UDC 519.1,514.128

**Vasyl Ustimenko**[1,2]**,** Doctor of Physical and Mathematical Sciences, Professor, Head of the information security department
ORCID ID: https://orcid.org/0000-0002-2138-2357 ***e-mail:*** vasulustimenko@yahoo.pl

**Oleksandr Pustovit**[2]**,** Candidate of Technical Sciences, Researcher
ORCID ID: https://orcid.org/0000-0002-3232-1787 ***e-mail:*** sanyk_set@ukr.net

[1]University of Royal Holloway, London, Great Britain
[2]Institute of Telecommunications and Global Information Space of the National Academy of Sciences of Ukraine, Kyiv, Ukraine

# ON SECURITY OF GIS SYSTEMS WITH N-TIER ARCHITECTURE AND FAMILY OF GRAPH BASED CIPHERS

***Abstract.*** *Discovery of q-regular tree description in terms of an infinite system of quadratic equations over finite field Fq had an impact on Computer Science, theory of graph based cryptographic algorithms in particular. It stimulates the development of new graph based stream ciphers. It turns out that such encryption instruments can be efficiently used in GIS protection systems which use N-Tier Architecture. We observe known encryption algorithms based on the approximations of regular tree, their modifications defined over arithmetical rings and implementations of these ciphers. Additionally new more secure graph based ciphers suitable for GIS protection will be presented.*

*Algorithms are constructed using vertices of bipartite regular graphs D(n,K) defined by a finite commutative ring K with a unit and a non-trivial multiplicative group K*. The partition of such graphs are n-dimensional affine spaces over the ring K. A walk of even length determines the transformation of the transition from the initial to the last vertex from one of the partitions of the graph. Therefore, the affine space Kn can be considered as a space of plaintexts, and walking on the graph is a password that defines the encryption transformation.*

*With certain restrictions on the password the effect when different passwords with K*)2s, s <[(n+5)/2]/2 correspond to different ciphertexts of the selected plaintext with Kn can be achieved.*

*In 2005, such an algorithm in the case of the finite field F127 was implemented for the GIS protection. Since then, the properties of encryption algorithms using D(n, K) graphs (execution speed, mixing properties, degree and density of the polynomial encryption transform) have been thoroughly investigated.*

*The complexity of linearization attacks was evaluated and modifications of these algorithms with the resistance to linearization attacks were found. It turned out that together with D(n, K) graphs, other algebraic graphs with similar properties, such as A(n, K) graphs, can be effectively used.*

*The article considers several solutions to the problem of protecting the geological information system from possible cyberattacks using stream ciphers based on graphs. They have significant advantages compared to the implemented earlier systems.*

***Keywords:*** *Stream ciphers; GIS protection; Multivariare Cryptography; Graph Based Cipher; graphs given by equations; regular tree approximations.*

**https://doi.org/10.32347/2411-4049.2023.3.113-132**

## 1. Introduction

Graph Based Cryptography (GBC) area is moving with great speed into the main stream of computer design, Information sciences, Information and Computer programming, Artificial Intelligence and design, Artificial Intelligent and various field of research. Application of GBC is in diverse area such as Data structures, Communication networks and their security. A Graph-based approach centres on conserving the environment of security events by breaking down factors of observable data into a graph representation of all cyber vestiges, from all data aqueducts, counting for all once and present data. For secret communication, Ciphers can be converted into graphs. The Application of Graph Theory plays a vital role in various field of Engineering and Sciences. GBC is used for the key exchange, development of Multivariate Public Keys, key dependent message authentication codes and algorithms of Noncommutative Cryptography (see [30]-[47]).

Especially Graph theory is commonly used as a tool of symmetric encryption. First cryptographical applications of Graph Theory appeared in the areas of Symmetric Cryptography and Network Security. This paper reflects some results in the area of applications of families of algebraic graphs of large girth of Extremal Graph Theory to the development of fast and secure encryption tools to process Big Data files. The vertices and edges of algebraic graphs form algebraic varieties defined over the field. The girth is the length of the minimal cycle in the graph. This parameter defines the size of the key space of corresponding cipher. The girth of several known families of algebraic graphs of large girth is not computed. It just evaluated via the lower bounds.

Observed and presented new ciphers have a multivariate nature. The space of plaintexts is an affine variety $K^n$ defined over finite commutative ring $K$. Bijective encryption map $F$ can be given by nonlinear multivariate polynomials $f_1, f_2, \ldots, f_n$ from the multivariate commutative ring $K[x_1, x_2, \ldots, x_n]$. It acts on the affine space accordingly the rule $(x_1, x_2, \ldots, x_n) \rightarrow (f_1(x_1, x_2, \ldots, x_n), f_2(x_1, x_2, \ldots, x_n), \ldots, f_n(x_1, x_2, \ldots, x_n))$, where $f_i$ are given via corresponding list of monomial terms. Trapdoor accelerator (see [27]) is a piece of information $A$ such that the knowledge of $A$ allows to compute the reimage of $F$ in time $O(n^2)$.

In presented ciphers correspondents Alice and Bob shares file $A$ (the password) and encrypt according to the robust procedure in time $O(n)$ or $O(n^2)$. The adversary does not have a password he/she can intercept large amount of pairs plaintext/corresponding ciphertext and try to approximate maps $F^{-1}$ and $F$. So degree of $F$ is an important parameter for the cryptanalytical studies. The most important (active) part of password are is the information about the walk in the algebraic graph.

Section 2 is dedicated to discussion of the applications of algebraic graphs to protection of Geological Information Systems. We discuss the known successful example [2] of such application in 2005. It was based on idea of usage walks on regular graphs approximating infinite $127$-regular. In fact the first description of selected graph based stream cipher based on approximations of $q$-regular tree where $q$ is a prime power was presented at the beginning of 2001. During last twenty years many new results on the construction of new encryption tools and there cryptanalysis were obtained. They lead to understanding of multivariate nature of these algorithms and necessity of usage of infinite algebraic graphs defined over infinite commutative rings of kind $F_q [x_1, x_2, \ldots, x_n]$ or more general $K[x_1, x_2, \ldots, x_n]$ where $K$ is a finite commutative ring. Implemented in [2] encryption map is a polynomial map of degree 3 such that their inverse is also cubical transformation. So, adversary can use linearisation attacks and after the interception of $O(n^3)$ pairs of kind plaintexts/corresponding ciphertext he/she can approximate the encryption map in time $O(n^{10})$. So, Section 2 is dedicated to observation of ciphers based on algebraic graphs and resistant to such linearisation attacks.

The general scheme of flexible encryption algorithm based on special family of algebraic graphs defined over commutative ring is presented there. The theory of approximations of regular trees is presented in Section 3 which contains description of $q$-regular forest approximation $D(n, q)$, $n\rightarrow\infty$ and tree approximations $CD(n, q)$ and $A(n, q)$, $n=2,3,\ldots$ Analogues of these families of graphs over an arbitrary commutative ring are presented there together with the known results on their properties and applications.

Precise description of observed graph based algorithm is given in the Section 4 together with evaluation of the degrees of encryption map and its inverse.

The special case of $A(n, 256)$ defined over the finite field $F_{256}$ is selected for an implementation. Parameters of corresponding computer simulations are given at the end of Section 4.

Last Section 5 is the conclusion.

## 2. On GIS and approximations of infinite regular trees

Security aspects of using geospatial data and Geographical Information Systems (GIS) are vital topics for current research. A number of publications on applications of GIS to Cybersecurity, National security, and Intelligence operations are rapidly growing (see [1], [2], [3]). Currently, GIS is an essential instrument of Decision Making. Despite these facts, questions on the security of GIS are relatively unexplored topics. Grows in the community of GIS users and the area of GIS applications search for new security solutions, a critical research direction. One of the first surveys [1], [2] with analyses of the quality and efficiency of such solutions published in 2005. The authors suggested using $N$-Tier GIS architecture.

This paper observes the application of the primary database security categories for managing spatial data. These categories are analyzed from the point of view of application within GIS in the Global Information Space. A File System within a Database (FSDB) with traditional and new encryption algorithms has been proposed as a new GIS Security solution. An FSDB provides more safe and secure storage for spatial files and supports centralized authentication and access control mechanism in legacy DBMS. Cryptography solutions, as a topic of central importance to many aspects of network security, are observed in detail.

The paper describes several traditional and new symmetric, fast and nonlinear encryption algorithms' implementation with fixed and adjustable key sizes, which uses methods of graph-based cryptography. This article is well cited during 2005-2018 (see proceedings of conferences on Geographical Information Systems Theory, Applications and Management – GISTAM). Several authors agree on the effectiveness of N-Tier architecture [2] and suggested methods of its usage. They implemented one case from the family of graph based stream ciphers defined in [4]. It is the case of finite field $F_{127}$.

This is a family of graph based ciphers based on the well-known algebraic graphs $D(n, q)$ of Extremal Graph Theory (see [5], [6]) defined over the finite field $F_q$ where $q$ is a prime power. The space of plaintexts is a vector space $(F_q)$. Used in [1], [2] case of $q = 127$ demonstrated that corresponding DBMS is capable enough to provide sufficient security to spatial files. This encryption procedure can provide additional security to confidential and sensitive GIS information. Oracle Advanced Security of the Oracle DBMS supports industry-standard encryption algorithms, including RSA's RC4, DES and 3DES and can be used for spatial data encryption with graph based algorithms.

