Clustering for Networks of Moving Objects

Veselin Rakocevic

[1] Department of Electrical and Electronic Engineering
[2] City University London, United Kingdom

Abstract. This chapter presents the overview of the technical challenges and the currently available solutions to the problem of clustering of moving objects in ad hoc wireless networks. The networks of moving objects have been receiving increasing interest recently. These networks include networks of flying objects, networks of cars and other vehicles, networks of people moving in the cities, and networks of robots sensing the environment or performing coordinated actions. Clustering of such objects increases the scalability of the network and improves efficiency, enabling the objects to simplify the communication with their peers. Clustering of static network objects has been analysed in great detail in the literature. While most of the clustering algorithms and protocols are applicable in the networks of moving objects, there are specific challenges produced by the mobility. This document will present a rich body of currently available scholarly work on clustering for moving objects, focusing on the case when all network nodes (both clusterheads and cluster members) are moving. Most of the research works presented in this Chapter aim to predict the movement of the networked nodes, or to measure the relative mobility between the nodes, in order to optimise the processes of clusterhead election and cluster maintenance.

1 Introduction

The modern world has witnessed a significant increase in the number and the complexity of electronic objects which move and have a need to be networked. These objects include a variety of robot devices, sensors, small flying machines, vehicles and smart telephones and handheld devices carried around by millions of people. Networks of these devices are built on wireless communication channels, which are unstable. Further to this, the fact that the network nodes (objects) are constantly moving implies that their point of attachment to the network is constantly changing, and their neighbours are not fixed. Designing networks for moving objects is complex; it requires fresh solutions to the problems that have already been solved for networking of fixed objects.

The design of ad hoc networks for moving objects is the main subject of this Chapter. Clustering of distributed network nodes can be defined as the generation of groups of nodes which share some common features and communicate to the rest of the network via their leader (often called the clusterhead), rather than individually. The common feature of the network elements is typically their geographic location, although other features, such as speed of movement, application interest or established trust agreements can be considered. Clustering is important in distributed networks, as it contributes to improved network efficiency, connectivity and saves the cost of networking communication with regard to resource utilisation, power consumption and signalling overhead. In the process of clustering we can identify the following two processes: (1) clusterhead election process; (2) cluster maintenance. Clusterheads are typically elected using a distributed algorithm in which all nodes follow a predefined sequence of steps and elect the clusterhead based on a predefined election rule. Cluster maintenance process deals with the life of cluster, with the routing of packets inside the cluster and the processes that take place when cluster members have to leave the cluster or new cluster members join the cluster.

Clustering in wireless ad hoc networks and more specifically in wireless sensor networks has been analysed extensively over the last few decades. A number of clustering solutions for sensor networks have been designed, showing major performance improvements in terms of network efficiency (i.e. the number of routing messages, or the number of retransmissions due to unnecessary collisions), or power consumption. Significant theoretical work in leader election algorithms and connected dominating set establishments have paved the way for numerous protocols and algorithms that have been designed and tested in the simulation and experimental environments over the past decades. Some of the most important of these solutions will be presented in this Chapter as they typically form the basis of the mobility-aware clustering network designs. It is well known that clustering combined with data aggregation mechanisms can significantly improve the overall network efficiency.

During the last decade we have witnessed the emergence of a new generation of electronic systems which have potential to have a significant impact on everyday life and industry. This new generation includes various so-called cyber-physical systems: robots, small quadcopters or similar unmanned small flying objects, networked intelligence integrated in our cars, wearable sensor networks, etc. These devices form what is known today as the Internet of Things. These devices typically require communication and naturally form networks. Networks that can be small, consisting of only a few moving robots investigating a difficult terrain, or can be large, with thousands or tens of thousands little sensors forming huge network, very difficult to manage. Clustering these objects to simplify their communication is not only a useful add-on to the operation of these networks, but an absolute necessity.

The objective of this Chapter is to present the problem of clustering in distributed networks of objects that may be moving. Contrary to many existing research surveys, we focus here on the movement of all network nodes, both the clusterhead nodes and the 'standard' nodes. The movement generates the following problems:

- The clusters become unstable, because members of the clusters move and can often disappear out of the range of the clusterhead. For this reason, it is often that the slowest moving element is elected as the clusterhead

– In the case the clusters perform some data aggregation, the lack of stability of the cluster can have a major impact on the quality of the aggregated data
– The movement of the networked objects means that clusterheads cannot count on the connectivity of the nodes, and the amount of time the nodes are not connected into the networks increases.

Considering the listed problems, it will come as no surprise that the majority of researchers today tend to think towards predicting the movement of the networked objects, in order to predict the next step of the clustering process before any damage is created by the movement of the nodes.

However, before we get more into the detail of clustering for mobility, it is necessary to remind ourselves about the nature of the distributed networks, the principles and ideas for clustering, and the modelling techniques we can use to model the networks we discuss.

2 Topology Management in Distributed Networks

We define distributed networks as networks without a clear point of centralised control. In communication literature, such networks are often called ad hoc networks, to emphasise the temporary and random process of network formation. In such networks, typically all nodes are at equal hierarchical level. Depending on the geographical distribution of nodes and the communication requirement of such a network, the number of potential direct communication links between nodes can be very large. In these networks, when a node needs to transmit data packets to another node, it can do this directly in the case the receiving node is in the range of the sending node, or indirectly, by using other network nodes as routers. For this reason, routing procedures and protocols are required in ad hoc networks. Considering the fact that ad hoc networks can grow very large, one of the main challenges is how to operate an efficient network service in such a network. It is clearly not optimal to broadcast all data packets to all nodes in the networks, as this will have major implications on the resource utilisation and energy consumption (often these networks are energy-limited). A certain level of the control over the network structure (topology) is required.

