Special Proceeding of 20th Annual Conference of SSCA
held at Pondicherry University, Puducherry during January 29-31, 2018; pp 01-08

Applications of the Theory of Block Designs and Probability Sampling in Cryptology

Mausumi Bose
Applied Statistics Division, Indian Statistical Institute, Kolkata 700108, India

Final Version Received on 23 August, 2018

Abstract
Cryptology is an ancient subject, with its early forms traceable to the Phraonic period of Egypt. Over the passage of time, cryptological methods have grown in sophistication and complexity. Currently, cryptology is a field of fundamental importance for protecting the confidentiality and integrity in communication. In this note we briefly highlight how concepts and results from the statistics literature on theory of block designs and varying probability sampling can be applied to different areas of cryptology to enhance the efficiency of the cryptographic schemes in these areas. Some open problems are also posed.

Key words: Anonymity in Shared Key Operations, Anti-collusion digital fingerprinting, Key pre-distribution in distributed sensor networks, Visual Cryptography

1 Introduction

In Sections 1 to 3, we focus on three different areas of cryptology and show how ideas from combinatorics and design of experiments can be applied in each of these areas to get efficient new schemes. In Section 4 we show how in another different area, ideas from sampling theory, in particular, sampling with probability proportional to size, can be used to get efficient schemes. A representative list of references is included and many more references may be obtained from these and also from a literature search. Some open problems are also stated at the end of the sections.

2 Visual Cryptography

In (k, n) visual cryptographic schemes (VCS), a secret image (e.g., a page of text, picture, diagram, etc., printed or handwritten) is encrypted into n pages of cipher text, each printed on

A synopsis based on the Key-Note lecture delivered by Mausumi Bose at IPECS 2018 at Pondicherry University
Corresponding author: Mausumi Bose
Email: mausumi.bose@gmail.com

a transparency sheet. These n sheets are distributed among n participants. Each single share is indistinguishable from random noise. The image can be visually decoded if any k of these sheets are stacked on top of one another, while this is not possible by stacking any $k-1$ or fewer sheets. Thus, visual cryptography concerns the encoding of a secret image in such a way that the decoding may be done simply by the human visual system without elaborate computations as required in usual cryptographic schemes.

Here, for simplicity, let us first focus on $k = 2$. We consider black and white images only, and thus the secret image is a collection of black and white pixels. During encryption, each pixel of the original image is encrypted into a number of subpixels which are distributed to the n shares. For any VCS, two important features are its *pixel expansion* (i.e., the number of subpixels in each share of a VCS which are needed to encode one pixel of the original image) and the *relative contrast* (i.e., the clarity with which a reconstructed image is visible.) One would like to have a small pixel expansion and a large relative contrast for a VCS.

Two $n \times m$ Boolean matrices, one for black pixels and another for white pixels, say B and W, respectively, are suitably constructed to perform the encryption for a VCS, m being the pixel expansion. While W is of a simple form, it has been shown how partially balanced incomplete block (PBIB) designs may be used to construct the matrix B which lead to VCS with varying relative contrasts, some of them attaining the maximum value possible for such contrasts. Next, after introducing three meaningful optimality criteria for evaluating different schemes, it has been been demonstrated how several broad classes of combinatorial designs, such as balanced incomplete block designs, PBIB designs and regular graph designs, can yield a large number of black and white $(2, n)$ schemes that are optimal with respect to all these criteria. For a practically useful range of n, optimal schemes can be obtained with the smallest possible pixel expansion.

Generalizing to the case of $k \geq 2$, it has been shown that a Kronecker algebra can be used to obtain necessary and sufficient conditions for the existence of a (k, n) VCS for given levels of relative contrasts. A connection of these conditions with an L_1-norm formulation as well as a convenient linear programming formulation has been established to settle certain conjectures on contrast optimal VCS for the cases k = 4 and 5. Furthermore, for $k = 3$, block designs can be used to construct VCS which achieve optimality with respect to the average and minimum relative contrasts but require much smaller pixel expansions than the existing VCSs..

