New upper bounds for anti-collusion schemes: A graph theoretical approach

Yujie Gu

Graduate School of Systems and Information Engineering,
University of Tsukuba

s1530147@u.tsukuba.ac.jp

1 Introduction

To protect the copyright of digital contents, a dealer who possesses a large amount of copyrighted data would charge for the copyright of data. The one who paid for the copyright can get the access to data, and the guy who did not pay would be held back. In broadcast encryption, the dealer encrypts the copyrighted data and uploads to a public cloud. Anyone can download the public encrypted contents, but a decryption key is required to recover the original data.

To hinder the illegal redistribution of the decryption key, the dealer would assign each authorized user, who purchased the copyright of data, a personal decoder, which is a collection of base decryption keys and can be used to recover the data (maybe with the help of some devices). However, several dishonest users (traitors) may work together to generate a new decoder (pirate) and distribute it to several unauthorized users. Anti-collusion schemes were introduced to help the dealer design the judicious key-distribution strategy and trace back to traitors once a pirate copy is confiscated [2, 3].

Based on a threshold secret sharing scheme, Stinson *et al.* [9] proposed the traceability scheme for the anti-collusion key-distribution in broadcast encryption and studied it from a combinatorial viewpoint. In this setting, a traceability scheme(v,k) is a set system $(\mathcal{X},\mathcal{B})$ with the desired properties, where $\mathcal{X}$ is a finite set of size v and $\mathcal{B}$ is collection of k-subsets of $\mathcal{X}$. The ground set $\mathcal{X}$ corresponds to the set of v base keys. Each authorized user, who paid for the copyright, is assigned with a k-subset of $\mathcal{X}$, which can be used to decrypt the encrypted contents. Thus the family $\mathcal{B}$ of k-subsets of $\mathcal{X}$ represents all the authorized users.

A *t-collusion* means that t dishonest users (traitors) $B_1,\dots,B_t \in \mathcal{B}$ work together to generate a k-subset (pirate) $T \subseteq \bigcup_{1\le i\le t} B_i$ and redistribute T to some unauthorized users. Stinson *et al.* [9] showed that their traceability scheme can ensure that once a pirate in a t-collusion is confiscated, at least one traitor can be traced back. Also in the same setting, parent-identifying set system was investigated in [4] with the advantage that can accommodate more users than traceability schemes, where the required properties are weaker than that of traceability schemes. The idea of parent-identifying property was introduced by Hollmann *et al.* in [8]. We first state the definition of parent-identifying set systems as follows.

Definition 1. *A (w,v) t-parent-identifying set system (or t-IPPS(w,v), for short) is a pair $(\mathcal{X},\mathcal{B})$ such that $|\mathcal{X}|=v$, $\mathcal{B}\subseteq\binom{\mathcal{X}}{w}$, with the property that for any w-subset $T\subseteq\mathcal{X}$, either $P_t(T)$ is empty, or*

$$\bigcap_{\mathcal{P}\in P_t(T)}\mathcal{P}\neq\emptyset,$$

where

$$P_t(T)=\{\mathcal{P}\subseteq\mathcal{B}:\ |\mathcal{P}|\leq t,\ T\subseteq\bigcup_{B\in\mathcal{P}}B\}.$$

When a pirate T generated by a s-collusion, $1\leq s\leq t$, is confiscated, t-IPPSs ensure that at least one traitor can be traced back. In fact, one could check each subset of $\mathcal{B}$ with size at most t and then get $P_t(T)$. By Definition 1, the intersection of all members in $P_t(T)$ is nonempty, and each guy in the intersection is a traitor.

The cardinality of $\mathcal{B}$ is called the *size* of the set system. Since the size of the set system corresponds to the number of authorized users in this scheme, we expect that the size can be as large as possible. Denote $I_t(w,v)$ as the maximum size of a t-IPPS(w,v). A t-IPPS(w,v) is called *optimal* if it has size $I_t(w,v)$. Given parameters t,w and v, the goal is to explore the exact value of $I_t(w,v)$ and to construct optimal t-IPPS(w,v). In the next section, we will argue the bounds of $I_t(w,v)$.

The following is one example of 2-IPPS.

