Investigation of Keyless Cryptosystem Proposed by Dean and Goldsmith

Valery Korzhik, Vladimir Starostin, Kseniia Akhrameeva
The Bonch-Bruevich Saint-Petersburg State University of Telecommunications
Saint-Petersburg, Russia
val-kotzhik@yandex.ru, star_vs_47@mail.ru,oklaba@mail.ru

***Abstract*—We consider keyless cryptosystem proposed recently by Dean and Goldsmith. We show by simulation that this scheme is secure even so the eavesdropper uses suboptimal decoding (zero-forcing attack) with linear complexity but if the number of legitimate and eavesdropper antennas are equal to one another. But if eavesdropper has even small advantage in number of antennas, the proposed cryptosystem occurs insecure that we prove both theoretically and experimentally. We investigate a modified scheme with inverse precoder and show that it requires untractable values of the legitimate user transmitter power. Thus Dean-Goldsmith cryptosystem seems to be insecure with practical point of view.**

I. Introduction

Recently, keyless cryptosystem [1] was introduced by two scientists from Stanford University. The term "*keyless*" means that both encryption and decryption along this cryptosystem do not require any keys at all. The main difference in the knowledge of legitimate users and eavesdropper is only their different locations in space. But of course such cryptosystem (*CS*) can be used not in all possible scenarios, because it requires secure information transmission only in the case when the encrypted messages are transmitted over fading channels with the use of a *massive multiple-input multiple-output* (*MIMO*) technology. We note that such scenario is known as a particular case of "*Physical Layer Security*" (see detail survey in [2]). But nevertheless, if Dean-Goldsmith Cryptosystem (DGC) would be proved as secure cryptosystem then it could be called by revolution in cryptography at least for MIMO-based scenarios.

In Section II we present model of DGC system and show by simulation that this system is secure if the number of eavesdropper's antennas n_r' and legitimate user antennas n_r are equal to one another.

In Section III we investigate DGC system both experimentally and theoretically in the case when $n_r' > n_r$ and show that under such condition this cryptosystem is insecure.

Section IV presents the results of inverse precoding application to DGC system.

Section V concludes the paper.

II. Model of DGC with the use of zero-forcing attack

Let us remind the DGC model [1] for the case $n_r' = n_r = n$. Legitimate channel from Alice (*A*) to Bob (*B*) is:

$$\boldsymbol{z} = \mathbf{A}\boldsymbol{y} + \boldsymbol{e} \tag{1}$$

where $\boldsymbol{z} \in \boldsymbol{R}^n$ is vector received by *B*, $\boldsymbol{y} \in \boldsymbol{R}^n$ is vector transmitted by *A*, $\boldsymbol{e} \in \boldsymbol{R}^n$ is additive noise vector at the receiver *B*, $\mathbf{A} \in \boldsymbol{R}^{n\times n}$ legitimate channel matrix.

It is assumed that $\boldsymbol{e}$ are i.i.d. vectors based on Gaussian distribution $N(0,\sigma_e^2)$, $\mathbf{A} = (a_{ij})$, $i = \overline{1,n}$, $j = \overline{1,n}$ with matrix elements a_{ij} which are i.i.d. ones based on Gaussian distribution $N(0,\sigma^2)$.

Eavesdropper channel from *A* to *E* is:

$$\boldsymbol{z}' = \mathbf{B}\boldsymbol{y} + \boldsymbol{e}' \tag{2}$$

where $\boldsymbol{z}' \in \boldsymbol{R}^n$ is vector received by *E*, $\boldsymbol{e}' \in \boldsymbol{R}^n$ is additive noise vector at the receiver of *E*, $\mathbf{B} \in \boldsymbol{R}^{n\times n}$ eavesdropper channel matrix.

It is assumed that $\boldsymbol{e}'$ are i.i.d. vectors based on Gaussian distribution $N(0,\tilde{\sigma}_e^2)$, $\mathbf{B} = (b_{ij})$, $i = \overline{1,n}$, $j = \overline{1,n}$ with . b_{ij} are i.i.d. based on Gaussian distribution $N(0,\sigma_w^2)$.

All entries of matrices $\mathbf{A}$ and $\mathbf{B}$ are assumed (to be) mutually independent. Next three conditions are very important for further discussion:

- channel matrices $\mathbf{A}$ and $\mathbf{B}$ are constant for the time as user A employs encoder,
- matrix $\mathbf{A}$ is known exactly by legitimate users,
- matrix $\mathbf{A}$ and $\mathbf{B}$ are known exactly by eavesdropper *E*.

(We can note that the model above is more-less valid in practice for fading channels based on *MIMO* technology if space distance between legitimate users and eavesdropper is at least several wavelengths of communication).

Encoding (encryption) procedure in line with [1] is the following:

$$\boldsymbol{y} = \mathbf{V}\boldsymbol{x} \tag{3}$$

where $\mathbf{V} \in \boldsymbol{R}^{n \times n}$ is orthogonal matrix taken from *singular value decomposition* (*SVD*) of matrix $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}$, $\boldsymbol{x} \in \boldsymbol{R}^{n}$, with binary entries x_i, $i = \overline{1,n}$. (We note that in [1] the values are not necessary binary ones but we have taken them as binary values for simplicity and without loss of generality).

