Clustering Heuristics in Wireless Networks: A Survey

Ameer Ahmed Abbasi
Al-Hussan Institute of Management
& Computer Science
Dammam-31411 Saudi Arabia
ameer_abbasi@hussan.edu.sa

M I Buhari
Dept. of Info. & Computer Science
University of Petroleum & Minerals
Dhahran-31261 Saudi Arabia
mibuhari@ccse.kfupm.edu.sa

M Akbar Badhusha
Faculty of Information Technology
University of Technology
Sydney, Australia
akbar@it.uts.edu.au

ABSTRACT

The subject of wireless networks has recently gained a lot of attention from research community to investigate. The well-organized subdivision of a wireless network into coherent, mostly non-overlapping clusters of physically close nodes is a significant building block in the design of well-organized upper layer network functions such as energy saving, routing, stability, coverage and data aggregation. In our paper, we present a snapshot of different clustering techniques for wireless networks. We also present comparative analysis of different clustering mechanisms based on metrics such as load balancing, efficiency of coverage, energy-awareness, cluster stability, routing overhead control, dynamic changes of topology and locality.

1. INTRODUCTION

Wireless networks are playing a fundamental role in emerging pervasive platforms that will host a wide range of next generation civil and military applications. With highly capable sensors and wireless network nodes available in the market, we need to efficiently utilize them with proper power consumptions. Configuration and maintenance of these nodes might impact their energy reserves. For example, having a centralized base station (BS), communications between the nodes and the centralized BS might consume more energy. But with the case of ad hoc networks, self-configuration and maintenance of these nodes might be energy ineffective. Clustering of wireless network nodes into groups with proper cluster head (CH) selection will impose a regular structure in the network. It is easier to control a CH with the summarized information instead of handling information from every node individually. Thus, CHs aggregate the data from various nodes and communicate with the far distant BS. In order to form clusters and efficiently control the network, clusters need to be formed in a fast and quick manner. Formation of clusters in one sector should not impact the formation of clusters in another sector. We need to make sure that the CHs are not over-burdened or under-utilized and the clusters not highly overlapped. Thus, appropriate equal-sized clusters have to be formed to evenly distribute the control. A wireless network without any clustering, every node has to send its information to a far away BS.

If a node has problems directly sending information to a far away BS, it has to communicate via other nodes by the concept of multi-hopping. This would create extra burden on those nodes that are close to the BS. Clustering is a logical solution to reduce this extra burden from those nodes. Another alternate is to have "Gateways" that are high-energy nodes which can pass the information to the BS (Younis et al. 2003). Furthermore, these gateways need to be considered based on their load and their communication with various sensors, to form clusters. These gateways can further act like CHs.

In a wireless network, communication between nodes could be single-hop or multi-hop. In single-hop communication (Younis and Fahmy 2004), each node has direct connection to the CH. There is no need for relaying or multi-hopping. Due to the direct connection, only one node should transmit at a time, and a contention-less Medium Access Control (MAC) protocol (Rappaport 1996) is needed because these protocols do not allow the possibility of frames colliding. We cannot constraint that all the nodes should be equally farther away from the CH. Thus, the energy depletion will vary from one node to another based on how far it is from the CH. Therefore, when deciding the battery lifetime, we need it to be based on the worst-case scenario. In a multi-hop communication environment (Younis and Fahmy 2004), the nodes are allowed to have multiple hops between the nodes and the CH. The number of hops that the packets can move from the source node to the destination CH could be indicated by a notation like K-hop, where K is the number of hops that the packet can travel towards the CH. If a node is unable to reach a CH within K hops, then the node (after deferring some random time) could claim itself as a CH.

The rest of paper is organized as follows. Various clustering algorithms are grouped under various sub-sections in section 2. A comparative analysis of the various algorithms discussed in section 2 is provided in section 3. Finally, section 4 concludes the paper.

2. CLUSTER ESTABLISHMENT HEURISTICS

2.1 Hierarchical Clustering Algorithms

The wireless networks could embrace large number of nodes. Proficient management of these nodes has crucial effect on traffic load, bandwidth resource availability,

dynamic changes of topology, etc. of a wireless network. Hierarchical or tree-based clustering is an effective way to administer these nodes. The management of these nodes could be in a single hierarchy with few hundred nodes which may or may not have CHs or multi-level hierarchies with thousands of nodes organized in several levels incorporating ordinary nodes, CHs and super-CHs. The examples of such multi-level hierarchies are presented in (Lauer 1986; Shacham and Westcott 1987; Lauer 1995).