Custom external encryption algorithms can be integrated into that security schema as well. The data encryption can significantly degrade system performance and application response time. For performance testing, the Oracle DBMS_OBFUSCATION.TOOLKIT was investigated [2]. Different key length gives different time results, e.g. difference in time between 16 and 24-byte keys is about 10-20%, but the time difference between 24 and 32-byte keys is only about 5%. It means that the new graph based stream cipher for GIS has to be compared well with ciphers DES, 3DES, stream ciphers RC4. Particular approaches were developed to encrypt large files in Oracle DBMS for the GIS objects.

The procedure splits the data into smaller binary chunks to encrypt large data objects, then encrypts and appends them back to the large data object (LDO). Once the encrypted spatial data files have been allocated into LDO segments, they can be decrypted by chunks and written back to LDO. Additional LOB objects, once encrypted, should always be kept for the read-only spatial data files. It will save time for the encryption procedure during log-off. The decrypted spatial data files will be replaced by read-only encrypted spatial data files in the permanent primary storage during log-off. The implanted cipher gives a more robust binary and text file encryption algorithm than DES, 3DES.

We have to report that the implemented case of $D(n, q)$ based encryption $E(n, q)$ is far from being optimal. As it was showed in [7] the increase of parameter $q$ leads to faster encryption of files of the same size. Noteworthy that the usage of loaded multiplication tables makes immaterial the difference between case of prime $q$ and composed prime powers. Such tables allow to use $q=128$ corresponding to the alphabet ASCEE with the essential speed increase comparably to implemented in [1], [2] $q=127$, where operator of taking modulo $127$ is used cn times where constant c depends on the length of the password. The multivariate nature of $D(n, q)$ encryption was noticed in [7] (see also [28] for the case of arbitrary ring $K$), described their symbolic computations turned out to be cubic.

This fact was mathematically proved in [8] for arbitrary parameters $n$ and $q$. The standard usage of multivariate transformation $E(n, q)$ with two affine transformation $T_1$ and $T_2$ in the form $T_1 E(n, q)T_2$ allow us to improve drastically the mixing properties of the cipher. Noteworthy that in the implemented case of $E(n, 127)$ encryption the change of single characters of the plaintext leads to the change of 48-52 percents of characters of corresponding ciphertexts. The experiment with special linear transformations $T_1$ and $T_2$ was described in [9, Ustim Kotorowicz]. To preserve linear time $O(n)$ of the encryption we have to select sparce transformations, i. e. those with $O(n)$ nonzero entries of corresponding matrices. Special sparce transformations allow us to improve drastically mixing properties of $E(n, q)$ encryption. For selected in [9, Ust Kotorow] cases the single change of a plaintext character leads to the change of more than $98$ percents of characters of corresponding ciphertext. As it was shown in [ust linguistic] transformation $E(n, q)$ with the password of length less than $[(n+5)]$ has no fixed points. This property holds for the case of ciphers of kind $T_1 E(n, q)(T_1)^{-1}$.

More general graphs *D(n, K)* defined over arbitrary commutative ring *K* can be obtained via simple change of $F_q$ for *K* (see [10]). Investigation of dynamical systems corresponding to these graphs showed the similarity of general graphs *D(n, K)* of the graphs defined for the case of fields (see [11], [12] and [13]). If passwords corresponds to tuples of characters from the multiplicative group $K^*$ of the ring *K* then different passwords of length < $[(n+5)/2]$ produce distinct ciphertext from the selected plaintext. It means that case of arithmetic rings $Z_m$ of integers of modulo *m* is attractive for the implementations.

Noteworthy that the cases of fields $F_q$, q = $2^m$ of characteristic two and rings $Z_q$, $q=2^m$ are most convenient for implementations because of files in the computer are presented in the form *0, 1*-sequences.

Recall that the girth of a graph is the length of its minimal cycle. The connected components *CD(n, q), n=2, 3,…* of algebraic graphs *D(n, q), q>1* form a family of tree approximations, i. e well defined projective limit of them is an infinite q-regular tree. Graphs *D(n, q)* are edge transitive. So, their connected components are isomorphic. The system of quadratic equations which defines *CD(n,q)* were presented in [14]. The union of these equations gives an algebraic description of *q*-regular tree. Existence of such description is very important for Computer Science because a q-regular tree is the deterministic part of branching process.

Noteworthy that the plaintext and the ciphertext of *E(n, q)* encryption are located in the same connected component of *D(n, q).* Graphs *CD(n, q)* have a natural analogue *CD(n, K)* defined over arbitrary commutative ring K with at least two elements, *CD(n, K)* is an induced subgraph of *D(n, K)* (see [10]). The description of *CD(n, K)* in terms of the system of recurrent quadratic equations is given in [10] together with the description of *CD(n, K)* based encryption *CE(n, K).*

It works with the space of plaintexts $K^m$, $m=3/4n +c$ where *c*, $c<3$ is some nonnegative integer constant. It is important that group of transformations of *CE(n,K)* corresponding to various passwords acts transitively on the space of plaintexts while the group generated by various transformations of kind *E(n, K)* is intransitive. It leads to better mixing pro$^n$perties of *CE(n, K)* in comparison with those of *E(n, K).* In fact we have to use $T_1CE(n, K)(T_1)^{-1}$ where $T_1$ is a special sparce transformation of $AGL_m(K)$.

Another q-regular tree approximation *A(n, q), q=2,3, …*were defined in [15]. It has some advantages in comparison with graphs *CD(n, q).* For instance the graphs are defined by simple homogeneous equation with two linear and one quadratic monomial terms. Finite field $F_q$ can be substituted by general commutative ring *K* and graphs *A(n, K)* can be obtained this way (see [15] or Extremal [13]). The girth *g(A(n, q))* of the graphs $A(n, q)=A(n, F_q)$ can be bounded from below via inequality $g(A(n, q))\geq[(n+2)/2]$ [16]. The computer simulation support the conjecture that $A(n, Z_m)$ based encryption with passwords from $((Z_m)^*)^t$, $m>2$, *t* is an even parameter $<[(n+2)/4$ is such that different passwords produce distinct ciphertext from the selected plaintext. We will use notation *AE(n, K)* for the *A(n, K)* based ciphers.

To summarise written above we discuss some properties of three graph based steam ciphers *E(n, K), CE(n, K)* and *AE(n, K)* defined in the case $K=F_q$, $q>m$ and $K=Z_m$, $m>2$. All of them can be used for GIS protection with described above *N*-tier architecture. For practical implementation case of large finite fields and arithmetic rings $Z_t$, $t=2^m$ is preferable.

The families of graphs $D(n, K), A(n, K)$ defined over arbitrary commutative ring $K$ are bipartite graphs of type $(1, 1, n-1)$ with partition sets which are two copies of $K^n$ (see [17] or [11]), i.e. graphs with the incidence $I=I(K)={}^nI(K)$ between points $(x_1, x_2,\ldots, x_n)$ and lines $[y_1, y_2,\ldots, y_n]$ given by the system of equations $a_2x_2-b_2y_2= f_2(x_1, y_1), a_3x_3-b_3y_3= f_2(x_1, x_2 ,y_1, y_2),\ldots, a_nx_n-b_ny_n= f_2(x_1, x_2 ,\ldots, x_{n-1}, y_1, y_2 ,\ldots, y_{n-1})$ where parameters $a_2, a_3 ,\ldots, a_{n-1}$ and $b_2, b_3 ,\ldots, b_{n-1}$ are taken from the multiplicative group $K^*$ of the commutative ring $K$. Parameters $\rho((x_1, x_2,\ldots, x_n))=x_1$ and $\rho([y_1, y_2,\ldots, y_n])=y_1$ serve as colours of the point and the line. The following linguistic property holds. Each vertex of the graph has a unique neighbour of the chosen colour.

Graph $CD(n,K)$ after the elimination of computed recurrently parameters also can be written as linguistic graphs of type $(1, 1, m-1)$ where $m=[3/4n]+c$.

In fact the architecture require a partition of information into blocks of the same size. So, parameters $n$ and $m$ equals to some selected constant. the length of the password is another even constant which has an impact on the speed of encryption. Other option to increase speed of execution is the increase the cardinality of the ground field or ring. Let us consider the general scheme of creating the cipher based on the family of linguistic graphs ${}^nI(K), n=2, 3, \ldots$

Noteworthy that we can expand defined above $I(K)$ to the infinite linguistic graph $I(K[x_1, x_2,\ldots, x_n])$ defined over the ring $K[x_1, x_2,\ldots, x_n]$ of all multivariate polynomials with coefficients from $K$ and the variables $x_i, i=1,2,\ldots, n.$ So points and lines of this graph are $X=(X_1(x_1, x_2,\ldots, x_n), X_2(x_1, x_2,\ldots, x_n),\ldots, X_n(x_1, x_2,\ldots, x_n)$ *and* $Y=[Y_1(x_1, x_2,\ldots, x_n), Y_2(x_1, x_2,\ldots, x_n),\ldots, Y_n(x_1, x_2,\ldots, x_n)]$. The incidence of this bipartite graph is given by equations $a_2X_2-b_2Y_2 = f_2(X_1, Y_1), a_3X_3-b_3Y_3= f_2(X_1, X_2, Y_1, Y_2),\ldots, a_nX_n-b_nY_n= f_2(X_1, X_2 ,\ldots, X_{n-1}, Y_1, Y_2 ,\ldots, Y_{n-1} )$, where parameters $a_2, a_3 ,\ldots, a_{n-1}, b_2, b_3 ,\ldots, b_{n-1}$ and polynomials $f_i, i=2, 3,\ldots, n$ with coefficients from $K$ are taken from the equations in the definition of the linguistic graph $I(K)$.