Example 1. *Let $\mathcal{X}=\{1,2,\ldots,11\}$ and $\mathcal{B}=\{B_1=\{1,2,3,4\},B_2=\{3,5,6,7\},B_3=\{4,7,8,9\},B_4=\{2,7,10,11\}\}$. By Definition 1, $(\mathcal{X},\mathcal{B})$ is a 2-IPPS$(4,11)$.*

For instance, if $T=\{2,3,5,7\}$, we have

$$\{2,3,5,7\}\subseteq B_1\cup B_2,$$
$$\{2,3,5,7\}\subseteq B_2\cup B_4.$$

Then $P_2(T)=\{\{B_1,B_2\},\{B_2,B_4\}\}$ and $\bigcap_{\mathcal{P}\in P_2(T)}\mathcal{P}=\{B_2\}\neq\emptyset$.

One can check that for each 4-subset $T\subseteq\mathcal{X}$, the desired property in Definition 1 can be satisfied. Thus $(\mathcal{X},\mathcal{B})$ is a 2-IPPS$(4,11)$.

2 Bounds for IPPS

2.1 Known results

In the literature, a combinatorial structure called *own-subset* by Erdős, Frankl and Füredi [5] was used to derive upper bounds for IPPS. In a set system $(\mathcal{X},\mathcal{B})$, $B\in\mathcal{B}$, a subset $B_0\subseteq B$ is called a $|B_0|$-*own-subset* of B if for any $B'\in\mathcal{B}\setminus\{B\}$, we have $B_0\nsubseteq B'$.

The first upper bound for IPPS was given by Collins in [4] by investigating own-subsets with size $\lceil\frac{w}{\lfloor t^2/4\rfloor+\lceil t/2\rceil}\rceil$.

Theorem 1 ([4])**.** *Let $v\geq w\geq 2,\ t\geq 2$ be integers. Then*

$$I_t(w,v)\leq\frac{\binom{v}{\lceil\frac{w}{\lfloor t^2/4\rfloor+\lceil t/2\rceil}\rceil}}{\binom{\lceil\frac{w}{\lfloor t/2\rfloor+1}\rceil-1}{\lceil\frac{w}{\lfloor t^2/4\rfloor+\lceil t/2\rceil}\rceil-1}}=O(v^{\lceil\frac{w}{\lfloor t^2/4\rfloor+\lceil t/2\rceil}\rceil}).$$

In [7], Gu and Miao improved the above upper bound by showing that some block of a t-IPPS must contain at least one own-subset with a smaller size than $\lceil \frac{w}{\lfloor t^2/4 \rfloor + \lceil t/2 \rceil} \rceil$, that is, own-subsets with size $\lceil \frac{w}{\lfloor t^2/4 \rfloor + t} \rceil$. Obviously, $\lceil \frac{w}{\lfloor t^2/4 \rfloor + t} \rceil \leq \lceil \frac{w}{\lfloor t^2/4 \rfloor + \lceil t/2 \rceil} \rceil$ holds for all $v \geq w \geq 2$ and $t \geq 2$.

Theorem 2 ([7]). *Let $v \geq w \geq 2,\ t \geq 2$ be integers. Then*

$$I_t(w,v) \leq \binom{v}{\lceil \frac{w}{\lfloor t^2/4 \rfloor + t} \rceil} = O(v^{\lceil \frac{w}{\lfloor t^2/4 \rfloor + t} \rceil}).$$

Furthermore, Gu *et al.* [6] provided a lower bound for IPPS by virtue of the probabilistic methods, which shows that the upper bound in Theorem 2 has the best possible exponent for certain cases.

Theorem 3 ([6]). *Let w and t be fixed positive integers such that $t \geq 2$. Then there exists a constant c, depending only on w and t, with the following property. For any sufficiently large integer v, there exists a t-IPPS(w,v) with size at least $cv^{\frac{w}{u-1}}$, that is, $I_t(w,v) \geq cv^{\frac{w}{u-1}}$, where $u = \lfloor (\frac{t}{2}+1)^2 \rfloor$.*

2.2 New results

We can see that the upper bound in Theorem 2 and the lower bound in Theorem 3 have the same order of magnitude, $\frac{w}{\lfloor t^2/4 \rfloor + t}$, when $\lfloor t^2/4 \rfloor + t$ is a divisor of w and v is sufficiently large. However, when $\lfloor t^2/4 \rfloor + t$ is not a divisor of w, there is a gap between the order of magnitude in Theorem 2 and in Theorem 3. In the following, we consider the case that $t = 2$ and $w = 4$, where $3 \nmid 4$.

