Symmetric-Key Broadcast Encryption: The Multi-Sender Case*

Cody Freitag† Jonathan Katz‡ Nathan Klein§

Abstract

The problem of (stateless, symmetric-key) *broadcast encryption*, in which a central authority distributes keys to a set of receivers and can then send encrypted content that can be decrypted only by a designated subset of those receivers, has been the focus of a significant amount of attention. Here, we consider a generalization of this problem in which *all* members of the group have the ability to act as both sender and receiver. The parameters of interest are the number of keys stored per user and the bandwidth required per transmission, as a function of the total number of users n and the number of excluded/revoked users r.

As our main result, we show a multi-sender scheme allowing revocation of an arbitrary number of users in which users store $O(n)$ keys and the bandwidth is $O(r)$. We prove a matching lower bound on the storage, showing that for revoking an arbitrary number of users $\Omega(n)$ keys are necessary (regardless of the bandwidth) for *unique predecessor* schemes, a class of schemes capturing most known constructions in the single-sender case. Previous work has shown that $\Omega(r)$ bandwidth is needed when the number of keys per user is polynomial, even in the single-sender case; thus, our scheme is optimal in both storage and bandwidth.

We also show a scheme with storage $\text{polylog}(n)$ and bandwidth $O(r)$ that can be used to revoke any $\text{polylog}(n)$ users.

1 Introduction

In the classical setting of *broadcast encryption* [16], there is a group of n users to which a sender periodically transmits encrypted data. At times, the sender requires that only some designated subset S of the users should be able to decrypt the transmission and recover the original plaintext; the remaining users R—who should be unable to learn anything about the underlying plaintext, even if they all collude—are said to be *revoked* from that transmission. We are interested here in *symmetric-key* schemes that use no public-key operations, and which are also *stateless*, i.e., in which the keying material stored by each user remains fixed even as different subsets of users are revoked. This problem is motivated by applications to secure content distribution, but has applications to secure multicast communication more generally.

To the best of our knowledge, all previous considerations of broadcast encryption explicitly consider the case in which there is one, designated sender, and each of the n users acts only as a (potential) receiver. (A case that *has* been considered previously is the "point-to-point" setting in which each user should be able to communicate securely with every other user. See further

*This research was supported in part by the NSF REU-CAAR program, award #1262805.

†University of Texas, Austin. **Email:** cody.freitag@utexas.edu.

‡Dept. of Computer Science, University of Maryland. **Email:** jkatz@cs.umd.edu.

§Oberlin College. **Email:** nklein@oberlin.edu.

discussion in Section 1.1.) But in the setting of multicast communication it makes sense to assume that each of the n users might want to communicate with any subset of the others; that is, each of the n users might sometimes act as a sender and sometimes as a receiver. We refer to this as the *multi-sender* setting. Multi-sender broadcast encryption is applicable whenever there is some group of users all of whom wish to jointly communicate, yet from time-to-time some users' devices are compromised and so those users must be revoked. Or, users in the group may each have different access privileges, and so the set of revoked users for any particular transmission (being made by any one of the n users) may vary depending on the context. We initiate a study of multi-sender broadcast encryption in this paper.

As in the single-sender case, the main parameters of interest are the storage per user and the bandwidth overhead per transmission, as a function of the total number of parties n and the number of revoked users r. There is a trivial solution in which each user shares a key with every other user, and uses the appropriate keys to encrypt to any desired subset. This solution requires each user to store $n-1$ keys and has bandwidth $n-r-1$. The natural questions are whether it is possible to achieve storage and/or bandwidth sublinear in n. (We remark that traditionally $r \ll n$ is considered the interesting case, as it is assumed that the number of revoked users will be small in normal operation of the scheme.) As our main results, we show:

- There is a multi-sender scheme supporting revocation of arbitrarily many users, in which each user stores $O(n)$ keys and the bandwidth is $O(r)$. Moreover, we prove a lower bound (when revocation of arbitrarily many users must be supported) showing that $\Omega(n)$ storage is necessary, regardless of the bandwidth, for *unique predecessor* schemes [2], a class capturing all recent constructions in the single-sender setting [29, 20, 19].

  Austrin and Kreitz [2] have previously shown that the bandwidth must be $\Omega(r)$, even in the single-sender case, when polynomially many keys are used; thus, our scheme is asymptotically optimal in both storage and bandwidth.

- There is a multi-sender scheme supporting revocation of any set of $r \leq \text{polylog}(n)$ users, having storage $\text{polylog}(n)$ and bandwidth $O(r)$.

We refer to Section 1.2 for a more complete discussion of our results.