Topology control in ad hoc networks can be done in many ways. Over the past few decades, two main directions of topology control in ad hoc networks have been identified: hierarchical topology organisation and power control.

In the hierarchical topology organisation the network nodes are grouped together (clustered) in smaller networks. One of the network nodes is elected to be a clusterhead, and all packets communicated from or to other nodes in the group (cluster) have to go through the clusterhead. We can see an example of a clustered network in Fig.1. The white nodes on Fig. 1 represent clusterheads, and the blue nodes represent the standard nodes. We can see how the clustering process introduces a clear structure into the network.

Topology control using power control is based on the idea that network nodes can have different transmission ranges - this can be controlled by varying the signal strength at transmission. Performing careful power control can establish

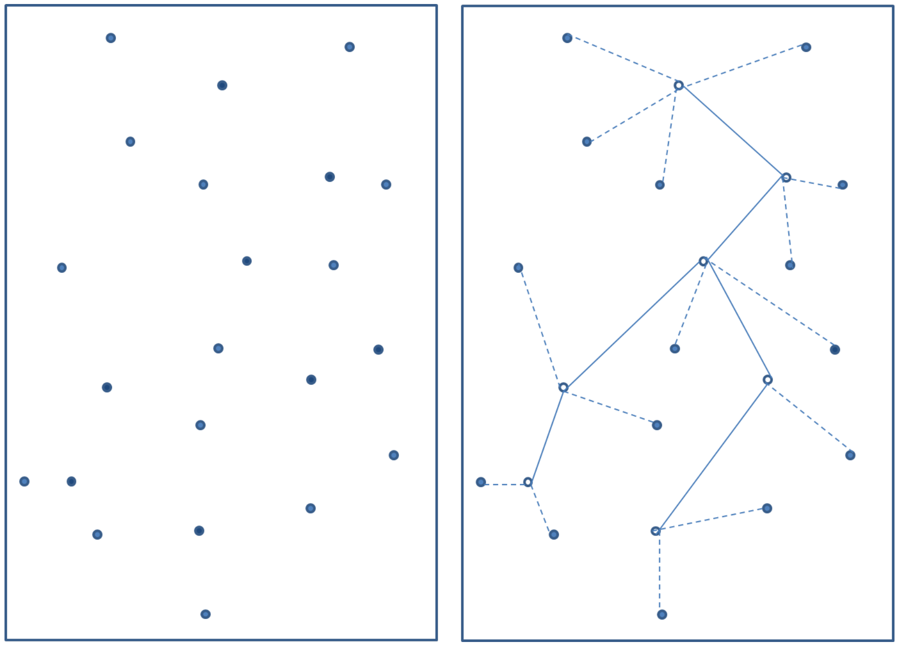

Fig. 1. Ad hoc network: unconnected (left) and fully connected and clustered (right)

communication patterns in the network that can increase the efficiency of the network communication. We can consider that the fewer the incidents when network nodes receive packets that are not destined for them, the more efficient the network is.

It is worth remembering that clustering is not a technique that is only used in distributed communication networks. As described in the work of [1] [2], given any data set of connected elements, the goal of clustering is to divide the data set into clusters such that the elements assigned to a particular cluster are similar or connected in some predefined sense.

In general, the topology of an ad hoc network can be presented by an undirected graph $G = (V; E)$, where V is the set of network nodes, and $E \subseteq V \times V$ is the set of links between nodes. Nodes in an ad hoc network communicate through a common broadcast channel using omni-directional antennas with the same transmission range. For any two nodes u and v that are within the packet-reception range of each other, u and v are called one-hop neighbours of each other. Two nodes that are not connected but share at least one common one-hop neighbour are called two-hop neighbour of each other. The challenge in topology control can also then be defined as [3]: to identify a subgraph of the unit disk graph, such that network features such as bounded node degree are preserved, and advance routing methods such as localized routing are enabled. Examples of localized routing include greedy routing [4] [5] [6], and compass routing [7]. A routing protocol is localized if the routing decision is based on the packet header information (i.e. destination node ID) and the local information from a small neighbourhood.

In graph theory, the minimum dominating set problem and the relevant minimum connected dominating set (MCDS) problem most closely represent the clustering approach to topology control. The dominating set problem can be described as finding a subset of nodes with the following property: each node is either in the dominating set, or is adjacent to a node in the dominating set. The MCDS problem consists of obtaining a minimum subset of nodes in the original graph, such that the nodes compose a dominating set of the graph, and the induced subgraph of an MCDS has the same number of connected components as the original graph. Although attractive, finding the MCDS is a well-known NP-complete problem in graph theory [8] [9].

Another point is important to make here. From the application point of view, we can identify two types of distributed networks. The first type includes networks where nodes act as routers/relays, delivering packets to the other network nodes. In such networks, hierarchical routing is applied to ensure the data reaches their destination. Clustering in such networks is very important, as it enables hierarchical routing and optimises the routing process. An example of such a network would be a standard wireless sensor network, or a large-scale ad hoc network of everyday electronic appliances. When it comes to the issue of mobility in such networks, the challenge is great, as there is a critical requirement for fast reclustering of moving nodes to ensure packets are routed through the network in an efficient way.

The second network type is a network in which the nodes are required to communicate their location information and potentially some basic information about their environment, in order to help other nodes to get a better understanding of the environment around them. Typical examples of such networks are vehicular ad hoc networks for traffic information dissemination, or dissemination of safety-related information from a particular geographic location.

