Generation of random keys for cryptographic systems

Mariusz Borowski\textsuperscript{1}\textsuperscript{*}, Marek Leśniewicz\textsuperscript{1}\textsuperscript{†}, Robert Wicik\textsuperscript{1}\textsuperscript{‡}, Marcin Grzonkowski\textsuperscript{1}\textsuperscript{§}

\textsuperscript{1}Cryptology Division, Military Communication Institute
Warszawska 22A, 05-130 Zegrze, Poland

Abstract – Military and government institutions need security services for storing and exchanging classified information among them. Security of such information is important for independence of the State. At present, cryptography provides a lot of methods for information security. A one-time pad cipher may be used to ensure perfect (unconditional) security. There are many ciphers and other cryptographic transformations, which are not perfect, but ensure conditional security adequate to needs. All cryptosystems require keys and other crypto materials. These keys should satisfy numerous conditions. The basic is randomness. One of the best sources of random bit sequences used in the production of keys for special cryptosystems is a hardware generator. Now we have an electronic device, where it is possible to generate binary random sequences with the potential output rate of 100 Mbit/s. It gives us the capability of building an efficient key generation equipment for the cryptosystems to rely on the one-time pad cipher, where we need very long keys and for those based on symmetric or asymmetric transformations where many relatively short keys are needed.

1 Introduction

Classified information, as defined by the Polish law: “Protection of classified information act”, is divided into four clauses: top secret, secret, confidential and restricted. Diplomacy, military top commands and some special government agencies need perfect (unconditional) security for storing and exchanging top secret information among them. Other government and military institutions classify stored and exchanged information as secret, confidential and restricted, if such information should be protected.
Security of top secret information is not limited in time. Interception of a plaintext by a hostile state or organization can prove destructive in two months as well as in a hundred years, then the need for a perfect cipher is obvious. It is important to recall that the messages encrypted in 1950's with the 'state of the art' not perfect cipher machines kept archived by the adversary (which actually happened) are now generally broken within a few seconds, minutes or some hours at most. On the other hand, the messages sent 60 years ago with any realization of perfect ciphering will stay unbreakable for ever if the keys are destroyed.

The methods of perfect ciphering realization has changed during decades from a pencil-and-paper version to a today's PC computer system equipped with modern software and providing other than confidentiality cryptographic services. It is interesting that all the methods have the same perfect security. Obviously, perfect security is not free. The perfect cipher requires random keys as long as the plaintext, a data management system and a robust, trusted key distribution system.

Conditional security is provided by not perfect ciphers where the security level is estimated and depends on the complexity of the employed cryptographic transformations, time and computational potential of an adversary for cryptanalysis. Asymmetric ciphers with public keys, symmetric ciphers with private keys and other cryptographic transformations use keys, which are much shorter than a plaintext. There should be considered many aspects of planning, managing and distributing key material to the cryptosystems built on such type of ciphers. Development of secret communications causes growing demand for cryptographic keys, especially in the systems which use symmetric ciphers with private keys. One of the basic problems in the key generation systems is an efficient source of random bit sequences.

In the Military Communication Institute (MCI), since the nineties we have been developing hardware random number generators and key generation systems for cryptographic systems based on perfect and no perfect ciphers utilizing one-time, private and public keys. We observe continuous growth of demand for random bits in the key generation systems. We started from random generators with a bit rate 115 kbit/s, then we constructed 1 Mbit/s and 8 Mbit/s random generators to meet requirements. Now we have possibility of hardware generation of binary random sequences with the potential bit rate 100 Mbit/s. It will eliminate the restrictions connected with availability of very long one-time keys for perfect ciphering and many short keys for non perfect ciphering.

2 One-Time Pad

The one-time pad (OTP), also called the Vernam-cipher or the perfect cipher, is a crypto algorithm where a plaintext is combined with a random key. The one-time pad was developed in 1917 by Gilbert Vernam for the use on the telex machines. Each transmitted 5-bit Baudot code was mixed with a random 5-bit code on a paper tape.
The tape with one-time key ran synchronously on both the sender’s and the receiver’s telex. The Vernam’s invention was the basis for several cryptosystems.

