Asymptotic Enumeration of Binary Orthogonal Arrays*

Konstantin N. Pankov
Department of Postgraduate Studies
Moscow Technical University of Communications and Informatics
Moscow, Russia
pankov_kn@mtuci.ru

***Abstract*. Orthogonal arrays are basic combinatorial structures, which appear in various cryptographic applications designed to ensure information security in electronic financial services. Asymptotic estimates for the number of orthogonal arrays are proved in this paper.**

***Key words:* cryptography, authentication codes, orthogonal arrays, asymptotic enumeration.**

I. INTRODUCTION

It was proposed launching a large-scale system-wide program to develop an economy of a new technological generation, the so-called digital economy, in the President Putin annual address to Federal Assembly in December 2016 [1]. In compliance with the list of instructions of the President for the implementation of the Address in 2017, the Ministry of Communications and Mass Media of Russia prepared the program "Digital Economy of the Russian Federation" [2]. The program was approved by the Government of the Russian Federation in its resolution No. 1632-r dated July 28, 2017. The Sub-commission for digital economy of the governmental commission for the use of information technologies to improve the quality of life and the conditions of doing business was established by the Government of the Russian Federation in its resolution No. 969 dated August 15, 2017. Maxim Akimov, the Deputy Prime Minister of the Russian Federation, heads up the Sub-commission.

According to the roadmap of the program in the field of regulatory regulation, the task is to ensure the legal conditions for the introduction and use of innovative technologies in the financial market. To accomplish this task, it is supposed to promote the introduction and use of innovative technologies in the financial market, including improving the mechanisms for providing financial services in electronic form and ensuring their information security.

The most powerful tool for ensuring information security in the modern world is the use of cryptography. Cryptographic tools have become an integral part of the services provided in electronic form by financial and credit institutions. Cryptographic algorithms based on the use of mathematical methods for transforming protected information solve the problem of ensuring confidentiality, integrity, authentication, non-repudation and untraceability.

In ISO 27000 [3], confidentiality is the property, that information is not made available or disclosed to unauthorized individuals, entities, or processes. Confidentiality is a component of privacy that implements to protect data from unauthorized viewers. Data integrity means maintaining and assuring the accuracy and completeness of data over its entire lifecycle. This means that data cannot be modified in an unauthorized or undetected manner. Authentication is the confirmation of the authenticity of various aspects of information interaction: the content and source of transmitted messages, communication session, interaction time, etc. It is an important part of the problem of ensuring the reliability of the information received. This problem is especially acute in the case of non-trusting parties, when not only the third party (adversary-enemy) can serve as a source of threats, but also the party with whom the informational interaction takes place (adversary-violator). In law, non-repudiation implies one's intention to fulfill their obligations to a contract. In cryptographic protocol it also implies that one party of a transaction cannot deny having received a transaction, nor can the other party deny having sent a transaction. And, finally, untraceability is a property meaning that it is impossible for an adversary to obtain information about the actions of the protocol participants.

Users of electronic financial services must be sure of the authenticity of messages arriving at their address. This is part of the task of authentication, which plays an important role, for example, in the organization of machine-to-machine interaction on the Internet of things and in the implementation of a private blockchain network. The solution to this problem is not generally provided by the use of encryption systems designed to perform the task of confidentiality. It follows from paper [4]. To protect against active adversary attacks, the message is supplied with a tag or message authenticity code. Information authentication systems are not required to maintain confidentiality, although in some systems both of these tasks are performed [5, chapter 14].

II. DEFINITIONS

The mathematical model of authentication systems is an authentication code or A-code, originally proposed in [6] and discussed in detail in the works of A.Yu. Zubov (for example. [5; 7]).

A Cartesian authentication code or systematic authentication code is a four-tuple $(\mathcal{S},\mathcal{T},\mathcal{K},\{E_k : k\in\mathcal{K}\})$ where $\mathcal{S}$ is the source state space associated with a probability distribution, $\mathcal{T}$ is the tag space, $\mathcal{K}$ is the key space associated with a probability distribution, and $E_k : \mathcal{S}\to\mathcal{T}$ is called an encoding rule. A transmitter and a receiver share a key k for authentication purpose. If the transmitter wants to send a source state $s\in\mathcal{S}$ to the receiver, he computes $t = E_k(s)$, $t\in\mathcal{T}$, and sends the message $m=(s,t)$ to the receiver through a public communication channel. When receiving $m'=(s',t')$, the receiver will compute $E_k(s')$ and checks whether $t' = E_k(s')$. If it does, the receiver will accept it as authentic. Otherwise, the receiver will reject it [8].

