Информационные процессы, Том 24, № 2, 2024, стр. 136–139.

Data protection in a network from both channel errors and unauthorized access using one code [1]

V.V. Zyablov[*], V.G. Potapov[*], V.R. Sidorenko[**]

[*] *Institute for Information Transmission Problems, Russian Academy of Sciences, B. Karetniy per. 19, Moscow, Russia*
[**] *Technical University of Munich, Theresienstr. 90, Munich, Germany*

Received July 5, 2024, Accepted August 5, 2024

Abstract—A communication star-network with a central station and users is considered. Transmission over noisy channels from the central station to users and back uses error-correcting coding. Typically, the error-correcting code is open to the world (it is not a secret). This allows an illegal user to gain unauthorized access to the transmitted data. We propose modifying the code and the transmission protocol to protect the network from channel errors and unauthorized access using the same code. The modified network assumes that every user has a secret key, also known to the central station.

KEYWORDS: Data protection, error-correcting coding, unauthorized access.

DOI: 10.53921/18195822_2024_24_2_136

1. INTRODUCTION

Consider a star-network shown in Fig. 1.

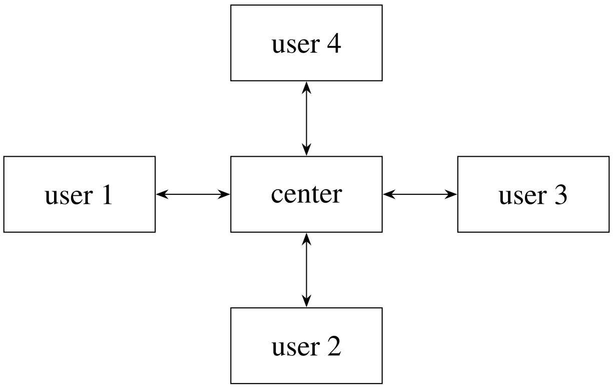

Fig. 1 The original star-network

The connection of each user with the center is shown in Fig. 2. The back connection Center → User is similar.

[1] The work of V. Potapov and V. Zyablov was supported by project no. FFNU-2022-0022.

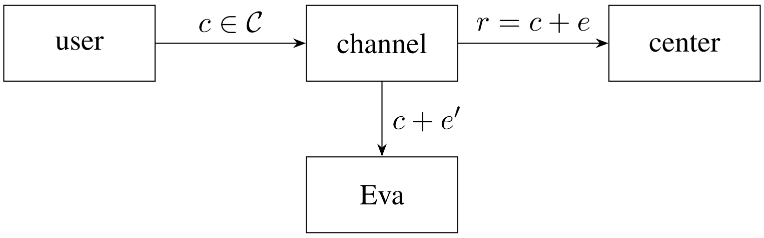

Fig. 2 The transmission User → Center

Here:

- $\mathcal{C}$ is a binary (n, k, d) linear block code with generator matrix G, which is not a secret;
- the channels are binary symmetric channels without memory with crossover probability p;
- a simple minimum distance decoder $\mathcal{D}_t$ for the code $\mathcal{C}$ is available correcting error vectors e of the Hamming weight up to $t = \lfloor (d-1)/2 \rfloor$ that provides a reliable transmission.

*O*ur goal:

- Transform the original network, where an illegal user (eavesdropper Eva) can read and decode any message (since she knows the code used), into a network, where Eva cannot decode the message of any user, since she does not know the code used for the transmission.
- The new coding system should use the decoder $\mathcal{D}_t$ and fully retain the correcting properties of the original network in relation to errors in the channel.

2. TRANSMISSION WITH DATA PROTECTION

The following figure shows transmission from a user to the center. Transmission from the center to a user is the same.

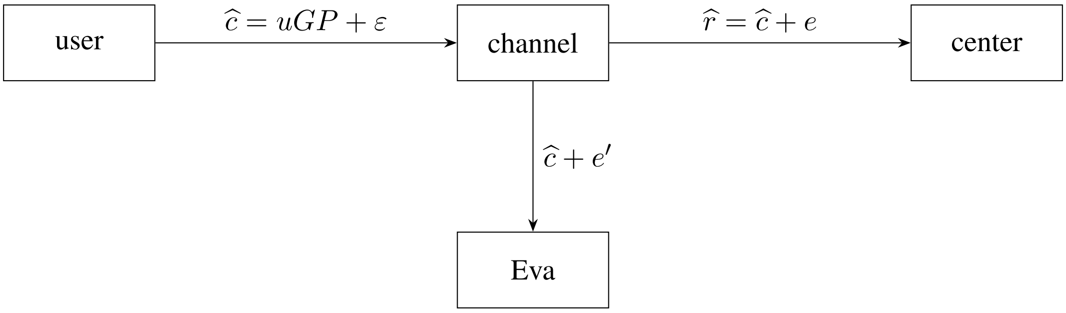
user
$\widehat{c} = uGP + \varepsilon$
channel
$\widehat{r} = \widehat{c} + e$
center
$\widehat{c} + e'$
Eva

Fig.3 Transmission User → Center with data protection

Here:

The transformed code $\widetilde{\mathcal{C}}$ is generated by the generator matrix $\widetilde{G}$ like in [1]

$$\widetilde{G} = GP, \tag{1}$$

where P is an $n \times n$ permutation matrix.

