Chapter 1
Introduction

Most practical questions can be reduced to problems of largest and smallest magnitudes ... and it is only by solving these problems that we can satisfy the requirements of practice which always seeks the best, the most convenient.

P. L. ČEBYŠEV

In this chapter, we introduce basic concepts, fundamental results and applications of connected dominating sets.

1.1 Connected Domination Number

Consider a graph $G = (V, E)$. A subset of vertices, D, is called a *dominating set* if every vertex is either in D or adjacent to a vertex in D. If D, in addition, induces a connected subgraph, then it is called a *connected dominating set (CDS)*. The *connected domination number* of a graph G is the minimum cardinality of a CDS, denoted by $\gamma_c(G)$. A CDS that has the size equal to the domination number is called a *minimum CDS*.

The connected domination number is a classical subject studied in graph theory for many years [94]. Some interesting results are obtained in those earlier efforts. The following are two examples.

Let $\ell(G)$ denote the *max leaf number* of a graph G, which is the maximum number of leaves in a spanning tree of G.

Theorem 1.1.1 (Douglas [35]). *For any graph of order n,*

$$\gamma_c(G) = n - \ell(G).$$

Proof. It is easy to see that for any tree T, $\gamma_c(T) = |V(T)| - \ell(T)$. Moreover, a CDS for a spanning tree T of G is also a CDS for G. Therefore, $\gamma_c(G) \leq n - \ell(G)$.

D.-Z. Du and P.-J. Wan, *Connected Dominating Set; Theory and Applications*, Springer Optimization and Its Applications 77, DOI 10.1007/978-1-4614-5242-3_1,

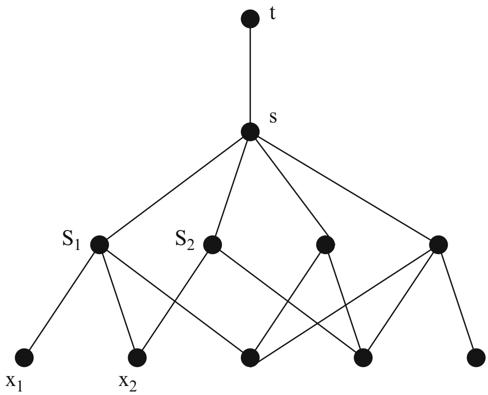

Fig. 1.1 Graph G in the proof of Theorem 1.1.1

Now, consider a minimum CDS D of G. Let H be a spanning tree of $G[D]$ where $G[D]$ is the subgraph of G induced by D. Connect H to every vertex in $V-D$ to obtain a spanning tree T of G. Then, every vertex in $V-D$ is a leaf of T. Conversely, every leaf of T is in $V-D$. Otherwise, if T has a leaf x not in $T-D$, then $D-\{x\}$ would be a CDS for T and hence a CDS for G, contradicting the minimality of D (Fig. 1.1). □

Theorem 1.1.2 (Sampathkumar and Walikar [94]). *Let G be a graph of order* $n \geq 4$. *Suppose both graph G and its complement* $\bar{G}$ *are connected. Then*

$$\gamma_c(G) + \gamma_c(\bar{G}) \leq n(n-3).$$

Proof. Note that a tree has at least two leaves. By Theorem 1.1.1, we have $\gamma_c(G) \leq n-2$. Moreover, G is connected and hence $n-1 \leq |E(G)|$. Therefore

$$\gamma_c(G) \leq n-2 = 2(n-1) - n \leq 2|E(G)| - n.$$

Similarly,

$$\gamma_c(\bar{G}) \leq 2|E(\bar{G})| - n.$$

Thus,

$$\gamma_c(G) + \gamma_c(\bar{G}) \leq 2(|E(G)| + |E(\bar{G})|) - 2n = 2\binom{n}{2} - 2n = n(n-3).$$ □

Laskar and Pfaff [71] showed the NP-hardness of computing the connected domination number or the minimum CDS. Namely, the following problem is NP-hard.

MIN-CDS: Given a graph $G = (V, E)$, find a CDS with minimum cardinality.

Remark. We make a different usage of MIN-CDS from minimum CDS that MIN-CDS is for a problem while the minimum CDS is for a subset of vertices.

Theorem 1.1.3. MIN-CDS *is NP-hard.*

Proof. Consider the following problem.

> SET COVER: Given a collection $\mathcal{C}$ of subsets of a base set X and a positive integer $k \leq |X|$, determine whether $\mathcal{C}$ contains a set cover with cardinality at most k, where a *set cover* is a subcollection $\mathcal{A}$ of $\mathcal{C}$ such that every element of X appears in at least one subset in $\mathcal{A}$.

