

Lower Bounds for Private Broadcast Encryption

Aggelos Kiayias and Katerina Samari

Department of Informatics and Telecommunications, University of Athens
{aggelos,ksamari}@di.uoa.gr

Abstract. Broadcast encryption is a type of encryption where the sender can choose a subset from a set of designated receivers on the fly and enable them to decrypt a ciphertext while simultaneously preventing any other party from doing so. The notion of *private* broadcast encryption extends the primitive to a setting where one wishes to thwart an attacker that additionally attempts to extract information about what is the set of enabled users (rather than the contents of the ciphertext).

In this work we provide the first lower bounds for the ciphertext size of private broadcast encryption. We first formulate various notions of privacy for broadcast encryption, (**priv-eq**, **priv-st** and **priv-full**) and classify them in terms of strength. We then show that any private broadcast encryption scheme in the sense of **priv-eq** (our weakest notion) that satisfies a simple structural condition we formalize and refer to as “atomic” is restricted to have ciphertexts of size $\Omega(s \cdot k)$ where s is the cardinality of the set of the enabled users and k is the security parameter. We then present an atomic private broadcast encryption scheme with ciphertext size $\Theta(s \cdot k)$ hence matching our lower bound that relies on key privacy of the underlying encryption. Our results translate to the setting **priv-full** privacy for a ciphertext size of $\Theta(n \cdot k)$ where n is the total number of users while relying only on KEM security. We finally consider arbitrary private broadcast encryption schemes and we show that in the **priv-full** privacy setting a lower-bound of $\Omega(n+k)$ for every ciphertext is imposed. This highlights the costs of privacy in the setting of broadcast encryption where much shorter ciphertexts have been previously attained with various constructions in the non-privacy setting.

1 Introduction

Consider the setting of an encrypted file system. Each file is encrypted so that only a designated subset of the set of users of the system can retrieve it. An attacker, who may be controlling a set of system users should be incapable of recovering the contents of the file provided that none of the controlled users belong to the enabled set for the file.

This setting is one of the application domains for broadcast encryption, a cryptographic primitive introduced by Fiat and Naor [9]. Broadcast encryption is also suitable for application to the setting of content distribution and is indeed widely used as the encryption system of DVDs (for example in the form of the AACS [1]) and other media content carrying mechanisms. A variety of schemes

have been developed over the years with the main objective of reducing the ciphertext length. Currently in the private key setting (see e.g. [14]) there are schemes that achieve a ciphertext length of $\Theta(r \cdot k)$ where r is the number of revoked users and k is the security parameter; in the public-key setting, using bilinear maps the scheme of [4] achieves a ciphertext length of $O(k)$ with public key of $O(n \cdot k)$ for any set of enabled users and the scheme of Delerablée [6] achieves a ciphertext length $O(k)$ while the public-key is of size $\Theta(s \cdot k)$ assuming that sets of enabled users never exceed cardinality s .

Barth, Boneh and Waters [3] put forth the notion of private broadcast encryption. Their objective is to consider another class of attacks for broadcast encryption where the goal of the attacker is to discover information about the set of enabled users rather than decrypting a ciphertext for which it is not enabled. Protecting the privacy of the users in the enabled set can be an equally and some times perhaps an even more important goal than the privacy of the message. Indeed, hiding the information that one is a recipient of a message, from other users and even from other recipients of the same message, is a critical security feature in any setting where the fact of receiving a message at a certain time or frequency reveals sensitive personal characteristics of the recipient. For example, in a file system, an encrypted system file under a certain account may reveal that the said account has a certain level of system privileges and this fact can assist an attacker in a more complex attack vector.

To address this important problem, Barth et al. [3] introduced a security model for private broadcast encryption and provided a first solution. The scheme of [3] applies to the public-key setting and has the characteristic of being linear in the number of users, i.e., has a ciphertext of length $\Theta(s \cdot k)$ where s is the number of enabled users. Given that, as shown above, previously known (non-private) schemes achieve much better ciphertext lengths, it is an important open question to improve this efficiency characteristic for private broadcast encryption schemes or demonstrate that no further improvement is possible.

In this work, motivated by the above, we provide various results suggesting the latter state of affairs by proving tight lower bounds for the ciphertext length of private broadcast encryption schemes. We outline our results below.

First, we study the formalization of the notion of privacy in the context of private broadcast encryption. We introduce three security formulations. The first notion we consider is inspired by that in [3]: it allows the adversary to interact with the broadcast encryption system by obtaining encryption and decryption queries as well as corrupting recipients. Upon completion of a first stage the adversary provides two target sets of users to be revoked R_0, R_1 . Then, provided that $|R_0| = |R_1|$, the adversary receives as a challenge a message M and an encryption of M with the set of users R_b revoked where b is a random bit. The adversary has to guess the bit b under the constraint that it does not submit the challenge ciphertext to a decryption oracle and does not control any user in the symmetric difference $R_0 \Delta R_1$. We call this level of privacy `priv-eq`.

We observe `priv-eq` is quite insufficient for many reasonable attack settings. Specifically, for a certain ciphertext the adversary may be absolutely certain that

the set of users R is revoked and only wishes to test whether an additional target user i is also revoked or not. Clearly this attack objective is not captured by the above definition since in this case it holds that $R_0 = R$ and $R_1 = R \cup \{i\}$, two sets of different cardinality. We formalize this notion of privacy as **priv-st**. It is very easy to see that there exist schemes that satisfy **priv-eq** and fail **priv-st**; in particular, any scheme that leaks the cardinality of the set of revoked users is such a candidate and in fact the scheme of [3] is one such scheme.

Taking this one step further we introduce *full privacy* to be the property where the adversary cannot distinguish any two sets R_0, R_1 ; we term this notion as **priv-full**. We then prove that in fact **priv-st** and **priv-full** are equivalent.

