软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2011,22(5):1053−1066 [doi: 10.3724/SP.J.1001.2011.00586]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

Ad Hoc 网络和无线传感器网络中连通支配集的分布式构造*

郑 婵[1,2+], 孙世新[1], 黄天云[3]

[1](电子科技大学 计算机科学与工程学院,四川 成都 610054)
[2](华南农业大学 信息学院,广东 广州 510642)
[3](西南民族大学 计算机科学与工程学院,四川 成都 610041)

Constructing Distributed Connected Dominating Sets in Wireless Ad Hoc and Sensor Networks

ZHENG Chan[1,2+], SUN Shi-Xin[1], HUANG Tian-Yun[3]

[1](College of Computer Science and Engineering, University of Electronic Science and Technology, Chengdu 610054, China)
[2](College of Information, South China Agricultural University, Guangzhou 510642, China)
[3](College of Computer Science and Technology, Southwest University for Nationalities, Chengdu 610041, China)
+ Corresponding author: E-mail: zheng.chan@163.com

Zheng C, Sun SX, Huang TY. Constructing distributed connected dominating sets in wireless ad hoc and sensor networks. *Journal of Software*, 2011,22(5):1053−1066. http://www.jos.org.cn/1000-9825/586.htm

Abstract: Ad hoc and sensor networks have a wide range of potential, practical and useful applications. Efficient hierarchical cluster organization plays an important role in wireless network systems without centralized control or infrastructure. In cluster-based wireless networks, in order to select a few dominating nodes to form a virtual backbone that supports routing and other tasks, such as broadcasting, area monitoring, etc., most of the published research regarding construction of connected dominating sets (CDS) in the literature has focused on selecting a small nodes set for high efficient performance based on different assumptions and design objectives. A review on more than 20 CDS construction mechanisms based on various assumptions, design objectives and performance results is provided in this paper. Some future research directions in this area are also pointed out.
Key words: connected dominating set (CDS); performance ratio; mobile ad hoc networks; wireless sensor networks

摘 要: Ad hoc 网络和无线传感器网络具有广泛的应用,但对于这样自组性的网络须采用分层结构的聚簇来有效管理.通过选择具有支配属性的节点构成虚拟主干以支持路由、广播及覆盖等应用.大部分的研究都集中在高效选择较小的连通支配集.全面阐述了连通支配集构造的研究进展,并依据不同的网络假设、设计目标和性能对超过 20 种连通支配集的构造算法进行分类和总结.指出这一领域的研究方向.
关键词: 连通支配集;性能比率;移动 Ad Hoc 网络;无线传感器网络
中图法分类号: TP393 **文献标识码**: A

---

* 基金项目: 国家高技术研究发展计划(863)(2006AA10Z246); 西南民族大学重点基金项目(09NZD001)
收稿时间: 2007-07-26; 定稿时间: 2010-03-15

1 Introduction

Mobile ad hoc networks (MANETs) consist of a large number of wireless hosts that communicate with each other in the absence of a fixed infrastructure. Wireless sensor networks (WSNs) can be seen as a special case of ad hoc networks with lower mobility and tighter energy budget. In a WSN, the interconnected units are battery-operated microsensors, each of which is integrated in a single package with low-power signal processing, computation, and a wireless transceiver. Wireless ad hoc and sensor networks can be used wherever a wired backbone is infeasible and/or economical in convenience, for example, to provide communications during emergencies, biological detection, special events (expos, concerts, etc.), or in hostile environments[1,2].

1.1 Application of connected dominating sets in wireless networks

Organizing a network into a hierarchical structure could make management efficient. Clustering offers such a structure, and it suits networks with relatively large numbers of nodes. The structure of clustering consists of three types of nodes: normal nodes, clusterheads (CHs), and gateway nodes. The presence of gateway node is not compulsory in cluster scheme. In this way, clusterheads are well-spread throughout the network. In general, most of the protocols proposed by researchers end up generating a clustering and a corresponding backbone whose nodes form a dominating set (DS) in network topology. Each node that is not in the backbone has at least a backbone node as its neighbor. Moreover, Connected Dominating Sets (CDSs) are prominently used due to the property of connectivity in these large-scale wireless ad hoc and sensor networks. For efficiency reasons, many of networks protocols are thus organized by CDSs. These protocols address media access[3], clustering[4−7], flooding/routing[8−11], multicast/broadcast[12−16], power management[17−21] and coverage/monitoring[22,23].

In addition, wireless ad hoc and sensor networks usually lack a central control instance, and the large problem size (e.g. hundreds of thousands of sensors) also prohibits the centralized CDS computation. Therefore, local and distributed algorithms are designed to operate in such a scenario.

1.2 Network model and graph notation

From a theoretical graph point of view, dominating sets that model wireless communication networks are of outstanding interest. The graph notations of these theoretical aspects, which may be mentioned in subsequent sections, are described and listed in detail in Table 1 of this section.

Table 1 A list of graph notations involved

Notations	Descriptions
General graph	$G=(V,E)$, where V represents the nodes and E the set of wireless links, i.e. there is an edge $(u,v)\in E$ if and only if a transmission from node $u\in V$ can be received by a neighbor $v\in V$.
Unit disk graph (UDG)	Suppose that all nodes send with the same signal strength and thus have the same maximum circular coverage area with radius $c>0$. For each node $v\in V$, its location $f(v)$. The set of edges then satisfies the following simple characterization: $(u,v)\in E\Leftrightarrow f(u)-f(v)\leqslant c$.
Open neighbor set	$N(u)=\{u\|(u,v)\in E\}$, is the set of nodes that are neighbors of u.
Closed neighbor set	$N[u]=N(u)\cup\{u\}$, is the set of neighbors of u and u itself.
Maximum degree	Δ, is the maximum count of edges emanating from a single node.
Independent set (IS)	A subset of V such that no two vertices within the set are adjacent in V.
Maximal independent set (MIS)	An independent set such that adding any vertex not in the set breaks the independence property of the set. Thus, any vertex outside of the maximal independent set must be adjacent to some node in the set.
Dominating set (DS)	S is defined as a subset of V such that each node in $V-S$ is adjacent to at least one node in S. Thus, every MIS is a dominating set. However, not every dominating set is an MIS. Finding a minimum-sized dominating set or minimal dominating set (MDS) is NP-Hard.
Connected dominating set (CDS)	C is a dominating set which induces a connected subgraph of G.
Minimum connected dominating set (MCDS)	MCDS is the CDS with minimum cardinality. Finding the MCDS is also NP-Hard.