This section is based on the papers Adhikari and Bose (2004), Adhikari et al. (2007), Bose and Mukerjee (2006, 2010). Two open problems are of interest:

1. Does there exist a unified method that may yield optimal VCS, in the senses considered here, for arbitrary n and m?

2. How does one extend the results on optimal black and white VCS in the above papers to color VCS?

3 Anti-collusion Digital Fingerprinting

Digital fingerprinting is a technique for tracing consumers who use their multi-media contents for illegitimate purposes, such as redistribution. A digital fingerprint can be thought of as a word (vector) of length v over an alphabet of size m. These fingerprints are embedded in multimedia content (through a variety of watermarking techniques), each user being sold copy of the data uniquely marked according to his assigned unique codeword. However, two or more users with the same content but with different marks can collude to combine their information and generate a new illegitimate version of the content for unauthorized redistribution in which the original identifying fingerprints are removed.

Anti-collusion codes (ACCs) aim at deterring such unauthorized utilization by a coalition of users, and have been of considerable recent interest. Let X be a code of n binary code vectors, each of length v and let these code vectors in X be used to watermark the digital contents of n consumers. Then, if some of these consumers collude and use their own contents to produce an illegitimate content, then the watermark of this illegitimate content can be detected in the form of the element-wise AND of the code vectors used for these colluders. Hence, an attractive class of ACCs are called AND-ACCs where, the code X is called a K-resilient AND-ACC if the element-wise ANDs of all distinct subsets of K or fewer code vectors in X are distinct. Therefore, if the number of colluders is K or fewer, then they can be uniquely identified whenever X is a K-resilient AND-ACC.

For a given resilience K , one prefers an AND-ACC with relatively large n(users) and small v and so, given K , the efficiency of an AND-ACC is measured by $\beta = n/v$, a larger β implying higher efficiency.

In Bose and Mukerjee (2013), it has been shown that the duals of certain block designs are useful in this context with the columns of the incidence matrices of these duals being used to form the AND-ACC codes. Moreover, orthogonal arrays of index unity and partially cover-free codes also give very efficient AND-ACC codes. Again, on introducing a new concept of partial cover-free families, it has been shown how they can be used to construct AND-ACCs with larger β than the AND-ACCs available in the literature till then.

This section is based on Bose and Mukerjee (2013). Two open problems are of interest:

1. Can one find further improved AND-ACCs based on code-vectors that do not all have the same weight?

2. Will it help if we start from suitable designs with blocks of variable sizes?

4 Key Pre-distribution in Distributed Sensor Networks

Distributed sensor networks (DSN) are used in civilian and military contexts where sensor nodes are distributed in a random manner over a sensitive area and, once deployed, these nodes

need to communicate with each other in order to collect and relay information. It is important that this communication be done in a secret manner and so secure keys need to be established between the nodes in the system. The use of key pre-distribution schemes (KPSs) is a popular method where secret keys are installed in each sensor node before the nodes are deployed.

The key metrics for a KPS are the network size (i.e., number of nodes in the network), key ring size (i.e., number of keys used in the network), storage (i.e., number of keys per node), and the intersection threshold q (where two nodes can communicate only if they have q common keys). In addition, there are also the connectivity measures of interest, namely, (a) Pr_1 =Probability that any two nodes can communicate directly and (b) Pr_2: Probability that any two nodes cannot directly communicate but can do so via a third node, and (c) $\mathrm{Pr} = \mathrm{Pr}_1 + \mathrm{Pr}_2$: Probability that any two nodes can communicate either directly or via one-hop.

There are other metrics of a KPS which are related to an attack on the network. In an attack, some nodes are compromised and all keys in the captured nodes are revealed and cannot be used for communication anymore. A good KPS needs to be resilient to such attacks and this is measured by some measures, one of them being the resiliency measure: fail(s)= Probability that nodes A and B fail to directly communicate given that any s other nodes are compromised, where A and B are two uncompromised nodes with q keys in common.

There are two other properties of a KPS of interest based on the level of difficulty in communication. Two nodes in the same neighborhood need to determine if they share q common keys before they can communicate, this is the shared-key discovery phase; and if they do not, then they try to establish a secure one-hop path for communication; this is the path-key establishment phase. The difficulties involved in these two phases are also used to assess the usefulness of a KPS.

In Bose et al. (2013), a new method for constructing KPSs based on a suitable combination of partially balanced incomplete block designs has been proposed and the properties of the resulting schemes have been studied. Unlike previous schemes, all the metrics have been studied for the proposed KPSs. Interestingly, this one general method of construction works for any given intersection threshold q and by varying the choices of the component designs, one can construct KPSs for distributed sensor networks with varying numbers of nodes, key-pool sizes and numbers of keys per node. Thus, this method provides more flexibility in choosing a scheme suitable for the requirements of a situation. Moreover, this method allows for a large number of nodes, while keeping the key ring size (k) in check.

