A recommender system for team formation in MANET

Waleed M. Al-Adrousy a,*, Hesham A. Ali b, Taher T. Hamza a

Received 15 February 2014; revised 16 April 2014; accepted 4 June 2014
Available online 3 April 2015

KEYWORDS
Mobile Ad-Hoc Networks (MANET); Social networks; Recommender systems; Group formation; Zone routing protocol (ZRP); Peer-to-peer (P2P)

Abstract Mobile social networking is a new trend for social networking that enables users with similar interests to connect together through mobile devices. Therefore, it possesses the same features of a social network with added support to the features of a Mobile Ad-hoc Network (MANET) in terms of limited computing power, limited coverage, and intermittent connectivity. One of the most important features in social networks is Team Formation. Team Formation aims to assemble a set of users with a set of skills required for a certain task. The team formation is a special type of recommendation which is important to enable cooperative work among users. Team formation is challenging since users’ interaction time is limited in MANET. The main objective of this paper is to introduce a peer-to-peer team formation technique based on zone routing protocol (ZRP). A comparison was made with Flooding and Adaptive Location Aided Mobile Ad Hoc Network Routing (ALARM) techniques. The suggested technique achieves fast successful recommendations within the limited mobile resources and reduces exchanged messages. The suggested technique has fast response time, small required buffering and low power consumption. The testing results show better performance of the suggested technique compared to flooding and ALARM technique.

© 2015 The Authors. Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Understanding the formation, evolution and utilization of online social networks becomes important due to the increased interest and usage of social networks in people’s way of life. A social network is a set of organizations or other social entities connected by a set of social relationships. It can be depicted as a large graph, with users represented as node, which are connected by edges that represent relationships. Normally, social networks assume a large degree of location stability, and do not focus on the mobility of users who are using laptops and mobile devices to access the social network. Mobile users need to exchange resources and experience in many times even if they do not have a long previous relationship.

The mobility of users has introduced the mobile social networks (MSN) which according to Tian (2012) have three types:
web-based, ad-hoc and hybrid networks. The web-based social networks use client server architecture with some possible variation at server implementation. The mobile nodes use web browsers to access web-based social networks. The ad-hoc networks are more similar to peer-to-peer (p2p) networks with no centralized server. This type of network is also called Mobile Ad-Hoc Networks (MANET) according to Li and Khan (2009).

This research focuses on mobile ad-hoc social networks. MANET algorithm design has many challenges, as illustrated in Singh (2012) and Sridhar and Baskaran, 2010. A MANET’s mobile nature dictates that the network be infrastructure-less, which results in a dynamic topology. This dynamic topology is constantly changing with time due to the node mobility and join/leave operations. This also affects the organization of node collaboration. Mobile nodes have power constraints and limited processing capabilities, affecting their lifetime and connectivity in the network. Nodes’ lifetime and quality of service are further impacted by the limited link bandwidth and the transmission quality for complex media. The variety in mobile devices in the network, as well as the self-configurable nature of the different nodes, leads to a heterogeneous network. The wireless coverage of nodes in MANETs leads to the well-known hidden terminal problem, which results in nodes being hidden from each other although in the direct coverage range of intermediate nodes. Also, the network supports a many-to-one communication pattern, which needs sophisticated synchronization and redundancy elimination techniques. The open nature of the wireless transmission media raises security and privacy concerns.

Before discussing the specific problem of this research, a quick overview shall be made in that sub-section to explore the general problems in recommender systems and MANET. Social network study in general requires several techniques from distributed systems and data mining sciences.

Generally, social networking has some common problems with many distributed systems like scalability, transparency and fault tolerance. It has also some specific problems according to its sociality nature. Building user and relationships model is one of the major problems in social networks. The user and relationships model must be storable in hosting machine media. The model also must enable searching for users and detecting similarity of users and group recommendation according to Brauer and Schmidt (2012) and Konstan and Riedl (2012).

The recommendation of users or groups is a problem of performance and accuracy. Indexing and directory building are used techniques to achieve the speed of recommendations and data retrieval in general as in Ben Mokhtar et al. (2010). The recommended items should be reasonable from the user-view and also fast and effective during computation. Trust of the resulted recommendation is also another problem since some recommendations may be inappropriate for some users as illustrated in Golbeck (2009). The Collaborative Filtering (CF) process aims to handle this problem by using the cumulative experience of user’s friends to recommend new friends or items as in Pham et al. (2011). Collaborative filtering must handle some privacy and security attacks to protect users and their profiles and prevent deceiving recommendations. Some CF attacks are known and have been studied in order to be eliminated and preserve the efficiency of the original social network in the same time according to Su and Khoshgoftaar (2009).

In many recommender systems, a lot of researches have been made assuming a general nature of communication between users and some sort of time stability of users’ existence in the target network. This assumption is valid in traditional web based social networks, but may be not in mobile ad-hoc based social networks where the communication intervals among users are relatively small. The main problem in this research is to enable group formation in MANET social networks. This research does not assume pre-existing groups in advance, but it focuses on building groups of skilled individuals in MANET in a relatively short interval of time. This research focuses on building groups of skilled individuals in MANET in a relatively short interval of time, rather than assume pre-existing groups in advance. The presented research can be used to perform some ad-hoc teamwork projects like developing a website, designing a UML diagram, testing of a large website, or even writing codes. Other benefit of that kind of network is resource exchanging. The exchanged resources can be books, articles, spreadsheets or code artifacts. When code artifacts are exchanged, this can enable reusability of software components to assemble a product as in Ralyté and Rolland (2001). Component reusability is an important target for software engineers. The main focus is performance of recommendation and building groups, not to recommend pre-existing groups. The security and privacy problems are out of this research scope. Moreover, it focuses on the mobile networks which have dynamic topology. The research handles the short co-existence periods of neighbors in MANET.

