Understanding Periods in Edon80
- Response on *Remarks on the Period of Edon80, by Jin Hong* -

D. Gligoroski[*] [**], S. Markovski[**], L. Kocarev[***] and M. Gušev[**]

[*] Centre for Quantifiable Quality of Service in Communication Systems, Norwegian University of Science and Technology, O.S.Bragstads plass 2E, N-7491 Trondheim, NORWAY, e-mail: gligoroski@yahoo.com [1]

[**] "Ss Cyril and Methodius" University , Faculty of Natural Sciences and Mathematics, Institute of Informatics, P.O.Box 162, 1000 Skopje, Republic of Macedonia, e-mail: {smile,marjan}@ii.edu.mk

[***] University of California, San Diego, Institute for Nonlinear Science CMRR Build., 9500 Gilman Drive, La Jolla, CA 92093-0402, USA, e-mail: lkocarev@ucsd.edu

1 Introduction

In this note we responde to the remarks of Jin Hong given in [1] regarding our stream cipher Edon80 [2]. Our mathematical understanding of the periods of streams that are produced by Edon80 is now much deeper then in the time of the design of the cipher. According to that mathematical knowledge, we discuss two scenarios how to deal with the further treatment concerning Edon80 submission to ECRYPT.

The attack presented in [1] is based on analyzing (*key, state*) pairs (concrete assignment of working quasigroups $*_i$ and initial values for a_i) that give small periods. Hong experimentally counts all possible (*key, state*) pairs for periods 4, 8 and 16, with the number d of rows from 5 to 18 and summarizes the results in Table 1 of [1]. These numbers are then extrapolated for the

[1]This work was carried out during the tenure of an ERCIM fellowship of D. Gligoroski visiting Q2S - Centre for Quantifiable Quality of Service in Communication Systems at Norwegian University of Science and Technology - Trondheim, Norway.

value $d = 40$ in Table 2 of [1]. Then, by repeating the sequence of obtained working quasigroups $*_i$ from the first 40 e-transformations to the last 40 e-transformations (as it is done in Edon80) and by giving a freedom of 80 bits for choosing the leaders for those transformations $a_i, i = 39, 40, \ldots, 79$ he computes that the probabilities of obtaining periods with the lengths in the range $2^{53} - 2^{55}$ are in the range from $2^{-88} - 2^{-71}$. By reducing the initial extrapolation not to the 40-th row, but to the 34-th row, he computes the probabilities of obtaining periods in the range $2^{61} - 2^{63}$ with even much higher values in the range $2^{-75} - 2^{-60}$. He even discuss the possibility of the existence of a (*key, state*) pair that will give a very short period of only 2^{20} but does not give the projection of the probability for obtaining that period. The summary of his findings are given in Table 1. He ends his note by the conclusion:

> "Even though these results do not break Edon80 completely and does not give us any information about how to recover keys, it does show that the period of Edon80 is far from being well understood."

Stream length less then	Probability
2^{20}	?
2^{53}	2^{-88}
2^{54}	2^{-78}
2^{55}	2^{-71}
2^{61}	2^{-75}
2^{62}	2^{-66}
2^{63}	2^{-60}

Table 1: Summary table from [1] with the projections of the probabilities for obtaining streams with the lengths shorter then indicated in the first column.

2 Building probabilistic model for periods produced by Edon80

In this section we briefly describe our findings about periods produced by Edon80. Much brother and detailed explanation about the mathematical

i	X_i				i	X_i			
1	1	2	3	4	9	1	2	3	4
	0.1271	0.5000	0.3729	0.0000		0.2505	0.2510	0.3416	0.1569
2	1	2	3	4	10	1	2	3	4
	0.1485	0.1875	0.3522	0.3118		0.2503	0.2536	0.3397	0.1564
3	1	2	3	4	11	1	2	3	4
	0.2369	0.3355	0.2539	0.1738		0.2502	0.2510	0.3407	0.1581
4	1	2	3	4	12	1	2	3	4
	0.2536	0.2661	0.3115	0.1688		0.2516	0.2461	0.3445	0.1577
5	1	2	3	4	13	1	2	3	4
	0.2457	0.2512	0.3448	0.1584		0.2479	0.2524	0.3429	0.1568
6	1	2	3	4	14	1	2	3	4
	0.2498	0.2484	0.3457	0.1561		0.2500	0.2502	0.3421	0.1577
7	1	2	3	4	15	1	2	3	4
	0.2474	0.2518	0.3432	0.1576		0.2538	0.2515	0.3378	0.1569
8	1	2	3	4	16	1	2	3	4
	0.2488	0.2493	0.3451	0.1568		0.2491	0.2499	0.3448	0.1561

Table 2: The distribution of the random variables X_i for the first 16 values of i.

