Computer Networks 75 (2014) 177–191

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

RCA: Efficient connected dominated clustering algorithm for mobile ad hoc networks

Dina S.M. Hassan *, Hossam M.A. Fahmy, Ayman M. Bahaa-ElDin

Faculty of Engineering, Ain Shams University, Department of Computer Engineering and Systems, Cairo, Egypt

ARTICLE INFO

Article history:
Received 3 May 2012
Received in revised form 22 September 2014
Accepted 1 October 2014
Available online 12 October 2014

Keywords:
Connected Dominated Set (CDS)
Clustering algorithms
Mobile ad hoc networks
Ring Clustering Algorithm (RCA)

ABSTRACT

Clustering of mobile ad hoc networks is a largely growing field. The perceived benefits of clustering are comprehensively analyzed in open literature. This paper considers the development of a new distributed connected-dominated-set clustering algorithm called Ring Clustering Algorithm (RCA). RCA is a heuristic algorithm that has three phases: ring-formation phase, members-joining phase and CDS-nodes selection phase. In the ring-formation phase, each ring consists of three ring-nodes. The ring is formed if it has the highest priority. The priority of the ring is based on the total ring-degree rather than the individual node-degree. The degree of a ring is the number of neighbors that the three ring-nodes have all together without repetition. Nodes that cannot form rings join neighboring rings as members in the members-joining phase. In the CDS nodes selection phase, the decision is made for a node to remain or leave the CDS. This paper presents the proof that the maximum number of rings that can be formed by RCA in any disk area equals the maximum number of independent nodes that create non-overlapping circles in a corresponding area. This allowed RCA to achieve the lowest fixed approximation ratio (5.146). Moreover, RCA has $O(n)$ for both time and message complexities. Thus, RCA algorithm outperforms the current-best CDS algorithms that are investigated in this paper.

1. Introduction

Mobile Ad hoc Networks (MANETs) have attracted a growing interest in recent years. The motivation behind this is the aptitude to connect people anywhere and anytime without any type of infrastructure, except for the mobile units themselves. Therefore, MANETs can be termed as instantaneous, temporal, economical and "on-request" networks [1–3]. The basic idea, which allowed that, was the ability to use the overlapped transmission ranges of adjacent mobile units to relay a message from a node to another till that message reaches its destination node. In other words, a mobile unit in such a network acts as a node and as a router as well. By exploiting this facility, the formation of wireless networks becomes possible in areas where conventional infrastructure cannot be established as in battlefields, during catastrophic events, and in uninhabited areas for scientific researches such as in deserts or at sea.

However, packets relayed through the network may suffer from collisions, which introduce retransmissions and increase end-to-end delays. As network size grows, the negative effect of retransmissions over nodes energy level and network throughput become increasingly evident. Hence, clustering the network to limit the number of nodes that perform inter-cluster communication represents a practical solution to maintain an acceptable performance level and optimize the use of the network scarce resources. One of the main methods used to cluster ad hoc networks is to form a Connected Dominated Set (CDS). CDS clustering creates a virtual backbone, which is a chain of connected nodes that are responsible for

* Corresponding author.
E-mail address: dina_samir1@yahoo.com (D.S.M. Hassan).

http://dx.doi.org/10.1016/j.comnet.2014.10.010
1389-1286/

handling communication requests in the network. Reducing the size of the CDS is a main target in CDS clustering; however, computing the minimum CDS is an NP-hard problem [4,5]. Accordingly, most CDS clustering algorithms are heuristic algorithms.

This paper proposes a novel greedy heuristic CDS clustering algorithm called Ring Clustering Algorithm (RCA). The main objective of RCA is to create the maximum number of rings. Each ring consists of three ring-nodes. Ring-nodes from neighboring rings can connect to each other directly or through ring-members. A preliminary idea of this work was previously published in [6]. Section 2 presents a background about CDS clustering algorithms. Algorithm overview is presented in Section 3. Section 4 discusses RCA performance analysis. Simulation results are discussed in Section 5 and finally, Section 6 concludes this paper.

2. Related work

In backbone or dominated set clustering, network nodes can only communicate through a set of nodes called the connected dominated set CDS. The CDS has two node types, independent nodes, also called dominators, and connectors, which link dominators together. These dominators are selected so that the hop distance between any two dominators is at least two hops. Dominators and connectors form the backbone of the network. However, as mentioned in the previous section, finding the minimum CDS (MCDS) was proved to be an NP-hard problem [4,5]. Therefore, CDS clustering algorithms tend to determine CDS nodes heuristically.