SET-COVER is a well-known NP-complete problem [57]. We construct a reduction from SET-COVER to MIN-CDS as follows.

For input collection $\mathcal{C}$ and base set X in SET-COVER we first construct a bipartite graph H with $n+m$ vertices labeled by all elements $x_1, u_2, \ldots, x_n$ in X and all subsets $S_1, S_2, \ldots, S_m$ in $\mathcal{C}$. An edge exists between two vertices a and b if and only if $a \in b$ or $b \in a$. Graph G is obtained from H by adding two new vertices s and t and connecting s to t and every S_i for $i = 1, 2, \ldots, m$.

Suppose $\mathcal{C}$ has a set cover $\mathcal{A}$ of at most size k. Then the vertices with labels in $\mathcal{A}$ together with s form a CDS with cardinality at most $k+1$.

Conversely, suppose G has a CDS C of size $k' \leq k+1$. Note that C must contains node s in order to dominating node t or connection t to other vertices in C. Furthermore, we claim that if $a \in C$ for some $a \in X$, then $C - \{a\}$ is still a CDS. In fact, to have a path connecting a and s, there must exist $A \in$ C such that $a \in A$. Thus, a can be dominated by A. Moreover, all vertices dominated by a are also dominated by s. Thus, $C - \{a\}$ is still a CDS. Now, let us denote by C' the CDS obtained from C by deleting t and all elements in X. Then, C' contains s and some vertices labeled by subsets $A_1, A_2, \ldots, A_h$ ($h \leq k' - 1$) in $\mathcal{C}$. These h subsets $A_1, A_2, \ldots, A_h$ must cover all elements in X. Therefore, G has a CDS of size at most $k+1$ if and only if $\mathcal{C}$ has a set cover of size at most k. □

While MIN-CDS in general is NP-hard, a lot of earlier efforts were made on design of polynomial-time algorithms for special class of graphs, such as series-parallel graphs [113] and permutation graphs [25].

This situation was changed after applications of CDS were found in wireless networks and optical networks. Since then, the study of CDS is toward application-oriented research. A plenty of issues are involved which generate many research problems in theory.

1.2 Virtual Backbone in Wireless Networks

To keep nodes in a wireless network being able to communicate each other, the network is required to have certain connectivity. Such a task is called *topological control*. Inspired by physical backbone in classical wired networks, the virtual backbone has been introduced to involve in topological control for wireless networks to reduce the utilization of resource. When the wireless network is formulated as a disk graph, the virtual backbone is a CDS of the graph so that all communications between nodes can be executed through the virtual backbone. In fact, the virtual

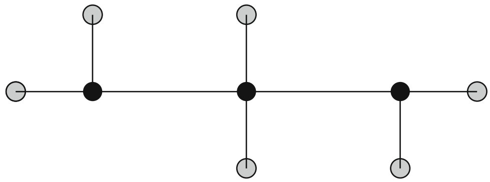

Fig. 1.2 Black nodes form a virtual backbone

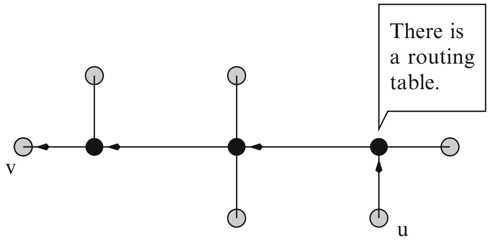

Fig. 1.3 Node u sends a data to node v through a virtual backbone

backbone is required to have two properties: (1) Every node not in virtual backbone should be able directly to communicate with (adjacent to) a node in the virtual backbone. (2) All nodes in the virtual backbone should be able to communicate each other within the virtual backbone, that is, the virtual backbone induces a connected subgraph.

To see the performance of using the virtual backbone, let us consider a network as shown in Fig. 1.2. Note that wireless network has no physical infrastructure. Without using the virtual backbone, every node has to store a routing table in order to be able to communicate with others. With the virtual backbone, only nodes in the virtual backbone need to store a routing table. For nodes not in the virtual backbone, each of them needs to know only an adjacent node in the virtual backbone. In Fig. 1.3, an example is presented to show how node u sends a data to node v through a virtual backbone. Node u first sends the data to its adjacent node w in the virtual backbone C and tells node w that the data is for node v. Since there is a routing table stored at node w, w will figure out a routing path through C to u and then delivers the data along this routing path to node v.

