J. Cryptology (1995) 8: 39–64

Journal of
CRYPTOLOGY

Graph Decompositions and Secret Sharing Schemes*

C. Blundo and A. De Santis
Dipartimento di Informatica, Università di Salerno,
84081 Baronissi (SA), Italy

D. R. Stinson
Computer Science and Engineering Department and
Center for Communication and Information Science,
University of Nebraska, Lincoln, NE 68588-0115, U.S.A.

U. Vaccaro
Dipartimento di Informatica, Università di Salerno,
84081 Baronissi (SA), Italy

Communicated by Ernest F. Brickell

Received 20 January 1992 and revised 9 November 1993

Abstract. In this paper we continue a study of secret sharing schemes for·access structures based on graphs. Given a graph G, we require that a subset of participants can compute a secret key if they contain an edge of G; otherwise, they can obtain no information regarding the key. We study the information rate of such schemes, which measures how much information in being distributed as shares compared with the size of the secret key, and the average information rate, which is the ratio between the secret size and the arithmetic mean of the size of the shares. We give both upper and lower bounds on the optimal information rate and average information rate that can be obtained. Upper bounds arise by applying entropy arguments due to Capocelli *et al.* [15]. Lower bounds come from constructions that are based on graph decompositions. Application of these constructions requires solving a particular linear programming problem. We prove some general results concerning the information rate and average information rate for paths, cycles, and trees. Also, we study the 30 (connected) graphs on at most five vertices, obtaining exact values for the optimal information rate in 26 of the 30 cases, and for the optimal average information rate in 28 of the 30 cases.

Key words. Secret sharing scheme, Graph access structure, Linear programming.

* The research of C. Blundo, A. De Santis, and U. Vaccaro was partially supported by the Italian Ministry of University and Research (M.U.R.S.T.) and by the National Council for Research (C.N.R.) under Grant 91.02326.CT12. The research of D. R. Stinson was supported by NSF Grant CCR-9121051.

1. Introduction

A secret sharing scheme is a method of dividing a secret S among a set $\mathscr{P}$ of participants in such a way that: if the participants in $A \subseteq \mathscr{P}$ are qualified to know the secret, then by pooling together their information, they can reconstruct the secret S; but any set $A \subseteq \mathscr{P}$, which is not qualified to know S, has absolutely no information on the secret.

Secret sharing schemes are useful in any important action that requires the concurrence of several designated people to be initiated, such as launching a missile, opening a bank vault, or even opening a safety deposit box. Secret sharing schemes are also used in the management of cryptographic keys and multiparty secure protocols (see [19], for example).

The first secret sharing schemes considered were threshold schemes, introduced by Blakley [4] and Shamir [31]. A (k, n) threshold scheme allows a secret to be shared among n participants in such a way that any k of them can recover the secret, but any $k - 1$, or fewer, have absolutely no information on the secret (see [34] for a comprehensive bibliography on (k, n) threshold schemes).

Ito *et al.* [21] (see also [22]) described the general method of secret sharing. An access structure is a specification of all the subsets of participants who can recover the secret and it is said to be monotone if any set which contains a subset that can recover the secret can itself recover the secret. Ito *et al.* gave a methodology to realize secret sharing schemes for arbitrary monotone access structures. Subsequently, Benaloh and Leichter [2] gave a simpler and more efficient way to realize secret sharing schemes for any given monotone access structure. Other general techniques for handling arbitrary access structures are given by Simmons *et al.* [37] and Martin [27].

An important issue in the implementation of secret sharing schemes is the size of shares since the security of a system degrades as the amount of information that must be kept secret increases. If it is required that a nonqualified set of participants should have no information on the secret, then the size of the shares cannot be less than the size of the secret [25]. In [2] it is proved that an access structure (namely, the path of length three) exists for which any secret sharing scheme must give to some participant a share which is from a domain larger than that of the secret. This was improved by Brickell and Stinson [12], who showed that, for the same access structure, the number of elements in the domain of the shares must be at least $2|S| - 1$ if the cardinality of the domain of the secret is $|S|$. Finally, Capocelli *et al.* [15] proved, for the same access structure, that the number of elements in the domain of the shares must be at least $|S|^{1.5}$, and they showed that the bound is tight.

Ideal secret sharing schemes, that is, schemes where the shares are taken from the same domain as that of the secret, were characterized by Brickell and Davenport [11] in terms of matroids. The uniqueness of the associated matroid is established by Martin in [26]. Brickell constructed some classes of ideal schemes in [10], and an interesting nonexistence result was proved by Seymour [30]. Beimel and Chor [1] investigated the access structures for which an ideal scheme can be constructed for every possible size of the set of secrets. Finally, equivalence of ideal threshold schemes and orthogonal arrays is shown independently in [17] and [23].

We also mention some "extended capabilities" of secret sharing schemes that have been studied.

- The idea of protecting against cheating by one or more participants is addressed in [28], [41], [29], [32], [13], and [16].
- Prepositioned schemes are studied in [34].
- Threshold schemes that permit disenrollment of participants are investigated in [5].
- Secret sharing schemes in which the dealer has the feature of being able (after a preprocessing stage) to activate a particular access structure out of a given set and/or to allow the participants to reconstruct different secrets (in different time instants) by sending to all participants the same broadcast message have been analyzed in [7].
- Schemes for sharing several nonindependent secrets simultaneously have been analyzed in [9].
- Schemes where different secrets are associated with different subsets of participants are considered in [24].
- The question of how to set up a secret sharing scheme in the absence of a trusted party is solved in [20].

Different measures are possible for the amount of secret information that must be given to participants. When we are interested in the maximum size of the shares, we can use the information rate [14], which is the ratio between the secret size and the maximum size of the shares. When we are interested in the total size of all the shares (and not just the maximum one), it is preferable to use as a measure the average information rate [6], [26], [27], which is the ratio between the secret size and the arithmetic mean of the size of all the shares.

In this paper we study secret sharing schemes in the case where the access structure consists of the closure of a (connected) graph. We consider all 30 connected graphs on at most five vertices, and determine the exact value of the optimal information rate in all but four cases and optimal average information rate in all but two cases. For these remaining cases, we give quite good upper and lower bounds. For two infinite classes of graphs—cycles of even length (≥ 6) and paths of arbitrary length (≥ 3)—we prove that the value of optimal information rate is $\frac{2}{3}$. For paths and for cycles of even length (≥ 4), we show how to realize secret sharing schemes with optimal *average* information rate. For any tree, we present a secret sharing scheme with information rate at least $\frac{1}{2}$, and a scheme with average information rate at least $\frac{2}{3}$, both of which improve previous results.

The main tool for proving upper bounds on the information rate is the entropy approach of Capocelli *et al.* [15]. Lower bounds are obtained by construction methods based on graph decompositions. The main idea of our new method is to use different constructions for different bits of the secret and different subsets of participants. Application of these constructions requires solving a suitable linear programming problem.

The paper is organized as follows. In Section 2 we give the formal definition of secret sharing schemes and recall some basic results. In Section 3 we give our general graph decomposition construction. In Section 4 we discuss the methods for bounding information rates and prove the results mentioned above concerning cycles, paths, and trees. In Section 5 we discuss the methods for bounding average information rates and prove the results concerning paths, cycles, ad trees. Then, in Section 6, we investigate the information rate and the average information rate for the connected graphs on at most five vertices.

2. Secret Sharing Schemes

We recall some definitions and notation from [14]. Suppose that $\mathscr{P}$ is the set of participants. Denote by Γ the set of subsets of participants which we desire to be able to determine the key; hence $\Gamma \subseteq 2^{\mathscr{P}}$. Γ is called the *access structure* of the secret sharing scheme. It seems reasonable to require that Γ be *monotone*, i.e., if $B \in \Gamma$ and $B \subseteq C \subseteq \mathscr{P}$, then $C \in \Gamma$.

For any $\Gamma_0 \subseteq 2^{\mathscr{P}}$, define the *closure* of Γ_0 to be

$$cl(\Gamma_0) = \{C\colon \exists B \subseteq \Gamma_0, B \subseteq C \subseteq \mathscr{P}\}.$$

Note that the closure of any set of subsets is monotone.

Let $\mathscr{K}$ be a set of q elements called *keys*. For every participant $P \in \mathscr{P}$, let $\mathscr{S}_P$ be a set of s_P elements. Elements of the set $\mathscr{S}_P$ are called *shares*. Suppose a *dealer* D wants to a share the secret key $K \in \mathscr{K}$ among the participants in $\mathscr{P}$ (we assume that $D \notin \mathscr{P}$). He does this by giving each participant $P \in \mathscr{P}$ a share from $\mathscr{S}_P$. We say that the scheme is a *perfect scheme* (with respect to access structure Γ) if the following two properties are satisfied:

1. If a subset $\mathscr{B}$ of participants pool their shares, where $\mathscr{B} \in \Gamma$, then they can determine the value of K.
2. If a subset $\mathscr{B}$ of participants pool their shares, where $\mathscr{B} \neq \Gamma$, then they can determine nothing about the value of K (in an information-theoretic sense), even with infinite computational resources.

Remark. In [14] Brickell and Stinson required every participant to have shares taken from the same set, say $\mathscr{S}$. This can easily be done, if desired, by taking a set $\mathscr{S}$ of cardinality $\max\{s_P\colon P \in \mathscr{P}\}$ and defining injections $\varphi_P\colon S_P \to \mathscr{S}$ for every $P \in \mathscr{P}$.

Throughout this paper we confine our attention to perfect schemes, so the term "secret sharing scheme" can be taken to mean "perfect secret sharing scheme."

