***Yevseiev Serhii**, Simon Kuznets Kharkiv National University of Economics, Associate professor, Senior Research, Ph.D, Department of Information Systems*
***Sverdlo Tamara**, Simon Kuznets Kharkiv National University of Economics, Lecturer researcher, Department of Information Systems*
***Korol Olha**, Simon Kuznets Kharkiv National University of Economics, Associate professor, Ph.D, Department of Information Systems*

Integrated security mechanisms and reliability of data in information systems based on the theoretical coding scheme

Abstract: In this article, we are interested in the study of cryptography to protect information to ensure the secure transmission of data in information systems. Thus, a comparative study of integrated security and reliability of the transmission based on the use of asymmetric encryption systems McEliece and Niederreiter based on error correcting codes m-ary data mechanisms.

Keywords: Asymmetric cryptosystems McEliece and Niederreiter, security, reliability.

***Yevseiev Serhii**, Université nationale d'économie de Kharkiv Simon Kuznets, maître de conférences, chargé de recherche, Ph.D, département des systèmes d'information*
***Sverdlo Tamara**, Université nationale d'économie de Kharkiv Simon Kuznets, enseignante chercheur, département des systèmes d'information*
***Korol Olha**, Université nationale d'économie de Kharkiv Simon Kuznets, maître de conférences, Ph.D, département des systèmes d'information*

Mécanismes intégrés de sécurité et de fiabilité des données dans les systèmes d'information basés sur la théorie des codes correcteurs d'erreurs

Résumé: Dans cet article, on s'intéresse à l'étude des moyens cryptographiques de la protection de l'information afin d'assurer la sécurité de la transmission de données dans les systèmes d'information. On réalise ainsi une étude comparative des mécanismes intégrés de sécurité et de fiabilité de la transmission de données basés sur l'utilisation de systèmes de chiffrement asymétrique de McEliece et de Niederreiter basés sur les codes correcteurs d'erreurs m-aires.

Mots clés: cryptosystèmes asymétriques de McEliece et de Niederreiter, sécurité, fiabilité.

1. Problématique et cadre de référence. Le développement rapide des technologies de l'information et de la communication pose aujourd'hui de manière cruciale le problème de la protection de l'information. Ce développement permet ainsi de construire des systèmes d'information sophistiqués avec une architecture distribuée, réunissant un grand nombre de segments situés à une certaine distance les uns des autres. Tout cela entraîne une augmentation du nombre de noeuds de réseau et le nombre de différentes lignes de communication entre eux, ce qui accroît les risques d'une connexion non autorisée au système d'information et de l'accès aux informations confidentielles [1].

2. Étude des mécanismes de protection de l'information dans les systèmes d'information. Les mécanismes de sécurité de l'information dans les systèmes d'information sont basés dans la plupart des cas sur des techniques de cryptographie assurant d'une part les modèles des systèmes secrets de stabilité temporelle (les systèmes cryptographiques symétriques où le cryptogramme est formé par l'exécution multiple des mêmes groupes de transformation, il en résulte un haut niveau de mélange et de dispersion des blocs d'information de données) et, d'autre part, la résistance prouvable (les systèmes de chiffrement asymétriques où la tâche de briser les données clés est réduite à la résolution de problèmes mathématiques complexes). Le principal inconvénient de systèmes cryptographiques symétriques (systèmes de chiffrement traditionnel) est l'absence d'une justification mathématique rigoureuse de la résistance cryptographique. Par contre, le principal inconvénient de systèmes cryptographiques asymétriques (systèmes de chiffrement à clé publique) est la vitesse de chiffrement qui est de 3-5 fois plus faible par rapport à la vitesse des algorithmes de chiffrement symétriques. Une voie prometteuse dans le développement des moyens

cryptographiques de protection de l'information de résistance prouvable est des mécanismes de chiffrement basés sur les schémas cryptographiques de McEliece et de Niederreiter, dont la construction est fondée sur la réduction du problème d'attaque de données clés à l'analyse théorique et à la résolution numérique du problème de décodage d'un code aléatoire [2, 3]. L'analyse montre en effet que l'utilisation de ces schémas permet de réaliser la transformation cryptographique rapide tout en assurant la résistance prouvable, les résultats de l'analyse sont présentés dans le tableau 1.