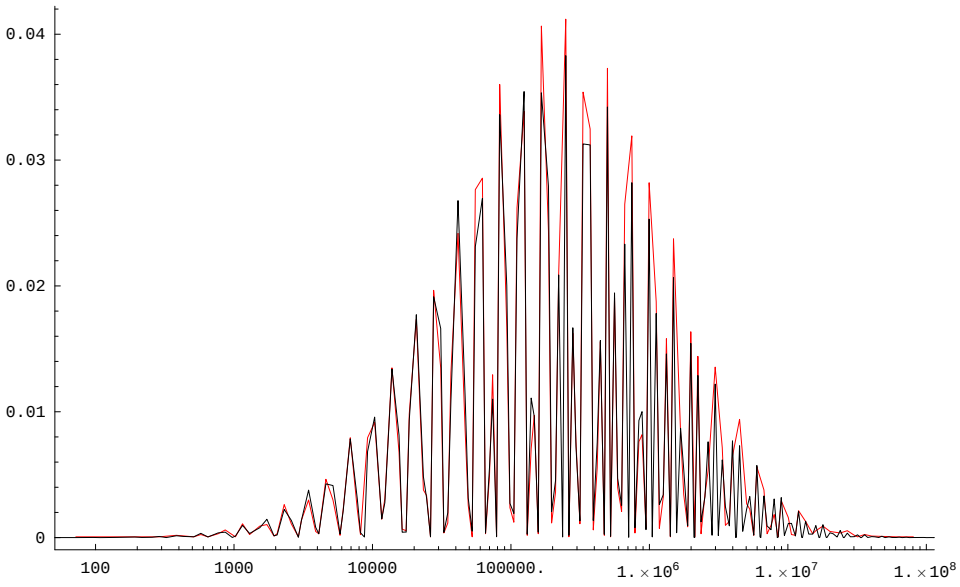

Figure 1: Comparison between our mathematical model and concrete experimental results for the periods of Edon14. The black line represents values from the model and red line represents obtained results after exhaustive search for all 2^{14} keys for Edon14.

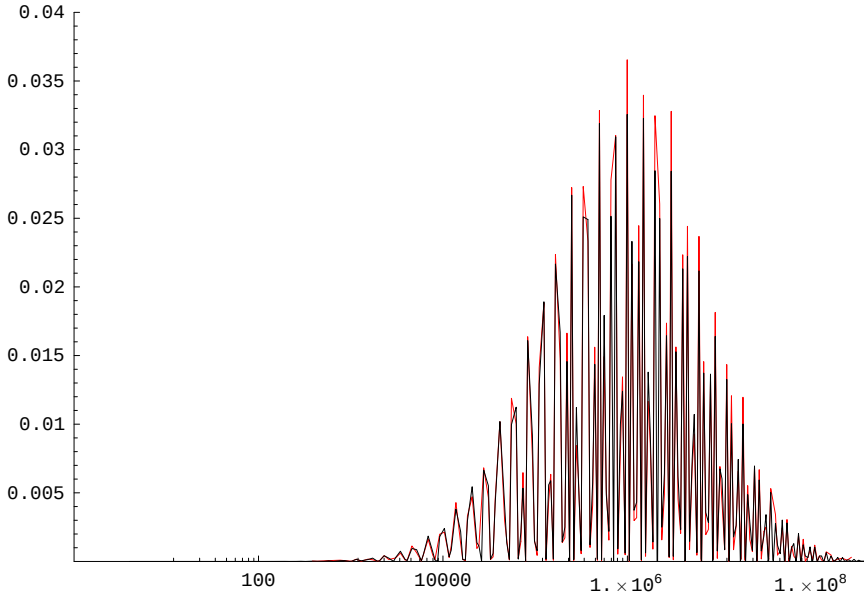

Figure 2: Comparison between our mathematical model and concrete experimental results for the periods of Edon16. The black line represents values from the model and red line represents obtained results after exhaustive search for all 2^{16} keys for Edon16.

probabilistic model that explains the distribution of the periods obtained by quasigroup string transformations like those used in Edon80 will be given in a forthcoming paper.