Whatever the service the network delivers and the requirement for clustering is, all clustering methods have several features in common. Cluster members all share a common feature, which can be location (cluster members are close to each other, often within a transmission range of the clusterhead), speed and direction of movement, or some application-level information. Clustering in mobile ad hoc networks is an area that has been analysed in the literature. Yu and Chong give a good survey on clustering algorithms in [11]. They investigate the cost of clustering and identify the following types of clustering solutions: (1) Dominating Set based clustering; (2) Low maintenance clustering; (3) Mobility-aware clustering; (4) Energy-efficient clustering; (5) Load Balancing clustering and (6) Clustering based on combined metrics. Their paper gives a good overview of representative algorithm that fit this classification. Similarly, Vodopivec et al [12] give a short survey of clustering schemes focused on vehicular networks. One-hop clustering algorithms and their performance are surveyed in [13].

This Chapter will focus on how movement and mobility is represented in the clustering algorithms and protocols. We start the analysis by introducing the basic principles of cluster formation (section 3). This is followed by an overview of mobility-aware clustering solutions in Section 4. Section 5 introduces cluster maintenance processes, and briefly discusses the performance evaluation methods.

3 Cluster Formation

This section analyses the most important methods for cluster formation. While clustering - as we have already mentioned earlier in this document - is a generic method with many applications in data processing and in the analysis of live organisms and biological processes, here we focus on the cluster formation in wireless ad hoc networks. As we have seen earlier in the Chapter, clusters are typically formed in a distributed process of leader (clusterhead) election. This process requires all nodes to be able to identify themselves and all nodes to follow a pre-defined procedure. Early leader election algorithms chose the clusterheads on the basis of a given ID, or on the basis of the number of neighbours nodes have. We will see that mobility-aware clustering algorithms typically use the detailed information about the movement of the nodes to elect a clusterhead. But, firstly we need to introduce the basic leader election algorithms that have been proposed to use in ad hoc networking.

The *leader election* algorithm is a standard algorithm for distributed systems, often found in theory and practice. The classical definition of the leader election problem is to elect a unique leader from among the elements of a distributed system. In a mobile ad hoc network, we can expect that the network topology will change frequently, so the definition of the leader election algorithm can be modified, as it was done by Vasudevan et al in [14]: the requirements for the leader election algorithm are: after topological changes stop sufficiently long, every connected component will eventually have a unique leader with maximum identifier from among the nodes in that component. Vasudevan, Kurose and Towsley in [14] present a leader election algorithm which is based on a process of growing and then shrinking of a spanning tree that is rooted at the node that initiates the leader election process. In this algorithm, in the Election phase, the node sends Election message to its neighbours. The receiving nodes identify the sending node as a parent node. For each node there can be only one parent node. In the acknowledgement phase, an ack message is sent to each node from which Election message is received, apart from the parent node. Nodes respond to their parents only when all of their children have responded to them. In these acknowledgement messages nodes will announce to their parents the maximum identity node among all downstream nodes. Finally, in a phase that is called Leader, once the root node has received all acknowledgements rom all the children, it will broadcast a leader message to all nodes announcing the identity of the leader. Other proposals for leader election algorithms can be found in the work of Royer and Perkins [16], and Malpani et al. [17] .

One of the best known simple clustering algorithms in the *lowest-ID* algorithm. In this algorithm, the node with the lowest identification number has the highest priority to be selected as the clusterhead. The lowest ID algorithm was originally proposed by Baker and Ephremides in [18] [19]. The neighbouring nodes with higher IDs assume the role of cluster members and form the cluster. The clusterhead selection procedure is repeated for the remaining nodes until either each node is selected as a cluster-head or a cluster member. The lowest ID is known to be a two-hop cluster formation algorithm, since the distance between each node and every other node in a cluster is at most two hops. In its basic form, the algorithm assumes that all nodes are given some IDs before the network is set up. The nodes exchange the IDs and the nodes with larger IDs back up and declare themselves as network nodes, while those with higher ID become clusterheads. This algorithm in its basic form show no appreciation of the topology changes in the networks and it is expected that the level of reclustering is larger compared to the e.g. highest degree clustering.

A somehow natural extension of the lowest ID algorithm is the so-called *node-weight clustering* concept, where nodes are given weights (in a sense equivalent concept to Ids), but not randomly - the weights are given based on some specified feature of the nodes. Basagni [20] generalized the lowest-ID algorithm proposed by Gerla and Tsai [22], by using a generic weight as a criterion for cluster-head selection. In his distributed and mobility adaptive clustering (DMAC) algorithm, a weight is associated with each host in the network. This weight corresponds to the suitability of the host to be selected as a clusterhead. The weight was originally thought to represent the residual energy of the host, but as we will see in the next section, it can also represent the mobility level. In [21], Ghosh and Basagni investigated the impact of the different mobility degrees and mobility patterns of the mobile hosts on the performance of DMAC protocol. They showed that the cluster reorganization rate of DMAC protocol considerably increases in the presence of the host mobility. To address the negative impact of the host mobility on the performance of DMAC, Ghosh and Basagni proposed a generalization of DMAC protocol called GDMAC in which the clusters are more stable against the host mobility. GDMAC reduces the rate of unnecessary cluster updates by applying two limiting rules: (1) controlling the rate of reclustering by reclustering only when the weight of the new clusterhead exceeds the weight of the current one; (2) controlling the spatial density of the cluster-heads.

Another standard method for clustering is the*highest degree* algorithm. This algorithm, with its application to mobile ad hoc networks originally proposed by Gerla [22] and Parekh [23], is based on the idea that the node with a maximum number of neighbours is chosen as clusterhead. In this algorithm, the clusterhead is directly connected to all nodes in the cluster, so the maximum distance between two nodes is two hops, similarly to the lowest ID algorithm. This results in a typically lower throughput than other clustering algorithms, as the degree (the number of neighbours) for some of the clusterheads may be large.