Table 1 A list of graph notations involved (continue)

Notations	Descriptions
k-Connected *k*-dominating set (simply called *k*-CDS)	A node set is *k-dominating set* (simply called *k*-DS) if every node is either in the set or has *k* neighbors in the set. A graph is *k-vertex connected* if removing any *k*−1 nodes from it dose not cause a partition. *k-CDS* is a *k*-dominating set if its induced subgraph is *k*-vertex connected.
Steiner tree	Steiner tree is a minimum weight tree connecting a given set of vertices in a weighted graph. Finding the Steiner tree is also NP-Hard.
Performance ratio	The performance ratio is the maximum value of $\|C\|/opt$, $\|C\|$ is the size of the computed CDS, and *opt* is the size of optimal MCDS of *G*.

1.3 Overview

In this paper, we survey some recent contributions addressing CDS constructions in the context of MANETs and WSNs. We present and classify various CDS construction according to formulations, their assumptions of networks, their design objectives as well as the solutions proposed. The following characteristic parameters are taken into considerations: Performance ratios/bounds; time and message complexities; degree of localization; energy-efficient topology; nodal movement; deterministic, probabilistic or hybrid scheme, backbone robustness, and designing for k-CDS or for d-hop clustering. We conclude with a discussion of these CDS problems.

The rest of the paper is organized as follows. We give classifications of the CDS constructions as a whole in Section 2, followed by detailed explanations and analyses of these typical distributed construction mechanisms in Section 3. In Section 4, we conclude and outline some future directions that motivate us to undertake further studies in this field.

2 A Classification of Distributed CDS Construction

Table 2 summarizes the CDS construction approaches covered in this paper according to several classificatory characteristics presented above, and simply describes related design techniques of every classification.

Table 2 Categorization of CDS construction approaches

CDS construction type	Typical algorithms & Refs.	Main design techniques & objectives
Greedy-Based	Das's-I, Das's-II[8]	1) Greedy method with maximum effective degree selection 2) Need to be provided global node information of network, at least 2-hop neighborhood
MIS-Based	Single leader algorithm: Alzoubi and Wan[24,25]; Cheng, *et al.*[26,27] Multiple leader algorithm: Alzoubi, *et al.*[28]; Bao, *et al.*'s TMPO[18]	1) Compute and connect an MIS to forming CDS 2) Be divided into single leader and multiple leaders selection 3) Usually have good performance bound and time/message complexities, but not purely localization in a strict sense
Pruning-Based	Wu's rule 1 & rule 2[9]; Dai's rule-*k*[29]; Span[17]; Ingelrest's enhance CDS[30]	1) Marking process and pruning rule to reduce redundant 2) The first purely localized CDS construction heuristic, and CDS maintenance easier
Multipoint relaying based	MPR[12]; MPR-CDS[31]	1) Selecting CDS from a multipoint relay set 2) Be generally used for broadcasting control to decrease overhead
Steiner tree based	ST-MSN[32]	1) Using a Steiner tree with minimum number of Steiner points to connect a dominating set, usually an MIS 2) Getting the best performance ratio to our knowledge with 6.8-approximation algorithm of CDS 3) Requiring large number of message exchange
Probability based	Gossip[10,33]	1) A typical probabilistic scheme 2) Generally has a satisfying scalability, required few operations and the information of the network 3) Not sure to form a CDS, but construct a CDS with very high probability if value *p* is selected advisable
k-CDS	*k*-Gossip[34]; *k*-coverage[34]; CBKC[34]	1) Constructing a *k*-connected *k*-dominating set (*k*-CDS) 2) Based on the coverage condition proposed by Wu, *et al.* 3) CBKC is a hybrid algorithm of probabilistic and deterministic schemes
d-Hop CDS	*d*-LowestID/*d*-CONID[6]; AC-LMST[4]	1) Forming d-hop cluster adopted local MST technique 2) Decreasing large number of clusters

Table 2 Categorization of CDS construction approaches (continue)

CDS construction type	Typical algorithms & Refs.	Main design techniques & objectives
Considering energy-efficient topology maintenance	LEACH[21]; Span[17]; Extended rules[35]; Combined metrics[19]; TMPO[18]; gossip[10]	1) One of designs is to rotate the role of clusterheads to prolong the average lifespan of each node 2) Residual energy level can be used as node priority
Considering nodal mobility	(α,t)-Clustering[36]; ALM[37]; MobDHop[38]	1) High nodal mobility adaptive 2) Different of mobility model and mobility metric 3) Suitable for moving in group

Here n and m are the number of vertices and edges respectively, Δ_2 is the maximum number of 2-hop nodes a 1-hop node may cover, and the harmonic function: $H(\Delta)=\sum_{i=1}^{\Delta} 1/i$. Other symbols are refered to Table 1.

In Table 3, the performance features of CDS constructions are summarized and compared. All the surveyed methods have a common objective that constructing a CDS is to form backbones of clusters. They also have different design approaches that are determined by their applications or assumptions. Therefore, a fair performance comparison among the surveyed mechanisms should be taken into consideration.

Table 3 Features of approaches listed in Table 2 (1)