Precoding (preencryption) procedure in line with [1] is the following:

$$\begin{aligned} \boldsymbol{z}'' &= \mathbf{U}^{\mathrm{T}}\boldsymbol{z} = \mathbf{U}^{\mathrm{T}}\mathbf{A}\boldsymbol{y} + \boldsymbol{e}'' = \\ &= \mathbf{U}^{\mathrm{T}}\mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}\mathbf{V}\boldsymbol{x} + \mathbf{U}^{\mathrm{T}}\boldsymbol{e} = \mathbf{S}\boldsymbol{x} + \boldsymbol{e}'' \end{aligned} \tag{4}$$

where $\mathbf{U} \in \boldsymbol{R}^{n \times n}$ is orthogonal matrix taken from SVD of matrix $\mathbf{A}$, $\boldsymbol{e}'' = \mathbf{U}^{\mathrm{T}}\boldsymbol{e}$. (We note that all transforms in (4) follow due to the property of matrices $\mathbf{V}$ and $\mathbf{U}$ to be orthogonal). Since $\mathbf{S}$ is diagonal matrix, we get from (4) the following optimal decoding (decryption) rule

$$x_i' = \arg\min_{x_i} \left| z_i'' - x_i \cdot S_i \right|, \; i = \overline{1,n} \tag{5}$$

where S_i is *i*-th element of diagonal matrix $\mathbf{S}$, x_i – are binary entries of vector x. We can see from (5) that decoding procedure has linear complexity on the number of antennas n.

Eavesdropper E following to strategy of legitimate users performs the optimal decoding procedure:

$$\boldsymbol{z}''' = \mathbf{U}'^{\mathrm{T}}\boldsymbol{z}' \tag{6}$$

where

$$\boldsymbol{z}' = \mathbf{B}\mathbf{V}\boldsymbol{x} + \boldsymbol{e}' = \mathbf{U}'\mathbf{S}'\mathbf{V}'^{\mathrm{T}}\mathbf{V}\boldsymbol{x} + \boldsymbol{e}', \tag{7}$$

where $\mathbf{U}', \mathbf{V}', \mathbf{S}'$ are SVD of matrix $\mathbf{B}$.

Substituting (7) into (6), we get

$$\boldsymbol{z}''' = \mathbf{C}\boldsymbol{x} + \tilde{\boldsymbol{e}} \tag{8}$$

where $\mathbf{C} = \mathbf{S}'\mathbf{V}'^{\mathrm{T}}\mathbf{V}$, $\tilde{\boldsymbol{e}} = \mathbf{U}'^{\mathrm{T}}\boldsymbol{e}'$.

Since matrix $\mathbf{C}$ is not a diagonal one in this case, then optimal decoding for an eavesdropper is the following

$$\tilde{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \left\| \boldsymbol{z}''' - \mathbf{C}\boldsymbol{x} \right\| \tag{9}$$

where $\|\cdot\|$ is Euclidean norm in $\boldsymbol{R}^{n}$.

Solution of problem (9) is known as *hard CVP problem* and it was proved in [1] that it has *exponential complexity* with respect to the number of antennas n if the following condition holds

$$M \cdot \sigma_w^2 \cdot \tilde{\sigma}_e^2 > n^{1/2}$$

However in the paper [3] a suboptimal decoding method was proposed that was also a subject of investigation in [4] and called there *zero-forcing attack*.

Assuming that matrix $\mathbf{C}$ in (8) is non singular we get after a multiplication of both sides (8) by $\mathbf{C}^{-1}$:

$$\mathbf{C}^{-1}\boldsymbol{z}''' = \boldsymbol{x} + \mathbf{C}^{-1}\tilde{\boldsymbol{e}}$$

Thus suboptimal decryption (decoding) method can be implemented as follows:

$$\tilde{x}_i = \arg\min_{x_i} \left| \tilde{z}_i - x_i \right|, \; i = \overline{1,n} \tag{10}$$

where $\tilde{z}_i$ is the *i*-th entry of vector $\mathbf{C}^{-1}\boldsymbol{z}'''$.

We can see from (10) that complexity of suboptimal decoding procedure is linear on the number n of antennas. The efficiency of DGC can be estimated by a calculation the symbol error probabilities with optimal decoding of legitimate users by (5) and suboptimal decoding by (10) for eavesdropper. In fact, if for some chosen DGC parameters the first probability is satisfactory, while the second is close to $1/2$, then this cryptosystem can be called secure.

In Table I are presented the results of simulation obtained in [3]. We can see from this Table that for all 5 sets of parameters except of the set 1, DGC is looking as acceptable cryptosystem because if $p' \geq 0.3$ it is impossible to recover a meaningfull text.

TABLE I. RESULTS OF SIMULATION FOR SYMBOL ERROR PROBABILITIES WITH DGC WITH CHOSEN PARAMETERS

№	System Parameters	n	Symbol error probability for legitimate users (p)	Symbol error probabilities for eavesdropper (p')
1	$\sigma^2 = \sigma_w^2 = 2,\ \sigma_e^2 = \tilde{\sigma}_e^2 = 1$	100	0.02	0.2
2	$\sigma^2 = \sigma_w^2 = 4,\ \sigma_e^2 = \tilde{\sigma}_e^2 = 8$	100	0.037	0.3
3	$\sigma^2 = \sigma_w^2 = 1,\ \sigma_e^2 = \tilde{\sigma}_e^2 = 30$	100	0.02	0.42
4	$\sigma^2 = 2, \sigma_w^2 = 1,\ \sigma_e^2 = \tilde{\sigma}_e^2 = 4$	1000	$5.6 \cdot 10^{-3}$	0.3
5	$\sigma^2 = 4, \sigma_w^2 = 8,\ \sigma_e^2 = \tilde{\sigma}_e^2 = 12$	1000	0.01	0.33

In fact, let us consider 32-ary symmetric noisy channel (in line with 32-ary alphabet of Russian language). Then it is easy to show that Shannon capacity of such channel, if every letter is presented by five bits and binary symbols are transmitted over *BSC* with error probability p', can be expressed as follows:

$$\begin{aligned} C = 5 &+ (1-p')^5 \log_2 (1-p')^5 + 5p'(1-p')^4 \log_2 (p'(1-p')^4) + \\ &+ 10p'^2(1-p')^3 \log_2 (1-p')^3 + 10p'^3(1-p')^2 \log_2 (p'^3(1-p')^2) + \\ &+ 5p'^4(1-p') \log_2 (p'^4(1-p')) + p'^5 \log_2 (p'^5) \end{aligned}$$

In Table II are presented the results of capacity C calculations for different values p'.