For this note we can say that every application of a quasigroup e-transformation to a string of period P, by quasigroup of order 4 produces a string P' with the length which is 1, 2, 3 or 4 times longer then the original string. Thus, every application of an e-transformation in the cipher like Edon80 can be seen as a random variable X that receives values from the set $\{1, 2, 3, 4\}$. Since Edon80 has 80 e-transformations, we have 80 random variables $X_1, X_2, \ldots, X_{80}$ (that can be treated as statistically independent under the assumption that one-way *IVSetup* procedure is well defined and maps the initial 144 bits of the *Key* and *IV* without bias into 160 bits). Furthermore, the period produced by the whole cipher can be seen as a random variable Y_{80} that is a product of 80 independent random variables X_i, i.e. $Y_{80} = \frac{1}{2} P X_1 X_2 \cdots X_{80}$. Here, $P = 4$ is the constant initial period of the initial sequence 012301230 . . . that is feeding the Edon80 pipeline of e-transformers, and 1/2 means that every second produced element is taken. The most im-

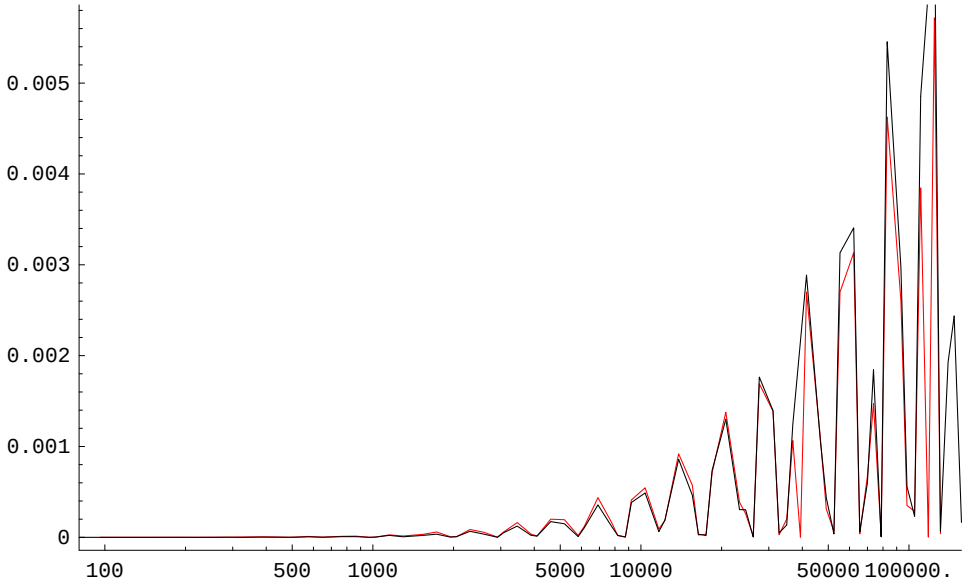

Figure 3: Comparison between our mathematical model and concrete experimental results for the periods of Edon18. The black line represents values from the model and red line represents obtained results after exhaustive search for all 2^{18} keys and all 2^2 values for IV, and counting only periods with length less then 2^{17}, for Edon18.

portant question is to find the distribution of the variables $X_i,\ i = 1, \ldots, 80$.

After many experiments of performing e-transformations on the strings obtained as in Edon80 cipher, we have found the numerical values for the distributions of the random variables $X_i, i = 1, 2, \ldots, 16$ and they are shown in Table 2.

We have taken, in the modelling of the periods of Edon80, that the rest of the random variables $X_{17}, X_{18}, ...$ have the same distribution as the last numerically obtained distribution X_{16}.

$$X_i = X_{16} : \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0.2491 & 0.2499 & 0.3448 & 0.1561 \end{pmatrix},\ i = 17, 18, \ldots .$$

To compute the distribution of the random variable Y_{80}, that describes the periods of Edon80, we have used the recurrent relation $Y_i = Y_{i-1}X_i,\ i = 2, \ldots, 80$ and $Y_1 = \frac{1}{2}PX_1$. There is a well developed theory about distributions that are result of multiplication of continuous random variables, and the resulting distribution is the so called LogNormal distribution. In our case

we have discrete variables and still, as the number of variables is increasing, the distribution is approaching the LogNormal distribution. See, for example, Figures 1, 2 and 4. Obtained discrete distributions are very close to a continuous LogNormal distribution.