We define the polynomial map $F$ from $K^n$ *to* $K^n$ via the following scheme (see [29]). Take the special point $X=(x_1, x_2,\ldots, x_n)$ of $I(K[x_1, x_2,\ldots x_n])$ and consider the list of colours $g_1(x_1), g_2(x_1), \ldots, g_t(x_1)$. We compute the path $v_0Iv_1Iv_2\ldots Iv_t$ where $v_0=X$ and $v_{i+1}$ is the neighbour of $v_i$ with the colour $g_i(x_1), i=1,2, \ldots, t$ and $I=I(K[x_1, x_2,\ldots, x_n])$. Then the destination point $v_t$ of this path can be written as $(g_t(x_1), F_2(x_1, x_2), \ldots, F_n(x_1, x_2,\ldots, x_n))$. The map $F$ is given by the rule $x_1\to g_t(x_1), x_2\to F(x_1, x_2),\ldots, x_n\to F(x_1, x_2,\ldots, x_n)$. It is easy to see that $F=F(g_1, g_2,\ldots, g_t)$ is a bijective map if and only if the equations of kind $g_t(x_1)=b$ have unique solutions for unknown $x_1$ for each $b$ from $K$.

So family of linguistic graphs ${}^nI(K), n=2, 3,\ldots$ together with family of affine transformations $T_n\epsilon AGL_n(K)$ can be used as a cipher with the space of plaintexts $\mathrm{K}^n$ and the password $g_1(x), g_2(x),\ldots, g_t(x)$ and the encryption map $T_n(F(g_1, g_2,\ldots, g_t)(T_n)^{-1}$.

Correspondents Alice and Bob share the password given by $g_1, g_2,\ldots, g_t$ and the sequence of transformations $T_n , n=2, 3,\ldots$ We assume that inverse maps $(T_n)^{-1}$ are computed and presented explicitly. For the encryption of potentially infinite plaintext $(p)=(p_1, p_2,\ldots, p_n)$ they will use transformation $T_nF(g_1, g_2,\ldots, g_t)(T_n)^{-1}$. One of them creates the plaintext $(p)$ and computes the ciphertext $T_n(F(g_1, g_2,\ldots, g_t)(T_n)^{-1}(p)=c$ recurrently. The procedure is the sequence of the following steps.

$S_1$. He/she computes $(T_n)^{-1}(p_1, p_2,\ldots, p_n) =(r(1), r(2),\ldots, r(n))=(r)$.

$S_2$. He/she computes $a(1)=g_1(r_1), a(2)=g_2(r_1),\ldots, a(t)=g(r_1)$.

S$_3$. Let $N_a(x_1, x_2,..., x_n)$ be the operator of taking the neighbour of point $(x_1, x_2,..., x_n)$ with the colour $a$ in the linguistic graph ${}^nI(K)$ and ${}^aN(y_1, y_2,..., y_n)$ be an operator of taking the neighbour of line $[y_1, y_2,..., y_n]$ with the colour $a$. He/she executes the following operation. Computation of $v_1=N_{a(1)}(r)$, $v_2= {}^{a(2)}N(v_1)$, $v_3=N_{a(3)}(v_2)$, $v_4={}^{a(4)}N(v_3),..., v_{t-1}= N_{a(t-1)}(v_{t-2}), v_t={}^{a(t)}N(v_{t-1})=u=(u_1, u_2,..., u_n)$.

S$_4$. He/she computes ciphertext as $T(u)=c$.

DECRYPTION PROCEDURE.

Assume that one of correspondents received the ciphertext **c.** He/she decrypts via the following steps.

**D$_1$.** Computation of $u$ as $(T_n)^{-1}(c)=u$ and getting the solution $x=r(1)$ of equation $g(x)=u_1$.

**D$_2$.** Computation of parameters $a(1)=g_1(r(1)), a(2)=g_2(r(1)), ...,a(t\text{-}1)=g_{t\text{-}1}(r(1))$ and the completion of the recurrent procedure $v_{t\text{-}1}=N_{a(t\text{-}1)}(u)$, $v_{t\text{-}2}= {}^{a(t\text{-}2)}N(v_{t\text{-}1})$, $v_{t\text{-}3}=N_{a(t\text{-}3)}(v_{t\text{-}2})$, $v_{t\text{-}4}={}^{a(4)}N(v_{t\text{-}3}),..., v_1= N_{a(1)}(v_{t\text{-}2})$, ${}^{r(1)}N(v_{4t\text{-}1})=r$.

**D$_3$.** *Computation of the plaintext (p) as* $T(r)$.

OBFUSCATION OF THE ALGORITHM.

Let us consider the colour jump operator $J_a$ which transforms point $(p_1, p_2,..., p_n)$ of the graph $I(K)$ to the point $(a, p_2, p_3,..., p_n)$.

We can change the encryption map $T_nF(g_1, g_2,..., g_t)(T_n)^{-1}$ for the $T_nF(g_1, g_2,..., g_t)J_g(T_n)^{-}$, where $J_g$ is a colour jump operator acting on points of $I(K[x_1, x_2,...x_n]$ with the colour $g(x_1)\epsilon K(x_1)$ such that the equation of kind $g(x_1)=b$ has a unique solution for each parameter $b$ from $K$.

After this change assumption the bijection of $g_t$ on $K$ is immaterial.

Encryption procedure requires computation of $(T_n)^{-1}(p_1, p_2,..., p_n) =(r(1), r(2),..., r(n))=(r)$, the computation of u accordingly step S$_2$. the computation of $J_g(u)=u'$ and application of affine transformation $T_n$ to the tuple $u'$.

For the decryption of ciphertext c the user has to compute $u'=(u'_1, u'_2,..., u'_n)$ as $(T_n)^{-1}(c~)$, solve for $x$ the equation $g(x)=u'_1$, use the solution $x=r(1)$ of this equation for the computation of $a(1)=g_1(r(1))$, $a(2)=g_2(r(1)),..., a(t)=g_t(r(1))$, compute $J_{a(t)}(u')=(u)=(u_1, u_2,..., u_n)$ in the graph $I(K)$ and execute procedures $D_3$ and $D_4$ to get the original plaintext.

## 3. On families of algebraic graphs of large girth

### 3.1. General remarks

Girth and diameter of a graph are the minimal length of its cycle and the maximal distance of the graph. We can consider girth indicator *Cind(v)* of vertex *v* of the graph $\Gamma$ as the minimal length the cycle through *v* and introduce cycle indicator *Cind(Γ)* of the graph as the maximal value of *Cind*(v) for its vertices.

The constructions of finite or infinite graphs with prescribed girth and diameter is an important and difficult task of the Graph Theory. Noteworthy that the incidence of classical projective geometry over various fields is a graph of girth 6 and diameter 3. J. Tits defined generalised *m*-gons as bipartite graphs of girth *2m* and diameter *m*. Feit and Higman proved that finite generalised *m*-gons with bi-degrees > 2 exist only in the cases of *m=3, 4, 6, 8* and *12*. Geometries of finite simple groups of rank 2 are natural examples of generalised *m*-gons for *m=3,4,6, 8*. Classification of flag transitive generalised *m*-gons of Moufang type were obtained by J. Tits and R. Weiss.

Infinite families of graphs of large girth of bounded degree are important objects of Extremal Graph Theory which were introduced by P. Erdős'. He proved the existence of such families via his well-known probabilistic method. Nowadays few explicit constructions of such families are known. The concept of infinite family of small world graphs of bounded degree turns out to be very important for various applications of graph theory.

Noteworthy that only one family of small world graphs of large girth is known. This is the family *X(p, q)* of Ramanujan graphs introduced by Gregory Margulis [18] and investigated via the computation of their girth, diameter and the second largest eigenvalue by A. Lubotsky, R. Phillips and P. Sarnak [19].

We have to admit that studies of families of graphs $\Gamma_i$ with well defined projective limit $\Gamma$, which is isomorphic to infinite tree, is well motivated.

We refer to such family as tree approximation. There is only one approximation by finite graphs which is a family of large girth. This is the mentioned above family of *CD(n, q)* defined by F. Lazebnik, V. Ustimenko and A. Woldar. The question whether or not *CD(n, q)* form a family of small world graphs has been still open since 1995.

In 2013 the tree approximation by finite graphs *A(n,q)* which is a family of small world graphs was presented (see [15]). It was proven that the graph from the family has maximal known cycle indicator (in fact $Cind\ (A(n,q)) \geq 2n+2$).

One of the main statements of this paper is *A(n, q)* where *n=2, 3,...* is a family of large girth.

We generalise these results in terms of the theory of algebraic graphs defined over arbitrary field and consider properties and applications of above mentioned graphs.

