Secret sharing on infinite graphs

László Csirmaz*
Central European University

Abstract

We extend the notion of perfect secret sharing scheme for access structures with infinitely many participants. In particular we investigate cases when the participants are the vertices of an (infinite) graph, and the minimal qualified sets are the edges. The (worst case) *information ratio* of an access structure is the largest lower bound on the amount of information some participant must remember for each bit in the secret – just the inverse of the information rate. We determine this value for several infinite graphs: infinite path, two-dimensional square and honeycomb lattices; and give upper and lower bounds on the ratio for the triangular lattice.

It is also shown that the information ratio is not necessarily *local*, i.e. all finite spanned subgraphs have strictly smaller ratio than the whole graph. We conclude the paper by posing several open problems.

Key words. Secret sharing scheme, information theory, infinite graph, lattice.

1 Introduction

Secret sharing scheme is a method of distributing a secret data among a set of participants so that only qualified subsets are able to recover the secret. If, in addition, unqualified subsets have no extra information, i.e. their joint shares is statistically independent of the secret, then the scheme is called *perfect*. The description of the qualified subsets among all possible subsets of participants is the *access structure*. In this paper only perfect secret sharing schemes are considered; when we speak about a secret sharing scheme, it is assumed to be perfect.

The most frequently investigated property is the efficiency of the system: how many bits of information the participants must remember for each bit of the secret in the worst case. This amount is the (worst case) *information ratio* of the system, which is just the inverse of the most commonly used information rate. (The name comes from the analogy to noisy channels.) Next to the the worst case, the average is also a good measure of the efficiency; in this paper by *information ratio* we mean the worst case one.

Determining the (worst case or average) information ratio, even in special cases, turned out to be extremely hard, both theoretically and technically. Several particular classes of access structures were investigated separately: access structures on four or five element sets [15, 19], access structures with three or four minimal sets [16], but most importantly access structures where the minimal qualified subsets are of size two — the so-called graph access structures [2, 3, 4, 7, 11, 12, 20].

Here we start the research in a new direction: we consider access structures with infinitely many participants. In the exposition we restrict ourselves to graph access structures. Our definitions, and some of the results, generalize easily for arbitrary access structures. [5] investigated secret sharing systems on infinite domain with finite access structures.

We determine the exact value of the (worst case) information ratio for a couple of infinite graphs. This part uses a new result on the exact information ratio for a particular family of graphs (Theorem 4.2). We also show that the information ratio is not necessarily *local*, i.e. there are cases when this number for the whole graph is larger than that of any of its finite subgraph (Corollary 4.6).

*The author can be reached at `csirmaz AT renyi DOT hu`

The paper is organized as follows. In the next session we recall the definition of a perfect secret sharing system, define the worst case and average information ratio, and introduce the so-called entropy method [3, 4, 6, 19]. Section 3 defines secret sharing systems on infinite structures. We introduce a generalizations of the decomposition technique for the infinite case (Theorem 3.6), and show that Stinson's celebrated bound works in the infinite case as well (Corollary 3.9). In section 4 we determine the exact information ratio of a particular family of graphs. It is used to prove the optimality of several constructions in section 5. Finally section 6 concludes the paper, and list some problems. For undefined notions and for an introduction to secret sharing schemes see [2] or [7]; for those in information theory consult [10].

All logarithms in this paper are of base 2.

2 Definitions

This section defines some of the most important notions which are used in the paper. First we recall some graph properties, then give a formal definition of a (finite) perfect secret sharing scheme based on graphs. Finally we connect perfect secret sharing schemes to certain submodular functions.

Let $G = \langle V, E\rangle$ be a (finite or infinite) graph with vertex set V and edge set E. A subset A of V is *independent* if there is no edge between vertices in A. A *covering* of the graph G is a collection of subgraphs of G such that every edge is contained on one of the (not necessarily spanned) subgraphs in the collection. The collection is k-*covering* if every edge of G is covered exactly k times. For subsets of vertices we usually omit the $\cup$ sign, and write AB for $A \cup B$. Also, it $v \in V$ is a vertex then Av denotes $A \cup \{v\}$.

A *perfect secret sharing scheme* $\mathcal{S}$ for a finite graph G is a collection of random variables ξ_v for each $v \in V$ and a ξ_s (the secret) with a joint distribution so that

(i) if vw is an edge in G, then ξ_v and ξ_w together determine the value of ξ_s;

(ii) if A is an independent set, then ξ_s and the collection $\{\xi_v : v \in A\}$ are statistically independent.

The *size* of the discrete random variable ξ is measured by its entropy, or information content, and is denoted by $\mathbf{H}(\xi)$, see [10]. This amount has to be well defined and finite, consequently *all random variables in this paper are assumed to be finite*, i.e. they can take only finitely many different values with positive probability. This is the main obstacle one has to overcome when defining a secret scheme on infinite domain.

The *information ratio* for a vertex (or participant) $v \in G$ is $\mathbf{H}(\xi_v)/\mathbf{H}(\xi_s)$. This value tells how many bits of information v must remember for each bit in the secret. The worst case (or average) *information ratio* of $\mathcal{S}$ is the highest (resp. average) information ratio among all participants.

Given a graph G its information ratio is the infimum of the corresponding value for all perfect secret sharing schemes $\mathcal{S}$ defined on G.