Approaches	Graph model	Ngh. info.	Priority of CH selection	Local maintenance	Deterministic method
Ref.[8]-I	General	Global	Effective degree	Distributed, non-local	Deterministic
Ref.[8]-II	General	2-Hop	Effective degree	Distributed, non-local	Deterministic
Ref.[24,25]	UDG	1-Hop	ID-Based/Level-Based	Distributed, non-local	Deterministic
Ref.[26]	UDG	1-Hop	Effective degree	Distributed, non-local	Deterministic
Ref.[26,28]	UDG	1-Hop	ID-Based	Distributed, non-local	Deterministic
TMPO[18]	UDG	2-Hop	Base on "willingness"	Distributed, local	Deterministic
Rule 1 & 2[9]	General	2-Hop	ID-Based	Distributed, local	Deterministic
Rule-*k*[29]	General	2-Hop/Global	ID-Based	Distributed, local	Deterministic
Span[17]	General	3-Hop	Base on "backoff delay"	Distributed, local	Deterministic
Enhance CDS[30]	UDG	2-Hop	Energy-base	Distributed, local	Deterministic
MPR[12,31]	UDG	2-Hop	ID-Based	Distributed, local	Deterministic
ST-MSN[32]	UDG	2-Hop	Effective degree and ID	Distributed, non-local	Deterministic
Gossip[10,33]	General	None	Probability p	Distributed, local	Probabilistic
k-Gossip[34]	General	None	Probability p	Distributed, local	Probabilistic
k-Coverage[34]	General	2-Hop	ID-Based	Distributed, local	Deterministic
CBKC[34]	General	2-Hop	ID-Based	Distributed, local	Hybrid
d-LowestID/*d*-CONID[6]	UDG	*d*-Hop	LowestID-Based/Connectivity-Based	Distributed, local	Deterministic
AC-LMST[4]	General	2d+1 hop	ID-Based	Distributed, local	Deterministic
LEACH[21]	General	None	Desired percentage	Distributed, local	Probabilistic
ALM[37]	General	2-Hop	Relative mobility	Distributed, local	Deterministic
MobDHop[38]	General	*d*-Hop	Relative mobility	Distributed, local	Deterministic

Table 3 Features of approaches listed in Table 2 (2)

Energy efficiency	Network mobility	Network scalability	Time complexity	Msg. complexity	Approx. performance ratio
Low	Weak	Small	$O(\|C\|(\Delta+\|C\|))$	$O(n\|C\|)$	$2H(\Delta)$
Low	Weak	Small	$O(\Delta(n+\|C\|))$	$O(n\|C\|+m+n\log n)$	$2H(\Delta)+1$
Low	Weak	Small	$O(n)$	$O(n\log(n))$	$8opt+1$
Low	Weak	Small	$O(n)$	$O(n)$	$8opt$
Moderate	Moderate	Moderate	$O(n)$	$O(n)$	$192opt+48$
Strong	Moderate	Large	N/A	N/A	N/A
Moderate	Moderate	Large	$O(\Delta^3)$	$\theta(m)$	$O(n)$
Moderate	Moderate	Large	$O(\Delta^2)$ (restricted)	$O(\Delta)$	$O(n)$
Strong	Moderate	Large	$O(\Delta^3)$	N/A	N/A
Strong	Moderate	Large	N/A	N/A	N/A
Moderate	Moderate	Large	$O(\Delta_2^2)$	N/A	$O(\log\Delta_2)$
Moderate	Weak	Small	High	High	$6.8opt$
Strong	Strong	Large/Easy	Very Low	None	No hard guarantee
Strong	Strong	Large/Easy	Very Low	None	No hard guarantee

Table 3 Features of approaches listed in Table 2 (2) (continue)

Moderate	Moderate	Moderate	$O(k\Delta^4)$	$O(\Delta)$	Non-Proved
Moderate	Strong	Large	$O(\Delta^3)$	$O(\Delta)$	$O(1)$
Moderate	Moderate	Moderate	N/A	N/A	N/A
Moderate	Moderate	Moderate	N/A	N/A	N/A
Strong	Strong	Large/Easy	Very low	None	N/A
Moderate	Strong	Moderate	N/A	N/A	N/A
Moderate	Strong	Moderate	$O(n)$	$O(n)$	N/A

3 Typical Distributed Construction Algorithms by Classification

Several typical CDS construction methods based on our classification are presented in detail in this section.

1. What are common assumptions of the wireless networks? Assumptions mainly include (1) network models (e.g., general graph, UDG, and Quasi-UDG; 2-D and 3-D topology; directed and undirected graphs), (2) sensor initial deployment strategy (e.g., randomly and uniformly distributed), (3) total number of sensor nodes (e.g., finite and infinite), (4) the level of nodes density (e.g., sparse, dense, and highly dense), (5) mobility (e.g., stationary and mobile network), and (6) location information (whether need geographic location);
2. What are designing objectives of an algorithm? These objectives involve the hop numbers of neighbor information; priority of clusterheads selection; degree of localization; difficulty of maintenance and scalability; deterministic, probabilistic or hybrid approach; energy-efficient topology, nodal movement, or backbone robustness; k-dominating clusters or d-clusters ($2d$+1 diameter);
3. What are performance metrics of an algorithm? The performance metrics of an algorithm include performance ratio (the number of CDS constructed), time and message complexity, and so on.

In the following subsections, these CDS construction methods are analyzed in detail based on above questions.

3.1 Greedy CDS construction

Das, *et al*. proposed two greedy distributed algorithms for general graphs[8]. The first algorithm selects a CDS from one or two nodes with maximum effective degree that is the number of unmarked neighbors of a node. Then each iterative step selects either a one- or two-edged path emanating from the current CDS, which is added to the extensional nodes with the highest combined effective degrees to the current fragment. Therefore, obviously in first step, a node must know the degrees of all nodes in the graph, while in the iterative step the nodes in the CDS must know the number of unmarked neighbors for all nodes one and two hops from the CDS. This algorithm generates a CDS with a performance ratio proximately of $2H(\Delta)$ in $O(|C|(\Delta+|C|))$ time, using the $O(n|C|)$ messages. The second algorithm first finds a small dominating set S. After S is determined, the edges $(u,dom(u))$ form a spanning forest of G. The next stage of the algorithm connects the fragments by using a distributed minimum spanning tree (MST) algorithm. A connected dominating set C consists of the interior nodes of the resulting spanning tree. This algorithm has performance ratio of $2H(\Delta)+1$ in $O((n+|C|)\Delta)$ time, using $O(n|C|+m+n\log n)$ messages.

3.2 MIS based CDS construction

The MIS based algorithms take the advantage of the relationship between an MIS and MCDS in UDGs. Algorithms in this category usually have good performance bound and time/message complexities. They only need 1-hop neighborhood information. Some algorithms compute an MIS based on either single leader[24−27], or multiple leaders[28]. However, the single leader based algorithms require leader election that makes them difficult to support localized CDS maintenance, although provided better performance bounds. Compared to single leader algorithms, multiple leader based algorithms are optimal in message complexity and relatively more practical in local maintenance since they obviate the rooted spanning tree construction.