TABLE II. CAPACITY OF 32-ARY SYMMETRIC CHANNEL AGAINST THE BINARY SYMBOL ERROR PROBABILITIES p'

p'	0.1	0.11	0.15	0.18	0.19
C	2.65	2.5	1.95	1.6	1.5
p'	0.2	0.25	0.3	0.35	0.4
C	1.39	0.94	0.59	0.33	0.15
p'	0.41	0.42	0.43	0.44	0.45
C	0.12	0.093	0.071	0.052	0.036
p'	0.46	0.47	0.48	0.49	0.5
C	0.023	0.013	0.0058	0.0014	0

It is well known that entropy H of Russian language lies within interval $1.5 \div 2.5$ bit/letter. Then in line with Shannon theorem [5] it is impossible to recover meaningfull text after corruption if the inequality $\boldsymbol{H} > \boldsymbol{C}$ holds. In our case this means that a reading (say a decrypting) of meaningfull text is impossible if $p' > 0.19$, that corresponds to all sets of parameters in Table I. We simulated also a corruption of Russian meaningfull text by errors and we have got that only very short words dispersed widely one from another can be readable. In the same time legitimate users can read the decrypted text easily. Moreover, legitimate users can execute *Wyner's wire-tap channel concept* [6] in order to make very reliable decryption over legitimate channel and practically "break of channel" for eavesdropper. In fact, it was shown in [6] that there exists so called *secrecy capacity* equal to

$$C_s = h(p) - h(p')$$

where $h(x) = -(x \log_2 x + (1-x)\log_2(1-x))$ is entropy function, p is the symbol error probability in the legitimate channel, p' is the symbol error probability in the eavesdropper channel.

It has been proved in [6] that there are some encoding and decoding methods providing under the condition $R < C_s$, where R is the transmission rate, that the block error probability after decoding approaches to zero for legitimate users, whereas it approaches to break of channel for eavesdropper channel, if the block length approaches to infinity. In Table III are presented the values of secret capacity C_s for different pairs of bit error probabilities p and p'. We can see from this Table that for real probabilities p and p' which occur more-less in line with results presented in Table II, it is possible to provide DGC security under sufficiently large transmission rate R. Thus it seems to be all fine with DGC at least for the case $n_r' \le n_r$.

TABLE III. SECRECY CAPACITY AGAINST PAIRS OF SYMBOL ERROR PROBABILITIES IN LEGITIMATE (p) AND IN EAVESDROPPER (p') CHANNELS

p	0.0207	0.0224	0.0248	0.0261
p'	0.2119	0.2240	0.2382	0.2569
C_s	0.5997	0.6127	0.6244	0.6476

However it would be interesting to investigate how works DGC under the conditions that the number of eavesdropper antennas n_r' is more than the number n_r of legitimate users antennas? Researches along this line were performed in the paper [4] for asymptotic case when $n_r, n_r' \to \infty$. Results of our investigations are presented in the next Section.

III. INVESTIGATION OF DGC FOR THE CASE $n_r' > n_r$

Initially we investigated the case $n_r' > n_r$ by simulation. Results of such experiment for typical parameters $\sigma^2 = \sigma_w^2 = 4$, $\sigma_e^2 = \tilde{\sigma}_e^2 = 7$, $n_r = 100$ are shown in Table IV. It was used there the suboptimal decision rule (10) where inverse matrix $\mathbf{C}^{-1}$ for rectangular $n_r \times n_{r'}$ matrix $\mathbf{C}$ was calculated as *Moore-Penrose pseudo-inverse of C*.

TABLE IV. RESULTS OF SIMULATION FOR EAVESDROPPER SYMBOL ERROR PROBABILITIES p' AGAINST THE NUMBER OF ITS ANTENNAS n_r', WHEN THE NUMBER OF LEGITIMATE USER ANTENNAS $n_r = 100$

n_r'	100	101	102	103	105	107
p'	0.31	0.22	0.16	0.12	0.07	0.048
n_r'	108	109	110	120	150	
p'	0.039	0.03	0.024	0.003	$7 \cdot 10^{-4}$	

We can see from this Table that even small increases of n_r' by 9 antennas (10% over n_r) results in a drastic degradation of DGC because the symbol error probability p' for eavesdropper becomes very close to the same probability p for legitimate users!

In order to find out that such "paradox" that contradicts to our intuition appears not due to some errors during numerical calculations with simulation, that maybe caused by *bad-posed*, inverse matrices, we consider theoretical proof of the bound for the symbol correct probabilities.

It is easy to see that decision rule (5) for legitimate users when $x_i \in (0,1)$ is equivalent to the following relation

$$x_i' = \begin{cases} 0, & \text{if } z_i'' \le S_i/2 \\ 1, & \text{if } z_i'' > S_i/2 \end{cases}.$$

Then for the symbol correct probability we get the following lower bound

$$P\{x_i' = x_i\} \ge P\{|e_i| \le S_i/2\} \qquad (11)$$

where e_i is the *i*-th entry of additive noise in (1).

