Special Proceedings: ISBN #: 978-81-950383-8-1
27th Annual Conference, 21-23 February 2025; pp 77-86

Use of Designs for Statistical Experiments in Constructing Cryptographic Schemes

Mausumi Bose

Indian Statistical Institute and St. Xavier's College, Kolkata

Received: 23 June 2025; Revised 08 July 2025; Accepted 10 July 2025

Abstract

In this article we highlight how the combinatorial properties of statistical designs of experiments have been used by many researchers for constructing various types of cryptographic schemes. In particular, we discuss key predistribution schemes for distributed sensor networks in some detail and show through examples, how useful schemes can be constructed from the duals of certain block designs.

Key words: Balanced incomplete block designs; Partially balanced incomplete block designs; Steiner's triple systems; Distributed sensor networks; Resilience.

AMS Subject Classifications: 05B05, 94A6.

1. Introduction

Combinatorial structures of different kinds have been extensively studied over the years by mathematicians, for example, Hadamard Matrices, orthogonal arrays, Latin squares, Steiner's triple systems, *etc.* The construction and existence of these structures have been well-developed and a considerable literature is available on such structures.

Later, statisticians found that many of these structures are useful in the field of Design of Experiments. Subsequently, optimality properties of the designs based on these structures, were also proved. For example, it was found that Hadamard matrices were useful in obtaining optimal weighing designs using the chemical balance, Steiner's triple systems were useful as incomplete block designs for one-way elimination of heterogeneity, Latin squares and mutually orthogonal Latin squares were useful as optimal designs for eliminating heterogeneity in two or three directions, orthogonal arrays were useful in obtaining fractional factorial designs, and the list goes on. For a comprehensive discussion on these designs, their combinatorial properties and construction, and their statistical optimality aspects, we refer to Raghavarao (1971), Street and Street (1987), Shah and Sinha (1989) and Hedayat, Stufken and Sloane (1999).

Much later, cryptographers found that many of these statistical designs of experiments

Corresponding Author: Mausumi Bose
Email: mausumi.bose@gmail.com

based on combinatorial structures can also be used to generate good cryptographic schemes. For some details of such use, we refer to Stinson (2004) and Stinson and Patterson (2023).

Cryptography is the practice of scrambling communications so that only the intended recipient can access them. In modern times, cryptography is used to protect confidentiality of sensitive information and protect it from hackers and other cyber criminals. It can be used to obscure various forms of digital communication, including text, images, video, or audio; protect confidentiality and integrity in communication. *e.g.*, computer passwords, email, online transactions, transmitting confidential information, *etc.* For a historical perspective of the development of the subject since ancient to recent times, we refer to Kahn (1996) and Bauer (2021). For a technical perspective of some schemes, we refer to Stinson and Patterson (2023).

In this paper we mention some cryptographic schemes which can be obtained from combinatorial structures. In Section 1, we give a brief description of two such schemes and mention the combinatorial structures which lead to these schemes. In Section 2 we focus on distributed sensor networks and describe how they can be obtained from statistical designs. References are given for all these results and the reader may obtain the details from these references.

2. Some cryptographic schemes and related combinatorial structures

In this section we mention two cryptographic schemes, error-correcting codes and visual cryptographic schemes, and mention the designs that may be used to construct these schemes. Our objective here is to only give a flavor of the versatility of the application of designs to cryptography. There are many other schemes which are not mentioned here for the sake of brevity.

2.1. Error correcting codes and Hadamard matrices

Error-correcting codes are used to detect and correct errors that can occur when transmitting data over noisy channels. They add extra bits, *i.e.*, redundant information, to the original data in such a way that the recipient of the data can compare the received data with the redundancies and identify the errors which arise due to noise or other factors. Each code is a collection of codewords, or k-tuples, say, with symbols from a set of symbols or an alphabet.

It is well known that optimal weighing designs are given by Hadamard matrices, *e.g.*, to optimally weigh 8 objects using 8 weighings with a chemical balance, the optimal design matrix will be given by a Hadamard matrix of order 8. In the cryptography context, the rows of this same Hadamard matrix, after replacing -1 by 1 and 1 by 0, will give an error-correcting code for transmitting a binary message of 3 bits as a message of 8 bits, and it can correct one error. More generally, using Hadamard matrices, one can construct the first-order Reed-Muller code over the binary alphabet which is useful in transmitting messages over noisy channels. For some applications in this context, we refer to Serberry, Wysocky and Wysocki (2005) and Yarlagadda and Hershey (1997).