## 3.2. On graphs *A(n, q)*, their properties and generalisations

All graphs we consider are simple, i. e. undirected without loops and multiple edges. Let $V(\Gamma)$ and $E(\Gamma)$ denote the set of vertices and the set of edges of $\Gamma$, respectively. The parameter $|V(\Gamma)|$ is called the order of $\Gamma$, and |E(G)| is called the size of $\Gamma$. A path in $\Gamma$ is called simple if all its vertices are distinct. When its convenient we shall identify $\Gamma$ with the corresponding antireflexive binary relation on $V(\Gamma)$, i.e. $E(\Gamma)$ is a subset of $V(\Gamma) \times V(\Gamma)$. The length of a path is a number of its edges. The girth of a graph $\Gamma$, denoted by $g=g(\Gamma)$, is the length of the shortest cycle in $\Gamma$. Let $k \geq 3$ and $g \geq 3$ be integers. The distance between vertices *v* and *u* of the graph $\Gamma$ is a minimal length of the path between them. The diameter of the graph is maximal distance between its vertices.

Graph is connected if its diameter is finite. Graph is *k*-regular if each vertex of the graph is incident exactly to *k* other vertexes. A tree is a connected graph which does not contain cycles.

(1) An infinite family of simple regular graphs $\Gamma_i$ of constant degree $k$ and order $v_i$ such that *diam* $(\Gamma_i) \leq c \log_{k-1}(v_i)$, where $c$ is the independent of $i$ constant and *diam* $(\Gamma_i)$ is diameter of $\Gamma_i$, is called a *family of small world graphs.*

(2) Recall that infinite families of simple regular graphs $\Gamma_i$ of constant degree $k$ and order $v_i$ such that $g(\Gamma_i) \geq c \log_{k-1}(v_i)$,where $\boldsymbol{c}$ is the independent of $i$ constant and $g(\Gamma_i)$ is a girth of $\Gamma_i$ are called *families of graphs of large girth.*

Let $\Gamma$ be a simple graph. Assume that *Cind(x)* is the minimal length of cycle through vertex $x$ of the graph $\Gamma$. Let C*ind(G)* stand for the maximal *value* of *Cind(x)* via all vertices $x$ of $\Gamma$. We refer to parameter *Cind(G)* as a cycle indicator of $\Gamma$.

One of the main purposes of the paper is to present a special interpretations of $q$-regular tree ($q$-regular simple graph without cycles) in terms of algebraic geometry over finite field $F_q$.

**Theorem 1** [16]. *For each prime power $q$, $q>2$ there is a family of $q$-regular graphs $\Gamma_i$ satisfying following properties*

*(i) $\Gamma_i$ is a family of small world graphs,*

*(ii) $\Gamma_i$ is a family of large girth,*

*(iii) Projective limit of graphs $\Gamma_i$ is well defined and coincides with q-regulat tree $T_q$.*

*(iv) Cind $\Gamma_i \geq 2 \log_q(v_i/2)+2$.*

We refer to family of graphs $\Gamma_i$ satisfying condition (iii) as *tree approximation.* The prove of Theorem 1 is given via explicit construction of graphs $\Gamma_i = A(i,q)$, $i \geq 2$ satisfying requirements of the statement. Noteworthy that $A(i,q)$ is a unique known example of the family satisfying conditions (i), (ii) an (iii).

In fact, there is exactly one other known construction of the $q$-regular family satisfying (i) and (ii), i.e. explicit construction of the family of regular simple small world graphs of large girth and with an arbitrarily large degree $q$.

This family $X(p, q)$ formed Cayley graphs for $PSL_2(p)$, where $p$ and $q$ are primes, had been defined by G. Margulis [18] and investigated by A. Lubotzky, Sarnak and Phillips [19]. As it is easy to see the projective limit of $X(p, q)$ does not exist.

**The construction of *A(n, q)*.**

Let $K$ be a finite field $F_q$. We define $A(n, K)=A(n,q)$ as bipartite graph with the point set $P=K^n$ and line set $L=K^n$ (two copies of a Cartesian power of $K$ are used). We will use brackets and parenthesis to distinguish tuples from $P$ and $L$.

So $(p)=(p_1, p_2, \dots , p_n) \epsilon P_n$ and $[l]=[l_1, l_2, \dots , l_n] \epsilon L_n$.

The incidence relation $I=A(n,K)$ (or corresponding bipartite graph $I$) is given by condition $p\, I\, l$ if and only if the equations of the following kind hold.

$p_2 - l_2=l_1p_1$,

$p_3 - l_3= p_1 l_2$,

$p_4 - l_4 = l_1p_3$,

$p_5 - l_3 = p_1 l_4$,

$\dots$ ,

$p_n - l_n = p_1 l_{n-1}$ for odd $n$ and $p_n - l_n = l_1 p_{n-1}$ for even $n$.

We can consider an infinite bipartite graph $A(K)$ with points $(p_1, p_2, \dots, p_n, \dots)$ and lines $[l_1, l_2, \dots, l_n, \dots]$.

**Proposition 1** [16]. *If $K=F_q$, $q>2$ then $A(n, F_q)$ is a family of small world graphs and tree approximation with $Cind(A(n, F_q))\geq 2n+2$.*

Let $K$ be an arbitrary field. We define $A(n, K)$ via simple change of $F_q$ on $K$ and announce the following statement.

**Proposition 2** [16]. *Let K be a field. Then the girth of $A(n,K)$ is $\geq 2[n/2]+2$.*

Symbol *[x]* stands for the flow function from ***x***. Theorem 1 follows from propositions 1 and 2.

### 3.3. On homogeneous algebraic graphs of large girth

Let *F* be a field. Recall that a projective space over *F* is a set of elements constructed from a vector space over *F* such that a distinct element of the projective space consists of all non-zero vectors which are equal up to a multiplication by a non-zero scalar. Its subset is called a quasiprojective variety if it is the set of all solutions of some system of homogeneous polynomial equations and inequalities.

An algebraic graph $\varphi$ over *F* consists of two things: the vertex set *Q* being a quasiprojective variety over *F* of nonzero dimension and the edge set being a quasiprojective variety $\varphi$ in Q × Q such that $(x, x)$ is not element of $\varphi$ for each $x \in Q$ and $x\varphi y$ implies $y\varphi x$ ($x\varphi y$ means $(x, y) \in \varphi$). The graph $\varphi$ is homogeneous (or *M*-homogeneous) if for each vertex $v \in Q$ the set $\{x \mid v\varphi x\}$ is isomorphic to some quasiprojective variety *M* over *F* of nonzero dimension. We further assume that *M* contains at least 3 elements.

**Theorem 2** [20]. *Let $\Gamma$ be homogeneous algebraic graph over a field F of girth g such that the dimension of neighborhood for each vertex is N, $N \geq 1$. Then $[(g-1)/2] \leq dim(V)/N$.*

The following corollary is an analog of Even Circuit Theorem by Erdős' for finite simple graphs.

**Corollary 1**. *Let $\Gamma$ be a homogeneous graph over a field F and $E(\Gamma)$ be a variety of its edges. Then $dim(E(\Gamma)) \leq dim V(\Gamma)(1 + [(g-1)/2]^{-1}$.*

We refer to a family of homogeneous algebraic graphs $\varphi_n$ for which dimension of neighborhood for each vertex is independent constant *N*, $N \geq 1$ as a family of *small world graphs* if diameter of each graph $\varphi_n$ is bounded from above by linear function $\alpha n + \beta$ defined by constants $\alpha$ and $\beta$.

We refer to a family of homogeneous algebraic graphs $\varphi_n$ for which the dimension of neighborhood for each vertex is independent constant *N*, $N \geq 1$ as a *family of large girth* if girth of each graph $\varphi_n$ is bounded from below by linear function $\alpha n+\beta$ defined by constants $\alpha$ and $\beta$.

We refer to a homogeneous algebraic graph as algebraic forest if it does not contain cycles. Their term algebraic tree stands for the connected algebraic forest.

We say that family of homogeneous algebraic graphs $\varphi_n$ is a forest (tree) approximation if projective limit of $\varphi_n$ is an algebraic forest (tree) and formulate thaw following statement.

**Theorem 3** [16]. *For each field F, $F \neq F_2$ there exists a tree approximation which is a family $\varphi_n$ of small world algebraic graphs of large girth with the vertex set of dimension n and cycle indicator $\geq 2n+2$.*

Family of graphs $\varphi_n = A(n, F)$ provides explicit construction of objects described in the theorem. As it follows from Theorem 2 homogeneous algebraic graphs $A(n, F)$ form a family with maximal possible girth indicator.

**Remark 1**. Graphs $A(n, F_2)$ are disconnected. So they are disjoint union of cycles. Graph $A(F_2)$ is 2-regular forests with trees presented on the following diagram …. -----*-----*-----*--- .... . Girth indicator of $A(n, F_2)$ coincides with its girth of size $\geq 2n+2$. So, formally $A(n, 2)$ are algebraic graphs of large girth. Noteworthy that cycles can be defined via the system of equations.