Because we assumed before that $e_i \in N(0, \sigma_e^2)$ we get from (11) that

$$P\{x_i' = x_i\} \ge 2\Phi(S_i/2\sigma_e) \qquad (12)$$

where $\Phi(a) = \dfrac{1}{\sqrt{2\pi}} \int_0^a \exp\left(-\dfrac{t^2}{2}\right) dt$.

The decision rule (10) for eavesdropper will be equivalent to the following one:

$$\tilde{x}_i = \begin{cases} 0, & \text{if } \tilde{z}_i \le 1/2 \\ 1, & \text{if } \tilde{z}_i > 1/2 \end{cases} \qquad (13)$$

From relation (13) we get the lower bound for correct symbol probability

$$P\{\tilde{x}_i = x_i\} \geq P\{|e_i''| \leq 1/2\} = 2\Phi\left(\frac{1}{2\sqrt{\mathrm{Var}\{e_i''\}}}\right) \qquad (14)$$

where e_i is the i-th entry of vector $e'' = \mathbf{C}^{-1}\tilde{e}$.

In order to find $\mathrm{Var}(e_i'')$ we accomplish some matrix transforms:

$$\begin{aligned} &\mathbf{C} = \mathbf{BV} = \mathbf{U}'\mathbf{S}'\mathbf{V}'^{\mathrm{T}}\mathbf{V}, \\ &\mathbf{C}^{-1} = \mathbf{V}^{\mathrm{T}}\mathbf{V}'(\mathbf{S}')^{-1}\mathbf{U}'^{\mathrm{T}}, \\ &e'' = \mathbf{C}^{-1}\tilde{e} = \mathbf{V}^{\mathrm{T}}\mathbf{V}'(\mathbf{S}')^{-1}\mathbf{U}'^{\mathrm{T}}\tilde{e} \end{aligned} \qquad (15)$$

Taking into account that $\mathbf{U}'$ is orthogonal matrix and $\mathbf{S}'$ is diagonal one, we get from (15)

$$\mathrm{Var}(e_i'') = \tilde{\sigma}_e^2 \cdot \sum_{k=1}^{n_t} \frac{V_{ik}^2}{S_k'^2} \qquad (16)$$

where V_{ik} are elements of matrix $\mathbf{V}^{\mathrm{T}}\mathbf{V}'$ and S_k' elements of matrix $\mathbf{S}'$.

Substituting (16) into (14) we obtain finally

$$P\{\tilde{x}_i = x_i\} \geq 2\Phi\left(\frac{1}{2\tilde{\sigma}_e\sqrt{\sum_{k=1}^{n_t} \frac{V_{ik}^2}{S_k'^2}}}\right) \qquad (17)$$

(We note that bound (17) is valid also in a general case when $n_r' \geq n_t$).

In order to calculate theoretically the average value of symbol correct probabilities, it would be necessary to average relations (12) and (17) on probability distributions of singular values S_k, $k = 1,2,...,n$ and also on elements of channel matrix $\mathbf{V}^{\mathrm{T}}\mathbf{V}'$. Solutions to this problem requires to execute a very crude approximations.

Therefore we used simulation for calculations by (12) and (17). In Table V are presented such results for average symbol correct probabilities of both legitimate users (q) and eavesdropper (q') with channel parameters: $\sigma^2 = \sigma_w^2 = 7$, $\sigma_e^2 = \tilde{\sigma}_e^2 = 4$, $n_t = n_r = 100$ against different numbers n_r' of eavesdropper antennas.

TABLE V. THE SYMBOL CORRECT PROBABILITIES OBTAINED BY SIMULATION OF AVERAGED BOUNDS (12) AND (17)

n_r'	100	101	102	103	104	105
q	0.95	0.95	0.95	0.95	0.95	0.95
q'	0.71	0.75	0.8	0.87	0.9	0.95
n_r'	106	107	108	109	110	
q	0.95	0.95	0.95	0.95	0.95	
q'	0.96	0.97	0.98	0.985	0.99	

We can see from this table that an increasing of the eavesdropper's antennas even till $n_r' = 105$ results in equality of values q and q', that is in line with our previous claiming. In order to make more clear why small increasing of eavesdropper's antennas numbers results in a degradation of DGC, let us note that singular values S_i of matrix $\mathbf{A}$ do not depend on the number of eavesdropper antennas n_r', whereas the values $1\Big/\sqrt{\sum_{k=1}^{n_t}\frac{V_{ik}^2}{S_k'^2}}$ almost do not depend on "i" but are increasing with n_r' that demonstrate results of simulations presented in Tables VI, VII, VIII, IX.

TABLE VI. RANDOM VALUES S_i, $i = \overline{1,35}$ FOR FIVE DIFFERENT REALIZATIONS OBTAINED BY SIMULATION WITH $n_r' = 100$