Tableau 1

Les résultats des études comparatives de l'efficacité des méthodes cryptographiques de protection de l'information à un niveau fixe de résistance

Méthodes de transformations cryptographiques	Modèle de sécurité	Longueur des données clés, bit	Vitesse de cryptage bit/s	Fonctions supplémentaires
Algorithmes de chiffrement symétrique par bloc	Sécurité pratique	128, 256, 512		Aucune
Algorithmes de chiffrement symétrique par flot	Sécurité pratique	128, 256, 512	$10^7 – 10^{10}$	$10^6 – 10^9$
Algorithmes de chiffrement RSA	Sécurité prouvable	3248 (128), 15424 (256)	$10^2 – 10^3$	Aucune
Algorithmes cryptographiques asymétriques basés sur les courbes elliptiques	Sécurité prouvable	283 (128), 571 (256)	$10^3 – 10^4$	Aucune
Algorithmes de chiffrement asymétrique basés sur schémas cryptografiques	Sécurité prouvable	$0,5 \cdot 10^6$ (128), $2 \cdot 10^6$ (256)	$10^6 – 10^8$	Contrôle des erreurs, augmentation de fiabilité

La complexité de leur mise en œuvre est comparable aux algorithmes cryptographiques symétriques (algorithmes de chiffrement symétrique par bloc). En outre, l'utilisation pratique de ces schémas cryptographiques permet d'appliquer l'infrastructure à clés publiques et d'établir des mécanismes intégrés de transformation cryptographique de données et du codage canal afin d'assurer la sécurité et la fiabilité de la transmission de données.

Ainsi, d'après les résultats de l'analyse comparative ci-dessus, le chiffrement asymétrique basé sur les schémas cryptographiques permet de réaliser la protection cryptographique de l'information à l'aide de la technique à clé publique et d'assurer le cryptage de l'information à la vitesse du chiffrement symétrique par bloc. De plus, les chercheurs soulignent dans ses travaux que l'utilisation pratique des schémas cryptographiques de la protection de l'information permet d'assurer d'une manière

cohérente la sécurité et la fiabilité de données transmises à l'aide des mécanismes de codage canal et de cryptage [4, 5]. Par conséquent, l'utilisation des schémas cryptographiques est économiquement plus avantageuse que l'utilisation de l'ensemble de différents mécanismes de chiffrement et de codage canal, visant à résoudre les tâches isolées. On observe aussi une réduction significative du coût total des services informatiques par unité d'information traitée et transmise.

Le classement général des méthodes de conception des schémas cryptographiques est représenté sur la Fig. 2.