2.2. Error-correcting codes and orthogonal arrays

Orthogonal arrays are well-known structures which are useful in statistics for obtaining suitable fractions of factorial experiments for experimentation. These orthogonal arrays also give useful error-correcting codes, namely MDS (Maximum distance separable) codes, the Reed Solomon Codes, Hamming codes, *etc.* These codes have optimal properties of various kinds.

2.3. Visual cryptographic schemes and BIBD, PBIBD

In a (k, n) visual cryptography scheme, a secret image (or text) is encoded to form n *'shares'* and each share is printed on a transparency sheet. There are n participants, each of whom gets one share. The encryption is such that only when $k(\geq 2)$ participants get together and stack their sheets one above another, the secret image is revealed, no set of $k-1$ or fewer participants can decode the secret image. This scheme is useful as decoding can be done simply by the human eye without the need of any computers or equipment. More details can be obtained from Naor and Shamir (1994) and Kang, Arce and Lee (2011), Ibrahim, Teh and Abdullah (2021) and Climato, Prisco and Santis (2005).

It has been shown in Blundo, Santis and Stinson (1999) that balanced incomplete block designs (BIBDs) are useful in encoding the secret image and forming the shares. Adhikari and Bose (2004) and Adhikari, Bose, Kumar and Roy (2007) showed that partially balanced incomplete block Designs (PBIBDs) lead to schemes where the sharpness of the recovered image is better for certain set of participants. Bose and Mukerjee (2006, 2010) showed that various other incomplete block designs like regular graph designs, symmetrical unequal block designs may also be used to obtain schemes with many desirable properties.

There are several other schemes in the literature which have been developed from designs of experiments and which have not been mentioned here, *e.g.*, general threshold access structures, anti-collusion digital fingerprinting, *etc.* Some references on these are Kang, Sinha and Lee (2006), Yagi, Matsushima and Hirasawa (2007), Bose and Mukerkee (2013, 2014). Moreover, there could be many other possibilities of using designs to construct useful cryptographic schemes of various types in future.

3. Distributed sensor networks

Distributed sensor networks (DSNs) are used in a wide range of applications. Some examples of their use are in air quality monitoring, water quality monitoring, wildlife tracking, seismic activity detection *etc.* These are also used in military applications such as battlefield surveillance, target tracking, perimeter security, reconnaissance missions, *etc.* Another interesting use of this system is in smart cities where they prove useful in traffic management, congestion monitoring, parking availability detection, street lighting control, *etc.*

This wide applicability of these network schemes is due to the ability of DSNs to collect real-time data from geographically dispersed sensors, enabling comprehensive monitoring and analysis of various physical phenomenon across large areas.

We now discuss key-predistribution schemes for DSNs in some detail, based on results from Bose, Dey and Mukerjee (2013); more references may be found therein.

3.1. Key predistribution schemes(KPS) for DSNs

We begin with an example of a situation where sensor nodes are pre-distributed in several locations. Suppose in a military operation, several sensor nodes, each with some secret keys installed in them, are randomly scattered over a sensitive area. The keys in each node are taken from a large set of keys. Each node can send or receive signals only over a certain wireless communication range or neighbourhood. Once deployed, these nodes have to communicate with each other through secure keys in order to gather and relay information.

In this context, some metrics of the KPS are important:

1. Network size, *i.e.*, the number of nodes deployed, say, n.
2. Key storage, *i.e.*, the number of keys stored per node, say, k.
3. Intersection Threshold *i.e.*, the number of keys common between 2 nodes, say, q.
4. Communication rule *i.e.*, if two nodes are within each other's neighbourhood, they can communicate with each other
   (a) directly, if they have $q \geq 1$ common keys, or
   (b) *via* one hop if there is a third node within the intersection of their neighbourhoods which shares q common keys with each of them. If needed, multiple secure links can also be used if there is a sequence of nodes connecting them such that every pair of successive nodes in this sequence share $q(\geq 1)$ common keys.

Now, after deployment, some nodes may be captured in an attack. In that case, all the keys in these captured nodes are considered to be lost and cannot be used for communication by the other nodes. However, if the remaining nodes can still communicate using their remaining keys as per 4 (a) or 4 (b) above, then the KPS is said to be 'resilient'. Resilience is a desirable property of a KPS.

3.2. Correspondence with block designs

Now we introduce a correspondence between some terms used in the context of block designs with the terms used in the context of the KPS as introduced in section 3.1.

The set of all keys of the KPS corresponds to the set of all treatments in a block design.

The sensor nodes of the KPS correspond to the blocks in a block design. Here, since we would like a large number of nodes in the system, we need a large number of blocks in the designs, as opposed to fewer blocks preferred in designs of experiments.