Low Energy Adaptive Clustering Hierarchy (LEACH) (Heinzelman et al. 2002) is one of the most popular hierarchical routing algorithms for sensor networks. It is a cluster-based protocol that uses local data fusion and classification to greatly reduce the amount of information that must be transmitted to the BS. LEACH forms clusters by using a distributed algorithm, where nodes make autonomous decisions without any centralized control. The energy load of being a CH is evenly distributed among the nodes by incorporating randomized rotation of the high energy CH position among the nodes. Nodes elect themselves as CHs based on this probabilistic rotation function and broadcast their decisions. Each non-CH node determines its cluster by choosing the CH that requires the minimum communication energy. LEACH is completely distributed and requires no global knowledge of network. However, LEACH uses single-hop routing where each node can transmit directly to the CH and the sink.

Priority-based Adaptive Topology Management (PATM) (Tan et al. 2005) is another example that decreases the routing overhead notably and enhances the data forwarding services by using cluster-based control mechanism. PATM also demonstrates an obvious tendency toward dynamic changes of topology, bandwidth availability and traffic load. Analytical comparison of PATM with other cluster formation algorithms like Geographic Adaptive Fidelity (GAF) (Xu et al. 2001), Pre-defined Intelligent Lightweight tOpology managemenT (PILOT) (Srinidhi et al. 2003), and Topology Management by Priority Ordering (TMPO) (Bao and Garcia-Luna-Aceves 2003), etc. reveals the following unique features of PATM:
(1) Synchronization among nodes or node position information is not required by PATM.
(2) Centralized control over the ad hoc network is not required by PATM. Each node, based on its local information, is free to makes its own decisions.
(3) Without exchanging control messages sporadically, PATM proactively maintains a connected backbone.

In addition, using the following two optimizations, PATM demonstrates dramatic decrease in topology updates: (i) Piggyback the small control messages to ongoing data traffic and (ii) Adapt the topology update intervals based on the network mobility.

A clustering scheme is presented in (Banerjee and Khuller 2001) to create hierarchical control structure for wireless networks. They show how clustering can be obtained for wireless network topologies by obeying certain constraints that are useful for management and scalability of the hierarchy. They also present a distributed implementation of their clustering algorithm for a set of wireless nodes to create the desired set of clusters. A drawback in using their algorithm in sensor networks is that only one node needs to instigate the clustering, and that the protocol still entails $O(n)$ time in linear networks.

(Zhang and Arora 2002) also presented a tree-based algorithm (GS^3) for self-configuring a network into cells of tightly bounded geographic radius and low overlap between cells. GS^3 enables network nodes to organize themselves into a cellular hexagon structure with a set of proved properties. GS^3 is a self-healing, scalable, local, and low convergence time algorithm. It is applicable to static networks as well as to the networks with high degree of dynamics and mobility. The only negative aspect in their centralized scheme is the assumption that each node is aware of its location precisely, which is possibly complex in sensor networks to accomplish.

(Bandyopadhyay and Coyle 2003) propose Energy Efficient Hierarchical Clustering (EEHC), a distributed, randomized clustering algorithm to organize the sensors in a wireless sensor network into clusters. They also extend this algorithm to generate a hierarchy of CHs and observe that the energy savings increase with the number of levels in the hierarchy. EEHC algorithm organizes sensors into a hierarchy of clusters with an objective of minimizing the total energy spent in the system to communicate the information gathered by these sensors to the information-processing center. By assuming the contention and error free environment, the algorithm has a time complexity of $O(k_1 + k_2 + ... + k_h)$, a significant improvement over the many $O(n)$ clustering algorithms. This makes the EEHC algorithm suitable for networks of large number of nodes.