Based on single leader selection, Alzoubi and Wan's algorithms provided two versions (the ID-based approach and the level-based approach) to construct a CDS[24,25]. In ID-based approach, the rank of each node is its own ID; while in Level-Based approach, an arbitrary spanning tree (ST) is constructed before the rank assignment. Fig.1 illustrates the execution scenario of two phases of the algorithm (the construction of MIS: Fig.1(a), Fig.1(b) and dominating tree: Fig.1(b), Fig.1(c)). The distributed algorithm gets a size of CDS at most 8*opt*+1, that is, it has an approximation constant factor of at most 8, time complexity $O(n\Delta)$, and message complexity of $O(n\log(n))$[25]. Compared with the work of Alzoubi, *et al*.[24,25], Cheng, *et al*.[26,27] introduced a new active state for vertices to describe the current labeling status. The latter has the same time and message complexity and performance ratio.

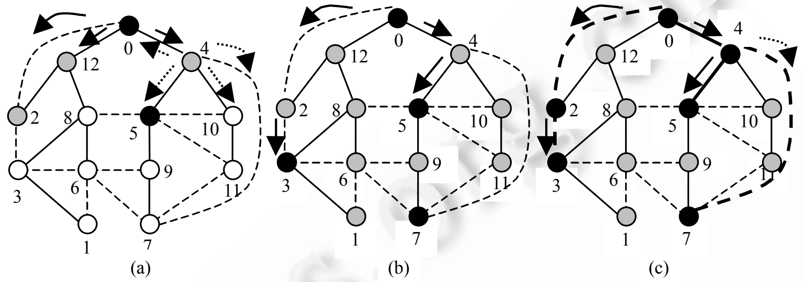

Fig.1 An example of single leader algorithms for CDS construction (from Ref.[25])

Multiple leader selection can decrease the protocol overhead caused by the large volume of message exchange, and improve the structure of the computed CDS. The algorithm introduced by Alzoubi, *et al*. is a representative[28]. The algorithm constructs a CDS in a UDG with the size at most 192*opt*+48, $O(n)$ time complexity, and $O(n)$ message complexity. It is generally faster than the single leader algorithm. Since the worst case time complexity of the multiple leader is $O(n)$, the algorithm is not purely localized in a strict sense. In Bao, *et al*. (called TMPO)[18], each node determines locally for itself based on 2-hop neighbor information whether is the membership of the MDS, similar multiple leader algorithm relying on MIS mentioned above. And then, each node determines itself doorway node or gateway node to connect MDS with a CDS.

3.3 Pruning based CDS construction

Wu et al. proposed a completely localized marking process to construct CDS in general graphs[9]. The marking process is as follows: Every v assigns itself marker T (marked) if there exist two unconnected neighbors. Such a simple marking rule makes the formed CDS with a lot of redundant nodes. Wu et al. also provided two pruning principles based on the neighborhood subset coverage to post-process the redundant result, called Rule 1 and Rule 2, respectively. These two pruning ideas were generalized by Dai, *et al*. to Rule k (based on k-neighbor coverage)[9]: A node u can be deleted from S if there exist k connected neighbors with higher IDs in S that can cover all u's neighbors. Rule 1 and Rule 2 are special cases of Rule k.

Since $N[v]\subset N[u]$ (Fig.2(a)), or $N[v]=N[u]$ (Fig.2(b)), vertex v is removed if $id(v)<id(u)$, and u is the only dominating node. To ensure one and only one is removed, the one with smaller ID is picked, called the selective removal process based on node ID. In Fig.2(c), the condition $N(v)\subseteq N(u)\cup N(w)$ in Rule 2 implies that u and w are connected. Therefore, node v can be deleted applying Rule 2 if $id(v)<id(u)$ and $id(v)<id(w)$.

It was claimed that the total message complexity is $O(n\Delta)$ and the time complexity at each node is $O(\Delta^2)$[9]. It was proved that Wu's complexity analysis is not accurate or not correct[24]: The performance ratio for Wu's algorithm is $O(\Delta)$, the corrected time complexity is $O(\Delta^3)$, and more accurate message complexity is $\Theta(m)$, where m

is the number of edges. In fact, one major advantage of Wu's algorithm is its locality of CDS maintenance for mobile hosts' movements. His algorithm only requires neighbors of a node to update their status when it changes switching status (off or on) or location.

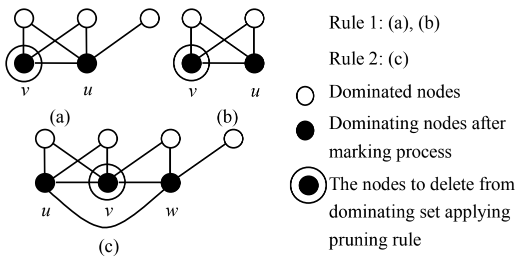

Fig.2 An example of two pruning principles to eliminate redundant nodes (from Ref.[35])

Chen, *et al*. proposed the Span protocol[17] to construct a set of forward nodes (called coordinators), whose selective rule is similar to Ref.[9]. Ingelrest, *et al*. presented a new enhanced notion[30] that some 2-hop neighbors may be used to cover 1-hop neighbors for computing a dominating set. They also proved that the obtained set is a subset of Wu's heuristics[9], and is always dominating and connected for any connected graph.

3.4 Multipoint relaying based CDS construction

Multipoint relaying (MPR) is mainly used for broadcast control to decrease flooding protocol overhead[12]. It's proved that the computation of a multipoint relay set with minimal size is NP-complete[12]. The idea of the multipoint relay technique is to compute local (2-hop neighborhood) dominating sets. Each node *u* computes a multipoint relay set that is included in the 1-hop neighborhood of the node *u*, and covers all the 2-hop neighbor of the node *u*, in other words, each node *u* select a forwarding set from $N(u)$ to cover $N(N(u))$ (see Fig.3).

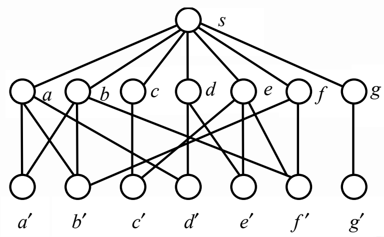

Fig.3 An example of a multipoint relay set for node s is the set {*a*,*e*,*g*}

A greedy heuristics with the ratio of optimal minimum MPR set was designed in Ref.[12]. Its approximation performance factor is $O(\log\Delta_2)$, and computation takes $O(\Delta_2^2)$ time, where Δ_2 is the maximum number of 2-hop nodes a 1-hop node may cover. Adjih, *et al*. proposed a localized heuristic to generate a CDS based on MPR[31]. The correctness of this heuristic was proved but no performance analysis was available. Their idea is sketched that each node first computes a multipoint relay set, a subset of 1-hop neighbors that can cover all the 2-hop neighbors, and then a CDS is constructed by two rules (called MPR-CDS rule). The MPR determining algorithm is source-independent, and involves only node ID and neighborhood of nodes. Wu noticed that some nodes selected by using his two rules were not essential for a CDS. Hence, in Wu's literature, two extended versions of MPR (MPR-E) were proposed for obtaining a smaller CDS than the original MPR[31] without imposing any additional communication overhead[39].