The key storage of a KPS corresponds to the block size of a design.

With the above correspondence, it is clear that the 'intersection threshold' of a KPS corresponds to the number of treatments that are common to two blocks. This means that the block intersection number of a block design becomes important. We will consider the duals of block designs where the roles of treatment and block in the original block design are reversed, and so, the incidence between the treatments and blocks is also reversed.

We give examples of two designs, one PBIBD(d_1) wih 2 associate classes, and one BIBD(d_2), and their corresponding dual designs d_1^* and d_2^*, as shown below. These duals will be used subsequently to construct KPS. For the design d_i, let v_i, b_i, r_i and k_i denote the number of treatments, number of blocks, replication number, and block size, respectively, $i = 1, 2$. Then, from the properties of BIBD and PBIBD it may be noted that these duals $d_i^*,\ i = 1, 2$, are such that:

1. every symbol occurs at most once in any block
2. every symbol occurs in k_i blocks, $2 \leq k_i < b_i$
3. every block contains r_i symbols, $v_i > r_i \geq 2$, and
4. there is an association scheme with 2 associate classes on the sets of blocks of d_i^*, $i = 1, 2$. Any 2 distinct blocks will either have no symbol in common (then we call these blocks 1st associates of each other) or they will have exactly one symbol in common (then we call these blocks 2nd associates of each other). Each block is called the 0th associate of itself. Clearly, any 2 distinct blocks of d_2^* will be 2nd associates, while any two distinct blocks of d_1^* may be either 1st or 2nd associates.

Example 1: PBIB design with GD scheme $d_1(v_1 = 6, b_1 = 9, r_1 = 3, k_1 = 2, \lambda_1 = 0, \lambda_2 = 1)$, blocks shown as columns labeled $1, \ldots, 9$.

	1	2	3	4	5	6	7	8	9
d_1 :	1	1	1	2	2	2	3	3	3
	4	5	6	4	5	6	4	5	6

Dual of d_1: $d_1^*(v_1^* = 9, b_1^* = 6, r_1^* = 2, k_1^* = 3)$, blocks shown as columns labeled $B_1, \ldots, B_6$.

	B_1	B_2	B_3	B_4	B_5	B_6
d_i^* :	1	4	7	1	2	3
	2	5	8	4	5	6
	3	6	9	7	8	9

Example 2: BIB design d_2 $(v_2 = 9, b_2 = 12, r_2 = 4, k_2 = 3, \lambda = 1)$, blocks shown as columns labeled $1, \ldots, 12$.

	1	2	3	4	5	6	7	8	9	10	11	12
d_2 :	4	7	1	5	8	2	6	9	3	1	4	7
	7	1	4	8	2	5	9	3	6	2	5	8
	2	5	8	3	6	9	1	4	7	3	6	9

Dual of d_2: $d_2^*(v_2^* = 12, b_2^* = 9, r_2^* = 3, k_2^* = 4)$, blocks shown as columns labeled $C_1, \ldots, C_9$.

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9
	2	1	4	1	2	5	1	3	6
d_2^* :	3	5	8	3	4	7	2	4	7
	7	6	9	8	6	9	9	5	8
	10	10	10	11	11	11	12	12	12

3.3. Use of block designs

We can build useful key predistribution schemes based on block designs because using the combinatorial structures of the designs we can

1. study the connectivity property of the scheme
2. study the resilience property of the scheme, and
3. carry out shared-key discovery and path-key establishment in a structured manner.

Schemes are evaluated on the basis of their connectivity and resilience using the measures Pr_1, Pr_2 and $fail(s)$ as proposed by Lee and Stinson (2004) and defined below:

For any 2 randomly chosen nodes in each other's neighbourhood, let Pr_1 be the probability that the 2 nodes can securely communicate directly with each other, *i.e.*, they have q keys in common.

Again, for any 2 randomly chosen nodes in each other's neighbourhood, let Pr_2 be the probability that these 2 nodes do not share q common keys but there is a third key in the neighbourhood of both of them which shares q common keys with both these nodes. So these 2 nodes can communicate securely via this third node.

Finally, $Pr_1 + Pr_2$ is used to study the connectivity of a KPS, either through a secure direct path, or through a secure path *via* a third node. The larger the value of $Pr_1 + Pr_2$, the better is the connectivity of the KPS.

In the event of an attack a number of nodes are compromised and the keys in the compromised nodes are rendered unusable for communication. Let A and B be 2 uncompromised nodes which share q common keys. Then, the resilience of the KPS is measured by $fail(s)$ which is equal to the conditional probability that the link between A and B will fail, when out of the other $n-2$ nodes, s randomly chosen nodes are compromised. A smaller $fail(s)$ means a larger resilience property for the KPS.