### 3.4. Graphs *A(n,K)* as homomorphic images of *D(n,K)*

Graphs *A(n,q)* obtained as homomorphic images of graphs *D(n,q)* which defines projective limit *D(q)* with points

$(p)=(p_{01}, p_{11}, p_{12} , p_{21}, p_{22}, p'_{22} , \ldots, p'_{ii}, p_{i\,i+1}, p_{i+1,i} , p_{+i+1,i+1} \ldots ),$

lines

$[l]=[ l_{10}, l_{11}, l_{12} , l_{21}, l_{22}, l'_{22} , \ldots, l'_{ii}, l_{i\,i+1}, l_{i+1,i} , l_{+i+1,i+1} \ldots ]$ and incidence relation given by equations

$l_{ii}-p_{ii}=l_{10}\, p_{i-1,i}$ ;

$l'_{ii} - p'_{ii} = l_{i,i-1}\, p_{01}$;

$l_{i,i+1} - p_{i,i+1} =l_{ii}\, p_{01}$ ;

$l_{i+1i} - p_{i+1,i} = l_{10}p'_{ii}$ .

This four relations are defined for $i\geq 1$, ($p'_{11}= p_{11}$, $l'_{11}= l_{11}$).

**Remark 2**. *You can see that indexes of vectors correspond to coordinates of positive roots of root system $A_1$ with a wave.*

Historically graph *D(q)* is not the first example of description of *q*-regular forest in terms of Algebraic Geometry. Geometries of buildings (see [21] and further references) corresponding to extended Dynkin diagram $A_1$ as incidence structures are $q+1$-regular trees or $q+1$-regular forests. As a result we get a description of a tree in group theoretical terms.

In [22] it was noticed that the restriction of this incidence relation on orbits of Borel subgroup $B^-$ acting on maximal parabolics are *q*-regular bipartite graphs. So we get a description of a *q*-regular tree in terms of positive roots of $A_1$ with a wave.

In [5] authors proved that *D(n,q)* defined via first *n-1*equations of *D(q)* form a family of graphs of large girth. The general point and line of these graphs are projections of *(p)* and [*l]* onto the tuples of their first n coordinates.

Unexpectedly it was discovered that these graphs are disconnected if $n \geq 6$. So forest *D(q)* contains infinitely many trees and the diameter is an infinity. F. Lazebnik conjectured that connected components of graphs *D(n,q), n =3,4, …* form a family of small world graphs. This conjecture is still open.

In 1994 it was found out how to describe connected components *CD(n, q)* of graphs *D(n, q)* in terms of equations (see [14], [6]).

Graphs *A(n, q)* were obtained in 2007 as homomorphic images of graphs *D(n, q)* ([11]). Corresponding homomorphism *ή* is a procedure to delete coordinates of points and lines with indexes *(i+1, i)* and *(i,i)'*.

The self importance of these graphs have been justified in joint research with U. Romanczuk (see [13] and further references) and M. Polak [23] via applications to Cryptography and Coding Theory.

In the case of families of graphs of large girth we would like to have "speed of growth" *c* of the girth "as large as it is possible".

P. Erdos' proved the existence of such a family with arbitrary large but bounded degree *k* with $c=1/4$ by his probabilistic method.

In the case of families *X(p,q)* and *CD(n,q)* the constant *c* is *4/3*. In the case of *A(n,q)* we just get inequality $1\leq c<2$. So exact computation of the girth is the area of the future research.

There are essential differences between family of graphs *X(p, q)* and tree approximations. Recall that the projective limit of *X(p, q)* does not exist.

It was proved that bipartite graphs $A(n,q)$ are not edge-transitive and not vertex transitive (transitivity on points and intransitivity on lines). Noteworthy that their projective limit $T$ (the tree) is obviously an edge-transitive infinite graph.

The usage of generalizations and modifications of graphs $A(n,q)$ allows us to construct postquantum cryptosystem of El Gamal type with encryption procedure for potentially infinite vector from $F_q$ with the execution speed $O(n^{1+2/n})$ (see [24]).

In fact the diameter of $A(n,q)$ is growing slower than diameter of $X(p,q)$. So, $A(n,q)$ are the best known small world graphs among known families of large girth. Recall the girth of $A(n,q)$ is not yet computed precisely.

So, the comparison of growth of the girth for $A(n,q)$ and $X(p,q)$ is the interesting task for the future research.

In the case of finite fields both families are expanding graphs, the second largest eigenvalue of $A(n, q)$ tends to $2q^{1/2}$, they are not Ramanujan graphs for which the second largest eigenvalue has to be bounded above by $2(q-1)^{½}$.

The family $X(p,q)$ is formed by Ramanujan graphs, so they are better expanding graphs than $A(n, K)$.

Families $X(p,q)$, $CD(n,q)$ and $A(n,q)$ can be used for the constructions of LDPC codes for noise protection in satellite communications. D. MacKay and M. Postol [25] proved that $CD(n, q)$ based LDPC codes have better properties than those from $X(p,q)$ for the constructions of LDPC codes.

Together with Monika Polak we proved that $A(n,q)$ based LDPC codes even better than those from $CD(n,q)$ (see [23]).

Cayley nature of $X(p,q)$ does not allow to use these graphs in multivariate cryptography. Various applications of graphs $D(n,q)$, $CD(n,q)$ and $A(n,q)$ have been known since 1998.

The most recent postquantum cryptosystem based on noncommutative multivariate group associated with $A(n,q)$ is described in [24], IACR e-print Archive 2021/1466.

### 3.5. On the equations for graphs *CD(n, K)*

Let $K$ stand for an arbitrary commutative ring. Noteworthy that graphs $A(n, K)$ and $D(n, K)$ are defined over arbitrary commutative ring $K$ have been already presented.

To facilitate notation in the future results on "*connectivity invariants*" of $D(n, K)$, it will be convenient for us to define $p_{-1,0} = l_{0,-1} = p_{1,0} = l_{0,1} = 0$, $p_{0,0} = l_{00} = -1$, $p'_{0,0} = l'_{0,0} = -1$, $p_{1,1} = p'_{1,1}$, $l_{1,1} = l'_{1,1}$ and to assume that our equations are defined for $i \geq 0$.

Graphs $CD(k,K)$ with $k \geq 6$ were introduced in [11] for as induced subgraphs of $D(k,K)$ with vertices $u$ satisfying special equations $a_2(u)=0, a_3(u)=0,\dots, a_t(u)=0$, $t=[(k+2)/4]$, where $u = (u_\alpha, u_{11}, u_{12}, u_{21}, \dots, u_{r,r}, u'_{r,r}, u_{t\,t+1}\, u_{r,r+1}, u_{r+1,r}, \dots)$, $2 \leq r \leq t$, $\alpha \epsilon \{ (1, 0), (0,1)\}$ is a vertex of $D(k, K)$ and $a_r = a_r(u)=\Sigma_{i=0,r}(u_{ii}\, u'_{r-i, r-i} - u_{i,i+1}\, u_{r-i,r-i-1})$ for every $r$ from the interval $[2,t]$ for every $r$ from the interval $[2,t]$.

We set $a=a(u)=(a_2, a_3, \dots, a_t)$ and assume that $D(k, K)=CD(k,K)$ *if* $k=2,3,4,5$.

As it was proven in [11] graphs $D(n, K)$ are edge transitive. So their connected components are isomorphic graphs. Let ${}^vCD(k,K)$ be a solution set of system of equations $a(u)=(v_2,v_3,\dots,v_t)=v$ for certain $v \epsilon K^{t-1}$. It is proven that each ${}^vCD(k,K)$ is the disjoint union of some connected components of graph $D(n,K)$.

It is easy to see that sets of vertices of ${}^{v}CD(k,K)$, $v \epsilon K^{t-1}$ form a partitions of the vertex set of $D(n,K)$.

The concept of quasiprojective variety over commutative ring $K$ can be introduced via simple substitution of $K$ instead of field $F$. It leads to concepts of homogeneous algebraic graphs over $K$, forest and tree approximations and families of graphs of large girth over $K$. It was proven that for the case of commutative ring $K$ with unity of odd characteristic graphs $CD(n,K)$ are connected (see [26]). So graph $CD(n,q)=CD(n, F_q)$ for odd $q$ is a connected component of $D(n,q)$.

As it follows from definitions the image of restriction of homomorphism ή from $D(n, K)$ onto $CD(n, K)$ coincides with $A(n, K)$.

So graphs $A(n,K)$ are connected for the case of $K$ with unity of an odd characteristic.

**Theorem 4** [16]. *For each commutative integrity ring K the families of graphs CD(n, K), n=2,3,... and A(n, K),n=2,3,.. are forest approximations and families of graphs of large girth.*

## 4. On the description of selected algorithms based on algebraic graphs of large girth

To achieve linear speed $O(n)$ of the encryption described in Section 1 functions $g_i$, $i=1,2,..,t$ are selected in the form $x_1+c(i)$, $c(i)\epsilon K$ and the parameter $t$ will be selected within the interval $[2, [(n+5)/2])$ when $I(K)=D(n, K)$ or $I(K)=CD(n, K)$ and interval $[2, [n/2]+1)$ in the case when $I(K)=A(n, K)$.

Additionally we take parameters $b(1), b(2), ...,b(k)$, $a(1), a(2),...,a(k)$, $k=t/2$ from $K^*$ *to* construct $c(i)$ recurrently via the following rules $c(1)=b(1)$, $c(2)=a(1)$, $c(i)=c(i-2)+b(i)$ if $i$, $i\geq 3$ is odd $n$ and $c(i)=c(i-2)=a(i)$ if $i$, $i\geq 4$ is even. We refer to the tuple $(b(1), b(2),..., b(k), a(1), a(2),...,a(k))$ as active password and affine transformation $T$ as passive password.

