needed to ensure the probability (3) to be $\geq 95\%$. AES-128 is originally a 10-round *Substitution Permutation Network*. The full-round AES is one of the most cryptographically resistant state-of-the-art ciphers. The majority of the known attacks with the runtime significantly smaller than that of the exhaustive search are designed only for AES with reduced numbers of rounds. We studied the 2.5-round version of AES-128 (following the notation of (Bouillaguet, Derbez, and Fouque 2011), "x.5r" means x full rounds and the final round).

Finally, we studied the problem of finding a secret key given 12 blocks of the known plaintext (*12KP*) and the corresponding ciphertext for the reduced version of the Magma cipher (GOST 28147-89). This cipher was used in the USSR and Russia from 1989 to 2015. Originally, Magma is a 32-round cipher based on the *Feistel network* architecture. Similar to AES, significant improvements in cryptanalysis of Magma compared to the brute force attacks are known only for reduced-round variants of the Magma cipher. We studied the 8-round variant of Magma. For ensuring (3) to be $\geq 95\%$, we need 12 blocks of known plaintext.

The estimations of the guess-and-determine attacks constructed by our method for the considered ciphers are showed in Table 1. The sizes of the corresponding IBS sets are $|B| = 131$ (out of $|X| = 288$) for Trivium, $|B| = 63$ (out of $|X| = 256$) for Magma and $|B| = 42$ (out of $|X| = 128$) for AES.

Trivium. In (Borghoff, Knudsen, and Matusiewicz 2010) a guess-and-determine attack on Trivium is a result of solving a discrete optimization problem over a Boolean hypercube. In contrast to our generic method, the approach of Borghoff et al. targets only the cryptanalysis equations for Trivium and does not apply a general technique or method for constructing guess-and-determine attacks on a large class of cryptographic functions. Also, in (Borghoff, Knudsen, and Matusiewicz 2010) the effectiveness of a guess-and-determine attack is estimated in a completely different way if compared with the proposed resistance function. Borghoff et al. proposed an attack with the estimation of $4.31e{+}55$ seconds. A guess-and-determine attack on Trivium with the smallest runtime estimation is described in (Huang and Lin 2011), which uses the "Characteristic Set method" (CS-method) for solve the Trivium cryptanalysis equations. Although we acknowledge that the method proposed in (Huang and Lin 2011) outperforming the approach of the present paper (applied to Trivium) is somewhat discouraging, note that (Huang and Lin 2011) did not propose a versatile automatic procedure applicable to other ciphers (in contrast to our approach of resistance function minimization).

AES-128. To our best knowledge, the state-of-the-art guess-and-determine attacks on AES-128 with a reduced number of rounds were proposed in (Bouillaguet, Derbez, and Fouque 2011). Bouillaguet et al. considered a set of all possible sets of guessed bits as a tree traversed in a way similar to the branch-and-bound method. They claim that their method "... is reminiscent of the DPLL procedure implemented in many SAT-solvers". However, their approach does not use SAT solvers. Also, they do not estimate the runtime of a guess-and-determine attack similarly to what is done in the present paper: by analyzing the performance of an algorithm for solving a set of weakened cryptanalysis instances. (Bouillaguet, Derbez, and Fouque 2011) considered the cryptanalysis of a truncated AES-128 (with 2.5 rounds), in which 2KP are analyzed. The main disadvantage of this method consists in an enormous amount of memory consumption. In all our estimations the amount of memory required is a tiny fraction of the required runtime.

Table 1: Estimated hardness of the guess-and-determine attacks for weakened variants of Trivium, AES, and Magma compared to Previous Best Attacks (PBA), i.e. see (Huang and Lin 2011) for Trivium, (Bouillaguet, Derbez, and Fouque 2011) for AES, and (Courtois, Gawinecki, and Song 2012) for Magma. Time complexity is measured in seconds scaled to one core of the Intel Xeon E5-2695 v4 CPU while memory is measured in bits. Here by "negligible" we mean the amount of memory, which is standard for a modern PC.

Cipher	Time	Memory	Reference
Trivium	2.04e+41	negligible	present paper
	2.50e+34	negligible	PBA
AES-128	1.45e+15	negligible	present paper
	3.08e+16	2^{80}	PBA
Magma	3.55e+22	negligible	present paper
	1.17e+23	negligible	PBA

Magma. In (Courtois, Gawinecki, and Song 2012) the Magma cipher was studied by SAT solvers. To estimate the performance of guess-and-determine attacks, Courtois et al. introduced the notions of SAT-immunity and UNSAT-immunity, which, however, were not strictly formalized. Our notion of resistance function can be seen as a *concretization* of the notion of SAT-immunity. The attack of Courtois et al. (see its running time in Table 1) was constructed as a result of a thorough analysis of the Magma design features without using automatic algorithms for constructing guess-and-determine attacks. Note, that in (Courtois, Gawinecki, and Song 2012) 4KP were analyzed, while our attack requires 12KP in accordance with the aforementioned reasons. However, the runtime estimation of our attack is lower.

Conclusions

The paper studies a new class of Backdoors Sets for SAT (Inverse Backdoor Sets, IBS), which aims at facilitating efficient cryptographic attacks, namely guess-and-determine attacks. The values of the backdoor variables are used as bits to guess in the proposed guess-and-determine attack. The efficiency/hardness of the attack is defined as a value of a specific resistance function, which is estimated statistically