We can only talk about OTP, if four important rules are followed. When rules are applied correctly, the one-time pad can be proved unbreakable. However, if only one of these rules is disregarded, the cipher is no longer unbreakable.

1. The key is as long as the plaintext.
2. The key is truly random (not generated by simple computer Rnd functions or whatever!).
3. There should only be two copies of the key: one for the sender and one for the receiver (some exceptions exist for multiple receivers).
4. The keys are used only once, and both the sender and the receiver must destroy their key after the use.

Electro-mechanical OTP cipher machines were manufactured from the fifties to the seventies and are widely used in diplomacy and army on the highest levels of command. A famous example of one-time pad’s security is the Washington-Moscow hotline with the ETCRRM II installed in 1963. Although simple and cheap, it provided absolute security and unbreakable communications between Washington and the Kremlin, without disclosing any crypto technology secret. There were other cipher machines that used the principle of one-time pad, for example the Polish T-352/T-353 DUDEK.

Wide usage of microprocessors, personal computers, magnetic data storage made it possible to replace the electro-mechanical crypto machines in the nineties. The newly designed OTP cipher machines should ensure unconditional information confidentiality by the OTP cipher usage. Moreover, it should provide the additional cryptographic services:

- integrity of messages;
- cryptographic confidentiality of one-time keys;
- integrity of one-time keys;
- secret sharing of keys;
- authentication of correspondent machines;
- authentication of the key generation station;
- authentication of operators;
- an automatic key generation according to the plan of secure connections.

The newly designed OTP cipher machine should also support:

- compression of data to be ciphered;
- electronic accountability;
- electromagnetic emanation protection.

Today the operation of OTP cipher machine needs a lot of long random keys. Many secret connections enhance demands for keys and this causes problems with generation. Now we take up a challenge of extending the key generation system for the OTP machines [1, 2] constructed in our Institute over ten years ago. Hardware generation of binary random sequences with the potential bit rate of 100 Mbit/s eliminates the
restrictions connected with availability of many very long one-time keys for perfect ciphering. Large amount and size of one-time keys also force a change of carriers used for transferring keys from the generation system to the ciphering devices.

3 Hardware random generator with an output rate of 100 Mbit/s

Binary random sequences have numerous applications in many fields of science and security (military) usage. Due to the lack of trusted sources of truly random sequences the Military Communication Institute investigated, implemented and developed a family of hardware random bit generators, first in the nineties. The devices can generate random sequences with the output rate of 115.2 kbit/s (SGCL-1) up to 8 Mbit/s (SGCL-1MB). Both were certified by the Polish national security authority according to the “The protection of classified information act” and can be used in the cryptographic systems up to the top secret level. We implemented these generators in many cryptographic systems designed for generation of keys. In Fig. 1, the overall structure of the electronic device is presented.

![Electronic device of random generation.](image)

In 2012 MCI decided to start the project of 100 Mbit/s hardware generator. The theoretical goal of the project is to develop mathematical and technical methods of generation, giving rise to the physical structure of the generator, implementing the hardware generation of binary random sequences with the potential throughput (amount of data per unit time) 100 Mbit/s, supported by a mathematical proof of their randomness, which guarantees a set of sequences with required probabilistic characteristics and parameters, confirmed by statistical research. The generator (a practical part of the project), will be allowed to be used in the cryptographic systems up to the top secret level. It will also be able to be used in any scientific and technical applications.

As a scientific tool the SGCL-100M generator will be used in advanced research in the field of probability theory, the theory of stochastic signals and information theory. The assumptions of such high bit-rate output of the generator are caused by the fact, that in most modern applications very large samples of random sequences are needed, reaching gigabytes on one calculation or simulation. At the rate 100 Mbit/s a sample of 1 GB size is generated in approximately 90 seconds.
OTP cipher machines use one-time keys as long as a plaintext (and only once) so key accessibility is critical [1]. Possibility for hardware generation of binary random sequences with the potential bit rate 100 Mbit/s eliminates restrictions connected with availability of very long one-time keys for the OTP cipher. The SGCL-100M will be able to generate continuously the one-time keys with a bit rate 100 Mbit/s. The keys can be recorded by a data management system for the OTP cipher machines to mass storage. The generator will be able to produce a little more than 1 TB one-time keys per day and act as a practically “infinite” source of one-time keys.