The target problem of this research is Team-Formation recommendation. There are similar problems to this problem like community detection (Nguyen et al., 2011), predicting links in social networks using random walks (Backstrom and Leskovec, 2011), detecting influential nodes in Zhang et al. (2011), and Database querying in social networks (Papapetrou et al., 2012). However, the MANET Team-Formation problem has some differences from those other problems such as:

- The unpredictable skill demands to respond.
- Requesting online results not offline.
- Selecting top members from same field and consistent members from different fields to build teams.
- The preference of members having several skills over the members with single skill.
- Handling the departure of nodes in dynamic MANET.
- The possible conflict between the power and distance factors that need short routing paths and the variable skill levels of mobile nodes which can lead to choosing distant nodes.

From all those factors, this paper discusses a relatively new problem with a few studies made before. This paper discusses a new style of team-working within mobile environment.

The contribution of this paper can be stated as following: (i) Discussing a relatively new problem in mobile social networks, (ii) Presenting a suggested peer-to-peer searching technique based on a (ZRP) with some modifications to form teams in MANET, (iii) Limiting the team formation process with practical constraints of time and hardware, (iv) Suggesting a
A recommender system for team formation in MANET

2. Related work

2.1. Social networks over internet

Mobile social networks inherit the traditional problems of wireless communications and wireless networking; as introduced in Akyildiz et al. (2002): fault-tolerance, scalability, hardware constraints, topology, production and deployment costs and transmission media constraints. In addition, mobile social networks have some additional problems/constraints compared to traditional social networks according to Li and Khan (2009). These constraints can be concluded as follows: (i) Mobile social networks must be decentralized and load balanced, (ii) Mobile social networks do not depend on the real users' relationships only, but have to consider the geographic location of users and neighborhood relations at the same time, (iii) The limitations of mobile power, memory and processing are a great challenge for many MANET algorithms, (iv) Routing is an important problem as introduced earlier in this research and discussed in Senthilkumaran and Sankaranarayanan (2013) and Khamaysheh et al. (2011), and (v) The balancing of the need for Quality-of-Service (QoS) and best-effort principles is a major problem in almost all MANET algorithms in Sesay et al. (2004) and Rikli and Almogari (2013). QoS metrics according to Mohapatra et al. (2003) are defined as a set of service requirements that needs to be met by the network while transporting a packet stream from a source to its destination. QoS metrics are like: delay, bandwidth, probability of packet loss, delay variance (jitter), power consumption and service coverage area.

The contents of users’ models can vary from a social network to another. However, Recommender systems need to have a profile for each user as a part of his/her model as illustrated in Gamah et al. (2011). User’s profile can be static or dynamic as illustrated in Zayani (2012). A static profile contains user’s basic information; preferences given by user explicitly such as preferences and selected friends. A dynamic profile is inferred by analyzing user’s actions such as types of daily browsed pages. Both profile types have several modeling techniques, including simple databases and vector representation for user and his/her corresponding interests as illustrated in Ghaht and Abdullah (2011) and Hassanpour and Zahmatkesh (2012). Vector Space Modeling (SVM) was the start of a newer technique that focuses on semantic which is named Latent Semantic Analysis in Fernández-luna et al. (2011). Most recommender systems focus on dynamic profile because of its updatable nature. However, building a dynamic profile is a challenging problem to maintain accuracy of future predictions and allow speed calculations. In this research, the static profile is used to simplify the main technique testing.

Usually, users’ profiles contain users’ skills vector. If the same skill has different indices in multiple node vectors, this can lead to comparison problems. Another similar problem is to have the same skill with different names in the network. This problem needs to have a semantic ontology of used terms and alternative meanings to treat different names for the same skill as same concept. A similar case can happen in multilingual environment where the same skill has different words from one language to another.

2.2. Social networks over MANETs

Although there is a lot of work on both of social networks and ad-hoc networks individually, a small set of researches combine both of them. There are many challenges when social networks are implemented over MANET compared to Internet as introduced in the previous section. The target of the work proposed by Mavromoustakis (2012) was resource sharing problem. He considered the probabilistic social interactions in order to assign available resources to communicating nodes according to a combined mobility model and the users’ social relations. He proposed the using of Fractional Random Walk (FRW) on weight graphs to spread updates. He extended his work to study social networking in vehicular ad hoc networks in Mavromoustakis (2012). Moreover, Li and Khan (2009) proposed another mobile social network model that is based on semantic analysis to detect similarity in ontology. They proposed a semantic-aware user profile model and a related concept model used for routing between mobile nodes. Li et al. (2011) have suggested a similar technique to enhance data sharing in large-scale Peer Data Management Systems (PDMS) based on semantic and optimized query evaluation and limiting the lifetime of queries by using Time-To-Life (TTL) technique. A similar technique has been suggested in Li et al. (2012) using Semantic Zone Discovery Protocol (SZDP) and semantic path discovery. A similar model has been suggested in Konstan and Riedl (2012) to build a social group middleware with an Application Programming Interface (API) in C++ , C#, Windows Mobile. Tian has also suggested a similar middleware designed for marketing in social networks (Tian, 2012).