Clearly, we would like to have the virtual backbone as small as possible. This gives a motivation to study MIN-CDS and to design constructions of CDS with small size in various cases [8, 26, 28, 38, 45, 93–95, 98–100, 111, 117–122, 128]. There are several remarkable theoretical results in the literature; each of them made an important progress on the study of CDS as mentioned in the following.

Guha and Khuller [62] designed the first polynomial-time approximation with guaranteed performance ratio $\ln n + O(1)$. They also showed a result on the inapproximability that MIN-CDS cannot have a polynomial-time $\rho \ln n$-approximation for $0 < \rho < 1$ unless $NP \subseteq DTIME(n^{O(\log \log n)})$ where n is the number of vertices in input graph. Improving Guha–Khuller's approximation introduced a study on analysis of greedy approximation with nonsubmodular potential functions. Ruan et al. [92] and Du et al. [40] made a significant contribution in this research direction.

Wan et al. [104] designed the first polynomial-time constant-approximation for MIN-CDS in the unit disk graph which is a mathematical model for homogeneous wireless sensor networks. A plenty of follow-up efforts have been made along this direction.

Cheng et al. [22] designed the first PTAS for MIN-CDS in unit disk graphs, that is a group of polynomial-time $(1+\varepsilon)$-approximation for any $\varepsilon > 0$. This initiates a series of research work on CDS with partition techniques.

Many variations of MIN-CDS or new problems on CDS are proposed recently, motivated from special needs in developments of wireless networking technology. For example Li et al. [75] and Thai et al. [101] proposed the directed CDS; Kim et al. [67] constructed the diameter-bounded CDS; Willson et al. [115] and Ding et al. [32] initiated a study on the routing-cost constrained CDS. Especially, Huang and Gao et al. [53, 66] discovered a technique, double partition, to design better approximations for the weighted CDS problem and related problems. We will study them in later chapters.

1.3 Converter Placement in Optical Networks

One of expectations on next-generation of communication network is to enable people to do remote data gathering and remote scientific experiments. Those applications demand high-speed communication networks with flexible deployment and/or mobile connectivity. One of proposed infrastructures with such properties is the wireless access network on top of the optical core network. Indeed, the optical network in core provides the efficient high-speed communication with high bandwidth and the wireless network in access provides mobile communication or/and flexible deployment. The advantage of fiber-optical backbone network combined with wireless technology has gained more and more interests in the study of the next generation communication network.

An optical network can be considered as a graph $G = (V,E)$ that each edge is associated with a set of wavelengths [60, 73, 81, 82, 88]. The multicast/broadcast/unicast routing requires the existence of a spanning subgraph of G. If a message from an edge to another edge uses different wavelengths, then a converter is required at the common endpoint of the two edges.

Let us use a color to represent a connected component in a subgraph induced by all edges with a certain wavelength. Each converter would connect two connected components into one. To save resource, a minimization problem is formulated [90, 91] as follows.

> CONVERTER PLACEMENT: Given a graph $G = (V,E)$ and color-sets for each edge of G such that for every color all edges in the color form a connected subgraph, find the minimum number of vertices such that placing converters on them would connect some colors into a connected spanning subgraph of G.

CONVERTER PLACEMENT can be reduced to MIN-CDS. To do so, we construct another graph G' with vertex set V. Two vertices u and v are connected with an edge if and only if they are in the same color of G. Without loss of generality, we may assume that no color covers all vertices because in such a case, no converter is required. Under the assumption, we can show that a vertex subset C is a feasible solution for CONVERTER-PLACEMENT if and only if C is a CDS in G'.

First, suppose C is a feasible solution. Then every vertex x must be adjacent to a converter; otherwise, it cannot communicate with any converter which is also a vertex in G. Moreover, C must induce a connected subgraph in G' since, otherwise, two converters in different connected components cannot communicate each other. Therefore, C is a CDS in G'.

Conversely, if C is a CDS in G', then there is a spanning tree T of G' with all internal vertices in C. Thus, placing converters at all vertices in C would connect all colors appearing in T together, which is clearly covering T. Therefore, C is a feasible solution for CONVERTER PLACEMENT.

In optical networks, there is also an amplifier placement problem related to CDS. In fact, an optical network usually consists of passive optical star couplers as nodes which are linked with undirectional fibers. When a signal is traveling too long or splits at some couplers, its power may become too weak and hence it needs an amplifier to increase its power to certain level. The minimization of number of amplifiers under certain network connectivity constraint can also be reduced to a special case of MIN-CDS [87].

1.4 Connected Domatic Number

The *domatic number* of a graph G, $\kappa(G)$, is the maximum number of disjoint dominating sets in G. The *connected domatic number* of a graph G, $\kappa_c(G)$, is the maximum number of disjoint CDS in G. $\kappa(G)$ and $\kappa_c(G)$ are very different. The following theorem indicates this fact.