Clustering algorithms do not have to support two-hop solutions only. We can take the example of a d-hop cluster formation given by Amis, Prakash, Vuong, and Huynh. In their paper [9] they first give a proof that the minimum d-hop dominating set problem is NP-complete, and then present an interesting heuristic for the construction of d-hop clusters in a network. In their heuristic, the clusters are formed following $2d$ rounds of broadcasting hello-type messages in the network. The basic idea is that each node maintains two arrays of nodes - WINNER and SENDER. Initially, each node sets WINNER to be equal to its ID, and then messages are broadcasted to all neighbours. Each node chooses the largest ID from all the received messages and puts it into the WINNER array. This process is called $Floodmax$. After d rounds of $Floodmax$, a separate process called $Floodmin$ is performed, where, following a similar message broadcast, the smallest IDs are chosen and remembered. The $Floodmin$ phase allows the relatively smaller clusterheads the opportunity to regain the nodes from the neighbourhood. The clustering algorithm follows these two message propagation processes with the next two stages: (a) determination of clusterheads and (b) linking of clusters. The determination of clusterheads happen based on simple rules - e.g. if a node received its own ID in the second round of flooding, the node immediately declares itself as a clusterhead. Otherwise, if a node finds another node in the WINNER arrays for both rounds of flooding, the node declares will declare that node as clusterhead. Finally, if neither of these two cases happens, the node will declare the maximum ID from the first round of flooding as clusterhead.

Another example of clusterhead formation where cluster size is not limited to two hops is the work of Err and Seah [33]. In their solution, the diameter of clusters is flexible and determined by the stability of clusters.

It is easy to see from these cluster formation examples, what may be the natural design idea for mobility-aware clusterhead formation. Rather than basing the clusterhead election decision on the ID of the node, or the node's degree, or an arbitrary weight, the idea would be to identify the nodes by their mobility level and/or movement direction, which can be measured in many different ways, as we will see later in this chapter. The idea then is to choose for the clusterheads those nodes that have a specific mobility pattern; in most of the works we can notice that there is an attempt to choose for the clusterhead the node that is the least mobile.

Finally, it is worth noting that clustering and cluster formation are not only applications of complex mathematical theory, but are networking techniques that needs to be designed to satisfy an application requirement. The work of Reumerman et al [24] gives an interesting insight into what is called application-level clustering. They present a concept where each application can set up its own virtual cluster in order to optimise the information exchange relevant for that application. It is interesting to note that they distinguish between two cluster types (they mostly observe vehicular networks) - the 'moving cluster', where all cluster members move more or less together, in a group; and the 'quasi static cluster', where the identity of the cluster head and the cluster members is not important - the application requires a clusterhead and cluster members at a certain geographical location, to be able to disseminate certain information. This would typically be applied for applications on traffic information and safety-related information in vehicular networks.

4 Considering Mobility in Clustering Solutions

In the previous section we have seen that clustering formation algorithms typically support a distributed election process where all nodes follow a pre-defined rule when electing the clusterhead. In this section we will review the existing solutions for mobility-aware clustering. We will see that the majority of solutions attempt to predict the position of mobile nodes and perform clustering with this knowledge. This typically means that the clusterhead election focuses on the choice of a clusterhead that will ensure maximal stability for the network. The estimation of mobility focuses on the *local* mobility, i.e. the relative position of the nodes compared to their neighbours, rather than on the *global* mobility which can be accurately measured using GPS or other satellite-based global systems.

In the analysis given in this section, we focus on the case when all nodes in the network (both clusterhead nodes and non-clusterhead nodes) are moving. With this in mind, the focus on the analysis is to identify the existing solutions for clusterhead selection and cluster formation in the case when all nodes in the network are moving. A similar analysis can be made on the network connectivity for moving nodes when clusterheads are fixed and uniformly distributed. For more details about this, we refer to the excellent paper of Wang et al. [42].

In this section different methods for mobility consideration will be presented. We are interested to analyse the ways mobility can be measured and what parameters can be used in the process of clustering. We are looking at a simple basic case of nodes as presented in Fig.2. There we have two nodes, x and y, which are changing their locations $l(x,t)$ and $l(y,t)$, using speed $v(x,t)$ and $v(y,t)$. We will see that mobility is used in clustering solutions in one of the following three ways: (1) by measuring the relative mobility of node x in comparison to node y; (2) by identifying similarities in the movemenet pattern of x and y; (3) by predicting the next locations of nodes x and y.

The general idea fo rusing the relative mobility measure for cluster formation in network of moving objects typically follows a well-known MOBIC protocol [10], where estimation of nodes speed variance is used in the clusterhead election process, with nodes with low speed variance having a better chance of becoming clusterheads. Basu et al. point out in [10] that correct estimation of local mobility is critical for the clustering process. A number of works investigate global mobility, by assuming the existence of the GPS or some other method of accurate positioning of the nodes. As clusters need to form and reform quickly, especially in the presence of high mobility, it is the local mobility, the relative speed/location difference between the nodes that is of interest. Basu et al. use the received power as the estimate of the distance between the nodes. While the received power is not the most accurate method of estimating the distance considering the wireless channel fading and other constraints to signal propagation, it is usually assumed that clustered nodes share the environmental constraints and the received power is often used as the estimate of the distance. Basu et al. introduce what they call the *relative mobility metric*, which is in essence the ratio of the received signal power between the old and the new packet that was exchanged between two nodes. The relative mobility metric can be defined as

$$M_y^{Rel}(x) = 10log_{10}\frac{Pr_{x->y}^{new}}{Pr_{x->y}^{old}} \tag{1}$$

For this to work, regular exchange of the packets is necessary. Regular exchange of packets is a feature of ad hoc networks, where periodic 'Hello' packets are exchanged between the neighbours. The aggregated local mobility for any node is then the variance with regard to zero of all relative mobilities for the neighbouring nodes:

$$M_y = var_0(M_y^{Rel}(x_1), M_y^{Rel}(x_2), M_y^{Rel}(x_3), ..., M_y^{Rel}(x_m)) = E[(M_y^{rel})^2] \tag{2}$$

Basu et al. continue to suggest that the node with the smallest variance of relative mobilities should be identified as the clusterhead, as this would maximise the cluster stability.