In the field of recommender systems, most of researches have been made; assuming a general nature of communication between users and some sort of time stability of users’ existence in the target network. One of the most effective techniques in recommender systems is to use grouping for users. The hierarchical clustering is used in Pham et al. (2011) to enhance...
recommendation. A similar technique using subgroups was used in Xu et al. (2012) to make social recommender systems. In both researches, grouping was made in application logic, and did not depend on the communication protocol of the used network.

The problem of group formation was discussed in several researches in traditional social networks. A recent research has been applied on developers’ networks by Surian et al. (2011). In that research, a dependency graph was built to represent the relation of Developer-Project-Property (DPP) in open source environment. Each developer’s history is analyzed to obtain the related projects that the developer worked in and its characteristics. The main process of that work must have a training process to compute DPP graph and some pre-computed matrices. In fact, their work is great but not suitable to mobile ad-hoc environment which is frequently changing. Their work assume central server node which might not be available in MANET. A similar research has been made by Brauer and Schmidt (2012) but in e-learning social networks. In that research, the dependency graphs between the users, groups, topics and contents are studied. The group formation strategy was built on the following three similarity factors: common learning style, high score in knowledge ranking and low distance on social networks.

Recently, some studies are made about the mobile social networks. Bulut and Szymanski (2012) have presented a new metric of friendship quality between mobile nodes. They introduced the Friendship-Based Routing algorithm. They also considered the periodicity of relationships per day. The Friendship-Based Routing depends on finding both direct and indirect relationships. They introduced a metric called Social Pressure Metric (SPM) that considers two nodes to be friends when their connections have three features: frequent, long and regular. SPM handles direct communications and another variation of it called relative-SPM (RSPM) handles indirect communications. Bulut and Szymanski have studied the effects of day periods on connection frequency between nodes and adapted their algorithm to that factor. Moreover, Li and Li (2011) have studied the trust models in MANET and proposed a model named MobileTrust, to establish decentralized and reliable trust relationships between mobile ad hoc social network users. They studied some scenarios with users who are experienced with the network and users who are unfamiliar with the environment. Their trust models cover some important factors of trust relationships in social networks, like the similarity of user profile, reputation, and history of friendship.

2.3. Overview of zone routing protocol (ZRP)

Zone routing protocol (ZRP) is a hybrid routing protocol. The section gives more details about it since it would be a base for our work. According to Garg et al. (2012) and Haas and Pearlman (2001), ZRP divides the entire network into zones based on a given threshold distance M. Each zone applies a proactive routing inside it using the proactive Inter-zone Routing Protocol (IARP). Routing among zones is done by another protocol named Inter-zone Routing Protocol (IERP). ZRP uses a helper protocol named Border-cast Resolution Protocol (BRP) (Pandey and Swaroop, 2011). BRP communicates to the nodes that lay of the border of the zone to initiate the IERP reactive searching. Another variation of ZRP named Multicast Zone Routing Protocol (MZRP) had been developed by Zhang and Jacob (2003). MZRP is different than ZRP since it is using on-demand flooding and multicast tree construction rather than unicast IERP. Other algorithms of routing named Shared Tree-Based Multicast Protocols construct a single tree that is rooted at a central control point called the Rendezvous Point (RP) (Pandey and Swaroop, 2011).

3. The proposed technique for MANET team formation

3.1. The proposed technique

Our suggested technique can be classified as a hybrid ZRP multicast mesh-based model. Generally, Neighbor mobile nodes in a MANET try to exchange profiles with each other. Some nodes may welcome to start exchanging process while other nodes may ignore or reject that. After the exchanging process is done in that MANET, a process of collaborative filtering and matching is done to suggest suitable teams according to user required skills to a certain task. The network is divided into zones based on geographic locations. More details about that step will be given in the next section.

Based on the illustration of Brauer and C. Schmidt (2012), Zhang et al. (2005) and Cacciapuoti et al. (2011), there are three main strategies for finding experts in traditional social networks: (a) Breath First Search: Breath First Search (BFS) or (flooding) Starting from the initiator node, BFS would probably find the nearest candidates, (b) Random Walk: Random Walks (RW) traverse the social network by random paths. Random Walk’s distance to the start node increases very fast. And (c) Best Connected Search: Best Connected Search (BCS) performs well at networks with a power-law distribution of nodal degrees. The strategy is to select nodes by the number of neighbors.

The definition of power law according to Newman (2001) is:

“If one makes a similar plot for the number of connections (or ‘links’) z to or from sites on the World Wide Web, the resulting distribution closely follows a power law: \( P(z) \approx z^{-\tau} \), where \( \tau \) is a constant exponent with (in that case) a value of about 2.5.”

The application of power law in that case considers the degree of connections as a measure of activeness of users. This cannot be always applied in the MANET since there are physical constraints on logical connections of users that can separate a user from some other desired users. The power law needs a large time frame of communications in a stable network to build new logical relations between users. Since the MANETs are not proofed to have the power law in the Ad-Hoc relationships between users, that type of search is ignored. Since Random walk can lead to a far distant nodes or lost paths in searches, that type of search is also ignored. In this research, the suggested technique is compared to the first type of search (Flooding).