2.2 On-Demand (Passive) Clustering

The clustering that involves the concept of using on-going data traffic to piggyback "cluster related information" (e.g., the IP address of the node, the state of a node in a cluster) is called passive or on-demand clustering (Lee et al. 2000; Yi et al. 2002; Yi et al. 2003). Such a clustering protocol develops and maintains the cluster structure utilizing on-going data packets and works only if there is live data traffic in the network. Since, neighboring information is bundled in the data packets that are received by a node, only minor control overhead and setup latency is required in passive clustering. *First Declaration Wins* rule and *Gateway Selection Heuristic* (Yi et al. 2002; Yi et al. 2003) are

the two novel cluster formation mechanisms in passive clustering. According to the *First Declaration Wins* rule, claiming first to be a CH, any node can "govern" the rest of the nodes in its radio coverage (cluster area). Unlike all weight-driven clustering mechanisms, there is no waiting time to make sure all the adjacent nodes have been checked (Gerla and Tsai 1995; Lin and Gerla 1997). The *Gateway Selection Heuristic* offers a method to elect the minimal number of gateways (including distributed gateways) required to maintain the connectivity in a distributed manner. Implicit timeout duration is a key factor to maintain clusters in passive clustering. If a node that had previously communicated with other nodes in the cluster is not able to send data again within its timeout duration due to any expected or unexpected event would be assumed dead or out of its locality by other nodes in the cluster. With reasonable offered load, nodes in passive clustering can easily keep track of dynamic topology changes by virtue of this timeout.

Based on passive, on-demand clustering, a new flooding mechanism is proposed by (Yi et al. 2002). Without any loss in the network performance this new mechanism can reduce redundant flooding by up to 70% with negligible extra protocol overhead. They also presented the simulation results to prove the claim. Additionally, according to (Yi et al. 2002), substantial performance could be achieved if passive clustering is applied to several reactive, on-demand routing protocols such as Ad hoc On Demand Distance Vector (AODV) (Perkins and Royer 1999), Dynamic Source Routing (DSR) (Johnson and Maltz 1996) and On-demand Multicast Routing Protocol (ODMRP) (Lee et al. 2000).

2.3 Adaptive Clustering Mechanisms

Weight-driven or adaptive clustering is an equally appropriate architecture for either a fixed infrastructure or an atmosphere without any infrastructure at all. Adaptive clustering algorithm is robust in nature so it is well tolerated to mobility. An appropriate example of such a clustering is presented by (Lin and Gerla 1997). They propose adaptive clustering architecture for multimedia support in multi-hop mobile network and to administer different tasks such as spatial reuse of bandwidth, Quality of Service (QoS) provisioning by resource allocation within clusters. The data traffic in such a network is interwoven of two types of packets: data packets, generated by data sources and real-time packets, transmitted by real-time sources. Top priority is given to real-time sources therefore, in each cluster, Time Division Multiple Access (TDMA) permits to reserve bandwidth for real-time data traffic. Furthermore, within each cluster only one code is used to reduce the control overhead and to overcome the limitation of the number of orthogonal codes. Spread Spectrum provided the facility to the neighboring clusters to transmit at the same time on different codes. The network architecture, presented in (Lin and Gerla

1997), has the following three main advantages: (1) It provides spatial reuse of the bandwidth due to node clustering (2) Bandwidth can be shared or reserved in a controlled fashion in each cluster (3) The cluster algorithm is robust in the face of topological changes caused by node motion, node failure and node insertion/removal. Simulation experiments have also revealed many advantages of code separation and spatial reuse, such as this architecture provides an efficient, stable infrastructure for the integration of different types of traffic in a dynamic radio network. The performance of the proposed cluster scheme is similar to that of Cluster TDMA (Gerla and Tsai 1995), with less implementation complexity (less burdened by synchronization requirements).

Random Competition based Clustering (RCC) scheme mainly focuses at simplicity and stability (Xu and Gerla 2002). Although RCC is an active clustering algorithm but it utilizes the idea of *First Declaration Wins* rule of passive clustering scheme (section 2.2). Since it is possible that during the delay time period many neighboring nodes may claim concurrently to be CH, therefore to overcome on such a concurrent broadcast conflict RCC introduces an explicit random timer. Every time before claiming to be a CH, each node reschedules a random time and it gives up simply its CH claim if it hears a CH claim during this random time. RCC also uses the node ID to resolve further the concurrent broadcast problems as random timer is not a complete solution. When the concurrent broadcasts happen, the node with lower ID will become the CH.

2.4 Solid-Disc Clustering Structure

A clustering technique that produces approximately equal-sized clusters with minimum over-lapping among clusters is highly desirable in the cluster formation of sensor networks. Solid-disc clustering is a technique that does this job quite appropriately. According to solid-disc clustering, a node is to be considered in the cluster only if it is within a unit distance of the cluster's CH. This condition creates clusters in which the CH is surrounded by only those nodes that belong to its cluster. In other words, CH does not have to listen to those nodes that are not in its cluster. It listens to only those nodes that are in its cluster and that leads to a significant drop in the intra-cluster signal arguments. Since the CH in solid-disc clustering is at the center of the cluster and listens from all directions of the cluster, this phenomenon makes enhanced space coverage possible. Furthermore, aggregation is more significant at the CH because it is not biased to only one direction.