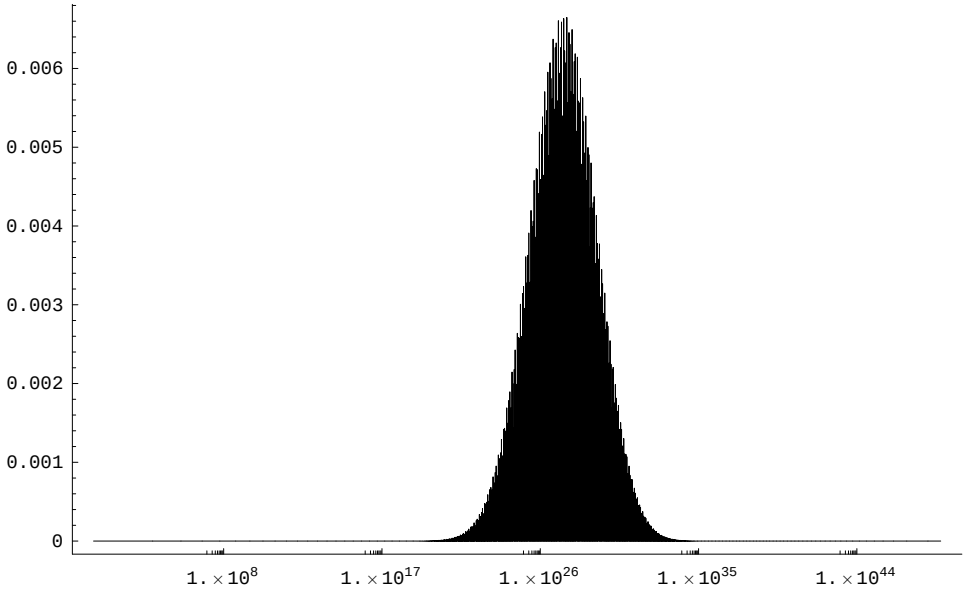

Figure 4: LogNormal distribution for periods of Edon80.

On figures 1, 2 and 3 we show by black line the distribution of the periods obtained by the probabilistic model for Y_{14}, Y_{16} and Y_{18} and by red line we show the concrete distribution obtained by exhaustive search in the whole space of the corresponding cipher. For Edon14 and Edon16 the periods with the length less then 2^{28} were computed. For Edon18 we just counted all (*key, state*) pairs that give streams with periods less then 2^{17}. By comparing the theoretical probabilistic model that we have developed and concrete numerical results, we are pretty confident that the model can be used for describing the distribution of the periods for Y_{80}, i.e. for Edon80. From technical aspect (if someone wants to repeat the experiments) for Edon14 we have used 14 bits that will represent the Key, the padding constant was: 3210012 (same as in Edon80), and there ware no bits for the IV. For Edon16 we have used 16 bits for the Key, the padding constant was: 32100123 (same as in Edon80), and there ware no bits for the IV. For Edon18 we have used 18 bits for the Key, the padding constant was: 32100123 (same as in Edon80), and IV had only 2 bits.