There are several variations of flooding to select paths for optimization (Boukerche et al., 2011). A simple flooding variation is Adaptive Location Aided Mobile Ad Hoc Network Routing (ALARM) (Boleng and Camp, 2004). ALARM uses link duration between two nodes as a metric for selecting forwarding direction. Instead of forwarding
messages to all neighbors in range, ALARM algorithm forwards message only to neighbors with link duration values greater than a given threshold. The ALARM algorithm decreases flooding messages depending on the used life threshold value which is arbitrary. ALARM has an advantage of other flooding variation that it does not require global state construction algorithm.

Although MANET can have small interaction time periods in general, the Team-Formation problem nature assumes a relative stability for short or medium term collaboration time. The suggested technique is not useful for on-the-fly users in MANET with quick departure nature. For this reason the problem of short time interaction in some networks is beyond the scope of this research for simplicity. In practice, when a node with a requested skill leaves network, some other alternative nodes can exist with the same skill in different paths. However, sometimes there is a possibility that some skills are rare and thus the departure of those skill holders can be a problem. The same is applied on intermediate nodes on paths from source node to another destination node. Having alternative paths can overcome this problem in the majority of cases. A detailed study for that point is beyond the scope of this research.

3.2. The proposed technique phases

The main four phases in the suggested technique are as follows:

1. **Profile generation and update phase**: its role is to estimate the mobile node user’s preferences and skills and periodically update them.

2. **Location-based zones building phase**: its role is to periodically sense the neighbor mobile nodes and exchange location details with them to build zones of nodes.

3. **Skill-based member searching and recommendation phase**: its role is to listen for requests of group building with a specific set of skills. The searching starts after that in networks for mobile nodes whose users’ profiles match the required skills. The search results can be full match, partial match, failure or timeout.

4. **Groups aggregation phase**: its role is to receive match results and rank users based on their experience level.

The first and fourth phases are implemented identically in both flooding and the suggested techniques for simplicity. In this research, the static profiling is used in the first phase. Of course the dynamic profiling is better than the static profiling but this point is beyond the scope of this research. The fourth phase can be implemented as a simple sorting routine based on the geographic distance between skill seeker and skill provider. This phase can be enhanced later to include other factors. The current research is focusing on the second and third phases as explained in the following sub-sections.

### 3.2.1. Model description

A brief discussion will be introduced about the investigated model for the suggested technique. Assuming that the radius for the wireless connection range for any mobile node $k$ is $R_k$, and the distance between two mobile nodes $j$ and $k$ is $Dist(j,k)$, then $j$ and $k$ visibility to each other can be defined by the boolean function in (Eq. (1)) applied in bi-directional connections.

$$\text{Visibility}(j,k) = \begin{cases} \text{Visible}, & \text{Dist}(j,K) \leq R_k \text{ and } \text{Dist}(j,K) \leq R_j \\ \text{NotVisible}, & \text{otherwise} \end{cases}$$ (1)

Moreover, assuming that network $N$ contains $n$ of nodes, where $N = \{j_0,j_1,\ldots,j_n\}$ and $n > 0$, then any zone $Z$ that contains a set of nodes in the network must satisfy that $Z \subseteq N$. The zone geographical dimensions can be defined by an input parameter $D_z$ that is updated after applying zone formation algorithm in Algorithm 2. $D_z$ is defined as illustrated in (Eq. (2)).

$$D_z = \text{Max}(\text{Dist}(j,k)), \forall j,k \in Z$$ (2)

The skill set for each node $j$ can be defined as $S_j$. Each zone $Z$ has a union skill set which can be defined as $\cup S_j$ where $j \in Z$.

All mobile nodes should act as a peer-to-peer to other nodes. Thus, based on the most important components, the suggested design has two main models: node model and message model. The node model for each mobile node $j$ has the illustrated following parameters in Table 1.

The synchronization time mechanism is a simple form of time stamping form in the work (Lamport, 1978). When node $j$ receives a message from $k$, then $T_j$ can be calculated using (Eq. (3)).

$$T_j = \text{Max}(T_j, T_k) + 1.$$ (3)

The $W_{max_j}$ can be calculated using (Eq. (4)).

$$W_{max_j} = T_j + \text{TTL}_j.$$ (4)

The $Z_{sj}$ value can be calculated as the illustrated in (Eq. (5)). Sometimes, same skill can have different names at multiple nodes. However, this semantic analysis problem is out of the current research scope. The proposed technique assumes that each skill has a unique name.

$$Z_{sj} = \cup S_j, \forall j \in Z.$$ (5)

The message model has the following parameters illustrated in Table 2.

Matching a request vector $On$ that contains a set of required skills and node $j$ skills $S_j$ can be defined as literal function $M(On, S_j)$ as follows in (Eq. (6)).

$$M(V, S_j) = \begin{cases} \text{FullMatch}, & \text{if } S_j = On \\ \text{PartialMatch}, & \text{if } S_j \subseteq On \\ \text{FailedMatch}, & \text{if } S_j \cap On = \emptyset \end{cases}$$ (6)

A local temporary variable set $newUn$ is calculated before updating the value of $Un$. After each match the value of $Un$ is updated by (Eq. (8)), and the Mat value is updated by (Eq. (9)). The reason for not directly modifying $Un$ value in (Eq. (7)) is to have both old and new values of $Un$ to test whether there is a new match or not to be able to apply (Eq. (8)).

$$newUn = Un - S_j$$ (7)

$$\text{Mat} = \begin{cases} \text{Mat} + j, newUn! = Un \\ \text{Mat}, & \text{otherwise} \end{cases}$$ (8)

$$Un = newUn$$ (9)
When $Un$ is empty, then no more requests are needed and this can affect the value of $Mt$ as follows in the literal function in (Eq. (10)).