Armed with this definitional basis we proceed to our lower bounds. We first consider the case of *atomic* broadcast encryption schemes. Atomic schemes have the characteristic that the ciphertext can be broken to a number of discrete components and each recipient when it is decrypting it applies a decryption function to one or more of those components. The private schemes of [3] satisfy this condition and it is also quite common in the wide class of combinatorial broadcast encryption schemes; a partial list of non-private atomic schemes is the following ([14],[12],[11],[16],[2]).

For such atomic schemes, we prove that any scheme that satisfies the **priv-eq** condition is susceptible to an attack against privacy in the case when the ciphertext drops below $s \cdot k$ where s is the cardinality of the set of enabled users. This means that a lower bound of $\Omega(s \cdot k)$ is in place. We then present an atomic private broadcast encryption scheme with this complexity hence showing the lower bound is tight. The scheme itself is a standard linear length construction; the scheme applies equally to the symmetric and public-key setting and abstracts the necessary properties needed for privacy to the existence of secure key-private encryption mechanism in the KEM sense [15]. We present a similar set of results for the **priv-full** level of privacy; in this case KEM security is sufficient and the corresponding tight bound is $\Theta(n \cdot k)$.

Having settled the case of atomic broadcast encryption, we switch our focus to the setting of general private broadcast encryption schemes (that are not necessarily atomic). We first show using an information theoretic argument that any broadcast encryption scheme should exhibit some ciphertexts of length $\Omega(n+k)$. Using this as a stepping stone we then prove that if a broadcast encryption scheme is assumed to be private in the sense of **priv-st**, **priv-full**, it will have to provide a ciphertext of length $\Omega(n+k)$ for any set of revoked users R hence a complexity bound sublinear in the number of users is impossible to be achieved if full privacy is desired.

Related Work. Independently of the present work, Libert, Paterson and Quaglia [13] have studied the problem of “anonymous broadcast encryption” where the main focus is to enable efficient decryption in the setting where the ciphertext is of length $\Theta(s \cdot k)$. In this case the known schemes that satisfy privacy require from the users to test sequentially until they find the proper element they can decrypt. In the public-key setting this can be an arduous task if the number of enabled users is large; by using some randomized tagging mechanism it is

possible to improve the decryption time complexity. Our modeling is consistent with that of [13] and our lower bounds readily apply to their setting as well.

Fazio and Perera in [8], introduce a weaker notion of anonymity compared to the one considered here and in previous works, called *outsider-anonymity*. An *Outsider-anonymous broadcast encryption scheme* ensures that a user in the revoked set can gain no information about the enabled set while a member of the enabled set may extract information about some other users in it. Taking advantage of this relaxation to the anonymity definition, the authors employ an atomic scheme, i.e. the public key variant of Complete Subtree method [7], in order to achieve sublinear ciphertext size.

2 Privacy Notions for Broadcast Encryption

Broadcast encryption is a triple $\langle \text{KeyGen}, \text{Encrypt}, \text{Decrypt} \rangle$ where **KeyGen** generates a set of n keys for any given n and **Encrypt** receives a set of revoked users $R \subseteq [n]$ that should be barred from decrypting. We define privacy in broadcast encryption using an experiment between a challenger and an adversary. The adversary is given access to an Encryption Oracle which means that he is capable of obtaining ciphertext-message pairs that can be decrypted by an enabled set of his choice. Also, he is able to derive the secret keys of a selected set of users, by submitting a number of queries to a Corruption Oracle. We will distinguish three levels of privacy in our formalization. In the most general type (full privacy), **priv-full**, the adversary should be unable to distinguish between any two sets of revoked users as long as the corrupted users do not cover the symmetric difference of the two sets. In the case of “single target” privacy, **priv-st** the adversary wishes to understand whether a single (target) user is included in an (otherwise) known revoked set. Finally, in privacy among equal sets, **priv-eq**, is identical to the case of **priv-full** with the additional restriction that the adversary should challenge on two sets with equal cardinality. Formally, we have the following:

EncryptionOracle(R)	CorruptOracle(u)	DecryptionOracle(u, c)
<i>retrieve</i> ek	$T \leftarrow T \cup \{u\}$	$D \leftarrow D \cup \{(u, c)\}$
$m \xleftarrow{r} M$	<i>return</i> K_u	<i>retrieve</i> K_u
$c \leftarrow \text{Encrypt}(ek, m, R)$		<i>return</i> $\text{Decrypt}(K_u, c)$
<i>return</i> (c, m)		

Experiment $\text{Exp}_A^{\text{priv-x}}(1^n, 1^\lambda)$
 $(ek, K_1, \dots, K_n) \leftarrow \text{KeyGen}(1^n, 1^\lambda)$
 $T \leftarrow \emptyset$
 $(state, R_0, R_1) \leftarrow \mathcal{A}^{\text{CorruptOracle}(\cdot), \text{EncryptionOracle}(\cdot), \text{DecryptionOracle}(\cdot)}(1^\lambda)$
 $b \xleftarrow{r} \{0, 1\}$
 $m \xleftarrow{r} M$
 $c^* \leftarrow \text{Encrypt}(ek, m, R_b)$
 $b^* \leftarrow \mathcal{A}^{\text{CorruptOracle}(\cdot), \text{EncryptionOracle}(\cdot), \text{DecryptionOracle}(\cdot)}(guess, (c^*, m), state)$
 if $(\exists i \in T \text{ such that } i \in (R_0 \Delta R_1)) \vee$

$(\exists(i, c) \in D$ such that $i \in (R_0 \Delta R_1)$ and $c = c^*$)
 then output a random bit else if $b = b^*$ then return 1 else 0;

Definition 1 (Privacy). *Let Φ be a fully exclusive broadcast encryption scheme with n receivers. We say that Φ is private priv-x , if for all PPT adversaries \mathcal{A} ,*

$$\text{Prob}[\text{Exp}_{\mathcal{A}}^{\text{priv-x}}(1^n, 1^\lambda) = 1] \leq \frac{1}{2} + \varepsilon,$$

where ε is a negligible function of λ and λ is the security parameter.