We depict a secret sharing scheme as a matrix M. This matrix is not secret, but is known by all the participants. There are $|\mathscr{P}| + 1$ columns in M. The first column of M is indexed by D, and the remaining columns are indexed by the members of $\mathscr{P}$. In any row of M, we place the key K in column D, and a possible list of shares corresponding to K in the remaining columns. When D wants to distribute shares corresponding to a key K, he chooses uniformly at random a row r of M having K in column D, and distributes the shares in that row to the participants (i.e., $M(r, P)$ is given to participant P, for all $P \in \mathscr{P}$).

With this matrix representation, we can present combinatorial conditions on the matrix M that ensure that the two properties above are satisfied. These conditions are equivalent to conditions presented in [14].

1. If $\mathscr{B} \in \Gamma$ and $M(r, P) = M(r', P)$ for all $P \in \mathscr{B}$, then $M(r, D) = M(r', D)$.
2. If $\mathscr{B} \notin \Gamma$, then for every possible assignment f of shares to the participants in $\mathscr{B}$, say $f = (f_P\colon P \in \mathscr{B})$ (where $f_P \in S_P$ for all $P \in \mathscr{B}$), a nonnegative integer $\lambda(f, \mathscr{B})$ exists such that
$$|\{r\colon M(r, P) = f_P \forall P \in \mathscr{B}, M(r, D) = K\}| = \lambda(f, \mathscr{B}),$$
independent of the value of K.

An important issue in the implementation of secret sharing schemes is the size of shares, since the security of a system degrades as the amount of the information that must be kept secret increases. Define $s = \max\{s_P\colon P \in \mathscr{P}\}$. The *information rate* [14] of the secret sharing scheme is defined to be

$$\rho = \frac{\log q}{\log s}.$$

(We use the term "information rate" because the concept is similar to that of the information rate of an error-correcting code.) It is not difficult to see that $q \leq s$ in a perfect scheme, so the information rate satisfies $\rho \leq 1$. If a secret sharing scheme is to be practical, we do not want to have to distribute too much secret information as shares. Consequently, we want to make the information rate as close to 1 as possible. A perfect secret sharing scheme with information rate $\rho = 1$ is called *ideal.*

In many cases it is preferable to limit the sum of the size of shares over all participants. To analyze such cases we use the *average information rate* [6], [27] defined as

$$\tilde{\rho} = \frac{|\mathscr{P}| \log q}{\sum_{P \in \mathscr{P}} \log s_P}.$$

In a perfect secret sharing scheme, $q \leq s_P$ for all $P \in \mathscr{P}$, and thus $\tilde{\rho} \leq 1$. Also, $\rho = 1$ if and only if $\tilde{\rho} = 1$. It is clear that $\tilde{\rho} \geq \rho$ for any scheme. Equality holds if and only if $s_P = s_{P'}$ for all $P, P' \in \mathscr{P}$.

2.1. *Basic Results*

We present some basic terminology from graph theory. Graphs do not have loops or multiple edges; a graph with multiple edges is termed a *multigraph.* If G is a graph, we denote the vertex set of G by $V(G)$ and the edge set by $E(G)$. We consider undirected graphs only. In an undirected graph the pair of vertices representing any edge is unordered. Thus, the pairs (u, v) and (v, u) represent the same edge. To avoid overburdening the notation we often describe a graph G by the list of all edges $E(G)$ and each edge $(u, v) \in E(G)$ is represented by uv. G is *connected* if any two vertices are joined by a path. The *complete graph* K_n is the graph on n vertices in which any two vertices are joined by an edge. The *complete multipartite graph* $K_{n_1,n_2,\ldots,n_t}$ is a graph on $\sum_{i=1}^{t} n_i$ vertices, in which the vertex set is partitioned into subsets of size

n_i $(1 \leq i \leq t)$ called *parts*, such that vw is an edge if and only if v and w are in different parts. An alternative way to characterize a complete multipartite graph is to say that the complementary graph is a vertex-disjoint union of cliques. Note that the complete graph K_n can be thought of as a complete multipartite graph with n parts of size 1.

A *stable set* or *independent set* of G is a subset of vertices $A \subseteq V(G)$ such that no two vertices in A are joined by an edge in $E(G)$. The *stability number* or *independence number* $\alpha(G)$ is defined to be the maximum cardinality of a stable set of G. A *vertex cover* of G is a subset of vertices $A \subseteq V(G)$ such that every edge in $E(G)$ is incident with at least one vertex in A. The *vertex covering number* $\beta(G)$ is defined to be the minimum cardinality of a vertex cover of G.

The *girth* of a graph G is defined to be the length of the smallest cycle in G. If G is acyclic, the girth is defined to be ∞. A *regular* graph is a graph where each vertex has degree d, for a fixed d.

We use the notation $PS(G, \rho, q)$ to denote a perfect secret sharing scheme with access structure $cl(E(G))$ and information rate ρ for a set of q keys. Analogously, a perfect secret sharing scheme with access structure $cl(E(G))$ and average information rate $\tilde{\rho}$ for a set of q keys is denoted by $\widetilde{PS}(G, \tilde{\rho}, q)$. Throughout this paper we restrict our attention to connected graphs. If a graph is not connected, it suffices to find schemes for each of its connected components. The following theorem was proved for information rate in [14]; the proof for average information rate is similar.

Theorem 2.1. *Suppose G is a graph having as its connected components G_i, $1 \leq i \leq t$. Suppose that there is a $PS(G_i, \rho, q)$, $1 \leq i \leq t$. Then there is a $PS(G, \rho, q)$. Similarly, if there is a $\widetilde{PS}(G_i, \tilde{\rho}, q)$ for $1 \leq i \leq t$, then there is a $\widetilde{PS}(G, \tilde{\rho}, q)$.*

Ideal schemes for connected graphs were characterized by Brickell and Davenport [11].

Theorem 2.2. *Suppose G is a connected graph. Then a $PS(G, 1, q)$ (and, equivalently, a $\widetilde{PS}(G, 1, q)$) for some q exists if and only if G is a complete multipartite graph.*

The following result from [14] specifies some values of q for which ideal schemes can be constructed.

Corollary 2.3. *Suppose $q \geq t$ is a prime power. Then there is a $PS(K_{n_1,n_2,\ldots,n_t}, 1, q)$.*

Proof. Let $V_1, \ldots, V_t$ be the parts of the graph $K_{n_1,n_2,\ldots,n_t}$. Let $x_1, \ldots, x_t$ be distinct elements of $GF(q)$. We construct a matrix M having q^2 rows and $1 + \sum_{i=1}^{t} n_i$ columns. The rows of M are indexed by $GF(q) \times GF(q)$, and the columns are indexed by $\{D\} \cup V_1 \cup \cdots \cup V_t$. Define the entries of M by the following rule:

$$M((a, b), D) = a,$$

$$M((a, b), v) = ax_i + b,$$

where $a, b \in GF(q)$ and $v \in V_i$. □

Remark. If we start with a complete graph K_t, then the matrix M, as constructed above, is a structure from combinatorial design theory known as an orthogonal array $OA(t+1, q)$.

We recall two basic results from [14]. The first result indicates that the information rate is an appropriate measure of the efficiency of a secret sharing scheme. It states that the existence of one scheme with a specified rate immediately implies the existence of a scheme with the same rate handling as many keys as desired. The result was proved for information rate in [14] and for average information rate in Corollary 2.4 of [27].

Theorem 2.4. *Suppose there is a $PS(G, \rho, q)$. Then, for any positive integer n, there is a $PS(G, \rho, q^n)$. Similarly, if there is a $\widetilde{PS}(G, \tilde{\rho}, q)$, then, for any positive integer n, there is a $\widetilde{PS}(G, \tilde{\rho}, q^n)$.*

If G is a graph, then G_1 is said to be a *subgraph* of G if $V(G_1) \subseteq V(G)$ and $E(G_1) \subseteq E(G)$. If $V_1 \subseteq V(G)$, then we define the graph $G[V_1]$ to have vertex set V_1 and edge set $\{uv \in E(G): u, v \in V_1\}$. We say that $G[V_1]$ is an *induced subgraph* of G. The following theorem is obvious.

Theorem 2.5. *Suppose G is a graph and G_1 is an induced subgraph of G. If there is a $PS(G, \rho, q)$, then a $PS(G_1, \rho, q)$ exists.*

Observe that the statement of the above theorem is *not* true for average information rate.

Let $\mathscr{A}$ be an access structure such that there are four participants, A, B, C, D, such that $\{A, B\}, \{B, C\}, \{C, D\} \in \mathscr{A}$ but $\{A, C\}, \{A, D\}, \{B\} \notin \mathscr{A}$. Capocelli *et al.* [15] proved that for any secret sharing scheme for $\mathscr{A}$ the sum of the entropies of the two random variables defined by the shares given to B and C cannot be less than three times the entropy of the secret. By taking all probability distributions to be uniform, the upper bound can be stated as follows:

Theorem 2.6. *Let $\mathscr{A}$ be an access structure. If there are four participants A, B, C, and D such that*

$$\{A, B\}, \{B, C\}, \{C, D\} \in \mathscr{A}$$

and

$$\{A, C\}, \{A, D\}, \{B\} \notin \mathscr{A},$$

then any secret sharing scheme for $\mathscr{A}$ satisfies

$$\log s_B + \log s_C \geq 3 \log q.$$

Examples of access structures that satisfy the hypotheses of the above theorem are the closure of P_3 (the path of length three), which is the graph having edge set

$$\{AB, BC, CD\};$$

and the closure of H, the graph having edge set

$$\{AB, BC, CD, BD\}.$$

Theorem 2.6 is the main tool we use for proving upper bounds on information rate and average information rate for paths, cycles, and general graphs.

3. Graph Decomposition Constructions

Suppose G is a graph and $G_1, \ldots, G_n$ are subgraphs of G, such that each edge of G occurs in at least one of the G_i's. Suppose also that each G_i is a complete multipartite graph. Then we say that $\Pi = \{G_1, \ldots, G_t\}$ is a *complete multipartite covering* (or CMC) of G. The following construction utilizing CMCs is a special case of Theorem 3.5 of [14]. The extension to average information rate is straightforward.