Definition 2.1 The *information ratio* of the (finite) graph G, denoted as $R(G)$, is defined as

$$R(G) = \inf_{\mathcal{S}} \max_{v\in V} \frac{\mathbf{H}(\xi_v)}{\mathbf{H}(\xi)}. \quad \blacksquare$$

The widely used *information rate* is the inverse of this value. While the "information rate" is the customary measure in the literature, we found its inverse, the ratio, to be more intuitive, furthermore certain expressions are easier to write and understand using the ratio.

As it has been pointed out in [1], it is not evident that the "infimum" in Definition 2.1 should actually be taken by some scheme $\mathcal{S}$, i.e. whether the infimum is always a minimum. [1] presents a general access structure where the infimum is not taken by any scheme. For access structures based on graphs the question whether inf = min is an open problem.

Let $\mathcal{S}$ be a perfect secret sharing scheme based on the (finite) graph G with the random variable acting ξ_s as secret, and ξ_v for $v \in V$ acting as shares. For each subset A of the vertices one can define the real-valued function f as

$$f(A) \stackrel{\text{def}}{=} \frac{\mathbf{H}(\{\xi_v : v \in A\})}{\mathbf{H}(\xi)}. \tag{1}$$

Clearly, the information ratio of $\mathcal{S}$ is the maximal value in $\{f(v) : v \in V\}$, while the average information ratio is the average of these values. Using standard properties of the entropy function, cf. [10], following inequalities hold for all subsets A, B of the participants:

(a) $f(\emptyset) = 0$, and in general $f(A) \geq 0$ (positivity);

(b) if $A \subseteq B \subseteq V$ then $f(A) \leq f(B)$ (monotonicity);

(c) $f(A) + f(B) \geq f(A \cap B) + f(A \cup B)$ (submodularity).

It is well known that for two (finite) random variables η and ξ, the value of η determines the value of ξ iff $\mathbf{H}(\eta\xi) = \mathbf{H}(\eta)$, moreover η and ξ are (statistically) independent iff $\mathbf{H}(\eta\xi) = \mathbf{H}(\eta) + \mathbf{H}(\xi)$. Using these facts and the definition of the perfect secret sharing scheme, we also have

(d) if $A \subseteq B$, A is an independent set and B is not, then $f(A)+1 \leq f(B)$ (strong monotonicity);

(e) if neither A nor B is independent but $A \cap B$ is so, then $f(A)+f(B) \geq 1+f(A\cap B)+f(A\cup B)$ (strong submodularity).

The celebrated *entropy method*, see, e.g., [2], can be rephrased as follows. Prove that for *any* real-valued function f satisfying properties (a)–(e), for some vertex $v \in G$, $f(v) \geq r$. Then, as functions coming from secret sharing schemes also satisfies these properties, conclude, that the (worst case) information ratio of G is also at least r.

Note that this method is not necessarily universal, as properties (a)–(e) are too weak to capture exactly the functions coming from entropy see [17]. However, for graphs all existing lower bound proofs use the entropy method, and no example is known where the entropy method would not work.

3 The case of infinite graphs

Trying to define secret sharing on an infinite object one faces several problems. As there are infinitely many participants, one has to define infinitely many random variables with a joint distribution. But infinitely many pairwise random bits (probably needed for any construction) require huge event space, where the standard entropy function does not exist. We used entropy as a tool to define the relative size of a share compared to the secret, but even finding such a weaker notion is problematic; see Problem 6.1

Rather than defining the information ratio directly, we choose an indirect way. In case of graphs the set of participant might be infinite, but the minimal qualified subsets are finite, namely pairs. Thus it seems quite natural to consider *finite restrictions*. Our starting point is the following easy, but very useful fact about secret sharing schemes on finite graphs. The fact generalizes easily to other access structures as well.

Fact 3.1 *Suppose G' is a spanned subgraph of G. The information ratio of G' is at most as large as the information ratio of G, i.e. $R(G') \leq R(G)$.* ∎

In general, this claim is not true for arbitrary subgraphs. By Shamir's result in [18], $R(K_n) = 1$ where K_n is the complete graph on n vertices, while by [7], there is a graph $G' \subseteq K_n$ where $R(G') \geq 0.25 \log n$.

Looking at an infinite graph as the "limit" of its finite spanned subgraphs, Fact 3.1 suggests the following definition:

Definition 3.2 The *information ratio* $R(G)$ for the infinite graph G is

$$R(G) = \sup\{R(G') : G' \text{ is a finite, spanned subgraph of } G\}. \ ∎$$

By fact 3.1 this is a sound definition, and applying to a finite G gives back the original value.

If the (finite) graph G is the disjoint union of G_1 and G_2, i.e. there are no cross edges between G_1 and G_2, then any secret sharing scheme on G is simply a composition of a secret sharing scheme on G_1, and another one on G_2.

Claim 3.3 *If G has several connected components, then*

$$R(G) = \sup\{R(G') : G' \textit{ is a connected component of } G\}. \quad \blacksquare$$

Consequently, in definition 3.2 it is enough to consider connected finite subgraphs of G only.

We have defined the information ratio of an infinite graph as a supremum. It is a natural question whether this value is actually taken, or is it a proper one.

Definition 3.4 The graph G is *local* if there is a finite spanned subgraph G' of G such that $R(G) = R(G')$. Otherwise G is *not local.* ∎

Of course, when $R(G)$ is infinite then G cannot be local as no finite graph has infinite information ratio. Locality is interesting only when $R(G)$ is finite.

