Information Processing Letters 115 (2015) 40–44

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

A note on quantum related-key attacks

Martin Roetteler [a,*], Rainer Steinwandt [b]

[a] Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA
[b] Florida Atlantic University, Boca Raton, FL 33431, USA

ARTICLE INFO

Article history:
Received 13 November 2013
Received in revised form 24 July 2014
Accepted 20 August 2014
Available online 26 August 2014
Communicated by D. Pointcheval

Keywords:
Cryptography
Quantum computing
Block ciphers
Related-key attacks
Hidden subgroup problems

ABSTRACT

In a basic related-key attack against a block cipher, the adversary has access to encryptions under keys that differ from the target key by bit-flips. In this short note we show that for a quantum adversary such attacks are quite powerful: if the secret key is (i) uniquely determined by a small number of plaintext–ciphertext pairs, (ii) the block cipher can be evaluated efficiently, and (iii) a superposition of related keys can be queried, then the key can be extracted efficiently.

1. Introduction

The availability of scalable quantum computers would jeopardize the security of many currently deployed asymmetric cryptographic schemes [1]. For symmetric cryptography the expectations for a post-quantum setting tend to be more optimistic, see, e.g., [2], from which we quote

> "quantum computers seem to have very little effect on secret-key cryptography, hash functions, etc. Grover's algorithm forces somewhat larger key sizes for secret-key ciphers, but this effect is essentially uniform across ciphers; today's fastest pre-quantum 256-bit ciphers are also the fastest candidates for post-quantum ciphers at a reasonable security level."

Related-key attacks are a powerful cryptanalytic tool when exploring block ciphers. In such attacks, the adversary is granted access to encryptions and/or decryptions of messages under secret keys which are related to the target key in a known or chosen way. As argued in [3], this type of attack is of practical interest, despite the assumptions made. When Winternitz and Hellman described this attack model more than 25 years ago, they focused on key relations given by bit-flips [4]. An illustrative example for an application of this attack model is an attack against 9 rounds of Rijndael with a 256-bit key, invoking 256 related keys with a particular choice of the bit-flips [5].

Current approaches to formalize related-key attacks allow more general key relations [6,7], and restricting to bit-flips can be considered to be a rather conservative choice. Below we show that for a quantum adversary such a basic form of related-key attack is quite powerful. We show that the possibility to query a superposition of related keys to a block cipher enables the efficient extraction of the secret key, if some rather mild conditions are met:

1. the block cipher can be implemented efficiently as a quantum circuit, and
2. the secret key is uniquely determined by a small number of available plaintext–ciphertext pairs.

The attack we describe is unlikely to pose a practical threat as querying a superposition of secret keys may not

* Corresponding author.
E-mail addresses: martinro@microsoft.com (M. Roetteler), rsteinwa@fau.edu (R. Steinwandt).

http://dx.doi.org/10.1016/j.ipl.2014.08.009
0020-0190/

be feasible for a typical implementation. Basically we require that the attacked honest user grants access to an implementation of the block cipher as a quantum circuit. Nonetheless, from the structural point of view our observation indicates an interesting limitation for the security guarantees of a block cipher that one can hope to prove in a post-quantum scenario, and our setting can be seen in the line of Boneh et al.'s quantum-accessible random oracle model [8].

2. Preliminaries

A block cipher with *key length* k and *block length* n is a family of 2^k permutations $\{E_K : \{0,1\}^n \to \{0,1\}^n\}_{K\in\{0,1\}^k}$ on bitstrings of length n. Popular block ciphers limit the possible choices of the key length k — e.g., for the Advanced Encryption Standard [9] we have $n = 128$ and $k \in \{128, 192, 256\}$. To characterize the efficiency of certain types of attacks, it can nonetheless be convenient to consider families of block ciphers, interpreting the key length k as a scalable security parameter. Measuring the running time of an adversary as a function of k, it is meaningful to speak of an expected polynomial time attack.

2.1. Related-key attacks

The attack model we consider goes back to [4]. After a key $K \in \{0,1\}^k$ has been chosen uniformly at random, the adversary has access to two oracles:

- $\mathcal{E}$: On input a bitmask $L \in \{0,1\}^k$ and a bitstring $m \in \{0,1\}^n$, this oracle returns the encryption $E_{K\oplus L}(m)$ of m under the key $K \oplus L$.
- $\mathcal{E}^{-1}$: On input a bitmask $L \in \{0,1\}^k$ and a bitstring $c \in \{0,1\}^n$, this oracle returns the decryption $E^{-1}_{K\oplus L}(c)$ of c under the key $K \oplus L$.