Our choice insures that in the case of constant passive password the single change of a single character of active password leads to a change of the ciphertext produced from the selected plaintext.

We choose an affine transformation $T$ in the form of linear map given by the following rule

$T(x_1)=x_1+m(1)x_2+...+m(n-1)x_{n-1}$ where $m(i)$, $i=1,2,..., n-1$ are elements of $K^*$. $T(x_i)=x_i$ for $i=2,3,..., n$. So $T^{-1}(x_1)=x_1-m(1)x_2-m(2)x_3-...-m(n-1)x_n$. $T^{-1}(x_i)=x_i$ for $i=2,3,..., n$.

Recall that explicit description of linguistic graphs $D(n, K)$ and $A(n,K)$ is given in the previous section and general encryption algorithm is described *in section 2*. So, ciphers $T\,E(n,K)\,T^{-1}$ and $TEA(n, K)\,T^{-1}$ have full description.

In the case of graph $CD(n, K)$ we will use in fact the induced subgraph ${}^{h}CD(n, K)$, $h=(h_2, h_3,..., h_t)$, $t=[(n+2)/4]$ of $D(n, K)$ of all points and lines $u=(u_\alpha, u_{11}, u_{12}, u_{21}, ..., u_{r,r}, u'_{r,r}, u_{t\,t+1}\, u_{r,r+1}, u_{r+1,r}, ...)$ satisfying conditions $a_i(u)=h_i$.

Linguistic graph ${}^{h}CD(n, K)$ can be thought as bipartite graph with points

$(p)=(p_{01}, p_{11}, p_{12}, p_{21}, ..., , p_{i\,i+1}, p_{i+1,i}, p_{+i+1,i+1} ... ), i=2,3,..., t-1$

and lines

$[l]=[ l_{10}, l_{11}, l_{12}, l_{21}, l_{22}, ..., l_{i\,i+1}, l_{i+1,i}, l_{+i+1,i+1} ... ], i=2,3,..., t-1$ of length $n-t$.

Their incidence is given by the following system of equations

$$l_{ii}-p_{ii}=l_{10}\,p_{i-1,i}\ ;$$
$$l_{i,i+1}-p_{i,i+1}=l_{ii}\,p_{01};$$
$$l_{i+1i}-p_{i+1,i}=l_{10}p'_{ii},$$

where $p'_{22}$ is defined by the equation $a_2(p_{01}, p_{11}, p_{12}, p_{21}, p_{22}, p'_{22})=h_2$ and can be written as $p'_{22}=a_2(p_{01}, p_{11}, p_{12}, p_{21}, p_{22}, p'_{22})-h_1+p'_{22}=b_2(p_{01}, p_{11}, p_{12}, p_{21}, p_{22})$, *other parameters are* $p'_{33}=a_3(p_{01}, p_{11}, p_{12}, p_{21}, p_{22}, p'_{22}, p_{2,3}, p_{3,2}, p_{3,3}\, p'_{3,3})-h_3+p'_{33}=b_3(p_{01}, p_{11}, p_{12}, p_{21}, p_{22}, p'_{22}, p_{2,3}, p_{3,2}, p_{,33}),\ldots, p'_{tt}=a_t(p_{01}, p_{11}, p_{12}, p_{21}, p_{22}, p'_{22},\ldots, p'_{t-1,t-1}, p_{t-1,t}, p_{t,t-1}, p_{t,t}, p'_{t,t})-h_t+p'_{t,t}=b_t(p_{01}, p_{11}, p_{12}, p_{21}, p_{22}, p'_{22},\ldots, p'_{t-1,t-1}, p_{t-1,t}, p_{t,t-1}, p_{t,t})$.

The computation of symbolic expressions $p'_{i,i}$ recurrently *and their explicit* substitution in the system of equations give us the equations of the linguistic graph.

We assume that corresponding cipher has the space of *plaintexts* $K^{n-t}$. We use active passwords $(b(1), b(2),\ldots, b(k), a(1), a(2),\ldots,a(k))$ an linear transformations $T$ *of* $K^{n-t}$ constructed via described above rules. We assume that parameters $h_2, h_3,\ldots,h_t$ *will be* considered as part of active password and denote the cipher as $TCE(n, K)T^{-1}T_nF(g_1, g_2,\ldots, g_t)J_g(T_n)^{-}$.

We will use presented in Section 2 obfuscation scheme for each cipher $TE(n, K)T^{-1}$ $TAE(n, K)T^{-1}$ and $TCE(n, K)T^{-1}$ in the case $K=F_q$, $q>2$. We use special disturbance function $g$ of $I_g$ selected as $x\to x^e+b$ where $b\epsilon F_q$, $e\epsilon Z_d$, $d=q-1$ and $(e, d)=1$. So, the notations $DE(n, K)=TE(n, K)I_gT^{-1}$ and $DA(n,K)=TAE(n, K)I_gT^{-1}$ and $DC(n,K)=TCE(n, K)I_gT^{-1}$ will be used for these encryption schemes with the disturbance.

Algorithms with the encryption maps $TE(n, K)T^{-1}$ and $TAE(n, K)T^{-1}$ independently on the choice of active and passive passwords have multivariate encryption and decryption functions of degree 3. In [45] the linearisations attacks on these ciphers with the interception of $O(n^3)$ pairs plaintext/cipheretext are presented. They can be executed in polynomial time $O(n^{10})$.

The ciphers $DE(n, K)$ and $DA(n, K)$ use cubical encryption maps as well but the usage of disturbance map $D: x\to x^e$ lead to the increase of the degree $r$ of inverse maps. Parameter $r$ can be evaluated from below by the polynomial degree of transformation $D^{-1}$ acting on the elements of multiplicative group $K^*$. So, if $K=F_q$, $q=2^{32}$ then the order of polynomial decryption map is at least $2^{31}$. It justifies that direct linearisation attacks are not feasible.

Case $TCE(n, K)T^{-1}$ is principally different. As it follows the results of [46] the encryption function corresponding to selected active password has degree $[(n+2)/4]+2$. Recall that active password is formed by tuples $(b(1), b(2),\ldots, b(k), a(1), a(2),\ldots,a(k))$ and $(h_2, h_3,\ldots, h_t)$ where *h_i* are internal parameters of subgraph ${}^{h}CD(n, K)$. If $k$ is less than half of the girth then different active passwords produce distinct ciphertexts.

High degree of the transformation insures that a generation of standard form for the encryption function can not be done in polynomial time.

So the directed linearisation attacks are theoretically impossible. Principle difference of $DC(n, K)$ and $TCE(n, K)T^{-1}$ is the fact that the usage of disturbance implies the fact that the degree of inverse function is essentially higher than those for encryption function.

**The implemented case**

For the first implementation we select *the case of encryption function of $DA(n,K)= TAE(n, K)I_gT^{-1}$* for finite field $F_{256}$ with g of kind g = $x^2+b$. In this case the degree of encryption map will be at least 128 (see [27]). The linearisation attacks by adversary requires the interception $O(n^3)$ pairs of kind plaintext/ciphertext. After that he/she need approximate the map of degree ≥ 128 with the possibility to choose the plaintext an get corresponding ciphertext. In practical case of n ≥ 64 such linearisation attacks are unfeasible.

CRYPTALL 4 software is written in C++ programming language and therefore it is portable and runs in many platforms such as Unix/Window. Thecontext diagram is depicted in Fig. 1. The interface is friendly. It allows users to enter active and passive password of selected length. The program is supported by key exchange protocol based on Eulerian transformations of $(F^*_{256})$. It allows the elaboration of tuple of nonzero field elements of sufficient length to form both passwords.

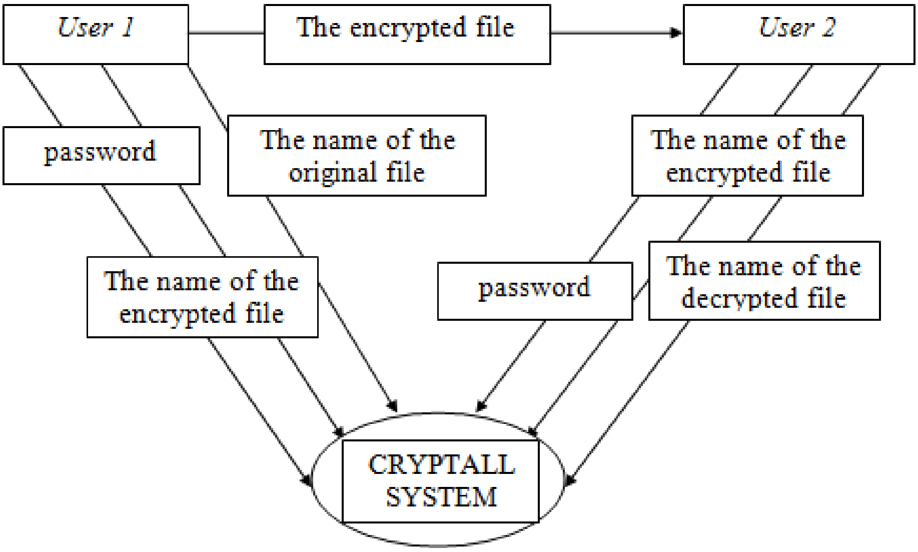
User 1
The encrypted file
User 2
password
The name of the original file
The name of the encrypted file
The name of the encrypted file
password
The name of the decrypted file
CRYPTALL SYSTEM