3.5 Steiner tree based CDS construction

Among all proposed approximation for the MCDS in a UDG, Wan, *et al*. achieved the best previously known performance ratio of 8[24]. Subsequently, a 6.8-approximation algorithm of CDS construction based on Steiner tree was further presented by Min, *et al*.[32], which also can be implemented distributed.

The Steiner Tree with Minimum Number of Steiner Nodes (ST-MSN) problem in UDGs had not been studied much. However, its geometric version in the Euclidean plane had been studied extensively, such as ST-MSN in Euclidean plane is NP-hard. Fortunately, a 3-approximation for ST-MSN can be extended from the Euclidean plane to UDGs, which becomes the fundament of Min, *et al*.'s algorithm[32]. His algorithm consists of two steps, which obtain CDS from dominating set by selecting connector nodes, similar to Wan's[24] and Cheng's[26]. The main difference is to employ the Steiner tree to do the job in the second step as the principle of selecting connector nodes. Combined with these two steps, the obtained CDS would have the size bounded by 6.8*opt*. It is clear that the distributed implementation has very high message complexity.

3.6 Probabilistic method based construction

A typical probabilistic scheme is the gossip-based algorithm[10,33]. A probabilistic idea is well applied to large networks with a satisfactory scalability, few operations and little information of the network. Each node randomly goes to sleep for some time with gossip sleep probability p[10]. When the value p is selected advisable, the network stays connected. The scheme was originally proposed as one of topology control schemes in wireless networks, but can be extended and modified for virtual backbones construction: each node v has a backbone status with probability p. The selection of backbone nodes in Gossip is purely in random without any neighborhood information. When p is larger than a threshold, these backbones form a CDS with very high probability. Simulation shows that threshold is determined based on experimental data, depending on network size and density.

Some analytical study has provided an upper bound of p_k that would be almost sufficient to guarantee the probability of 1-coverage or k-coverage[40] in a network. Kumar, *et al*. achieved the sufficient condition for asymptotic k-coverage and non-k-coverage with three kinds of deployment in square region: $\sqrt{n}\times\sqrt{n}$ grid deployment, random uniform deployment and Poisson deployment[40]. However, these upper bounds are conservative estimations of the perfect p_k, which usually need adjustments based on experimental data.

3.7 *k*-CDS construction

Much previous works in construction of CDS is focused on selecting a smallest or approximating smallest backbone set for high performance efficiency in wireless network. However, sometimes, it is desirable to have several sensors monitor the same target even if changes of topology occur; it is needed to let each sensor report via different routes to avoid losing important data. Therefore, it is equally important to maintain a certain degree of redundancy in the virtual backbone for fault tolerance or routing flexibility[41−45]. A model of cone-based was introduced and shown for preserving k-connectivity with $\alpha=5\pi/(6k)$[45]. The model was considered of extending to 3-dimensions[41]. A centralized algorithm was proposed for moving a block of nodes to new locations such that the resulting network becomes biconnected[43]. A lot of works suggests maintaining k-vertex connectivity for fault tolerance and high throughput in topology control in wireless networks[41,42,44]. Most of these methods were focused on decreasing nodes degree in turn. The idea was based on a result from Ref.[46]: If a graph has minimum degree k with high probability, it is k-connected with high probability. As far as we know, the localized construction of a k-CDS has not been comprehensively discussed, but a hybrid algorithm of constructing a k-connected k-dominating set (simply called k-CDS) has been proposed by Dai, *et al*.[34].

Virtual backbone nodes are represented by darkened nodes in Fig.4. Every non-backbone node has at least k

neighboring dominating node, and removing any $k-1$ nodes from a k-CDS, the remaining nodes still form a CDS. Therefore, a k-CDS as a virtual backbone can survive failures of at least $k-1$ nodes.

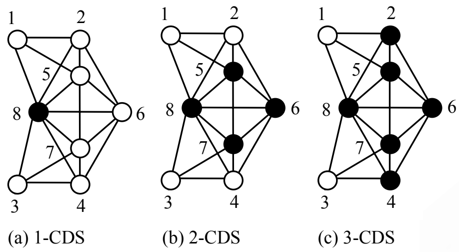

Fig.4 k-CDS constructed with k=1, 2, and 3 (from Ref.[34])

Three localized k-CDS construction schemes: k-Gossip, k-Coverage and hybrid color-base k-CDS construction (CBKC) were proposed by Dai, *et al*.[34]. k-Coverage (Fig.5(b)), which is a deterministic approach, is extended from the 1-CDS Coverage Condition (1-Coverage, Fig.5(a))[16]. It has been proved that applying the k-coverage condition to a k-connected network G, the resultant virtual backbone is a k-CDS[34].

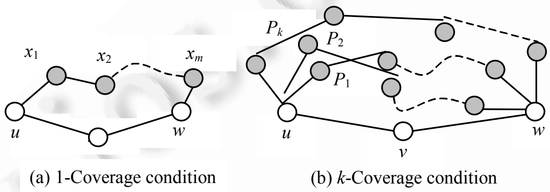

Fig.5 Replacement paths between two neighbors u and w of node v with higher priorities (gray nodes) (from Ref.[34])

As shown in Fig.6, a hybrid algorithm called color-base k-CDS construction (CBKC) was presented[34]. It also proved that the generic scheme almost always constructs a k-CDS in dense networks, with the worst case computation complexity of $O(\Delta^3)$.