Theorem 3.1 (CMC Construction). *Suppose G is a graph and $\Pi = \{G_1, \ldots, G_n\}$ is a complete multipartite covering of G. For $1 \le i \le n$, denote by t_i the number of parts in G_i, and let*

$$t = \max\{t_i \colon 1 \le i \le n\}.$$

For every vertex v, define

$$R_v = |\{i \colon v \in G_i\}|$$

and

$$R = \max\{R_v \colon v \in V(G)\}.$$

Then, for $\rho = 1/R$ and

$$\tilde{\rho} = \frac{|V(G)|}{\sum_{v \in V(G)} R_v},$$

there is a $PS(G, \rho, q)$ and a $\widetilde{PS}(G, \tilde{\rho}, q)$ for any prime power $q \ge t$.

Proof. Let $q \ge t$, and, for $1 \le i \le n$, let M_i be the matrix representing $PS(G_i, 1, q)$, which exists by Corollary 2.3. Let $\mathscr{K}$ denote a set of q keys and let $\mathscr{S}$ denote a set of q shares (which we can assume are the same for all the schemes). Then define a matrix M as follows: for every key K, and for every n-tuple of rows $(r_i \colon 1 \le i \le n)$ such that r_i is a row of M_i $(1 \le i \le n)$ and $M_i(r_i, D) = K$ $(1 \le i \le n)$, define a row $(r_i \colon 1 \le i \le n)$ of M by the rule

$$M((r_1, r_2, \ldots, r_n), v) = (M_i(r_i, v) \colon v \in V(G_i)),$$

$$M((r_1, r_2, \ldots, r_n), D) = K. \qquad \square$$

Remark. It is not actually necessary to construct the matrix M of the above proof. When D wishes to share a secret K, it suffices for him to choose, for each i, $1 \le i \le n$, a random row r_i of M_i such that $M_i(r_i, D) = K$. Then, for $1 \le i \le n$ and for each $v \in G_i$, D gives $M_i(r_i, v)$ to participant v. Hence, each participant v gets a share corresponding to each G_i such that $v \in V(G_i)$.

The main result of this section is a generalization of the CMC construction. The idea is to use several decompositions, rather than just one.

Theorem 3.2 (Multiple CMC Construction). *Suppose G is a graph and for $1 \leq j \leq l$, suppose $\Pi_j = \{G_{j1}, \ldots, G_{jn_j}\}$ is a complete multipartite covering of G. Denote by t_{ji} the number of parts in G_{ji} $(1 \leq j \leq l, 1 \leq i \leq n_j)$ and define*

$$t = \max\{t_{ji}\colon 1 \leq j \leq l, 1 \leq i \leq n_j\}.$$

For every vertex v and for $1 \leq j \leq l$, define

$$R_{jv} = |\{i\colon v \in G_{ji}\}|,$$

$$R_v = \sum_{j=1}^{l} R_{jv},$$

and

$$R = \max\{R_v\colon v \in V(G)\}.$$

Then, for $\rho = l/R$, there is a $PS(G, \rho, q^l)$ for any prime power $q \geq t$.

Proof. Carry out the construction of Theorem 3.1 independently for each of l keys. The details are left to the reader. □

Remark. In the case $l = 1$, we recover the original CMC construction. Also, we observe that we cannot improve the lower bound on $\tilde{\rho}$ by taking $l > 1$.

Example 3.1. Recall that P_3, the path of length 3, has edges AB, BC, CD. Using one CMC, the best information rate that can be obtained for the access structure $cl(E(P_3))$ is $\frac{1}{2}$. However, using two CMCs, we can get $\rho = \frac{2}{3}$ (a result first obtained by Capocelli *et al.* [15]). The two CMCs are

$$\{\{AB\}, \{BC, CD\}\}$$

and

$$\{\{AB, BC\}, \{CD\}\}.$$

Then $R_A = R_D = 2$ and $R_B = R_C = 3$. Hence $R = 3$ and $\rho = \frac{2}{3}$. A $PS(P_3, \frac{2}{3}, 4)$ can be constructed. Note that if we implement the scheme, we get precisely the scheme presented in [15]. Also, either of these two CMCs yields a scheme with average information rate $\tilde{\rho} = \frac{4}{5}$.

Example 3.2. The graph H has edges AB, BC, CD, BD. From the two CMCs

$$\{\{AB\}, \{BC, BD, CD\}\}$$

and

$$\{\{AB, BC, BD\}, \{CD\}\},$$

we can construct a $PS(H, \frac{2}{3}, 9)$. Using Corollary 2.3, this scheme could be implemented as follows. Take $\mathscr{K} = GF(3) \times GF(3)$. The dealer will choose four random elements (independently) from $GF(3)$, say b_{11}, b_{12}, b_{21}, and b_{22}. Given a key (K_1, K_2), the dealer distributes shares as follows: participant A receives $(b_{11} + K_1, b_{21} + K_2)$; participant B receives (b_{11}, b_{12}, b_{21}); participant C receives $(b_{12} + K_1, b_{21} + K_2, b_{22})$; and participant D receives $(b_{12} + 2K_1, b_{21} + K_2, b_{22} + K_2)$. Hence, $\mathscr{S}_A = GF(3) \times GF(3)$ and $\mathscr{S}_B = \mathscr{S}_C = \mathscr{S}_D = GF(3) \times GF(3) \times GF(3)$. Finally, observe that the first CMC yields a scheme with average information rate $\tilde{\rho} = \frac{4}{5}$, while the second CMC would give $\tilde{\rho} = \frac{2}{3}$.

4. Optimal Information Rates

For a graph G, define

$$\rho^*(G) = \sup\{\rho\colon \exists PS(G, \rho, q) \text{ for some integer } q\}.$$

Note that the definition does *not* require that a $PS(G, \rho^*(G), q)$ exists for any integer q. However, in all cases where we know the value of $\rho^*(G)$, we can actually construct a scheme having that information rate.

Of course, $\rho^*(G) \leq 1$ for all graphs, and $\rho^*(G) = 1$ if G is a complete multipartite graph. The first nontrivial upper bounds on ρ^* were proved by Capocelli *et al.* [15]. Using Theorem 2.6, they proved that $\rho^*(P_3) = \frac{2}{3}$ and $\rho^*(H) \leq \frac{2}{3}$. In view of the construction given in Example 3.2, we have the following theorem.

Theorem 4.1. *Let P_3 be the graph having edges AB, BC, CD and let H be the graph having edges AB, BC, CD, BD. Then $\rho^*(P_3) = \frac{2}{3}$ and $\rho^*(H) = \frac{2}{3}$.*

We can also prove the following general upper bound.

Theorem 4.2. *Suppose G is a connected graph that is not a complete multipartite graph. Then $\rho^*(G) \leq \frac{2}{3}$.*

Proof. We will prove that any connected graph that is not a complete multipartite graph must contain four vertices w, x, y, z such that the induced subgraph $G[w, x, y, z]$ is isomorphic to either P_3 or H (from Examples 3.1 and 3.2). The desired result then follows from Theorem 2.5.

Let G^C denote the complement of G. Since G is not a complete multipartite graph, three vertices x, y, z must exist such that $xy, yz \in E(G^C)$ and $xz \in E(G)$. Define

$$d = \min\{d_G(y, x), d_G(y, z)\},$$

where d_G denotes the length of a shortest path (in G) between two vertices. Then $d \geq 2$. Without loss of generality, we can assume that $d = d_G(y, x)$ by symmetry. Let $y_0, y_1, \ldots, y_{d-1}, x$ be a path in G, where $y_0 = y$. We have that $y_{d-2}z, y_{d-2}x \in E(G^C)$ and $y_{d-2}y_{d-1}, y_{d-1}x, xz \in E(G)$. It follows that $G[y_{d-2}, y_{d-1}, x, z]$ is isomorphic to either P_3 or H, as desired. □

Hence, $\rho^*(G) = 1$ if and only if G is a complete multipartite graph; and $\rho^*(G) \leq \frac{2}{3}$ if and only if G is not a complete multipartite graph. Thus, there is a "gap" in the possible values for $\rho^*(G)$.

4.1. *A Linear Programming Problem*

We are also interested in the best possible information rate that can be obtained by applying the multiple CMC construction, Theorem 3.2. We define the quantity $\rho_C^*(G)$ which will denote this optimal rate for graph G. In view of the nature of the construction, we can construct a $PS(G, \rho_C^*(G), q^l)$ for some fixed integer l and for all sufficiently large prime powers q. Of course, $\rho_C^*(G) \leq \rho^*(G)$.

Our main observation is that $\rho_C^*(G)$ can be computed by solving a suitable linear programming problem. We describe how this can be done in the remainder of the section.

Suppose G is a graph. We define a partial order on the CMCs of G as follows. Suppose $\Pi_j = \{G_{j1}, \ldots, G_{jn_j}\}, j = 1, 2$, are two CMCs of G. For every vertex v and for $j = 1, 2$, define $R_{jv} = |\{i\colon v \in G_{ji}\}|$. Then we define $\Pi_1 \preceq \Pi_2$ if $R_{1v} \leq R_{2v}$ for all $v \in V(G)$. Define a CMC, Π, to be *minimal* if there is no Π' such that $\Pi' \preceq \Pi$ and $\Pi' \neq \Pi$.