Another advantage of this method of construction over many other schemes in the literature is that it allows one to obtain unified and explicit algebraic expressions for the metrics for evaluating the connectivity and resiliency of these schemes, all for general values of $q(\geq 1)$. Using these expressions, the metrics can be easily calculated from the parameters of the particular designs used in the construction. It has also been shown that the KPSs obtained by this method have good connectivity with high levels of resiliency and the combinatorial structure of the underlying designs make the shared-key discovery and path-key establishment phases particularly simple. The many advantages of this method compared to the other existing methods have also been established via a large number of examples.

This section is based on Bose et al. (2013).

5 Anonymity in Shared Key Operations

Unequal probability schemes have a long history in estimating parameters in finite population sampling. This has been applied to a new area of application, namely shared key operations where the issue is not estimation or variance reduction, but ensuring greater anonymity. It has been shown how the use of unequal probability selection can result in appreciable gains in anonymity.

In shared key operations there are a number of keys, say, v, each key consisting of multiple components. These key components are to be distributed among n participants, such that (a) every group of t participants has access to at least one key when the members of the group pool their key components together, while (b) no group of $t-1$ or fewer participants has access to any key even after such a pooling. A distribution of key components, satisfying (a) and (b), is called a (t, n)-threshold access structure (TAS). So, every group of t participants will collectively have at least one key for use.

Now, given any (t, n)-TAS, any one group of t participants is activated by probabilistic selection. The t members of this activated group then pool their key components to perform a shared key operation, such as encryption or authentication, using a key at their collective disposal.

Two issues of anonymity are crucial here. An adversary must be prevented from learning which group of participants performed the operation (group anonymity) and whether an individual participant was a member of that group (participant anonymity). These anonymities should hold even if the adversary has knowledge about the way the key components were distributed among the participants and also knows the key used for the shared operation. These group and participant anonymities can be quantified via appropriate conditional probabilities, depending on the selection mechanism for a group of participants from the TAS.

While the existing literature had advocated that all groups have the same probability of being selected for activation, it has been shown in [14] how unequal probability selection can perform better in the sense of leading to enhanced anonymities, both for the group and the participants.

Perfect hash families (PHFs) have been seen to play a key role in constructing these TASs, the combinatorial structures underlying the PHFs ensuring the requirements of a TAS. Then, introducing the method that any group of t participants uses a key with probability proportional to the number of keys at its disposal, and next, if a group uses a key, then it uses any of its available keys with equal probability, it has been shown that the resulting TAS has better values for both the anonymity measures. For implementing this method, a simple algorithm has been given based on Lahiri's (1951) well-known method of sampling with probability proportional to size. The two anonymity measures have been derived in terms of the parameters of the underlying PHF, thus making it possible to choose a suitable PHF for use in a given context.

The improvement in anonymity by using this method continues to hold whether the underlying

PHF is balanced or not. When a balanced PHF is used, this proportional scheme has some further attractive features. Moreover, it has been shown that this proportional scheme works even for (t, n)-TASs which are not obtained through PHFs and examples are given to highlight that these schemes can again outperform equal probability schemes.

This section is based on Bose and Mukerjee (2014). An open problem is to further explore the behavior of the proportional schemes based on structures other than PHFs.

References

Adhikari, A. and Bose, M. (2004). Construction of new visual threshold schemes using combinatorial designs. *IEICE Transactions,* **5**, 1198-1202.

Adhikari, A., Bose, M., Kumar, D. and Roy B.(2007). Applications of partially balanced incomplete block designs in developing (2, n) visual cryptographic schemes. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences,* E90-A, 949-951.

Blundo, C., De Santis, A. and Stinson, D.R. (1999). On the contrast in visual cryptography schemes. *Journal of Cryptology,* **12**, 261-289

Blundo, C., De Santis, A. and Naor, M. (2001). Visual cryptography for grey-level images. *Information Processing Letters,* **75**, 255-259

Boneh, D. and Shaw, J. (1998). Collusion-secure fingerprinting for digital data. *IEEE Transactions on Information Theory,* **44**, 1897-1905.

Bose, M. and Mukerjee, R. (2006). Optimal (2, n)visual cryptographic schemes. *Designs, Codes and Cryptography* **40**, 255-267.