1.1 Prior Work

As noted earlier, to the best of our knowledge all prior work treating (symmetric-key) broadcast encryption focuses only on the case of a single sender. (Nevertheless, as we discuss below, some prior work is applicable to the multi-sender setting.) We briefly survey this body of work here, without intending to be exhaustive.

We remark that in some formulations of broadcast encryption, security is defined to hold with respect to all coalitions $R \subseteq [n] \setminus S$ containing at most r' users, for some bound r', rather than with respect to $R \stackrel{\text{def}}{=} [n] \setminus S$ as here. The former offers reduced security, but (potentially) allows for security/efficiency tradeoffs depending on the assumed number of colluding users. For simplicity, and following recent work in this area (e.g., [29, 20, 19, 18]), we assume $R = [n] \setminus S$ in the discussion below and throughout the paper.

Single-sender broadcast encryption. The work of Blundo et al. [7], which extends the work of Blom [6], can be used to construct a scheme in which a group of n users is given keying material

that allows any subset S' of size t to compute a shared key that is information-theoretically hidden from the $r = n - t$ other users. Their work implies a multi-sender broadcast encryption scheme with bandwidth 1: user i can transmit to a set of $n - r - 1$ other users S by encrypting with the key shared by users in $S' = S \cup \{i\}$. Unfortunately, the storage per user in their scheme is $\binom{n-1}{n-r-1}$, which they prove is optimal for their setting. Blundo et al. [8] also consider a more careful application of these ideas to the problem of broadcast encryption, trading off higher bandwidth for lower storage. For most interesting settings of the parameters, however, this work is subsumed by the schemes discussed below.

Fiat and Naor [16] introduce the term "broadcast encryption," and show a scheme with storage $O(r_{\max} \log r_{\max} \log n)$ and bandwidth $O(r_{\max}^2 \log^2 r_{\max} \log n)$, where $r_{\max}$ denotes a pre-determined upper bound on the size of r. Further improvements were given by [1, 17, 24, 25, 18]. Note that in all these schemes both the storage and the bandwidth depend on $r_{\max}$, so either $r_{\max}$ must be small or the parameters of the scheme are high even if only few users are actually revoked.[1]

Most recent work has focused on schemes that directly have the flexibility to communicate with arbitrary subsets of users while revoking all others. The following general approach can be used for constructing such schemes: Fix a set of keys K held by the sender. Each user i is given some subset $K_i \subset K$ of these keys. For the sender to securely send a message to a group S, it suffices if there is a set $K_S \subset K$ of keys such that (1) each user in S knows at least one key in K_S, and (2) no user in the revoked set $R = [n] \setminus S$ knows any of the keys in K_S. This implies a solution with bandwidth $|K_S|$ in which the sender encrypts the content independently using each of the keys in K_S, and each intended receiver decrypts the appropriate ciphertext using a key they know. Following [26], we refer to this as the *OR approach*. Naor, Naor, and Lotspiech [29] propose the *complete subtree (CS) scheme* that uses this approach, and has storage $\log n$ and bandwidth $r \log n/r$.

The schemes above have information-theoretic security. One can use *key derivation* to reduce the per-user storage, at the expense of achieving only computational security. When using key derivation, roughly speaking, users need not explicitly store all the keys they have access to; instead, they may derive one key from another, or derive multiple keys from a single predecessor, using a hash function (possibly modeled as a random oracle), a pseudorandom generator, or a pseudorandom function. Naor, Naor, and Lotspiech [29] present the *subset difference (SD) scheme* that uses the OR approach and key derivation, and achieves storage $O(\log^2 n)$ and bandwidth $O(r)$. This was improved in subsequent work [20, 5, 4], culminating in the *SSD scheme* of Goodrich et al. [19] that achieves storage $O(\log n)$ and bandwidth $O(r)$, though at the expense of requiring computation linear in n. (Hwang et al. [21] show how to improve the computation to $O(\log n)$ at the expense of a small increase in bandwidth.) Jho et al. [22] show a scheme with storage $O(c^p)$ and bandwidth $O(\frac{r}{p} + \frac{n-r}{c})$, where c, p are parameters; the scheme fares best (and beats [29, 20] in terms of both storage and bandwidth) when r is a large constant fraction of n. This scheme was further improved in [21], but even in that case either the bandwidth is $\Omega(\sqrt{n})$ when $r = O(1)$ or else the scheme requires storage $\Omega(\sqrt{n})$. Other relevant work in the single-sender case includes [12, 31].

Lower bounds for broadcast encryption schemes following the OR approach have been studied in both the information-theoretic setting [26, 29, 18] and when key derivation is used [2].