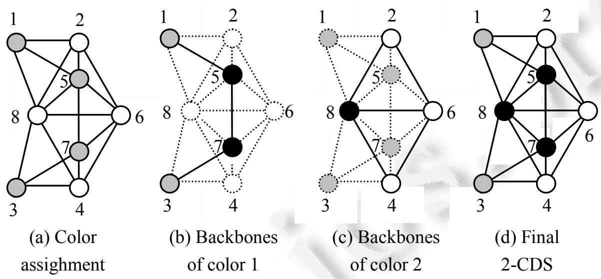

Fig.6 Color-Based coverage condition (from Ref.[34])

3.8 *D*-Hop domination based clustering construction

Most of the existing heuristic clustering algorithms restrict themselves to 1-hop clusters (diameter≤2), which may not be too useful in very large and relatively sparse MANETs since these algorithms may form a large number of clusters. They have been advocated for larger, possibly overlapping clusters. The generalized concept of d-hop cluster was defined by Kim, *et al*.[11] that a cluster is composed of all nodes within d-hop distance from its clusterhead. A d-hop dominating set is a subset that a node either becomes a clusterhead belonging to, or is at most *d* wireless hops away from its clusterhead.

A variant of scheme based on lowest ID was proposed independently by Amis, *et al*.[5] and Nocetti, *et al*.[6] The former showed that the minimum *d*-hop dominating set problem is NP-complete and also generalized a clustering heuristics. It allowed more control and flexibility in determination of clusterhead density. The latter proposed two *d*-hop clustering algorithms, assuming that all nodes are aware of their *d*-hop neighbors through broadcasting. Lou, *et al*. studied a d-hop zone-based broadcast protocol and constructed a CDS during the forward nodes set selection process[47]. Ref.[4] extended the "2.5"-hop coverage theorem[13] by developing a local minimum spanning tree (LMST)[48]. At most 2*d*+1 hops neighborhood information is needed. The parameter *d* is usually small and tunable such that clusters adapting to sufficiently fast reflect the topology changes.

3.9 Considering energy-efficient topology maintenance in CDS construction

Energy saving is one critical issue for sensor networks and ad hoc networks since most nodes are equipped with non-rechargeable batteries that have limited lifetime. Furthermore, nodes in CDS consume generally more energy in order to handle various bypass traffics than nodes outside the set. Therefore, in cluster-based networks, clusterheads are usually selected in a way to minimize the total energy consumption, and they are rotated among all the sensors to balance energy consumption. Many works is spread on the issue to lengthen the network lifetime as an overall goal[10,17−21,35].

Wang, *et al*.[20] surveyed and classified 15 distributed scheduling mechanisms in the literature by design assumptions and design objectives. Heinzelman, *et al*. proposed Low-Energy Adaptive Clustering Hierarchy (LEACH)[21], which utilizes randomized rotation of clusterheads to evenly distribute work load among the sensors. Subsequently, many algorithms are enhanced based on the LEACH to further improve in some major aspects. Span formed a backbone through rotating coordinators in time[17], not only reducing energy consumption but also preserving network connectivity. Considering the power property for all the nodes, Wu, *et al*. extended their work to calculate power-aware CDS as a criterion for the post-pruning[35]. A gossip-based sleep management was proposed to balance energy consumption both in sensor and mobile ad hoc networks[10].

3.10 Considering nodal mobility in CDS construction

Most of the previously related algorithms have attempted to address nodal movements by incorporating a periodical maintenance phase in which the CDS can be constructed when nodes move. However, in the case of high nodal mobility, this reactive approach may not yield a stable infrastructure.

McDonald, *et al*. designed for dynamically organizing mobile nodes into clusters in an MANET[36]. A probability-based (α,t)-clustering that adaptively changes its variable diameter was proposed. Its diameter depends on the degree of current mobility of the nodes: the more the nodes move, the smaller the cluster (easier to maintain), and vice-versa. A weight-based clustering algorithm similar to Lowest-ID is proposed[37]. A mobile awareness based clustering mechanism was proposed for those moving nodes without speed measuring equipments like GPS[49]. A distributed clustering algorithm, named Mobility-based D-Hop (MobDHop)[38], was designed to form dynamically d-hop stable clusters based on the concept of relative mobility[37].

4 Conclusions

CDS constructions play an important role in large-scale wireless ad hoc and sensor networks for efficient hierarchical self-organization. To not only find the commonalities and differences, but also identify the pros and cons within each class, more than 20 recent CDS construction solutions are categorized and analyzed in our classification of network assumptions, design objectives, and performances of algorithms in the paper. Our research shows various design assumption either explicitly or implicitly, design objectives, and various performance results. In general, several relevant metrics of assessing CDS formation algorithms are summarized as follows:

(a) *Degree of localization* is a measure of how much the decision of a node being part of CDS, depending on nodes that are possibly several hops away from it. As representatives of these mechanisms, pruning algorithms and MPR are higher localized than greedy-based, MIS-based, and Steiner Tree based, etc., because of the very nature of sequential implementation of the latter types. The latter proceed in rounds. At each round when a node decides whether to be part of the DS or not, it must keep waiting for decisions made by neighboring nodes at previous round. The degree of localization is thus limited for the possible presence of "chains of dependency" that prevent a node to make a faster decision;

(b) *Performance* metrics are much closely related to the degree of localization. Algorithms including MIS-based and Steiner Tree based possess theoretical properties of producing a CDS with a constant approximation factor, but with a lower degree of localization. The degree of localization does not only affect construction time, but also message complexity. Otherwise, algorithms based on probability do not assure properties of connectivity and domination but implemented fast without any message exchanges of neighboring information. *D*-hop clustering aims at reducing the numbers of clusters and forming larger clusters;

(c) *Backbone robustness* is defined as the numbers of CDS nodes failures (because of failure or energy depletion) which lead to either disconnection or uncovering all the other nodes. It provides an indirect metric on how long the network will be operational before a CDS recomputation is required. All representative algorithms designed for *k*-CDS have strong robustness;

(d) *Energy consumption* and *nodal mobility* are considered as two important designing factors in large-scale mobile wireless networks with energy-constrained nodes. If the maintenance cost of energy-efficiency and mobility is high, clustering and backbone re-organization may impose a non-negligible burden to the network. Therefore, techniques in terms of CHs and gateways rotation are needed to prevent nodes from depleting their energy.

