A sufficient condition for key-privacy

Shai Halevi
IBM T.J. Watson Research Center
shaih@alum.mit.edu

January 7, 2005

Abstract

The notion of *key privacy* for encryption schemes was defined formally by Bellare, Boldyreva, Desai and Pointcheval in Asiacrypt 2001. This notion seems useful in settings where anonymity is important. In this short note we describe a (very simple) sufficient condition for key privacy. In a nutshell, a scheme that provides data privacy is guaranteed to provide also key privacy if the distribution of *a random encryption of a random message* is independent of the public key that is used for the encryption.

1 Introduction

The motivation for the notion of key-privacy is anonymous communication, where eavesdroppers (or even active attackers) are prevented from learning the identities of the communicating parties. Assuming that the communicating parties are using public-key encryption, anonymous communication requires that an attacker cannot determine the public keys that were used to generate the ciphertexts that it sees.

This notion was defined formally by Bellare, Boldyreva, Desai and Pointcheval in Asiacrypt 2001 [BBDP01]. They defined this notion via probabilistic games, similar to the standard games that are used to define secrecy of plaintext. The difference is that the goal of the attacker is not to determine what message was encrypted under a known public key. Rather, the attacker tries to determine what public key was used in an encryption of a known message. Specifically, the attacker is given two public keys, it generates a message, and then it sees the encryption of that message under one of these keys. The attacker wins if it can guess what key was used to encrypt the message, and the scheme provides key-privacy if feasible attackers only have insignificant advantage over a random guess.

With this notion in mind, it seems clear that encryption schemes that are based on the Decision Diffie-Hellman assumption, such as ElGamal [ElG85] and Cramer-Shoup [CS98], provide key-privacy "right out of the box". After all, the ciphertext is such schemes consists of several random-looking group elements, regardless of the public key. Similarly, it seems clear that RSA-based schemes do not provide key-privacy, since a ciphertext looks like a random element modulo N, and in particular a ciphertext gives some information about the modulus N. To fix that, Bellare et al. described some techniques to ensure that ciphertexts will always end up in a common domain.

The arguments from above hint on the possibility that key-privacy can be proven information-theoretically, but this was not the case in any of the key-privacy proofs in the literature. Although we seem to have an "information theoretical intuition" for what scheme does or doesn't provide key-privacy, all the actual proofs were computational, essentially replicating the arguments that were used to prove secrecy of plaintext for the corresponding schemes. To understand why, recall that in the game as defined in [BBDP01], the attacker knows the message that is "hidden inside the ciphertext". With this knowledge, the ciphertext does provide information on the public key (in the information-theoretic sense), hence one must rely on some computational hardness to show key-privacy.

The simple observation in this note is that since the schemes in question are known to provide secrecy of plaintext, then in particular the attacker cannot distinguish between "an encryption of the right message" and "an encryption of a random message". Hence, for such schemes to provide also key privacy, it is sufficient that the attacker cannot distinguish between encryptions of random messages under the two public keys. In particular, in the DDH-based constructions from [BBDP01] the distribution of a random encryption of a random message is independent of the public key. Hence our "information theoretical intuition" for their providing key-privacy.

2 Observation

We have an encryption scheme $\mathcal{E} = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$. We assume that the message domain is implied by the public key, and denote it by $\mathrm{Domain}(pk)$. We assume that the reader if familiar with the standard notions of CPA-secure and CCA-secure encryption for data privacy. The corresponding notions for key privacy were defined in [BBDP01] by considering the following game:

1. The key generation algorithm is run twice with the security parameter to generate $(pk_0, sk_0) \leftarrow \mathrm{Gen}(1^k)$, $(pk_1, sk_1) \leftarrow \mathrm{Gen}(1^k)$, and a bit b is chosen at random in $\{0, 1\}$.

2. The attacker A is run with the two public keys pk as input, and it also has access to the two matching decryption oracles $\mathrm{Dec}_{sk_b}(\cdot)$.

3. The attacker A produces a message m, that belongs to the domains of both public keys, and it gets back the "target ciphertext" $c^* \leftarrow \mathrm{Enc}_{pk_b}(m)$.