First, we have the following corollary directly from Theorem 2 and Theorem 3.

Corollary 1. *For sufficiently large v, we have*

$$cv^{4/3} \leq I_2(4,v) \leq \frac{1}{2}v^2,$$

where c is a positive constant.

One interesting problem is to determine the order of magnitude of the size of 2-IPPS$(4,v)$. By using a graph theoretic method, we show that

Theorem 4. $\lim\limits_{v \to \infty} I_2(4,v) = o(v^2)$.

The tool exploited in the argument of Theorem 4 is the well-known graph removal lemma proved by Alon, Duke, Lefmann, Rödl and Yuster in [1].

Lemma 1 ([1]). *For every $\gamma > 0$ and every positive integer k, there exists a constant $\delta = \delta(k,\gamma) > 0$ such that every graph G on n vertices, containing less than δn^k copies of the complete graph K_k on k vertices, contains a set of less than γn^2 edges whose deletion destroys all copies of K_k in G.*

Theorem 4 can be generated to the case $t = 3$ and $w = 6$ by a similar argument.

Theorem 5. $\lim\limits_{v\to\infty} I_3(6,v)=o(v^2)$.

However, for $t\geq 4$ and $w=\lfloor(\frac{t}{2}+1)^2\rfloor$, we may cannot have a similar argument as that of Theorem 4. Since in a graph, we can only get $2t$ points from t distinct edges, and the fact $w=\lfloor(\frac{t}{2}+1)^2\rfloor>2t$ for any $t\geq 4$ implies that $2t$ points are not enough to form a w-subset. But we believe that this obstacle can be removed by virtue of hypergraphs or some elaborate analyses. To be precise, we have the following conjecture.

Conjecture 1. *Suppose $t\geq 4$ is a positive integer, then*

$$\lim_{v\to\infty} I_t(w,v)=o(v^2),$$

where $w=\lfloor(\frac{t}{2}+1)^2\rfloor$.

Moreover, we conjecture that the upper bound in Theorem 4 is the best possible for 2-IPPS$(4,v)$. To be exact, we have

Conjecture 2. *For any constant $\epsilon>0$ and sufficiently large v, there exists a 2-IPPS$(4,v)$ with size $cv^{2-\epsilon}$, where c is a positive constant.*

We remark that to prove Conjecture 2, some techniques or tools in number theory and additive combinatorics may be required.

References

[1] N. Alon, R. A. Duke, H. Lefmann, V. Rödl, and R. Yuster, The algorithmic aspects of the regularity lemma, *J. Algorithms*, vol. 16, pp. 80–109, 1994.

[2] B. Chor, A. Fiat, and M. Naor, Tracing traitors, *in Cryto'94 (Lecture Notes in Computer Science),* Berlin, Heidelberg, New York: Springer-Verlag, vol. 839, pp. 480–491, 1994.

[3] B. Chor, A. Fiat, M. Naor, and B. Pinkas, Tracing traitors, *IEEE Trans. Inf. Theory,* vol. 46, no. 3, pp. 893–910, May 2000.

[4] M.J. Collins, Upper bounds for parent-identifying set systems, *Des. Codes Cryptogr.,* vol. 51, pp. 167–173, 2009.

[5] P. Erdős, P. Frankl, and Z. Füredi, Families of finite sets in which no set is covered by the union of r others, *Israel J. Math.,* vol. 51, pp. 79–89, 1985.

[6] Y. Gu, M. Cheng, G. Kabatiansky, and Y. Miao, Probabilistic existence results for parent-identifying schemes, preprint.

[7] Y. Gu and Y. Miao, Bounds on traceability schemes, *IEEE Trans. Inf. Theory,* to appear.

[8] H. D. L. Hollmann, J. H. van Lint, J.-P. Linnartz, and L. M. G. M. Tolhuizen, On codes with the identifiable parent property, *J. Combin. Theory Ser. A*, vol. 82, pp. 121–133, 1998.

[9] D.R. Stinson and R. Wei, Combinatorial properties and constructions of traceability schemes and frameproof codes, *SIAM J. Discrete Math.*, vol. 11, pp. 41–53, 1998.