Accordingly, the produced CDS will always be larger than the minimum CDS (MCDS). In [7], it was proved that the size of the Maximum Independent Set $(\text{MIS}) \leqslant 4$ MCDS + 1. However, in [8], the authors derived an inequality, in which they proved that MIS $\leqslant$ 3.8 MCDS + 1.2. A tighter expression was lately presented in [9], it states that MIS $\leqslant 3\frac{2}{3}$ MCDS + 1. Finally, the best-known relation, which is MIS $\leqslant$ 3.4306 MCDS + 4.8185, was derived in [10]. The approximation ratio was used to reflect the size of the generated CDS over the MCDS. Several methods were used to identify the independent set and to form the CDS [11–13]. One of the well-known backbone algorithms is Rule-k algorithm [12], which comes as an enhancement for Rule-1 and Rule-2 algorithms [14,15]. Rule-k algorithm has no fixed approximation ratio; however, it was proved in [12] that an upper bound for the approximation ratio has a very small probability of being infinitely large. The algorithm in [16] is an enhancement for Rule-k algorithm. It uses nodes degree rather than ID for assigning CDS nodes. The message optimal CDS algorithm [17] has an approximation ratio bounded by 192 while the CDS algorithm in [18] has a fixed approximation ratio, which equals 44. Algorithms in [19,20] focus on reducing routing costs rather than minimizing the total CDS size. However, these algorithms have no fixed approximation ratios. Moreover, in [21], a special case of the CDS is studied which is the Shortest Path CDS (SPCDS). In which, all intermediate nodes inside every pairwise shortest path is included in the CDS. Although finding the minimum SPCDS is solvable in polynomial time, the approximation ratio of the proposed algorithm is not fixed.

The performance of RCA is compared to four recently published CDS clustering algorithms [7,9,22,23] because they all have low and fixed approximation ratios. The first algorithm is Zone clustering algorithm [22]. In Zone algorithm, the dominator node is the node with the highest priority in its neighborhood. The priority of nodes can be determined according to various factors such as node ID, degree, mobility pattern or energy level. The algorithm was executed in two versions: the lowest ID node (Zone-Min-ID) and the highest degree node (Zone-Max-Degree) as dominator selecting factors. For either version, all nodes in the network start in the initial state. Then, the node with the highest priority assigns itself as a seed dominator and broadcasts a Dominator message to its one-hop neighbors. Each node receiving this message considers itself as dominatee and replies by broadcasting a Dominatee message. When a one-hop neighbor that has a lower priority than a dominatee node, and in the same time has the highest priority among its initial state one-hop neighbors, receives a Dominatee message from that dominatee, it assigns itself as a non-seed dominator. Accordingly, the network is divided into separate zones where in each zone there is only one seed dominator. These zones take the IDs of their seed dominators. Determined dominatees and non-seed dominators are members in these zones. In order to identify connectors, each dominatee broadcasts the One-Hop-Dominator message, which has the IDs of all one-hop neighboring dominators. The highest-priority node between two dominators, in the same zone, is considered a connector node and broadcasts One-Hop-Connector message. Each node receiving messages that have different zone IDs, considers itself a zone probable border node. Zone probable border nodes send Two-Hop-Dominator messages to inform their dominators about the IDs of neighboring zones and dominators. According to aggregated information, dominators assign border nodes.

The second algorithm is the Connected Dominated Sets-Bounded Diameters-Distributed (CDS-BD-D) clustering algorithm [23]. The CDS-BD-D algorithm is a distributed clustering algorithm that comprises two phases. The first phase applies the distributed Breadth First Search (BFS) algorithm [24]. The second phase selects dominators and connectors. After constructing the BFS tree the second phase starts. In this phase, the node may have one state out of three: the dominator state, which is colored black, the connector state, colored in blue, and an ordinary state, colored in white. The state of a node is determined according to its level in the BFS tree and its weight. The weight is calculated according to three prioritizing parameters.

The first parameter is the node energy level, the second one is its degree and the third parameter is the node ID. The node with the highest energy level is considered highest weight node. In case of a tie, the second parameter is checked then, if the tie still holds, the third parameter is used to break the tie. The root node is selected as a black