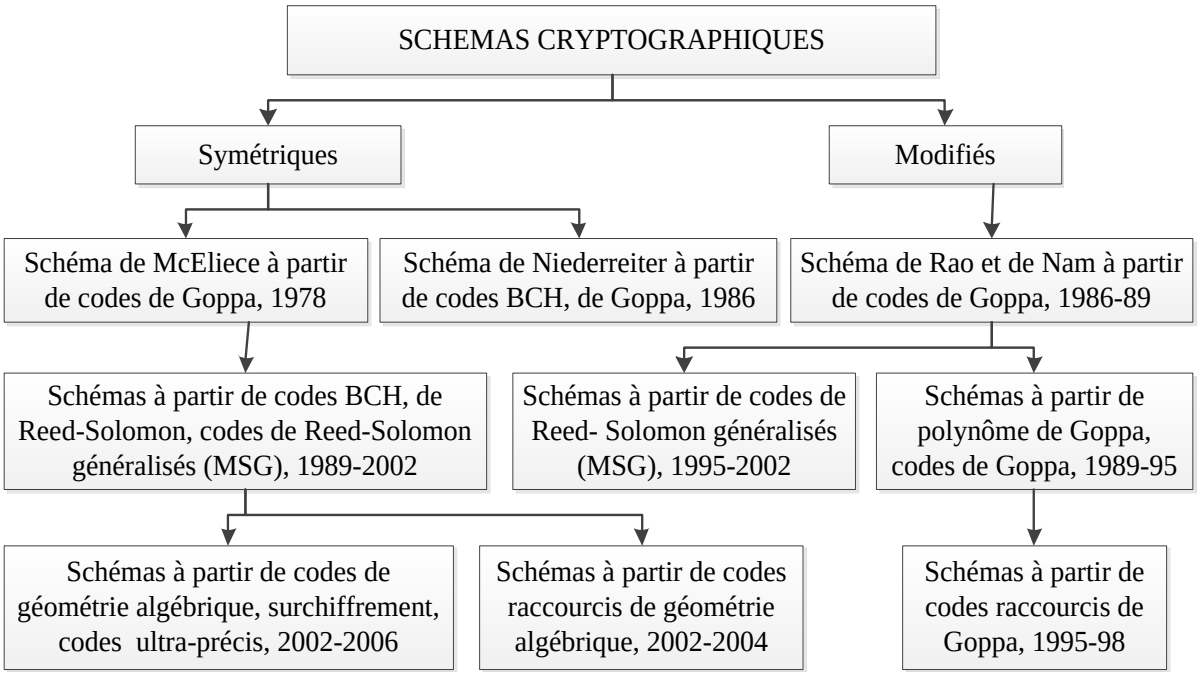

Fig. 2. Classement général des méthodes de conception des schémas cryptographiques

L'analyse accomplie et les études comparatives ont montré que les moyens cryptographiques asymétriques de protection de l'information sont conçus sur: la dissimulation du code de matrice génératrice (schéma de McEliece) et la dissimulation du code de matrice de contrôle (schéma de Niederreiter) [5]. Or, les schémas cryptographique à partir de codes en bloc linéaires non binaires représentés sur des courbes algébriques (codes de géométrie algébrique) sont les plus résistants [4, 5]. D'une part, les constructions similaires sont résistantes aux attaques proposées par Sidelnikov [4], et, d'autre part, elles fournissent une grande fiabilité et efficacité de transmission de données.

3. Construction générale des schémas cryptographiques. Considérons la structure générale des schémas cryptographiques. Fixons le corps fini *GF(q)*.

Considérons l'espace vectoriel $GF^n(q)$ comme l'ensemble des *n*-suites d'éléments de *GF (q)* muni de l'addition composante par composante et la multiplication par un scalaire. *Le code linéaire C (n, k, d)* est un sous-espace en $GF^n(q)$, c'est-à-dire l'ensemble non vide de *n*-suites (*mots de code*) sur *GF (q), k – la dimension* du sous-espace linéaire, *d – la distance minimale de code* (poids minimum d'un mot de code non nul).

Le but principal de codage de l'information est de contrôler (détecter et corriger) les erreurs qui sont produites lors de la transmission d'un message à travers un canal bruité. Pour contrôler les erreurs, l'encodeur ajoute de la redondance (partie de contrôle de longueur *r*, $r = n - k$) dans les données transmises. Pendant la réception, tout en analysant les propriétés de la somme de contrôle et sa conformité avec les données transmises, le décodeur réduit l'impact des erreurs apparues lors de la transmission.

La tâche de décodage peut être effectuée d'une manière efficace, mais à complexité polynômiale, pour une catégorie restreinte de codes, comme les codes BCH et les codes de Reed-Solomon. L'un des algorithmes plus efficaces pour le décodage algébrique des codes BCH est l'algorithme de Berlekamp-Massey et ses modifications (améliorations). Il est bien connu que l'algorithme de Berlekamp-Massey contient le nombre de multiplications de l'ordre t^2, soit, formellement, la complexité de l'algorithme est de $O(t^2)$, où *t* – la capacité de correction du code. Si *t* est grand, on utilise algorithme rapide de Berlekamp-Massey permettant de réduire la complexité de calcul de l'algorithme. Encore plus efficace, en termes de complexité de calcul, est l'algorithme récurrent de Berlekamp-Massey. La complexité asymptotique de décodage des codes de Reed-Solomon dans ce cas ne dépasse pas $O(n\log^2 n)$, et elle est très proche de la valeur $O(n\log n)$.

Le décodage d'un code générique est un des problèmes de calcul assez complexes, et la complexité de ces solutions est en croissance exponentielle. Ainsi, pour le décodage corrélation du code générique *(n, k, d)* sur *GF (q)*, il est nécessaire, dans le cas général, de comparer la suite reçue avec tous les mots de code q^k et de sélectionner le plus proche (en métrique de Hamming). Même si la taille de *n, k, d* et *q* n'est pas grande, la tâche de décodage corrélation est très laborieuse. Cette thèse est la base de tous les systèmes cryptographiques fondés sur le codage algébrique en bloc.