When constructing secret sharing schemes the most frequently used tool is Stinson's decomposition technique from [20]. For our case it can be worded as

Theorem 3.5 (Stinson) *Let $G_i \subseteq G$ be arbitrary subgraphs of G, and assume that each edge of G is in at least k of the subgraphs. Let $\mathcal{S}_i$ be a perfect secret sharing scheme on G_i such that $\mathcal{S}_i$ assigns $\mathcal{S}_i(v)$ bits to $v \in G$ for each bit in the secret ($\mathcal{S}_i(v) = 0$ if $v \notin G_i$). Then there is a scheme $\mathcal{S}$ on G which assigns*

$$\mathcal{S}(v) = \frac{1}{k}\sum \mathcal{S}_i(v)$$

bits to v for each bit in the secret. ∎

This theorem is meaningful for finite graphs only. The following generalization, however, holds for infinite graphs as well.

Theorem 3.6 *Let $G_i \subseteq G$ be arbitrary (finite or infinite) subgraphs of G, and assume that each edge of G is in at least k of the subgraphs. For a vertex $v \in G$ define $r_i(v) = 0$ if $v \notin G_i$, and $r_i(v) = R(G_i)$, i.e. the information ratio of G_i otherwise. Then*

$$R(G) \le \sup_{v\in G} \frac{\sum r_i(v)}{k}.$$

Proof Let us denote the value of the sup on the right hand side by r; we may assume that r is finite otherwise there is nothing to prove. Let $G' \subseteq G$ be a finite *spanned* subgraph of G. According to Definition 3.2, we need to check that there is a perfect secret sharing scheme $\mathcal{S}$ on G' which assigns at most r bits to each vertex of G' for each bit in the secret.

As G' has finitely many edges, we can choose a finite set I of the indices of the subgraphs G_i such that each edge of G' is in at least k of the subgraphs in the family $\{G_i : i \in I\}$. For $i \in I$ let G'_i be the spanned subgraph of G_i restricted to the vertices of G'. As $R(G_i) = r_i$ and G'_i is a spanned subgraph of G_i, by Fact 3.1 there is a secret sharing scheme $\mathcal{S}_i$ on G'_i which assigns at most r_i bits to all $v \in G'_i$ for each bit in the secret. By Theorem 3.5 then there is a scheme $\mathcal{S}$ which assigns

$$\mathcal{S}(v) = \frac{1}{k}\sum_{i\in I}\mathcal{S}_i(v) \le \frac{1}{k}\sum_{i\in I} r_i(v) \le \frac{\sum r_i(v)}{k} \le r$$

bits to $v \in G'$, which was wanted. ∎

We close this section by a generalization of Stinson's celebrated result [20]. For the proof we need some well-known facts. The first statement is a folklore, the proof is an easy application of the entropy method, see, e.g. [2], or the results in section 4.

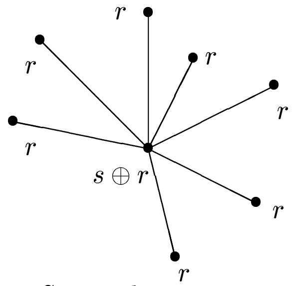

Figure 1: Secret sharing on a star

Claim 3.7 *If G is empty (independent), then $R(G) = 0$. Otherwise $R(G) \geq 1$.* ∎

The complete graph on (countably) infinitely many points is denoted by K_∞, and the (infinite) graph where one point is connected to an infinite independent set is denoted as Star_∞.

Claim 3.8 $R(K_\infty) = R(\mathrm{Star}_\infty) = 1$.

Proof All finite spanned subgraph of K_∞ is the complete graph. By Shamir's result in [18] all of them have ratio 1, thus their sup is also 1.

As for the other graph, $R(\mathrm{Star}_\infty) \geq 1$ by Claim 3.7. We show that this value is also ≤ 1. Each finite, connected spanned subgraph of Star_∞ is a (finite) star. Let the secret be the random bit $s \in \{0, 1\}$, and let $r \in \{0, 1\}$ be selected randomly and independently from s. The center of the star will get $s \oplus r$, and all other vertices get r. This is a perfect secret sharing scheme; all participants get 1 bit, and the secret is 1 bit as well (see figure 1). Thus the ratio in this case is 1 as well. ∎

Corollary 3.9 *If the maximal degree of G is d, then $R(G) \leq (d + 1)/2$.*

Proof For each vertex v in G consider the star G_v with center v and all edges outgoing from v as rays. These subgraphs G_v cover all edges twice, and each vertex is in at most $d + 1$ of these subgraphs (once as center, and d times as endpoint of a ray). Now $R(G_v) = 1$ by Claim 3.8, and apply Theorem 3.6. ∎

4 Lower bounds: the comb

Almost all lower bounds use the celebrated entropy method, see [3, 4, 6, 19] which has been outlined at the end of section 2. We shall use this method for a particular family of graphs, which will then be used to determine the exact ratio of several infinite graphs as well.

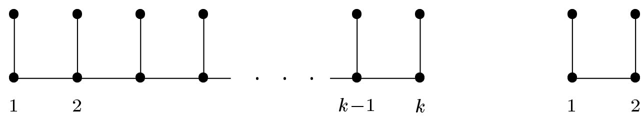

Figure 2: The graphs Comb_k and Comb_2

Definition 4.1 For $k \geq 2$ Comb_k is the graph on $2k$ vertices as indicated on figure 2, in particular Comb_2 is the path of length 3. Comb_∞ is the infinite comb with no 2-degree vertex. ∎

The main result of this section is

Theorem 4.2 *For $k \geq 2$, $R(\mathrm{Comb}_k) = 2 - 1/k$.*

Proof By using Stinson's decomposition method we show that the ratio is $\leq 2 - 1/k$. Figure 3 indicates the first two of the k different covers of Comb_k; each component in a cover is a star on two, three or four vertices, thus has information ratio 1. The numbers below the vertices indicate how many bits that vertex receives. Putting together all k covers, each edge is covered k times, the bottom vertices receive a total of $2k - 1$ bits, while the top vertices receive k bits. Using Theorem 3.5 we conclude that $R(\mathrm{Comb}_k) \leq (2k - 1)/k$.