Now, suppose $\Pi_j = \{G_{j1}, \ldots, G_{jn_j}\}$, $1 \leq j \leq L$, comprise a complete enumeration of the minimal CMCs of G. For every vertex v and for $1 \leq j \leq L$, define

$$R_{jv} = |\{i\colon v \in G_{ji}\}|.$$

Consider the following optimization problem $\mathcal{O}(G)$:

$$\text{Minimize} \quad R_0 = \max\left\{\sum_{j=1}^{L} a_j R_{jv}\colon v \in V(G)\right\}$$

$$\text{subject to} \quad a_j \geq 0, \qquad 1 \leq j \leq L,$$

$$\sum_{j=1}^{L} a_j = 1.$$

Theorem 4.3. *Let R^* be the optimal solution to $\mathcal{O}(G)$. Then $\rho_C^*(G) = 1/R^*$.*

Proof. Suppose

$$R^* = \max\left\{\sum_{j=1}^{L} a_j R_{jv}\colon v \in V(G)\right\},$$

where $a_j\,(1 \leq j \leq L)$ satisfy the constraints of $\mathcal{O}(G)$. It is clear that the a_j are rational, so denote $a_j = b_j/c_j$, where $b_j, c_j \in \mathbf{Z}$, $1 \leq j \leq L$. Let C denote the least common multiple of $c_1, \ldots, c_L$. Then take Ca_j copies of Π_j for $1 \leq j \leq L$, and apply the multiple CMC construction. We get a scheme with information rate $1/R^*$; hence $\rho_C^*(G) \geq 1/R^*$.

Conversely, suppose we start with an application of the multiple CMC construction that yields the information rate $\rho_C^*(G)$. We can assume without loss of generality that only minimal CMCs are used. Suppose there are b_j copies of Π_j, $1 \leq j \leq L$. Let $B = \sum_{j=1}^{L} b_j$, and define $a_j = b_j/B$, $1 \leq j \leq L$. Then $(a_1, \ldots, a_L)$ satisfy the constraints of $\mathcal{O}(G)$, and yield $R_0 = 1/\rho_C^*(G)$. Hence, $\rho_C^*(G) \leq 1/R^*$, and we are done. □

The difficulty with problem $\mathcal{O}(G)$ is that the objective function is the maximum of several linear functions. However, we can easily obtain an "equivalent" linear programming problem $\mathcal{O}'(G)$:

$$\begin{aligned} &\text{Minimize } T \\ &\text{subject to} \quad a_j \geq 0, \qquad 1 \leq j \leq L, \\ &\sum_{j=1}^{L} a_j = 1, \\ &T \geq \sum_{j=1}^{L} a_j R_{jv}, \qquad v \in V(G). \end{aligned}$$

It is easy to see that $\mathcal{O}(G)$ and $\mathcal{O}'(G)$ have the same optimal solution. Hence, we obtain the following result.

Theorem 4.4. *Let T^* be the optimal solution to $\mathcal{O}'(G)$. Then $\rho_C^*(G) = 1/T^*$.*

4.2. *Information Rate for Paths and Cycles*

We next establish some general results when G is a path or a cycle. P_n denotes a path of length n, that is, the graph with edges $X_1X_2, \ldots, X_nX_{n+1}$; and C_n denotes a cycle of length n, that is, the graph with edges $X_1X_2, \ldots, X_{n-1}X_n, X_nX_1$.

Theorem 4.5. *If $n \geq 3$, then $\rho^*(P_n) = \frac{2}{3}$.*

Proof. If $n \geq 3$, $\rho^*(P_n) \leq \frac{2}{3}$ by Theorem 4.2. First, suppose $n + 1$ is odd. Then $\rho^*(P_n) \geq \frac{2}{3}$ by using the following two CMCs:

$$\Pi_1 = \{\{X_1X_2, X_2X_3\}, \{X_3X_4, X_4X_5\}, \ldots, \{X_{n-1}X_n, X_nX_{n+1}\}\}$$

and

$$\Pi_2 = \{\{X_1X_2\}, \{X_2X_3, X_3X_4\}, \ldots, \{X_{n-2}X_{n-1}, X_{n-1}X_n\}, \{X_nX_{n+1}\}\}.$$

If $n + 1$ is even, then $\rho^*(P_n) \geq \frac{2}{3}$ by using

$$\Pi_3 = \{\{X_1X_2, X_2X_3\}, \{X_3X_4, X_4X_5\}, \ldots, \{X_{n-2}X_{n-1}, X_{n-1}X_n\}, \{X_nX_{n+1}\}\}$$

and

$$\Pi_4 = \{\{X_1X_2\}, \{X_2X_3, X_3X_4\}, \ldots, \{X_{n-1}X_n, X_nX_{n+1}\}\}.$$ □

Theorem 4.6. *If $n \geq 3$, then $\rho^*(C_{2n}) = \frac{2}{3}$.*

Proof. If $n \geq 3$, $\rho^*(C_{2n}) \leq \frac{2}{3}$ by Theorem 4.2. $\rho^*(C_{2n}) \geq \frac{2}{3}$ by using the following two CMCs:

$$\{\{X_1X_2, X_2X_3\}, \{X_3X_4, X_4X_5\}, \ldots, \{X_{2n-1}X_{2n}, X_{2n}X_1\}\}$$

and

$$\{\{X_2X_3, X_3X_4\}, \{X_4X_5, X_5X_6\}, \ldots, \{X_{2n}X_1, X_1X_2\}\}.$$ □

Theorem 4.7. *If $n \geq 2$, then $\rho_C^*(C_{2n+1}) = (2n+1)/(3n+2)$.*

Proof. Here, we appeal to Theorem 4.4. First, we enumerate the minimal CMCs for C_{2n+1}. Take the vertices to be $X_1, X_2, \ldots, X_{2n+1}$, and perform all arithmetic operations on indices mod$(2n+1)$. Define

$$\Pi_0 = \{\{X_1X_2, X_2X_3\}, \ldots, \{X_{2n-1}X_{2n}, X_{2n}X_{2n+1}\}, \{X_{2n+1}X_1\}\}.$$

For $0 \leq j \leq 2n$, define Π_j by "adding" j to indices of Π_0 and reducing mod$(2n+1)$ to the interval $1, 2, \ldots, 2n+1$. Then Π_j, $0 \leq j \leq 2n$, are the $2n+1$ minimal CMCs. We get a $(2n+1) \times (2n+1)$ matrix of values R_{jX_v}, where $R_{jX_v} = 1$ if and only if $v - j$ mod$(2n+1)$ is odd (where $v - j$ is reduced mod$(2n+1)$ to the interval $1, 2, \ldots, 2n+1$). For example, in the case $2n+1 = 5$, we get the matrix

$$\begin{bmatrix} 2 & 1 & 2 & 1 & 2 \\ 2 & 2 & 1 & 2 & 1 \\ 1 & 2 & 2 & 1 & 2 \\ 2 & 1 & 2 & 2 & 1 \\ 1 & 2 & 1 & 2 & 2 \end{bmatrix}.$$

The optimal solution to $\mathcal{O}'(C_{2n+1})$ is obtained when $a_1 = \cdots = a_{2n+1} = 1/(2n+1)$; then $T = (3n+2)/(2n+1)$ and $\rho_C^*(C_{2n+1}) = (2n+1)/(3n+2)$. In applying the multiple CMC construction, we take one copy of each Π_j. □

4.3. *Information Rate for Trees*

Brickell and Stinson proved in Theorem 3.8 of [14] that, for any graph G of maximum degree d, a secret sharing scheme can be realized with information rate

$$\rho \geq \frac{1}{\lceil d/2 \rceil + 1}.$$

This was proved using the CMC construction, by decomposing G into complete bipartite graphs $K_{1,m}$ (such a decomposition is called a *star decomposition*, since $K_{1,m}$ if often called a *star*). In the case where G is regular and has girth at least 5, this result is the best that can be obtained using star decompositions [14, Theorem 3.9]. However, we can improve the lower bound whenever G is acyclic. We use star decompositions to obtain information rate equal to $\frac{1}{2}$ in this case.

We now describe the algorithm used to obtain this decomposition. First, we need some definitions. Let G be a connected graph and $v \in V(G)$. $Inc(v)$ denotes the set of edges incident with v:

$$Inc(v) = \{uv\colon uv \in E(G)\}.$$

By $Adj(v)$ we denote the set of vertices adjacent to v:

$$Adj(v) = \{u \in V(G)\colon uv \in E(G)\}.$$

For any vertex $v \in V(G)$, let G_v denote the graph with vertex set $V(G_v) = \{v\} \cup Adj(v)$ and edge set $E(G_v) = Inc(v)$.

The algorithm **Covering** constructs a star decomposition of G by calling the recursive algorithm **Cover**. The algorithms are as follows:

```
Covering($G$)
    Let $X \in V(G)$
    $\Pi \leftarrow \emptyset$
    Cover($X$)
    Output the star decomposition $\Pi$

Cover($X$)
    $\Pi \leftarrow \Pi \cup \{G_X\}$
    $B \leftarrow \{Y \in Adj(X)\colon |Inc(Y)| = 1\}$
    $E(G) \leftarrow E(G) - Inc(X)$
    $V(G) \leftarrow V(G) - (B \cup \{X\})$
    For all $X' \in Adj(X) - B$ do Cover($X'$)
```

The algorithm **Covering** always finds a complete multipartite covering of G. It is easy to see that if G is acyclic, then each of its vertices belongs to at most two different connected subgraphs of the covering, as stated in next lemma.

Lemma 4.8. *Let Π be a complete multipartite covering of a tree G obtained by applying* **Covering** *to G. Then each vertex $X \in V(G)$ belongs to at most two different subgraphs $G', G'' \in \Pi$.*

The following result is immediate from Lemma 4.8 and Theorem 3.1.

Corollary 4.9. *For any tree, a secret sharing scheme exists with information rate $\rho \geq \frac{1}{2}$.*

There is only one case in which G is connected and **Covering** gives a secret sharing scheme with information rate greater than $\frac{1}{2}$. This case arises when G is itself a star graph and X is chosen to be the vertex of maximum degree in G.