Further researches in the area of CDS constructions in wireless networks may go into the following directions:

(1) Concerning more realistic graph models for wireless networks, there are still many open questions in the field;

(2) A higher level of localization offers many challenges. It still leaves room for major improvements of performances, optimization of algorithms, and further understanding;

(3) Issues that are will likely to arise include security, reliability, and survivability of CDS-based applications;

(4) Topology control of energy-efficient CDS-based scheduling and the impact of group mobility should be carefully investigated;

(5) Implementation is perhaps another open issue in the field. Despite the considerable researches devoted to the field, and many theoretical and simulation-based techniques presented, to date, there is little experimental evidence that CDS applications can actually be used.

An extensive study on multiple-connectivity and multiple-domination is still required to improve the reliable broadcasting in WSN. Further designs and simulations need to be conducted.

Acknowledgement The authors wish to thank the referees for the many constructive suggestions that helped improving the presentation quality of the article.

References:

[1] Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E. Wireless sensor networks: A survey. Computer Networks, 2002,38(4): 393−422. [doi: 10.1016/S1389-1286(01)00302-4]

[2] Arampatzis T, Lygeros J, Manesis S. A survey of applications of wireless sensors and wireless sensor networks. In: Proc. of the 13th Mediterranean Conf. on Control and Automation Limassol. Cyprus, 2005. [doi: 10.1109/.2005.1467103]

[3] Blum J, Ding M, Thaeler A, Cheng X. Connected Dominating Set in Sensor Networks and MANETs. Handbook of Combinatorial Optimizationed. Kluwer Academic Publisher, 2004. 329−369.

[4] Yang S, Wu J, Cao J. Connected *k*-hop clustering in ad hoc networks. In: Proc. of the 2005 Int'l Conf. on Parallel Processing (ICPP 2005). 2005. [doi: 10.1109/ICPP.2005.25]

[5] Amis AD, Praknsh K, Vuong THP, Huynh DT. Max-Min *d*-cluster formation in wireless ad hoc networks. In: Proc. of the IEEE INFOCOM. 2000. [doi: 10.1109/INFCOM.2000.832171]

[6] Nocetti FG, Gonzalez JS, Stojmenovic I. Connectivity based *k*-hop clustering in wireless networks. Telecommunication Systems, 2003,22(1/4):205−220.

[7] Banerjee S, Khuller S. A clustering scheme for hierarchical control in multi-hop wireless networks. In: Proc. of the 20th IEEE Infocom 2001. Anchorage, 2001. [doi: 10.1109/INFCOM.2001.916296]

[8] Das B, Bharghavan V. Routing in ad-hoc networks using minimum connected dominating sets. In: Proc. of the IEEE Int'l Conf. on Communications. Montreal, 1997. [doi: 10.1109/ICC.1997.605303]

[9] Wu J, Li H. On calculating connected dominating set for efficient routing in ad hoc wireless networks. In: Proc. of the 3rd Int'l Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications. 1999. [doi: 10.1145/313239.313261]

[10] Hou X, Tipper D. Gossip-Based sleep protocol (GSP) for energy efficient routing in wireless ad hoc networks. In: Proc. of the Wireless Communications and Networking Conf. (WCNC). 2004. [doi: 10.1109/WCNC.2004.1311631]

[11] Kim D, Ha S, Choi Y. *K*-Hop cluster-based dynamic source routing in wireless packet radio network. In: Proc. of the IEEE VTC. 1998.

[12] Qayyum A, Viennot L, Laouiti A. Multipoint relaying for flooding broadcast message in mobile wireless networks. In: Proc. of the 35th Annual Hawaii Int'l Conf. on System Sciences (HICSS-35 2002). 2002. [doi: 10.1109/HICSS.2002.994521]

[13] Wu J, Lou W. Forward-Node-Set-Based broadcast in clustered mobile ad hoc networks. Wireless Communications and Mobile Computing, 2003,3(2):141−154. [doi: 10.1002/wcm.109]

[14] Stojmenovic I, Seddigh M, Zunic J. Dominating sets and neighbor elimination-based broadcasting algorithms in wireless networks. IEEE Trans. on Parallel and Distributed Systems, 2002,13(1):14−25. [doi: 10.1109/71.980024]

[15] Dai F, Wu J. Performance analysis of broadcast protocols in ad hoc networks based on self-pruning. IEEE Trans. on Parallel and Distributed Systems, 2004,15(11):1027−1040. [doi: 10.1109/TPDS.2004.69]

[16] Wu J, Dai F. A generic distributed broadcast scheme in ad hoc wireless networks. IEEE Trans. on Computers, 2004,53(10): 1343−1354. [doi: 10.1109/TC.2004.69]

[17] Chen B, Jamieson K, Balakrishnan H, Morris R. Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. In: Proc. of the ACM/IEEE Int'l Conf. on Mobile Computing and Networking (MobiCom 2001). 2001. [doi: 10.1023/A:1016542229220]

[18] Bao L, Garcia-Luna-Aceves JJ. Topology management in ad hoc networks. In: Proc. of the MobiHoc. 2003. [doi: 10.1145/778415. 778432]

[19] Shaikh J, Solano J, Stojmenovic I, Wu J. New metrics for dominating set based energy efficient activity scheduling in ad hoc networks. In: Proc. of the WLN. 2003. [doi: 10.1109/LCN.2003.1243206]

[20] Wang L, Xiao Y. A survey of energy-efficient scheduling mechanisms in sensor networks. Mobile Networks and Applications, 2006,11:723−740. [doi: 10.1007/s11036-006-7798-5]

[21] Heinzelman WR, Chandrakasan A, Balakrishnan H. Energy-Efficient communication protocols for wireless microsensor networks. In: Proc. of the Hawaaian Int'l Conf. on Systems Science. Hawaii, 2000.

[22] Cardei M, Wu J. Coverage in Wireless Sensor Networks. Handbook of Sensor Networksed. Boca Raton: Florida Atlantic University, 2004.

[23] Lou W, Wu J. Double-Covered broadcast (DCB): A simple reliable broadcast algorithm in MANETs. In: Proc. of the IEEE Infocom. Institute of Electrical Engineers Inc. (IEEE), 2004. [doi: 10.1109/INFCOM.2004.1354616]

[24] Wan PJ, Alzoubi KM, Frieder O. Distributed construction of connected dominating set in wireless ad hoc networks. In: IEEE Infocom. 2002. [doi: 10.1109/INFCOM.2002.1019411]

[25] Alzoubi KM, Wan PJ, Frieder O. Distributed heuristics for connected dominating sets in wireless ad hoc networks. Journal of Communications and Networks, 2002,4(1):22−29.