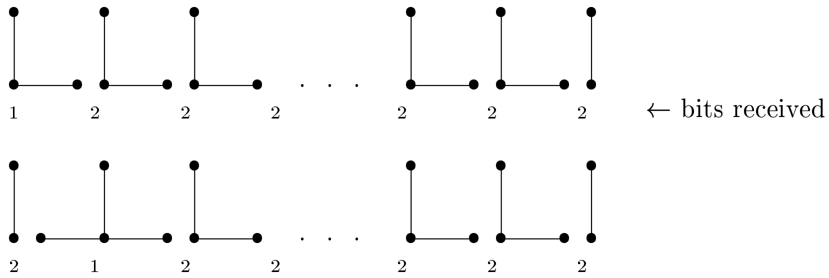

Figure 3: Two out of k covering of Comb_k with stars

For the other direction we use the entropy method. Label the bottom vertices of Comb_k from left to right as $A_1, A_2, \ldots, A_k$, the top vertices as $a_1, \ldots, a_k$ so that a_i is connected to A_i only. Let f be any real valued function which satisfies properties (a)–(e) from section 2 for the particular access structure which is defined by the edges of Comb_k. Then

$$\sum_{i=1}^{k} f(A_i) \geq 2k - 1. \tag{2}$$

If we can show this, then we are done. The sum of k terms is at least $2k-1$, thus at least one of them is $\geq (2k-1)/k$. Given *any* secret sharing scheme on Comb_k, the particular choice in (1) for the function f satisfies (a)–(e), consequently for at least one A_i we have

$$f(A_i) = \frac{\mathbf{H}(\xi_{A_i})}{\mathbf{H}(\xi_s)} \geq \frac{2k-1}{k},$$

i.e. someone *must remember* at least $2 - 1/k$ bits for each bit in the secret.

To finish the proof, we state and prove two lemmas, 4.3 and 4.4. Inequality (2) is just the sum of the claims in the lemmas. ∎

Lemma 4.3 $\sum_{i=1}^{k} f(A_i) \geq f(A_1A_2\ldots A_k) + k - 2.$

Proof In general for any $2 \leq \ell \leq k$, $\sum_{i=1}^{\ell} f(A_i) \geq f(A_1A_2\ldots A_\ell) + \ell - 2$ which we will prove by induction on ℓ. When $\ell = 2$ the claim is $f(A_1) + f(A_2) \geq f(A_1A_2)$, which is just the submodularity (c).

Now suppose we know the claim to be true for $\ell - 1$; to conclude it for ℓ it is enough to check that whenever $\ell \geq 3$ then

$$f(A_1A_2\ldots A_{\ell-1}) + f(A_\ell) \geq f(A_1\ldots A_\ell) + 1. \tag{3}$$

As $\ell \geq 3$, both $A_1\ldots A_{\ell-1}$ and $A_{\ell-1}A_\ell$ contain edge, i.e. they are a qualified sets. Then property (e) says that

$$f(A_1A_2\ldots A_{\ell-1}) + f(A_{\ell-1}A_\ell) \geq f(A_1\ldots A_\ell) + f(A_{\ell-1}) + 1$$

as the singleton $A_{\ell-1}$ is not qualified. By the submodularity (c) we have

$$f(A_{\ell-1}) + f(A_\ell) \geq f(A_{\ell-1}A_\ell).$$

Adding up the last two inequalities we get (3), as required. ∎

Lemma 4.4 $f(A_1A_2\ldots A_k) \geq k + 1.$

Proof Let $X = \{A_1A_2\ldots A_k\}$, and consider the differences

$$d_i = f(Xa_1\ldots a_i) - f(a_1\ldots a_i).$$

As $f(\emptyset) = 0$, the value we are interested in is d_0. The trick is to consider these differences in reverse order. As $\{Xa_1\ldots a_k\}$ is qualified, while $\{a_1\ldots a_k\}$ is not, condition (d) gives $1 \leq d_k$. Furthermore, for all $1 \leq i \leq k$, $d_i + 1 \leq d_{i-1}$ which implies $k + 1 \leq d_0$ as the lemma states.

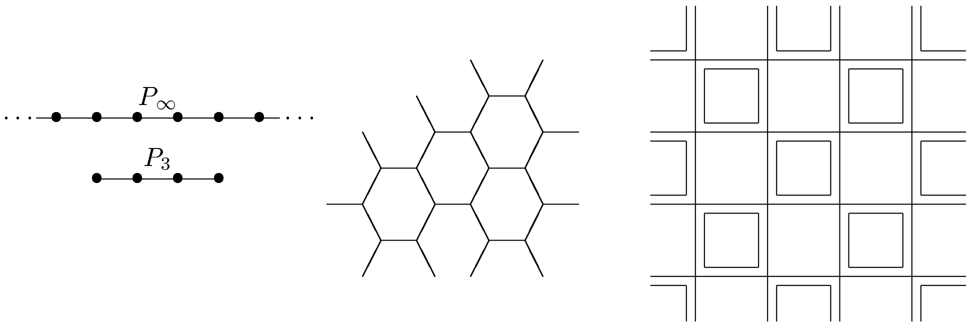

Figure 4: Paths, honeycomb, and 2-lattice

Now both $\{A_i a_1 \dots a_i\}$ and $\{X a_1 \dots a_{i-1}\}$ are qualified, their intersection, which is $\{A_i a_1 \dots a_{i-1}\}$, is not, thus (e) gives

$$f(A_i a_1 \dots a_i) + f(X a_1 \dots a_{i-1}) \geq f(A_i a_1 \dots a_{i-1}) + f(X a_1 \dots a_i) + 1.$$