Based on the definition above, we provide three different definitions for privacy whose differences concern the form of the challenge (R_0, R_1) .

- We call $\text{Exp}^{\text{priv-full}}$ the experiment in which R_0, R_1 can be any set which is subset of $[n]$.
- With $\text{Exp}^{\text{priv-st}}$, we define the experiment where R_0, R_1 have to be of the form R and $R \cup \{i\}$, accordingly.
- With $\text{Exp}^{\text{priv-eq}}$, we define the experiment where R_0, R_1 have to be of equal size. Consequently, it is necessary to add one more or-factor, $(|R_0| \neq |R_1|)$, in the condition of the last line of the experiment, to guarantee that the experiment outputs a random bit in case the adversary's challenge sets are of unequal size.

We then proceed to show relations between the three notions of privacy.

- Theorem 1.**
1. *Privacy definitions priv-st and priv-full are equivalent.*
 2. *Privacy definition priv-full implies the privacy definition priv-eq .*
 3. *Privacy definition priv-eq does not imply privacy definition priv-st .*

Proof. 1. We need to prove two directions in order to show that these definitions are equivalent. The easy direction is the one which says that privacy definition priv-full implies privacy definition priv-st . If we assume that there exists a PPT adversary \mathcal{A} that breaks privacy definition priv-st challenging a pair $(R, R \cup \{i\})$ with non-negligible advantage α , this adversary also breaks privacy definition priv-full considering that $R_0 = R$ and $R_1 = R \cup \{i\}$. The opposite direction will be derived from lemma 1.

2. Assuming that there exists a PPT adversary that breaks privacy definition priv-eq having advantage α , then the same adversary does also break privacy definition priv-full with non-negligible advantage α .
3. It suffices to provide a broadcast encryption scheme which satisfies the definition priv-eq but not private according to the definition priv-full . Let Φ be a broadcast encryption scheme which is priv-eq . Now consider Φ' to be exactly like Φ but with the added feature that the encryption algorithm appends at the end of all ciphertexts the cardinality of the revoked set. It is obvious that this scheme is inherently incapable of satisfying privacy definition priv-full (while it remains priv-eq). Such schemes exist under standard cryptographic assumptions as we will see in section 4. ■

Lemma 1. *Let Φ be a broadcast encryption scheme with n receivers. If there exists a PPT adversary that has advantage α in breaking privacy definition priv-full , then there exists a PPT adversary that breaks privacy definition priv-st with probability at least $1/2 + \alpha/n$.*

Sketch of Proof: Let \mathcal{A} be a PPT adversary that breaks priv-full definition with advantage α . Conditioning on the fact that \mathcal{A} breaks privacy for a pair of sets (R_0, R_1) , we consider a sequence of sets P_0, \dots, P_{k-1} , where $k = |R_0 \Delta R_1| + 1$, $P_0 = R_0$ and $P_{k-1} = R_1$. We set $m = |R_0 \setminus R_1|$ and we define P_i as follows: if $i \in \{0, \dots, m\}$ $P_i = P_{i-1} \setminus \{j\}$, for some user $j \in R_0 \setminus R_1$, otherwise $P_i = P_{i-1} \cup \{j'\}$ for some user $j' \in R_1 \setminus R_0$. Namely, all the members of this sequence are supersets of $R_0 \cap R_1$ and every pair of consecutive sets are of the form $(R, R \cup \{i\})$ for some R . We denote as \mathcal{A}_1 the part of the algorithm \mathcal{A} that corresponds to the training stage of the experiment, i.e. before the output of challenge, while with \mathcal{A}_2 we denote \mathcal{A} 's steps after the receipt of the response. Together with the challenge pair (R_0, R_1) , \mathcal{A}_1 outputs a random variable *state*.

We construct a PPT adversary \mathcal{B} that breaks definition priv-st as follows: \mathcal{B} runs \mathcal{A}_1 until he outputs the challenge pair (R_0, R_1) together with *state*. Then \mathcal{B} makes a guess $j \in \{0, \dots, k - 2\}$ and challenges the corresponding pair. Due to the structure of the sequence, if $j \in \{0, \dots, m - 1\}$ \mathcal{B} challenges (P_{j+1}, P_j) , otherwise challenges (P_j, P_{j+1}) . The received response is provided together with *state* to \mathcal{A}_2 . Then, if $j \in \{0, \dots, m - 1\}$ \mathcal{B} outputs the complement of \mathcal{A}_2 's output, otherwise outputs \mathcal{A}_2 's output. We conclude that \mathcal{B} breaks definition priv-st with advantage $\alpha/(k - 1)$ which is at least α/n . ■

3 Lower Bounds for Atomic Broadcast Encryption Schemes

Definition 2. *An atomic broadcast encryption scheme with n receivers is defined as a tuple of algorithms (KeyGen, Encrypt, Decrypt) :*

- **KeyGen:** On input $1^n, 1^\lambda$, it generates the set of keys (ek, SK_1, \dots, SK_n) , where ek is the encryption key and SK_i is the decryption key assigned to a user i . Each decryption key SK_i is a set which consists of elements $\{sk_{ij}\}_{j=1}^\ell$ (we call those atomic keys) for some value ℓ which is not necessarily the same for each user. It also produces the description of a language \mathcal{L} which encodes all the possible subsets of users that may be provided as input to the encryption function.
- **Encrypt:** On input a message m , the encryption key ek and a revocation instruction $R \in \mathcal{L}$, it outputs a ciphertext C such that $C \leftarrow \text{Encrypt}(ek, m, R)$ which among possibly other values, contains a number of components c_1, \dots, c_p (we call those the atomic ciphertexts of C).
- **Decrypt:** On input a ciphertext C , such that $C \leftarrow \text{Encrypt}(ek, m, R)$ and a decryption key SK_i : It outputs m if $i \notin R$ and some value $x \neq m$ if $i \in R$. Depending on the instantiation, x could be the symbol \perp , or some plaintext sampled independently of m .

