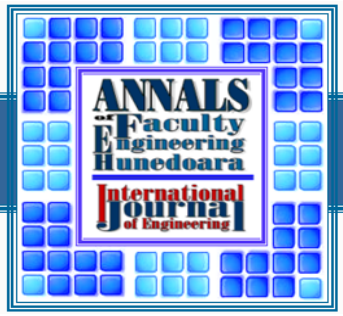
ANNALS OF FACULTY ENGINEERING HUNEDOARA
- INTERNATIONAL JOURNAL OF ENGINEERING
Tome IX (Year 2011). Fascicule 1. (ISSN 1584 - 2665)



[1.]Mária FRANEKOVÁ, [2.]Karol RÁSTOČNÝ, [3.]Aleš JANOTA, [4.]Peter CHRTIANSKY

SAFETY ANALYSIS OF CRYPTOGRAPHY MECHANISMS USED IN GSM FOR RAILWAY

[1,2,3.] UNIVERSITY OF ŽILINA, FACULTY OF ELECTRICAL ENGINEERING, DEPARTMENT OF CONTROL AND INFORMATION SYSTEMS, ŽILINA, SLOVAKIA
[4.] TEMPEST, A. S., GBC IV, GALVANIHO 17/B, 821 04 BRATISLAVA, SLOVAKIA

ABSTRACT: The paper deals with problems related to safety analysis of cryptography mechanisms that are applied in the GSM-R system. Within introduction the authors briefly describe necessary background and position of the GSM-R and Euroradio in the European Train Control System. To ensure required safety level, in the context of OSI Reference Model an additional safety layer must be implemented consisting of two sub-layers: Euroradio Safety Layer and Safety Application Interface. The authors address only the former when paying attention to safety analysis of cryptographic mechanisms applied. To demonstrate and verify some of theoretical findings, an experimental part has been involved to show results of the particular attack to the DES algorithm, in this case an attack based on the birthday paradox realised via UML.
KEYWORDS: safety-related communication, Euroradio, GSM-R, cryptoanalysis, block cipher, CBC-MAC, 3-DES

❖ INTRODUCTION

Any railway infrastructure operator operating in the Central Europe area should endeavour to modernize trans-European corridor lines as fast as possible, with the highest investment priority. Requirement for corridors modernization results from a need to provide the best quality of railway infrastructure respecting both technological and legislative bases according to the latest technologies and European standards.

To fulfil really this requirement means to implement the European Train Control System (ETCS) as a part of the European Rail Traffic Management System (ERTMS) [1] that has been developed since early 90-ties of the 20th century. The main objective of the ERTMS programme is to design a standardized European rail system, common for all EU countries, which will make possible movement of trains equipped with the ETCS wherever within the European railway network. Based on the track-side ERTMS/ETCS equipment the ETCS may be built on one of three basic application levels L1, L2 and L3 [2]. For the ETCS level L2 and higher the technical solution also must inevitably comprise the Global System for Mobile Communications – Railway (GSM-R) which provides radio information transmission between a stationary and mobile part of the ETCS system.

The GSM-R network, as a technological base for open communication system in the railway transport was chosen and specified within the EIRENE and MORANE projects solved under the auspices of the International Union of Railways (UIC). The project EIRENE led to specification of system and functional requirements representing a fundamental interoperability frame for mobile radio communication at lines of the European conventional railway system according to TSI CCS (Technical Specifications for Interoperability Control-Command and Signalling). The final documents [3] and [4] define a set of requirements to the railway radio communication system. The GSM-R system specification results from a technological platform of digital public cell radiophone GSM system extended for specific requirements of railways and properties required from the professional radio system dedicated to railway operation. From the topology point of view, the GSM-R system is a line system, unlike the public GSM system with an area topology. The GSM-R cells are typically overlapped, sometimes up to a half cell; due to ability to serve a mobile station MS reliably at any place. Assumed velocities within the GSM-R network are up to 350 km/h and frequency bands reserved for data transmission are „uplink“ (876 – 880 MHz) and „downlink“ (921 – 925 MHz).