5. Optimal Average Information Rates

Recall that we use the notation $\widetilde{PS}(G, \tilde{\rho}, q)$ to denote a perfect secret sharing scheme with access structure $cl(E(G))$ and average information rate $\tilde{\rho}$ for a set of q keys. For a graph G, define

$$\tilde{\rho}^*(G) = \sup\{\tilde{\rho}\colon \exists \widetilde{PS}(G, \tilde{\rho}, q) \text{ for some integer } q\}.$$

As before, the definition does *not* require that a $\widetilde{PS}(G, \tilde{\rho}^*(G), q)$ exists for any integer q.

The following lemma is the analogue of Theorem 4.2 for the average information rate. It is a generalization of Lemma 4.3.5 of [26].

Lemma 5.1. *Let G be a connected graph with n vertices. If G is a complete multipartite graph, then $\tilde{\rho}^*(G) = 1$; otherwise $\tilde{\rho}^*(G) \leq n/(n + 1)$.*

Proof. Assume G is a complete multipartite graph. By Theorem 2.2 an ideal scheme exists; this scheme has an average information rate equal to 1. If G is not a complete multipartite graph then, from Theorem 4.2 and Theorem 2.6, two vertices in $V(G)$, X and Y with $XY \in E(G)$, exist such that $\log s_X + \log s_Y \geq 3 \log q$. Thus

$$\sum_{X \in V(G)} \log s_X \geq (n+1) \log q,$$

so the average information rate is not greater than $n/(n+1)$. □

5.1. *A Linear Programming Problem*

With respect to the information rate $\rho^*(G)$, we solved a linear programming problem to obtain a lower bound. Now, for average information rate $\tilde{\rho}^*(G)$, we obtain an *upper* bound by solving a linear programming problem.

Let G be a graph, and define a subgraph G_1 of G as follows: $xy \in E(G_1)$ if and only if vertices $w, z \in V(G)$ exist such that $G[w, x, y, z] = \{wx, xy, yz\}$ or $G[w, x, y, z] = \{wx, xy, yz, xz\}$. We take $V(G_1)$ to consist of all vertices in $V(G)$ that are incident with at least one edge in $E(G_1)$ (i.e., we delete all isolated vertices from G_1). We say that G_1 is the *foundation* of G.

For example, the path P_4, having edges $\{AB\}, \{BC\}, \{CD\}, \{DE\}$, has a foundation consisting of the two edges $\{BC\}, \{CD\}$.

If xy is an edge in the foundation of a graph G, then by Theorem 2.6, $\log s_x + \log s_y \geq 3 \log q$ for any secret sharing scheme with access structure $cl(E(G))$. Consider the following linear programming problem $\mathscr{A}(G)$:

$$\begin{aligned} \text{Minimize} \quad & C = \sum_{v \in V(G)} a_v \\ \text{subject to} \quad & a_v \geq 0, \qquad v \in V(G), \\ & a_v + a_w \geq 1, \qquad vw \in E(G_1). \end{aligned}$$

Then we have the following upper bound on the average information rate.

Theorem 5.2. *Let G be a graph with foundation G_1. Let C^* be the optimal solution to problem $\mathscr{A}(G)$. Then*

$$\tilde{\rho}^*(G) \leq \frac{|V(G)|}{C^* + |V(G)|}.$$

Proof. Consider any secret sharing scheme realizing the access structure $cl(E(G))$. For every vertex $v \in V(G)$, define

$$a_v = \frac{\log s_v}{\log q} - 1.$$

Suppose vw is an edge of the foundation G_1. Now, from Theorem 2.6, we get $\log s_v + \log s_w \geq 3 \log q$, or $a_v + a_w \geq 1$. For any $v \in V(G)$, we have $s_v \geq q$, so $a_v \geq 0$. Hence, the a_v's, as defined above, are a feasible solution for problem $\mathscr{A}(G)$. Hence,

$$C^* \leq \sum_{v \in V(G)} a_v,$$

where C^* is the optimal solution to $\mathscr{A}(G)$. It follows that

$$C^* \leq \frac{\sum_{v \in V(G)} \log s_v}{\log q} - |V(G)|.$$

However, then we have

$$\begin{aligned}\tilde{\rho}(G) &= \frac{|V(G)| \log q}{\sum_{v \in V(G)} \log s_v} \\ &\leq \frac{|V(G)|}{C^* + |V(G)|},\end{aligned}$$

which is the bound to be proved. □

Remark. Given a graph G, the foundation G_1 can be determined in polynomial time. One way to do this is to check all 4-subsets of $V(G)$. Every time we get an induced subgraph isomorphic to P_3, we can add one edge to the foundation; and every time we find an induced subgraph isomorphic to H, we can add two edges to the foundation. This algorithm requires time $O(n^4)$, where $n = |V(G)|$. Since the linear programming problem $\mathscr{A}(G)$ can be solved in polynomial time, so the bound of Theorem 5.2 can be computed in polynomial time too.

We have a couple of general observations on the linear programming problem $\mathscr{A}(G)$. Let d be a positive ineger. A *d-factor* of a graph is a spanning subgraph that is regular of degree d.

Lemma 5.3. *Let G be a graph having foundation G_1. If G_1 has a d-factor for some integer $d \geq 1$, then the optimal solution to $\mathscr{A}(G)$ is $C^* = |V(G_1)|/2$.*

Proof. Let the edges in the d-factor be $x_j y_j$, $1 \leq j \leq dn_1/2$, where without loss of generality $V(G_1) = \{1, \ldots, n_1\}$. Then we obtain the following:

$$\begin{aligned}\frac{dn_1}{2} &\leq \sum_{j=1}^{dn_1/2} (a_{x_j} + a_{y_j}) \\ &= d \sum_{i=1}^{n_1} a_i.\end{aligned}$$

Hence, $C^* \geq n_1/2$. To obtain $C^* \leq n_1/2$, let $a_i = \frac{1}{2}$ for $1 \leq i \leq n_1$, $a_i = 0$, otherwise. □

Next, note that $C^* \leq \beta(G_1)$. To see this, let W be a minimum vertex cover of G_1, and define $a_v = 1$ if $v \in W$; $a_v = 0$, otherwise. This gives a feasible solution for which

$$\sum_{v \in V(G)} a_v = \beta(G_1).$$

In the case where G_1 is bipartite, this will in fact be the optimal solution, as follows.

Lemma 5.4. *Let G be a graph having foundation G_1. If G_1 is bipartite, then the optimal solution to $\mathscr{A}(G)$ is $C^* = \beta(G_1)$ and the optimal solution is given by $a_v = 1$ if $v \in W$, $a_v = 0$, otherwise, where W is a minimum vertex cover of G_1.*

Proof. It is well known that the incidence matrix of a bipartite graph is a totally unimodular matrix (that is, the determinant of any square submatrix is 0, 1, or -1). Hence, if G_1 is bipartite, the linear programming problem $\mathscr{A}(G)$ and the corresponding integer programming problem have the same optimal solution. However, an optimal solution to the integer programming problem is obtained from a vertex cover, as described above. □

Hence, we have the following bound as an immediate consequence.

Theorem 5.5. *Let G be a graph with foundation G_1, and suppose G_1 is bipartite. Then*

$$\tilde{\rho}^*(G) \leq \frac{|V(G)|}{\beta(G_1) + |V(G)|}.$$

5.2. *Vertex Covers and Secret Sharing Schemes*

From Theorem 3.1, a secret sharing scheme for a graph G exists with average information rate

$$\tilde{\rho} = \frac{|V(G)|}{\sum_{v \in V(G)} R_v}.$$

Suppose we construct a scheme by using a star decomposition, as in Section 4.3. Let W denote the set of centers of the stars used in the decomposition. Then W must be a vertex cover of G. Conversely, if W is a vertex cover of G, then we can use it to construct a star decomposition of G and hence a secret sharing scheme. The algorithm to do this is as follows:

```
Algorithm
  Let $W = \{v_1, \ldots, v_n\}$ be a vertex cover of $G$
  $\Pi \leftarrow \emptyset$
  For $i \leftarrow 1$ to $n$ do
    $X \leftarrow v_i$
    $\Pi \leftarrow \Pi \cup \{G_X\}$
    $B \leftarrow \{Y \in Adj(X): |Inc(Y)| = 1\}$
    $E(G) \leftarrow E(G) - Inc(X)$
    $V(G) \leftarrow V(G) - (B \cup \{X\})$
  Output the star decomposition $\Pi$
```

We now show that if we construct a scheme from the star decomposition Π, then we can express $\tilde{\rho}$ as a function of $|V(G)|$, $|E(G)|$, and $|W|$. Let $\Pi = \{G_1, \ldots, G_n\}$, where $n = |W|$. Consider a star $G_i = K_{1,m}$ in the decomposition. The total number of shares in the scheme $PS(G_i, 1, q)$ is $m + 1 = |E(G_i)| + 1$. Hence, the total number of shares in the scheme for G is

$$\sum_{v \in V(G)} R_v = \sum_{i=1}^{n} (|E(G_i)| + 1)$$
$$= |E(G)| + |W|.$$

Hence, applying Theorem 3.1, we have the following result.

Theorem 5.6. *Let G be a graph and let $W \subseteq V(G)$ be a vertex covering. Then a secret sharing scheme for G exists with average information rate*

$$\tilde{\rho} = \frac{|V(G)|}{|E(G)| + |W|}.$$

Since $\tilde{\rho}$ depends only on $|W|$, finding the maximum rate among all vertex coverings is equivalent to minimizing $|W|$, i.e., determining the vertex covering number $\beta(G)$. Unfortunately, the problem of computing $\beta(G)$ is NP-hard [18]. However, for certain classes of graphs, such as bipartite graphs and chordal graphs, $\beta(G)$ can be computed in polynomial time (see [18]). We return to this in Section 5.4.