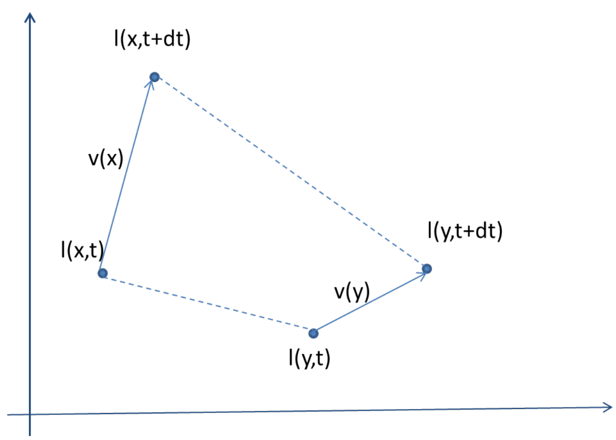

Fig. 2. Simple model of movement of two nodes, x and y

Other research presents similar approach to the problem. One of the first mobility analyses was performed in the well-known work of Johansson et al [25]. They present a global mobility metric where the speed of node is measured relative to the other moving nodes. A mobility metric is proposed which is geometric in the sense that the speed of a node in relation to other nodes is measured. Johansson et al propose a simple model where the relative velocity $v(x, y, t)$ at time t between node x which is at location $l(x, t)$ and node y which is at location of $l(y, t)$ (like in Fig.2) can be expressed as $v(x, y, t) = \frac{d}{dt}[l(x, t) - l(y, t)]$. The mobility measure M_{xy} is defined as the absolute relative speed taken as average over the time period:

$$M_{xy} = \frac{1}{T} \int_{t_0}^{t_0+T} |v(x, y, t)| dt \quad (3)$$

Many other variations of the basic mobility analysis given by Basu et al. exist. For example, Wu et al [37] use mobility index to characterize the mobility of each node. This index is shared in the hello messages and can be used to improve the stability of the clustering process. The mobility index is expressed as

$$M_y = \sum_{i=1}^{n} W_i * D_y^{x_i} \quad (4)$$

where W_i is a weight parameter defined as

$$W_i = \frac{(M_{x_i})^{-1}}{\sum_{j=1}^{n} (M_{x_j})^{-1}}, \sum_{i=1}^{n} W_i = 1, \quad (5)$$

and $D_y^{x_i}$ is the relative geometric distance between node y and node x_i, calculated from two consecutively received hello messages. Wu et al. propose the use of the mobility index in the cluster formation process.

An and Papavassiliou [30] use a similar approach to measure mobility. They assume that all nodes are able to identify their location and define relative velocity between two nodes x and y at time t as $v(x, y, t) = v(x, t) - v(y, t)$. They then define the relative mobility $M_{x,y,T}$ between any pair (x, y) of nodes during time period T as absolute relative speed:

$$M_{x,y,T} = \frac{1}{N} \sum_{i=1}^{N} |v(x, y, T)| \quad (6)$$

where N is the number of discrete times the speed is calculated during the period T. Then they define two cluster mobilities - the first one represents the motion behaviour of a cluster as a whole, $\frac{1}{M} \sum v(i, T)$, and the second one represents the motion behaviour of nodes within the cluster, $\frac{1}{N} \sum M_{x,y,T}$, where M is the number of nodes in the cluster, and N is the number of node pairs in the cluster. The clusters are then formed by nodes exchanging the velocity information, calculating their mobility metrics and then the clusterhead can be elected only if its relative mobility is below a threshold.

Measuring relative mobility is done in a slighlty different way by Err and Seah in [33]. They attempt to measure the variation of distance between nodes over time in order to estimate the relative mobility of two nodes. Their idea is to cluster together nodes that have a similar moving pattern. This idea is later used for clustering in vehicular networks, where there are several attempts to cluster the nodes based on their speed (e.g. [43] [44]), so that the nodes that move at similar speeds are clustered together. Err and Seah, similarly to some other works we have seen here, use the received power strength to estimate the distance between two nodes in the network. The estimated distance between two nodes x and y:

$$E[D_{xy}] = \frac{k}{\sqrt{P_r}} \tag{7}$$

is used only as an approximation of the 'closeness' of the nodes. The relative mobility is defined as:

$$M_{xy}^{t-1} = E[D_{xy}^{t}] - E[D_{xy}^{t-1}] \tag{8}$$

In the algorithm they estimate the stability of a node by observing the standard deviation of the variations in mobility to all other nodes in the network.

We can see in all of the presented examples that the objective of clusterhead election process in mobile networks is to elect for clusterheads the nodes that are the least mobile, in expectation that the relative lack of movement of such nodes will increase the stability of the clusters.

Idetifying similarities in the moving patterns of networked nodes often leads to *group mobility*, which is a frequent event in the real world and is also naturally linked to the concept of clustering. Within group mobility modelling, the work of Hong et al [26] is especially important. Hong et al investigate the movement of groups of mobile nodes and make an assumption that each group needs to have a logical centre, with the group movement being defined by the movement of the centre node. They introduce a novel group mobility model - Reference Point Group Mobility (RPGM) - to represent the relationship among mobile hosts.

Work of McDonald and Znati [27] is also very important, as they attempt to bound the probability of path availability for moving nodes. They look at the problem that is often found is the research on the networks of moving objects: developing a model that derives expressions for the probability of path availability as a function of time. In other words, the aim was to determine the conditional probability that the nodes will be within range of each other at some time $t + \delta t$ given that they are located within range at time t.