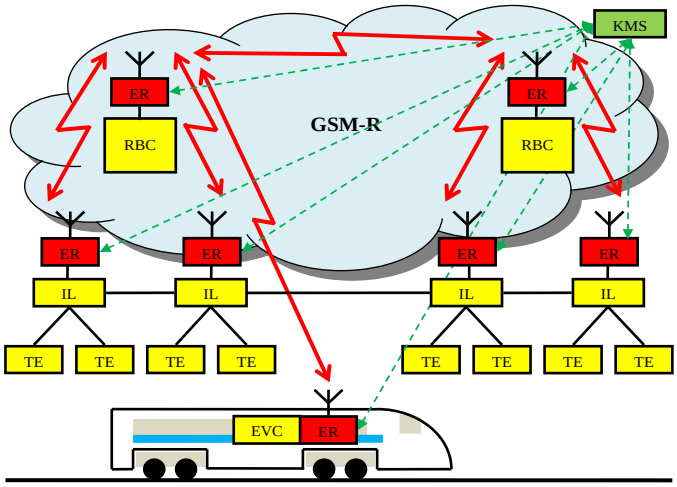


Figure 1. Localisation of GSM-R and ER systems in ETCS architecture

Figure 1 shows usage of the GSM-R system within the ETCS L2. Information needed for railway traffic control (train position, vacancy of track sections, etc.) obtained from Trackside Elements (TE) is concentrated to classical railway interlocking and signaling systems (IL) and from there transmitted via Euroradio (ER) system and the GSM–R communication network. On the base of that information the Radio-Block Centre (RBC) transmits move permissions to individual trains together with other information, again using Euroradio (ER) and the GSM-R communication network. A train sends backward information on its position and other train data to RBC. All functions related to supervision and controls of train velocity are performed by the central European Vital Computer (EVC) located on board the locomotive. Euroradio needs a key management system (KMS) due to management of keys used in cryptographic algorithms.

❖ SAFETY OF GSM-R NETWORK

Since the GSM-R transmission system is classified to open transmission systems (classes 6 and 7 according to [5]) it is necessary to assess risk of unauthorised access and consider all threads listed in the standard [5]. According to [5] messages transmitted by ETCS system correspond to the message model of A1 type utilizing the secure cryptographic code with a secret key. Communication between system components is based on layer principle and meets standardised demands for safety-related communication. Additional safety-related layer, added to standard layers of the Open System Interconnection Reference Model (RM OSI) is formed by two sub-layers:

- ❖ Euroradio Safety Layer (Euroradio SL) [6];
- ❖ Safety Application Interface (SAI) [7].

Within the RM OSI they are integrated above transport layer, having an adaptation layer between them. Figure 2 shows a reference structure of the message for functional specification of interface between trackside subsystems A and B. Euroradio Safety Layer is responsible for secure data transmission which implies protection against threats such as corruption, masking or inserting a message, establishment and release of secure communication link together with error handling. Among secure procedures of the layer there are procedures ensuring message authentication and integrity during transmission. They are realized with the help of the cryptographic technique MAC (Message Authentication Code) which is a function of the message M and the shared key Kc, when applying operation of ciphering C. The formal notation of MAC calculation is:

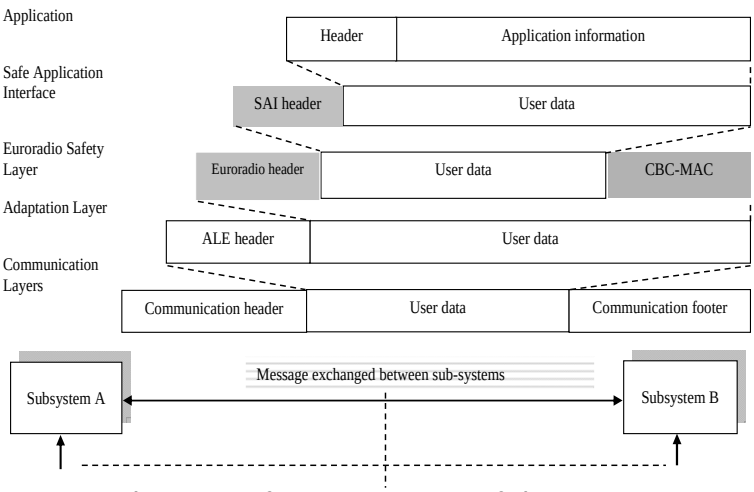


Figure 2. Reference structure of the message

$$MAC = C_{K_c}(M). \quad (1)$$

MAC is calculated both at the side of transmitter which adds it to the message being sent, and the side of receiver which verifies coincidence of received and self-calculated authentication codes.

If the codes are equivalent it may be assumed the message has not been corrupted (message integrity), and the message has been sent by the original sender because no one else shares the secret key. To increase safety of procedure for MAC calculation in Euroradion Safety Layer the chained mode of Cipher Block Chaining MAC (CBC-MAC) is used together with the algorithm Triple-DES in EDE mode (Encryption – Decryption - Encryption), also known as a Triple Data Encryption Algorithm (TDEA) defined in ANSI X9.52 [8]. Another safety procedure of the Euroradio SL is procedure for peer entity authentication, which also uses the algorithms CBC-MAC and Triple-DES. In addition to mutual

authentication of communicating partner entities the procedure also outputs the session key K_S. The paper only contains a detailed analysis of the Euroradio SL.

❖ SAFETY ANALYSIS OF CRYPTOGRAPHY MECHANISMS IN EURORADIO SL

Fundamental requirement of the cryptosystem applied in Euroradio is that implemented cryptographic mechanisms must be able to resist crypto-analytic attacks during the whole life-cycle of the system. To make assessment of safety and effectiveness of applied cryptographic algorithm those crypto-analysis methods may be used that are based on complexity theory. Computationally complexity of the algorithm can be determined on the base of asymptotic complexity describing how behaviour of the algorithm changes in dependence on the size of *n* input data. Operation complexity is usually notated *O* (called Landau notation or Bachmann-Landau notation) and is a function of input data $O(f(n))$ [9]. It is a limit description of the function curve, so called asymptotic upper limit of the magnitude *MG* of the function $f(n)$ expressed by other (usually simpler) function $g(n)$. Computation complexity is usually determined by three parameters: space *S*, time *T*, and data *D*.

Algorithm optimisation is then related to minimisation of one out of these three parameters. Algorithms applied in computer science most often have one of the following complexities (*m* is a real number, $m > 1$):

- Linear complexity: $O(n)$
- Logarithmic – linear complexity: $O(n.log\ n)$
- Polynomial complexity: $O(n^m)$
- Exponential complexity: $O(m^n)$
- Combinatorial complexity: $O(n!)$

The fastest algorithms are considered algorithms with linear, logarithmic-linear or polynomial complexities, algorithms with exponential or combinatorial complexities are realizable in real-time only for low number of inputs *n*.

Safety analysis of cryptographic algorithms used to secure GSM-R communication via Euroradio system must be concentrated on safety assessment of the CBC-MAC algorithm based on Triple-DES, which is applied within the Euroradio SL.

As inputs to safety procedure of CBC-MAC calculation based on Triple-DEC the following entities can be seen: the session key K_S with sub-keys K_1, K_2, K_3, message *M* and cryptographic key K_C shared by transmitter with the source address S_A and receiver with destination address D_A. Addresses S_A and D_A represent entities of the ETCS (e.g. RBC and EVC). Safety procedure of CBC-MAC calculation in the Euroradio SL can be described in the following way:

1. A flag is set to the value log. 0 in the case of communication initiator (transmitter) or the value log. 1 in the case of respondent (receiver).
2. Destination address D_A is added to beginning of the message: D_A_M (the symbol „_“ represents concatenation).
3. A length of the chain D_A_M (denoted as *d*, in the form of two octets added ahead the chain: $d\_D_A\_M$.
4. If a length of the chain $d\_D_A\_M$ in bits is not a multiple of 64 (block size), at the end of the message padding *p* is added: $d\_D_A\_M\_p$.
5. The authentication code MAC is calculated from the chain $d\_D_A\_M\_p$ and the shared key K_C using the chained mode CBC-MAC based on symmetric Triple-DES cipher which can be written in the following way: MAC(*M*) = CBC-MAC(K_C, $d\_D_A\_M\_p$).