Let us mention a couple of general bounds that can be proved by this technique. It is obvious that $W \subseteq V(G)$ is a vertex covering of G if and only if $V - W$ is a stable set of G. Hence, $\beta(G) = |V(G)| - \alpha(G)$. Using known lower bounds on the stability number of a graph, we can obtain the following corollaries to Theorem 5.6.

Corollary 5.7. *Let G be a graph with $|V(G)| = n$ and $|E(G)| = m$. Then*

$$\tilde{\rho}^*(G) \geq \frac{n(2m + n)}{m(2m + 3n)}.$$

Proof. Use Theorem 5.6 and the bound $\alpha(G) \geq n^2/(2m + n)$ [3, Corollary 2, p. 279]. □

Corollary 5.8. *Let G be a graph with $|V(G)| = n$ and maximum degree d. Then*

$$\tilde{\rho}^*(G) \geq \frac{1}{((d + 2)/2) - (1/n)\lceil n/(d + 1)\rceil}.$$

Proof. Use Theorem 5.6 and the bound $\alpha(G) \geq \lceil n/(d + 1)\rceil$ [3, Corollary 2, p. 276]. □

Note that the bound on average information rate given by Corollary 5.8 exceeds the bound on information rate proved in Theorem 3.8 of [14].

5.3. *Average Information Rate for Paths and Cycles*

In this section we give an upper bound for average information rate for P_n, the path of length n. Then we show how to construct secret sharing schemes with optimal average information rate.

If n is equal either to 1 or to 2, then P_n is a complete multipartite graph and a secret sharing scheme with an average information rate equal to 1 exists. If

n is greater than 2, then the next theorem provides the optimal average information rate.

Theorem 5.9. *The optimal average information rate of a secret sharing scheme for P_n, where $n \geq 3$, is given by*

$$\tilde{\rho}^*(P_n) = \begin{cases} \dfrac{2(n+1)}{3n} & \textit{if n is even,} \\[2ex] \dfrac{2(n+1)}{3n+1} & \textit{if n is odd.} \end{cases}$$

Proof. It is easy to see that the foundation of P_n consists of the edges

$$X_2X_3, \ldots, X_{n-1}X_n,$$

so it is isomorphic to P_{n-2}. P_{n-2} is bipartite, an $\beta(P_{n-2}) = \lfloor(n-1)/2\rfloor$. First suppose that n is even and $n \geq 4$. By applying Theorem 5.5 we know that $\tilde{\rho}^*(P_n) \leq 2(n+1)/3n$. We have $\tilde{\rho}^*(P_n) \geq 2(n+1)/3n$ by using the CMC Π_1 from Theorem 4.5. If n is odd and $n \geq 3$, then $\tilde{\rho}^*(P_n) \leq 2(n+1)/(3n+1)$ by Theorem 5.5: We obtain a secret sharing scheme with average information rate equal to $2(n+1)/(3n+1)$ by using the CMC Π_3 from Theorem 4.5. □

We now consider average information rate for cycles. If n is equal either to 3 or to 4, then C_n is a complete multipartite graph and a secret sharing scheme exists with an average information rate equal to 1. If n is greater than 4, then the next theorem gives the optimal average information rate for even length cycles, while for odd length cycles it gives upper and lower bounds.

Theorem 5.10. *The optimal average information rate of a secret sharing scheme for C_n, where $n \geq 5$, satisfies*

$$\tilde{\rho}^*(C_n) = \tfrac{2}{3} \qquad \textit{if n is even,}$$

$$\frac{2n}{3n+1} \leq \tilde{\rho}^*(C_n) \leq \tfrac{2}{3} \qquad \textit{if n is odd.}$$

Proof. It is easy to see that the foundation of C_n is again C_n. C_n is a 2-factor of itself, so $C^* = n/2$, by Lemma 5.3. Applying Theorem 5.2, we get $\tilde{\rho}^*(C_n) \leq \frac{2}{3}$. First, suppose n is even, $n \geq 6$. We have already shown in Theorem 4.6 that $\rho^*(C_n) = \frac{2}{3}$. Since $\tilde{\rho}^*(C_n) \geq \rho^*(C_n)$ and since $\tilde{\rho}^*(C_n) \leq \frac{2}{3}$, we obtain $\tilde{\rho}^*(C_n) = \frac{2}{3}$. Next, let n be odd, $n \geq 5$. From Theorem 4.7, we have $\rho^*(C_n) \geq 2n/(3n+1)$. Since $\tilde{\rho}^*(C_n) \geq \rho^*(C_n)$ and since $\tilde{\rho}^*(C_n) \leq \frac{2}{3}$, the stated bounds follow. □

5.4. *Average Information Rate for Trees*

In this section we discuss upper and lower bounds on the average information rate of secret sharing schemes for trees.

For a graph G, let *degree_one*(G) denote the set of vertices in $V(G)$ having degree one. Our first observation is that the foundation of a tree T can be constructed by deleting all degree one vertices from T.

Lemma 5.11. *Let T be a tree; then the foundation of T is*

$$T_1 = T[V(T) - degree\_one(T)].$$

Proof. Let xy be an edge of T. If $\{x, y\} \cap degree\_one(T) \neq \emptyset$, then, clearly, $xy \notin E(T_1)$. So assume $\{x, y\} \cap degree\_one(T) = \emptyset$. Let wx, $yz \in E(T)$, where $w \neq y$, $z \neq x$. Since T is a tree, $wy, wz, xz \notin E(T)$. Hence, $T[w, x, y, z] = \{wx, xy, yz\}$ and $xy \in E(T_1)$. □

Remark. It is not difficult to see that the conclusion remains true if T is any bipartite graph having girth at least six.

Here now are our upper and lower bounds on the average information rate for trees.

Theorem 5.12. *Let T be a tree and let $T_1 = T[V(T) - degree\_one(T)]$. Then we have*

$$\frac{|V(T)|}{\beta(T) + |V(T)| - 1} \leq \tilde{\rho}^*(T) \leq \frac{|V(T)|}{\beta(T_1) + |V(T)|}.$$

Proof. By Lemma 5.11, T_1 is the foundation of T. Hence, the upper bound on $\tilde{\rho}^*$ follows from Theorem 5.5. The lower bound follows from Theorem 5.6, since $|E(T)| = |V(T)| - 1$ for any tree T. □

Remarks. 1. Since T and T_1 are bipartite graphs, the vertex covering numbers can be computed in polynomial time. In fact, by Konig's theorem, the vertex covering number of a biparitite graph equals the size of a maximum matching.

2. The reader can check that, in the special case where T is a path, the upper and lower bounds of Theorem 5.12 coincide, and they agree with Theorem 5.9.

Now, we give a general lower bound on the average information rate for trees.

Theorem 5.13. *Let T be a tree with n vertices. Then*

$$\tilde{\rho}^*(T) \geq \frac{2n}{3n - 2}.$$

Proof. In a bipartite graph G with vertex bipartition V_1, V_2, both V_1 and V_2 are vertex covers. Hence,

$$\beta(G) \leq \min\{|V_1|, |V_2|\} \leq \frac{|V_1| + |V_2|}{2}.$$

A tree is bipartite, so $\beta(T) \leq n/2$. Apply Theorem 5.12 to obtain the stated result. □

6. The Connected Graphs on at Most Five Vertices

In this section we give upper and lower bounds on the information rate and average information rate for the connected graphs on at most five vertices (excluding the trivial graph K_1). First, there are nine connected graphs on at most four vertices.

Table 1. Information rate and average information rate.

Graph	Information rate	Average information rate
$G_1, \ldots, G_9$	$\rho^* = \frac{2}{3}$	$\tilde{\rho}^* = \frac{5}{6}$
G_{10}, G_{11}	$\rho^* = \frac{2}{3}$	$\tilde{\rho}^* = \frac{5}{7}$
G_{12}	$\frac{5}{8} \le \rho^* \le \frac{2}{3}$	$\frac{5}{8} \le \tilde{\rho}^* \le \frac{2}{3}$
G_{13}	$\frac{3}{5} \le \rho^* \le \frac{2}{3}$	$\frac{5}{7} \le \tilde{\rho}^* \le \frac{10}{13}$
G_{14}	$\frac{3}{5} \le \rho^* \le \frac{2}{3}$	$\tilde{\rho}^* = \frac{5}{7}$
G_{15}	$\frac{4}{7} \le \rho^* \le \frac{2}{3}$	$\tilde{\rho}^* = \frac{5}{7}$

Seven of these are complete multipartite graphs and admit ideal schemes: K_2, K_3, $K_{1,2}$, K_4, $K_{1,3}$, $K_{2,2}$, $K_{1,1,2}$. The remaining two graphs are P_3 (the path of length 3) and the graph H (from Example 3.2). We have already shown that $\rho^*(P_3) = \frac{2}{3}$ (Theorem 4.5) and $\tilde{\rho}^*(P_3) = \frac{4}{5}$ (Theorem 5.9). With regard to H, we have $\rho^*(H) = \frac{2}{3}$ (Theorem 4.1) and $\tilde{\rho}^*(H) = \frac{4}{5}$ (Example 3.2 and Theorem 5.1).

So, let us move on to the connected graphs on five vertices. There are 21 nonisomorphic connected graphs on five vertices. Of these 21 graphs, six are complete multipartite graphs and admit ideal schemes. These graphs are $K_{1,4}$, $K_{2,3}$, $K_{1,1,3}$, $K_{1,2,2}$, $K_{1,1,1,2}$, and K_5. The remaining 15 graphs are depicted in the Appendix, where we also show the minimal CMCs for each graph.

The bounds on information rate and average information rate are summarized in Table 1. The lower bounds are obtained by making use of CMC constructions. Upper bounds on information rate are give by Theorem 4.2, whereas upper bounds on average information rate are given by application of Theorem 5.2.