$$Mt = \begin{cases} 
\text{acknowledging}, & Un = \emptyset \\
\text{searching, otherwise} & 
\end{cases}$$ (10)

If full match occurred at any node in searching path, then it returns its response to requester. Otherwise, a forwarded message is transferred to available neighbors in range that are not in the Routing node vector $Rt$. There are three cases that prevent forwarding messages at any node in general:

i) Full Match of skills ($Un$ is empty)

ii) No more neighbors that do not exist in message routing node vector.

iii) Expired Time-To-Live ($TTL = 0$)

When any requester $k$ has a waiting message with $W_{max} > k$ condition achieved, then this message is assumed to fail due to time out consideration. When Node $k$ receives a match response message $x$, then the simple aggregation function in Algorithm 1 is applied.

### 3.2.2. Simple introductory example

A simple example is presented in Fig. 1 to demonstrate the main process of ZRP based skill searching. The example assumes 9 mobile nodes named: A, B, C, D, E, F, J and H and 4 sample skills named: w, x, y and z. Each mobile node – except node A - has its skills below its name in two braces. There are 4 zones with union sets of skills in braces below the name of each zone. Suppose the node A is requesting skills x, y and z. Let us first start with the brute force (flooding) technique to perform that search: A will send a broadcast message to

---

**Table 1** The proposed node model for each mobile node $j$.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ID_j$</td>
<td>Node unique identifier (ID)</td>
</tr>
<tr>
<td>$Loc_j$</td>
<td>Geographic location defined in three dimensional space $(x, y, z)$</td>
</tr>
<tr>
<td>$T_j$</td>
<td>Logical timing value which is synchronized with other zone clocks to be unified at all network nodes</td>
</tr>
<tr>
<td>$R_j$</td>
<td>Communication range radius</td>
</tr>
<tr>
<td>$S_j$</td>
<td>Skill vector from profile</td>
</tr>
<tr>
<td>$TTL_j$</td>
<td>Default TTL that is given for any future request message as Time-To-Live value for it</td>
</tr>
<tr>
<td>$Nd_j$</td>
<td>Requirement vector</td>
</tr>
<tr>
<td>$ZID_j$</td>
<td>Current zone identifier which is obtained from zone building phase. As explained later in Algorithm 2</td>
</tr>
<tr>
<td>$Buff_j$</td>
<td>On-demand request buffer that contains the requests sent by the holder node and still waiting for responses from some nodes</td>
</tr>
<tr>
<td>$Res_j$</td>
<td>Maximum waiting time that is an integer value to avoid waiting for ever for lost responses due to network traffic or vanishing of some intermediate nodes in some routing paths</td>
</tr>
<tr>
<td>$W_{maxj}$</td>
<td>Zone skill hash table that contains each skill name as a key and a list of holder nodes of that skill in the current zone</td>
</tr>
</tbody>
</table>

---

**Table 2** The proposed message model parameters.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$IDs$</td>
<td>Source node ID (requester ID)</td>
</tr>
<tr>
<td>$TTL$</td>
<td>Current time-to-live (TTL) value of the message. It is calculated as the number of hops passed on in the current path of message</td>
</tr>
<tr>
<td>$Mt$</td>
<td>Message type which has two possible literal values: searching (request) when message has a valid TTL and is still searching for at least one needed skill, and acknowledging (response) when returning back to requester with partial or full successful match results</td>
</tr>
<tr>
<td>$Un$</td>
<td>Unsatisfied needed skill vector so far: starts with an empty vector and when reaching any node in path that contains at least one unfound skill so far, that node is added to the vector</td>
</tr>
<tr>
<td>$On$</td>
<td>Original needed skill vector which contains required skill vector issued by requester node</td>
</tr>
<tr>
<td>$Mat$</td>
<td>Matching node vector which is a set of nodes that has full or partial match with original needed skill sector</td>
</tr>
<tr>
<td>$Rt$</td>
<td>Routing node vector that contains nodes that has received the current message instance</td>
</tr>
<tr>
<td>$Rts$</td>
<td>Request time stamp (RTS) which is logical starting time of requester start of seeking skills. That value is used as a key to this message at the requester on-demand requests buffer</td>
</tr>
<tr>
<td>$Rbl$</td>
<td>Return back path list which is always the inverse of the current routing path. It can be used to optimize returning to requester in case of stable intermediate node state instead of re-routing back to requester all times</td>
</tr>
<tr>
<td>$Brb$</td>
<td>Broken Return Back list to indicate that return path is still valid or not. In case of having at least a missing node in returning path, a new routing procedure is applied by using ZRP routing. In that case, the Brb is set to true to prevent further usage of return path reverse vector</td>
</tr>
<tr>
<td>$Ts$</td>
<td>Sending node logical time value at the moment of message transmission start</td>
</tr>
</tbody>
</table>

---

**Algorithm 1** Aggregation function at requester nodes.

**Fun Aggregate ($k$, $x$):**

**Inputs:**

- $k$: Requester node ID.
- $x$: Received response message at node $k$.

**Outputs:**

- The modified $Res_k[x, Rt]$ and $Buff_k$ after receiving message $x$.

**Steps:**

1. $Buff_k = x$
2. If (sizeof $(x, Mat) > 0$) then $Res_k[x, Rt] += x, Mat$.
3. Else drop $x$.

