On the Information Rate of Secret Sharing Schemes*

Carlo Blundo, Alfredo De Santis, Luisa Gargano, Ugo Vaccaro

Dipartimento di Informatica ed Applicazioni, Università di Salerno, 84081 Baronissi (SA), Italy

Preliminary version

Abstract

We derive new limitations on the information rate and the average information rate of secret sharing schemes for access structure represented by graphs. We give the first proof of the existence of access structures with optimal information rate and optimal average information rate less that $1/2+\epsilon$, where ϵ is an arbitrary positive constant. We also consider the problem of testing if one of these access structures is a sub-structure of an arbitrary access structure and we show that this problem is NP-complete. We provide several general lower bounds on information rate and average information rate of graphs. In particular, we show that any graph with n vertices admits a secret sharing scheme with information rate $\Omega((\log n)/n)$.

1 Introduction

A secret sharing scheme is a method to distribute a secret s among a set of participants $\mathcal{P}$ in such a way that only qualified subsets of $\mathcal{P}$ can reconstruct the value of s whereas any other subset of $\mathcal{P}$, non-qualified to know s, cannot determine anything about the value of the secret. We briefly recall the results on secret sharing schemes that are more closely related to the topics of this paper.

Shamir [36] and Blackley [3] were the first to consider the problem of secret sharing and gave secret sharing schemes where each subset A of $\mathcal{P}$ of size $|A| \geq k$ can reconstruct the secret, and any subset A of participants of size $|A| < k$ have absolutely no information on the secret. These schemes are known as (n, k) *threshold schemes*; the value k is the threshold of the scheme and n is the size of $\mathcal{P}$.

Ito, Saito and Nishizeki [25] considered a more general framework and showed how to realize a secret sharing scheme for any access structure. An access structure is a family of all subsets of $\mathcal{P}$ which are qualified to recover the secret. In case of (n, k) threshold schemes the access structure consists of all subsets of $\mathcal{P}$ that have size greater than or equal to k. Their technique requires that the size of set where the shares are taken be very large compared to the size of the set where the secret is chosen. Benaloh and Leichter [2] proposed a technique to realize a secret sharing scheme for any access structure more efficient than Ito, Saito and Nishizeki's methodology. Benaloh and Leichter showed that there exist access structures for which any secret sharing scheme must give to some participant a share which is from a domain strictly larger than that of the secret.

Brickell and Davenport [12] analyzed ideal secret sharing schemes in terms of matroids. An ideal secret sharing scheme is a scheme for which the shares are taken from a set that has the same

*Partially supported by Italian Ministry of University and Research (M.U.R.S.T.) and by National Council for Research (C.N.R.) under grant 91.02326.CT12.

size of the set where the secret is chosen. In particular, in case the access structure consists of only those subsets of participants containing an edge of a given graph G, Brickell and Davenport [12] proved that an ideal secret sharing scheme exists if and only if G is a complete multipartite graph. Equivalently, if we define the information rate of an access structure as the ratio between the size of the secret and that of the largest share given to any participant, Brickell and Davenport's result can be stated saying that a graph has information rate 1 if and only if it is a complete multipartite graph.

The problem of establishing bounds on the size of the shares to be given to participants in secret sharing schemes, or equivalently on the information rate, is one of the basic problem in the area and has received considerable attention by several researchers. The practical relevance of this issue is based on the following observations: Firstly, the security of any system tends to degrade as the amount of information that must be kept secret, i.e., the shares of the participants, increases. Secondly, if the shares given to participants are too long, the memory requirements for the participants will be too severe and, at the same time, the shares distribution algorithms will become inefficient. Therefore, it is important to derive significative upper and lower bounds on the information rate for classes of access structures.

Brickell and Stinson [14] gave several upper and lower bounds on the information rate of access structures based on graphs. Stinson in [41] presented new lower bounds on general access structures. Capocelli, De Santis, Gargano, and Vaccaro [15] gave the first example of access structures with information rate bounded away from 1.

Blundo, De Santis, Stinson, and Vaccaro [9] analyzed the information rate and the average information rate of secret sharing schemes based on graphs. The average information rate is the ratio between the secret size and the arithmetic mean of the size of the shares for such schemes. They proved the existence of a gap in the values of information rates of graphs, more precisely they strengthened the above quoted result of Brickell and Davenport [12], proving that if a graph G with n vertices is not a complete multipartite graph then any secret sharing scheme for it has information rate not greater than $2/3$ and average information rate not greater than $n/(n+1)$. These upper bounds arise by applying entropy argument due to Capocelli, De Santis, Gargano, and Vaccaro [15].

A discussion of the best bounds known so far and of our improvements is presented in the technical sections of the paper.

The recent survey by Stinson [40] contains an unified description of recent results in the area of secret sharing schemes. For different approaches to the study of secret sharing schemes, for schemes with "extended capabilities" as disenrollment, fault-tolerance, and pre-positioning and for a complete bibliography we recommend the survey article by Simmons [39].

We also mention some "extended capabilities" of secret sharing schemes that have been studied. In [1] and [8] has been addressed the problem of designing secret sharing schemes having the additional feature that qualified minorities can forbid any other set of participants from reconstructing the secret. These schemes are referred to as secret sharing schemes with "veto" capability. Ingemarsson and Simmons [24] solve the question of how to set up a secret sharing scheme in the absence of a trusted party. Prepositioned schemes are studied in [38]. The idea of protecting against cheating by one or more participants is addressed in [30, 43, 35, 37, 13, 16]. In [4] the authors investigated threshold schemes that permit disenrollment of participants. 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 [6]. Schemes for sharing several non-independent secrets simultaneously have been analyzed in [10]; whereas, schemes where different secrets are associated with different subsets of participants are considered in [26] and [7]. Recently Naor and Shamir [33] considered a type of cryptographic scheme that is able to decode concealed images without any cryptographic computation. They extended it into a visual variant of the (n, k) secret sharing problem.

In this paper we derive new limitations on the information rate and the average information rate for access structures represented by graphs. The paper is organized as follows. In Section 2 we formally define secret sharing schemes using an information theoretical framework[1]. We also define the optimal (average) information rate of an access structure $\mathcal{A}$ by using the entropy approach. In Section 3 we prove new upper bounds on the information rate and the average information rate. These bounds are obtained by using the entropy approach introduced in [15] and are the best possible for the considered structures since we exhibit secret sharing schemes that meet the bounds. In particular, we give the first proof of the existence of access structures with information rate and average information rate strictly less that 2/3. This solves a problem of [9]. In Section 3.1 we also consider the problem of efficiently testing if one of these low–information–rate access structures is a sub-structure of an arbitrary access structure. This is important since it would immediately give an efficient way to get upper bounds on the information rate for classes of access structures. Unfortunately, we show that above decision problem is NP–complete. In Section 4 we consider the problem of finding good lower bounds on the information rate and the average information rate for access structures based on graphs and we give several general lower bounds that improve on previously known results. In particular, we show that any graph on n vertices of maximum degree d admits a secret sharing scheme with information rate $1/(\lceil d/2 \rceil + 1 - \lceil d/2 \rceil /n)$. We provide a scheme for any tree with n internal vertices having information rate equal to $n/(2n-1)$. Finally, we show that any graph with n vertices admits a secret sharing scheme with information rate $\Omega((\log n)/n)$ and any graph with n vertices and m edges admits of a secret sharing scheme with average information rate $\Omega\left(\frac{n \log n}{m \log \frac{n^2}{m}}\right)$.

2 Secret Sharing Schemes

A secret sharing scheme permits a secret to be shared among a set $\mathcal{P}$ of n participants in such a way that only qualified subsets of $\mathcal{P}$ can recover the secret, but any non-qualified subset has absolutely no information on the secret. An access structure $\mathcal{A}$ is the set of all subsets of $\mathcal{P}$ that can recover the secret.

Definition 2.1 *Let $\mathcal{P}$ be a set of participants, a* monotone *access structure $\mathcal{A}$ on P is a subset $\mathcal{A} \subseteq 2^{\mathcal{P}}$, such that*

$$A \in \mathcal{A}, A \subseteq A' \subseteq P \Rightarrow A' \in \mathcal{A}.$$

Definition 2.2 *Let $\mathcal{P}$ a set of participants and $\mathcal{A} \subseteq 2^{\mathcal{P}}$. The* closure *of $\mathcal{A}$, denoted by $cl(\mathcal{A})$, is the set*

$$cl(\mathcal{A}) = \{C | B \in \mathcal{A} \text{ and } B \subseteq C \subseteq \mathcal{P}\}.$$

For a monotone access structure $\mathcal{A}$ we have $\mathcal{A} = cl(\mathcal{A})$.

Let S be the set of secrets, $\{p_S(s)\}_{s \in S}$ be a probability distribution on S, and let a secret sharing scheme Σ for secrets in S be fixed. For any participant $P \in \mathcal{P}$, let us denote by $K(P)$ the set of all possible shares given to participant P. Suppose a *dealer* D wants to a share the secret $s \in S$ among the participants in $\mathcal{P}$ (we will assume that $D \notin \mathcal{P}$). He does this by giving each participant $P \in \mathcal{P}$

[1]All the necessary information theoretical definitions are listed in Appendix A, together with the basic terminology in graph theory.

a share from $K(P)$ chosen according to some, non necessarily uniform, probability distribution. Given a set of participants $A = \{P_{i_1}, \ldots, P_{i_r}\} \subset \mathcal{P}$, denote by $K(A) = K(P_{i_1}) \times \cdots \times K(P_{i_r})$.