One of the mathematical objects associated with A-codes are orthogonal arrays. In [9], the design of the authentication code using orthogonal arrays, which is resistant to impersonation and substitution attacks, is described in detail.

Now we shall give the following definition. An orthogonal array $OA_\lambda(t,k,v)$ is a $\lambda v^t \times k$ array whose entries are chosen from a set X with υ points such that in every subset of t columns of the array, every t-tuple of points of X appears in exactly λ rows. These parameters are given the following names: υ is the number of levels, k is the number of factors, t is the strength, and λ is the index. An orthogonal array is simple if it does not contain any repeated rows [10]. This combinatorial structure was first introduced by an Indian statistician C. R. Rao [11] for use in design of experiments.

For cryptography applications the most frequently used orthogonal arrays are those with all factors at two levels, which we usually refer to as 2-level or binary orthogonal arrays, i.e. let $v=2$. It is binary orthogonal arrays that were used to construct authentication codes in [9].

This is example of a binary orthogonal array – orthogonal array $OA_3(2,11,2)$ from [10]:

$$\begin{pmatrix}
0&0&0&0&0&0&0&0&0&0&0\\
1&1&1&0&1&1&0&1&0&0&0\\
0&1&1&1&0&1&1&0&1&0&0\\
0&0&1&1&1&0&1&1&0&1&0\\
0&0&0&1&1&1&0&1&1&0&1\\
1&0&0&0&1&1&1&0&1&1&0\\
0&1&0&0&0&1&1&1&0&1&1\\
1&0&1&0&0&0&1&1&1&0&1\\
1&1&0&1&0&0&0&1&1&1&0\\
0&1&1&0&1&0&0&0&1&1&1\\
1&0&1&1&0&1&0&0&0&1&1\\
1&1&0&1&1&0&1&0&0&0&1
\end{pmatrix}$$

Correlation-immune functions were first introduced by Siegenthaler [12] as a class of suitable functions for combining the outputs of several linear feedback shift registers. They lead to the construction of running-key generators for stream ciphers which resist the correlation attack. The equivalence between t-th order correlation-immune binary function $f(x_1,...,x_k)$ with weight λv^t and simple binary orthogonal arrays $OA_\lambda(t,k,v)$ was proved in [13]. It follows that the study of orthogonal arrays is closely related to the cryptosyntesis and the cryptoanalysis of encryption systems.

III. RESULTS ABOUT ENUMERATION OF ORTHOGONAL ARRAYS

Enumeration of orthogonal arrays have been intensively investigated in [14–18]).

In [14] the enumeration of binary orthogonal arrays is studied, and a closed expression for the enumeration of binary orthogonal arrays of strength 1 is given using the inclusion – exclusion principle and the edge-induced subgraph.

In [15] there was specified an algorithm to enumerate a minimum complete set of combinatorially non-isomorphic orthogonal arrays of given strength t, run-size N, and level-numbers of the factors. The algorithm was the first one handling general mixed-level and pure-level cases. Using an implementation in C, there was generated most non-trivial series for $t = 2$, $N < 29$, $t = 3$, $N < 65$, and $t = 4$, $N < 169$. The exceptions defined limiting run-sizes for which the algorithm returns complete sets in a reasonable amount of time.

In [16] there was given an application of orthogonal arrays to construct D-optimal discrete choice experiments. Also there was addressed the existence problem of orthogonal arrays and described new algorithms to enumerate all orthogonal arrays with given parameters. Key techniques include eliminating significant space requirements, using previously-computed information on substructures in an efficient way. Computational results show these algorithms to be significantly faster than the prior state-of-the-art. Also there was defined the linear independence of Latin squares and stated a basis for a linear space which contains all Latin squares. Furthermore, there was used this theory to study an asymptotic formula for the number of orthogonal arrays with three columns.

In [17] the construction and enumeration of mixed orthogonal arrays (MOA) are described to produce optimal experimental designs. A MOA is a multiset whose rows are the different combinations of factor levels, discrete values of the variable under study, having very well defined features such as symmetry and strength three (all main interactions are taken in consideration). The applied methodology blends the fields of combinatorics and group theory by applying the ideas of orbits, stabilizers and isomorphisms to array generation and enumeration. Integer linear programming was used in order to exploit the symmetry property of the arrays under study. The backtrack search algorithm was used to find suitable arrays in the underlying space of possible solutions. To test the performance of the MOAs, an engineered system was used as a case study within the stage of parameter design. The analysis showed how the MOAs were capable of meeting the fundamental engineering design axioms and principles, creating optimal experimental designs within the desired context.