4. Conception des schémas cryptographiques à la base des courbes elliptiques. Fixons le corps fini *GF (q)*. Soit *X* – courbe algébrique projective lisse d'un espace projectif P^n sur *GF (q)*, *g = g (X)* – genre de courbe, *X (GF (q))* – ensemble de points sur un corps fini, *N = X (GF (q))* – leur nombre. Soit *C* – classe de diviseurs sur *X*

de degré $\alpha > g - 1$. Alors C définit une application $\varphi: X \to P^{k-1}$, où $k \geq \alpha - g + 1$. Le code est spécifié par $y_i = \varphi(x_i)$.

Le nombre de points à l'intersection $\varphi(X)$ avec l'hyperplan égal à α, c'est-à-dire $n - d \leq \alpha$. Ce schéma permet de construire des codes aux paramètres $k + d \geq n - g + 1$, dont la longueur n est inférieure ou égale au nombre de points de la courbe X. Si $2g < \alpha \leq n$, le code de géométrie algébrique comporte des paramètres $(n, n - \alpha + g - 1, d^{\perp})$, $d^{\perp} \geq \alpha - 2g + 2$. Son code dual est aussi le code de géométrie algébrique et comporte des paramètres $(n, n - \alpha + g - 1, d^{\perp})$, $d^{\perp} \geq \alpha - 2g + 2$. Introduisons ainsi la notion de code de géométrie algébrique.

Définition 1. Soit X – courbe algébrique projective lisse dans l'espace projectif P^n, c'est-à-dire l'ensemble des solutions de l'équation algébrique irréductible homogène de degré $degX$ avec des coefficients de $GF(q)$. Considérons les multiciplicités correspondant aux hyperplans projectifs définis en P^n par équations $F = 0$, où F – monômes homogènes de degré $degF$. Soit $I(i_1, i_2, \ldots, i_n)$ – suite d'information. *Le code de géométrie algébrique sur une courbe X* sur $GF(q)$ – est un code linéaire de longueur $n \leq N$, dont les mots de code $C(c_1, c_2, \ldots, c_n)$ sont définis par l'égalité $\sum_{i=0}^{k-1} i_j F_j(P_i) = c_i$, où $P_i(X_i, Y_i, Z_i)$ – points projectives de la courbe X, c'est-à-dire (X_i, Y_i, Z_i) – solutions des équations algébriques homogènes définissant la courbe X, $i = \overline{1,n}$; $F_j(P_i)$ – valeurs des fonctions génératrices dans les points de la courbe.

Cette définition est équivalente à la représentation de matrice de code de géométrie algébrique : $G(i_0, i_1, \ldots, i_{k-1})^T = (c_0, c_1, \ldots, c_{n-1})$, où G – matrice génératrice de dimension $k \times n$, $k = \alpha - g + 1$, $\alpha = degX \cdot degF$.

$$G = \begin{pmatrix} F_0(P_0) & F_0(P_1) & \ldots & F_0(P_{n-1}) \\ F_1(P_0) & F_1(P_1) & \ldots & F_1(P_{n-1}) \\ \ldots & \ldots & \ldots & \ldots \\ F_{k-1}(P_0) & F_{k-1}(P_1) & \ldots & F_{k-1}(P_{n-1}) \end{pmatrix} = \left\| F_j(P_i) \right\|_{n,k}.$$

Définition 2. La courbe elliptique (CE) dans l'espace affine A^2 sur le corps fini $GF(q)$ est une courbe lisse définie par l'équation :

$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$,

ou dans P^2 définie par l'équation homogène :

$y^2z + a_1xyz + a_3yz^2 = x^3 + a_2x^2z + a_4xz + a_6z^3$,

$a_i \in GF(q)$, genre de courbe $g = 1$.

Affirmation 1. Le code de géométrie algébrique *(n, k, d)* sur une courbe elliptique *GF(q)* (code elliptique) construit par l'application du type $\varphi: EC \to P^{k-1}$, est lié par les caractéristiques $k + d \geq n$, dans laquelle : $n \leq 2\sqrt{q} + q + 1$, $k \geq \alpha$, $d \geq n - \alpha$, $\alpha = 3 \cdot degF$.

Définition 3. Soit *X* – courbe algébrique projective lisse dans l'espace P^n, soit ensemble des solutions de l'équation algébrique irréductible homogène de degré *degX* avec des coefficients de *GF (q), F* – monômes homogènes de degré *degF*. *Le code de géométrie algébrique sur une courbe X* sur *GF (q)* est un code linéaire de longueur $n \leq N$, dont les mots de code $C(c_1, c_2, \ldots, c_n)$ sont définis par l'égalité $d + g - 1$ des équations

$$\sum_{i=0}^{n-1} c_i F_j(P_i) = 0,$$

où $c_i \in GF(q)$, $d \geq \alpha - 2g + 2$, $\alpha = degX \cdot degF$.