Based on solid-disc clustering technique, (Demirbas et al. 2004) propose a Fast Local Clustering Service (FLOC) that produces non-overlapping and approximately equal-sized clusters. In FLOC, CH uses a variable m (where the value of m could be ≥ 2) to decide which node belongs to its cluster. A node that is m unit away from the CH is not in its cluster. All nodes

that belong to the CH's cluster must be within the unit distance. By declaring the value of $m \geq 2$ FLOC creates a sense of locality as the entire network related problems are extremely restricted within at most 2 units distance. FLOC demonstrates self-healing capability in a manner that after facing problems it re-clusters itself in a constant time with locality. Also, with any number of nodes in the network, FLOC achieves clustering in a constant time, where constant time is $O(1)$, that characterize it fast and scalable. FLOC reveals the double-band character of wireless radio-model and introduces a communication and energy efficient program, by taking m to be the maximum communication radius and unit radius to be the reliable communication radius of a node.

2.5 Max-Min D-Cluster Formations

Max-Min Leader Election in ad hoc networks is a new clustering approach that takes special measures to eradicate the overhead of highly synchronized clocks and operates asynchronously (Amis et al. 2000). In the determination of the CH density, control and flexibility is achieved by generalizing the maximum distance to be d hops, where d is a value selected for the heuristic, from a node to its CH.

The approach has time complexity of $O(d)$ rounds which compares favorably to $O(n)$ for earlier heuristics (Baker and Ephremides 1981; Das and Bharghavan 1997) for large mobile networks. This reduction in time complexity is obtained by increasing the concurrency in communication. Since the number of messages is a multiple of d rounds, it helps to provide better efficiency at the network level. Furthermore, to provide enhanced stability and to minimize the transferal of network structure information, re-election of CHs is preferred. Data structure is simple that helps to minimize the local resources at each node. Redundant backbone architecture is formed by using CHs and multiple gateway nodes to provide communication between clusters. The approach is scalable as it tends to generate less CH than other clustering algorithms and provides better load balance among the CHs as it has a low variance in cluster sizes.

Mobility-based *d-hop* distributed clustering algorithm (MobDHop) generates variable-diameter clusters based on mobility pattern in Mobile Ad hoc Networks (MANETs) to achieve the desired scalability (Er and Seah 2004). The MobDHop algorithm allows cluster members to be more than two hops away from their CH and determines the diameter of clusters by measuring the stability of clusters. It measures the stability of clusters by using the relative mobility of cluster members and proposes a new metric to measure the variation of distance between nodes over time in order to estimate the relative mobility of two nodes. The MobDHop ensures maximum cluster stability by grouping the nodes with similar moving pattern into one cluster. It creates less clusters and CH change is low.

The performance of MobDHop is comparable to other existing algorithms.

2.6 Algorithm for Cluster Establishment (ACE)

A clustering algorithm that divides the network into clusters using minimum rounds of feedback would be characterized as efficient in term of coverage. Among such algorithms, Algorithm for Cluster Establishment (ACE) is an evolving algorithm that is proposed by (Chan and Perrig 2004). Cluster formation in ACE is highly efficient and uniform; it covers the entire network using just three rounds of feedback. The efficiency of ACE in terms of communication or coverage approaches that of hexagonal close-packing (Zhang and Arora 2002). ACE is powerful, flexible, fast, and robust against packet loss and node failure. Geographic location information, directional or any other kind of distance estimation between nodes is not required by ACE and no matter what is the number of nodes in the network, it completes in constant time using only local communication between nodes. It is a good demonstration of *emergent* algorithms in large-scale distributed systems. The definition of an emergent algorithm is outlined by (Fisher and Lipson 1999): "*An emergent algorithm is any computation that achieves formally or stochastically predictable global effects, by communicating directly with only a bounded number of immediate neighbors and without the use of central control or global visibility*".