Let the session key consists of three parts of the same length (64 bits) $K_C = (K_1, K_2, K_3)$ with the total length 192 bits, including parity bits (every eighth bit). Let data chain (message) consists of bit blocks $X_1, X_2, X_3, \ldots, X_q$. Every block has a size of 64 bits. Encryption function *E* is a ciphering operation of the DES algorithm in the basic mode, which performs ciphering of data chain with use of the key K_i for $i = 1, 2, 3$, and decryption E^{-1} is decryption operation of the DES algorithm. Then individual parts of the key are applied when creating the authentication code MAC in the CBC operation mode for data chain parts X_i using the formula (2). Result of the iteration procedure is H_q = CBC-MAC(*M*).

$$\begin{aligned} H_0 &= 0, \\ H_i &= E(K_1, H_{i-1} \oplus X_i); i = 1,2,...,q-1, \\ H_q &= E(K_3, E^{-1}(K_2, E(K_1, H_{q-1} \oplus X_q))). \end{aligned} \tag{2}$$

From the crypto-analytic view, calculation of the authentication code in the chained CBC-MAC mode of the cryptographic Euroradio protocol implies several essential facts:

- 192 bit key K_S is used sliced to 3 equally long parts K_1, K_2, K_3, while every eighth bit of the key is the parity bit.
- Function $E(K,P)$ represents encryption of input *P* by the block cipher DES with a key *K*. Function $E^{-1}(K,C)$ is decryption of the *C* by the cipher DES with a key *K*. Thus input to procedure is divided to 64 bit blocks $X_1, X_2, \ldots, X_q$.

- Message M is extended using a prefix „$d \mid D_A$“, where d is a length of the chain „$D_A \mid M$“ in octets, and p is padding of the last block X_q to the size 64 bits.
- Hash function CBC-MAC is initialised by zero initialisation vector (IV), formula (2).
- The first q-1 steps of the procedure represent a simple DES cipher in a chained CBC mode.
- The last step q is a cipher Triple-DES in so called EDE mode (Encryption - Decryption - Encryption) only.

It is apparent that using the Triple-DES or DES algorithms in a chained mode (multiple operation mode) increases safety of the DES algorithm but what also should be considered is the fact that in the formula (2) the simple DES is used in q-1 steps.

Quite a lot of crypto-analytic works have been published, describing theoretical attacks to the DES, faster than brute force attack with computing complexity O (2^{56}). There are known several attacks to all 16 rounds of the DES cipher (there are also attacks to the reduced version of the cipher having less than 16 rounds of calculation). There are known attacks based on differential crypto-analysis [10] and linear crypto-analysis [11]. It seems that so called Davies attack [12] is very efficient, based on assumption of knowledge of a certain number of pairs of input and output data that can be used for determining so called empirical distribution of certain characteristics. In this way several bits of the key may be found, the rest of them can be detected using a brute force attack. For 2^{52} known pairs of input and output data it is possible to determine 24 bits of the key with probability of success about 53%.

Usage of the TDEA or DES algorithms in a chained mode (multiple operation mode) increases safety of the DES algorithm. However, as shown in [13], theoretical power of the algorithm is only $O(2^{84})$ at 228 known encrypted data, or $O(2^{112})$ at the use of meet-in-the-middle attack which is under present conditions, permanently growing computing capabilities and in certain applications sufficient only for several coming years. The National Institute of Standard Technology (NIST) confirmed the Triple-DES as a cipher applicable concurrently with the AES for sensitive government data till 2030 year [14]). Meet-in-the-middle attack is a standard technique for crypto-analysis of the TDEA algorithm, in principle it is similar to a birthday paradox.