Zhang et al in [31] propose a revised group mobility metric, the linear distance based spatial dependency (LDSD), which is derived from the linear distance of a nodes movement instead of its instantaneous speed and direction. Their work also assumes the nodes node their exact physical location. The linear distance is a square root of the sum of the changes in geometric x and y coordinates of a node. If the node's linear distance D is greater than some threshold D_{thr}, the mobility history information is udpated correspondingly. Based on the information in the history cache, a node calculates its linear distance based total spatial dependency with respective to its neighbours.

In a number of works researchers attempt to predict the mobility of the moving nodes, in order to improve the cluster stability. In their excellent paper, Konstantopoulos et al [36] present an algorithm that estimates the probability of neighbourhood stability (which is for them the product of nodes' stabilities). They make an assumption that a mobile node can be considered a good candidate for the clusterhead if its neighborhood is relatively stable in comparison to the neighborhoods of other candidate hosts.In their paper a detailed algorithm for estimating neighbourhood stability is presented.

Leng et al [35] use the node connectivity and node mobility jointly to select clusterheads. Their approach is based on availability of position information using GPS or similar technology. They then consider the *link expiration time*, defined as the time after which two nodes will leave each other's transmission range. They use simple geometry for this, assuming that the velocity does not change during the link expiration process. After this, they define average link expiration time for each node, by averaging link expiration times. The node with a large value of average link expiration time is able to maintain relatively long connection with their neighbouring hosts. Such a host should more likely be selected as a clusterhead than the host with a short link expiration time. In the clusterhead election process, the nodes use the combined metric of node degree (i.e. the number of neighbours) and the node's average link expiration time.

Finally, it is worth noting that, while the clustering solutions are required to consider the mobility of the nodes, for the correct clustering solution it is important that the mobility of the nodes in modelled the right way. Mobility modelling has been observed over many decades in the research community. An excellent survey of mobility modelling for ad hoc network research has been done by Camp, Boleng and Davies in [38]. In addition to this, a lot work has been done on modelling specific mobility patterns, such as the one identified in vehicular networks, where cars and other vehicles are constrained by the transport network. The work of Fiore and Harri in [39] is typical for mobility analysis of vehicular network. This paper gives a detailed analysis of the topological properties of a vehicular network. Similarly, Wang and Tsai in [42] present an interesting study of the mobility of vehicles in urban environments and develop algorithms for the estimation fop traffic congestion based on vehicle mobility patterns.

Comparing the approaches presented here is not easy. The accuracy of mobility measurements is only one performance aspect, because design of the networks of moving objects often has to follow a tailored-made approach, where rarely we have the case when one solution fits all problems. This is why full understanding of all presented approaches is important in the design process, to help us make the correct design decision.

5 Cluster Maintenance and Reclustering

For the majority of clustering solutions, the clustering process is based on the fundamental algorithms presented earlier in section 3. The distributed process begins with nodes broadcasting the information about themselves. This is followed by an election process where all nodes identify their position as clusterheads or cluster members, depending on the node features which were included in the broadcasted message. While in the traditional clustering algorithms this feature was typically node ID or the node degree (number of neighbours), in the mobility-aware clustering solutions, the nodes distribute information about their mobility, and clusterheads are chosen on the basis of the mobility information. Typically, the clusterhead is chosen as a node that is the least mobile; this is done in attempt to increase cluster stability. A typical example of the mobility aware clustering algorithm is the DMAC algorithm presented by Basagni in [20].

Once the clusters are set up, and each node knows its role in a clustered network (clusterhead or cluster member), the critical issue with regard to the operation of network of moving objects is cluster maintenance. If we consider a network with moving nodes, we can assume that the cluster members will frequently leave the transmission range of clusterheads. This can cause a significant problem, especially if we are observing a network of nodes expecting support for complex data transfer application. For these networks, to minimise the effect of clustering on the quality of service, it is essential that alternative routes are identified when the nodes change clusters. For the type of network where routing is not the main challenge (i.e. the networks where nodes exchange location information, rather than acting as relays for data traffic), the problem is smaller, but is still important.

Basagni [20] gives a simple procedure for dealing with moving nodes. The movement of the nodes can result in link failure, and the idea is that the clustering process reacts on link failure, as a trigger that a node has left the cluster. When this happens, the clusterhead removes that node from a list of nodes in the cluster. For the departed cluster member, there is an immediate need to find a new clusterhead. This node would then listen to incoming messages from clusterheads, and would perform a clusterhead check, similar to the one that was done at the cluster formation process. In the case of the DMAC algorithm, for example, this means the weight assigned to new clusterheads needs to be checked. If the weight of the newly arrived node is greater than the weights of the potential new clusterheads, the new arrival will announce itself as clusterhead. Otherwise, it will accept the clusterhead with the greatest weight as its new clusterhead.

Similar approach is taken by other algorithms. For example, Konstantopoulos [36] present a solution where the moving node chooses the clusterhead that has the highest probability of still being the neighbour in the immediate future. An et al [30] expect the moving node to choose the clusterhead with the highest mobility index and Zhang et al [31] expect the node to choose the clusterhead with the largest linear distance based spatial dependency. We can notice that for the majority of existing solutions the idea is that the choice of the new clusterhead is done on the basis of the mobility metric and/or weight assigned to the nodes. For Ni et al [34] in the initial clustering stage, the nodes having the smallest relative mobility in their neighbourhoods are selected as the clusterheads. In the cluster maintaining stage, mobility prediction strategies are introduced to handle the various problems caused by node movements, such as possible association losses to current clusterheads and clusterhead role changes, for extending the connection lifetime and providing more stable clusters.