For atomic broadcast encryption schemes we further assume the existence of a deterministic algorithm called **Decryptmatching** which matches the atomic ciphertexts of a ciphertext tuple C with the atomic keys under which they are decrypted. In all cases we know, this algorithm is in part of the Decryption algorithm.

Proposition 1. *The broadcast encryption schemes that rely on the subset cover framework [14] are atomic. The private schemes of [3] are atomic.*

Given that in this section we will provide lower bounds, we provide a weaker definition of privacy which departs from definition **priv-eq** in the existence of the **CorruptOracle** and **DecryptionOracle** in the security experiment. More precisely, the adversary is not given access to a Decryption Oracle and instead of being provided access to a Corruption Oracle, he is given access to an Atomic Decryption Oracle which operates as follows:

$$\text{AtDecOr}(j, t, C) = \begin{cases} 0 & \text{if no atomic ciphertext in } C \text{ is supposed to be decrypted} \\ & \text{under the key } sk_{jt} \\ \perp & \text{if the number of keys in the set } SK_j \text{ are less than } t \\ 1 & \text{if there exists an atomic ciphertext that can be decrypted} \\ & \text{under the key } sk_{jt} \end{cases}$$

EncryptionOracle (R) <i>retrieve</i> ek $m \xleftarrow{r} M$ $c \leftarrow \text{Encrypt}(ek, m, R)$ <i>return</i> (c, m)	AtDecOr (j, t, C) $E \leftarrow E \cup \{(j, t)\}$ <i>return</i> $x \in \{0, 1, \perp\}$
--	---

Experiment $\text{Exp}_{\mathcal{A}}^{\text{priv-eq-at}}(1^n, 1^\lambda)$
 $(ek, K_1, \dots, K_n) \leftarrow \text{KeyGen}(1^n, 1^\lambda)$
 $T \leftarrow \emptyset$
 $(state, R_0, R_1) \leftarrow \mathcal{A}^{\text{AtDecOr}(\cdot), \text{EncryptionOracle}(\cdot)}(1^\lambda)$
 $b \xleftarrow{r} \{0, 1\}$
 $m \xleftarrow{r} M$
 $c^* \leftarrow \text{Encrypt}(ek, m, R_b)$
 $b^* \leftarrow \mathcal{A}^{\text{AtDecOr}(\cdot), \text{EncryptionOracle}(\cdot)}(\text{guess}, (c^*, m), state)$
 if $(\exists(i, \cdot) \in E \text{ such that } i \in (R_0 \Delta R_1)) \vee (|R_0| \neq |R_1|)$
 then output a random else if $b = b^*$ then return 1 else 0;

The experiment $\text{Exp}_{\mathcal{A}}^{\text{priv-eq-at}}$ is defined identically to $\text{Exp}_{\mathcal{A}}^{\text{priv-eq}}$ with the oracle **AtDecOr** substituting the corruption and decryption oracles.

Definition 3. *Let Φ be a broadcast encryption scheme with n receivers. We say that Φ is private **priv-eq-at**, if for all PPT adversaries \mathcal{A} ,*

$$\text{Prob}[\text{Exp}_{\mathcal{A}}^{\text{priv-eq-at}}(1^n, 1^\lambda) = 1] \leq \frac{1}{2} + \varepsilon,$$

where ε is a negligible function of λ and λ the security parameter.

The following proposition is easy:

Proposition 2. *Any broadcast encryption scheme Φ that satisfies privacy definition priv-eq , does also satisfy privacy definition priv-eq-at .*

Proof. It is easy to see that assuming the existence of a PPT adversary \mathcal{A} that has non-negligible advantage in breaking privacy definition priv-eq-at , there is a PPT adversary \mathcal{B} that breaks privacy definition priv-eq with the same advantage as \mathcal{A} executing \mathcal{A} inside him. The proof relies on the fact that \mathcal{B} can perfectly answer the queries submitted by \mathcal{A} to the Atomic Decryption Oracle because of his access to a Corruption Oracle.

Theorem 2. *(Lower bound for atomic schemes) Let Φ be an atomic broadcast encryption scheme and suppose that there exists an enabled set $S \subseteq [n]$ such that the number of atomic ciphertexts included in the prepared ciphertext C_S are less than $|S|$. Then, the scheme is not private according to definition priv-eq-at .*

Proof. We will assume that for every R the atomic ciphertexts produced by the algorithm **Encrypt** are always decrypted under the same set of atomic keys (in the other case, if the algorithm **Encrypt** flips a number of coins in order to decide the atomic keys that will be used, then the same argument we present below can take place with the only difference that in this case the adversary will have to run a number of times the algorithm **Encrypt** for the set R_0 to approximate the distribution). Let us assume that there exists such a set S_0 and let C_{S_0} be a ciphertext produced by the algorithm **Encrypt** on input ek, m, R_0 with $R_0 = [n] \setminus S_0$. Then, according to the pigeonhole principle, there exists at least one atomic ciphertext c_k in the ciphertext C_{S_0} that can be decrypted by at least two users $i, j \in [n]$. As a result, the ciphertext c_k can be decrypted under an atomic key sk_m which is a member of both sets SK_i, SK_j , where SK_i, SK_j are the sets of atomic decryption keys of i and j accordingly. Given this, an adversary \mathcal{A} that breaks privacy can be constructed following the logic presented below:

1. If $i, j \in [n]$ are two users which decrypt the same atomic ciphertext in a ciphertext tuple C_{S_0} , where $C_{S_0} \leftarrow \text{Encrypt}(ek, m, R_0)$, select a set R_1 such that $|R_1| = |R_0|$, $i \in R_1$ and $j \notin R_1$. Choose arbitrarily the other $|R_1| - 1$ members of R_1 and challenge R_0, R_1 .
2. When the response C^* is received, issue a query R_0 to the Encryption Oracle which is replied with a ciphertext C .
3. Submit a number of queries of the form (j, t, C) to the Atomic Decryption Oracle, for all the possible values of t , starting from $t = 1$, until **AtDecOr** returns \perp . If we ignore the symbol \perp , the output of this procedure is a bitstring x_1 of length s , where s is the number of atomic keys included in the decryption key of SK_j .
4. Repeat the same procedure submitting queries on inputs of the form (j, t, C^*) and obtain a bitstring x_2 of length k (note that this is allowed since j is enabled in both challenge ciphertexts). If $x_1 \neq x_2$, then answer 1 else 0. ■

Corollary 1. *Any atomic broadcast encryption scheme with n receivers and ciphertext length less than n cannot be private according to definition `priv-full`.*

Proof. If $R = \emptyset$ and the atomic ciphertexts are less than n , the assumption of the Theorem 2 takes place for $S = [n]$. It is easily observed that the fact that the challenge sets R_0, R_1 were of equal length played no crucial role in the proof of Theorem 2. Thus, we can apply exactly the same arguments with $R = \emptyset$ being the one set in the challenge.

Corollary 2. *For any atomic broadcast encryption scheme Φ with n receivers which is private according to `priv-eq` definition, it holds that for any enabled set $S \subseteq [n]$, the ciphertext length is $\Omega(k \cdot |S|)$ bits, where k is the maximum size of an atomic ciphertext. For any broadcast encryption scheme which is private according to `priv-full` definition, the ciphertext length is $\Omega(k \cdot n)$ for all the enabled sets $S \subseteq [n]$.*

4 Constructions of Atomic Private Broadcast Encryption Schemes

In this section, we present matching schemes for the lower bounds of the previous section. We focus on CCA-1 security for simplicity but our results can be easily extended to CCA-2 security. Due to lack of space most of our results are presented without proofs; full proofs are presented in the full version. We first consider security in the sense of key encapsulation mechanisms (KEM) defined with the aid of the following experiment:

Experiment $\text{Exp}_A^{KEM}(1^\lambda)$
 Select k at random.
 $aux \leftarrow \mathcal{A}^{\text{Enc}_k(\cdot), \text{Dec}_k(\cdot)}$
 $m_0, m_1 \xleftarrow{r} M$;
 $b \xleftarrow{r} \{0, 1\}$; $c \leftarrow \text{Enc}_k(m_b)$
 $b^* \leftarrow \mathcal{A}^{\text{Enc}_k(\cdot)}(m_1, c)$
 if $b = b^*$ then return 1 else 0;

Definition 4. *We say that the symmetric encryption scheme $(\text{Gen}, \text{Enc}, \text{Dec})$ is `KEM-secure` if for any probabilistic polynomial time adversary \mathcal{A} it holds that*

$$\text{Prob}[\text{Exp}_A^{KEM}(1^\lambda)] \leq \frac{1}{2} + \varepsilon,$$

where ε is a negligible function of λ .

Experiment $\text{Exp}_A^{BE-KEM}(1^n, 1^\lambda)$
 $(ek, K_1, \dots, K_n) \leftarrow \text{KeyGen}(1^n, 1^\lambda)$
 $T \leftarrow \emptyset$
 $R \leftarrow \mathcal{A}^{\text{CorruptOracle}(\cdot), \text{EncryptionOracle}(\cdot), \text{DecryptionOracle}(\cdot)}(\cdot)$
 $b \xleftarrow{r} \{0, 1\}$

$m_0, m_1 \xleftarrow{r} M$
 $c^* \leftarrow \text{Encrypt}(ek, m_b, R)$
 $b^* \leftarrow \mathcal{A}^{\text{EncryptionOracle}(\cdot)}(c^*, m_1)$
 If $T \not\subseteq R$ then output a random bit
 else if $b = b^*$ then return 1 else 0;

Definition 5. Let Φ be a broadcast encryption scheme with n receivers. We say that a broadcast encryption scheme Φ is *KEM-secure* if for any probabilistic polynomial time adversary \mathcal{A} it holds that

$$\text{Prob}[\text{Exp}_{\mathcal{A}}^{BE-KEM}(1^n, 1^\lambda) = 1] \leq \frac{1}{2} + \varepsilon,$$

where ε is a negligible function of λ .

Experiment $\text{Exp}_{\mathcal{A}}^{\text{key-priv}}(1^\lambda)$
 Select $k_0 \leftarrow \text{Gen}(1^\lambda); k_1 \leftarrow \text{Gen}(1^\lambda)$
 $aux \leftarrow \mathcal{A}^{\text{Enc}_{k_0}(\cdot), \text{Enc}_{k_1}(\cdot), \text{Dec}_{k_0}(\cdot), \text{Dec}_{k_1}(\cdot)}$
 $m \xleftarrow{r} M$
 $b \xleftarrow{r} \{0, 1\}; c \leftarrow \text{Enc}_{k_b}(m)$
 $b^* \leftarrow \mathcal{A}^{\text{Enc}_{k_0}(\cdot), \text{Enc}_{k_1}(\cdot)}(m, c)$
 if $b = b^*$ then return 1 else 0;

Definition 6. We say that the symmetric encryption scheme $(\text{Gen}, \text{Enc}, \text{Dec})$ is *key private* if for any probabilistic polynomial time adversary \mathcal{A} it holds that

$$\text{Prob}[\text{Exp}_{\mathcal{A}}^{\text{key-priv}}(1^\lambda)] \leq \frac{1}{2} + \varepsilon,$$

where ε is a negligible function of λ .

Scheme 1. This scheme is defined as a tuple of algorithms $(\text{KeyGen}, \text{Encrypt}, \text{Decrypt})$ which are described below. A basic component of the scheme is the underlying symmetric encryption scheme $(\text{Gen}, \text{Enc}, \text{Dec})$.