The prototype of the generator and the necessary documentation will be forwarded to the certification. The generator will have to possess a “certificate of type” up to the top secret level issued by the national security authority. The data management system for the OTP cipher machines is a perfect place to use the SGCL-100M generator.

3.1 Theory of hardware generation of binary random sequences

The Military Communication Institute has already an outline of theory of hardware generation of binary random sequences, which involves generation of many binary (little) imperfectly random component sequences and their post-processing using an XOR sum to the form perfectly random output sequences, then their superposition into one sequence.

![Outline of electronic realization of the SGCL-1MB generator.](image)

The outline of the electronic realization of the SGCL-1MB generator is presented in Fig. 2. $G_n$ denotes imperfectly random generators (the source of noise and comparators). $D_n$ denotes the digital converters (flip-flops) and $C_n$ denotes the tests of randomness. The XOR sum produces an output stream of bits and almost eliminates
a lack of balance between ‘0’ and ‘1’ coming from 8 G generators. MCI has published reviewed monograph [5]. The monograph describes the problem of generation of sequences of 8 Mbit/s rate.

The essence of the SGCL-1MB generator work is based on following principles:

- Source random sequences are passed from exits of Dn flip-flops on Kn inspector controllers testing entropies of all sequences in the real time.

- The entropy of every of these sequences is obviously not isolated and depends on the so-called randomness errors of the sequence on which set the relative bias of „0” and „1” number, \( s = \left| n(0) - 1/2 \right| = \left| n(1) - 1/2 \right| \) (practically in the range from \( s = 10^{-3} \) to \( s = 10^{-2} \)) and correlations between next bits in the sequence, results from sampling of the Poisson signal and expresses a dependence of the correlation coefficient defined by \( K = e^{-2\lambda/fp} \), where \( 2\lambda \) marks the frequency of the pass in the Poisson signal (from „0” to „1” and with the return - in typical avalanche diodes from \( 2\lambda = 35\)MHz to \( 2\lambda = 55\)MHz), meanwhile the \( fp \) sampling frequency of testing equals 8.192 MHz (the correlation errors are in the range from \( K = 10^{-3} \) to \( K = 10^{-2} \)). One can show [5], that entropy of sequence about such parameters carries out \( H = 1 - a(4s^2 + b \cdot K^2) \), where \( a \) and \( b \) are certain constants close the unity.

- The creation of the output sequence follows in the arrangement of the digital processing of source sequences, subjected the parallel operation XOR, which minimizes the randomness errors of the output sequence in the relation \( s_\oplus = 1/2(2s)^M \) and \( K_\oplus = K^M \), and \( H \) imports entropy to the value \( H \cong 1 - a((2s)^{2M} + bK^{2M}) \), where \( M \) is the number of sequences used for this operation [5]. If we accept the above mentioned, even the poorest values of the randomness errors \( (s = 10^{-2}, K = 10^{-2}) \) and \( M = 8 \) source sequences, then the entropy of the output sequence will not be smaller than \( H = 1 - 4.7 \cdot 10^{-28} \), which causes that the sequence with such entropy possesses farther not equals to one entropy. Therefore such fact could be identified just on the basis of the investigations of the test of the sequence about the \( L > 1.4 \cdot 10^{27} \) bits long.

The generation of such long sequence with a bit rate \( BR = 8\)Mbit/s would have to last over \( T = L/BR > 5.6 \cdot 10^{12} \) years, which corresponds to about five hundred times of the Universe existence. The proof is obvious, but one can also try to verify it using the test. On the basis generated so far and the examined tests of the above 1TB sequences which come from about thirty various generators of this type, there is no justification for rejection of the
hypothesis about the randomness of sequences produced by our constructed generator.

- In the case, when any randomness errors of any source sequences grow up unacceptably \((s > 10^{-2} \text{ or } K > 10^{-2})\), reducing the admissible entropy of the given sequence, and as a consequence, of the output sequence, the decision circuit raises the alarm and the generator switches off.