We represent, as in [42], a secret sharing scheme Σ by a collection of distribution rules. A *distribution rule* is a function

$$f : \mathcal{P} \cup \{D\} \rightarrow K(\mathcal{P}) \cup S$$

which satisfies the conditions $f(D) \in S$ and $f(P_i) \in K(P_i)$, for $i = 1, 2, \ldots, n$. A distribution rule f represents a possible distribution of shares to the participants, where $f(D)$ is the secret being shared, and $f(P_i)$ is the share given to P_i. If $\mathcal{F}$ is a family of distribution rules and $s \in S$, then $\mathcal{F}_s = \{f \in \mathcal{F} \ : \ f(D) = s\}$ is the family of all distribution rules having s as the secret. If $s \in S$ is the value of the secret that D wants to share, then D will randomly choose a distribution rule $f \in \mathcal{F}_s$, according to some probability distribution, and use f to distribute shares to the participants.

The family of distribution rules $\mathcal{F}$ can also be depicted as a matrix M, each row of which corresponds to one distribution rule. One column of M will be indexed by D, and the remaining columns are indexed by the members of $\mathcal{P}$.

Any secret sharing scheme for secrets in S and a probability distribution $\{p_S(s)\}_{s \in S}$ naturally induce a probability distribution on $K(A)$, for any $A \subseteq \mathcal{P}$. Denote such probability distribution by $\{p_{K(A)}(a)\}_{a \in K(A)}$. Finally, denote by $H(S)$ the entropy of $\{p_S(s)\}_{s \in S}$ and by $H(A)$ the entropy of $\{p_{K(A)}(a)\}_{a \in K(A)}$, for any $A \in \mathcal{P}$.

In terms of the probability distribution on the secret and on the shares given to participants, we say that a secret sharing scheme is a *perfect* secret sharing scheme, or simply a secret sharing scheme, for the monotone access structure $\mathcal{A} \subseteq 2^{\mathcal{P}}$ if

1. *Any subset $A \subseteq \mathcal{P}$ of participants enabled to recover the secret can compute the secret:*
   If $A \in \mathcal{A}$ then for all $a \in K(A)$ with $p_{K(A)}(a) > 0$ a unique secret $s \in S$ exists such that $p(s|a) = 1$.

2. *Any subset $A \subseteq \mathcal{P}$ of participants not enabled to recover the secret has no information on the secret value:*
   If $A \notin \mathcal{A}$ then for all $s \in S$ and for all $a \in A$, it holds $p(s|a) = p_S(s)$.

Property 1. means that the value of the shares held by $A \in \mathcal{A}$ completely determines the secret $s \in S$. Notice that the property 2. means that the probability that the secret is equal to s given that the shares held by $A \notin \mathcal{A}$ are a, is the same as the *a priori* probability of the secret s. Therefore, no amount of knowledge of shares of participants not qualified to reconstruct the secret enables a Bayesian opponent to modify an *a priori* guess regarding which the secret is.

Following the approach of [27], [29], and [15] we can restate above conditions 1. and 2. using the information measures listed in Appendix A. Therefore, we say that a secret sharing scheme is a sharing of the secrets in S among participants in $\mathcal{P}$ such that

1′. *Any subset $A \subseteq \mathcal{P}$ of participants enabled to recover the secret can compute the secret:*
   Formally, for all $A \in \mathcal{A}$, it holds $H(S|A) = 0$.

2′. *Any subset $A \subseteq \mathcal{P}$ of participants not enabled to recover the secret has no information on the secret value:*
   Formally, for all $A \notin \mathcal{A}$, it holds $H(S|A) = H(S)$.

Notice that $H(S|A) = 0$ means that each set of values of the shares in A corresponds to a unique value of the secret. In fact, by definition, $H(S|A) = 0$ is equivalent to the fact that for all $a \in K(A)$ with $p_{K(A)}(a) > 0$ a unique $s \in S$ exists such that $p(s|a) = 1$. Moreover, $H(S|A) = H(S)$

is equivalent to state that S and $K(A)$ are statistically independent, i.e., for all $a \in K(A)$ and for all $s \in S$, it holds $p(s|a) = p_S(s)$ and therefore the knowledge of a gives no information about the secret.

2.1 The Size of the Shares

One of the basic problems in the field of secret sharing schemes is to derive bounds on the amount of information that must be kept secret. This is important from the practical point of view since the security of any system degrades as the amount of secret information increases.

Let $\mathcal{P}$ be a set of n participants and $\mathcal{A} \subseteq 2^{\mathcal{P}}$ be an access structure on $\mathcal{P}$. Different measures of the amount of secret information that must be distributed in a secret sharing scheme are possible. If we are interested in limiting the maximum size of shares for each participant (i.e., the maximum quantity of secret information that must be given to any participant), then a worst-case measure of the maximum of $H(P)$ over all $P \in \mathcal{P}$ naturally arises. To analyze such cases we use the *information rate* of $\mathcal{A}$ defined below. Given a set of secrets S, a non-trivial probability distribution Π_S on S, and a fixed secret sharing scheme Σ for $\mathcal{A}$, we define

$$\rho(\mathcal{A}, \Pi_S, \Sigma) = \frac{H(S)}{\max_{P \in \mathcal{P}} H(P)}.$$

This measure was introduced by Brickell and Stinson [14] when the probability distributions over the secret and the shares are uniform. In such a case the information rate reduces to $\log|S|/\max_{P\in\mathcal{P}} \log|K(P)|$, and correspond to the ratio between the size of the secret (measured in bits) and that of the largest share given to any participant. The *optimal information rate* of the access structure $\mathcal{A}$ is then defined as:

$$\rho^*(\mathcal{A}) = \sup_{\mathcal{Q},\mathcal{T}} \rho(\mathcal{A}, \Pi_S, \Sigma),$$

where $\mathcal{Q}$ is the space of all non-trivial probability distributions Π_S and $\mathcal{T}$ is the space of all secret sharing schemes for the access structure $\mathcal{A}$. In [27] and [15] it has been proved that in any secret sharing scheme the relation $H(P) \geq H(S)$ holds for any $P \in \mathcal{P}$. Since $H(P) = H(S)$, for any $P \in \mathcal{P}$, is the optimal situation we refer to such a scheme as an *ideal* scheme.

In many cases it is preferable to limit the sum of the size of shares given to all participants. In such a case the arithmetic mean of the $H(P)$, for $P \in \mathcal{P}$, is a more appropriate measure. We define the *average information rate* as follows. Given a set of secrets S, a non-trivial probability distribution Π_S on S, and a fixed secret sharing scheme Σ for $\mathcal{A}$, we define

$$\widetilde{\rho}(\mathcal{A}, \Pi_S, \Sigma) = \frac{H(S)}{\sum_{P \in \mathcal{P}} H(P)/|\mathcal{P}|}.$$

This measure was introduced in [5], [31], and [32] when an uniform probability distribution on the set of secrets is assumed. In such a case the average information rate reduces to $|\mathcal{P}|\log|S|/\sum_{P\in\mathcal{P}}\log|K(P)|$. Blundo, De Santis, Stinson, and Vaccaro [9] analyzed secret sharing schemes by means of this measure, when the probability distributions over the secret and the shares are uniform. If the secret and the shares are chosen under a uniform probability distribution, considering previous measure is equivalent to consider the "average size" of the shares assigned to each participant to realize a secret sharing scheme. The *optimal average information* rate of the access structure $\mathcal{A}$ is then defined as:

$$\widetilde{\rho}^*(\mathcal{A}) = \sup_{\mathcal{Q},\mathcal{T}} \widetilde{\rho}(\mathcal{A}, \Pi_S, \Sigma),$$

where $\mathcal{Q}$ is the space of all non-trivial probability distributions Π_S and $\mathcal{T}$ is the space of all secret sharing schemes for the access structure $\mathcal{A}$.

It is clear that, for the same secret sharing scheme and non-trivial probability distribution $\mathcal{P}_S$ on the secret, the information rate ρ is no greater than the average information rate $\widetilde{\rho}$, that is $\widetilde{\rho} \geq \rho$ and $\widetilde{\rho} = \rho$ if and only if all $H(P)$, for $P \in \mathcal{P}$, have the same value. In case the access structure $\mathcal{A}$ coincides with the closure of the edge-set of some graph $G(V(G), E(G))$, we will identify $\mathcal{A}$ with the graph G. As done in [9] we denote, for a graph G, the optimal information rate with $\rho^*(G)$ and the average information rate with $\widetilde{\rho}^*(G)$.

2.2 Auxiliary Results

In this section we recall some auxiliary results. We will improve some of them in the next sections and we will use others in our constructions.

Brickell and Stinson [14] proved the following lower bound on the information rate for any graph of maximum degree d. We denote with U_s the uniform probability distribution on the set of secrets S.

Theorem 2.1 *Let G be a graph with maximum degree d. Then for any set of secrets S of size $q \geq 2$, there exists a secret sharing scheme Σ with information rate*

$$\rho(G, U_s, \Sigma) = \frac{1}{\lceil d/2 \rceil + 1}.$$

In Section 3 we will show how to improve this bound for odd d. Blundo, De Santis, Stinson, and Vaccaro [9] proved the following result for trees.

Lemma 2.1 *Let G be a tree. Then for any set of secrets S of size $q \geq 2$, there exists a secret sharing scheme Σ with information rate $\rho(G, U_s, \Sigma) = 1/2$.*

In Section 3 we will show how to improve this bound for any tree.

The following results, proved in [9] and [42] will be used to obtain good secret sharing schemes for graphs with maximum degree 3.

Theorem 2.2 *Let C_n be a cycle of length n, $n \geq 5$. For any set of secrets S of size q^2, with $q \geq n$, a secret sharing scheme for C_n exists with optimal information rate* $2/3$.

The following lemmas have been proved by Capocelli, De Santis, Gargano, and Vaccaro [15]; we will use them to find new upper bounds on the information rate of access structures. Since their proofs are simple, we report them for reader's convenience.