- **KeyGen** : On input $1^n, 1^\lambda$:
 - For any user $i \in [n]$ run the algorithm $\text{Gen}(1^\lambda)$ which generates a key k_i . The encryption key is $ek = \{k_j\}_{j \in [n]}$.
- **Encrypt**: On input a message m and a revoked set R :
 - By employing the scheme $(\text{Gen}, \text{Enc}, \text{Dec})$ compute a ciphertext tuple c as follows: For each $i \in [n] \setminus R$ compute $\text{Enc}_{k_i}(m)$. Perform a random permutation f to the ciphertext components which results to a ciphertext tuple of length s , where s is the cardinality of the set $[n] \setminus R$.
- **Decrypt**: On input a ciphertext $c = \langle c_1, \dots, c_s \rangle$ and a key k_u :
 - Starting from c_1 , try to decrypt each ciphertext component under the key k_u . If there exists c_j that is supposed¹ to be decrypted by u , return $\text{Dec}_{k_u}(c_j)$.

¹ In order to determine this *strong* correctness is required; this notion means that applying a wrong key to a ciphertext results to a special fail message to be returned. This can be achieved e.g., by appending a value $H(M)$ to all plaintexts M (here H is a hash function); we omit further details.

Scheme 2. This scheme is defined as a tuple of algorithms (KeyGen, Encrypt, Decrypt) which we describe below. A basic component of the scheme is the underlying symmetric encryption scheme (Gen, Enc, Dec).

- KeyGen : On input $1^n, 1^\lambda$:
 - For any user $i \in [n]$ run the algorithm Gen(1^λ) which generates a key k_i . The encryption key is $ek = \{k_j\}_{j \in [n]}$.
- Encrypt: On input a message m and a revoked set R :
 - By employing a scheme (Gen, Enc, Dec) compute a ciphertext tuple c of length n as follows: For any user $i \in [n]$, if $i \in R$ choose randomly a message $m' \in M$, compute $E_{k_i}(m')$ and place $E_{k_i}(m')$ at the i -th position. If $i \notin R$, compute $\text{Enc}_{k_i}(m)$ and place it to the i -th position.
- Decrypt: On input a ciphertext $c = \langle c_1, \dots, c_n \rangle$ and a key k_u of a user u :
 - Compute $\text{Dec}_{k_u}(c_u)$.

Theorem 3. *If Scheme 1 satisfies that the underlying scheme (Gen, Dec, Enc) key-private then Scheme 1 is private according to the definition priv-eq.*

Theorem 4. *If Scheme 2 is a broadcast encryption scheme in which the underlying scheme (Gen, Dec, Enc) is KEM-secure, then Scheme 2 is private according to definition priv-full.*

It remains to show that the broadcast encryption schemes Scheme 1 and Scheme 2 are BE-KEM-secure, i.e. they are secure under the definition 5. The proofs of security are similar and we prove this only for Scheme 2.

Theorem 5. *If the underlying encryption scheme (Gen, Enc, Dec) is KEM-secure then Scheme 2 is BE-KEM secure.*

Proof. Let \mathcal{A} be a PPT adversary that breaks BE-KEM security such that $\text{Prob}[\text{Exp}_{\mathcal{A}}^{BE-KEM}(1^n, 1^\lambda) = 1] \geq \frac{1}{2} + \alpha$, for α non-negligible. We define a sequence of experiments $\text{Exp}_0^{\mathcal{A}}, \dots, \text{Exp}_n^{\mathcal{A}}$, where $\text{Exp}_0^{\mathcal{A}}$ is the experiment $\text{Exp}_{\mathcal{A}}^{BE-KEM}$. We define as $\text{Exp}_v^{\mathcal{A}}$ the experiment which operates exactly as $\text{Exp}_0^{\mathcal{A}}$ modified in a way that the first v enabled users will be given encryptions of randomly chosen plaintexts rather than the encryption of the appropriate plaintext. If s is the size of the enabled set, for $v = s, s + 1, \dots, n$ the experiments are the same.

Now, let $p_0 = \text{Prob}[\text{Exp}_0^{\mathcal{A}} = 1]$ and $p_1 = \text{Prob}[\text{Exp}_1^{\mathcal{A}} = 1]$. Moreover, let \mathcal{B} be an attacker against KEM-security of the scheme (Gen, Enc, Dec). \mathcal{B} guesses i to be the user he will play $\text{Exp}_{\mathcal{B}}^{KEM}$ and then running $n - 1$ times the algorithm Gen(1^λ) he generates the private keys for the other users. When \mathcal{A} challenges R , \mathcal{B} checks whether i is the first enabled user and returns 0 if this does not hold. Otherwise, when \mathcal{B} receives $(m_1, \text{Enc}_k(m_b))$, he places $\text{Enc}_k(m_b)$ at the first position and then chooses randomly a message m' from the plaintext space and flips a perfect coin b' . \mathcal{B} sets $m'_{b'} = m_1$ and $m'_{1-b'} = m'$ and encrypts the message $m'_{b'}$ for the enabled users except for i . \mathcal{B} encrypts a message m'' for the revoked users which is randomly chosen from the plaintext space. \mathcal{B} always sends to \mathcal{A} the message m'_1 together with the prepared ciphertext tuple.

Due to the fact that for all \mathcal{B} , $\text{Prob}[\text{Exp}_{\mathcal{B}}^{KEM}(1^\lambda) = 1] \leq \frac{1}{2} + \varepsilon$, it can be proven that $p_0 - p_1 \leq 2n \cdot \varepsilon$. Similarly, we have that for all $i \in \{0, 1, \dots, n\}$, $p_i - p_{i+1} \leq 2n \cdot \varepsilon$. Summing these relations for both sides, we have that $p_0 - p_n \leq 2n^2 \cdot \varepsilon$. Because of $p_n = 1/2$, it holds that $\text{Prob}[\text{Exp}_0^A = 1] - 1/2 \leq 2n^2 \cdot \varepsilon$, which contradicts the initial assumption. ■

5 Lower Bounds for General Broadcast Encryption Schemes

We now turn our attention to the setting of general, unrestricted broadcast encryption schemes. We will prove that any scheme that is private in the sense of *priv-st*, *priv-full* has ciphertext length that with reasonably high probability is linear. We denote as $|x|$, the number of bits of the value x .