Figure 1 Small simplified zone team formation example.
node B to search for target skills. Node B performs a match to skill X and broadcasts the request to nodes C and H. Node C forwards request to Node D with extra match to Y. Node H forwards request to node J which does not have a match. Node D forwards the request to node E which forwards the request to nodes G and F. Node G has a match for skill z. Node F will forward the request to node G (duplication). In that scenario there are two problems: overflowing the network with 9 request messages to get the required match, and duplication request sent to node G.

Let us node apply the suggested ZRP based searching. Node A sends a message to node B that has a match to skill x. Node B asks nodes C and H for their zone skills. As a result, Zone 4 will be ignored totally and only zone 2 will have a forwarded request. After a match in node D with skill y, zone 3 is asked for its union set which indicates that it has required remaining skill z. As explained earlier, this union has a supporting hash table with holder nodes of each skill in the zone. So, a response is generated by any node in zone 3 that node G has a match of skill z. In that scenario, there are only 5 request messages with no duplication of requests to same node.

3.2.3. Location-based zone building phase
A modification to ZRP aims to combine the semantic nature of exchanged users’ profiles and also to keep in mind the distance delay of communication between nodes. In this suggested modification, the main steps of location-based zone building algorithm are listed in Algorithm 2.

This algorithm is similar to k-means algorithm with the difference of last 2 steps in our suggested technique which tries to avoid small zones that would be easier to break or disappear due to mobility nature of nodes. After the initial a proactive periodic recalculations are made to maintain the state of connection inside each zone. Each zone would have 2 skill sets: an averaged location set and a union set. The average location set is used to maintain zone membership based on geographic location. The union set is used for inter-zone collaboration. The problems of skill indexing to keep same ordering in all network mobile nodes are beyond the scope of the current research.

The routing table formation technique in standard reactive protocols would be used but with modification. The modification in routing table is to make a skill routing table. In that skill routing table, routing would be made to target skill not target zone. In order to make that, each zone floods its union set of skills to its neighbor zones. For each zone i, when a neighbor set of skills is received Sj, the redundant skills are ignored. The skills in zone J that do not exist in zone i are put in skill routing table with the next hop value j, which means that in order to get that skill for any member in zone i, a collaboration has to be with zone j.

3.2.4. Skill-Based Members Searching and Recommendation Phase
The third phase of searching for required skills is based on modified ZRP. The suggested technique can be summarized using the UML activity diagram in Fig. 2. There are two main roles for each node in the mobile network: the requester who initiates the team searching process, and any other neighbor nodes (either direct or indirect neighbors) which can respond to requests for team formation. The two roles are displayed as two swimlanes in Fig. 2. The requester manually inputs the required skills to search for.

An important suggested step can be named Farthest Forward (FF). FF aims to forward messages to the farthest nodes in the current zone when no direct reachable nodes exit having a different zone. This step is a modification in original ZRP protocol. The original ZRP uses BRP as explained in the introduction section to directly discover zone border nodes and transmit messages for them as gateways to other zones. The suggested technique aims to simplify the implementation of team formation process by using FF instead of BRP. The effect of FF compared to BRP is beyond the research scope of this research since it is a detailed routing research point. The suggested technique needs extra type of messages to: (a) build zones as in Algorithm 2, (b) build zones hash tables and (c) synchronize logical timing or neighbor nodes.

4. Simulation and performance results
4.1. Simulation environment
Since we did not have access to a real testing dataset, we have to use a simulation environment. There are many standard environments for networks like NS-2, NS-3, Omnet++, Jist-Swan, DARS and QualNet. Some of them are free to use and have basic MANET simulation tools. We have tested many of those environments. We have chosen a simulation environment named NetLogo. NetLogo was free and suitable to our available resources. NetLogo has a simple scripting language that has a good learning curve.

NetLogo does not have a built in MANET library. So, we had to build a simple custom for MANET in NetLogo. We have contributed the basic MANET and flooding model to NetLogo Community then we have added our suggested ZRP based model in later version of that model. A screen shot is presented in Fig. 3. The simulation inputs are (some are shown in Fig. 3):

1) Width and Height of simulation screen using logical blocks (patches).
2) Count of mobile nodes (Drawn as small arrows in Fig. 3).
3) Count of requesters.
4) Request TTL value (Using count of path nodes as a measuring unit)
5) Maximum range of connection area for each node to limit the random function of initiating the radius of connection (Drawn as circles around nodes in Fig. 3).

---

1 http://www.isi.edu/nsnam/ns/.
2 http://www.nsnam.org/.
3 www.omnetpp.org/.
4 www.opnet.com/.
5 http://jist.ece.cornell.edu/.
6 https://sourceforge.net/projects/dars/.
7 https://sourceforge.net/projects/dars/.
9 http://ccl.northwestern.edu/netlogo/index.shtml.
6) Zone width and height in patches. For simplicity zones are assumed to be square shaped in simulation. Zone borders are marked as gray dots in Fig. 3.
7) The logical timing rate (Measured in Ticks/s).
8) Enable/Disable node movement during simulation, and the movement rate.
9) Technique to apply (Flooding or ZRP based).

Simulations are performed on randomly generated networks and tested on a desktop machine with core i3 processor and 2 GB RAM. During simulation each node periodically tries to fetch neighbor nodes. Nodes without any neighbors in range will remain with default arrow shapes and nodes with at least one neighbor will have star shapes as in Fig. 8 later. Ranges of nodes have a maximum value but they are not all equal. The simple Random Waypoint Mobility model (Divecha and Abraham, 2007) was applied with some simplifications to let each node choose random directions periodically and move to it.