Lemma 2.2 *Let $\mathcal{A}$ be an access structure on a set $\mathcal{P}$ of participants and $X, Y \subset \mathcal{P}$. Let $Y \notin \mathcal{A}$ and $X \cup Y \in \mathcal{A}$. Then $H(X|Y) = H(S) + H(X|YS)$.*

Proof. The conditional mutual information $I(X; S|Y)$ can be written either as $H(X|Y) - H(X|YS)$ or as $H(S|Y) - H(S|XY)$. Hence, $H(X|Y) = H(X|YS) + H(S|Y) - H(S|XY)$. Because of $H(S|XY) = 0$ for $X \cup Y \in \mathcal{A}$ and $H(S|Y) = H(S)$ for $Y \notin \mathcal{A}$, we have $H(X|Y) = H(S) + H(X|YS)$. □

Lemma 2.3 *Let $\mathcal{A}$ an access structure on a set $\mathcal{P}$ of participants and $X, Y \subset \mathcal{P}$. If $X \cup Y \notin \mathcal{A}$ then $H(Y|X) = H(Y|XS)$.*

Proof. The conditional mutual information $I(Y, S|X)$ can be written either as $H(Y|X) - H(Y|XS)$ or as $H(S|X) - H(S|XY)$. Hence, $H(Y|X) = H(Y|XS) + H(S|X) - H(S|XY)$. Because of $H(S|XY) = H(S|X) = H(S)$, for $X \cup Y \notin \mathcal{A}$, we have $H(Y|X) = H(Y|XS)$. □

Finally, we briefly recall a technique introduced in [9] to obtain lower bounds on the information rate of a graph G.

Suppose G is a graph, suppose that $\Pi = \{G_1, \ldots, G_t\}$ is a complete multipartite covering (or CMC) of G if $G_1, \ldots, G_n$ are subgraphs of G, each edge of G occurs in at least one of the G_i's and each G_i is a complete multipartite graph. 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 $R_{jv} = |\{i : v \in G_{ji}\}|$. 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 $\Pi' \neq \Pi$ such that $\Pi' \preceq \Pi$. Let $\Pi_j = \{G_{j1}, \ldots, G_{jn_j}\}$, $j = 1, \ldots L$, comprise a complete enumeration of the minimal CMCs of G. For every vertex v and for $j = 1, \ldots L$ define $R_{jv} = |\{i : v \in G_{ji}\}|$ and consider the following optimization problem $\mathcal{O}(G)$:

Minimize T subject to:

$$\begin{aligned} a_j &\geq 0, \;\; 1 \leq j \leq L \\ \textstyle\sum_{j=1}^{L} a_j &= 1 \\ T &\geq \textstyle\sum_{j=1}^{L} a_j R_{jv}, \;\; v \in V(G) \end{aligned}$$

In [9] it is proved that if T^* is the optimal solution to $\mathcal{O}(G)$ then $\rho^*(G) \geq 1/T^*$.

3 Upper Bounds on the Information Rate and Average Information Rate

In this section we will exhibit an access structure having optimal information rate less than 2/3. This solves an open problem in [9]. The result is obtained using the entropy approach of [15].

Consider the graph $\mathcal{AS}_k = (V(\mathcal{AS}_k), E(\mathcal{AS}_k))$, $k \geq 1$, where

$$V(\mathcal{AS}_k) = \{Y_0, X_0, X_1, \ldots, X_k, X_{k+1}, \ldots, X_{2k}\}$$

and

$$E(\mathcal{AS}_k) = \{(Y_0, X_0), (X_0, X_1), \ldots, (X_0, X_k), (X_1, X_{k+1}), \ldots, (X_k, X_{2k})\}.$$

As an example, the graph $\mathcal{AS}_k$ for $k = 3$ is depicted in Figure 1.*a*.

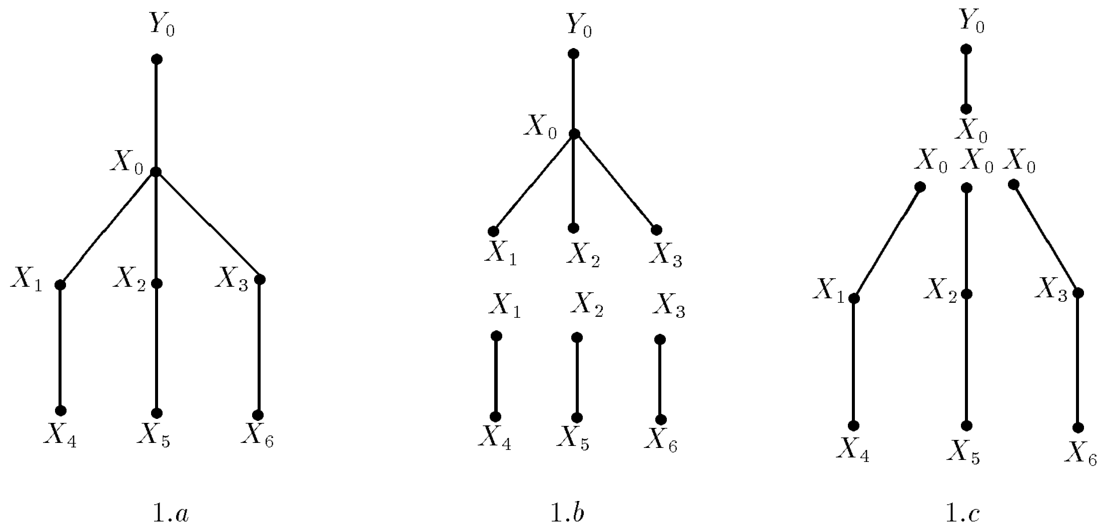


Figure 1

Theorem 3.1 *The optimal information rate of the graph* $\mathcal{AS}_k$, $k \geq 1$, *is*

$$\rho^*(\mathcal{AS}_k) = \frac{1}{2} + \frac{1}{4k+2},$$

and the optimal average information rate is

$$\widetilde{\rho}^*(\mathcal{AS}_k) = \frac{2}{3} + \frac{2}{9k+6}.$$

Proof: Consider the conditional entropy $H(X_1 \ldots X_k|Y_0)$. We have

$$\begin{aligned}
H(X_1 \ldots X_k|Y_0) &= H(X_1|Y_0) + H(X_2|X_1Y_0) + \cdots + H(X_k|X_1 \ldots X_{k-1}Y_0) \\
&\qquad\qquad \text{(from (4) of Appendix } A) \\
&\geq H(X_1|Y_0X_{k+1}) + H(X_2|X_1Y_0X_{k+2}) + \\
&\quad\; H(X_3|X_1X_2Y_0X_{k+3}) + \cdots + H(X_k|X_1 \ldots X_{k-1}Y_0X_{2k}) \\
&\qquad\qquad \text{(from (6) of Appendix } A) \\
&\geq kH(S) \quad \text{(from Lemma 2.2 and (3) of Appendix } A).
\end{aligned}$$

On the other hand, we have also

$$\begin{aligned}
H(X_1 \ldots X_k|Y_0) &= H(X_1 \ldots X_k|Y_0S) \quad \text{(from Lemma 2.3)} \\
&\leq H(X_0X_1 \ldots X_k|Y_0S) \quad \text{(from (4) and (3) of Appendix } A) \\
&\leq H(X_0|Y_0S) + H(X_1|X_0S) + \cdots + H(X_k|X_0S) \\
&\qquad\qquad \text{(from (4) and (6) of Appendix } A) \\
&= H(X_0|Y_0) - H(S) + \cdots + H(X_k|X_0) - H(S) \quad \text{(from Lemma 2.2)} \\
&\leq H(X_0) + \cdots + H(X_k) - (k+1)H(S) \quad \text{(from (5) of Appendix } A).
\end{aligned}$$

Therefore, we get

$$H(X_0) + H(X_1) + \ldots + H(X_k) \geq (2k+1)H(S). \tag{1}$$

From (1) it follows that there exists $i \in \{0, 1, \ldots, k\}$ such that

$$H(X_i) \geq \frac{2k+1}{k+1} H(S).$$

Therefore, the optimal information rate $\rho^*(\mathcal{AS}_k)$ is upper bounded by

$$\rho^*(\mathcal{AS}_k) \leq \frac{k+1}{2k+1} = \frac{1}{2} + \frac{1}{4k+2}.$$

From (1) and from Lemma 2.2 it follows that

$$H(Y_0) + \sum_{i=0}^{2k} H(X_i) \geq (3k+2)H(S).$$

Therefore, the optimal average information rate of $\mathcal{AS}_k$ is upper bounded by

$$\frac{2k+2}{3k+2} = \frac{2}{3} + \frac{2}{9k+6}.$$

Actually, $1/2 + 1/(4k+2)$ is the true value of the optimal information rate. This value can be attained by using the *CMC* technique presented in [9]. Consider the following two minimal complete multipartite coverings of $\mathcal{AS}_k$

$$\Pi_1 = \Big\{\{Y_0X_0, X_0X_1, \cdots, X_0X_k\}, \{X_1X_{k+1}, \cdots, X_kX_{2k}\}\Big\}$$

$$\Pi_2 = \Big\{\{Y_0X_0\}, \{X_0X_1, X_1X_{k+1}\}, \cdots, \{X_0X_k, X_kX_{2k}\}\Big\}.$$

(An example of these two covering of $\mathcal{AS}_k$ is depicted in Figures 1.*b* and 1.*c* for $k = 3$.) Taking k copies of Π_1 and one copy of Π_2 we can attain the rate $\rho^*(\mathcal{AS}_k) = (2k+1)/(k+1)$ for any $|S| \geq 2$. Thus, the optimal information rate of $\mathcal{AS}_k$ is $1/2 + 1/(4k+2)$. The optimal average information rate equal to $2/3 + 2/(9k+6)$ is attained by either Π_1 or Π_2. □

In case the probability distribution on the set of secrets is the uniform one, we obtain the following result, whose proof is immediate using Theorem 3.1 and inequality (2) of Appendix *A*.