Bose, M. and Mukerjee, R.(2010). Optimal (k, n)visual cryptographic schemes for general k. *Designs, Codes and Cryptography* **55**, 19-35.

Bose, M. and Mukerjee, R. (2013). Union distinct families of sets, with an application to cryptography. *Ars Combinatorica,* **110**, 179-192.

Bose, M., Dey, A. and Mukerjee, R. (2013) Key predistribution schemes for distributed sensor networks via block designs. *Designs, Codes and Cryptography*, **67**, 111-136.

Bose, M. and Mukerjee, R. (2014). An unequal probability scheme for improving anonymity in shared key operations. *Journal of Statistical Theory and Practice. Special Issue: Design of Experiments and Related Combinatorics in Memory of Professor Jagdish N. Srivastava. Part I.* **8**, 100-112.

Hedayat, A. S., Sloane N. J. A. and Stufken, J. (1999). *Orthogonal Arrays: Theory and Applications*. New York: Springer-Verlag.

Hofmeister, T., Krause, M. and Simon, H. U. (2000) Contrast-optimal k out of n secret sharing schemes in visual cryptography. *Theoretical Computer Science*, **240**, 471-485.

Iwamoto, M. (2008). *Weakly secure visual secret sharing schemes.* In: *Proc. ISITA*, 42-47. Auckland.

Kang, I., Sinha K. and Lee, H. K.(2006). New digital fingerprint code construction scheme using group-divisible design. *IEICE Trans. Fundamentals E89-A*, 3732-3735.

Koga, H. and Ueda, E. (2006). Basic properties of the (t , n)-threshold visual secret sharing scheme with perfect reconstruction of black pixels. *Des. Codes Crypt.* 40, 81-102.

Koga H., Iwamoto M., Yamamoto, H.(2001). An analytic construction of visual secret sharing scheme for color images. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* **E84-A**, 262-272.

Li, Q., Wang, X., Li, Y., Pan, Y. and Fan, P. (2009). Construction of anti-collusion codes based on cover-free families. *Sixth International Conference on Information Technology*: New Generations, ITNG 2009, Las Vegas, USA.

Lahiri, D. B. (1951). A method of sample selection providing unbiased ratio estimates. *Bulletin of International Statistical Institute*, **33**, 133-146.

Lee, J. and Stinson, D. (2005). *A combinatorial approach to key predistribution for distributed sensor networks*. In: *IEEE Wireless Communications and Networking Conference* (WCNC05), **2**, 1200-1205. IEEE Communications Society

Lee, J. and Stinson, D. (2005). *Deterministic key predistribution schemes for distributed sensor networks.* In: *SAC 2004 Proceedings. Lecture Notes in Computer Science,* **3357**, 294-307. Springer, New York.

Lee, J. and Stinson D. (2008). On the construction of practical key predistribution schemes for distributed sensor networks using combinatorial designs. *ACM Transactions on Information and System Security* **11**, 1-35.

Martin, K. (2009). *On the applicability of combinatorial designs to key predistribution for wireless sensor networks.* In: *Proceedings of the 2nd International Workshop on Coding and Cryptology.* Springer, Berlin.

Martin, K., Blackburn, S. R., Etzion, T. and Paterson, M.B. (2010). Distinct difference configurations: multihop paths and key predistribution in sensor networks. *IEEE Transactions on Information Theory* **56**, 3961-3972.

Martin, K., Stinson, D. R. and Paterson, M. B. (2011) Key predistribution for homogeneous wireless sensor networks with group deployment of nodes. *ACM Transactions on Sensor Networks.* **7** (2).

Raghavarao, D. (1971). *Constructions and Combinatorial Problems in Design of Experiments.* Wiley, New York.

Trappe, W., Wu, M., Wang, Z. J. and Liu, K. J. R. (2003). Anti-collusion finger-printing for multimedia. *IEEE Transactions on Signal Processing.* **51**, 1069-1087.

Yagi, H., Matsushima, T. and Hirasawa, S. (2007). Improved collusion-secure codes for digital fingerprinting based on finite geometries. *IEEE International Conference on System, Man and Cybernetics*, 948-953.

Walker, R. A. II. (2012). PHFtables.com. www.phftables.com. Accessed January, 2012.

Zaverucha, G. M. and Stinson, D. G. (2010). Anonymity in shared symmetric key primitives. *Designs, Codes and Cryptography*, *57*, 139-160.