Fig. 1. Context Diagram of CRYPTALL 4

***Experimental Measurements.*** To evaluate the performance of our algorithm, we use with different size of files. We denote by *t (k, L)* the time (in millisecond) that is needed to encrypt or decrypt (because of symmetry). The file size is in kilobytes for passwords of length *L*. Then the value of *t(k, L)* can be represented by the following matrix (Fig. 2).

| L\k | 3000 | 4000 | 5000 | 6000 |
|---|---|---|---|---|
| 4 | 1143 | 1535 | 1755 | 2120.75 |
| 8 | 2162.25 | 2999.75 | 3452.5 | 4150 |
| 12 | 3070.5 | 4108 | 5061.25 | 6053 |
| 16 | 4090.5 | 5429.25 | 6673.5 | 7945.75 |
| 20 | 5131.75 | 6778.75 | 8303 | 9873.75 |

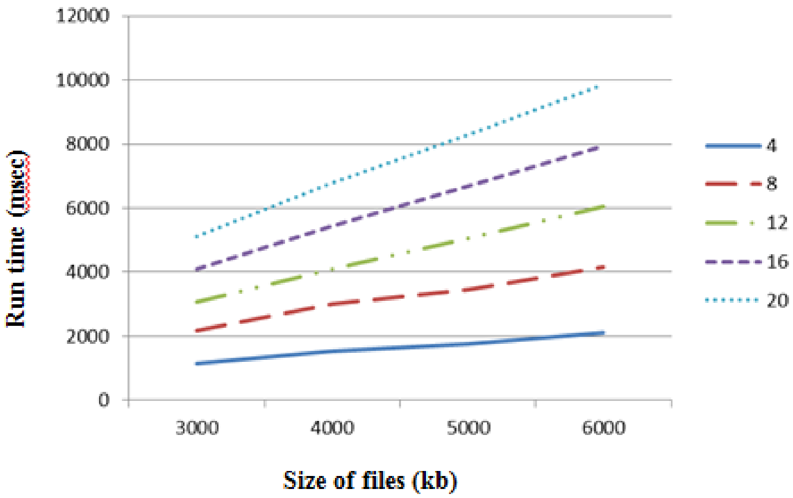

Fig. 2. Run time for CRYPTALL 4 System

Computer experiment justifies that in implemented case the speed of execution of decryption or encryption procedures are essentially higher than in the case of stream cipher of [2] used for GIS protection. New algorithm has essentially better mixing properties (see [47]).

## 5. Conclusion

The main theoretical result of the paper is explicit construction of the family of multivariate map of affine maps $F_n$ with the trapdoor accelerator of linear degree cn, $c=3/4$ acting on affine space $K^n$ defined over arbitrary commutative ring $K$ with at least 3 elements. Corresponding cipher has execution speed of kind $¼\ n^2+O(n)$ which is proportional to the length of active password of size $0(1)$. The decryption procedure takes the same time with the encryption process. In the case of choice of special linear conjugation $T$ it has nice mixing properties: change of single character of the plaintext or active password leads to the change of ≥ 98% of characters of corresponding ciphertext.

So $F_n$ based cipher can provide essentially better security than the cipher selected in [2]. The disadvantage of $F_n$ is speed of encryption $O(n^2)$ but not $O(n)$. So the usage of $F_n$ will drastically improve the security level of GIS protection but essentially slow down of speed of spatial information processing.

Noteworthy that speed of processing is very important parameter. That is why we suggest usage of ciphers $DE(n, K)$ and $DA(n, K)$ for GIS protection which are more robust in the comparison of cipher chosen in [2], they have essentially better mixing properties and practically resistant against linearisation attacks. Ciphers $DE(n, K)$ and $CE(n, K)$ can be chosen in the case of tasks where security aspects are more important than the execution speed.

**ACKNOLEDGEMENTS**. This research is partially supported by the Fellowship of British Academy for RaR.

## REFERENCES

1. S. Dhanjal, Y. Khmelevsky, M. Govorov, V. A. Ustymenko, P. N. Sharma, Security solutions for spatial data in storage - (Implementation case within oracle 9iAS), *8th World MultiConference on Systemics, Cybernetics and Informatics, Vol Ii, Proceedings*, pp. 318–323, 2004.

2. M. Govorov, Y. Khmelevsky, V. Ustimenko, A. Khorev, Security for GIS N-tier architecture, *Developments in Spatial Data Handling*, pp. 71–83, 2005.

3. Y. Khmelevsky, S. Dhanjal, Information Security and Data Protection in Computer Science Education, in *12th Western Canadian Conference Education on Computing Education (WCCCE-2007), Thompson Rivers University, Kamloops, Canada, May*, 2007, pp. 3–5.

4. V. Ustimenko, CRYPTIM: Graphs as tools for symmetric encryption, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 2227, 2001.

5. F. Lazebnik, V. Ustimenko (1993). Some Algebraic Constractions of Dense Graphs of Large Girth and of Large Size, DIMACS series in Discrete Mathematics and Theoretical Computer Science, 10, p. 75-93. https:/doi.org/10.1090/dimacs/010/07

6. Lazebnik F., Ustimenko V. A. and Woldar A. J. (1995). New Series of Dense Graphs of High Girth // Bull (New Series) of AMS, 32, No. 1, p. 73-79. https:/doi.org/10.1090/S0273-0979-1995-00569-0

7. V. Ustimenko, On graph-based cryptography and symbolic computations, *Serdica Journal of Computing*, vol. 1, no. 2, pp. 131–156, 2007.

8. Aneta Wroblewska, On some properties of graph based public keys, Albanian Journal of Mathematics, Albanian J. Math. 2(3), 229-234, (2008).

9. J. S. Kotorowicz and V. A. Ustimenko, On the implementation of cryptoalgorithms based on algebraic graphs over some commutative rings, in *Condensed Matter Physics*, vol. 11, no. 2, 2008.

10. V. Ustimenko (1998), Coordinatisation of Trees and their Quotients, in the Voronoj's Impact on Modern Science, Kiev, Institute of Mathematics, 2, p. 125-152.

11. V. Ustimenko (2007). Linguistic Dynamical Systems, Graphs of Large Girth and Cryptography, Journal of Mathematical Sciences, Springer, 140, No. 3, p. 412-434. https:/doi.org/ 10.1007/s10958-007-0453-2

12. V. A. Ustimenko, U. Romanczuk, On Dynamical Systems of Large Girth or Cycle Indicator and their applications to Multivariate Cryptography, in "Artificial Intelligence, Evolutionary Computing and Metaheuristics", In the footsteps of Alan Turing Series: Studies in Computational Intelligence, 427, 2012, p. 231-256, https:/doi.org/ 10.1007/978-3-642-29694-9_10

13. V. A. Ustimenko, U. Romanczuk, On Extremal Graph Theory, Explicit Algebraic Constructions of Extremal Graphs and Corresponding Turing Encryption Machines, in "Artificial Intelligence, Evolutionary Computing and Metaheuristics", In the footsteps of Alan Turing Series: Studies in Computational Intelligence, Vol. 427, Springer, January, 2013, 257-285.

14. F. Lazebnik, V. Ustimenko and A. J. Woldar (1996). A characterisation of the components of the graphs D(k,q), Discrete Mathematics, 157, p. 271-283. https:/doi.org/10.1016/S0012-365X(96)83019-6

15. V. A. Ustimenko. On the extremal graph theory and symbolic computations, Dopovidi National Academy of Sci, Ukraine, 2013, No. 2, p. 42-49.

16. V. Ustimenko, On new results on extremal graph theory, theory of algebraic graphs, and their applications, Reports of National Acad. of Sci. of Ukraine, 2022, N4, pp. 25-32.

17. V. Ustimenko, Maximality of affine group, hidden graph cryptosystem and graph's stream ciphers, Journal of Algebra and Discrete Mathematics, 2005, v. 1, pp. 51-65.

18. G. Margulis (1988). Explicit group-theoretical constructions of combinatorial schemes and their application to design of expanders and concentrators, Probl. Peredachi Informatsii, 24, No. 1, p. 51-60.

19. A. Lubotsky, R. Philips, P. Sarnak (1989). Ramanujan graphs, J. Comb. Theory, 115, No. 2, p. 62-89. https://doi.org/10.1007/BF02126799

20. T. Shaska, V. Ustimenko (2009). On the homogeneous algebraic graphs of large girth and their applications, Linear Algebra and its Applications, 430, No. 7, p. 1826-1837. https:/doi.org/10.1016/j.laa.2008.08.023

21. F. Buekenhout (editor), Handbook in Incidence Geometry, Ch. 9, North Holland, Amsterdam, 1995.

22. V. Ustimenko (1989). Affine system of roots and Tits geometries, Voprosy teorii grupp i gomologicheskoy algebry, Yaroslavl, p. 155-157.