The modified version of ACE is presented by (Chan et al. 2005), enhancing the regularity of the separation between CHs. The basic idea is "*to use clustering as a basis for determining the position information of sensor nodes*". First of all, they enhanced regularity at the cost of increased communication, by increasing the number of iterations for ACE from 3 to 5. By approximating a *spring effect* between neighboring clusters, they also modified the migratory mechanism. This effect triggers clusters that are too far apart to be attracted together, and clusters that are close together to migrate apart. The potential *fitness score* of each candidate node C in its neighborhood is evaluated by each CH during the migration phase. If C was to become the next CH, the score for each candidate C is calculated as the total number of sensor nodes belonging to the cluster, plus a modifier for each adjacent cluster of C. They assumed s to be the estimated separation between C and the neighboring CH in terms of maximum communication radius r. s can be estimated by counting the number of common nodes in both the clusters - the greater the number of common nodes, the closer the separation of the two clusters. They calculated final modifier via the function $g(s)$.

2.7 Linked Cluster Algorithm (LCA)

Proposing cluster formation and CH election algorithms, many papers focused on single-hop clustering and guaranteed that no node will be more

than one hop away from leader (Baker and Ephremides 1981; Baker et al. 1984; Gerla and Tsai 1995). The Linked Cluster Algorithm (LCA) takes special safeguard to avoid communication collisions among nodes and uses Time Division Multiple Access (TDMA) frames for inter-nodes communication with each frame has a slot for each node in the network to communicate. Every node requires *2n* TDMA time slots, where *n* is the number of nodes in the network, to have knowledge of all nodes in its neighborhood. If a node *x* has the highest identity among all nodes within 1 wireless hop of it *or* does not have the highest identity in its 1-hop neighborhood, but there exists at least one neighboring node *y* such that *x* is the highest identity node in *y's* 1-hop neighborhood, it becomes a CH.

To overcome these problems and to decrease the number of CHs that the original LCA produces, a revised version (LCA 2) is introduced in (Ephremides et al. 1987). According to LCA 2, a node would be considered "in the cluster" if it is in the 1-hop neighborhood of the CH. Furthermore, a node declares itself to be a CH if it has the lowest ID among the non-covered nodes in its 1-hop neighborhood. Basically, the LCA approach was designed to be used in the networks with less than 100 nodes. In such a small network, the delay between node transmissions is minor and may be accepted. On other hand, in the networks with more than 100 nodes, the delays between node transmissions in the TDMA communication scheme would be greater and unacceptable as the number of nodes increase in the network.

3. COMPARATIVE ANALYSIS

We have discussed various methods that have been used to handle clustering in ad hoc/wireless sensor networks. For the various methods proposed, the stated or claimed advantages with their respective disadvantages were given. To cater the aim of this paper, which is to provide chances for further research and updates in this developing area, we provide a comparative analysis of clustering methods discussed above.

Comparing to conventional clustering schemes, such as Lowest ID (LID) and Highest Degree (HD) algorithms (Lin and Gerla 1997), RCC (Xu and Gerla 2002) is more stable. Since CHs formed by RCC are usually at least two hops away from each other and a node only gives up its CH position when another CH moves near to it, clusters are much more stable. The algorithm presented in (Lin and Gerla 1997), cluster are chosen such that the CHs form a dominating set in the underlying graph topology and unlike (Banerjee and Khuller 2001), this makes the number and size of the clusters, largely dependent on the graph topology and the use is primarily for "spatial reuse" for channel spectrum. A drawback in using the (Banerjee and Khuller 2001) in sensor networks is that only one node needs to instigate the clustering, and that the protocol still entails *O(n)* time in linear networks. Adaptive

clustering (Lin and Gerla 1997; Xu and Gerla 2002), LCA 1 (Baker et al. 1984) and LCA 2 (Ephremides et al 1987) all use TDMA frames for inter-node communications but adaptive clustering is a multi-hop and LCA is a single hop scheme. GS^3 (Zhang and Arora 2002) differs from LEACH (Heinzelman et al. 2002) in terms that LEACH offers no guarantee about placement and the number of CHs in a system. The algorithm by (Banerjee and Khuller 2001), only considers logical radius (hops) of cluster, instead of their geographic radius, which makes long intra-cluster link possible. Also, its convergence under perturbations depends on multiple rounds of messages diffusion, instead of one-way diffusion within perturbed areas as in GS^3. Moreover, given certain node density in a network, the geographic radius in GS^3 implicitly guarantees the logical radius of clusters. GS^3 is a self-healing, scalable, local, low convergence time and applicable to static networks as well as to the networks with high degree of dynamics and mobility. The only negative aspect in this centralized scheme is the assumption that each node should be aware of its location precisely, which is possibly complex in sensor networks to accomplish.