In an experimental part of the paper there is realization of an attack based on the birthday paradox which reduces the power of the key by the square root of the key size. This attack can also be extended to the DES applied in a multiple operation mode with an initialisation vector. Then for m multiple mode of the cipher the computing complexity $O(2^{m.k/2})$ is assumed at $2^{k/2}$ known encrypted data, if used in combination with the meet-in-the-middle attack.

Supposed that a head (prefix) of the message (input data) A with a larger number of bits than a key is known, every message has it the same (on each occasion encrypted with a different key), it is possible to obtain one of the keys through this attack (actual for a certain modification of the algorithm) as early as for $2^{k/2}$ messages, where k is size of the encryption key. Let l is a number of casually chosen keys K, which will be used to encrypt a head of the message A and result will be inserted to the table as a pair [$E(K, A)$, K], while encryption result $E(K, A)$ will be an index of the table and K its value. Let's choose a number of obtained encrypted messages n, which will be used to extract encrypted head C. On each occasion when a message is received we can have a look into a table whether C is obtained or not. If yes the value for this row of table is the used encryption key. Another possibility is using two tables and analysing an encrypted message without pre-calculation of casually encrypted heads. Probability of finding one of n used keys is expressed by the formula (3).

$$P_S = 1-\left(1-l.2^{-k}\right)^n = 1-\left(1-\frac{1}{2^k/l}\right) \geq 1-e^{-l.n/2^k} \qquad (3)$$

If $l.n \geq 2^k$, probability of finding the key P_S is high (e.g. $l = 2^{k/2}$, $n = 2^{k/2}$). So under given circumstances (known prefix of the message, pre-calculation with casually chosen keys, searching in the table with a constant time) theoretical power of the DES algorithm is only $O(2^{28})$ at 2^{28} known encrypted data. This attack may also be extended to the DES applied in a multiple operation mode with the initialisation vector. Then for the m multiple mode of cipher there is assumed computing complexity $O(2^{m.k/2})$ at $2^{k/2}$ known encrypted data, if used in combination with meet-in-the-middle attacks.

❖ EXPERIMENTAL VERIFICATION

Within the experimental verification a simple software application has been developed making possible to verify success of a birthday attack to the DES cipher with effective size of the key 56 bits. For this purpose the Unified Modeling Language UML 2.0 has been used supported by the Enterprise Architect 6.0 tool. A chosen development tool for Java has been Oracle JDeveloper 10.1.3.3. The application has been primarily designed as a console-based with opportunity for later GUI implementation. Logging is solved through the *log4j* frame. Running scripts with pre-set parameters

have been written in the shell script. Class diagram of the model is shown in Figure 3. Package of cryptographic attacks pch.crypto.attack contains the following trends necessary for the attack realization:

- **ACryptoAttack** – abstract class creating an implementation frame for particular tasks, requires implementation of the method doAttack() from the descendent class returning instance of the class AttackInfo. Further there are methods implemented for writing to so called log (listing directed to file or to console, depending on configuration *log4j*) and to the standard system console.
- **BirthdayAttack** – implementing class of the birthday attck. Besides the overload method doAttack(*) performing one realization of attack set according to the class attributes, it also contains the overload method doAttacks(*) perfoming a chosen number of attacks and returning instance of the class AttackStats. This class is also the target class for running application from the console, so it contains the method main(String[]).
- **HalfSecureRandom** – descendant of the standard class of the language java.security.SecureRandom, covers the method nextBytes(byte[]) to return a mirror symmetrical field of pseudorandom generated bytes.
- **Routines –** a class of accessory statistic methods.

The package of processing (statistic) data from attacks pch.crypto.stat contains classes for saving mentioned data (beans) and their processing:

- **AttackInfo** – a class with attributes containing data about the course and result of the attack.
- **AttackStats** – a class with attributes containing statistic data obtained from several attacks.
- **AttackStatCollector** – a class containing a container for collectivisation of instances AttackInfo and a method calculateStats() for processing of obtained data and creating the instance AttackStats.