i	S_{i1}	S_{i2}	S_{i3}	S_{i4}	S_{i5}
1	53.1421	51.2926	52.7574	51.2378	51.3030
2	49.0113	49.2300	51.3601	49.0261	50.5698
3	47.7710	48.9185	49.0998	48.8155	48.5014
4	46.4833	47.3819	48.5190	47.3624	47.7105
5	45.9201	47.0209	47.2610	46.5177	46.5155
6	44.5610	45.5884	45.7601	45.9859	45.6272
7	44.3215	44.4919	45.2538	45.4024	45.0222
8	43.7468	44.3907	44.9980	44.3537	44.0759
9	43.0688	43.3646	42.8331	43.1716	43.6774
10	42.1501	42.7447	42.5853	42.4437	43.0632
11	41.9077	41.8793	41.7970	41.5423	42.4033
12	41.0089	41.1890	41.5498	40.5631	41.6260
13	40.0538	40.2417	40.9805	40.1696	41.5232
14	39.3225	39.2230	40.4994	39.6923	40.4332
15	38.9847	38.9259	40.2919	39.0572	39.2452
16	38.6270	38.5469	39.4584	38.6847	38.6931
17	37.9362	37.6582	38.8286	37.7011	37.9846
18	37.4425	36.6192	37.8178	37.4408	37.7040
19	37.2768	35.5700	37.2773	36.6693	37.0888
20	36.8424	35.4713	36.6765	35.7515	36.8850
21	35.2984	34.9491	36.4558	35.5875	35.7610
22	34.9989	34.7923	35.7722	35.1888	35.3799
23	34.6652	33.7961	35.1258	34.4821	34.8883
24	34.0018	33.5146	35.0249	34.1817	34.3303
25	33.2964	31.9934	34.3889	33.7225	33.7744
26	32.7586	31.9353	34.1112	33.5202	33.1398
27	32.5421	31.5150	33.2352	32.8792	32.9256
28	32.3027	31.2458	32.7062	31.8925	32.2116
29	31.8145	30.8377	32.2890	31.1338	31.9206
30	30.4265	30.5690	32.0480	30.8268	31.4086
31	30.0797	30.0090	31.3032	30.5448	30.9605
32	29.5904	29.3225	30.5034	29.7901	30.4697
33	29.2897	28.9221	29.6564	28.8425	39.9553
34	28.7577	28.7277	28.9913	28.6074	28.8208
35	28.0603	27.5379	28.3158	28.3992	28.5650

TABLE VII. RANDOM VALUES S_i, $i = \overline{1,35}$ FOR FIVE DIFFERENT REALIZATION WITH $n_r' = 110$

i	S_{i1}	S_{i2}	S_{i3}	S_{i4}	S_{i5}
1	50.9824	53.0538	51.9318	51.0807	51.8145
2	49.9056	50.6061	49.9129	49.7870	50.6450
3	48.9524	49.9480	49.1683	48.5569	49.7767
4	48.1527	49.6644	48.5795	47.4824	46.7860
5	47.4449	47.7428	46.5753	46.6238	46.3256
6	46.8592	46.8847	45.8667	45.2923	45.9752
7	45.8809	45.3978	45.3142	44.6132	44.9388

8	45.1494	44.7303	44.5100	43.6112	44.6423
9	44.2111	44.1368	43.8957	43.0458	43.9833
10	42.9884	43.1129	42.8597	42.4217	43.0394
11	42.2652	41.8919	42.4499	41.0177	42.4000
12	41.6356	41.6720	41.8295	40.6356	41.5470
13	41.3963	40.6965	40.8558	40.3816	41.3071
14	40.8802	40.5208	40.4595	39.3285	40.7017
15	39.6835	39.3599	39.3890	39.2097	40.3211
16	39.2449	39.2334	38.9739	38.4492	39.7391
17	38.6670	37.9198	38.4702	38.2563	39.0567
18	38.2243	37.0359	37.9270	38.1011	38.3523
19	37.3227	36.6517	37.5203	37.1555	37.6947
20	36.7111	36.0370	36.7697	36.5700	37.3387
21	36.5286	35.8650	36.0634	35.5483	36.4972
22	35.3623	34.5166	35.5966	35.0304	36.2223
23	34.9923	34.2696	35.2518	33.9324	35.1049
24	34.2928	33.9754	34.2315	33.4401	34.5635
25	33.3237	33.6524	34.0435	33.0659	34.5321
26	32.8997	33.1967	33.1023	32.7848	34.2210
27	32.7647	32.9741	32.3773	32.1794	33.1142
28	32.3728	32.6702	31.8248	32.0076	33.0427
29	31.6148	31.9982	31.6476	31.7610	32.6673
30	31.2276	31.6184	31.0695	31.0658	31.7917
31	30.7519	30.9305	30.6775	30.8500	31.5113
32	30.4182	30.3407	30.3726	30.1798	30.9599
33	23.9003	29.8950	29.7286	29.8344	30.4207
34	29.2256	29.4883	29.3305	28.9796	30.1205
35	28.8967	29.0893	28.8353	28.4064	29.4850

TABLE VIII. RANDOM VALUES $1\Big/\sqrt{\sum_{k=1}^{n_t}\frac{V_{ik}^2}{S_k'^2}}$, FOR $i=\overline{1,35}$ AND FIVE DIFFERENT SAMPLES WITH $n_r'=100$

i	*j*				
	1	**2**	**3**	**4**	**5**
1	1.9846	0.5222	3.2900	4.6527	2.0207
2	3.4308	1.2298	2.5761	1.6060	2.4357
3	3.3587	1.5144	2.8048	3.7926	2.6829
4	6.6143	1.1309	1.0924	0.6373	1.3335
5	2.8218	0.5050	1.6004	1.6562	2.5735
6	2.8608	2.0991	3.4095	0.5823	2.0196
7	3.6834	3.2847	3.2654	0.5395	3.9628
8	4.4297	0.4656	1.1136	3.3682	5.2492
9	4.4331	0.6591	3.9570	0.7569	4.1201
10	3.1958	0.7685	3.1320	0.8674	3.6958
11	3.4283	2.7693	2.1594	1.1643	1.4124
12	4.0915	0.8224	2.8227	0.5428	3.3620
13	3.1960	2.0685	2.7783	1.0022	2.2754
14	5.6759	1.9221	3.1338	1.4704	2.4892
15	1.8939	1.9314	1.7613	4.6215	2.6918
16	3.0903	1.7260	5.8338	0.4706	1.9392
17	1.8665	0.8096	1.8194	4.1651	1.7015
18	2.3706	0.9222	3.6568	4.5733	5.1911
19	1.9215	2.4882	4.0727	0.6510	5.1157
20	2.8299	0.6427	5.7474	1.3071	2.1649
21	1.8476	1.1147	3.4409	0.3493	5.6532
22	2.7517	0.5247	3.7741	0.6099	3.2229
23	4.4024	0.5002	3.8064	0.5944	6.2287
24	2.7851	1.1352	5.4927	2.7942	1.7500
25	3.3366	0.4097	4.6385	0.3326	4.2765
26	2.6730	0.9086	2.0913	2.8262	2.3091
27	3.8579	0.8057	1.4655	1.1352	3.8464
28	2.3699	0.4921	4.0721	0.6573	1.8491
29	5.2229	0.7782	2.8351	0.6173	2.9256
30	1.1227	0.7285	1.6397	0.9779	4.5044
31	3.3540	0.7621	2.1291	0.9221	3.1184
32	3.0100	1.6583	3.1495	0.5934	1.9128
33	4.3621	3.7209	3.4513	1.4781	3.6105
34	1.9256	1.5865	2.5176	1.2569	3.7876
35	1.3389	0.9054	1.5619	1.1241	1.9794