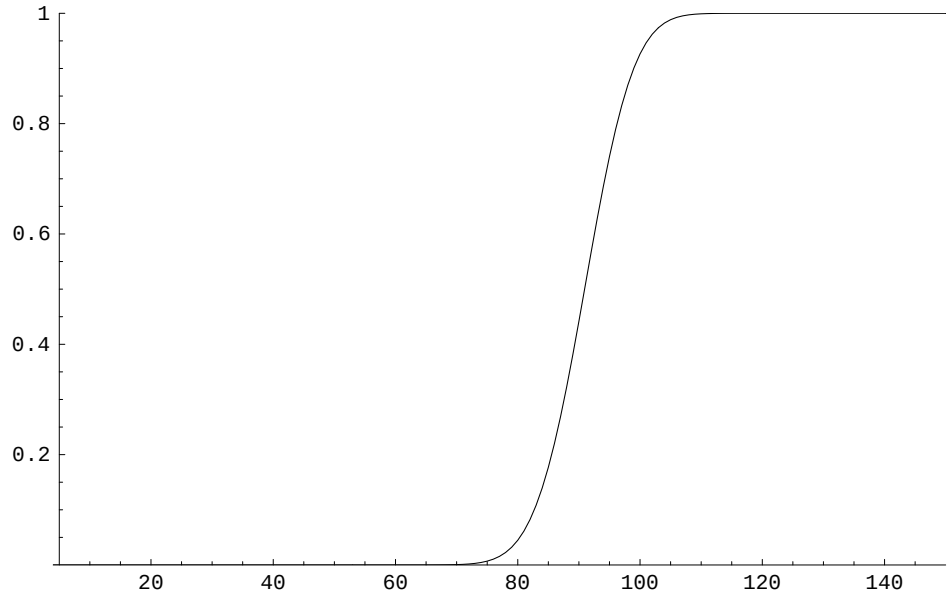

Figure 5: Log plot for the cumulative distribution function for periods of Edon80. The numbers on the x-axes have to be interpreted as powers of 2.

In the next Figure 4 we show the Log plot for the probability distribution of Y_{80}, in Figure 5 we have plotted the cumulative distribution function, and in Table 3 we give some numerical values from the cumulative distribution function of Y_{80}. Comparing these results with the values that Hong has obtained in his analysis we conclude that, although he was right to suspect that there are (*key, state*) pairs that will give periods with smaller length then 2^{80}, the probabilities for obtaining such (*key, state*) pairs are significantly higher then his projections. That is due to the fact that he missed to count a lot of periods that are between two periods of length which are power of 2.

In the following Table 4 we give some values from the cumulative distribution of the random variable Y_{160}, for some periods less then 2^{80}. The reasons why we give this table will be explained in Scenario 2 in the next section.

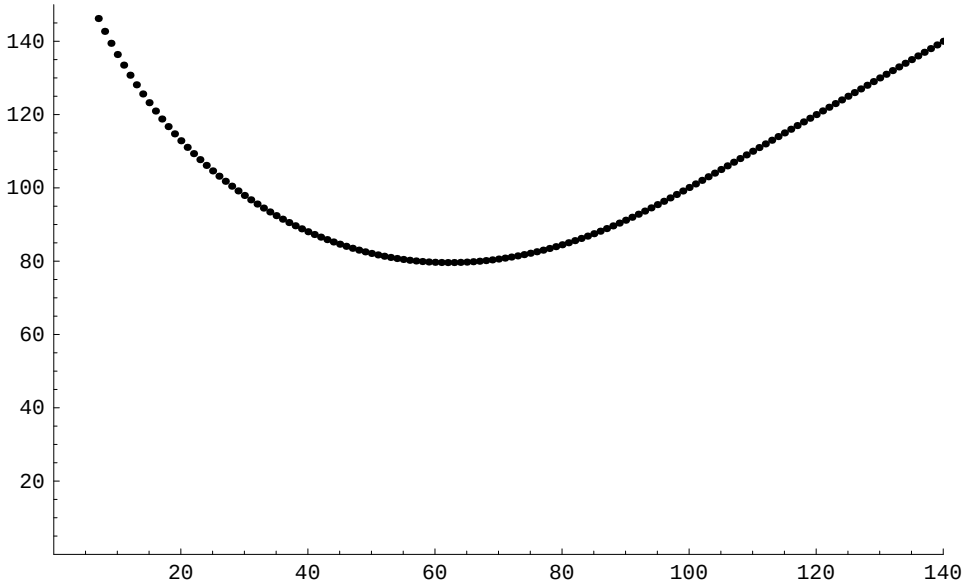

Figure 6: Trade–off log–log plot between the search for short periods vs total amount of collected data in bits. The numbers on the x-axes and y-axes have to be interpreted as powers of 2. The minimum is obtained for the periods of 2^{61} and the value is $2^{79.62}$.

3 Two scenarios how to deal with further status of Edon80 as a submission to ECRYPT

According to the above analysis we propose two scenarios how to treat Edon80 submission to ECRYPT. The first one does not eliminate the short periods in the initial version of Edon80, and the second one does. Here are those scenarios:

1. By this note we update the information in the documentation submitted to ECRYPT, about the periods of the streams produced by Edon80. The reader is urged to look at Table 3. This table updates the information in the original submission that the average period of the stream is 2^{103}. As a matter of fact, it can be seen form Figures 5 and 4, as well as from numerical data for the distribution of periods of Edon80, that the average period of Edon80 is 2^{91}. We have to note additionally that the results about the existence of the shorter periods, do not decrease the total workload for breaking Edon80. For example, if the attacker wants to exploit the fact that the probability for obtaining periods less then