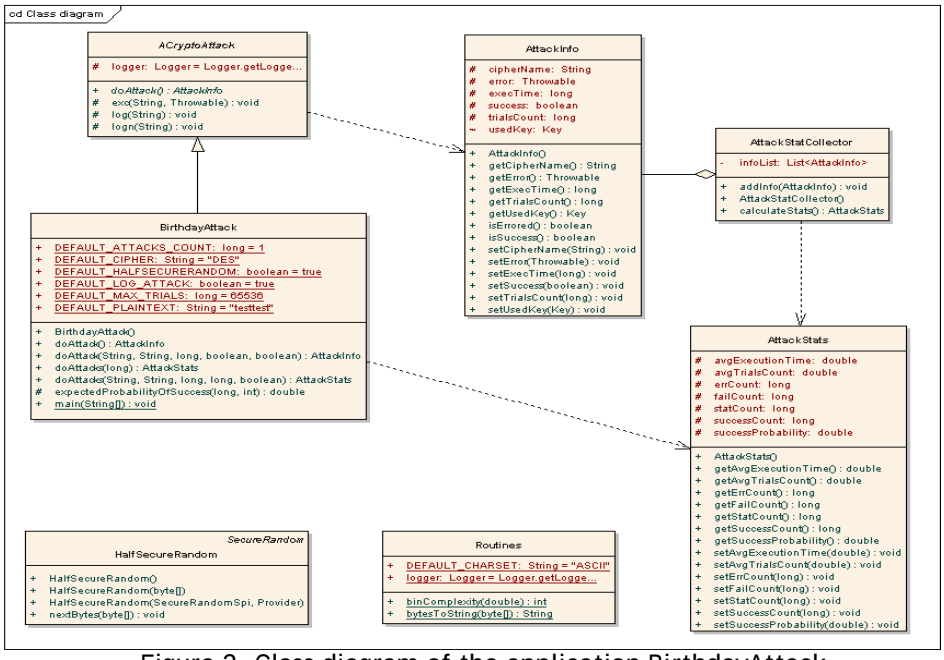


Figure 3. Class diagram of the application BirthdayAttack

The birthday attack according to variant with the use of two indexed tables without pre-calculation would need 2.2^{28} records (memory places) and the same number of encryptions for successful finding of the first key with high probability provided that within the application simulation of obtaining messages with different keys is also considered. Obviously time needed for writing and lookup in tables must also be calculated. Such large tables (either as objects in the operation memory or database) are practically unrealizable on the current hardware (the laptop HP Compaq nx7300 with two-core processor Intel Centrino Duo with clock frequency 2 GHz and operation memory 2 GB, frequency 997 MHz has been available), especially if feasibility is considered from statistic point of view, that is repeated running of the application with the same input parameters. Therefore it was necessary to reduce complexity of the task and thus decreasing effective size of the key. The simplest solution is to decrease effective size of the key to half so that cryptographic keys are generated in mirror symmetry (in bytes – to avoid problem with parity bites), so if n is a length of the key, the byte b_i is identical with the byte b_{n-i-1} for $i = 0, \ldots, n-1$. Then the size of necessary tables (and thus also computing complexity) should be on average $2^{14} = 16384$ records for the successful attack.

Experimental results are summarized in Table 1. Expected probability of success P_S has been calculated according to the formula (3), $k = 28$, $l = n = k / 2$. Probability of successful experiment P_S^* is determined by the empirical formula as a proportion of successful and whole realisations of the experiment.

Table 1. Results of verification of the birthday attack to the modified cipher DES

Characteristics	Maximum number of attempts of the 2nd realisation			
	2^{12}	2^{13}	2^{14}	2^{15}
Number of realizations	1000	1000	1000	1000
Number of successful realizations	68	226	639	984
Number of unsuccessful realizations	932	774	361	16
Effective size of the key k [bit]	28	28	28	28
Probability of success P_S	0,068	0,226	0,639	0,984
Expected P_S	0,061	0,221	0,632	0,982
Average number of attempts	3949,888	7228,148	10811,649	11353,897
Average time of attack realization [s]	0,077	0,147	0,234	0,282

On the base of experimental results we can state that theoretical assumes of the birthday attack have been successfully verified for the block cipher with intentionally decreased cryptographic power. On the other side it is necessary to emphasize that possibility of practical realization of such a type of attack in real cryptosystem is extremely low. However, the birthday paradox has been successfully used and applied in different types of theoretical considerations and as we can seen it is also applicable for construction of crypto-analytic attack.