An introduction to the further work will be devoted to the analysis and synthesis of the mathematical basis of the theory of perfect and imperfect binary random sequences and impact of requirements for generated sequences. Much attention area will devoted to the analysis of selecting a source of randomness, conducted on the basis of analytical investigations and results of the author’s experience. Theoretical support of the analysis is the theory of analog and binary stochastic noise signals. As a result of these studies, the conditions for selection of potential sources of randomness will be indicated, leading to a physical source of randomness in the form of avalanche diodes batteries, which generate Poisson signals with controlled randomness. The target theory of generation, however, there will be formulated on the basis of the author’s approach, using the original theory, based on integrated considerations, resulting from the above experiences. Experimental support for the scientific tools will come from the experiments and statistical measurements.

The proof of randomness of generated sequences will be based on the analysis and synthesis of Poisson signals, modelled as stochastic, binary Markov chains. The methodology of the proof will be based on the probabilistic-signal risk analysis of imperfectly random sequences generation \([4]\). In addition to assessing the quality of sequences in the above sense, the security analysis of the generator operation will be made from the viewpoint of electromagnetic compatibility and electromagnetic leakage of information.

Theoretical part of the work also requires to formalize the mathematical description and to show what properties and parameters will have generated sequences. Then, the prototypes of generators will be constructed, which will be used for the practical verification of the theory.

### 3.2 Realization of the SGCL-100M generator

Technical design problems connected with the SGCL-100M generator are encountered on two levels - the electronics and the programming. The electronic board of the generator will consist of 48 generators, which must be calibrated to the generation state consistent with the Poisson signal theory. The stability of the properties and parameters of such a signal as a function of time and climate-mechanical exposures must be tested. The electronic system will also consist of a programmable chip, in which all post-processing operations will be performed, including formatting of the sequence before its sending. Transmission of the sequence from the generator to a computer will take place through a standard 100Base-TX Ethernet. As handling of this interface with full throughput is a very difficult task, the dedicated Ethernet interface controller
Generation of random keys for cryptographic systems will be used and it will be controlled by the RISC microprocessor that will perform the data transfer between the programmable chip and a controller in the Direct Memory Access mode. In practice, only such solution allows to achieve full throughput of 100 Mbit/s. Outline of the electronics is presented in Fig. 3.

![Fig. 3. Outline of 100Mbit/s random generator.](image)

The essence of the 100Mbit/s generator work is based on the following principles:

- Every section is a copy of the technical solutions of the SGCL-1MB generator in the draft layout sense.
- Every section is a 16.384 Mbit/s random sequence source. This value results from duplication of the testing frequency of the Poisson signals. Such testing does not enlarge the randomness errors above established $s < 10^{-2}$ and $K < 10^{-2}$, because we used selected avalanche diodes (the raised frequency of changes crossing $\lambda = 70 \text{ MHz}$).
- The construction contains six sections, clocked by the same 16.384 MHz clock tact synchronically, which allows to get $6 \cdot 16.384 \text{ Mbit/s} = 98.304 \text{ Mbit/s}$.
- A multiplexation can be made on any principle, but the optimum algorithm takes six synchronic bits from all six generator sections in parallel and then formats them in frame boxes. Notice that difference from the systematic algorithm of taking bits will not lead to obtain a full 98.304 Mbit/s output. There are excluded any algorithms of multiplexation controlled by the values of bits from the generated sequences or using the same bits from any sequences many times.
• The standard Ethernet 100Base-TX is the optimum interface to send the 98.304 Mbit/s output stream. The interface enables an efficient 100 Mbit/s random sequences transmission from the generator to a computer.

The generator is a very complex hardware object requiring the software. The software is generally required by two circuits – a programmable chip (a program in AHDL, a VHDL language in the corporate version of Altera) and the RISC microprocessor (programs in C/C++ with "inserts" in the assembler). Both softwares must be optimized due to the efficiency of data transfer to avoid a conflict with the essential functions of random sequence generation. The correctness of theoretical assumptions and the correctness of technical solutions - including software - will be confirmed experimentally by statistical testing of generators in all stages of the development [6].