We have seen in this Chapter that various parameters can be used to identify optimal clusterheads and that various methods can be used to establish and maintain clusters in networks of moving objects. The current research attempts to evaluate the presented clustering solutions, by identifying the performance evaluation parameters most suited for these dynamic solutions. The main performance parameters that can be used to evaluate clustering techniques for networks of moving objects include the following: the number of formed clusters in a network; the clusterhead duration; the reaffiliation rate, and signalling overhead.

Clustering is done to simplify the network operation, and to find the optimal subset of the network graph which can deliver full network connectivity. Therefore, the objective of cluster formation is to strike a balance between the number of clusters (the fewer, the better) and achieving network connectivity. Measuring the number of active clusterheads is a good measurement of the efficiency of the clustering scheme. For example, the simulation results shown in [36] show that the lowest ID and the highest degree clustering solutions results in the largest number of clusterheads. This result is expected, as the two algorithms have not been designed to minimize the number of clusterheads.

Similarly, the clusterhead duration can be used as a performance parameter, as a measurement of how long on average clusterheads keep their role. Fewer changes in clusterheads are desirable in order to increase the network stability. In power-based topology control, changes in clusterheads may be desirable, to increase the fairness of the power consumption across the network, but in the hierarchical topology control we can look at this measurement in a different way. This is especially interesting when moving nodes are considered. In the simulation results presented in [36], for high average speeds of node movement, the 'traditional' clustering algorithms (the lowest ID, the highest degree), perform worse that the algorithms that are specifically designed to deal with mobility. This shows again why designing for mobility is important, as movement of nodes presents a new set of challenges.

Finally, measuring the reaffiliation rate, we can find out the average number of times a mobile node has to change the clusterhead. This is an important measurement, as it can show why when designing clustering schemes for moving nodes it is important to specify carefully when reafilliation can take place. Many traditional clustering algorithms are designed for nodes to identify the optimal clusterhead based on the ID, or node degree, or some other parameter, and are designed for static networks where the risk of mobility is minimal.

6 Conclusion

This Chapter presented the current research work on clustering for moving nodes in distributed wireless networks. Clustering is a process of grouping the nodes in network subsets, in order to simplify network operation and ensure network connectivity. This Chapter introduces topology control in ad hoc wireless networks, and introducsd the standard techniques for hierarchical topology control. It then presents how mobility is considered in clustering solutions. The first generation of clustering solutions was focused on optimisation of the choice of network subset, without detailed consideration of topology dynamics or node movement. The

Chapter presents a number of ways mobility can be considered. All the solutions have in common the design principle that they attempt to predict the mobility of the nodes and to integrate the mobility information in the clusterhead election process, typically by choosing the least moving node to be a clusterhead. For most of the presented solutions, the reafilliation and cluster maintenance should be based on minimising the number of clusterhead changes and simplifying the control overhead.

References

[1] Schaeffer S. E.: Graph clustering. Computer Science Review I: 27-64 (2007).

[2] Boccaletti S., Latora V., Moreno., Chavez M., Hwang D-U.: Complex Networks: Structure and Dynamics Physics Reports: Volume 424, Issues 45. 175-308. (2006)

[3] Basagni S., Conti M., Giordano S., Stojmenovic I. Mobile Ad Hoc Networking. IEEE Press (2004)

[4] Bose P., Morin P., Stojmenovic I., Urrutia J. Routing with Guaranteed Delivery in Ad Hoc Wireless Networks Wireless Networks 7, 609-616 (2001)

[5] Stojmenovic I. Position-Based Routing in Ad Hoc Networks IEEE Communications Magazine 40(7), 128-134. (2002)

[6] Karp B., Kung H.T. GPSR: Greedy Perimeter Stateless Routing in Wireless Networks in Proceedings of the 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking. (2000)

[7] Kranakis E., Singh H., Urrutia J. Compass Routing on Geometric Networks. In Proceedings of 11th Canadian Conference on Computatioonal Geometry (1999)

[8] Bao L., Garcia-Luna-Aceves J. J.: Topology Management in Ad Hoc Networks. Proceeding MobiHoc '03 Proceedings of the 4th ACM international symposium on Mobile ad hoc networking and computing Pages 129-140 (2003).

[9] Amis A. D., Prakash R., Vuong T. H. P., Huynh D. T. Max-Min D-Cluster Formation in Wireless Ad Hoc Networks in Proceedings of IEEE INFOCOM (2000)

[10] Basu P., Khan N., and Little T. D. C. , A Mobility Based Metric for Clustering in Mobile Ad Hoc Networks. in Proc. IEEE ICDCSW 01, pp. 41318. (2001)

[11] Yu J. Y., Chong P. H. J. Survey of Clustering Schemes for Mobile Ad Hoc Networks. IEEE Communication Surveys, 7(1), (2005)

[12] Vodopivec S., Bester J., Kos A. A Survey on Clustering Algorithms for Vehicular Ad-Hoc Networks Telecommunications and Signal Processing (TSP), 2012 35th International Conference on (2012)

[13] Chinara S., Rath S. K. A Survey on One-Hop Clustering Algorithms in Mobile Ad Hoc Networks Journal of Network System Management 17:183207. (2009)

[14] Vasudevan, S. ; Kurose, J. ; Towsley, D. Design and analysis of a leader election algorithm for mobile ad hoc networks, Network Protocols, 2004. ICNP 2004. Proceedings of the 12th IEEE International Conference on

[15] N. Malpani, J. Welch and N. Vaidya. Leader Election Algorithms for Mobile Ad Hoc Networks. In Fourth International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications. (2000).