Comparative analysis of passive clustering with "active" clustering and route aggregation techniques reveals several advantages. Specifically, passive clustering could be characterized as energy-aware clustering as it eliminates the periodic, background control packet exchange that reduces node power consumption. Furthermore, passive clustering is the only mechanism that can provide scalability and practicality for selecting the minimal number of forwarding nodes when topology changes dynamically. Max-Min D-cluster algorithm (Amis et al. 2000) partitions the wireless ad-hoc networks into d-hop clusters. Unlike FLOC (Demirbas et al. 2004), it does not guarantee solid-disc clustering and in the worst case, the number of clusters generated may be equal to the number of nodes in the network (for a connected network). LEACH forms 1-hop clusters but does not satisfy the FLOC's solid-disc property: Not all nodes within 1-hop of a CH *j* belong to *j*. Hence, in LEACH the CHs are susceptible to network contention induced by members of other clusters. FLOC complements LEACH since it addresses the network contention problem at the CHs by constructing solid-disc clusters. Moreover, LEACH style load balancing is readily applicable in FLOC by using the probabilistic rotation function for determining the waiting-times for the candidacy announcements at the nodes. Unlike LCA 1, LCA 2 and hierarchical clustering algorithm (Banerjee and Khuller 2001), EEHC (Bandyopadhyay and Coyle 2003) algorithm does not demand clock synchronization between the sensors and the simulation results prove that the authors have found the optimal parameter values for these algorithms that minimize the energy spent in the network.

Simulation results demonstrate that the efficiency of ACE (Chan and Perrig 2004) in terms of communication or coverage approaches that of GS[3]. ACE sometimes exceeds the ratio for hexagonal close packing (HCP) because it intelligently forms clusters around areas where nodes are most densely distributed, while choosing areas of overlap where the nodes are least densely distributed. ACE exhibits superior packing efficiency to either LCA or Adaptive algorithms (Lin and Gerla 1997; Xu and Gerla 2002). The variance in cluster sizes for ACE is small and only slightly larger than the baseline variance of the number of nodes in a given area (this is reflected in the variance of cluster sizes for HCP). The low variance and high average cluster sizes reflect that the ACE algorithm produces good packing efficiency. ACE maintains its large advantage over LCA and Adaptive algorithms even under conditions of heavy packet loss. ACE degrades at a higher rate and approaches the performance of LCA and Adaptive algorithms under conditions of high packet loss, but never actually performs worse than either algorithm under the simulated conditions. Unlike hierarchical control clustering algorithm (Banerjee and Khuller 2001), ACE algorithm is localized and requires no central direction that makes it highly resistant to transmission errors and random node failure. The overall communications overhead is small compared with the normal communications load for the sensor network.

Table 1: Comparison of Clustering Methods

Clustering Approaches	Cluster Overlapping	Location Awareness	Convergence Time	Energy Efficient	Failure Recovery	Balanced Clustering	Localization	Cluster Stability	Network Contention
LCA	No	Required	Variable	No	Yes	OK	Yes	Moderate	No
Adaptive Clustering	No	Required	Variable	N/A	Yes	OK	Yes	Low	No
Hierarchical Control Clustering	Low	Not Required	Variable	N/A	Yes	Good	Yes	Moderate	Limited
RCC	No	Required	Variable	N/A	Yes	Good	Yes	Moderate	No
LEACH	No	Not Required	Constant	No	Yes	OK	Yes	Moderate	Limited
GS[3]	Low	Required	Variable	N/A	Yes	Good	Yes	Moderate	Very Limited
EEHC	No	Required	Variable	Yes	N/A	OK	Yes	N/A	No
FLOC	No	Not Required	Constant	N/A	Yes	Good	Yes	High	No
ACE	Very Low	Not Required	Constant	N/A	Yes	Good	Yes	High	Very Limited
PATM	Low	Not required	Variable	N/A	Yes	Very Good	Yes	High	N/A
Passive Clustering	Low	Not Required	Variable	Yes	N/A	OK	Yes	Moderate	N/A
MobDHop	No	Required	Variable	Yes	Yes	Good	No	High	No

4. CONCLUSION

Clustering of various nodes in wireless networks has been vital to handle various upper layer network functions like energy saving, routing, etc. In this paper, we have provided a survey on various most frequently used clustering techniques. Pros and cons of various approaches are also provided. A comparative analysis to summarize the discussion and a comparative table that portrays the well-being any clustering technique is provided. Based on our analysis, we would prefer Solid-Disc clustering (section 2.4) because it could handle various requirements properly and efficiently. Furthermore, we wish to proceed with various theoretical, simulated and real-time scenario setup of Solid-Disc clustering approach.

