIT test bed and tested the performance using the MIT Reality trace. The GENI ORBIT test bed contains 400 nodes with 802.11 wireless cards. Nodes can communicate with each other through the wireless interface. We used real trace to simulate node mobility in ORBIT: two nodes can communicate with each other only during the period of time when they meet in the real trace.
3. Event-driven experiments with real trace. We also conducted event-driven experiments with two real traces.
4. Evaluation of enhancement strategies. We evaluated the effect of the enhancement strategies introduced in Sections 3.7, 3.8, and 3.9 through event-driven experiments.
5. NS2 experiments with synthetic mobility. We conducted experiments on NS-2 using a community-based mobility model to evaluate the applicability of SPOON in different types of networks.

We disabled the intelligent prefetching and the multicopy forwarding (i.e., _ ¼ 1 and _ ¼ 1) in SPOON to make the method comparable. Also, because the comparison methods can return only one file for a query, we set Smax ¼ 1 in SPOON. In each community, we used the node having the most contacts with other communities as the ambassador in SPOON and as the broker in MOPS. We also set the node with the most contacts with its community members as the coordinator in SPOON. Besides the Haggle trace, we further tested with the MIT Reality trace , in which 94 smartphones were deployed among students and staffs at MIT to record their encountering. The two traces last 0.34 million seconds

(Ms) and 2.56 Ms, respectively. As in MOPS, we used 40 percent of the two traces to detect groups in which nodes share frequent contacts. Here, we use "group" to represent a group of nodes with frequent contacts, and use "community" to represent a group of nodes with common interests and frequent contacts. We got seven and eight groups for the MIT Reality trace and the Haggle trace, respectively. Then, because there is no real trace for Peer to peer over mobile adhoc networks, we collected articles from different news categories (e.g., sports, entertainment, and technology) from CNN.com and mapped them to the identified communities. Each node contains 50 articles from the news category for its community. Each node extracts its interests from its stored files. The similarity threshold was set to 70 percent in AGNES for file classification. In experiments with the Haggle trace and the MIT Reality trace, we set the initialization period to 0.09 Ms and 0.3 Ms, the query generation period to 0.1 Ms and 1 Ms, and the TTL of a query to 0.15 Ms and 1.2 Ms, respectively.People generate queries based on their interests,so we have o set 70 percent of total queries searching for files located in the local community whre every query is for an article randomly selected from the article pool.The following metrics are measured:

1. Hit rate: The percentage of requests that are successfully delivered to the file holders.
2. Average delay:The average delay of the successfully delivered requests
3. Maintenance cost: The total number of all messages except the requests, which are for routing information establishment and update, or replication creation.
4. Total cost: The total number of messages, including maintenance messages and requests, generated in a test.

EVALUATION OF COMMUNITY CONSTRUCTION*:*

When spoon community construction is compared with Active and Centralized community construction. The active community construction collects node contacts and interests when they meet nodes to select three most active nodes. But in centralized community construction a super node collects all node contacts and interests. Both use Agnes to build communities with the information collected. Since there is no real trace of content sharing in peer to peer mobile adhoc networks it has to be tested in an indirect way. We first conducted the group construction and content distribution as previously described, and then removed the group identity of each node. Then, we run the three methods to create communities. After this, we matched each new community to the most similar old community and calculated the similarity value by $N^2{}_{sm}/(Np *Nn)$, where N_{sm} is the number of common nodes, N_p and N_n denote the size of the old community and the new community, respectively. In SPOON CC, we set TG to 1 to ensure interest closeness, and set h1 and h2 to 30. Active-CC and Centralized-CC used the same threshold for granting community membership as SPOON community construction. The average similarities in SPOON-CC, Active-CC, and Centralized-CC are 0.95, 0.87, and 1 with the Haggle trace and 0.91, 0.85, and 1 with the MIT trace, respectively. The Centralized-CC has inferior performance because active nodes can only collect information from nodes they have met, leading to less accurate community construction. The performance of SPOON-CC is close to that of Centralized-CC, which has the best performance.

EFFICIENCY AND COST IN THE EXPERIMENTS ON GENI

Method	Hit Rate	Ave.Delay(s)	Maintenance Cost	Total cost
SPOON	0.671	152731.3	258764	275312
MOPS	0.629	163282.5	310131	320412
Cache DTN	0.5712	219021.4	283210	298123
Pod Net	0.5932	183621	223218	240238
PDI+DIS	0.524	7418.9	298641	359841
Epidemic	0.8813	15621.2	669193	860475

When h_1 and h_2 varied from 20 to 50 to verify the effectiveness of Spoon community configuration.