Corollary 3.1 *Suppose $p_S(s) = 1/|S|$, for any $s \in S$. Then any secret sharing scheme for the access structure $\mathcal{AS}_k$ must give to at least a participant a share whose size is at least $2 - 1/(k+1)$ times the size of the secret.*

Theorem 3.1 is a generalization of Theorem 4.1 of [15]. In fact if we choose $k = 1$ the access structure $\mathcal{AS}_k$ is the closure of the edge-set of P_3, the path on four vertices.

In Appendix *B* are depicted all graphs on six vertices that have $\mathcal{AS}_2$ as induced subgraph and, therefore, have optimal information rate less than 3/5. It turns out that the optimal information rate for all those graphs is equal to 3/5, and all but one have also an optimal average information rate equal to 3/4.

Using Theorem 3.1 we can show the existence of access structures having *average information* rate less than 2/3, which represented the best upper bound known so far [15] on average information rate.

Consider the graph $\mathcal{M}_k$, where $V(\mathcal{M}_k) = \{X_1, X_2, \ldots, X_{2k+3}, X_{2k+4}\}$ and

$$E(\mathcal{M}_k) = \{X_1X_2\} \bigcup \{X_2X_i, X_iX_{k+i}, X_{k+i}X_{2k+3} | 3 \leq i \leq k+2\} \bigcup \{X_{2k+3}X_{2k+4}\}.$$

The graph $\mathcal{M}_3$ and a CMC that attains the optimal average information rate are depicted in Figure 2. The following theorem holds.

Theorem 3.2 *The optimal average information rate for $\mathcal{M}_k$, $k \geq 1$, is*

$$\widetilde{\rho}^*(\mathcal{M}_k) = \frac{1}{2} + \frac{1}{2k+2}.$$

Proof : From Lemma 2.2 we get $H(X_1) \geq H(S)$ and $H(X_{2k+4}) \geq H(S)$, whereas from Theorem 3.1 we have

$$\sum_{i=2}^{k+2} H(X_i) \geq 2k+1$$

and

$$\sum_{i=k+3}^{2k+3} H(X_i) \geq 2k+1.$$

Thus,

$$\sum_{i=1}^{2k+4} H(X_i) \geq 4k+4.$$

Hence,

$$\widetilde{\rho}^*(\mathcal{M}_k) \leq \frac{k+2}{2k+2} = \frac{1}{2} + \frac{1}{2k+2}.$$

It is easy to see that the following complete multipartite covering Π of the graph $\mathcal{M}_k$ meets previous bound.

$$\Pi = \Big\{ \{X_1X_2, X_2X_3, \ldots, X_2X_{k+2}\},$$
$$\{X_3X_{k+3}\}, \{X_4X_{k+4}\}, \ldots, \{X_{k+2}X_{2k+2}\},$$
$$\{X_{k+3}X_{2k+3}, \ldots, X_{2k+2}X_{2k+3}, X_{2k+3}X_{2k+4}\} \Big\}.$$

□

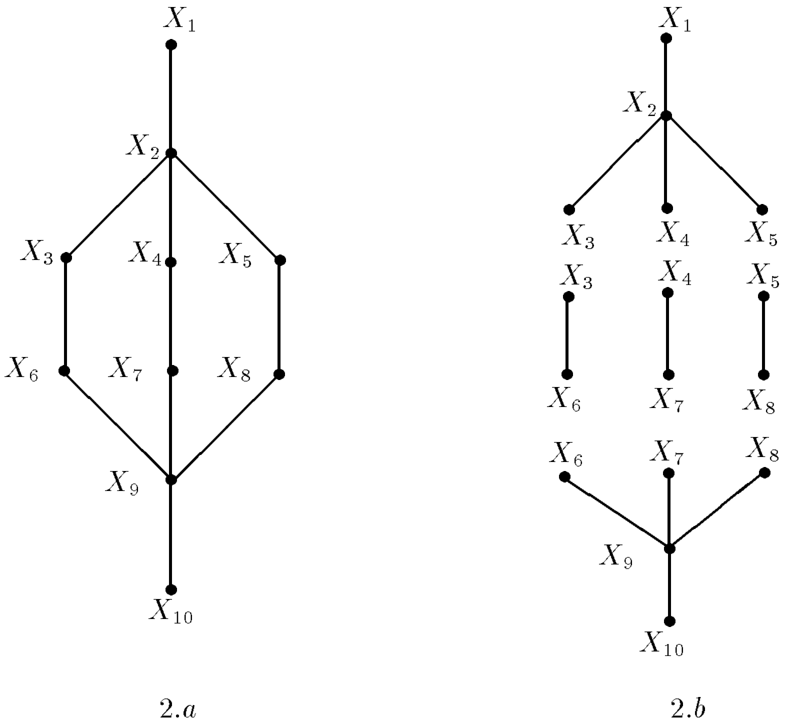


Figure 2

3.1 A NP-completeness Result

A close look to the proof of the upper bound in Theorem 3.1 shows that it can be applied also to any access structure $\mathcal{A}$ on $2k+2$ participants, $Y_0, X_0, X_1, \ldots, X_{2k}$, such that the set $\mathcal{A}$-*allowed* defined as

$$\mathcal{A}\text{-}allowed = \{Y_0X_0\} \bigcup \{X_0X_i, X_iX_{k+i} | 1 \leq i \leq k\}$$

is in the access structure, i.e., $\mathcal{A}$-*allowed* $\subseteq \mathcal{A}$, but the set $\mathcal{A}$-*forbidden* defined as

$$\mathcal{A}\text{-}forbidden = \{X_1X_2\ldots X_kY_0\} \bigcup \{Y_0X_{k+1}\} \bigcup \{X_1\ldots X_iY_0X_{k+i+1} | 1 \leq i \leq k-1\}$$

has no intersection with the access structure, i.e., $\mathcal{A}$-*forbidden* $\bigcap \mathcal{A} = \emptyset$. Let $\mathcal{B}_k$ be the set of all access structures which satisfy the above requirements. The sequence $(X_1, X_2, \ldots, X_k)$ is called the *children list* of access structure $\mathcal{A}$ (the name is inspired by the fact that the set $\mathcal{A}$-*allowed* has the form of a tree). To maintain simpler notation we denote a set $\{a_1, a_2, \ldots, a_n\}$ by the sequence $a_1a_2\ldots a_n$. In case the access structure is the closure of a graph, the set $\mathcal{A}$-*forbidden* can be written as

$$\mathcal{A}\text{-}forbidden\text{-}edges = \{Y_0X_i | 1 \leq i \leq 2k\} \bigcup \{X_iX_j | 1 \leq i < j \leq k\} \bigcup \{X_iX_{k+j} | 1 \leq i < j \leq k\}.$$

Let $\mathcal{A}$ be an access structure on a set $\mathcal{P}$ of participants. Given a subset of participants $\mathcal{P}' \subseteq \mathcal{P}$, we define the access structure *induced by* $\mathcal{P}'$ as the family of sets $\mathcal{A}[\mathcal{P}'] = \{x \in \mathcal{A} | x \subseteq \mathcal{P}'\}$. Extending Theorem 3.3 of [14] to general access structures and using Theorem 3.1 we can prove the following theorem.

Theorem 3.3 *Let $\mathcal{A}$ be an access structure on a set $\mathcal{P}$ of participants and $\mathcal{P}' \subseteq \mathcal{P}$. If $\mathcal{A}[\mathcal{P}'] \in \mathcal{B}_k$, where $k \geq 1$, then the optimal information rates for $\mathcal{A}$ and $\mathcal{A}[\mathcal{P}']$ satisfy*

$$\rho^*(\mathcal{A}) \leq \rho^*(\mathcal{A}[\mathcal{P}']) \leq \frac{1}{2} + \frac{1}{4k+2},$$

and optimal average information rate for $\mathcal{A}[\mathcal{P}']$ satisfies

$$\tilde{\rho}^*(\mathcal{A}[\mathcal{P}']) \leq \frac{2}{3} + \frac{2}{9k+6}.$$

Above theorem gives an upper bound on the information rate of access structures given that the access structure induced by a subset of participants is in $\mathcal{B}_k$. We will use above theorem to get upper bounds on the optimal information rate and on the optimal average information rate of several graphs with six vertices, extending the results of [9] that computed the information rate of all graphs with five vertices. Unfortunately, testing for above property in general is an hard computational problem, as we show that this is NP–complete. Let $\mathcal{A}$ be an access structure, a set $C \in \mathcal{A}$ is a *minimal* set of $\mathcal{A}$ if $A \notin \mathcal{A}$ whenever $A \subset C$. Define the $\mathcal{B}_k$–INDUCED–SUBSTRUCTURE problem as follows: Given a set of participants $\mathcal{P}$, an access structure $\mathcal{A}$ defined by the family of minimal sets which can recover the secret and a positive integer $k \geq 3$, determine if there is a subset $\mathcal{P}' \subseteq \mathcal{P}$ such that the induced access structure $\mathcal{A}[\mathcal{P}']$ is in $\mathcal{B}_k$.

Theorem 3.4 *$\mathcal{B}_k$–INDUCED–SUBSTRUCTURE is NP–complete.*

Proof. (For definition of NP–complete problems and notation used in this proof, we refer the reader to [23].) It is easy to see that $\mathcal{B}_k$–INDUCED–SUBSTRUCTURE $\in$ NP, since a nondeterministic algorithm needs only guess participants $Y_0, X_0, X_1, \ldots, X_{2k}$, and check in polynomial time whether the set $\mathcal{A}$-*allowed* is a subset of $\mathcal{A}$ and $\mathcal{A}$-*forbidden* $\bigcap \mathcal{A} = \emptyset$.

We transform 3SAT to $\mathcal{B}_k$–INDUCED–SUBSTRUCTURE. Let $U = \{u_1, u_2, \ldots, u_{k-1}\}$, $k \geq 3$, be a set of variables and $C = \{c_1, c_2, \ldots, c_m\}$ be a set of clauses, each containing 3 literals. We will construct an access structure $\mathcal{A}$ on a set $\mathcal{P}$ of participants, such that there is a subset of participants $\mathcal{P}' \subseteq \mathcal{P}$ and the induced access structure $\mathcal{A}[\mathcal{P}']$ is in $\mathcal{B}_k$ if and only if C is satisfiable.