REFERENCES

Amis, Alan D.; Ravi Prakash; T.H.P. Vuong; and D.T. Huynh. 2000. "Max-Min D-Cluster Formation in Wireless Ad Hoc Networks." Proceedings of IEEE INFOCOM'2000, Tel Aviv, (Mar).

Baker, D. J. and A. Ephremides. 1981. "The Architectural Organization of a Mobile Radio Network via a Distributed Algorithm." IEEE Transactions on Communications, COM-29(11):1694–1701, (Nov).

Baker, D.J.; A. Ephremides; and J. A. Flynn. 1984. "The Design and Simulation of a Mobile Radio Network with Distributed Control." IEEE Journal on Selected Areas in Communications, pages 226–237.

Bandyopadhyay, Seema and Edward J. Coyle. 2003. "An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks." In the Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM).

Bao, L. and J.J. Garcia-Luna-Aceves. 2003. "Topology management in ad hoc networks." In proceedings of the 4th ACM MobiHoc, Annapolis, Maryland, (Jun).

Banerjee S .and S. Khuller. 2001. "A Clustering Scheme for Hierarchical Control in Wireless Networks." In Proceedings of IEEE INFOCOM, (Apr).

Chan, H. and A. Perrig. 2004. "ACE: An Emergent Algorithm for Highly Uniform Cluster Formation." In European Workshop on Sensor Networks. pp. 154-171.

Chan, H. and Mark Luk, A. Perrig. 2005. "Using Clustering Information for Sensor Network Localization." In The International Conference on Distributed Computing in Sensor Systems (DCOSS 2005).

Das, B. and V. Bharghavan. 1997. "Routing in Ad-Hoc Networks Using Minimum Connected Dominating Sets." In Proceedings of IEEE International Conference on Communications, Implicit Computational Complexity (ICC), ISBN 0-7803-3925-8, Volume 1, pp.376-380.

Demirbas, M.; A. Arora and V. Mittal. 2004. "FLOC: A fast local clustering service for wireless sensor networks." Workshop on Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks (DIWANS/DSN).

Er, I.I. and W. K. G. Seah. 2004. "Mobility-based d-Hop Clustering Algorithm for Mobile Ad Hoc Networks." In Proceedings of Wireless Communications and Networking Conference (WCNC'04), (Mar).

Ephremides, A.; J. E. Wieselthier; and D. J. Baker. 1987. "A Design Concept for Reliable Mobile Radio Networks with Frequency Hopping Signaling." Proceedings of IEEE, 75(1):56–73.

Fisher, David A. and Howard F. Lipson. 1999. "Emergent Algorithms: A New Method for Enhancing Survivability in Unbounded Systems." In Proceedings of the Hawaii International Conference on System Sciences, (Jan).

Gerla, M. and J. T.-C. Tsai. 1995. "Multicluster, mobile, multimedia radio network." ACM Baltzer Journal of Wireless Networks, 1(3): pp255–265.

Heinzelman, W. B.; A. P. Chandrakasan; and H. Balakrishnan. 2002. "Application specific protocol architecture for wireless microsensor networks." IEEE Transactions on Wireless Networking.

Johnson, D.B. and D.A. Maltz. 1996. "Dynamic source routing in ad hoc wireless networks." Mobile Computing, Academic Publishers. pp. 153-181.

Lauer, G. 1986. "Hierarchical routing design for SURAN." In Proceeding of ICC. pp. 93–101.

Lauer, G. 1995. "Packet-radio routing." In: Routing in Communication Networks, ed. M. Steenstrup (Prentice-Hall, Englewood Cliffs, NJ).

Lee, S.-J.; W. Su; and M. Gerla. 2000. "On-Demand Multicast Routing Protocol (ODMRP) for Ad Hoc Networks." Internet Draft, DRAFT-IETF-MANET-ODMRP-02.txt, (Jan). Available at <http://www.cs.ucla.edu/NRL/wireless/PAPER/draft-ietf-manet-odmrp-02.txt>

Lin, C.R. and M. Gerla. 1997. "Adaptive clustering for mobile wireless networks." IEEE Journal on Selected Areas in Communications, vol. 15, no. 7, pages 1265-1275, (Sep).