Cette définition est équivalente à la représentation de matrice de code de géométrie algébrique : $H(c_0, c_1, \ldots, c_{n-1})^T = 0$, où *H* – matrice de contrôle de dimension $r \times n$, $r = n - k = d + g - 2$ de type :

$$H = \begin{pmatrix} F_0(P_0) & F_0(P_1) & \ldots & F_0(P_{n-1}) \\ F_1(P_0) & F_1(P_1) & \ldots & F_1(P_{n-1}) \\ \ldots & \ldots & \ldots & \ldots \\ F_{r-1}(P_0) & F_{r-1}(P_1) & \ldots & F_{r-1}(P_{n-1}) \end{pmatrix} = \left\| F_j(P_i) \right\|_{n,r}.$$

Les definitions 1-2 et le résultat de l'affirmation 1 permettent de définir le schéma cryptologique de McEliece à la base des codes elliptiques qui suit. Soit G^{EC} est la matrice génératrice du code elliptique *(n, k, d)* sur *GF(q)* de type :

$$G^{EC} = \begin{pmatrix} F_0(P_0) & F_0(P_1) & \ldots & F_0(P_{n-1}) \\ F_1(P_0) & F_1(P_1) & \ldots & F_1(P_{n-1}) \\ \ldots & \ldots & \ldots & \ldots \\ F_{k-1}(P_0) & F_{k-1}(P_1) & \ldots & F_{k-1}(P_{n-1}) \end{pmatrix} = \left\| F_j(P_i) \right\|_{n,k}.$$

et de dimension $k \times n$, $k = \alpha$, $\alpha = 3 \cdot degF$.

Soit *X* – la matrice $k \times k$ non dégénérée sur *GF (q), D* – matrice diagonale avec des éléments diagonaux sont non nuls, P – matrice de permutation de taille $n \times n$. Définissons

alors *le schéma asymétrique de McEliece avec le code elliptique* : la clé publique – la matrice $G_X^{EC} = X \cdot G^{EC} \cdot P \cdot D$, la clé privé – matrice *X, P, D.*

L'information fermée (codogramme) représente un vecteur de longueur *n* et est calculée par la formule: $c_X^* = i \cdot G_X^{EC} + e$, où le vecteur $c_X = i \cdot G_X^{EC}$ appartient au code elliptique *(n, k, d)* avec la matrice génératrice G_X^{EC}, *i* – vecteur de l'information *k* bits, le vecteur *e* – vecteur secret d'erreur de poids $\leq t$. Le système de transmission d'un message secret de l'abonné A à l'abonné B dans le schéma asymétrique de McEliece tout en utilisant des codes elliptiques est représenté sur la fig.3. Pour définir le schéma asymétrique de Niederreiter basé sur les codes elliptiques, utilisons une autre définition du code de géométrie algébrique.

Définition 3. Soit *X* – courbe algébrique projective lisse dans l'espace P^n, soit ensemble des solutions de l'équation algébrique irréductible homogène de degré *degX* avec des coefficients de *GF (q), F* – monômes homogènes de degré *degF*. *Le code de géométrie algébrique sur une courbe X* sur *GF (q)* est un code linéaire de longueur $n \leq N$, dont les mots de code $C(c_1, c_2, \ldots, c_n)$ sont définis par , l'égalité $d + g - 1$ des équations

$$\sum_{i=0}^{n-1} c_i F_j(P_i) = 0,$$

où $c_i \in GF(q)$, $d \geq \alpha - 2g + 2$, $\alpha = degX \cdot degF$.