In [18], a method is described for the construction of all geometrically non-isomorphic ternary orthogonal arrays in 18

runs. One representative from each geometric isomorphism class is provided in the electronic appendix of [18]. The approach that have taken in this paper can be extended in principle but the problem rapidly becomes very large. For instance, when increasing the number of levels from 3 to 4, and considering 32 runs, the number of incidence matrices for two factors that needs to be considered to determine all possible valid third columns is 282, and the number of valid third columns is 22,695, none of which are geometrically isomorphic. If we keep ternary factors but instead increase the number of runs to 27 then the number of incidence matrices is 55, giving 847 valid triples of incidence matrices, and 424 valid third columns. To extend this to approach to larger m would require the consideration of $6^9 \times 847$ potential fourth columns. For $N = 36$ there are 120 incidence matrices, 3921 valid triples corresponding to 1971 possible third columns.

In addition to these papers, there are others devoted to the topic of enumeration of orthogonal arrays.

Using the results of [13; 19], we can estimate the asymptotic number of simple binary orthogonal arrays.

Let $|OA_\lambda(t,k,2)|$ be the number of simple binary orthogonal arrays $OA_\lambda(t,k,v)$. Hereafter, we use the following notation: $\exp_2(x) = 2^x$. Using theorem 3.1 [13] and theorem 4 [19], we get the following:

Theorem 1. Suppose $0 < \varepsilon < \frac{1}{2}$ and $t(5+2\log_2 k) \le k\left(\frac{1}{2}-\varepsilon\right)$ for sufficiently large k; then

$$|OA_\lambda(t,k,2)| = \theta_1(\lambda,t)\exp\left(2^k \ln 2 - 0{,}05\cdot\exp_2\left(2^{k\varepsilon-\log_2 k}\right)\right) +$$

$$+\exp_2\left(2^k - (k-t)\binom{k}{t} - \left(\log_2\sqrt{\pi/2}\right)\sum_{i=0}^{t}\binom{k}{i}\right)\times$$

$$\left(\exp\left(-2^{1-k}\left(2^{k-1}-\lambda 2^t\right)^2\right)\times\right.$$

$$\left.\times\left(1+\theta_2(\lambda,t)k^{-2}\right)+\theta_3(\lambda,t)\exp\left(-0{,}4\cdot 2^{k\varepsilon+3t-\log_2 k}\right)\right),$$

where $|\theta_1(\lambda,t)| \le 1$; $|\theta_2(\lambda,t)| \le 3{,}62$; $|\theta_3(\lambda,t)| \le 0{,}8$.

Let $|OA_\vee(t,k,2)|$ be the number of all simple binary orthogonal arrays $OA_\lambda(t,k,v)$ with any index λ. From the proposition 3 [19], it is clear that the following theorem is true.

Theorem 2. If $0 < \varepsilon < \frac{\ln 2}{4}$ and $t < \frac{k}{\ln k}\left(\frac{\ln 2}{4}-\varepsilon\right)$ for sufficiently large k then, as $k \to \infty$ $|OA_\vee(t,k,2)|$:

$$\exp_2\left(2^k - \frac{1}{2}\left((k-t)\binom{k}{t}-k\right) - t - \left(\log_2\sqrt{\pi/2}\right)\sum_{i=1}^{t}\binom{k}{i}\right).$$

IV. CONCLUDING REMARKS

In conclusion, we can add that orthogonal arrays have found numerous application in cryptography. Partial de-randomization of randomized algorithms, for example, the Monte Carlo algorithm used in cryptanalysis [20], secret sharing schemes, universal hash functions, perfect local randomizers [21] are among their applications [22].

V. REFERENCES

[1] Ежегодное послание Федеральному собранию. 01.12.2016. [Presidential annual address to Federal Assembly. 01.12.2016].. URL: http://www.kremlin.ru/events/president/news/53379/.

[2] Об утверждении программы «Цифровая экономика Российской Федерации». 31.07.2017. [On approval of the program “Digital Economy of the Russian Federation”. 31.07.2017]. URL: http://government.ru/docs/28653/.

[3] ISO/IEC 27000:2009 Information security management systems — Overview and vocabulary. URL: http://pqm-online.com/assets/files/lib/std/iso _iec_27000-2009.pdf.

[4] Месси Дж. Л. Введение в современную криптологию // ТИИЭР. 1988. Т. 76, N 5. С. 24–42. [Messi J.L. Introduction to modern cryptology // Proceedings of the Institute of Electrical and Electronics Engineers. 1988. Vol. 76, N 5. P. 24–42].