Theorem 6. *For all the sets $R \subseteq [n]$, we define the random variable*

$$S_R : \text{Encrypt}(ek, m, R) \rightarrow |\text{Encrypt}(ek, m, R)|,$$

where ek is an encryption key and m is a plaintext chosen from a message space M . Suppose that Φ is a broadcast encryption scheme with n receivers, and let R, R' be two sets. If Φ is private according to *priv-full* definition, then for all $R, R' \subseteq [n]$ and for all the PPT statistical tests D , it holds that $\Delta_D[S_R, S_{R'}] < \varepsilon$.

Proof. Suppose that there exists a pair of sets R, R' and a PPT statistical test D such that $\Delta_D[S_R, S_{R'}] \geq \alpha$, with α non-negligible. Then, a PPT adversary \mathcal{A} breaks definition *priv-full* with advantage at least $\alpha/2$ following the steps below.

Phase 1: Challenge R, R' .

Phase 2: On input $\langle m, \text{Encrypt}(ek, m, R_b) \rangle$:

- Compute $|\text{Encrypt}(ek, m, R_b)|$.
- Run D on input $|\text{Encrypt}(ek, m, R_b)|$.
- Return the output of D .

The adversary can execute the algorithm D a number of times in order to understand whether it is biased to 1 on input S_R or vice versa. Without loss of generality we assume that D returns 1 with greater probability in case of input $|\text{Encrypt}(ek, m, R')|$. As a result, we have that

$$\text{Prob}[D(S_{R'}) = 1] - \text{Prob}[D(S_R) = 1] \geq \alpha.$$

We note that if D is biased to 1 on input S_R we can consider the adversary $\overline{\mathcal{A}}$ in order to obtain the same results.

$$\begin{aligned} \text{Prob}[\text{Exp}_{\mathcal{A}}^{\text{priv-full}}(1^\lambda) = 1] &= \frac{1}{2} \left(\text{Prob}[\text{Exp}_{\mathcal{A}}^{\text{priv-full}} = 1 | b = 0] + \text{Prob}[\text{Exp}_{\mathcal{A}}^{\text{priv-full}} = 1 | b = 1] \right) \\ &= \frac{1}{2} \left(\text{Prob}[D(S_R) = 0] + \text{Prob}[D(S_{R'}) = 1] \right) \\ &\geq \frac{1}{2} + \frac{\alpha}{2}. \end{aligned}$$

■

Next, we will prove a lower bound on the ciphertext size that any private broadcast encryption scheme can achieve. Our proof is based on a standard information theoretic fact (cf. [5]), which is presented below:

Fact 1. *Suppose there is a randomized procedure $Enc : \{0, 1\}^n \times \{0, 1\}^r \rightarrow \{0, 1\}^m$ and a decoding procedure $Dec : \{0, 1\}^m \times \{0, 1\}^r \rightarrow \{0, 1\}^n$ such that*

$$\text{Prob}_{r \in U_r}[Dec(Enc(x, r), r) = x] \geq \delta.$$

Then, $m \geq n - \log \frac{1}{\delta}$.

Theorem 7. *Let Φ be a broadcast encryption scheme with n receivers and let $\varepsilon(\lambda)$ be the upper bound of all the probabilities $\text{Prob}[E_{R,i}]$. For any $R \subseteq [n]$ and $i \in [n]$, we denote as $E_{R,i}$ the event*

$$(\text{Decrypt}(SK_i, c) \neq m \wedge i \notin R) \vee (\text{Decrypt}(SK_i, c) = m \wedge i \in R),$$

where $c = \text{Encrypt}(ek, m, R)$. If for any λ there exists some β for which $\varepsilon(\lambda) < \frac{1}{2n} - \frac{\beta}{n}$, then there exists a set $R \subseteq [n]$ such that $\text{Prob}[S_R \geq n] > \beta$.

Proof. Recall the definition of S_R :

$$S_R : \text{Encrypt}(ek, m, R) \rightarrow |\text{Encrypt}(ek, m, R)|.$$

We define a procedure f which is an encoding procedure of a set $R \subseteq [n]$, while f^{-1} is a decoding procedure. The procedure f is a randomized procedure that takes as input two arguments $\rho \in \{0, 1\}^r$ and $R \subseteq [n]$ and outputs ψ . We note that ρ depends on the security parameter λ and represents all the coins needed in order for the system to setup and the encryption. The procedures f and f^{-1} are defined as follows:

$f(\rho, R)$:

1. Using ρ , compute a message m and the key ek which will be used by the encryption algorithm.
2. Compute $\text{Encrypt}(ek, m, R)$.
3. If $|\text{Encrypt}(ek, m, R)| \geq n$, output 0^{n-1} else $\text{Encrypt}(ek, m, R)$.

$f^{-1}(\psi, \rho)$:

1. Use ρ to compute SK_1, \dots, SK_n .
2. Execute the following algorithm:
 - $R := \emptyset$.
 - For $i = 1$ to n
 - if $\text{Decrypt}(SK_i, \psi) \neq m$ then $R := R \cup \{i\}$ else R .

Considering the definition of the decoding procedure, we say that f^{-1} fails when its result is $R' \neq R$, given that R is the encoded set. This happens either in case an event $E_{R,i}$ takes place or the output of f is 0^{n-1} . With δ we denote the probability that the procedure f^{-1} succeeds.