GENI EXPERIMENTS

From the above table we find that Epidemic generates the highest rate with the highest total cost and alow average delay.This is the result of the dissemination nature of broadcasting.Spoon can produce highest hit rate at the second lowest total cost because it uses both contacts and content properties of networks to guide file querying and so it can locate the queried files without any information changes and request messages.

EVENT-DRIVEN EXPERIMENTS WITH REAL TRACE:

The total number of queries are varied from 5,000 to 25,000 to show the scalability of each method in terms of the amount of queries.

HIT RATE

Epidemic has the highest hit rate because of its broadcasting nature. In SPOON, coordinators and ambassadors facilitate intra- and intercommunity searching, while the IRA actively forwarded to the node with a high probability of meeting the destination. MOPS only depends on the encountering of mobile brokers for file searching. This probability is lower than that of SPOON, resulting in a lower hit rate. Pod Net and Cache DTN lack active query forwarding, leading to median hit rates. However, replicas on each node are more catered to the interests of nodes it can meet in Pod Net, while Cache DTN just caches files on network center. Therefore, Pod Net has slightly higher hit rate than Cache DTN

AVERAGE DELAY

We only measure the delay of successful requests. In PDI+DIS, most successful requests are resolved in the initial 3-hop broadcasting stage. Therefore, it generates the least average delay. In Epidemic, requests are rapidly distributed to nodes at the cost of multiple copies. As a result, requests can reach their destinations quickly. MOPS exhibits a large delay because requests in it usually have to wait for a long time for brokers or same community file holders. In contrast, SPOON always tries to find an optimal neighbor to send a request to the file holder with the interest-oriented routing algorithm. In addition, the designation of ambassadors in SPOON increases the possibility of relaying requests to foreign communities. As a result, SPOON has lower average query delay than MOPS. PodNet and CacheDTN generate high average delay because queries only wait for file holders passively on their originators. However, because PodNet create replicas that are more likely to be encountered by nodes that are interested in them, it has lower average delay than CacheDTN.

COST

When the total number of queries is larger than 10,000, the maintenance costs generally follow: Epidemic>PDIþDIS>MOPS>CacheDTN>SPOON>PodNet.
In PodNet, nodes only replica interested files when meet other nodes, leading to the least maintenance cost. In SPOON, nodes exchange node vectors for the update of history vector. Nodes also report its contents to coordinators for file indexing. In MOPS, brokers exchange the contents of all nodes from their home communities when meeting each other. Therefore, MOPS produces slightly higher cost than SPOON. The active replication of files to network centers in CacheDTN leads to a high cost. PDI+DIS needs to build content tables through reply messages and disseminated queries, so it has higher maintenance cost than above four methods when the number of queries becomes large. In Epidemic, two nodes need to inform each other requests already on them, which causes a lot of information exchange and leads to the highest maintenance cost. We see that when the number of queries increases, the maintenance costs of SPOON, PodNet, CacheDTN, and MOPS remain stable while those of Epidemic and PDI+DIS increase quickly. This is because the maintenance costs of the former four methods are determined by the information/replication exchanges among nodes and are irrelevant with the number of queries.

EVALUATION OF THE ENHANCEMENT STRATEGIES:

MULTICOPY FORWARDING AND PREFETCHING

Multicopy forwarding and prefetching denote the Spoon with the corresponding improvement strategy.In multicopy forwarding each query originator distribute two copies of its query.Whereas in prefetching we let ambassador store top 10 most popular files.
the multicopy forwarding strategy with only two copies enhances the hit rate greatly in the experiments with both traces. This is because when each query has two copies in the system, its probability of encountering the node containing the queried file increases. Such a result shows the effectiveness of the multiforwarding strategy.