The first CMC for each graph in the Appendix gives rise to the scheme that attains the given lower bound for the average information rate. For the graphs $G_1, \ldots, G_{11}$, the schemes with information rate equal to $\frac{2}{3}$ are obtained by taking one copy of each CMC shown in the Appendix. We next consider the lower bounds on ρ^* for the remaining four graphs, $G_{12}, \ldots, G_{15}$.

- First, let $E(G_{12}) = \{AB, BC, CD, DE, AE\}$. Then G_{12} is the cycle C_5 and $\rho_C^*(G_{12}) = \frac{5}{8}$ from Theorem 4.7.
- Let $E(G_{13}) = \{AB, BC, BE, EC, CD\}$. $\rho_C^*(G_{13}) = \frac{3}{5}$ is realized by using the three CMCs shown in the Appendix.
- Let $E(G_{14}) = \{AB, AD, BD, BC, DE, CE\}$. $\rho_C^*(G_{14}) = \frac{3}{5}$ is realized by using the three CMCs shown in the Appendix.
- Finally we consider $E(G_{15}) = \{AB, AC, AD, BC, BE, CD, CE\}$. The four minimal CMCs of G_{15} are depicted in the Appendix. The matrix of entries R_{jv} is

$$\begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 & 1 \\ 2 & 3 & 1 & 2 & 2 \\ 3 & 2 & 1 & 2 & 2 \end{bmatrix}.$$

Hence the linear programming problem to be solved is the following:

$$\begin{aligned} &\text{Minimize } T \\ &\text{subject to} \quad a_j \geq 0,\ 1 \leq j \leq 4, \\ &\qquad \sum_{j=1}^{4} a_j = 1, \\ &\qquad T \geq a_1 + 2a_2 + 2a_3 + 3a_4, \\ &\qquad T \geq 2a_1 + a_2 + 3a_3 + 2a_4, \\ &\qquad T \geq 2a_1 + 2a_2 + a_3 + a_4, \\ &\qquad T \geq a_1 + a_2 + 2a_3 + 2a_4. \end{aligned}$$

The optimal solution is

$$(a_1, a_2, a_3, a_4, T) = (\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{1}{4}, 0, \tfrac{7}{4}).$$

Hence, $\rho_C^*(G_{15}) = \frac{4}{7}$, and this rate can be attained by taking one copy of Π_1, two copies of Π_2, and one copy of Π_3.

Now we turn to the upper bounds on average information rate. Theorem 5.1 gives the upper bound $\tilde{\rho}^* \leq \frac{5}{6}$ for $G_1, \ldots, G_9$. So, the six graphs $G_{10}, \ldots, G_{15}$ remain to be considered.

- Consider the graph G_{10}. The foundation of G_{10} consists of the four edges *BC*, *BE*, *DC*, *DE*. This foundation is a 2-regular graph on four vertices, so $C^* = 2$ (Lemma 5.3). Hence, by Theorem 5.2, $\tilde{\rho}^* \leq \frac{5}{7}$.
- Consider the graph G_{11}. The foundation of G_{11} consists of the four edges *BC*, *BE*, *DC*, *DE*. As with G_{10}, we obtain $\tilde{\rho}^* \leq \frac{5}{7}$.
- G_{12} is the cycle of length 5, so $\tilde{\rho}^* \leq \frac{2}{3}$ (Theorem 5.10).
- Consider te graph G_{13}. The foundation of G_{13} consists of the three edges *BC*, *BE*, *CE*. This foundation is a 2-regular graph on three vertices, so $C^* = \frac{3}{2}$ (Lemma 5.3). Hence, by Theorem 5.2, $\tilde{\rho}^* \leq \frac{10}{13}$.
- Consider the graph G_{14}. The foundation of G_{14} consists of the five edges *AB*, *AD*, *BC*, *BD*, *DE*. The optimal solution to the linear programming problem is $C^* = 2$. Hence, by Theorem 5.2, $\tilde{\rho}^* \leq \frac{5}{7}$.
- Consider the graph G_{15}. The foundation of G_{15} consists of the five edges *AB*, *AC*, *BC*, *CD*, *CE*. The optimal solution to the linear programming problem is $C^* = 2$. As in the previous case, we get $\tilde{\rho}^* \leq \frac{5}{7}$.

7. Comments

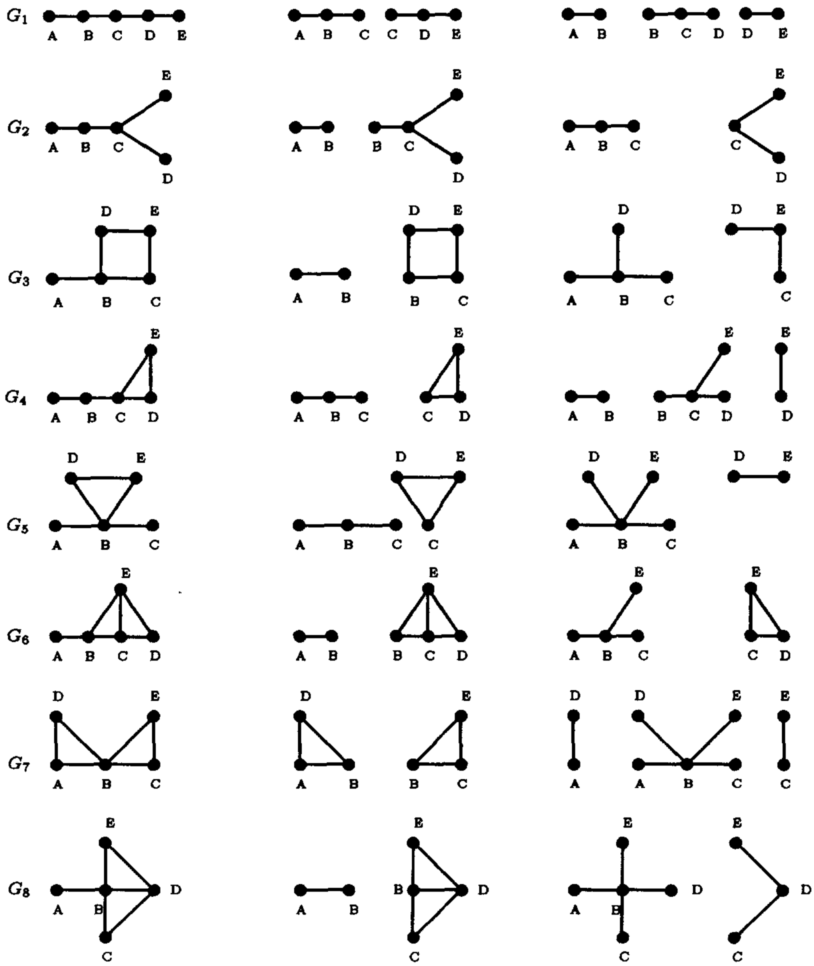
Since this paper was written in January 1992, some of the results in it have been improved. We briefly summarize some of these improvements now. In [8] the entropy technique is used to exhibit access structures for which the optimal information rate is $\frac{1}{2} + \varepsilon$ for arbitrarily small positive constants ε. The exact values of optimal information rate and optimal average information rate for the remaining graphs on five vertices are determined in [39]. The information rate of any cycle of odd length (≥ 5) is shown to be $\frac{2}{3}$ in Corollary 5.3 of [40]. The average information rate for a tree with n internal vertices is shown to be at least $n/(2n - 1)$ in Theorem

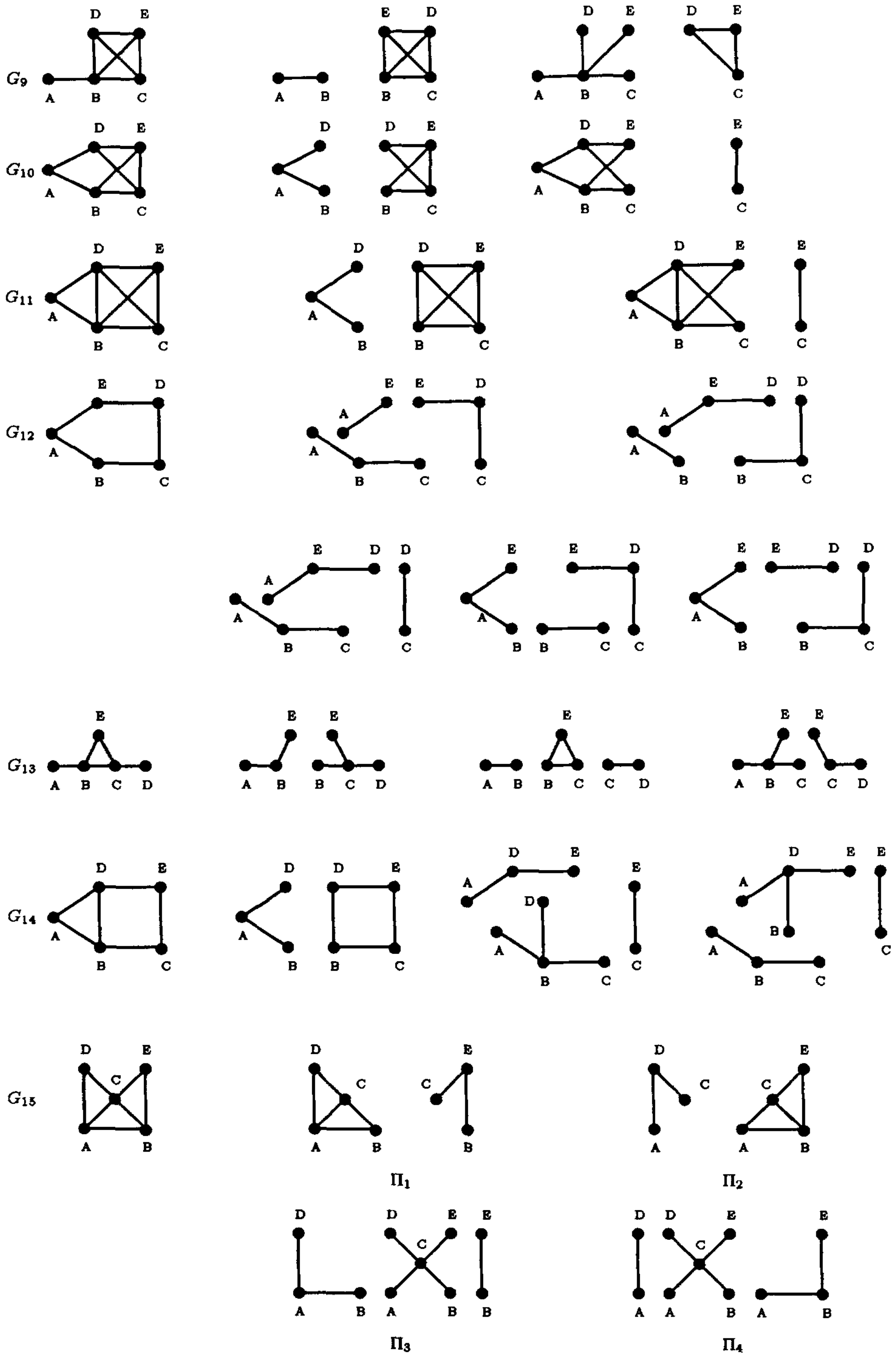
5.3 of [8]. It is proved in Theorem 5.2 of [40] that the average information rate for a graph on n vertices and m edges is at least $2n/(2m + n)$. This improves Corollary 5.7 for any connected graph. Finally, we also mention a recent expository paper on information rates of secret sharing schemes [38].