In order to prove the theorem, we assume that for any λ for which there exists a β such that $\varepsilon(\lambda) < \frac{1}{2n} - \frac{\beta}{n}$ it holds that $\text{Prob}[S_R \geq n] \leq \beta$ for all $R \subseteq [n]$. Let us fix a value λ . From the above assumption, we have that $\text{Prob}[f \text{ outputs } 0^{n-1}] \leq \beta$ which subsequently means that $\text{Prob}[f^{-1} \text{ fails}] \leq n \cdot \varepsilon(\lambda) + \beta$. Consequently, we have that $\delta \geq 1 - n \cdot \varepsilon(\lambda) - \beta$.

Due to the fact that the length of the encoding produced by f^{-1} is always $n - 1$ bits at most, using the fact 1, we have that

$$n - 1 \geq n - \log \frac{1}{\delta} \Rightarrow \varepsilon(\lambda) \geq \frac{1}{2n} - \frac{\beta}{n}, \tag{1}$$

which is a contradiction. ■

Lemma 2. *Let Φ be a private broadcast encryption scheme with n receivers and λ a security parameter for which $\beta < 1/2$ and β non-negligible in λ . Then, for all $R \subseteq [n]$, it holds that $\text{Prob}[S_R \geq n] \geq \alpha$, for α non-negligible.*

Proof. We assume that there exists a set R_0 such that $\text{Prob}[S_{R_0} \geq n] < \delta$, where δ is a negligible function of λ . We construct the following statistical test D :

D : On input S_R : If $S_R \geq n$ return 1 else return 0.

According to the Theorem 7, we have that there exists a set R_1 for which $\text{Prob}[S_{R_1} \geq n] > \beta$. As a result, we have that

$$\text{Prob}[D(S_{R_1}) = 1] - \text{Prob}[D(S_{R_0}) = 1] > \beta - \delta,$$

which is non-negligible. This contradicts to Theorem 6. ■

Corollary 3. *For any broadcast encryption scheme Φ which is private in the sense of definition priv-full,priv-st , the ciphertext is of length $\Omega(n + k)$.*

The additive factor k stems from the fact that at least one ciphertext should be present in the encryption of a message m for any enabled set S .

6 Conclusion

The provided lower bounds highlight the high costs that privacy may incur for broadcast encryption schemes. The fact that privacy for atomic schemes requires a linear number of ciphertexts in the number of users, leaves essentially no room for improvement in terms of the ciphertext size. If the objective is to attain full privacy, this result suggests that our attention should be turned to non-atomic schemes. In the non-atomic case, our lower bound is much weaker. It is thus an interesting open problem to design a fully private scheme with sublinear ciphertext size.

References

1. AACS, <http://www.aacsla.com/>
2. Attrapadung, N., Imai, H.: Graph-Decomposition-Based Frameworks for Subset-Cover Broadcast Encryption and Efficient Instantiations. In: Roy, B. (ed.) ASIACRYPT 2005. LNCS, vol. 3788, pp. 100–120. Springer, Heidelberg (2005)
3. Barth, A., Boneh, D., Waters, B.: Privacy in Encrypted Content Distribution Using Private Broadcast Encryption. In: Di Crescenzo, G., Rubin, A. (eds.) FC 2006. LNCS, vol. 4107, pp. 52–64. Springer, Heidelberg (2006)
4. Boneh, D., Gentry, C., Waters, B.: Collusion Resistant Broadcast Encryption with Short Ciphertexts and Private Keys. In: Shoup, V. (ed.) CRYPTO 2005. LNCS, vol. 3621, pp. 258–275. Springer, Heidelberg (2005)
5. De, A., Trevisan, L., Tulsiani, M.: Time Space Tradeoffs for Attacks against One-Way Functions and PRGs. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 649–665. Springer, Heidelberg (2010)
6. Delerablée, C.: Identity-Based Broadcast Encryption with Constant Size Ciphertexts and Private Keys. In: Kurosawa, K. (ed.) ASIACRYPT 2007. LNCS, vol. 4833, pp. 200–215. Springer, Heidelberg (2007)
7. Dodis, Y., Fazio, N.: Public Key Broadcast Encryption for Stateless Receivers. In: Feigenbaum, J. (ed.) DRM 2002. LNCS, vol. 2696, pp. 61–80. Springer, Heidelberg (2003)
8. Fazio, N., Perera, I.M.: Outsider-anonymous broadcast encryption with sublinear ciphertexts. In: Fischlin, et al. [10], pp. 225–242
9. Fiat, A., Naor, M.: Broadcast Encryption. In: Stinson, D.R. (ed.) CRYPTO 1993. LNCS, vol. 773, pp. 480–491. Springer, Heidelberg (1994)
10. Fischlin, M., Buchmann, J., Manulis, M. (eds.): PKC 2012. LNCS, vol. 7293, pp. 2012–2015. Springer, Heidelberg (2012)
11. Goodrich, M.T., Sun, J.Z., Tamassia, R.: Efficient Tree-Based Revocation in Groups of Low-State Devices. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 511–527. Springer, Heidelberg (2004)
12. Halevy, D., Shamir, A.: The LSD Broadcast Encryption Scheme. In: Yung, M. (ed.) CRYPTO 2002. LNCS, vol. 2442, pp. 47–60. Springer, Heidelberg (2002)
13. Libert, B., Paterson, K.G., Quaglia, E.A.: Anonymous broadcast encryption: Adaptive security and efficient constructions in the standard model. In: Fischlin, et al. [10], pp. 206–224
14. Naor, D., Naor, M., Lotspiech, J.: Revocation and Tracing Schemes for Stateless Receivers. In: Kilian, J. (ed.) CRYPTO 2001. LNCS, vol. 2139, pp. 41–62. Springer, Heidelberg (2001)
15. Shoup, V.: A proposal for an iso standard for public key encryption. IACR Cryptology ePrint Archive 2001, 112 (2001)
16. Wang, P., Ning, P., Reeves, D.S.: Storage-Efficient Stateless Group Key Revocation. In: Zhang, K., Zheng, Y. (eds.) ISC 2004. LNCS, vol. 3225, pp. 25–38. Springer, Heidelberg (2004)