(Note that we presented only 5 samples because the paper is limited in space. But in fact we have got several thousands of such samples).

TABLE IX. RANDOM VALUES $1\Big/\sqrt{\sum_{k=1}^{n_t}\frac{V_{ik}^2}{S_k'^2}}$, FOR $i=\overline{1,35}$ AND FIVE DIFFERENT REALIZATION WITH $n_r'=110$

i	*j*				
	1	**2**	**3**	**4**	**5**
1	4.0585	11.2842	8.8642	9.9303	10.4852
2	5.2736	10.6344	6.5744	6.7458	8.0473
3	10.3415	7.7069	8.2569	9.8411	12.8172
4	8.0853	12.8003	6.8681	7.1194	7.8092
5	5.1807	15.2354	10.1152	7.3922	7.8905
6	8.5214	9.7867	9.6368	9.6712	8.3813
7	11.8872	10.3566	10.1228	10.4157	7.1919
8	8.4294	9.7628	7.8912	10.0617	9.3328
9	8.0600	13.4725	8.0985	8.8789	8.6212
10	9.8815	8.4520	9.2557	7.4100	5.0882
11	9.1390	5.8559	8.9312	11.4078	7.7790
12	8.8064	11.1386	6.6912	10.6122	10.3782
13	9.2563	10.8738	9.0259	8.9749	8.5289
14	6.1006	8.9830	9.9791	8.2050	9.9487
15	11.2609	5.6981	11.3987	8.0006	7.4600
16	9.7094	7.8390	10.5191	10.0633	6.1619
17	6.6268	6.0898	6.2783	9.3037	6.9804
18	8.0474	7.5894	9.7629	10.7501	9.9889
19	7.2807	7.1295	9.5860	11.4094	8.2144
20	6.4332	8.9631	9.4696	8.6278	9.4293
21	12.6921	4.8959	8.2130	6.8257	8.3650
22	9.6247	7.2096	7.4602	10.1393	7.5571
23	6.8225	8.2650	9.2639	6.7344	6.7694
24	9.6913	11.0944	7.2127	9.1266	9.6910
25	5.6094	6.4283	7.3720	9.2079	9.2125
26	9.7341	8.3402	11.0075	7.7998	6.3047
27	6.8236	9.8864	9.2352	9.4150	7.1576
28	10.8530	6.4126	11.6173	9.9142	9.2040
29	3.8850	9.9808	8.0961	5.0994	7.7569
30	8.1758	7.7924	9.3919	8.1343	6.2250
31	7.4400	9.5434	7.7656	8.0532	8.9836
32	7.1387	9.8464	10.3588	8.3829	8.9897
33	8.4575	13.7233	7.4610	8.2168	10.3041
34	9.6480	8.4067	8.6245	8.0259	5.9852
35	9.3084	10.0877	6.9661	10.2476	4.0590

Thus we can conclude that a compromising of DGC with a small increment of the eavesdropper's antennas numbers (in comparison with legitimate antennas numbers) is the proved fact but not a consequence of bad conditioned matrix property. On the other hand, legitimate users executing DGC do not take for granted that the condition $n_r' \le n$ holds.

IV. INVERSE PRECODING WITH DGC

In the paper [4] has been proposed to change matrix $\mathbf{V}$ in "precoding" procedure to another matrix. In particular, authors of this paper have proved that a choice of matrix $\mathbf{A}^{-1}$ has some advantages, for the thing, a growth of *advantage* for legitimate users proportional to n^2 if $n = n_t = n_r = n_r'$. But it obviously that a choice of matrix $\mathbf{V}$ to be equal to matrix $\mathbf{A}^{-1}$ inverse to channel matrix $\mathbf{A}$ means a simply canceling of channel fading. Then instead of (1) we get

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{e}$$

and instead of relation (2) we get

$$\boldsymbol{z}' = \mathbf{B}\mathbf{A}^{-1}\boldsymbol{x} + \boldsymbol{e}'$$

that can be transformed as follows

$$\left(\mathbf{B}\mathbf{A}^{-1}\right)^{-1}\boldsymbol{z}' = \boldsymbol{x} + \mathbf{B}^{-1}\mathbf{A}\boldsymbol{e}'$$

Our experiments showed that for additive noise variances $\sigma_e^2 = \tilde{\sigma}_e^2 = 0.05$ we get for legitimate users $p \approx 4.5 \cdot 10^{-4}$, whereas for zero forcing eavesdropper rule (10) we get $p' \approx 0.4$.