4.2. Results

Since there are multiple simulation parameters, several experiments are made to study the effect of changing one parameter in team searching using the same required skill sets in requesters in Flooding and ZRP-based team formation techniques. Experiments are made first on some small networks then some larger networks. Fig. 4 presents the comparison between the two techniques with fixing all parameters except the network size. Multiple tests are made using variable number of nodes.
and the total search finish time is measured to compare performance of both techniques in MANET. In all cases the ZRP-based technique finish time is better than the flooding technique. The positioning and selected requesters are chosen randomly in each simulation start. This can explain the difference between consumed times in networks with size 40 to the network with size 45.

Table 3 presents another comparison based on changing the requester count while fixing the other parameters of simulation. The search finish time of the ZRP-based technique is also better in all cases. Using 15 requesters lead to a remarkable performance speedup. The ratio of speeds in that case is 14% of required time in the flooding technique. Again, the positioning and selected requesters are chosen randomly in each simulation start. This can explain the difference in times in cases of having 7, 9 and 11 requesters for example.

Table 3 also presents a different way of testing. Instead of changing only a single factor, the experiment is done using changes in 2 combined factors: network size and count of concurrent requests while fixing the TTL value of 5 nodes in path before expiration. The periodic messages in ZRP-based technique are put in count with other types of search and acknowledge messages. The interesting result is the flooding based techniques more efficient in small networks with small count of requesters (1 or 3) since it does not need any periodic messages for zone formation or timing unlike the ZRP-based technique. But in medium sized networks with increased number of requesters, the flooding performance is worse than the suggested ZRP-based technique.

Table 3 adds TTL change into count. The total exchanged messages in case of ZRP-based technique are remarkably less than its corresponding results in flooding techniques. ZRP-based messages are counted in tens or hundreds (348, 213 and 172 messages) whereas the flooding count of messages is counted in thousands or tens of thousands (51,220, 4212 and 29,118 messages). Moreover, Table 3 presents some detailed results of simulations to study the load on mobile nodes. The technique name is shorthanded to “F” for “Flooding” and “Z” for “ZRP-Based”. Each experiment in flooding is presented in a gray row flow by the same configuration experiment in ZRP-based experiment case in the next row. The results in that table confirm that the ZRP-based technique is much better than flooding in performance and network overhead. It deserves to point out that the synchronization messages needed for the ZRP-based technique are never needed in the flooding technique.

Previous tests compared the general form of the flooding technique. As mentioned earlier, the ALARM technique is a variation to flooding to reduce multicasting paths. The ALARM technique was used in the experiments to enhance the flooding technique compared to ZRP. In the following tests, three techniques were compared: Flooding, ALARM, and ZRP. The first comparison was about performance of the three techniques when operating on the same configurations of MANET simulator. The comparison results are presented in Fig. 5. The ALARM connection duration lifetime values were arbitrarily chosen in the range between 5 and 16 simulation ticks.

To focus on the space conservation between the two techniques, Table 4 shows the great difference in total exchanged message in all over the network samples. Increasing network parameters in values increased the load of flooding and ALARM techniques rapidly compared to ZRP. The experiments had shown that the average buffered messages in each node were: 25.6905238, 31.88827922, and 0.255214286 for ALARM, Flooding, and ZRP, respectively.

To compare the successfulness of each team formation technique, Fig. 6 compares the count of successful matches in the three techniques. Apparently, ALARM and flooding techniques are better than our suggested ZRP-based technique in most cases. This phenomenon shall be explained in the discussion section.

To focus on the effect of fast node movement in MANETs, some experiments are made on medium sized networks with 20–30 nodes having 5–7 concurrent requests. The nodes were adjusted to move too frequently with 1 patch per each logical time unit (tick). Some requesters had too small connection range which leads to having a few or no neighbors to provide skills. In order to have worst case tests, each requester had one non-existing skill in the entire network. This worst case testing is useful to force the tested techniques to keep searching as long as possible. Fig. 7 compares the partial successes counts...
for the same samples listed in Fig. 6. This comparison shows that the ALARM and flooding technique success rate is more than our suggested technique.

The simulation networks in Fig. 6 and Fig. 7 were small or medium sized and can be called sparse networks. The meaning of sparse networks is to have a few nodes per geographical area as shown in Fig. 3. However, testing in dense networks with a lot of nodes per geographical area as shown in Fig. 8 shows the importance of response time of searching. The flooding technique can overtake the whole network in order to get a high success rate. Nodes in dense networks also have a high probability of having many neighbors unlike sparse networks.

When testing at dense networks and limiting search time to be several seconds, the flooding technique always failed to form teams in a proper time. This testing had several samples within the first 20 s of simulation using MANET with 7 requesters, 200 mobile nodes, TTL = 7 nodes and Max Connection Range 20 patches. The mean value of ZRP successful requests of the 7 requesters was 2.166666667 results within 20 s of simulation using the same configurations. The corresponding flooding technique (and similarly ALARM) did not make any results for any sample during the first 20 s of simulation.