Stream length less then	Probability
2^{20}	$2^{-92.86}$
2^{24}	$2^{-82.14}$
2^{28}	$2^{-72.46}$
2^{32}	$2^{-63.59}$
2^{36}	$2^{-55.46}$
2^{40}	$2^{-48.03}$
2^{44}	$2^{-41.24}$
2^{48}	$2^{-35.02}$
2^{52}	$2^{-29.35}$
2^{56}	$2^{-24.25}$
2^{60}	$2^{-19.70}$
2^{64}	$2^{-15.67}$
2^{68}	$2^{-12.16}$
2^{72}	$2^{-9.13}$
2^{76}	$2^{-6.57}$
2^{80}	$2^{-4.48}$

Table 3: Some probabilities for obtaining streams with periods less then indicated in column 1 for Edon80.

2^{40} is around $2^{-48.03}$, then he would have to have at least 2^{47} sequences of length 2^{40} (or to collect in total information in the range of 2^{87} bits) to have the probability more then 0.5 that some of those sequences are shorter then 2^{40}. This type of trade–off attacks are most efficient for periods with length around 2^{61}, that appears with probability $2^{-18.62}$ (see the Figure 6). The total amount of collected bits in such a case would be $2^{79.62}$ bits. However, even in a case of having such a situation, it is still unclear how the attacker will obtain some information for the key or the internal state of the cipher. So, in total, the workload for performing such an attack (that will not recover the key, in fact) would exceed 2^{80}. Hence, the simple exhaustive search in the key space (2^{80}) is still best kind of attack on Edon80.

2. The second scenario is to reduce the probability of obtaining streams with periods less then 2^{80}. That can be done by simple incensement of the internal pipeline of Edon80 from 80 to 160 e-transformers. The speed of the whole cipher will be not affected at all. If non-shared RAM

Stream length less then	Probability
2^{40}	$2^{-175.01}$
2^{44}	$2^{-164.37}$
2^{48}	$2^{-154.24}$
2^{52}	$2^{-144.54}$
2^{56}	$2^{-135.25}$
2^{60}	$2^{-126.38}$
2^{64}	$2^{-117.92}$
2^{68}	$2^{-109.87}$
2^{72}	$2^{-102.19}$
2^{76}	$2^{-94.86}$
2^{80}	$2^{-87.85}$

Table 4: Some probabilities for obtaining streams with periods less then indicated in column 1 for Edon80 v2.0 (Edon80 with pipeline of 160 internal e-transformers).

is not used then the hardware requirements will be doubled, otherwise the hardware requirements will increase only slightly. The additional interventions in the submitted source code to ECRYPT would be also minimal. That intervention would mean that instead of 16 bit padding constant, we will use 176 bit padding constant. According to the values in the Table 4 the probability of obtaining periods with the length less then 2^{80} in such case will be less then $2^{-87.85}$. Soon, following the suggestions in the latest announcement from ECRYPT, we will place a link from where the updated version Edon80 v2.0 will be available for interested parties.

4 Conclusions

We have build a mathematical probabilistic model by which the periods produced by Edon80 can be modelled. The Hong attack on the initial Edon80 design gave us valuable inspiration to do that[2]. However, we have to stress that the weaknesses noticed by Hong are due to the short internal pipeline,

[2]We thank Jin Hong for sending us the source C code of the program by which the experiments in [1] were performed.

and not due to the weaknesses of the concept of Edon80. Those weaknesses do not reveal any information about the used key or the internal state of the cipher and are easily removable.

We have proposed two scenarios how to deal with the further status of Edon80 as a submission to ECRYPT. We prefer the Scenario 2 and we have principal reasons for that. In the design of Edon80 we have used simple iterative design, that is nicely mathematically describable. That design allowed us to analyze it much deeper even in the case when some weaknesses were found in the initial design. Furthermore, if the internal pipeline is doubled, then the property of speed asymmetry in software and hardware that Edon80 has as a hardware stream cipher, will be increased even more, making it even more "exponentially slower on massively produced CPUs" and keeping it "very fast" when implemented in hardware.

References

[1] J. Hong: Remarks on the Period of Edon80, ECRYPT database, June 2005.

[2] D. Gligoroski, S. Markovski, L. Kocarev, and M. Gušev, Edon80 - Hardware synchoronous stream cipher. Symmetric Key Encryption Workshop, Århus, Denmark, May, 2005.