[5] Зубов А.Ю. Математика кодов аутентификации. М.: Гелиос АРВ, 2007. [Zubov A.Yu. Mathematics of authentication codes. Moscow: Helios ARV, 2007].

[6] Gilbert E.N., MacWilliams F.J., Neil J. et al. Codes which detect deception // Bell System Technical Journal. 1974. Vol. 53, N 3. P. 405–424.

[7] Зубов А.Ю. Коды аутентификации. М.: Гелиос АРВ, 2017. [Zubov A.Yu. Authentication Codes. Moscow: Helios ARV, 2017].

[8] Ding C., Helleseth T., Klove T. et al. A Generic Construction of Cartesian Authentication Codes // IEEE Transactions on Information Theory. 2007. Vol. 53, N 6. P. 2229–2235.

[9] Тарапников Ю.В. Комбинаторные свойства дискретных структур и приложения к криптологии. М.: МССМЕ, 2011 [Tarannikov Yu.V. Combinatorial properties of discrete structures and applications to cryptology. Moscow: MCCME, 2011].

[10] Sloane N.J.A., Hedayat A.S., Stufken J. Orthogonal Arrays: Theory and Applications NY: Springer, 1999. (Springer Series in Statistics).

[11] Rao C.R. Factorial Experiments Derivable from Combinatorial Arrangements of Arrays // Supplement to the Journal of the Royal Statistical Society. 1947. Vol. 9, N 1. P. 128–139.

[12] Siegenthaler T. Correlation-immunity of nonlinear combining functions for cryptographic applications // IEEE Trans. Information Theory. 1984. Vol. 30. P. 776–780.

[13] Camion P., Carlet C., Charpin P. et al. On Correlation-immune functions // Advances in Cryptology — CRYPTO ’91. CRYPTO 1991. Lecture Notes in Computer Science.Vol. 576. Berlin; Heidelberg: Springer, 1992. P. 86–100.

[14] Zhang J.-Z., You Z.-S., Li Z.-L. Enumeration of binary orthogonal arrays of strength 1 // Discrete Mathematics. 2001. Vol. 239, N 1–3. P. 191–198.

[15] Demirkale F.Orthogonal Arrays; Enumeration and Applications: PhD Thesis / The University of Queensland: School of Mathematics and Physics. 2013. https://espace.library.uq.edu.au/view/UQ:310823/s42201049_phd_finalthesis.pdf.

[16] Schoen E.D., Eendebak P.T., Nguyen M.V. Complete enumeration of pure-level and mixed-level orthogonal arrays // Journal of Combinatorial Designs. 2010. Vol. 18, N 2. P. 123–140.

[17] Romero J. Enumeration of strength three orthogonal arrays and their application in parameter design: PhD Thesis / University of Canberra; Faculty of Education, Science, Technology & Maths. 2017.

http://www.canberra.edu.au/researchrepository/file/19e87ebf-9906-42a6-8c68-a72e4631bf28/1/full_text.pdf.

[18] Bird E.M., Street D.J. Complete enumeration of all geometrically non-isomorphic three-level orthogonal arrays in 18 runs // Australasian Journal of Combinatorics. 2018.Vol. 71, N 3.P. 336–350.

[19] Панков К.Н. Улучшенные асимптотические оценки для числа корреляуионно-иммунных и _эластичных двоимчных вектор-функций // Дискретная математика. 2018. Т. 30, № 2. С. 73–98.[Pankov K.N. Improved asymptotic estimates for numbers of correlation-immune and (n,m,k)-resilient vectorial boolean functions // Discrete Mathematics. 2018. Vol. 30, N 2. P. 73–98].

[20] Заикин О.С., Семенов А.А. Применение метода Монте-Карло к прогнозированию времени параллельного решения проблемы булевой выполнимости // Вычислительные методы и программирование. 2014. Т. 15, №1.С. 22–35. [Zaikin O.S., Semenov A.A. Application of the Monte Carlo method for estimating the total time of solving the SAT problem in parallel // Computational Methods and Programming. 2014. Vol. 15, N 1. P. 22–35].

[21] Maurer U.M., Massey J.L. Local Randomness in Pseudo-random Sequences // Journal of Cryptology. 1991. Vol. 4, N 2. P. 135–149.

[22] Gopalakrishnan K., Stinson D.R.Applications of orthogonal arrays to computer science // Proc. of ICDM. 2006. P. 149–164,