4 Data management for the OTP crypto machines

Data management systems have been subject to big changes over the time of cryptographic systems development. At the beginning they were simple elements – key generators – producing only keys in open (not encrypted) form. The other operations connected with data processing (i.e. protecting, storing) were carried out by a person. Such kind of the key management system was used by the OTP cipher machines in the seventieth [7].

In the next stages the tasks of system development generators were broadened to the recording results, protection (ciphering), and authentication. Such extended systems are called generation systems. As a result of a rising number of cryptographic devices and development of computer systems, generation systems were equipped with mechanisms of planning secure connections and an element responsible for distribution. Only such systems can be called cryptographic data management systems. These complex management systems has been built since the middle of the nineties. They raised efficiency of data processing and security. The data management systems are intended to deliver correct and reliable key data to proper cryptographic devices. The OTP cipher machines demand a data management system [1]. The system consists of: a secure connecting planning station and a key generation station. The OTP cipher machines can work in two modes: "in a direction way" and “in a circular way” These two modes of operation should be introduced by the secure connecting planning station.

4.1 Planning secure connections

The main aim of a planning secure connection station is to implement only really necessary connections in an OTP cipher machines net. The OTP cipher machine uses one-time keys and time of generating keys is an important factor of a key generating process. “In a direction way” mode needs generation of unique keys for each direction therefore an automatic making connection “each to each” is disabled in the planning station. “In a circular way” mode needs only generation of unique keys for a whole
circular. The information about the OTP cipher machine planned networks includes the number of OTP cipher machines, types of directions, number of one-time keys. Then the information goes to the key generation station. The planning station should be built with the use of a hardened, electromagnetic emanation leakage resistant computer set.

4.2 Generation of one-time keys

The key generation station generates keys on the basis of the information obtained from the planning secure connections station. The keys are generated for all algorithms used by the OTP cipher machine. Of course, the longest time is needed for generating one-time keys. One-time keys are automatically generated, ciphered and signed by the key generation station. Cryptographic keys do not leave the station unprotected: ciphered one-time keys are copied on carriers and symmetric and asymmetric keys needed to fulfill the additional cryptographic services of OTP cipher machines are transferred to temper-resistant smart cryptographic modules.

The quality of generated keys depends on a random number generator. The key generation station uses a hardware random bit generator. Basic characteristics and parameter of the generator:

- good statistical quality of generated binary random streams confirmed by appropriate statistical tests \([6, 8, 9, 10]\);
- user-friendly utilization and maintenance of generated bit streams quality; alarm activation while statistical defects are detected \([3]\).

The random bit generator will have a “certificate of type” issued by the national security authority. The certificate must determine that the generator is suitable for generating data for usage in cryptosystems up to the top secret level.

4.3 One-time pads in today’s world

In the PC computer era, modern algorithms such as symmetric block ciphers and asymmetric public key algorithms replaced one-time pads because of practical considerations and key distribution solutions. Modern crypto algorithms provide practical (not proved) security and privacy, essential to our economy and everyday life. However, top commands of the arm forces and some special military and government institutions need ever lasting absolute security and privacy, and in practice, that is only possible with one-time encryption.

Some experts argue that the distribution of large quantities of one-time pads or keys is impractical. This was indeed the limitation in the era of paper tapes on reels and paper pads. However, today’s electronics, such as the SGCL-100M generator will be able to act as a practically “infinite” source of one-time keys. Current data storage technology such as USB sticks, DVD’s, external hard disks, solid-state drives or dedicated carriers enable the physical transport of enormous quantities of truly random keys. Current sensitive communications are often limited to a small number
of important users. In such cases, one-on-one communications with the associated key distribution, possibly in the configuration with a star topology, is no longer a practical problem, especially considering the security benefits. By using the so-called sneakernet (transferring data on removable media by physical courierring), you can reach a throughput of one-time keys that is greater than what a network can process on encrypted data. In other words, it could take a few hours to drive a terabyte of key material, stored on an external drive, but it will take days or even weeks to consume that amount of keys on a broadband network. A terabyte sized key can easily encrypt e-mail traffic of special (military or diplomacy users) for a year, including attachments.