Furthermore the submodularity (c) tells

$$f(a_1 \dots a_i) + f(A_i a_1 \dots a_{i-1}) \geq f(A_i a_1 \dots a_i) + f(a_1 \dots a_{i-1})$$

Adding these up and rearranging we get $d_i + 1 \leq d_{i-1}$, as needed. ∎

Theorem 4.5 *The information ratio of* Comb_∞ *is* 2.

Proof Each connected component of Comb_∞ is a spanned subgraph of Comb_k for some k, consequently, by Theorem 4.2, its information ratio is < 2.

On the other hand, for each natural number k, Comb_k is a spanned subgraph of Comb_∞, thus $R(\mathrm{Comb}_\infty)$ is at least $\sup_k \{2 - 1/k\} = 2$. ∎

Corollary 4.6 Comb_∞ *is not local, i.e. all of its finite spanned subgraphs have smaller information ratio.* ∎

5 Examples

The path of length n (with $n+1$ vertices) is denoted as P_n, and P_∞ is the infinite path, or the 1-dimensional lattice. The *honeycomb* is the two-dimensional tiling of the plane with regular hexagons. The two-dimensional (square) *lattice* is the usual checkered paper like tiling (see fig 4), and the triagonal tiling is the tiling with regular triangles (fig 7).

Example 5.1 $R(P_\infty) = 3/2$.

Proof We have seen that Comb_2 is the same graph as the path of length 3, thus by Theorem 4.2 $R(\mathrm{Comb}_2) = R(P_3) = 3/2$. (For other proofs see, e.g. [2, 3, 6], or the Appendix.) As P_3 is a spanned subgraph of P_∞), we have $R(P_3) = 3/2 \leq R(P_\infty)$. For the other direction we use

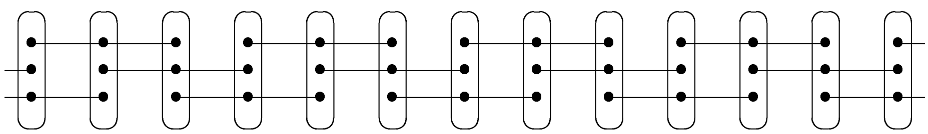

Figure 5: Covering P_∞ by stars

Theorem 3.6 for the 2-cover indicated on figure 5. All subgraphs in the cover are stars having ratio 1; each edge is covered twice and each vertex gets 3 bits. Thus $R(P_\infty) \leq 3/2$. ∎

Example 5.2 $R(\text{honeycomb}) = 2$.

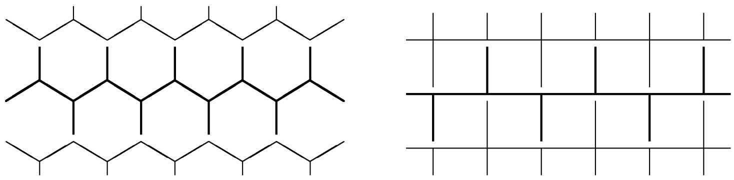

Figure 6: The comb as spanned subgraph

Proof As each vertex has degree 3, Corollary 3.9 says that the ratio is at most 2. On the other hand the, honeycomb contains the infinite comb as a spanned subgraph (left picture on figure 6). Consequently $2 = R(\mathrm{Comb}_\infty) \le R(\text{honeycomb})$. ∎

Example 5.3 $R(\text{2-lattice}) = 2$.

Proof Here each vertex has degree 4, thus Corollary 3.9 gives only 5/2. We could, however, apply Theorem 3.6 directly for the four-cycles C_4 indicated on figure 4. Each edge is covered once, and each vertex is in two cycles. As C_4 has information ratio 1 we proved that the information ratio for the 2-lattice is ≤ 2. The statement on C_4 can be shown as follows: let the random bit $s \in \{0,1\}$ be the secret, and pick $r \in \{0,1\}$ randomly and independently from s. Give r to the first and third node in C_4, and $r \oplus s$ to the two other nodes.

The lower bound follows from the fact that the infinite comb can be embedded to the 2-lattice as a spanned subgraph, see Figure 6. ∎

The proof given here does not tell whether the 2-lattice is local or not. In the Appendix we show that a particular graph on 8 vertices has information ratio 2. That graph is a spanned subgraph of the 2-lattice, thus the 2-lattice is *not* local.

Example 5.4 $2 \le R(\text{triangle lattice}) \le 12/5$.

Proof The lower bound follows again from the fact that the comb can be embedded into this lattice as well, see figure 7.