4. The attacker continues to run with access to the decryption oracles as before, and it outputs a bit b'. It is considered successful if it never queried any of the decryption oracles on the target ciphertext c^*, and yet $b' = b$.

The scheme $\mathcal{E}$ is CCA-secure for key-privacy if any efficient attacker A can only be successful with probability at most negligibly more than 1/2, and CPA security for key-privacy is defined similarly, except that the attacker is not given access to the decryption oracle.

Observation 1 *Let $\mathcal{E} = (Gen, Enc, Dec)$ be an encryption scheme that is CCA-secure (resp. CPA-secure) for data-privacy. Then a sufficient condition for $\mathcal{E}$ to be also CCA-secure (resp. CPA-secure) for key-privacy if that the statistical distance between the two distributions*

$$\mathcal{D}_0 = \{(pk_0, pk_1, Enc_{pk_0}(m)) \;:\; (pk_0, sk_0) \leftarrow Gen(1^k),\; (pk_1, sk_1) \leftarrow Gen(1^k),\; m \leftarrow Domain(pk_0)\}$$
$$\mathcal{D}_1 = \{(pk_0, pk_1, Enc_{pk_1}(m)) \;:\; (pk_0, sk_0) \leftarrow Gen(1^k),\; (pk_1, sk_1) \leftarrow Gen(1^k),\; m \leftarrow Domain(pk_1)\}$$

is negligible.

Proof (sketch) The proof is elementary. Due to data-privacy, the attacker A in the key-privacy game cannot distinguish between the distributions where the message encrypted under pk_0 is the real message m or a random message. Similarly it cannot distinguish between the distributions where the message encrypted under pk_1 is the real message m or a random message. But if they were both random messages, then the condition above says that the distributions are statistically close, so clearly A cannot distinguish between them. Hence an advantage of ϵ in the the key-privacy game can be transformed into an advantage of at least (negligibly close to) $\epsilon/2$ in the data-privacy game. ■

An obvious extension. Of course we do not really need statistical closeness of $\mathcal{D}_0, \mathcal{D}_1$, it is clear that computational indistinguishability suffices. But it seems that the observation is more useful (in simplifying proofs, at least) when we have statistical closeness, since it is in that case that we can replace computational arguments by information-theoretic ones.

Interpretation in the random oracle model. If the scheme $\mathcal{E}$ is analyzed in the random-oracle model, then the algorithms $\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec}$ as well as the adversary A are also given oracle access to a function H (say, from $\{0,1\}^*$ to $\{0,1\}^k$ where k is the security parameter), and the games are analyzed with respect to a random function H.

Some care must be used when interpreting the observation above in the random-oracle model. Specifically, here the distributions $\mathcal{D}_0, \mathcal{D}_1$ depend also on the function H (which is modeled as a random function). If it is true that $\mathcal{D}_0(H) \approx \mathcal{D}_1(H)$ *for every fixed function* H then the observation from above holds just the same. However, the weaker condition that $\mathcal{D}_0(H_0) \approx \mathcal{D}_1(H_1)$ for random functions H_0, H_1 does not seem to suffice. It is likely possible to formulate some condition on the proof of data-privacy, such that "IF the data-privacy reduction works like that, THEN the weaker condition from above is sufficient". (The condition would essentially say that the attacker only query the H-oracle in the "relevant points for the encryption of the target ciphertext" with negligible probability.) But formulating and checking such condition does not appear to be substantially easier than proving key-privacy from scratch.

References

[BBDP01] Mihir Bellare, Alexandra Boldyreva, Anand Desai, and David Pointcheval. Key-privacy in public-key encryption. In *Advances in Cryptology - ASIACRYPT 2001*, volume 2248 of *Lecture Notes in Computer Science*, pages 566–582. Springer, 2001.

[CS98] Ronald Cramer and Victor Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In *Advances in Cryptology – CRYPTO '98*, volume 1462 of *Lecture Notes in Computer Science*, pages 13–25. Springer-Verlag, 1998.

[ElG85] T. ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In *Advances in Cryptology – CRYPTO '84*, volume 196 of *Lecture Notes in Computer Science*, pages 10–18. Springer-Verlag, 1985.