There are $4k$ participants in $\mathcal{P}$: Four participants y_0, x_0, v, v', and for each variable $u_i \in U$ there are four participants $u_i, \overline{u}_i, u_i', \overline{u}_i'$ in $\mathcal{P}$.

The access structure $\mathcal{A}$ consists of three components, i.e., $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3$. The family $\mathcal{A}_1$ is defined as

$$\mathcal{A}_1 = \{y_0x_0, x_0v, vv'\} \bigcup \{x_0u_i, x_0\overline{u}_i, u_i\overline{u}_i', \overline{u}_i\, u_i', u_iu_i', \overline{u}_i\,\overline{u}_i' | 1 \leq i \leq k-1\}.$$

Note that the pairs of participants in $\mathcal{A}_1$ have been chosen so that if there is a set $\mathcal{P}' \subseteq \mathcal{P}$ such that $\mathcal{A}[\mathcal{P}'] \in \mathcal{B}_k$, then: 1) $y_0, x_0, v, v' \in \mathcal{P}'$; 2) for each pair $\{u_i, \overline{u}_i\}$, $i = 1, 2, \ldots, k-1$, exactly one element is in $\mathcal{P}'$; 3) for each pair $\{u_i', \overline{u}_i'\}$, $i = 1, 2, \ldots, k-1$, exactly one element is in $\mathcal{P}'$.
The set $\mathcal{A}_2$ is defined as

$$\mathcal{A}_2 = \{v'u_iu_{i-1}',\ v'u_i\,\overline{u}_{i-1}',\ v'\overline{u}_i\,u_{i-1}',\ v'\overline{u}_i\,\overline{u}_{i-1}' | 2 \leq i \leq k-1\}.$$

Note that the definition of the set $\mathcal{A}_2$ implies that if there is a set $\mathcal{P}' \subseteq \mathcal{P}$ such that $\mathcal{A}[\mathcal{P}'] \in \mathcal{B}_k$, then any children list $(w_0, w_1, w_2, \ldots, w_{k-1})$ of $\mathcal{A}[\mathcal{P}']$ satisfies $w_0 = v$ and $w_i \in \{u_i, \overline{u}_i\}$, for $i = 1, 2, \ldots, k-1$. Should it be otherwise, a set $A \in \mathcal{A}_2$ would belong to $\mathcal{A}[\mathcal{P}'] \cap \mathcal{A}$-*forbidden* and we could not have $\mathcal{A}[\mathcal{P}'] \in \mathcal{B}_k$, getting a contradiction.
The set $\mathcal{A}_3$ is defined as

$$\mathcal{A}_3 = \{l_{i,1}l_{i,2}l_{i,3} |\ l_{i,1}l_{i,2}l_{i,3} \text{ are the complements of the 3 litterals in } c_i \in C\}.$$

The construction can be accomplished in polynomial time. We now show that C is satisfiable if and only if there is a subset of participants $\mathcal{P}' \subseteq \mathcal{P}$ whose induced access structure $\mathcal{A}[\mathcal{P}']$ is in $\mathcal{B}_k$. Suppose $\mathcal{P}'$ is a set of participants such that $\mathcal{A}[\mathcal{P}'] \in \mathcal{B}_k$. Recalling the definition of $\mathcal{A}_1$, we have that $v \in \mathcal{P}'$ and for each pair $\{u_i, \overline{u}_i\}$, $i = 1, 2, \ldots, k-1$, exactly one element is in $\mathcal{P}'$. Consider the truth assignment $t : U \longrightarrow \{T, F\}$ defined as follows: If $u_i \in \mathcal{P}'$ then $t(u_i) = T$, else $t(u_i) = F$. Let $c_i \in C$ be a clause consisting of litterals $w_{i,1}, w_{i,2}, w_{i,3}$. Since $\overline{w}_{i,1}\,\overline{w}_{i,2}\,\overline{w}_{i,3}$ is in $\mathcal{A}_3$, then the three elements $\overline{w}_{i,1}, \overline{w}_{i,2}, \overline{w}_{i,3}$ cannot be all in $\mathcal{P}'$, otherwise $\mathcal{A}[\mathcal{P}'] \notin \mathcal{B}_k$ since $\{\overline{w}_{i,1}\,\overline{w}_{i,2}\,\overline{w}_{i,3}\} \in \mathcal{A}$-*forbidden*. If $\overline{w}_{i,j} \notin \mathcal{P}'$, for $j \in \{1, 2, 3\}$, then $t(w_{i,j}) = T$ and clause c_i is satisfiable.

On the other hand, assume that $t : U \longrightarrow \{T, F\}$ is a satisfying truth assignment for C. Define w_i and w_i' as follows: $w_i = u_i$ and $w_i' = u_i'$ if $t(u_i) = T$, and $w_i = \overline{u}_i$ and $w_i' = \overline{u}_i'$ otherwise. Let $\mathcal{P}'$ be the set $\{y_0, x_0, v, w_1, w_2, \ldots, w_{k-1}, v', w_1', w_2', \ldots, w_{k-1}'\}$. Then, $\mathcal{A}[\mathcal{P}'] \in \mathcal{B}_k$.

As an example, let $U = \{u_1, u_2, u_3\}$ and $C = \{\{u_1, u_2, u_3\}, \{u_1, u_2, \overline{u}_3\}, \{u_1, \overline{u}_2, \overline{u}_3\}, \{\overline{u}_1, u_2, u_3\}, \{\overline{u}_1, \overline{u}_2, \overline{u}_3\}\}$. The set of participants is $\{y_0, x_0, v, v', u_1, \overline{u}_1, u_1', \overline{u}_1', u_2, \overline{u}_2, u_2', \overline{u}_2', u_3, \overline{u}_3, u_3', \overline{u}_3'\}$. The graph representing the set $\mathcal{A}_1$ is depicted in Figure 3. Sets $\mathcal{A}_2$ and $\mathcal{A}_3$ are equal to $\mathcal{A}_2 = \{v'u_2u_1', v'u_2\overline{u}_1', v'\overline{u}_2u_1', v'\overline{u}_2\overline{u}_1', v'u_3u_2', v'u_3\overline{u}_2', v'\overline{u}_3u_2', v'\overline{u}_3\overline{u}_2'\}$ and $\mathcal{A}_3 = \{\overline{u}_1\,\overline{u}_2\,\overline{u}_3, \overline{u}_1\,\overline{u}_2u_3, \overline{u}_1u_2u_3, u_1\overline{u}_2\,\overline{u}_3, u_1u_2u_3\}$. There are three satisfying assignments for C: $u_1 = 0$, $u_2 = 1$, $u_3 = 0$; $u_1 = 1$, $u_2 = 0$, $u_3 = 1$; and $u_1 = 1$, $u_2 = 1$, $u_3 = 0$. The sets of participants $\mathcal{P}'$ such that $\mathcal{A}[\mathcal{P}'] \in \mathcal{B}_4$ are the following: $\{y_0, x_0, v, \overline{u}_1, u_2, \overline{u}_3, v', r_1, r_2, r_3\}$, $\{y_0, x_0, v, u_1, \overline{u}_2, u_3, v', r_1, r_2, r_3\}$, and $\{y_0, x_0, v, u_1, u_2, \overline{u}_3, v', r_1, r_2, r_3\}$, where each r_i can be either equal to u_i' or to $\overline{u}_i'$. $\square$

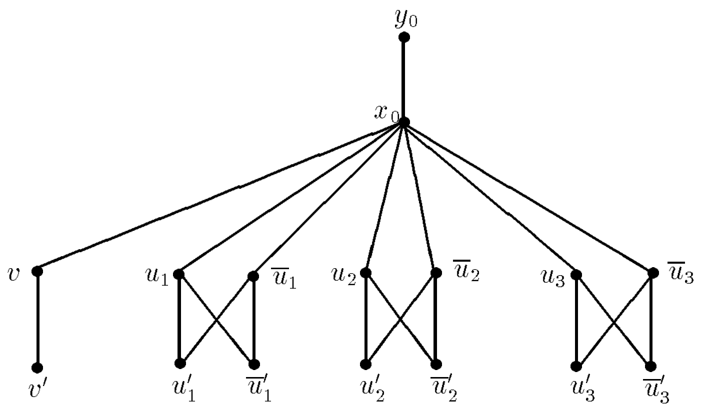


Figure 3

4 Lower Bounds on Information Rate and Average Information Rate

In this section we will give several general lower bounds on the information rate and on the average information rate of access structures represented by graphs. Our lower bounds are obtained, as customary, assuming an uniform probability distribution on the set of secrets. Let us denote the uniform probability distribution on the set of secrets S with U_s. It is an open problem to determine similar bounds when arbitrary probability distributions on the secrets are assumed. A few results in this directions are contained in [15]. We first recall the following theorem by Brickell and Davenport in [12] stating that a complete bibartite graph admits an ideal secret sharing scheme. Since we will use this result several times, we repeat the proof for the reader convenience[2].

Theorem 4.1 *Let G be a complete bipartite graph. Then, for any set of secrets S of size $q \geq 2$, there exists an ideal secret sharing scheme for G.*

Proof. Let V_1 and V_2 be the parts of G, an ideal secret sharing for G can be constructed as follows. Let $q \geq 2$ be an integer. Consider $S = \mathbf{Z}_q$. If the secret is $s \in S$, then the dealer randomly chooses an element $\alpha \in \mathbf{Z}_q$ and computes an element $\beta \in \mathbf{Z}_q$ such that $s = \alpha + \beta \bmod q$. The dealer gives the share α to all participants in V_1 and the share β to all participants in V_2. It is obvious that this realizes a secret sharing scheme with information rate equal to 1. □

We first improve on the bound of Theorem 2.1 for graphs with n vertices and odd maximum degree d.