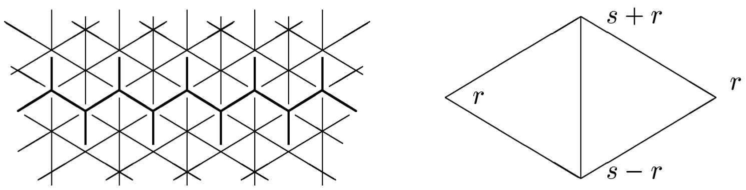

Figure 7: The triangle lattice

The construction which gives the upper bound 12/5 is due to Péter Gergely [13]. Consider the diamond-shaped graph on four vertices at the right hand side of figure 7. This graph has ratio 1, which can be shown as follows: pick the secret s from $\{0,1,2\}$ uniformly, and pick a random r from the same set. Give r to nodes on the left and on the right, give $s+r \bmod 3$ to the top vertex, and $s-r \bmod 3$ to the bottom vertex. Any two connected vertices can recover the secret s, moreover each assigned number is independent of s. The two unconnected vertices receive the same share, thus their joint information is independent of the secret as well. Both the secret and the shares have entropy $\log 3$, thus the ratio is 1.

The cover consists of *all* spanned subgraphs of the triangle lattice isomorphic to this graph. It is easy to check that this is a 5-cover (each edge is in 5 of such subgraphs), and each vertex is covered 12 times. By Theorem 3.6 this gives 12/5 as an upper bound. ∎

The countable *universal* or *random graph* is the unique (up to isomorphism) graph on countable many vertices which has the property that picking finitely many vertices v_i and numbers $\varepsilon_i \in \{0,1\}$, there exists a vertex in the graph which is connected to v_i just in case $\varepsilon_i = 1$.

Example 5.5 *The information ratio for the universal graph is* ∞.

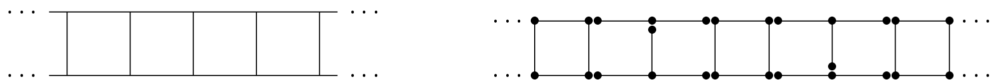

Figure 8: The ladder and a cover

Proof As we have remarked, there is a graph of n vertices with information ratio $\geq 0.25 \log n$, see [7]. As all finite graphs can be embedded into the universal graph as spanned subgraphs, its information ratio is at least $0.25 \log n$, i.e. not bounded. ∎

Example 5.6 *The information ratio for the infinite binary tree is* 2.

Proof Lower bound: for each k the graph Comb_k can be embedded into this graph, consequently $R \geq 2$.

Upper bound: we show that each tree has information ratio ≤ 2 by using the Theorem 3.6. We may assume that the tree is connected. Pick any vertex and consider it as "root." Direct the edges recursively away from the root. When all edges has been directed, each node has one invertex except for the root which has none. Consider the stars with center at a vertex consisting of all outgoing edges. Each edge is covered exactly once, and each vertex gets one or two bits (one bit for the root and leaves, and two bits for all other vertices). ∎

In [9] it has been proved that all finite trees have information ratio $2 - 1/k$ for some integer $k \geq 1$. Therefore the information ratio of an infinite tree is the sup of numbers of this form. If T is an infinite tree, then either $R(T) = 2$, and then T is not local (as, e.g., is the case for the complete binary tree in Example 5.6), or $R(T) = 2 - 1/k$ for some integer $k \geq 1$, and then T is local.

Example 5.7 *There is an infinite graph with information ratio* $5/3$.

Proof Consider the (rooted) complete binary tree from example 5.6, and insert a new vertex at the midpoint of every edge. This will be our graph T.

As three-tooth comb can be embedded into T, its information ratio is $R(T) \geq R(\mathrm{Comb}_3)$ $= 2 - 1/3$. The upper bound comes from Theorem 3.6: we define a 3-cover by stars so that each vertex gets 5 bits; this gives $R(T) \leq 5/3$. First take the stars with center at the inserted new vertices and both edges as rays. Second take stars at the old vertices with all incident edges as rays, but take them twice. That way an old vertex gets $2 \cdot 1 + 3$ bits (two from the double star, and one from each new neighbor), while a new vertex gets $2 \cdot 2 + 1$ bits, as required. ∎

Example 5.8 *The infinite ladder L on figure 8 has information ratio* $10/6 \leq R(L) \leq 11/6$.

Proof The upper bound comes from the following construction (see figure 8). The 1-cover on the right has period 6. It uses stars and C_4, both assigns a single bit for each bit in the secret. In a period all vertices get 2 bits, except for one in the top, and one in the bottom which get only 1 bit. Shifting this cover by 1, ..., 5 we get a 6-cover, and each vertex gets a total of 11 bits. By Theorem 3.6 the upper bound follows.

For the lower bound one can observe that the infinite path is a spanned subgraph. Unfortunately this gives only $9/6$. For the missing $1/6$ we prove in the Appendix that the graph G_1 on figure 9 has information ratio $10/6$. As this is a spanned subgraph of the ladder, we are done. ∎

We remark that the ladder of width 2 has information ratio 2. It is a spanned subgraph of the 2-lattice (≤ 2), and contains Comb_∞ as a spanned subgraph (≥ 2).

6 Conclusion and further problems

Determining the exact amount of information a participant must remember in a perfect secret sharing scheme is an important problem both from theoretical and practical point of view. Access structures based on graphs pose special challenges. They are easier to define, have a transparent, and sometimes trivial, structure.

In this paper we extended the definition of secret sharing for infinitely many participants, and gave a definition for its information ratio. We consider this extension to be an important contribution, and hope to see further, interesting applications.

We have used a compactness-type definition to overcome the difficulty of infinite entropy: the information ratio is defined as the sup of the ratio for the finite embedded structures. The first problem is to find an appropriate definition of the "relative information content" for arbitrary random variables.