Therefore, one-time key encryption is still well-suited in specific circumstances where absolute security is preferable to practical considerations, regardless of the cost of secure physical transport of keys by courierring.

In the future quantum key distribution (QKD) may be helpful as the alternative for secure physical transport of keys by courierring. The security of quantum key distribution relies on the foundations of quantum mechanics, in contrast to the traditional key distribution protocol which relies on the computational difficulty of certain mathematical functions. An interesting and promising method of QKD was presented in [11] with the usage of Professor Artur Ekert type of QKD [12]. But at present the ability of the efficient QKD usage is still an open problem.

5 Generation of keys for non perfect cryptosystems

Modern communication systems, which store, process and transmit classified information, consist of several hundred or even several thousand of cryptographic devices, which require huge amounts of cryptographic data. The generation of cryptographic data entails the performance of large amount of time-consuming calculations and does not only relate to the problem of generating cryptographic keys, but also to their appropriate protection against errors, disclosure, labelling and storage.

The currently applied systems and tools for generating cryptographic data are not very efficient for large communication systems, where symmetric keys are used. For every secure information relation, appropriate cryptographic data should be assumed, e.g. if there is \( n = 100 \) devices, at least \( \frac{1}{2} \cdot n \cdot (n-1) \), i.e. nearly five thousand cryptographic data for the “each to each” information relation model should be prepared. The planning, generation and distribution of cryptographic data for such a large network is a technically complicated system.

The cryptographic key generation subsystem for special networks consists of one or several combined computer stations. These stations perform various functions within a system.

A station for special network planning and cryptographic data distribution is to implement necessary connections in a secure network. Proper functioning of a secret data information system requires designing of a network made up of encryption devices and software as well as providing cryptographic data to every device and user.
Generation of random keys for cryptographic systems

This operation is carried out regularly at certain time intervals (every few/several months). When planning, the need to immediately generate data in particular emergency situations should be taken into account. Once generated, the cryptographic data should be combined into sets and distributed to loading stands or directly to the devices. The data ought to be delivered in a safe manner, so as to preclude its disclosure and unauthorized modification.

A cryptographic key generation station serves the cryptographic data generation for every cryptographic device operating within a secure communication network. The data is secured within the distribution period. The cryptographic key generation station is most often built based on a personal computer with the attached external devices such as the hardware random sequence generator, order station and data preparation for distribution in the system. The cryptographic data generation station should generate data necessary for the operation of various cryptographic algorithms such as stream and block ciphers, message signing and different passwords for cryptographic devices and systems.

Development of secret communications causes growing demand for cryptographic keys, especially in the systems which use symmetric ciphers with private keys. One of the basic problems in the key generation systems is an efficient source of random bit sequences. Currently used random generators with an output bit rate from 115 kbit/s to 8 Mbit/s in the key generation systems will be not sufficient in the future. Now we are capable of the hardware generating binary random sequences with the potential 100 Mbit/s bit rate. It will eliminate the restriction connected with availability of a large number of keys for non-perfect ciphering.

6 Conclusions

All cryptographic systems require keys and other crypto materials. These cryptographic keys should satisfy many conditions and randomness is the most obvious. The SGCL-100M generator shown in Fig. 3 and described in Section 3 of the article will be able to produce a little more than $2^{40}$ random bytes per day and act as a practically “infinite” source of random material for producing any keys. The generator will have a “certificate of type” issued by the national security authority. The certificate must determine if the generator is suitable for generating data for usage in cryptosystems up to the top secret level. A data management system for classified communication systems is a perfect place to use the SGCL-100M generator. Capability of one-time keys generation or generating random keys for non-perfect cryptosystems will be no limitation any longer.

As a scientific tool, the SGCL-100M generator can be used in advanced research in many fields of science and technology. The most important are cryptography, theory of stochastic signals, information theory, statistics, numerical computation, stochastic simulations using the Monte Carlo method, and many others. Since the generator is a
quite complex and costly device with a very high output rate, it can be assumed that it could be used as a source for random sequence servers in the R&D centers.

References