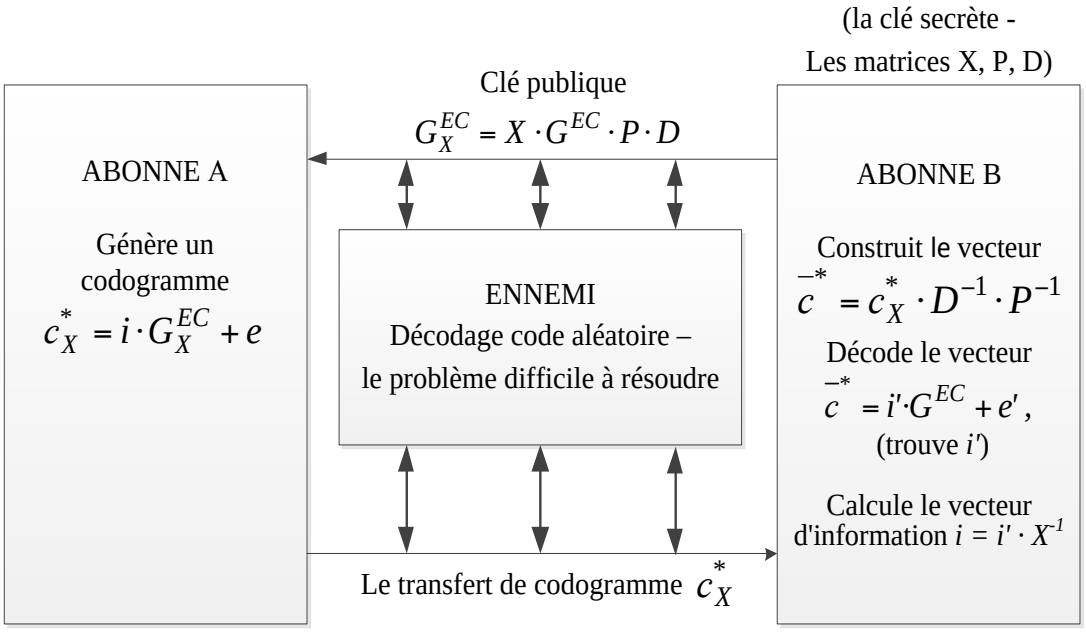

Fig.3. Schéma de transmission du codogramme dans le schéma asymétrique de McEliece tout en utilisant des codes elliptiques

Cette définition est équivalente à la représentation de matrice de code de géométrie algébrique : $H(c_0, c_1, ..., c_{n-1})^T = 0$, où H – matrice de contrôle de dimension $r \times n$, $r = n - k = d + g - 2$ de type :

$$H = \begin{pmatrix} F_0(P_0) & F_0(P_1) & \dots & F_0(P_{n-1}) \\ F_1(P_0) & F_1(P_1) & \dots & F_1(P_{n-1}) \\ \dots & \dots & \dots & \dots \\ F_{r-1}(P_0) & F_{r-1}(P_1) & \dots & F_{r-1}(P_{n-1}) \end{pmatrix} = \left\| F_j(P_i) \right\|_{n,r}.$$

La définition 3 et le résultat de l'affirmation 2 permettent de définir le schéma cryptologique de Niederreiter à la base des codes elliptiques qui suit. Soit H^{EC} est la matrice de contrôle du code elliptique *(n, k, d)* sur *GF(q)* de type :

$$H^{EC} = \begin{pmatrix} F_0(P_0) & F_0(P_1) & \dots & F_0(P_{n-1}) \\ F_1(P_0) & F_1(P_1) & \dots & F_1(P_{n-1}) \\ \dots & \dots & \dots & \dots \\ F_{r-1}(P_0) & F_{r-1}(P_1) & \dots & F_{r-1}(P_{n-1}) \end{pmatrix} = \left\| F_j(P_i) \right\|_{n,r}$$

Et de la dimension $r \times n$, $r = \alpha$, $\alpha = 3 \cdot degF$.

Soit X – matrice $k \times k$ non dégénérée sur *GF (q), D* – matrice diagonale avec des éléments diagonaux sont non nuls, P – matrice de permutation de taille $n \times n$. Définissons alors *le schéma asymétrique de Niederreiter avec le code elliptique* : la clé publique – matrice $H_X^{EC} = X \cdot H^{EC} \cdot P \cdot D$, la clé privé – matrice *X, P, D.*

L'information fermée (codogramme) représente un vecteur de longueur n et est calculée par la formule: $S_X = e \cdot \left(H_X^{EC}\right)^T$, où le vecteur e – le vecteur de longueur n et de poids $\le t$, ce vecteur possède l'information confidentielle (message d'information visant à être fermé). Le système de transmission d'un message secret de l'abonné A à l'abonné B dans le schéma asymétrique de Niederreiter tout en utilisant des codes elliptiques est représenté sur la fig.4.