Problem 6.1 Given two random variables ξ and η, define their *relative size*, which is the analog of $\mathbf{H}(\xi)/\,\mathbf{H}(\eta)$ when both ξ and η are finite.

In [5] the authors show that the secret and shares, as random variables, cannot be based on a countable domain, even if the number of participants is finite. The paper also contains the following using reals numbers: let the secret ξ_s be uniform in $[0,1)$, then choose the shares ξ_i of the first $k-1$ of the participants uniformly and independently in $[0,1)$, and choose $\xi_k \in [0,1)$ so that

$$\xi_s = \xi_1 + \ldots + \xi_{k-1} + \xi_k \pmod 1$$

It is easy to check that ξ_s is independent of any set of $k-1$ shares, and, of course, all shares determine the secret uniquely. Just as in the finite case, this scheme can be used as a building block to create a perfect secret sharing when all qualified subsets are finite: simply distribute ξ_s for each qualified subset independently. In this case each participant will receive as many shares as many minimal qualified sets it is in. In particular, if the scheme is based on a graph, then this number will be the degree of the vertex.

In case of finite complete graphs the above construction has extremely high information ratio. Shamir's construction from [18] is more efficient, it has the lowest possible information ratio 1. Can Shamir's construction be generalized for the infinite case? The beginning is easy. Pick elements x_i of a field for each participant, and pick the point x_s for the secret (these values are public). Choose a secret linear function $p(x) = ax + b$ according to a certain distribution. The value of the secret is $\xi_s = p(x_s)$; and the shares are $\xi_i = p(x_i)$. Clearly all pair of shares determine the function p, thus the value ξ_s. What is not clear why ξ_s and ξ_i should be independent.

Problem 6.2 Do there exists a distribution for which ξ_s and ξ_i are always independent when the linear functions are taken (a) over reals, (b) over some appropriately chosen field?

By the result of [5] the field cannot be countable. Here and in all problems "determines" can (and should) mean that ξ_s is determined uniquely with probability 1.

Existence of an "infinite" threshold scheme seems to be a pure probability theoretical question.

Problem 6.3 Do there exists a perfect scheme where the secret is independent of finitely many of the shares, but is determined by infinitely many of them? Or at least do there exist a *rump scheme* where the secret is independent of any finite collection of the shares, but is determined by any cofinite collection (i.e. all but finitely many) of shares?

We return to schemes based on graphs. In almost all examples we have used Theorem 3.6, the generalization of Stinson's decomposition theorem. Does it generalize for infinite sharing schemes as well?

Problem 6.4 Suppose that G_i is (an arbitrary) subgraph of G, and $\mathcal{S}_i$ is a perfect secret sharing scheme on G_i. Assume moreover that all edges of G is contained in at least t of the subgraphs. Does it follow that there is perfect secret sharing scheme on G?

The problem is that the secret in $\mathcal{S}_i$ might have arbitrary distribution, and it is not clear how to combine them.

In the Examples section we have computed the information ratio for several infinite graphs. In two cases: the triangle lattice (Example 5.4) for the ladder (Example 5.8) we could only give estimates.

Problem 6.5 Determine the exact information ratio of the triangle lattice and that of the ladder.

We conjecture that for the triangle the upper bound, for the ladder the lower bound is the right value.

We have seen two examples for non-local graphs: the comb, and the complete binary tree. Both of them have information ratio 2, but every finite spanned subgraph has smaller information ratio.

Problem 6.6 Is there any non-local graph with information ratio strictly below 2?

In fact is the following stronger conjecture true:
Problem 6.7 Is it true that if $R(G) < 2$ then $R(G) = 2 - 1/k$ for some integer k?

References

[1] A. Beimel, N. Livne: On matriods and non-ideal secret sharing, In: Proc. of the third Theory of Cryptography Conference, Lecture Notes in Computer Science, Vol 3876(2006) pp. 482–501

[2] C. Blundo, A. De Santis, D. R. Stinson, U. Vaccaro: Graph Decomposition and Secret Sharing Schemes *Journal of Cryptology*, Vol 8(1995) pp. 39–64.

[3] C. Blundo, A. De Santis, R. D. Simone, U. Vaccaro: Tight Bounds on the Information Rate of Secret Sharing Schemes *Designs, Codes and Cryptography*, Vol 11(1997) pp. 107–110

[4] R. M. Capocelli, A. De Santis, L. Gargano, U. Vaccaro: On the size of shares of secret sharing schemes *Journal of Cryptology*, vol 6(1993), pp. 157–168

[5] B. Chor, E. Kushilevitz: Secret sharing over infinite domains *Journal of Cryptology*, Vol 6(1993) pp. 87–96

[6] L. Csirmaz: The size of a share must be large *Journal of Cryptology*, vol 10(1997) pp. 223–231

[7] L. Csirmaz: Secret sharing schemes on graphs, *Studia Mathematica Hungarica*, in print – available as IACR preprint `http://eprint.iacr.org/2005/059`

[8] L. Csirmaz: Exact rate of secret sharing schemes on d-dimensional cube, *manuscript*

[9] L. Csirmaz, G. Tardos: Exact information rate of trees, *manuscript*

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

[11] M. van Dijk: On the information rate of perfect secret sharing schemes *Designs, Codes and Cryptography*, Vol 12(1997) pp. 143–169

[12] M. van Dijk, T. Kevenaar, G. J. Schrijen, P. Tuyls: Improved constructions of secret sharing schemes by applying (λ,ω)-decompositions Inf. Processing Letters, vol 99(2006) pp. 154–157