### Table 3

<table>
<thead>
<tr>
<th>Technique</th>
<th>Configurations</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Requesters count</td>
<td>TTL (count of nodes in path)</td>
</tr>
<tr>
<td>Flooding</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>ZRP</td>
<td>1.51</td>
<td>28</td>
</tr>
<tr>
<td>Flooding</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>ZRP</td>
<td>1.68</td>
<td>54</td>
</tr>
<tr>
<td>Flooding</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>ZRP</td>
<td>2.00</td>
<td>81</td>
</tr>
<tr>
<td>Flooding</td>
<td>21</td>
<td>5</td>
</tr>
<tr>
<td>ZRP</td>
<td>2.43</td>
<td>98</td>
</tr>
<tr>
<td>Flooding</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>ZRP</td>
<td>4.84</td>
<td>348</td>
</tr>
<tr>
<td>Flooding</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>ZRP</td>
<td>4.85</td>
<td>213</td>
</tr>
<tr>
<td>Flooding</td>
<td>7</td>
<td>30</td>
</tr>
<tr>
<td>ZRP</td>
<td>5.60</td>
<td>172</td>
</tr>
</tbody>
</table>

The presented results are rounded to two decimal place values.

When testing at dense networks and limiting search time to be several seconds, the flooding technique always failed to form teams in a proper time. This testing had several samples within the first 20 s of simulation using MANET with 7 requesters, 200 mobile nodes, TTL = 7 nodes and Max Connection Range 20 patches. The mean value of ZRP successful requests of the 7 requesters was 2.166666667 results within 20 s of simulation using the same configurations. The corresponding flooding technique (and similarly ALARM) did not make any results for any sample during the first 20 s of simulation.

### Figure 5

![Performance time comparison of three team formation techniques.](image)

In experiments with a lot of requesters, there was a large burst of exchanged messages in flooding techniques. This is clear from Table 3 when the increasing of requesters count which leads to a long response time at sizes 7 and 9 requesters compared to ZRP-based. The same performance result can be concluded when TTL value is fixed with a simultaneous increase in network size and requester count in small and medium networks. Even when the network size is small but having a high value of TTL as in Table 3, the performance drawback of the flooding technique has reached about more than 30 doubles of response time in ZRP-based comparatives in some cases. Those results also confirm the large gap and difference when increasing TTL, network size and requester count at the same time in small networks. Although the ALARM technique has some enhancement to the basic flooding technique, it still has the problem of local view in selective path selection which sometimes results in losing of requests. The ALARM has a similar problem to depth first search selection, which is considered a possible future study.
Although the suggested technique adds a new type of messages to synchronize time and zones as shown in Table 3, the performance is enhanced compared to flooding cases. The total exchanged messages in the flooding technique can be measured by thousands of messages whereas the comparatives in the ZRP-based technique. The flooding technique requirements are too high for P2P mobile networks. From Table 3 and Table 4, the average required buffer size for each mobile node in flooding (and similarly ALARM) can be over 10 doubles of the comparative ZRP-based cases in small and medium networks.

According to power measurements in mobile nodes in experiments with durations of five minutes (Agarwal et al., 2007), the mobile node consumption of power when no radio communications are made is 15.688 mW (milliWatt) whereas the Global System for Mobile Communications (GSM) mode consumes at least 27.33 mW. The Wi-Fi communication consumes from 441.82 mW to 1113.811 mW. From those power measures, it is clear that computation processing consumes less power than message transmission power. So, reducing the count of exchanged messages leads to reduction of power consumption.

4.3.3. Match success in stable and dynamic networks
Skill matching has been studied in two types of networks; stable networks with no moving nodes, and dynamic networks which have frequent movement of most/all of its nodes. The suggested technique has partial success compared to flooding and ALARM techniques to get all possible team recommendations in stable networks and in dynamic networks as in Figs. 6 and 7. Testing in small and medium sized sparse networks gives flooding and ALARM techniques the advantage. This situation is not practical when testing at dense large graph for real-time systems – even in stable networks. The suggested technique works better at these networks in a way that suits limited power and memory devices unlike the flooding and ALARM techniques which can need impractical requirements to achieve the ideal successful matches. Even when the suggested has partial successful recommendations, it does not consume the hardware resources of mobile devices Table 4. Although flooding and ALARM techniques are theoretically better in successful matching ratio than the suggested technique, it is not practical in time, power, processing, and memory space constraints.

5. Conclusion and future work
In this paper, we proposed a recommender system that is based on ZRP protocol to target a relatively new problem in the
MANET. The problem of team formation in MANET has time and hardware constraints that could be satisfied in a reasonable way. This technique could enable collaboration and team recommendation in mobile social networks. It was designed to be a peer-to-peer technique with no need to a centralized server. We performed extensive simulations to measure the performance of the system in terms of main three comparison categories: efficiency, memory space and successful matching rates. The constraints of time, power, processing, and memory space were well handled by the suggested technique.

However, the accuracy of the technique needs more work. The relation between users and skills needs to be measured in order to distinguish normal users from advanced ones. Some semantic analysis can be added later to enhance recommendation. The ranking of multiple skill providers can be improved to include other factors like ratings of other users to skill providers, the current power level, the length of routing path and the computing power of mobile nodes. The prediction of movement direction of nodes can enhance selection process of Farthest Forward (FF). A comparison between FF and BRP can be a future work to know difference in performance and accuracy between both of them. Future work can involve studying of the application of the suggested ZRP-based technique on software production and component assembly.

Acknowledgement

Many thanks to Ph.D Amira Rezk and Ph.D Mervat Abu-Elkheir (Mansoura University, Computer and Information Systems faculty), for their help in paper formatting and phrasing.

References