Several researchers have studied this problem. Lee and Stinson (2004) considered KPS with $q = 1$ and $q = 2$ and used transversal designs for their construction. Bose, Dey and Mukerjee (2013) studied KPS for general q and used various types of designs for their construction, *e.g.*, BIBD, PBIBD based on GD, LS and triangular association schemes, and suitable duals of these designs, for general q.

3.4. An illustration of the construction of KPS for $q = 2$

We now illustrate how duals of some suitable block designs can be used in the construction of the schemes. For our illustration, we use the designs shown in Section 3.2. For more examples, details and theoretical justifications, we refer to Bose, Dey and Mukerjee (2013). We only consider the case where $q = 2$; the case with $q = 1$ is easier and omitted here.

We can onstruct a KPS with $q = 2$ as follows:

(1) We start with 2 designs, each being either a PBIB design with $\lambda_1 = 0, \lambda_2 = 1$, or a BIB design with $\lambda = 1$, and then we consider their dual designs. *e.g.*, we start with d_1 and d_2 shown in Section 3.2 and take their duals d_1^* and d_2^*.

(2) We identify the symbols of d_1^* and d_2^* as the keys. So, the number of possible keys is $v_1^* + v_2^* = b_1 + b_2$, which equals to $9 + 12 = 21$ keys in our example.

(3) We take all possible selections of a block from each of d_1^* and d_2^*, and consider their union as a node. So, any node in our example is of the form: $B_i \cup C_j,\ i = 1, \ldots, 6,\ j = 1, \ldots, 9$. Thus we get the number of nodes as $n = b_1^* \times b_2^* = v_1 \times v_2$, which equals $6 \times 9 = 54$ nodes in our example. Each of these nodes have $k_1^* + k_2^* = r_1 + r_2$ keys, which equals $3+4 = 7$ keys in our example.

We can check the properties of the KPS from the properties of the constituent designs.

For example, by taking the union of block B_1 from d_1^* and block C_1 from d_2^*, and writing the symbols of d_2^* in italics to differentiate them from the symbols of d_1^*, we get the node as

$$B_1 \cup C_1 = 1\ 2,\ \ 3,\ \ \textit{2},\ \ \textit{3},\ \ \textit{7},\ \ \textit{10}$$

Similarly, taking union of block B_3 from d_1^* and block C_4 from d_2^*, and writing the symbols of d_2^* in italics, we get the node as

$$B_3 \cup C_4 = 7,\ \ 8,\ \ 9,\ \ \textit{1},\ \ \textit{3},\ \ \textit{8},\ \ \textit{11}$$

Note that B_1 and B_3 have no symbol in common and hence these blocks are 1st associates of each other. Again, blocks C_1 and C_4 have 1 symbol in common and hence these blocks are 2nd associates of each other. So we will say that the 2 nodes given by $B_1 \cup C_1$ and $B_3 \cup B_4$ are *12*th associates of each other.

Now, since d_1 is a PBIB design with 2 associate classes, blocks of d_1^* can be either 0, 1, or 2 associates. Again, as d_2 is a BIB design, blocks of d_2^* can be either 0, or 2 associates. So the association relationship between any 2 *distinct* nodes $B_{i_1} \cup C_{j_1}$ and $B_{i_2} \cup Cj_2$ in this KPS will be given by the set

$$\{02, 10, 12, 20, 22\}$$

Using this association structure between two nodes, we can deduce which two nodes can directly communicate with each other and which two nodes need a path via a third node to communicate.

It can be shown that with $q = 2$, all pairs of nodes except those which are *12* associates of each other can communicate directly with one another.

In this example, it can be checked that the number of *12* associates of any node in the KPS is 16. So the remaining $54 - 16 = 38$ nodes can directly communicate with each other.

Algebraic expressions for Pr_1, Pr_2 and $fail(s)$ can also be obtained using the combinatorial properties of the component designs. We omit the details here.

3.5. Evaluating Local connectivity and resilience for the above KPS

For the KPS obtained from the designs d_1^* and d_2^*, it may be shown that:

$$Pr_1 = 0.6981, \quad Pr_2 = \frac{16}{53}[1 - (1 - \frac{29}{52})^{\eta}]$$

where the intersection of the neighbourhoods of nodes A and B contain η nodes, excluding A and B themselves. So for $q = 2$ and for some choices of η, the probability that any 2 randomly chosen nodes in the KPS can communicate with each other is equal to

η	1	2	3	4	5	10	15	20
$Pr_1 + Pr_2$	0.867	0.941	0.974	0.988	0.995	0.9999	1.000	1.000

The above table shows that this KPS has quite high local connectivity. Different choices of the constituent designs will lead to different KPS and their metrics can be computed.