Theorem 4.2 *Let $G = (V(G), E(G))$ be a graph of n vertices and maximum degree d, d odd. Then for any set of secrets S of size q^n, with $q \geq 2$, there exists a secret sharing scheme Σ with information rate*

$$\rho(G, U_s, \Sigma) = \frac{1}{\lceil d/2 \rceil + 1 - \lceil d/2 \rceil / n}.$$

Proof. Let $Adj(X)$, $Inc(X)$, $degree\_one(X)$ be the following sets : $Adj(X) = \{Y : (X,Y) \in E(G)\}$ is the set of vertices adjacent to X. $Inc(X) = \{(X,Y) : (X,Y) \in E(G)\}$ is the set of edges incident

[2]Actually, Brickell and Davenport proved the theorem for the general case of complete multipartite graphs, but we use it only in the particular case of complete bipartite graphs.

to X. Finally, $degree\_one(X) = \{Y \in Adj(X) : |Inc(Y)| = 1\}$ is the set of vertices adjacent to X with degree 1. We will prove the theorem in the case $|S| = 2^n$, the construction can be easyly extended to the general case $|S| = q^n$ and $q \geq 2$.
For a vertex $X \in V(G)$ define G_x as the subgraph of G such that $V(G_x) = \{X\} \bigcup Adj(X)$ and $E(G_x) = Inc(X)$. The graph G_x is a complete multipartite graph and by Theorem 4.1 there is a secret sharing scheme for G_x with information rate 1. Let G' be the graph with vertices $V(G') = V(G) - (\{X\} \bigcup degree\_one(X))$ and edge-set $E(G') = E(G) - Inc(X)$. Assume that the secret consists of a single bit. If we use the secret sharing scheme described in Theorem 3.8 of [14] for G', then each vertex in $Adj(X) \bigcap V(G')$ gets at most $\lceil (d-1)/2 \rceil + 1$ bits while all other vertices get at most $\lceil d/2 \rceil + 1$ bits. We realize a secret sharing scheme for G by using both the scheme for G_x and the scheme for G'. In the resulting scheme the vertex X receives only one bit, the vertices in $Adj(X) \bigcap V(G')$ receive at most $\lceil (d-1)/2 \rceil + 2$ bits, while all remaining vertices get at most $\lceil d/2 \rceil + 1$ bits. Since d is odd then $\lceil (d-1)/2 \rceil + 2 = \lceil d/2 \rceil + 1$. Therefore, the above described secret sharing schemes for G gives to any predeterminated vertex only one bit, while all other vertices in G get at most $\lceil d/2 \rceil + 1$ bits. Now, assume that the secret consists of n bits. Consider the scheme that for each bit of the secret distributes it by choosing as a predetermined vertex X each vertex of G in turn. The resulting secret sharing scheme, for a secret of n bits, gives to each vertex at most $1 + (n-1)(\lceil d/2 \rceil + 1)$ bits. The information rate of the scheme is equal to

$$\frac{1}{\lceil d/2 \rceil + 1 - \lceil d/2 \rceil / n},$$

and the theorem follows. □

For a graph G of maximum degree 3, the bound of Theorem 2.1 gives $\rho(G, U_s, \Sigma) = 1/3$ while the bound of Theorem 4.2 gives $\rho(G, U_s, \Sigma) = 1/(3 - 2/n)$. The following theorem gives an improved bound.

Theorem 4.3 *Let $G = (V(G), E(G))$ be a graph of maximum degree 3 with n vertices. Then, for any set of secrets S of size q^2, with $q \geq n$, there exists a secret sharing scheme Σ with information rate $\rho(G, U_s, \Sigma) = 2/5$.*

Proof. Consider a partition of the edge set $E(G)$ in cycles $C_1, \ldots, C_r$ and trees $T_1, \ldots, T_m$. Such a partition exists for any graph G. Indeed, removing all the cycles from the graph we are left with a forest of connected acyclic graphs. From Theorem 2.2 we know that, for any cycle of lenght $n \geq 5$ and for any set of secrets S of size q^2, with $q \geq n$, there exits a secret sharing scheme with information rate equal to 2/3. For a secret of $2 \log q$ bits, the scheme gives only $3 \log q$ bits to all vertices of the cycle. If a cycle has lenght four then from Theorem 4.1 there exists an ideal secret sharing scheme for any set of secrets S of size ≥ 2; whereas if a cycle has lenght three, then from the main theorem of [12] there exists an ideal secret sharing scheme for any set of secrets S of size ≥ 3. From Theorem 2.1 we know that, for any set of secrets S of size ≥ 2, there is a secret sharing scheme for any tree with information rate equal to 1/2. For a secret of $2 \log q$ bits, the scheme given in [9] distributes only $2 \log q$ bits to the leaves of the tree while all other vertices get $4 \log q$ bits. We now realize a secret sharing scheme for G, by sharing a secret consisting of $2 \log q$ separately in each tree $T_1, \ldots, T_m$ and cycle $C_1, \ldots, C_r$. A vertex of G of degree one can only be a leaf of a three so it receives $2 \log q$ bits. If a vertex has degree two then either it belongs to a cycle, receiving $3 \log q$ bits, or it is an internal node of a three and it receives $4 \log q$ bits. If a vertex has degree three then it belongs to a cycle and it is the leave of a tree, receiving $5 \log q$ bits in total. Any vertex of the graph cannot be an internal vertex of a tree and belong to a cycle, would it be otherwise it should have degree four contradicting the hypothesis. Thus, we can construct a secret sharing scheme for G, giving to each vertex a share of at most $5 \log q$ bits for a secret of $2 \log q$ bits. This scheme has information rate $2/5$. $\square$

If the number of vertices in the graph G is known, then we can improve on the bound provided by Theorem 4.3 by employing the same technique used in Theorem 4.2. This gives the bound $\rho(G, U_s, \Sigma) = 2/(5 - c/n)$ for a constant $c > 0$.

Applying the same reasoning of Theorem 4.3 to graphs of odd degree d leads to the bound $\rho^*(G) \geq 1/(1.5\lfloor d/2 \rfloor + 1)$ which is worse than previous bounds.

Regardless of the degree, it is possible to obtain better bounds for trees. We recall that an internal node is a vertex of degree greater than one.

Theorem 4.4 *Let G be a tree with n internal vertices. Then for any set of secrets S of size q^n, with $q \geq 2$, there exists a secret sharing scheme Σ with information rate*

$$\rho(G, U_s, \Sigma) = \frac{n}{2n-1}.$$

Proof. We will prove the theorem in the case $|S| = 2^n$, the construction can be easyly extended to the general case $|S| = q^n$ and $q \geq 2$. In [9] it was showed how to obtain a secret sharing scheme for any tree with information rate equal to $1/2$. This scheme, for a secret consisting of a single bit, gives one bit to a predeterminated vertex $X \in V(G)$ and to all non-internal vertices, whereas each other vertex gets two bits. Assume that the secret consists of n bits. Consider the scheme that for each bit of the secret distributes it by choosing as a predetermined vertex X each vertex of G in turn. This scheme, for a secret of n bits, gives to each vertex at most $2(n-1) + 1 = 2n - 1$ bits. Thus

$$\rho(G, U_s, \Sigma) = \frac{n}{2n-1}.$$

$\square$

If only the number of vertices is known, what can we say on the information rate of a graph G? The maximum degree of G can be as bad as $n - 1$. Thus, the bound of [14] gives $\rho^*(G) \geq 1/(\lceil (n-1)/2 \rceil + 1)$, while the bound of Theorem 4.2 gives $\rho^*(G) \geq 1/(\lceil (n-1)/2 \rceil + 1 - \lceil (n-1)/2 \rceil / n)$, if n is even.

In this last part of the paper we present general lower bounds on the optimal information rate and optimal average information rate for *any* graph G with n vertices. The lower bounds are obtained by using known results on the covering of the edges of a graphs by means of complete bipartite graphs.

Tuza [44] proved that the edge-set of an arbitrary graph G can be covered by complete bipartite subgraphs $G_1(V(G_1), E(G_1)), \ldots, G_T(V(G_T), E(G_T))$ such that $\sum_{i=1}^{T} |V(G_i)| \leq 3n^2/2 \log n + o(n^2/\log n)$. We now use again Theorem 4.1, namely that there exists a secret sharing scheme for each G_i with information rate equal to 1. We can construct a secret sharing scheme for G by sharing the secret separetely in each G_i. In this way we need to generate a total of $3n^2/2 \log n + o(n^2/\log n)$ shares, each of them of the same size as the secret. Thus, we get that the average size of a share given to any participant is less than $3n/2 \log n + o(n/\log n)$. Therefore, we get that the optimal average information rate for any graph G with n vertices is greater than n times the inverse of $3n^2/2 \log n + f(n)$, where $|f(n)| < \epsilon n^2/\log n$, for all $\epsilon > 0$ and sufficiently large n. Thus, the average information rate is greater than $2\log n/3n + g(n)$, where $|g(n)| \leq (2\epsilon/3(\epsilon+3/2))\log n/n$, if $|f(n)| < \epsilon n^2/\log n$.

Feder and Motwani [21] proved that the problem of partitioning the edges of a graph G into complete bipartite graphs such that the sum of the cardinalities of their vertex sets is minimized is NP–complete. However, they proved that the edge set of a graph $G = (V, E)$, with $|V| = n$ and $|E| = m$ can be partitioned into complete bipartite graphs with sum of the cardinalities of their vertex sets $O(\frac{m \log \frac{n^2}{m}}{\log n})$, and presented an efficient algorithm to compute such a partition. Using their result and again sharing the secret in each complete bipartite graph with Brickell and Davenport's algorithm, it follows that there is a secret sharing scheme with average information rate at least $\Omega(\frac{n \log n}{m \log \frac{n^2}{m}})$.