[26] Butenko S, Cheng XZ, Du DZ, Pardalos PM. On the construction of virtual backbone for ad hoc wireless network. In: Proc. of the 2nd Conf. on Cooperative Control and Optimization. 2001.

[27] Cardei M, Cheng XY, Cheng XZ, Du DZ. Connected domination in multihop ad hoc wireless networks. In: Proc. of the 6th Int'l Conf. on Computer Science and Informaticss. 2002.

[28] Alzoubi KM, Wan PJ, Frieder O. Message-Optimal connected dominating sets in mobile ad hoc networks. In: Proc. of the MobiHoc. EPFL Lausanne, 2002. [doi: 10.1145/513800.513820]

[29] Dai F, Wu J. An extended localized algorithm for connected dominating set formation in ad hoc wireless networks. IEEE Trans. on Parallel and Distributed Systems, 2004,15(10):908−920. [doi: 10.1109/TPDS.2004.48]

[30] Ingelrest F, Simplot-Ryl D, Stojmenovic I. Smaller connected dominating sets in ad hoc and sensor networks based on coverage by two-hop neighbors. In: Proc. of the Communication Systems Software and Middleware (COMSWARE). 2007. 1−8.

[31] Adjih C, Jacquet P, Viennot L. Computing connected dominated sets with multipoint relays. In: Proc. of the Ad Hoc & Sensor Wireless Networks (AHSWN). 2005.

[32] Min M, Du H, Jia X, Huang C. Improving construction for connected dominating set with Steiner tree in wireless sensor networks. Journal of Global Optimization, 2006,35(1):111−119. [doi: 10.1007/s10898-005-8466-1]

[33] Haas ZJ, Halpern JY, Li LE. Gossip-Based ad hoc routing. IEEE/ACM Trans. on Networking, 2006,14(3):479−491. [doi: 10.1109/TNET.2006.876186]

[34] Dai F, Wu J. On constructing *k*-connected *k*-dominating set in wireless networks. In: Proc. of the 19th Int'l Parallel and Distributed Processing Symp. (IPDPS). 2005. [doi: 10.1109/IPDPS.2005.302]

[35] Wu J, Gaol M, Stojmenovic I. On calculating power-aware connected dominating sets for efficient routing in ad hoc wireless networks. Journal of Communications and Networks, 2002,5(2):169−178. [doi: 10.1109/ICPP.2001.952080]

[36] McDonald AB, Znati TF. A mobility-based framework for adaptive clustering in wireless ad hoc networks. IEEE Journal on Selected Areas in Communications, 1999,17(8):1466−1487. [doi: 10.1109/49.780353]

[37] Basu P, Khan N, Little TDC. A mobility based metric for clustering in mobile ad hoc networks. In: Proc. of the 21st Int'l Conf. on Distributed Computing Systems Workshops (ICDCSW 2001). 2001. [doi: 10.1109/CDCS.2001.918738]

[38] ER II, Seah WKG. Mobility-Based *d*-hop clustering algorithm for mobile ad hoc networks. In: Proc. of the WCNC 2004. 2004. [doi: 10.1109/WCNC.2004.1311457]

[39] Wu J. An enhanced approach to determine a small forward node set based on multipoint relay. In: Proc. of the 58th IEEE Seminannual Vehicular Technology Conf. (VTC 2003-Fall). Orlando, 2003. [doi: 10.1109/VETECF.2003.1286086]

[40] Kumar S, Lai TH, Balogh JO. On *k*-coverage in a mostly sleeping sensor network. In: Proc. of the MobiCom. 2004. [doi: 10.1145/1023720.1023735]

[41] Bahramgiri M, Hajiaghayi M, Mirrokni VS. Fault-Tolerant and 3-dimensional distributed topology control algorithms in wireless multi-hop networks. In: Proc. of the ICCCN. 2002. [doi: 10.1007/s11276-005-5265-z]

[42] Li XY, Wan PJ, Wang Y, Yi CW. Fault tolerant deployment and topology control in wireless networks. In: Proc. of the MobiHoc. 2003. [doi: 10.1145/778415.778431]

[43] Basu P, Redi J. Movement control algorithms for realization of fault tolerant ad hoc robot networks. IEEE Networks, 2004,18(4): 36−44. [doi: 10.1109/MNET.2004.1316760]

[44] Li N, Hou JC. Localized fault-tolerant topology control in wireless ad hoc networks. IEEE Trans. on Parallel and Distributed Systems, 2006,17(4):307−320. [doi: 10.1109/TPDS.2006.51]

[45] Li LE, Halpern JY, Bahl P, Wang YM. A cone-based distributed topology-control algorithm for wireless multi-hop networks. IEEE/ACM Trans. on Networking, 2005,13(1):147−159. [doi: 10.1109/TNET.2004.842229]

[46] Bettstetter C. On the minimum node degree and connectivity of a wireless multihop network. In: Proc. of the ACM MobiHoc. 2002. [doi: 10.1145/513800.513811]

[47] Lou W, Wu J. A *K*-hop zone-based broadcast protocol in mobile ad hoc networks. In: Proc. of the IEEE Globecom 2004. 2004. [doi: 10.1109/GLOCOM.2004.1378264]

[48] Li N, Hou JC, Sha L. Design and analysis of an MST-based topology control algorithm. IEEE Trans. on Wireless Communications, 2005,4(3):1195−1206. [doi: 10.1109/TWC.2005.846971]

[49] Wu HT, Ke KW, Chen CH, Kuan CW. Mobile awareness based cluster selection mechanisms in wireless ad hoc networks. In: Proc. of the Vehicular Technology Conf. (VTC2004-Fall). 2004. [doi: 10.1109/VETECF.2004.1400563]

ZHENG Chan was born in 1978. She is a Ph.D. candidate in Department of Computer Science at University of Electronic Science and Technology of China (UESTC). She is also a lecturer at South China Agricultural University. Her current research interests include wireless networks, distributed computing, and multimedia networks.

SUN Shi-Xin was born in 1940. He is a professor and doctoral supervisor in Department of Computer Science at UESTC. His researches include parallel algorithms, and computing complexity.

HUANG Tian-Yun was born in 1973. He is a professor in College of Computer Science and Technology at Southwest University for Nationalities. His research interests include distributed multimedia system and computer networks.