Nagpal. R. and D. Coore. 1998. "An algorithm for group formation in an amorphous computer." Proceedings of the Tenth International Conference on Parallel and Distributed Systems (PDCS), (Oct).

Perkins, Charles E. and Elizabeth M. Royer. 1999. "Ad hoc On-Demand Distance Vector (AODV) Routing." Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, LA, pp. 90-100, (Feb).

Rappaport T. 1996. "Wireless Communications: Principles and Practice." Prentice Hall, New Jersey, USA.

Shacham, N. and J.Westcott. 1987. "Future directions in packet radio architectures and protocols." In Proceeding of IEEE 75(1). pp 83–98.

Srinidhi, T.; G. Sridhar; and V. Sridhar. 2003. "Topology management in ad hoc mobile wireless networks." Real-Time Systems Symposium, Work-in-Progress Session, Cancun, Mexico, (Dec 3rd).

Tan, H.; W. Zeng; and L. Bao. 2005. "PATM: Priority-based Adaptive Topology Management for Efficient Routing in Ad Hoc Networks." The International Conference on Computational Science 2005 (ICCS), Atlanta, GA, (May)

Xu, Y.; J. Heidemann; and D. Estrin. 2001. "Geography-informed energy conservation for ad-hoc routing." In proceedings of MobiCom 2001, Rome, Italy, pp. 70-84.

Xu, Kaixin and Mario Gerla. 2002. "A Heterogeneous Routing Protocol Based on a New Stable Clustering Scheme." IEEE MILCOM 2002, Anaheim, CA, (Oct).

Younis, Mohamed; Moustafa Youssef; and Khaled Arisha. 2003. "Energy-Aware Management for Cluster-Based Sensor Networks." Journal of Computer Networks, Volume 43, Issue 5, Pages 539-694 (5 Dec).

Younis, S. and S. Fahmy. 2004. "Distributed clustering in ad-hoc sensor networks: A hybrid, energy efficient approach." In Proceedings of IEEE INFOCOM., Hong Kong, China, (Mar).

Yi, Y; Mario Gerla; and TaekJin Kwon. 2003 "Efficient Flooding in Ad Hoc Networks using On-Demand (Passive) Cluster Formation." In proceedings of the Second Annual Mediterranean Ad Hoc Networking Workshop (Med-hoc-Net 2003)

Yi, Y.; M. Gerla; and T.J. Kwon. 2002. "Efficient flooding in ad hoc networks using on-demand (passive) cluster formation". Proceedings of the 3rd ACM MobiHoc, Lausanne, Switzerland, (June 9-11).

Zhang, H. and A. Arora. 2002. "GS^3: Scalable Self-configuration and Self-healing in Wireless Networks." In 21st ACM Symposium on Principles of Distributed Computing (PODC 2002).

AUTHOR BIOGRAPHIES

AMEER AHMED ABBASI was born in Karachi, Pakistan. He went to the Sheikh Zayed Research Centre, University of Karachi, where he studied IS with computer technology and obtained his Master's degree in 2000. He worked one and half years for the Usman Business Solution Inc. Karachi, Pakistan as a software engineer. In 2001 he moved to the Al-Hussan Institute of Management & Computer Science, Dammam, Saudi Arabia. He is teaching computer science courses and participating in research in the fields of mobile computing, ad-hoc and wireless sensor networks. His e-mail address is: ameer_abbasi@hussan.edu.sa and his personal Web-page can be found on the Internet at http://ameerabbasi.topcities.com

M I BUHARI is currently working as lecturer at Information and Computer Science Department, King Fahd University of Petroleum and Minerals, Saudi Arabia. He holds PhD degree in Information Technology from Multimedia University [Malaysia], 2003 in addition to Cisco Certification of CCNA. He has more than 30 International publications. His areas of interest are Computer Networks, Operating Systems, and Parallel Processing. His e-mail address is: mibuhari@ccse.kfupm.edu.sa and his personal Web-page can be found on the Internet at http://www.ccse.kfupm.edu.sa/~mibuhari

M AKBAR BADHUSHA is a PhD candidate at University of Technology Sydney, Australia. He received his Bachelor of Engineering (Computer Engineering) degree from Madurai Kamaraj University and Master of Engineering (Computer Science and Engineering) degree from PSG College of Technology, Bharathiar University, Coimbatore, India. His research interest is in the area of Internet Service Architecture and specifically focused on Programmable Service Platform and programmable/active services. His e-mail address is: akbar@it.uts.edu.au