23. M. Polak, V. A. Ustimenko (2012). On LDPC Codes Corresponding to Infinite Family of Graphs A(k,K). Proceedings of the Federated Conference on Computer Science and Information Systems (FedCSIS), CANA, Wroclaw, p. 11-23.

24. Ustimenko, V. (2021). On semigroups of multivariate transformations constructed in terms of time dependent linguistic graphs and solutions of Post Quantum Multivariate Cryptography. Cryptology ePrint Archive: Report 2021/1466. Retrieved from https://eprint.iacr.org/2021/1466.pdf

25. D. MacKay and M. Postol (2003). Weakness of Margulis and Ramanujan - Margulis Low Dencity Parity Check Codes, Electronic Notes in Theoretical Computer Science, 74, p. 97-104. https:/doi.org/ 10.1016/S1571- 0661(04)80768-0

26. V. Ustimenko (2009). Algebraic groups and small world graphs of high girth, Albanian Journal of Mathematics, 3, No. 1, p. 25-33.

27. V. Ustimenko, On Extremal Algebraic Graphs and Multivariate Cryptosystems, Cryptology ePrint Archive, reprint 2022/1537.

28. V. Ustimenko, On the extremal graph theory for directed graphs and its cryptographical applications. In: T. Shaska, W.C. Huffman, D. Joener and V. Ustimenko, Advances in Coding Theory and Cryptography, Series on Coding and Cryptology, vol. 3, 181-200 (2007).

29. V. Ustimenko, Graphs in terms of Algebraic Geometry, symbolic computations and secure communications in Post-Quantum world, Editorial House of University of Maria Curie – Sklodowska, Lublin, November, 2022, 196 pages.

30. Geetha N K, Ragavi V, Graph Theory Matrix Approach in Cryptography and Network Security, Proceedings of 2022 Algorithms, Computing and Mathematics Conference (ACM), 29-30 Aug. 2022, https://ieeexplore.ieee.org/document/10202460

31. Costache, A., Feigon, B., Lauter, K., Massierer, M., Puskás, A. (2019). Ramanujan Graphs in Cryptography. In: Balakrishnan, J., Folsom, A., Lalín, M., Manes, M. (eds) Research Directions in Number Theory. Association for Women in Mathematics Series, vol. 19. Springer, Cham. https://doi.org/10.1007/978-3-030-19478-9_1

32. P.L. K. Priyadarsini, A Survey on some Applications of Graph Theory in Cryptography, Journal of Discrete Mathematical Sciences and Cryptography, https://www.tandfonline.com/doi/abs/10.1080/09720529.2013.878819

33. W. M. Al Etaiwi, Encryption Algorithm Using Graph Theory, Journal of Scientific Research & Reports, 3(19): 2519-2527, 2014; Article no. JSRR.2014.19.004.

34. Samid Gideon, Denial Cryptography based on Graph Theory, US patent 6823068-2004 http://www.patentstorm.us/patents/6823068.html

35. Lothrop Mittenthal, Sequencings and Directed Graphs with Applications to Cryptography, S.W. Golomb et al. (Eds.): *Springer-Varlag LNCS* 4893, pp. 70–81, 2007.

36. Moni Naor and Adi Shamir. Visual cryptography. In *Advances in Cryptology - EURO-CRYPT'94, LNCS*, vol. 950, pp. 1–12, 1994.

37. Steve Lu, Daniel Manchala, Rafail Ostrovsky, Visual Cryptography on Graphs, *COCOON 2008*: pp. 225–234, 2008.

38. William Stallings, Cryptography and Network Security Principles and Practices, Prentice Hall India, 2006.

39. Dawn Song, David Zuckermany, J. D. Tygar, Expander Graphs for Digital Stream Authentication and Robust Overlay Networks, *Proceedings of the 2002 IEEE Symposium on Security and Privacy (S&P.02)*, 2002.
40. M Yamuna, Meenal Gogia, Ashish Sikka and Md. Jazib Hayat Khan, "Encryption using graph theory and linear algebra", *International Journal of Computer Application*, pp. 2250-1797, 2012.
41. A Paszkiewicz et al., *Proposals of graph based ciphers theory and implementations. Research Gate*, 2001.
42. Cusack, B.; Chapman, E. Using graphic methods to challenge cryptographic performance. In Proceedings of the 14th Australian Information Security Management Conference, Edith Cowan University, Perth, Australia, 5–6 December 2016; pp. 30–36. [Google Scholar].
43. Chapman, E. Using Graphic Based Systems to Improve Cryptographic Algorithms. Ph.D. Thesis, Auckland University of Technology, Auckland, New Zealand, 2016. [Google Scholar].
44. Kinani, E.H.E. Fast Mapping Method based on Matrix Approach For Elliptic Curve Cryptography. *Int. J. Inf. Netw. Secur. (IJINS)* 2012, *1*, 54–59.
45. M. Klisowski. Zwiększenie bezpieczeństwa kryptograficznych algorytmów wielu zmiennych bazujących na algebraicznej teorii grafów. Rozprawa doktorska, Politechnika Częstochowska, Częstochowa, 2014.
46. Vasyl Ustimenko, Aneta Wroblewska, On the key exchange and multivariate encryption with nonlinear polynomial maps of stable degree, Annales of UMCS, Informatica, Vol 13, No 1 (2013), pp 63-80. http://dx.doi.org/10.2478/v10065-012-0047-6
47. Pustovit O.S. Application of the theory of extreme graphs to modern problems of information security. - Dissertation for obtaining the scientific degree of candidate of technical sciences in the specialty 05.13.06 - Information technologies. – Institute of Telecommunications and Global Information Space of the National Academy of Sciences of Ukraine, Kyiv, 2021.

*The article was received 11.05.2023 and was accepted after revision 25.08.2023*

**В.О. Устименко, О.С. Пустовіт**

## ПРО БЕЗПЕКУ ГІС-СИСТЕМ З N-РІВНЕВОЮ АРХІТЕКТУРОЮ ТА СІМЕЙСТВА АЛГОРИТМІВ ШИФРУВАННЯ, ВИЗНАЧЕНИХ ЗА ГРАФАМИ

**Анотація.** Відкриття опису *q*-регулярного дерева в термінах нескінченної системи квадратних рівнянь над скінченним полем *Fq* мало вплив на розвиток Інформатики, зокрема теорії криптографічних алгоритмів, що визначаються за графами. Це стимулювало розвиток конструкцій безпечних потокових алгоритмів шифрування. Виявилося, що такі інструменти шифрування можна ефективно використовувати в системах захисту ГІС, що вживають *N*-рівневу архітектуру. Ми оглянемо відомі алгоритми шифрування, засновані на апроксимаціях регулярних дерев, їх модифікації, визначені над арифметичними кільцями, та програмні реалізації цих алгоритмів. Крім того, будуть представлені нові більш безпечні алгоритми потокового шифрування, придатні для захисту ГІС.

Алгоритми будуються з використанням блукань на вершинах дводольних регулярних графів *D(n,K)*, визначених за скінченним комутативним кільцем *K* з одиницею та нетривіальною мультиплікативною групою *K**. Долі таких графів є *n*-вимірними афінними просторами над кільцем *K*. Блукання парної довжини визначає перетворення переходу від початкової до останньої вершини з однієї з долей графу. Отже, афінний простір *Kn* можна розглядати як простір відкритих текстів, а блукання на графі паролем, який визначає перетворення, що шифрує.

При певних обмеженнях на паролі досягається ефект, коли різним паролям з *(K*)2s, s <[(n+5)/2]/2* відповідають різні шифрограми обраного відкритого тексту з *Kn*. У 2005 році такий алгоритм у випадку скінченного поля F127 використовувався для захисту

ГІС. З цього часу властивості алгоритмів шифрування з використанням графів *D(n, K)* (швидкодія, властивості зміни, степінь та густина поліноміального перетворення шифрування) були ретельно досліджені. Було оцінено складність атак лінеаризації та знайдено модифікації цих алгоритмів із стійкістю до атак лінеаризації. Виявилося, що разом з графами *D(n, K)* можна ефективно використовувати й інші алгебраїчні графи з подібними властивостями, такі як графи *A(n,K)*.

У статті розглядаються кілька розв’язань задачі захисту геологічної інформаційної системи від можливих кібератак за допомогою потокових алгоритмів, що спираються на графи. Вони мають істотні переваги в порівнянні з реалізованими раніше алгоритмами.

*Стаття надійшла до редакції 11.05.2023 і прийнята до друку після рецензування 25.08.2023*

**Устименко Василь Олександрович**
доктор фізико-математичних наук, професор, завідувач відділу інформаційної безпеки Інституту телекомунікацій і глобального інформаційного простору НАН України, Університет Роял Холловей (Лондон)
**Адреса робоча:** 03186 Україна, м. Київ, Чоколівський бульвар, 13
ORCID ID: https://orcid.org/0000-0002-2138-2357 ***e-mail:*** vasulustimenko@yahoo.pl

**Пустовіт Олександр Сергійович**
кандидат технічних наук, науковий співробітник Інституту телекомунікацій і глобального інформаційного простору НАН України
**Адреса робоча:** 03186 Україна, м. Київ, Чоколівський бульвар, 13
ORCID ID: https://orcid.org/0000-0002-3232-1787 ***e-mail:*** sanyk_set@ukr.net