Theorem 1.4.1. *For any positive integer k, there exists a graph G such that*

$$\kappa(G) - \kappa_c(G) = k.$$

Proof. Let $K = (V,E)$ and $K' = (V',E')$ be two disjoint complete graphs of order $k+1$. Add an edge (v,v') between K and K' for a vertex $v \in V$ and a vertex $v' \in V'$. Denote by G the resulting graph. Then $\kappa(G) = k+1$ and $\kappa_c(G) = 1$. □

Computing $\kappa_c(G)$ is equivalent to the following problem.

MAX#CDS: Given a graph $G = (V,E)$, find the maximum number of disjoint CDS's.

This problem is also intractable.

Theorem 1.4.2 (Cardei et al. [16]). MAX#CDS *is NP-hard.*

Proof. The NP-completeness of the following problem is proved in [16].

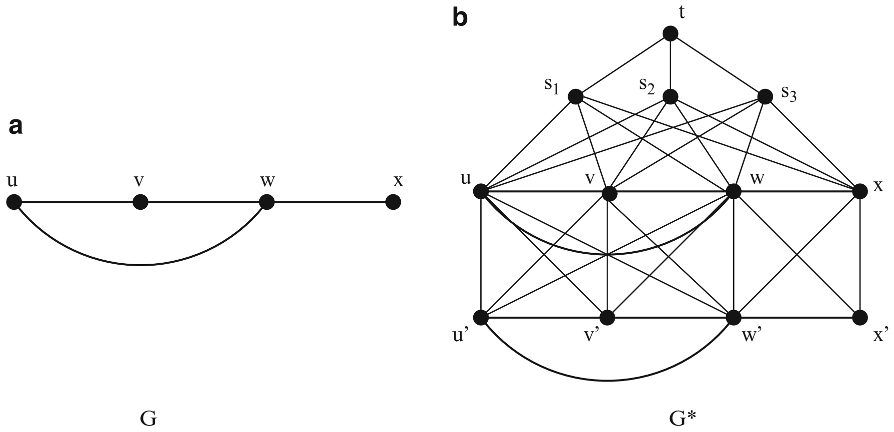

Fig. 1.4 Reduction in the proof of Theorem 1.4.2

> 3DDS: Given a graph $G=(V,E)$, determine whether G contains three disjoint dominating sets.

We now construct a polynomial-time reduction from 3DDS to MAX#CDS. For input graph $G=(V,E)$ of 3DDS, we make a copy of V, $V'=\{v' \mid v\in V\}$. Connect each vertex $u\in V$ to u' and v' for all $(u,v)\in E$. Add four new vertices s_1,s_2,s_3 and t. Connect every s_i to t for $i=1,2,3$ and connect every $u\in V$ to every s_i for $i=1,2,3$. Let $G^*=(V^*,E^*)$ be the graph obtained from the above construction (Fig. 1.4), that is,

$$
\begin{aligned}
V^* &= V\cup V'\cup\{s_1,s_2,s_3,t\}\\
E^* &= E\cup\{(u,v')\mid (u,v)\in E\}\cup\{(u,s_i)\mid u\in V, 1\le i\le 3\}\\
&\quad \cup\{(s_i,t)\mid 1\le i\le 3\}\cup\{(u,u')\mid u\in V\}.
\end{aligned}
$$

We show that G contains three disjoint dominating sets if and only if G^* contains three disjoint CDS's. To do so, we first assume that G contains three disjoint dominating sets D_1,D_2,D_3. Then $D_1\cup\{s_1\}$, $D_2\cup\{s_2\}$ and $D_3\cup\{s_3\}$ are three disjoint CDS's for G^*.

Conversely, assume G^* contains three disjoint CDS's C_1, C_2 and C_3. Define $D_i=C_i\cap V$ for $i=1,2,3$. Then D_1, D_2 and D_3 are disjoint. We claim that each D_i is a dominating set in G. In fact, if there exist a vertex $v\in V$ which cannot be dominated by D_i, then v' cannot be dominated by C_i because every vertex $v'\in V'$ can be dominated by only some vertices in V and v' is dominated by $u\in V$ if and only if v is dominated by u.