A binary n-vector ε (preferably of high weight) is used to insert errors to the word to be transmitted.

The secret key of the user. We assume that the permutation matrix P and the *secret error vector* ε form a *secret key* (P, ε) known to both the user and the center. In this case the generator matrix G is not included in the secret key and is open. Another variant is to include G in the secret key (G, P, ε).

Encoding at the user. Let u be a length k information vector. To transmit u the encoder computes the codeword of the code $\widetilde{\mathcal{C}}$

$$\widetilde{c} = u\widetilde{G} = uGP \tag{2}$$

and sends

$$\widehat{c} = \widetilde{c} + \varepsilon = uGP + \varepsilon. \tag{3}$$

Decoding at the center. The center receives via the noisy channel

$$\widehat{r} = \widehat{c} + e = uGP + \varepsilon + e, \tag{4}$$

where e is a error vector in the channel. Since ε is known to the decoder, it computes

$$\widetilde{x} = \widehat{r} - \varepsilon = uGP + e \tag{5}$$

and obtains the noisy codeword x of the code $\mathcal{C}$

$$x = \widetilde{x}P^{-1} = uG + eP^{-1}. \tag{6}$$

Clearly, $w(eP^{-1}) = w(e)$. Assume that the weight of the channel error vector is $w(e) \leq t$, then the minimum distance decoder $\mathcal{D}_t$ corrects the errors in x and outputs the correct codeword $c = uG$ and the information vector u.

Straightforward decoding by Eva.

Since Eva does not know the secret pair P, ε she is forced to decode the received vector

$$\widehat{r} = \widetilde{c} + \varepsilon + e', \tag{7}$$

which is a codeword $\widetilde{c}$ of the code $\widetilde{\mathcal{C}}$ corrupted by an error vector $\varepsilon + e'$ of high weight. The code $\widetilde{\mathcal{C}}$ is unknown to Eva. But even if the matrix $\widetilde{G}$ is known to Eva, this is computationally non-feasible to decode $\widetilde{\mathcal{C}}$ using the non-structured generator matrix $\widetilde{G}$.

How to select a secret error vector ε?

A. One way is to use as the secret error vector a constant (for many blocks) error vector ε known to both the user and the center.

B. We can also select secret error vectors ε_i for every block i recurrently, e.g., as follows. The information blocks $\ldots, u_{i-1}, u_i, \ldots$ are transmitted by encoded blocks

$$\ldots, \widehat{c}_{i-1}, \widehat{c}_i, \ldots \;=\; \ldots, \widetilde{c}_{i-1} + \varepsilon_{i-1},\; \widetilde{c}_i + \varepsilon_i, \ldots .$$

We use $\varepsilon_i = \widetilde{c}_{i-1}$. In fact, this means that instead of $\widetilde{c}_i + \varepsilon_i$ we transmit just another codeword $\widetilde{c}_i + \widetilde{c}_{i-1}$ without adding any secret error. In this case, Eva can not obtain $\widetilde{c}_i$ even if she knows the sum $\widetilde{c}_i + \widetilde{c}_{i-1}$.

Secret keys in the network.

- Every user has his own secret key (P, ε).
- The secret keys should be changed periodically or by a command from the center.

3. POSSIBLE ATTACKS.

Assume that Eva also receives the sequence

$$\dots, \widehat{r}_{i-1}, \widehat{r}_i, \dots = \dots, \widetilde{c}_{i-1} + \varepsilon_{i-1} + e_{i-1}, \; \widetilde{c}_i + \varepsilon_i + e_i, \dots .$$

Let us try to find the matrix $\widetilde{G}$ of the permuted code $\widetilde{\mathcal{C}}$.

If we use the variant A to select $\varepsilon = const$, than we can compute the sequence

$$\delta_i = \widehat{r}_i - \widehat{r}_{i-1} = (\widetilde{c}_i - \widetilde{c}_{i-1}) + (e_i - e_{i-1}).$$

Let us form the matrix M with δ_i as rows. If we find k linearly independent rows δ_i without errors, i.e., $(e_i - e_{i-1}) = 0$, then these rows form a generator matrix $\widetilde{G}'$ of the code $\widetilde{\mathcal{C}}$. Probability of this event is very small, and even knowing $\widetilde{G}'$ Eva still can not reconstruct transmitted information.

If we use the variant B to select time varying ε_i then we may construct the matrix M using $\widehat{r}_i$ without errors, $e_i = 0$ and also get a generator matrix $\widetilde{G}''$ of the code $\widetilde{\mathcal{C}}$.

Since a code with its cosets form the complete space $\mathbb{F}_2^n$, we obtain the following statement.

Statement 1. *A fixed vector $\widehat{r}$ can be received when any (n, k) binary code or its coset was used for transmission.*

CONCLUSIONS

- We propose modifying the code and the transmission protocol to protect a network from channel errors and unauthorized access. The modified network assumes that every user has a secret key, also known to the center.
- The proposed method does not change the protection against the channel errors and does not require an additional redundancy.
- The implementation of the method does not require changing encoders and decoders and comes down to only additional multiplication by permutation matrices.

REFERENCES

1. R.J. McEliece A public-key cryptosystem based on algebraic coding theory. DSN Progress Report, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, pp 114–116, (1978).