After interacting with these oracles, the adversary has to output a guess K' for K, and it is considered successful if and only if $K = K'$. For our attack we will also assume that the block cipher at hand can be evaluated efficiently, i.e., with a polynomial-size quantum circuit that has the secret key and a plaintext as input. For block ciphers that are actually used, this condition is of no concern.

The quantum attack below will not involve the decryption oracle, but we will allow the adversary to query the block cipher and also the oracle $\mathcal{E}$ with a superposition of keys. Finally, we require that the adversary has access to a polynomial number of plaintexts $m_1, \ldots, m_r$ such that for every pair of keys $(K, K') \in \{0,1\}^k \times \{0,1\}^k$ with $K \neq K'$ the condition

$$\big(E_K(m_1), \ldots, E_K(m_r)\big) \neq \big(E_{K'}(m_1), \ldots, E_{K'}(m_r)\big) \tag{1}$$

holds. As illustrated by the key schedule of SC2000-256, it is possible to have a block cipher where certain secret keys result in identical encryptions for all plaintexts [10], but this behavior is rather pathological. According to the strict key avalanche criterion [11,12], for a fixed plaintext each bit of the corresponding ciphertext should change with probability 1/2 if a key bit is flipped. So for two secret keys $K' \neq K$ we expect Inequality (1) to hold with probability about $1 - 2^{-rn}$, if the plaintexts m_i are pairwise different. Facing a total of $2^{2k} - 2^k$ key pairs (K, K') with $K \neq K'$, about $(2^{2k} - 2^k) \cdot 2^{-rn} \leq 2^{2k-rn}$ keys $K' \neq K$ violating Eq. (1) are expected. So it seems plausible to estimate that

$$r > \lceil 2k/n \rceil \tag{2}$$

plaintexts suffice to ensure that for every $K' \neq K$ at least one separating plaintext m_i is available. For the 128-bit version of AES, where $n = k = 128$, one can think of an r-value as small as 3. Throughout we will assume that r satisfies Inequality (2). Then the main idea to mount a quantum related-key attack is a reduction to a quantum algorithm described in [13] which we describe next.

2.2. Simon's problem

Let $f : \{0,1\}^k \to \{0,1\}^{k'}$ with $k \leq k'$ be a function such that one of the following two conditions holds:

(a) f is injective;
(b) there exists a bitstring $s \in \{0,1\}^k \setminus \{0^k\}$ such that for every two distinct $x, x' \in \{0,1\}^k$ we have

$$f(x) = f(x') \quad \Longleftrightarrow \quad x = x' \oplus s.$$

Simon's problem asks to decide for such a function f which of the two conditions holds, and in the case (b) to find s. Allowing the function f to be evaluated at a superposition of inputs, [13] establishes the following result:

Theorem 1. *Let $g(k)$ be an upper bound for the time needed to solve a $k \times k$ linear system of equations over the binary field $\mathbb{F}_2$, and let $t_f(k)$ be an upper bound for the time needed to evaluate the function f on (a superposition of) inputs from $\{0,1\}^k$. Then the above problem can be solved in expected time $O(k \cdot t_f(k) + g(k))$. In particular, for $t_f = t_f(k)$ being polynomial, the above problem can be solved in expected polynomial time.*

3. Description of the attack

Alluding to the Electronic Code Book mode of operation [14, Section 7.2.2], subsequently we will simply write $E_K(\vec{m})$ for the tuple of ciphertext blocks $(E_K(m_1), \ldots, E_K(m_r)) \in \{0,1\}^{rn}$. For a fixed, unknown secret key $s \in \{0,1\}^k \setminus \{0^k\}$ and messages $\vec{m} \in \{0,1\}^{rn}$ that characterize s uniquely as described in Section 2, we define the function

$$\begin{aligned} f_s : \quad \{0,1\}^k &\to 2^{\{0,1\}^{rn}} \\ x &\mapsto \big\{E_x(\vec{m}), E_{s\oplus x}(\vec{m})\big\}. \end{aligned}$$

We remark that for each x in the domain of f_s, the image is comprised of two different ciphertexts, i.e., it does not collapse to a singleton set. Indeed, this is the case due to the choice of the plaintexts $m_1, \ldots, m_r$ as the condition $s \neq 0^k$ implies that $E_x(\vec{m}) \neq E_{s\oplus x}(\vec{m})$. We next describe our core result, namely a reduction from the problem of finding the secret key s to an instance of Simon's problem which can then be solved efficiently on a quantum computer.