❖ CONCLUSION

Safety of the Euroradio SL communication, particularly algorithms CBC-MAC and associated power of encryption of the authentication key, may be theoretically lower than expected.

Operation modes have been projected to reduce propagation of bit errors, they better overlay certain characteristics of input data and protect against attacks with chosen input data (chosen plaintext). Obviously higher safety of operation modes against crypto-analytic attacks has been assumed. However, theoretical safety is the same in comparison with simple encryption (usage of the key with the same length). Potential shortcomings of the safety procedure MAC may lead to disclosure of the session key. Then it should be mentioned that cryptographic power of the first q-1 steps of the chained CBC-MAC applied in the Euroradio SL is not higher than it is in the case of a simple DES algorithm usage.

Mentioned shortcomings may be improved in several ways. For example the DES cipher or 3-DES could be substituted with the AES cipher which features higher safety limits, higher flexibility of use and higher software efficiency. What's more modifications of safety mechanisms would be minimal. Size of the block would increase to 128 bits and one 192 bit key would be used. Another alternative of how to increase safety of the authentication procedure MAC based on the Triple-DES is to use a proper triple cascade code, for example CBC|CBC|CBC^{-1} [15].

❖ ACKNOWLEDGEMENTS

This paper was supported by the scientific grant agency VEGA, grant No. VEGA-1/0023/08 "Theoretical apparatus for risk analysis and risk evaluation of transport telematic systems".

❖ REFERENCES

[1.] http://www.ertms.com
[2.] ZAHRADNÍK, J., RÁSTOČNÝ, K.: Safety of Railway Interlocking Systems. (in Slovak), EDIS Zilina, 2006, ISBN 80-8070-546-1.
[3.] UIC EIRENE SRS: System requirements specification. 2006, PSA167D006-15.
[4.] UIC EIRENE FRS: Functional requirements specification. 2006, PSA167D005-7.
[5.] EN 501 59-2: Railway applications: Communication, signalling, and processing systems. Part 2: Safety-related communication in open transmission systems, 2001.
[6.] Subset-037: Euroradio FIS. 2005, v 2.3.0
[7.] FIS SAI: Safe Application Service. 2002, v 8.0, SI/TRK/UP/2.
[8.] ANSI X9.52: Triple Data Encryption Algorithm Modes of Operation.
[9.] http://www.cs.cas.cz/portal/AlgoMath/MathematicalAnalysis/Inequalities/BachmannLandauNotation.htm
[10.] BIHAM, E., SHAMIR, A.: Differential Cryptanalysis of the Data Encryption Standard. Springer-Verlag, 1993, ISBN 3-540-97930-1.
[11.] MATSUI, M.: Linear Cryptanalysis Method for DES Cipher. Advances in Cryptology – EUROCRYPT'93, In: Lecture Notes in Computer Science, Vol. 765, Springer-Verlag, 1993, p. 386-397, ISBN 3-540-57600-2.
[12.] BIHAM, E., BIRYUKOV, A.: An Improvement of Davies' Attack on DES. Advances in Cryptology – EUROCRYPT'94, In: Lecture Notes in Computer Science, Vol. 950, Springer, 1995, p. 461-467, ISBN 3-540-60176-7.
[13.] BIHAM, E.: How to Forge DES-Encrypted Messages in 228 Steps. In: Technical Report, Department of Computer Science, Technion, Haifa, Israel, 1996, CS 884.
[14.] NIST SP 800-67: Recommendation for the Triple Data Encryption Algorithm (TDEA) Blok Cipher. 2004, v 1.0.0.
[15.] BIHAM, E.: Cryptanalysis of Multiple Modes of Operation. In: Technical Report, Department of Computer Science, Technion, Haifa, Israel, 1994, CS 0833.