Moreover, if the number of legitimate antennas is $n_r = 200$, then even for the number of eavesdropper antennas equals to $n_r' = 250$, the probability p' occurs to be more than the probability p.

However in the case of such precoding we face with a growing of a transmitter power. In fact, this power will be equal to $\left\|\mathbf{A}^{-1}\right\|^2$ instead of power $\|x\|^2$ for precoding with matrix $\mathbf{V}$, that for the case of $x_i \in (-1, +1)$ was equal to n_t^2.

In Table X are presented the results of the averaged transmitter power calculated for case of precoding with matrix $\mathbf{A}^{-1}$ and obtained by simulation for $n = n_t = n_r = n_r' = 100$ on different sessions iterations with 10000 realizations for each session.

TABLE X. AVERAGE REQUIRED TRANSMITTER POWER P_0 FOR PRECODING BY MATRIX $\mathbf{A}^{-1}$ AGAINST THE NUMBER OF LEGITIMATE ANTENNAS $n = n_t = n_r = n_r' = 100$ ON DIFFERENT SESSIONS

Sessions	1	2	3	4	5
P_0	$4.5 \cdot 10^5$	$1.8 \cdot 10^7$	$9.1 \cdot 10^5$	$2.1 \cdot 10^6$	$8.8 \cdot 10^4$
Sessions	6	7	8	9	10
P_0	$5.9 \cdot 10^5$	$3.4 \cdot 10^4$	$5.2 \cdot 10^4$	$2.5 \cdot 10^5$	$1.08 \cdot 10^5$

We can see from Table X that the use of "inverse encoding" results in a drastic growing of the required transmission power and large fluctuation on each of sessions that makes such approach impracticable. (We remember that such power was equal to $n^2 = 10^4$ only for an encoding by matrix $\mathbf{V}$!) However, experiment showed that for the most numbers of channel matrices $\mathbf{A}$ (about 9800) the transmitter power is at most 10^5. This result may prompt an idea – to transmit information by controlled *sessions* and do it only if the required transmitter power is less or equal than some chosen threshold. But such approach requires further investigation.

At a first glance seems to be attractively to use for precoding the matrix $\mathbf{V}_0 \neq \mathbf{V}$, $\mathbf{V}_0 \neq \mathbf{A}^{-1}$, but for which the required transmitter power is less than some threshold. Because the relation (3) cannot be used for encoding it is necessary to perform suboptimal rule like "zero forcing". This approach also requires further investigations.

One more problem with a practical application of DGC is a *correct estimation of the matrix* $\mathbf{A}$ *by legitimate users*. They may do it by a sending of special test signal from user *A* to *B* and back from *B* to *A* during *coherent time* of legitimate channel, when matrix $\mathbf{A}$ holds practically constant. But let us assume that this matrix may be slightly changed and estimated by user *A* as $\tilde{A} \neq A$. Then encoding procedure in place of (3) should be

$$\tilde{\boldsymbol{y}} = \tilde{\mathbf{V}}\boldsymbol{x}$$

where $\tilde{\mathbf{V}}$ is taken from SVD of matrix $\tilde{\mathbf{A}} = \tilde{\mathbf{U}}\tilde{\mathbf{S}}\tilde{\mathbf{V}}$. Transmission over channel $A \to B$ be instead of (1)

$$\boldsymbol{z} = \mathbf{A}\tilde{\boldsymbol{y}} + \boldsymbol{e}$$

User *B* share with user *A* the same channel matrix $\tilde{\mathbf{A}}$ and decodes message $\boldsymbol{x}$ as follows:

$$\tilde{x}_i' = \min_{x_i} \left| \tilde{z}_i' - x_i \cdot \tilde{S}_i \right|, \; i = \overline{1,n} \qquad (18)$$

where $\tilde{z}_i$ are entries of vector

$$\tilde{\boldsymbol{z}}' = \tilde{\mathbf{U}}^{\mathrm{T}}\mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}\tilde{\mathbf{V}}\boldsymbol{x} + \tilde{\mathbf{U}}^{\mathrm{T}}\boldsymbol{e}$$

We model incorrectly estimated elements $\tilde{a}_{ij}$ of matrix $\mathbf{A}$ as follows:

$$\tilde{a}_{ij} = a_{ij} + \varepsilon_{ij}$$

where a_{ij} are valid elements of matrix $\mathbf{A}$, ε_{ij} are components of additive Gaussian noise which appear as a result of incorrect matrix estimation. Assuming that ε_{ij} are zero mean, i.i.d. random values with equal variances σ_ε^2, we simulated decoding rule (18) and found the averaged symbol error probabilities $\tilde{p} = \bar{P}\{x_i \neq \tilde{x}_i'\}$. The result of such simulations for typical parameters σ_e^2, σ^2 and σ_ε^2 with $n_t = n_r = 100$ are presented in Table XI.

TABLE XI. THE VALUES OF SYMBOL ERROR PROBABILITIES $\tilde{p}$ FOR LEGITIMATE USERS UNDER THE CONDITION OF INCORRECT CHANNEL MATRIX $\mathbf{A}$ ESTIMATION

σ^2	σ_e^2	σ_ε^2				
		1	0.1	0.01	0.001	0
2	3	0.1677	0.0587	0.0271	0.0208	0.02
3	8	0.1332	0.053	0.0383	0.0378	0.037
50	4	0.0364	0.0126	0.0093	0.0088	0.0084

We can see from this Table that a correctness of matrix element estimation affects very strong on the symbol error probabilities. The probability of error $\tilde{p}$ under incorrect matrix estimation approaches to the error probabilities for ideal estimation (p) if and only if the variance σ_ε^2 of additive noise ε_{ij} is at most 0.001. This value corresponds to averaged

SNR in the legitimate channel equal to $30 \div 40$ dB that is sufficiently high requirement.