[16] E. Royer and C. Perkins. Multicast Operations of the Ad Hoc On-Demand Distance Vector Routing Protocol. In Proceedings of Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM), 207-218. (1999)

[17] Malpani N., Welch J. L., Vaidya N. Leader election algorithms for mobile ad hoc networks Proceeding DIALM '00 Proceedings of the 4th international workshop on Discrete algorithms and methods for mobile computing and communications Pages 96-103 (2000)
[18] Baker, D.J., Ephremides, A. The Architectural Organization of a Mobile Radio Network via a Distributed Algorithm IEEE Transactions on Communications, 29(11) (1981)
[19] Ephremides A., Wieselthier, J.E. ; Baker, D.J. A design concept for reliable mobile radio networks with frequency hopping signaling Proceedings of the IEEE, 75(1) (1987)
[20] Basagni, S.: Distributed clustering for ad-hoc networks. In: Proceedings of International Symposium on Parallel Architectures, Algorithms and Networks (I-SPAN99), pp. 310315 (1999)
[21] Ghosh, R., Basagni, S.: Mitigating the impact of node mobility on ad-hoc clustering. J. Wirel. Commun. Mob. Comput. 8, 295308 (2008)
[22] Gerla, M., Tsai, J.: Multicluster, mobile, multimedia radio network. ACM/Baltzer J. Wirel. Netw. 1(3), 255265 (1995)
[23] Parekh A.K. Selecting Routers in Ad Hoc Wireless Networks in Proceedings of the SET/IEEE International Telecommunication Symposium (1994)
[24] Reumerman H-J., Roggero M., Ruffini M. The Application-Based Clustering Concept and Requirements for Intervehicle Networks IEEE Communications Magazine 43(4). 108-113. (2005)
[25] Johansson P., Larsson T., Hedman N., Mielczarek B. and Degermark M. Scenario-based performance Analysis of Routing Protocols for Mobile Ad Hoc Networks. Proceedings of ACM Mobicom, Seattle WA, (1999).
[26] X. Hong, M. Gerla, G. Pei and C.-C. Chiang. A Group Mobility Model for Ad Hoc Wireless Networks. Proceedings of ACM/IEEE MSWiM 99 Workshop, Seattle WA, (1999).
[27] B. McDonald and T. F. Znati. A Mobility-Based Framework for Adaptive Clustering in Wireless Ad Hoc Networks. IEEE Journal of Selected Areas in Communications, 17(8), (1999).
[28] Lin, C.R., Gerla, M.: Adaptive clustering for mobile wireless networks. IEEE J. Sel. Areas Commun. 15(7), 12651275 (1997)
[29] Torkestani J. A., Meybodi M. R.: A mobility-based cluster formation algorithm for wireless mobile ad-hoc networks. Cluster Computing, 14(4), pp 311-324 (2011)
[30] An B., Papavassiliou S. A mobility-based clustering approach to support mobility management and multicast routing in mobile ad-hoc wireless networks. Int. J. Network Mgmt ; 11: 387 395 (2001)
[31] Yan Zhang, Jim Mee Ng, Chor Ping Low. A distributed group mobility adaptive clustering algorithm for mobile ad hoc networks. Computer Communications 32 189202 (2009)
[32] Chatterjee M., Das S.K., Turgut D. WCA: A Weighted Clustering Algorithm for Mobile Ad Hoc Networks Cluster Computing 5, 193-204 (2002)
[33] Inn Inn ER, Winston K G Seah. Mobility-based d-Hop Clustering Algorithm for Mobile Ad Hoc Networks IEEE Wireless Communications and Networking Conference 2004, pp 2359 - 2364 (2004)
[34] Minming Ni, Zhangdui Zhong, and Dongmei Zhao. MPBC: A Mobility Prediction-Based Clustering Scheme for Ad Hoc Networks, IEEE Transactions on Vehicular Technology, 60(9), (2011)

[35] Leng S., Zhang Y., Chen H-H., Zhang L., Liu K. A Novel k-Hop Compound Metric Based Clustering Scheme for Ad Hoc Wireless Networks. IEEE Transactions on Wireless Communications, 8(1). (2009)

[36] Konstantopoulos C., Gavalas D., Pantziou G. Clustering in mobile ad hoc networks through neighborhood stability-based mobility prediction Computer Networks 52. 17971824. (2008)

[37] Wu H-T., Ke K-W., Chen C-H., Kuan C-W. Mobile Awareness Based Cluster Selection Mechanisms in Wireless Ad Hoc Networks In Proceedings of IEEE Vehicular Tehcnology Conference (Fall), 2774 - 2778 Vol. 4. (2004)

[38] Camp T., Boleng J., Davies V. A Survey of Mobility Models for Ad Hoc Network Research Wireless Communications and Mobile Computing, 2(5), 483502. (2002)

[39] Fiore M., Harri J. The Networking Shape of Vehicular Mobility In proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (2008)

[40] Wang C., Tsai H-M. Detecting Urban Traffic Congestion with Single Vehicle In Proceedings of the 3rd International Conference on Connected Vehicles and Expo (2013)

[41] Gutting R. F., de Almeida T. Ding Z. Modeling and querying moving objects in networks The International Journal on Very Large Data Bases. 15(2). 165-190. (2006)

[42] Wang X, Lin X, Wang Q, Luan W. Mobility Increases the Connectivity of Wireless Networks IEEE/ACM Transactions on Networking, 21(2), (2013)

[43] Ucar S, Ergeny S C, Ozkasap O. VMaSC: Vehicular Multi-hop algorithm for Stable Clustering in Vehicular Ad Hoc Networks In Proceedings of IEEE Wireless Communications and Networking Conference (2013)

[44] Hassanabadi B, Shea C, Zhang L, Valaee S. Clustering in Vehicular Ad Hoc Networks using Affinity Propagation. Ad Hoc networks, Volume 13, Part B, Pages 535548 (2014)