Finally, we recall a result of Erdös and Pyber [20] (see also [34]) which states that edges of a graph G with n vertices can be partitioned into complete bipartite graphs such that each vertex of G is contained by at most $O(n/\log n)$ complete bipartite graphs. This result, together with Theorem 4.1, directly implies that the optimal information rate of G is $\rho^*(G) = \Omega\left(\frac{\log n}{n}\right)$.

These results can be summarized in the following theorem.

Theorem 4.5 *Let G be a graph with n vertices and m edges. Then, for any set of secrets S of size $q \geq 2$ there exist secret sharing schemes Σ_1 and Σ_2 with average information rate*

$$\widetilde{\rho}(G, U_s, \Sigma_1) > \frac{2\log n}{3n} + o\left(\frac{\log n}{n}\right),$$

and

$$\widetilde{\rho}(G, U_s, \Sigma_2) = \Omega\left(\frac{n \log n}{m \log \frac{n^2}{m}}\right),$$

respectively. Moreover, there exists a secret sharing scheme Σ_3 with information rate

$$\rho(G, U_s, \Sigma_3) = \Omega\left(\frac{\log n}{n}\right).$$

Acknowledgments

We are indebted to professor Capocelli for his constant encouragement and support. We would like to dedicate this paper to his memory as a sign of appreciation and love.

We would like to thank L. Pyber for providing us reference [34] and A. Marchetti–Spaccamela and E. Feuerstein for bringing to our attention reference [21]. Finally, we would like to thank the anonymous referees for their useful comments and suggestions that made the paper more readable.

References

[1] A. Beutelspacher, *How to Say 'No'* in "Advances in Cryptology - EUROCRYPT 89", "Lecture Notes in Computer Science", Vol. 434, Springer-Verlag, Berlin, pp. 491–496, 1990.

[2] J. C. Benaloh and J. Leichter, *Generalized Secret Sharing and Monotone Functions,* in "Advances in Cryptology - CRYPTO 88", S. Goldwasser Ed., "Lecture Notes in Computer Science", Vol. 403, Springer-Verlag, Berlin, pp. 27–35, 1985.

[3] G. R. Blakley, *Safeguarding Cryptographic Keys,* Proceedings AFIPS 1979 National Computer Conference, pp.313–317, June 1979.

[4] B. Blakley, G. R. Blakley, A. H. Chan, and J. L. Massey, *Threshold Schemes with Disenrollment,* in "Advances in Cryptology - CRYPTO '92", "Lecture Notes in Computer Science", Vol. 740, E. Brickell Ed., Springer-Verlag, Berlin, pp. 546–554, 1993.

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

[6] C. Blundo, A. Cresti, A. De Santis, and U. Vaccaro, *Fully dynamic secret sharing schemes*, in "Advances in Cryptology - CRYPTO 93", D.R. Stinson Ed., "Lecture Notes in Computer Science", Vol. 773, Springer-Verlag, Berlin, pp. 126–135, 1994.

[7] C. Blundo, A. De Santis, G. Di Crescenzo, A. Giorgio Gaggia, and U. Vaccaro, *Multi-Secret Sharing Schemes*, to appear in "Advances in Cryptology – CRYPTO 94", Y. Desmedt Ed., Lecture Notes in Computer Science, Springer-Verlag, Berlin.

[8] C. Blundo, A. De Santis, L. Gargano, and U. Vaccaro, *Secret Sharing Schemes with Veto Capabilities*, in "Proceedings of the French-Israeli Workshop in Algebraic Coding", "Lecture Notes in Computer Science", Vol. 781, Springer-Verlag, Berlin, pp. 82–89, 1994.

[9] C. Blundo, A. De Santis, D. R. Stinson, and U. Vaccaro, *Graph Decomposition and Secret Sharing Schemes*, in "Advances in Cryptology – EUROCRYPT '92", R. Rueppel Ed., Lecture Notes in Computer Science, Vol. 658, Springer-Verlag, Berlin, pp. 1–24, 1993. To appear in Journal of Cryptology.

[10] C. Blundo, A. De Santis, and U. Vaccaro, *Efficient Sharing of Many Secrets*, "Proceedings of STACS '93 (10*th* Symp. on Theoretical Aspects of Computer Science)", P. Enjalbert, A. Finkel, K. W. Wagner Eds., "Lecture Notes in Computer Science", Vol. 665, Springer–Verlag, Berlin, pp. 692–703, 1993.

[11] C. Blundo, A. De Santis, and U. Vaccaro, *Randomness in Distribution Protocols*, to appear in "21st International Colloquium on Automata, Languages and Programming" (ICALP '94), Serge Abiteboul and Eli Shamir Eds., "Lecture Notes in Computer Science".

[12] E. F. Brickell and D. M. Davenport, *On the Classification of Ideal Secret Sharing Schemes*, J. Cryptology, 4:123–134, 1991.

[13] E. F. Brickell and D. R. Stinson, *The Detection of Cheaters in Threshold Schemes*, SIAM J. on Discrete Math. Vol. 4, pp. 502–510, 1991.

[14] E. F. Brickell and D. R. Stinson, *Some Improved Bounds on the Information Rate of Perfect Secret Sharing Schemes,* J. Cryptology, 5:153–166, 1992.

[15] R. M. Capocelli, A. De Santis, L. Gargano, and U. Vaccaro, *On the Size of Shares for Secret Sharing Schemes,* J. Cryptology, 6:157–168, 1993.

[16] M. Carpentieri, A. De Santis, e U. Vaccaro, *Size of Shares and Probability of Cheating in Threshold Schemes*, in "Proceedings of EUROCRYPT '93, Advances in Cryptology", T. Helleseth Ed., Lecture Notes in Computer Science, Vol. 765, Springer-Verlag, Berlin, pp. 118–125, 1994.

[17] L. Csirmaz, *The Size of a Share Must be Large*, to appear in "Advances in Cryptology – EUROCRYPT '94", A. De Santis Ed., "Lecture Notes in Computer Science", Springer-Verlag, Berlin.

[18] I. Csiszár and J. Körner, *Information Theory. Coding Theorems for Discrete Memoryless Systems,* Academic Press, 1981.

[19] M. van Dijk, *On the Information Rate of Perfect Secret Sharing Schemes*, Preprint, 1994.

[20] P. Erdös and L. Pyber, unpublished.

[21] T. Feder and R. Motwani, *Clique Partition, Graph Compression and Speeding-up Algorithms,* Proceedings of the $23rd$ Annual ACM Symposium on Theory of Computing, New Orleans, pp. 123–133, 1991.

[22] R. G. Gallager, *Information Theory and Reliable Communications*, John Wiley & Sons, New York, NY, 1968.

[23] M. Garey and D. Johnson, *Computers and Intractability: a Guide to the Theory of NP-Completeness,* W. H. Freeman & Co., New York, 1979.

[24] I. Ingemarsson and G. J. Simmons, *A Protocol to Set up Shared Secret Schemes Without the Assistance of a Mutually Trusted Party,* in “Advances in Cryptology - CRYPTO 90”, “Lecture Notes in Computer Science”, Vol. 473, Springer-Verlag, Berlin, pp. 266–282, 1991.

[25] M. Ito, A. Saito, and T. Nishizeki, *Secret Sharing Scheme Realizing General Access Structure,* Proc. IEEE Global Telecommunications Conf., Globecom 87, Tokyo, Japan, 1987.

[26] W.-A. Jackson, K. M. Martin, and C. M. O'Keefe, *Multisecret Threshold Schemes*, in “Advances in Cryptology - CRYPTO '93”, D.R. Stinson Ed., “Lecture Notes in Computer Science”, Vol. **773**, Springer-Verlag, Berlin, pp. 126–135, 1994.

[27] E. D. Karnin, J. W. Greene, and M. E. Hellman, *On Secret Sharing Systems,* IEEE Trans. on Inform. Theory, vol. IT-29, no. 1, pp. 35–41, Jan. 1983.

[28] D.E. Knuth and A.C. Yao, *The Complexity of Nonuniform Random Number Generation*, in “Algorithms and Complexity”, J.F. Traub Ed., Academic Press, pp. 357–428, 1976.

[29] S. C. Kothari, *Generalized Linear Threshold Schemes*, in “Advances in Cryptology - CRYPTO 84”, G. R. Blakley and D. Chaum Eds., “Lecture Notes in Computer Science”, Vol 196, Springer–Verlag, Berlin, pp. 231–241, 1985.

[30] R. J. McEliece and D. V. Sarwate, *On Sharing Secrets and Reed-Solomon Codes*, Commun. of the ACM, Vol. 24, 583–584, 1981.

[31] K. M. Martin, *Discrete Structures in the Theory of Secret Sharing*, PhD Thesis, University of London, 1991.

[32] K. M. Martin, *New Secret Sharing Schemes from Old*, Journal of Combin. Math. and Combin. Comput., 14:65–77, 1993.

[33] M. Naor and A. Shamir, *Visual Cryptography*, to appear in “Advances in Cryptology – Eurocrypt '94”, A. De Santis Ed., Lecture Notes in Computer Science, Springer-Verlag, Berlin.

[34] L. Pyber, *Covering the Edges of a Graph by ...*, in Sets, Graphs and Numbers, Colloquia Mathematica Soc. János Bolyai, L. Lovász, D. Miklós, T. Szönyi, Eds., Nort-Holland, pp. 583–610, 1992.

[35] T. Rabin and M. Ben-Or, *Verifiable Secret Sharing and Multiparty Protocols with Honest Majority,* Proc. 21st ACM Symp. on Theory of Computing, pp. 73–85, 1989

[36] A. Shamir, *How to Share a Secret,* Communications of the ACM, vol. 22, n. 11, pp. 612–613, Nov. 1979.

[37] 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'*, Congressus Numer., Vol. 68, pp. 215–248, 1989.