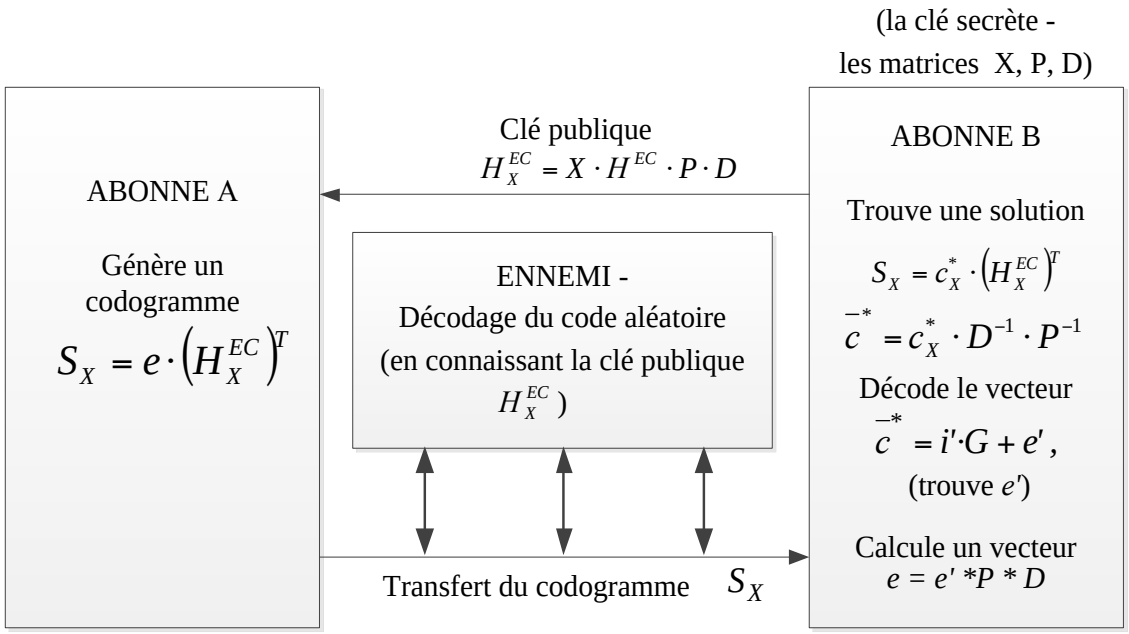

Fig.4. Schéma de transmission du codogramme dans le schéma asymétrique de Niederreiter tout en utilisant des codes elliptiques

3. Conclusions. L'analyse accomplie montre que les moyens cryptographiques asymétriques existants de la protection de l'information n'assurent pas les exigences actuelles. La complexité de la mise en œuvre de transformations cryptographiques dans des systèmes asymétriques est de 3-5 fois plus élevée que dans des systèmes symétriques similaires (algorithmes de chiffrement symétrique par bloc). Mais, puisque le volume de données traitées et transmises augmente de jour en jour, cette complexité reste inadmissible. Le principal inconvénient des moyens cryptographiques asymétriques de la protection de l'information est de manque de cohérence dans la mise en œuvre rapide des transformations cryptographiques de grands volumes de données tout en utilisant des infrastructures à clés publiques. En outre, l'utilisation des moyens cryptographiques symétriques de la protection de l'information implique la présence d'une infrastructure coûteuse de formation, stockage, distribution et utilisation de données à clé secrète ce qui constitue la condition inacceptable pour la plupart des systèmes et réseaux informatiques.

Références :

1. Шнайер Б. Прикладная криптография. –М.: «ТРИУМФ», 2003. – 816 с.

2. R.J. McEliece. A Public-Key Criptosystem Based on Algebraic Theory. // DGN Progres Report 42-44, Jet Propulsi on Lab. Pasadena, CA. January – February, 1978. – P. 114-116.
3. H. Niederreiter. Knapsack-Type Cryptosystems and Algebraic Coding Theory. // Probl. Control and Inform. Theory. – 1986. –V.15. – P. 19-34.
4. Сидельников В.М. Криптография и теория кодирования. Материалы конференции «Московский университет и развитие криптографии в России», МГУ. – 2002. – 22с.
5. Евсеев С.П. Исследование свойств несимметричных и симметричных теоретико-кодовых схем с эллиптическими кодами // Наукові праці НАУ. Серія: Електроніка та системи управління – Київ: НАУ. – 2006 – Вип. 2 (8). – С. 9-16.