Since above construction is done clearly in polynomial-time, MAX#CDS is NP-hard. □

1.5 Lifetime of Sensor Networks

When a very large number of sensors are randomly deployed in target field, the existence of redundant sensors implies the existence of disjoint CDS's. By properly scheduling activation/sleep time of sensors, those CDS's can be organized working in different period as virtual backbone so that the lifetime of the sensor networks is equal to the lifetime of a sensor multiplying the number of disjoint CDS's. Therefore, the maximization of the number of disjoint CDS's has impact in the lifetime maximization of sensor network. This gives an application of MAX#CDS. Actually, from study on the lifetime of sensor networks, more research problems on CDS have been promoted. The following are some of them.

An improvement of lifetime can be seen from the following example as shown in Fig. 1.5. The graph in Fig. 1.5 does not contain two disjoint CDS's. However, if we organize sensors working in the following way, then the lifetime of sensor network can reach 1.5 when every sensor is supposed to has lifetime one.

1. At the 1st 0.5 time period, CDS $\{v_1, v_2\}$ is active.
2. At the 2nd 0.5 time period, CDS $\{v_2, v_3\}$ is active.
3. At the 3rd 0.5 time period, CDS $\{v_3, v_1\}$ is active.

Motivated from this example, we may study the following problem [129].

CDS-SCHEDULING: Given a graph $G = (V, E)$ and a positive vector $b : V \to R^+$, find a sequence of pairs (C_1, t_1), (C_2, t_2), ..., (C_k, t_k) where $1 \le k \le |V|$ to maximize $t_1 + t_2 + \cdots + t_k$ under constraint that $\sum_{i:u \in C_i} t_i \le b(u)$ for every $u \in V$.

Although our scheduling mechanism is able to make the system lifetime longer, the control complexity is increased. Note that different orderings of those CDS $C_1, .., C_p$ give different control complexities. For example, suppose scheduling is in ordering of

$$C_1 = \{v_1, v_2\}, \quad C_2 = \{v_3, v_4\}, \quad C_3 = \{v_2, v_3\}.$$

Then, sensor v_2 should be activated twice. However, in ordering of C_1, C_3, C_2, none of sensors needs to activate twice. This fact raised a research problem on CDS permutation.

An interesting fact discovered in [5] is that putting a sensor alternatively in active and sleep modes in a proper way may double its lifetime since the battery could be

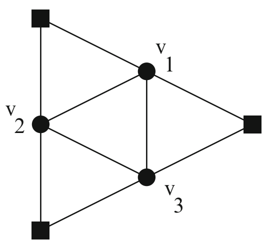

Fig. 1.5 An example

recovered in certain level during sleeping. This fact indicates that CDS permutation contains interesting issues. A proper number of changing between sleep and active modes is good to the lifetime. We may need to find a way to balance the control complexity and the lifetime.

If we allow partial domination, then the lifetime of the system can certainly be increased. A *partial CDS* with percentage p $(0 < p < 1)$ is a vertex subset C which dominates at least pn vertices and induces a connected subgraph. There are three types of problems on partial CDS [19, 20].

In the first type of problems, similar to MAX#CDS and CDS SCHEDULING, we want to maximum the lifetime of the system under constraint that the dominating percentage is always kept at least p.

In the second type of problems, the lifetime of network is given, we want to find a sequence of disjoint (or nondisjoint) partial CDS to maximize the minimum dominating percentage p.

In the third type of problems, the lifetime of network is also given, we want to find a sequence of disjoint (or nondisjoint) partial CDS to maximize the sum of products of dominating percentage p and working time of each partial CDS.

1.6 Theory and Applications

In previous sections, we have seen that two mathematical problems MIN-CDS and MAX#CDS have important applications in network technology. Moreover, motivated from those applications, many new mathematical problems and new issues about CDS have been proposed and studied. Especially, as wireless networks and optical networks are developing rapidly, theory of CDS is growing quickly. The aim of this book is to put together recent results on theory and applications of CDS in order to provide the state of arts in this research area for students, professors, researchers in applied mathematics, operations research and computer science.

In each chapter, we first give a motivation and overview, as well as existing open problems, for subject which is going to be studied in the chapter. Then we present theoretical developments. For convenience of the reader, we try to have this book almost self-contained and each chapter also almost self-contained. Therefore, the definition of notations may be defined repeatedly in different chapters.

Also for convenience of the reader, we restrict usage of brief names. Indeed, except DS (dominating set), CDS (connected dominating set), SCDS (strongly connected dominating set), WCDS (weakly connected dominating set), and names of problems, we rarely use brief names for others.

Although most of contents of this book come from research with motivations from applications in the real world, we have to admit that this is a theory book or mathematical book. Therefore, we do not put any computer experimental result in this book.

We wish that this book can be a useful tool in further developments on theory and applications of CDS to enrich contents of combinatorial optimization and computer and communication networks.