HIT RATE IMPROVEMENT WITH THE MIT REALITY TRACE

No.of packets	Original	Prefetching	Multi-copy forwarding
5000	0.761371	0.7671675	0.780413
10000	0.759941	0.762841	0.770838
15000	0.760135	0.763762	0.772843
20000	0.756418	0.759957	0.773901
25000	0.751835	0.754837	0.771963

ISSN 2348 – 7968

EFFECT OF QUERY-COMPLETION STRATEGY

TRACE	SPOON-OR	SPOON-QC	SPOON-QCLP
HIT RATE			
Haggle	0.8741	0.8701	0.8813
MIT Reality	0.9091	0.9012	0.9406
NUMBER OF QUERY FORWARDING OPERATIONS			
Haggle	569841	530516	304953
MIT Reality	721852	609863	249132

From the two tables it is known that the file prefetching strategy slightly improves the hit rate with the both the two traces.It is because there are only two ambassadors per community and the prefetched files only a small amount of queries because most queries are for contents in local community and the improvement on the hit rate still demonstrates the effectiveness of the file pre fetching strategy.

QUERYING-COMPLETION AND LOOP-PREVENTION

SPOON without querying-completion as "SPOON-OR", SPOON with the querying-completion only as "SPOON-QC", and SPOON with both querying-completion and loop-prevention as "SPOON-QCLP". We set the number of queries to 15,000. In SPOON-OR, queries do not stop searching until TTL expiration.From the table it is observed that Spoon=QCLP has slightly higher hit rate than Spoon-QR and Spoon-QC because of the loop-prevention which avoids forwarding a query to the same file holder repeatedly thereby using forwarding opportunities more efficiently. SPOON-QC has slightly lower hit rate than SPOON-OR because it stops querying when S_{max} files are fetched.

EFFET OF THE DETECTION OF COORDINATOR DEPARTURES

Trace	w/o churn consideration	w/churn consideration-Ab	w/churn consideration-Vo
Haggle	0.650643	0.684512	0.729942
MIT Reality	0.641053	0.677841	0.746842

NODE CHURN CONSIDERATION:

The node that contains a file matching a query as the primary matching node for the query. To demonstrate the performance of SPOON in node churn, for each queried file, we purposely created a file that has 70 percent similarity with it in a nonleaving node in the same community with the primary matching node. We name this node as the secondary matching node.

Without node churn consideration, queries just wait for coordinators until their TTL expiration if they need to be forwarded to coordinators, leading to a low hit rate and a high average delay.

5. Conclusion :

We explored a Social network based P2P content file sharing system in detached mobile adhoc Networks (SPOON). SPOON considers both node interest and contact frequency for efficient file sharing. We introduce four main components of SPOON: Interest extraction identifies nodes' interests; Community construction builds common-interest nodes with frequent contacts into communities. The node role assignment component exploits nodes with tight connection with community members for intra-community file searching and highly mobile nodes that visit external communities frequently for inter-community file searching; The interest-oriented file searching scheme selects forwarding nodes for queries based on interest similarities. SPOON also incorporates additional strategies for file prefetching, querying-completion and loop-prevention, and node churn consideration to further enhance file searching efficiency.

ACKNOWLEDGMENTS

This research was supported in part by US National Science Foundation grants IIS-1354123, CNS-1254006, CNS-1249603,
OCI-1064230, CNS-1049947, CNS-0917056, and CNS-1025652, Microsoft Research Faculty Fellowship 8300751, and the United States Department of Defense 238866. An early version of this work was presented in the Proceedings of MASS 2011 [40].

REFERENCES

[1] "The State of the Smartphone Market," http://www.allabout symbian.com/news/item/6671_The_State_of_the_Smartphone_
Ma.php, 2013.

[2] "Next Generation Smartphones Players, Opportunities & Forecasts

2008-2013," technical report, Juniper Research, 2009.
[3] "A Market Overview and Introduction to GyPSii," http:// corporate.gypsii.com/docs/MarketOverview, 2013.
[4] "Bittorrent," http://www.bittorrent.com, 2013.
[5] "Kazaa," http://www.kazaa.com, 2013.
[6] M. Papadopouli and H. Schulzrinne, "A Performance Analysis of 7DS: A Peer-to-Peer Data Dissemination and Prefetching Tool for Mobile Users," Proc. IEEE Sarnoff Symp. Digest Advances in Wired and Wireless Comm., 2001.
[7] A. Klemm, C. Lindemann, and O. Waldhorst, "A Special-Purpose Peer-to-Peer File Sharing System for Mobile Ad Hoc Networks," Proc. IEEE 58th Vehicular Technology Conf. (VTC '03), 2003.
[8] C. Lindemann and O.P. Waldhort, "A Distributed Search Service for Peer-to-Peer File Sharing," Proc. Int'l Conf. Peer-to-Peer Computing (P2P '02), 2002.