Acknowledgments

D. R. Stinson would like to thank Dean Hoffman, Wen-ai Jackson, U. S. R. Murty, John van Rees, and Qiu-rong Wu for helpful discussions. In particular, Dean Doffman provided the linear programming formulation given in Theorem 4.4, and Wen-ai Jackson and the referee pointed out the second minimal decomposition of graph G_{11} in the Appendix. Thanks also to the referee for carefully reading the manuscript and for providing numerous valuable suggestions, including a shortened proof of Theorem 4.2.

Appendix. Minimal CMCs for the Connected Graphs on Five Vertices




G_9
G_{10}
G_{11}
G_{12}
G_{13}
G_{14}
G_{15}
A
B
C
D
E
Π_1
Π_2
Π_3
Π_4

References

[1] A. Beimel and B. Chor. Universally ideal secret sharing schemes. *Proc. Crypto '92.* Lecture Notes in Computer Science, Vol. 740. Springer-Verlag, Berlin, 1993, pp. 185–197.

[2] J. Benaloh and J. Leichter. Generalized secret sharing and monotone functions. *Proc. Crypto '88.* Lecture Notes in Computer Science, Vol. 403. Springer-Verlag, Berlin, 1990, pp. 27–35.

[3] C. Berge. *Graphs,* 2nd rev. edn. North-Holland, Amsterdam, 1985.

[4] G. R. Blakley. Safeguarding cryptographic keys. *AFIPS Conf. Proc.* **48** (1979), 313–317.

[5] B. Blakley, G. R. Blakley, A. H. Chan, and J. L. Massey. Threshold schemes with disenrollment. *Proc. Crypto '92.* Lecture Notes in Computer Science, Vol. 740. Springer-Verlag, Berlin, 1993, pp. 546–54.

[6] C. Blundo. Secret Sharing Schemes for Access Structures Based on Graphs. Tesi di Laurea, University of Salerno, 1991 (in Italian).

[7] C. Blundo, A. Cresti, A. De Santis, and U. Vaccaro. Fully dynamic secret sharing schemes. *Proc. Crypto '93.* Lecture Notes in Computer Science, Vol. 773. Springer-Verlag, Berlin, 1994, pp. 110–125.

[8] C. Blundo, A. De Santis, L. Gargano, and U. Vaccaro. On the information rate of secret sharing schemes. *Proc. Crypto '92.* Lecture Notes in Computer Science, Vol. 740. Springer-Verlag, Berlin, 1993, pp. 149–169.

[9] C. Blundo, A. De Santis, and U. Vaccaro. Efficient sharing of many secrets. *Proc. STACS '93.* Lecture Notes in Computer Science, Vol. 665. Springer-Verlag, Berlin, 1993, pp. 692–703.

[10] E. F. Brickell. Some ideal secret sharing schemes. *J. Combin. Math. Combin. Comput.* **9** (1989), 105–113.

[11] E. F. Brickell and D. M. Davenport. On the classification of ideal secret sharing schemes. *J. Cryptology* **4** (1991), 123–134.

[12] E. F. Brickell and D. R. Stinson. Some Improved Bounds on the Information Rate of Perfect Secret Sharing Schemes. Department of Computer Science and Engineering Report Series # 106, University of Nebraska, May 1990.

[13] E. F. Brickell and D. R. Stinson. The detection of cheaters in threshold schemes. *SIAM J. Discrete Math.* **4** (1991), 502–510.

[14] E. F. Brickell and D. R. Stinson. Some improved bounds on the information rate of perfect secret sharing schemes. *J. Cryptology* **5** (1992), 153–166.

[15] R. M. Capocelli, A. De Santis, L. Gargano, and U. Vaccaro. On the size of shares for secret sharing schemes. *J. Cryptology* **6** (1993), 157–168.

[16] M. Carpentieri, A. De Santis, and U. Vaccaro. Size of shares and probability of chearing in threshold schemes. *Proc. Eurocrypt '93.* Lecture Notes in Computer Science, Vol. 765. Springer-Verlag, Berlin, 1994, pp. 118–125.

[17] E. Dawson, E. S. Mahmoodian, and A. Rahilly. Orthogonal arrays and ordered threshold schemes. *Austral. J. Combin.* **8** (1993), 27–44.

[18] M. R. Garey and D. S. Johnson. *Computers and Intractability. A Guide to Theory of NP-Completeness.* Freeman, New York, 1979.

[19] O. Goldreich, S. Micall, and A. Wigderson. How to play any mental game. *Proc. 19th ACM Symp. on Theory of Computing,* 1987, pp. 218–229.

[20] I. Ingemarsson and G. J. Simmons. A protocol to set up shared secret schemes without the assistance of a mutually trusted party. *Proc. Eurocrypt '90.* Lecture Notes in Computer Science, Vol. 473. Springer-Verlag, Berlin, 1991, pp. 266–282.

[21] M. Ito, A. Saito, and T. Nishizeki. Secret sharing scheme realizing general access structure. *Proc. IEEE Globecom '87,* 1987, pp. 99–102.

[22] M. Ito, A. Saito, and T. Nishizeki. Multiple assignment scheme for sharing secret. *J. Cryptology* **6** (1993), 15–20.

[23] W.-A. Jackson and K. M. Martin. On ideal secret sharing schemes. Preprint.

[24] W.-A. Jackson, K. M. Martin, and C. M. O'Keefe. Multisecret threshold schemes. *Proc. Crypto '93.* Lecture Notes in Computer Science, Vol. 773. Springer-Verlag, Berlin, 1994, pp. 126–135.

[25] E. D. Karnin, J. W. Greene, and M. E. Hellman. On secret sharing systems. *IEEE Trans. Inform. Theory* **29** (1983), 35–41.

[26] K. M. Martin. Discrete Structures in the Theory of Secret Sharing. Ph.D. thesis, University of London, 1991.

[27] K. M. Martin. New secret sharing schemes from old. *J. Combin. Math. Combin. Comput.* **14** (1993), 65–77.

[28] R. J. McEliece and D. V. Sarwate. On sharing secrets and Reed–Solomon codes. *Comm. ACM* **24** (1981), 583–584.

[29] T. Rabin and M. Ben-Or. Verifiable secret sharing and multiparty protocols with honest majority. *Proc. 21st ACM Symp. on Theory of Computing*, 1989, pp. 73–85.

[30] P. D. Seymour. On secret-sharing matroids. *J. Combin. Theory Ser. B* **56** (1992), 69–73.

[31] A. Shamir. How to share a secret. *Comm. ACM* **22** (1979), 612–613.

[32] G. J. Simmons. Robust shared secret schemes or "how to be sure you have the right answer even though you don't know the question." *Congr. Numer.* **68** (1989), 215–248.

[33] G. J. Simmons. How to (really) share a secret. *Proc. Crypto '88*. Lecture Notes in Computer Science, Vol. 403. Springer-Verlag, Berlin, 1990, pp. 390–448.

[34] G. J. Simmons. Prepositioned shared secret and/or shared control schemes. *Proc. Eurocrypt '89*. Lecture Notes in Computer Science, Vol. 434. Springer-Verlag, Berlin, 1990, pp. 436–467.

[35] G. J. Simmons. Shared secret and/or shared control schemes. *Proc. Crypto '90*. Lecture Notes in Computer Science, Vol. 537. Springer-Verlag, Berlin, 1991, pp. 216–241.

[36] G. J. Simmons. An introduction to shared secret and/or shared control schemes and their application. *Contemporary Cryptology*, IEEE Press, New York, 1991, pp. 441–497.

[37] G. J. Simmons, W. Jackson, and K. Martin. The geometry of shared secret schemes. *Bull. ICA* **1** (1991), 71–88.

[38] D. R. Stinson. An explication of secret sharing schemes. *Designs Codes Cryptography* **2** (1992), 357–390.

[39] D. R. Stinson. New general lower bounds on the information rate of secret sharing schemes. *Proc. Crypto '92*. Lectures Notes in Computer Science, Vol. 740. Springer-Verlag, Berlin, 1993, pp. 170–184.

[40] D. R. Stinson. Decomposition constructions for secret sharing schemes. *IEEE Trans. Inform. Theory* **40** (1994), 118–125.

[41] M. Tompa and H. Woll. How to share a secret with cheaters. *J. Cryptology* **1** (1988), 133–138.