[13] P. Gergely: Personal communication

[14] O. Ferràs, J. Martì-Farré, C. Padró: Ideal Multipartite Secret Sharing Schemes Preprint (2006) `http://www-ma4.upc.edu/ cpadro/papers/mltprtt.pdd`

[15] W. Jakson, K. M. Martin: Perfect secret sharing schemes on five participants, *Designs, Codes and Cryptography*, Vol 9(1996) pp. 233-250

[16] J. Martì-Farré, C. Padró: Secret sharing schemes with three or four minimal qualified subsets *Designs, Codes and Cryptography*, Vol 34(2005) pp. 17–34

[17] F. Matus: Matroid representations by partitions, *Discrete Mathematics*, vol 203(1999) pp. 169–194

[18] A. Shamir: How to share a secret, *Communications of the ACM*, vol 22(1979), pp. 612–613

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

[20] D. R. Stinson: Decomposition construction for secret sharing schemes, *IEEE Trans. Inform. Theory* Vol 40(1994) pp. 118-125.

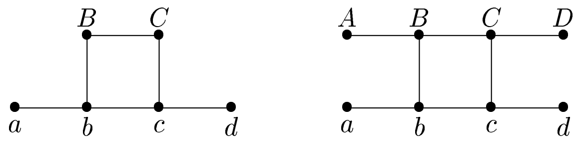

Figure 9: The graphs G_1 and G_2

Appendix

We determine the exact information ratio for the graphs G_1 and G_2 of figure 9. G_1 is a spanned subgraph of the infinite ladder in Example 5.8 while G_2 is a spanned subgraph of the 2-lattice (but not of the ladder).

Claim 6.8 *The information ratio of* G_1 *is* $5/3$.

Proof For the upper bound consider the following 3 covers of G_1. The first cover consists of the cycle $BbcCB$ (assigning 1 bit for each secret bit), plus the two edges ab and cd. Using this cover nodes b and c get two bits, all other nodes get one bit. The second cover contains the star with rays ba, bB, bc; the path Ccd and the edge BC. Here a, b, and d get one bit, all other nodes get two bits. The third cover is the mirror image of the second one: the star cb, cC cD, the path abB, and the edge BC. Using all three covers all edges are covered three times, and every node gets either three (a and d), or five bits (all the rest).

For the lower bound we use the entropy method. Assume f satisfies properties (a)–(e) listed at the end of section 2. We claim that

$$f(b) + f(c) + f(C) \geq 5, \tag{4}$$

i.e. at least one of b, c and C gets $5/3$ bits for each bit in the secret.

First we give a strengthening of the usual proof that the information ratio of the path of length 3 is at least $3/2$. That proof goes by showing that $f(bc) \geq f(abcd) - f(ad) + 2$. As $abcd$ is qualified and ad is not, $f(abcd) - f(ad) \geq 1$. That is, $f(b) + f(c) \geq f(bc) \geq 3$, therefore either $f(b)$ or $f(c)$ is $\geq 3/2$. Here we show that in this inequality $f(abcd)$ can be replaced by $f(acd)$:

$$\begin{array}{r}
f(a) + f(b) \geq f(ab) \\
f(ab) + f(bc) \geq f(b) + f(abc) + 1 \\
f(abc) \geq f(ac) + 1 \\
f(ac) + f(ad) \geq f(a) + f(acd) \\
\hline
f(bc) + f(ad) \geq f(acd) + 2
\end{array}$$

Second, we take into account the vertices B and C as well:

$$\begin{array}{r}
f(c) + f(C) \geq f(cC) \\
f(cd) + f(cC) \geq f(c) + f(cdC) + 1 \\
f(acd) + f(cdC) \geq f(cd) + f(acdC) \\
f(acdC) \geq f(adC) + 1 \\
f(adC) + f(adB) \geq f(ad) + f(adBC) \\
f(adBC) \geq f(adB) + 1 \\
\hline
f(C) + f(acd) \geq f(ad) + 3
\end{array}$$

As $f(b) + f(c) \geq f(bc)$, the sum of the two inequalities gives (4). ∎

Claim 6.9 G_2 *has information ratio* 2.

Proof $R(G_2) \leq 2$ as G_2 is a spanned subgraph of the 2-lattice, and the 2-lattice has information ratio 2. On the other hand, let f be again any function satisfying (a)–(e); we claim that

$$f(bc) + f(BC) \geq 8. \tag{5}$$

As $f(b) + b(c) + f(B) + f(C) \geq f(bc) + f(BC) \geq 8$, the lower bound 2 follows.

Each of the inequalities below are instances of one of the properties (a)–(e) of the function f:

$$\begin{array}{r}
f(a) + f(b) \geq f(ab) \\
f(ab) + f(bc) \geq 1 + f(b) + f(abc) \\
f(acAC) - f(acA) \geq f(acACD) - f(acAD) \geq 1 \\
f(acABC) - f(acAC) \geq 1 \\
f(ac) - f(a) \geq f(acB) - f(aB) \\
f(acB) - f(aB) \geq 1 + f(acABC) - f(aABC) \\
f(abc) - f(ac) \geq f(abcA) - f(acA) \\
\hline
f(bc) \geq 4 + f(abcA) - f(aABC)
\end{array}$$

Swapping lower case and upper case letters leaves the graph unchanged, thus we also have the "swapped" instance:

$$f(BC) \geq 4 + f(aABC) - f(abcA).$$

Adding these latter two inequalities we get (5), as required. ∎