IV. Conclusion

Keyless cryptosystem would be very requested by information security technology because it does not use at all any prior key distribution, neither secret or not public. The feature of DGC is that it executes only a fact of *different space locations of legitimate users* and *eavesdropper*. It is looking very attractively because such condition holds in the most of practical situations if carrier frequency f for a communication is high, for example $f = 5.366$ GHz for a common *MIMO*-based technology. This corresponds to wave length $\lambda = 6sm$ and in line with a claiming in [1] a difference in a space locations should be at least 0.4λ that is commonly even for indoor communication.

But there are more main requirements in design of cryptosystems to be practically secure. They can be formulated as following:

- adversary should know algorithm of cryptosystem operation,
- adversary may know exact space locations of legitimate users,
- legitimate users should be authenticated against attack of active adversary,
- the conditions in legitimate channel and MIMO technology should be not better than in eavesdropper channel (in particular, should be equal variances of matrices **A** and **B**, variances of additive noises at legitimate and eavesdropper receivers, the numbers of legitimate and eavesdropper antennas),
- the conditions presented in previous point should be kept even in the case of some technological advances at least in the nearest future.

As far as the first two requirements, they are obeyed for DGC. The third requirement may be provided by the use of additional authentication protocol executing with some *short-authentication key* technique. The forth requirement was used in the current paper investigation. Such, indeed, is the case with parameters $\sigma^2 = \sigma_w^2$, $\sigma_n^2 = \tilde{\sigma}_n^2$, $n_r = n_r'$ for which we have got a reliable decryption for legitimate users and "break of channel" for eavesdropper.

But as far as the last requirement we get an improper situation. In fact, if the qualities of both legitimate and eavesdropper channel would be imported in such a way to be $\sigma^2 = \sigma_w^2 = 100$, then even for large additive noise ($\sigma_e^2 = \tilde{\sigma}_e^2 = 8$) we have got by simulation that that $p = 0.0081$, $p' = 0.0869$. This means that although then the legitimate decoder occurs much better than eavesdropper one it is unacceptable situation because eavesdropper is able to decrypt messages with very small symbol error probability. We note that in the paper [1] was proposed "to add a small amount of Gaussian noise at the transmitter to ensure an adversary will receive the signal below the required SNR to ensure security". But such approach is questionable and requires further investigations.

The most "dramatic" situation is with the number of eavesdropper antennas. We showed in the current paper that if the number of legitimate antennas $n_t = n_r = 100$ but the number of eavesdropper antennas may be incremented even till $105 \div 110$, then DGC is compromised completely. The same simulation holds even if the number of legitimate antennas $n_t = n_r = 1000$. Our experiments showed that in this case an incrementing of eavesdropper antennas on 5 antennas results in a changing the symbol error probability of eavesdropper from $p' = 0.19$ till $p' = 0.05$ that compromises also DGC.

Thus we believe that has no sense to propose an application DGC into practice right now.

Theoretically would be interesting to investigate precoding by inverse matrix including session-based transmission and a consideration of modified DGC algorithm with "addition of small amount of Gaussian noise" mentioned above. But practically it seems to be better to use in frames of "physical-layer security" model a key sharing procedure elaborated in the papers [2,7,8,9] and other. A special interest has, by our opinion, a recently published paper [10] in which authors propose to execute interactive key sharing protocol between legitimate users with specially chosen commutative transforms. We note that the feature of this protocol is similar to idea to use a serial commutative encryption procedure known in cryptography [11] in order to share secret key between legitimate users. The difference is only in that the last transforms are cryptographic ones while in [10] they are based on physical properties of communication fading channel.

References

[1] T.Dean, A.J. Goldsmith, "Physical-layer cryptography through massive MIMO", *Proc. of IEEE Information Theory Workshop*, 2013, pp. 1-5

[2] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey", *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, 2014, pp. 1550-1573.

[3] V.I. Korzhik, V.A. Yakovlev, and S.V. Tikhonov, "Keyless cryptosystem secure on physical level: myth or reality", *Problems of information security. Computer systems,* vol. 4, 2015, pp. 79-89. (in Russian)

[4] R. Steinfeld and A. Sakzad, "On massive MIMO physical layer cryptosystem", *Proc. of IEEE Information Theory Workshop*, 2015, pp. 292-296.

[5] C.E. Shannon, *A mathematical theory of communication*. Bell system tech. J., vol. 27, 1948.

[6] A.D. Wyner, "The wire-tap channel". *Bell system tech. J.*, vol. 54, no. 8, 1975, pp. 1355-1387.

[7] U. Maurer, "Secret key agreement by public discussion from common information", *IEEE Trans Inf. Theory*, vol. 39, 1993, pp. 733-742.

[8] V.A. Yakovlev, V.I. Korzhik and G. Morales-Luna, "Key distribution protocols based on noisy channels in presence of an active adversary: Conventional and new versions with parameter optimization", *IEEE Trans Inf. Theory*, vol. 54, no. 6, 2008, pp. 2535-2549.

[9] V.A. Yakovlev, V.I. Korzhik, P. Mylnikov and G. Morales-Luna, "Outdoor secret key agreement scenario using wireless MIMO fading channels", *International Journal of Computer Science and Applications*, vol. 1, no. 1, 2017, pp. 1-25.

[10] D. Qin and Z. Ding, “Exploiting multi-antenna non-reciprocal channels for shared secret key generation”, *IEEE Transactions on information forensics and security,* vol. 11, no. 12, 2016, hh. 2693-2705.
[11] A. Menezes, P. Oorschot and S. Vanstone, *Handbook of applied cryptography*, CRC Press, 1997.