[38] G. J. Simmons, *Prepositioned Shared Secret and/or Shared Control Schemes*, in “Advances in Cryptology – CRYPTO '89”, “Lecture Notes in Computer Science”, Vol. 434, Springer-Verlag, Berlin, pp. 436–467, 1990.

[39] G. J. Simmons, *An Introduction to Shared Secret and/or Shared Control Schemes and Their Application*, Contemporary Cryptology, IEEE Press, pp. 441–497, 1991.

[40] D. R. Stinson, *An Explication of Secret Sharing Schemes*, Designs, Codes and Cryptography, 2:357–390, 1992.

[41] D. R. Stinson, *New General Lower Bounds on the Information Rate of Secret Sharing Schemes*, in "Advances in Cryptology – CRYPTO '92", E. Brickell, Ed., "Lecture Notes in Computer Science", Vol. 740, Springer-Verlag, Berlin, pp. 170–184, 1993.

[42] D. R. Stinson, *Decomposition Constructions for Secret Sharing Schemes*, IEEE Trans. Inform. Theory, 40:118–125, 1994.

[43] M. Tompa and H. Woll, *How to Share a Secret with Cheaters*, J. Cryptology 1:133–138, 1988.

[44] Z. Tuza, *Covering of Graphs by Complete Bipartite Subgraphs; Complexity of 0-1 matrices,* Combinatorica, vol. 4, n. 1, pp. 111–116, 1984.

Appendix A

In this appendix we review the basic concepts of Information Theory we will use. For a complete treatment of the subject the reader is advised to consult [18] and [22]. We will also recall some basic terminology from graph theory.

Given a probability distribution $\{p(x)\}_{x \epsilon X}$ on a set X, we define the *entropy* of X, $H(X)$, as

$$H(X) = -\sum_{x \epsilon X} p(x) \log p(x)^3.$$

The entropy $H(X)$ is a measure of the average uncertainty one has about which element of the set X has been chosen when the choices of the elements from X are made according to the probability distribution $\{p(x)\}_{x \epsilon X}$. The entropy satisfies the following property

$$0 \leq H(X) \leq \log |X|, \tag{2}$$

where $H(X) = 0$ if and only if there exists $x_0 \in X$ such that $p(x_0) = 1$; $H(X) = \log |X|$ if and only if $p(x) = 1/|X|$, for all $x \in X$.

Given two sets X and Y and a joint probability distribution $\{p(x, y)\}_{x \epsilon X, y \epsilon Y}$ on their Cartesian product, the *conditional entropy* $H(X|Y)$ is defined as

$$H(X|Y) = -\sum_{y \epsilon Y} \sum_{x \epsilon X} p(y) p(x|y) \log p(x|y).$$

From the definition of conditional entropy it is easy to see that

$$H(X|Y) \geq 0. \tag{3}$$

If we have $n+1$ sets $X_1, \ldots, X_n, Y$, the entropy of $X_1 \ldots X_n$ given Y can be expressed as

$$H(X_1 \ldots X_n|Y) = H(X_1|Y) + H(X_2|X_1 Y) + \cdots + H(X_n|X_1 \ldots X_{n-1} Y) \tag{4}$$

The *mutual information* $I(X;Y)$ between X and Y is defined by $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$, since it is always non negative one gets

$$H(X) \geq H(X|Y). \tag{5}$$

Given $n+2$ sets $X, Y, Z_1, \ldots, Z_n$ and a joint probability distribution on their Cartesian product, the *conditional mutual information* $I(X;Y|Z_1, \ldots, Z_n)$ between X and Y given $Z_1, \ldots, Z_n$ can be written as

$$I(X;Y|Z_1, \ldots, Z_n) \quad = \quad H(X|Z_1, \ldots, Z_n) - H(X|Z_1, \ldots, Z_n Y).$$

Since the conditional mutual information is always non negative we get

$$H(X|Z_1, \ldots, Z_n) \geq H(X|Z_1, \ldots, Z_n Y). \tag{6}$$

We now present some basic terminology from graph theory. A graph, $G = (V(G), E(G))$ consists of a finite non empty set of vertices $V(G)$ and a set of edges $E(G) \subseteq V(G) \times V(G)$. Graphs do not have loops or multiple edges. We consider only undirected graphs. In an undirected graph the pair of vertices representing any edge is unordered. Thus, the pairs (X, Y) and (Y, X) represent

[3] All logarithms in this paper are of base 2

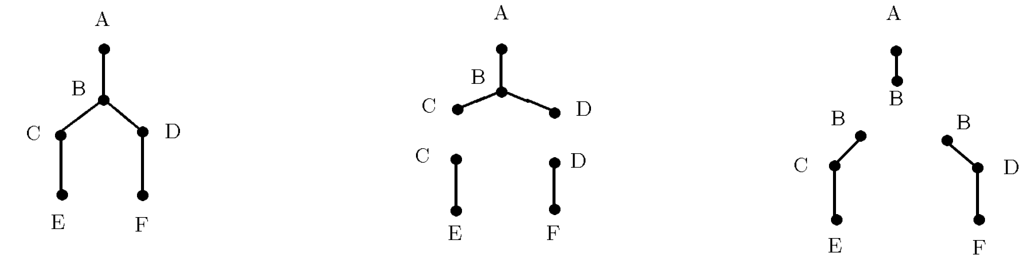
the same edge. To avoid overburdening the notation we often describe a graph G by the list of all edges $E(G)$. We will use reciprocally (X, Y) and XY to denote the edge joining the vertices X and Y. G is *connected* if any two vertices are joined by a path. The *complete graph* K_n is the graph on n vertices in which any two vertices are joined by an edge. The *complete multipartite graph* $K_{n_1,n_2,\dots,n_t}$ 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 XY is an edge if and only if X and Y are in different parts. If G is a graph, then G_1 is said to be a *subgraph* of G if $V(G_1) \subseteq V(G)$ and $E(G_1) \subseteq E(G)$.

Suppose G is a graph and $G_1, \dots, G_t$ are subgraphs of G, such that each edge of G occurs in at least one of the G_i's. We say that $\Pi = \{G_1, \dots, G_t\}$ is a covering of G and if each G_i, $i = 1, \dots, t$ is a complete multipartite graph then we say that Π is a *complete multipartite covering* (CMC) of G.

Appendix B

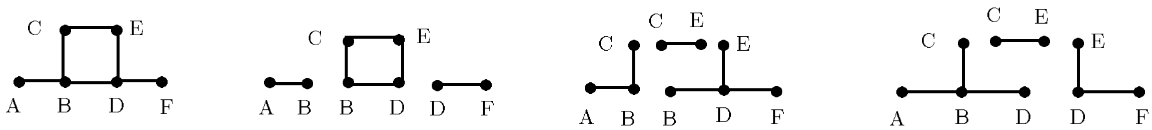
In this appendix we analyze all graphs who have optimal information rate less than 2/3 accordingly to Theorem 3.3. The schemes for these graphs are obtained by using the Multiple Construction Technique [9] based on complete multipartite coverings of the graph. The optimal information rate is not greater than 3/5 and the optimal average information rate is less than or equal to 3/4 for all graphs from Theorem 3.3. All these results are summarized in Table 1, and the first CMC of each graph gives the scheme with average information rate showed in Table 1. Below are depicted some of the minimal CMCs for 5 graphs on 6 vertices.

G_1



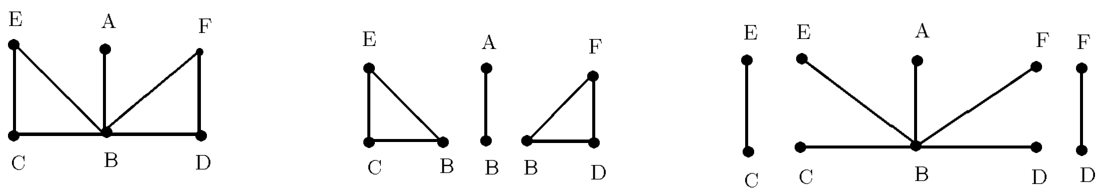
Π_1 Π_2

G_2



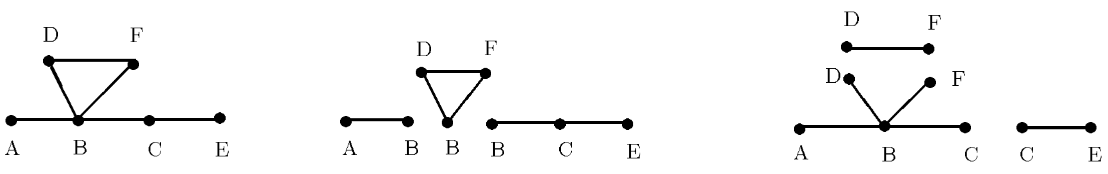
Π_1 Π_2 Π_3

G_3



Π_1 Π_2

G_4



Π_1 Π_2

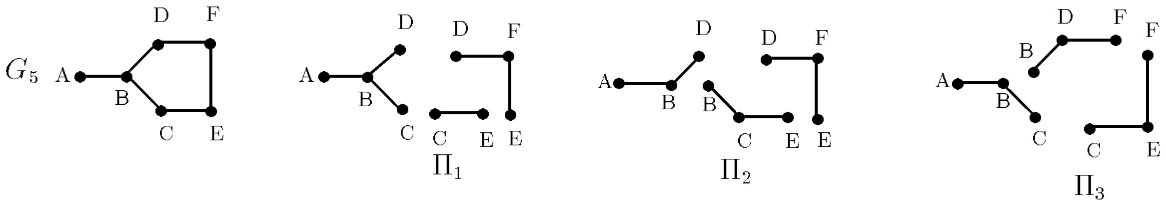


Table 1: Information Rate and Average Information Rate

Graph	Information Rate	Average information Rate
G_1, G_2, G_3, G_4	$\rho^* = 3/5$	$\widetilde{\rho}^* = 3/4$
G_5	$\rho^* = 3/5$	$2/3 \leq \widetilde{\rho}^* \leq 3/4$