This idea of construction for $q = 2$ can be extended to general $q(\geq 2)$ where we start with q suitable initial designs, take their duals, and then form KPS as in steps (1), (2) and (3) in Section 3.3. Each time, n is multiplicative in the b_i^* while k is additive in k_i^*. Thus, this method gives schemes with many nodes but small key storage. The properties of such KPS can be similarly ascertained from the properties of the designs.

Conflict of interest

The authors do not have any financial or non-financial conflict of interest to declare for the research work included in this article.

References

Adhikari, A. and Bose, M. (2004). Construction of new visual threshold schemes using combinatorial designs. *IEICE Transactions*, **E87-A**, 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*, **E90-A**, 949-951.

Bauer, C. P. (2013). *Secret History: The Story of Cryptology.* Chapman and Hall/CRC.

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

Bose, M, Dey, A., and Mukerjee, R. (2013). Key predistribution schemes distributed sensor networks via block designs. *Design, Codes and Cryptography*, **467**, 111-136.

Bose, M. and Mukerjee, R. (2006). Optimal $(2, n)$ visual cryptographic schemes. *Design, 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 Combinatoria,* **110**, 179-192.

Bose, M. and Mukerjee, R. (2014). An unequal probability scheme for improving anonymity in shared key operations. *Journal of Statistical Theory and Practice*, **8**, 100-112.

Bose, R. C. and Shrikhande, S. S. (1959). A note on result in the theory of code construction. *Information and Control*, **2**, 183–194.

Bush, K. A. (1952). Orthogonal arrays of index unity. *Annals of Mathematical Statistics*, **23**, 416-434.

Clatworthy, W. H. (1973). *Tables of Two-associate Partially Balanced Designs*. National Bureau of Standards, Applied Maths, series no **63**, Washington D.C.

Climato, S., Prisco, R. D., and Santis, A. D. (2005). Optimal coloured threshold visual cryptography schemes. *Designs Codes and Cryptography*, **35**, 311-335.

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

Ibrahim, D. R., Teh, J. S., and Abdullah, R. (2021). An overview of visual cryptography techniques. *Multimedia Tools and Applications*, **80**, 31927-31952.

Kahn, D. (1996). *The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet*. Simon and Schuster.

Kang, I., Arce, G., and Lee, H. K. (2011). Color extended visual cryptography using error diffusion. *IEEE Transactions on Image Processing*, **20**, 132-145.

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

Lee, J. and Stinson, D. R. (2004). Deterministic key predistribution schemes for distributed sensor networks. *SAC 2004 Proceedings, Lecture notes in Computer Science*, **3357**, 294-307.

Mukerjee, R. and Wu, C. F. J. (2006) *A Modern Theory of Factorial Design.* Springer, New York.

Naor, M. and Shamir, A. (1994). Visual cryptography. Advances in Cryptology, Eurocrypt'94. Lecture Notes in Computer Science, **950**, 1-12, Springer-Verlag.

Plotkin, M. (1960). Binary codes with specified minimum distance. *IRE Transactions*, **IT-6** 445–450.

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

Rao, C. R. (1947). Factorial experiments derivable from combinatorial arrangements of arrays. *Journal of Royal Statistical Society*, **9**, 128-139.

Rao, C. R. (1949). On a class of arrangements. *Proceedings of Edinburgh Mathematical Society*, **8**, 119-125.

Serberry, J., Wysocky, B. J., and Wysocki T. A. (2005). On some applications of Hadamard matrices. *Metrika*, **62**, 221-239.

Shah, K. and Sinha, B. K. (1989). *Theory of Optimal Designs*. Springer-Verlag, New York.

Stinson, D. R. (2004). *Combinatorial Designs / Constructions and Analysis*. Springer, ISBN 978-0-387-95487-5.

Stinson, D. R. and Paterson, M. B. (2023). *Cryptography: Theory and Practice*. 4th ed. CRC Press.

Street, A. P. and Street, D. J. (1987). *Combinatorics of Experimental Design.* Oxford. Clarendon Press.

Takeuchi, K. (1962). A table of difference sets generating balanced incomplete block designs. *Review of International Statistical Institute*, **30**, 361-366.

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.

Yarlagadda, R. K. and Hershey, J. E. (1997). *Hadamard Matrix Analysis and Synthesis: with Applications to Communications and Signal/Image Processing*. Kluwer.