Secure point-to-point communication. Motivated by achieving secure point-to-point communication, Dyer et al. [14] (see also [28]) consider the setting in which each user i holds a subset

[1] Another possibility is to run $\log r_{\max}$ independent copies of the scheme using powers of two for the maximum size of the revoked set. This allows the bandwidth to depend on the actual number of revoked users r, though increases space by a factor of $\log r_{\max}$.

Security	Storage	Bandwidth	Scheme
info. theoretic	$O(r_{\max}^2 \log n)$	$O(r_{\max}^2 \log n)$	Follows from prior work [14]
info. theoretic	$n \log n$	$r \log \frac{n}{r}$	Result 1 applied to CS scheme [29]
info. theoretic	$O(n^{1+1/k})$	$O(kr)$	Result 1 applied to Scheme 1
info. theoretic	$O(r_{\max}^4 n^{1/2} \log n)$	$2\, r_{\max}$	Result 3 applied to [18]
computational	$O(n)$	$O(r)$	Result 2 applied to Scheme 1
computational	$O(r_{\max}^2 \log^2 n)$	$O(r)$	Result 3 applied to SSD scheme [19]

Table 1: Constructions of multi-sender broadcast encryption schemes. Scheme 1 is described in Appendix A.

$K_i \subset K$ of keys, and each pair of users i, j has a set of keys $K_{i,j} = K_i \cap K_j$ in common that are not *all* known to any set of at most $r_{\max}$ other users. Although Dyer et al. do not explicitly treat the case of multi-sender broadcast encryption, we observe that their scheme can be used to solve that problem if the number of revoked users is bounded: Consider a user i who wishes to securely transmit a message to some set S of users. Let $R = [n] \setminus (S \cup \{i\})$, where $|R| \leq r_{\max}$. Then i and each user in S must share at least one key not known to any user in R; user i can encrypt its content using all such keys. This results in a multi-sender broadcast encryption scheme with[2] storage and bandwidth $O(r_{\max}^2 \log n)$.

Other related work. The case of *stateful* broadcast encryption has also received extensive attention (for the single-sender case), in terms of both constructions [30, 32, 3, 11, 10] and lower bounds [11, 27]. Here, some set of authorized users S is continually maintained by the sender; the authorized users always share a single key under which the sender encrypts its communication. From time to time, the sender *revokes* a user i, thus changing the set of authorized users. When this happens, the sender transmits rekeying information that allows all users in $S \setminus \{i\}$ to both compute a new, shared key as well as to update their individual keying material.

Broadcast encryption has also been studied in the *public-key* setting [13, 9]. Of course, in this setting every receiver can trivially also act as a sender. Public-key schemes inherently require stronger assumptions than symmetric-key schemes, and generally incur higher computational costs.

1.2 Our Results

We show general transformations from single-sender broadcast encryption (BE) schemes to multi-sender ones. Fix a single-sender BE scheme Π with storage s, bandwidth b, and where the sender stores s^* keys.[3] We show:

Result 1: There is a multi-sender BE scheme with storage $(n-1) \cdot s + s^*$ and bandwidth b that supports the same number of revoked users as Π does; if Π is information-theoretic then so is the derived scheme.

[2]These parameters are not stated explicitly by Dyer et al., who report only the total number of keys. However, Corollary 1 and the proof of Theorem 4 in their paper show that the per-user storage is $O(r_{\max}^2 \log n)$; the bandwidth is bounded by the number of keys held by any user acting as a sender.

[3]When computational security suffices, any single-sender BE scheme can be modified to have $s^* = 1$ by having the sender use a PRF to derive all the keys in the system. In the information-theoretic setting that is not the case.

Result 2: If Π is information-theoretic, there is a multi-sender BE scheme with storage s^* and bandwidth b that supports the same number of revoked users as Π does. The scheme uses key derivation, and so is no longer information-theoretically secure.

Result 3: For any bound $r_{\max}$ on the number of revoked users, there is a multi-sender BE scheme with storage $O((s \cdot r_{\max}^2 + s^* \cdot r_{\max}) \cdot \log n)$ and bandwidth b; moreover, if Π is information-theoretic then so is the derived scheme.

Applying the above to known single-sender schemes (see also Appendix A) gives the results in Table 1. Particularly interesting in practice, where computational security suffices, are:

- A scheme supporting revocation of arbitrarily many users, where each user stores $O(n)$ keys and the bandwidth is $O(r)$. (Although the storage may seem high, we prove that it is optimal for schemes of a certain class allowing arbitrary revoked sets.)
- A scheme with a pre-determined bound $r_{\max}$ on the number of revoked users that has storage $O(r_{\max}^2 \log^2 n)$ and bandwidth $O(r)$.

As noted earlier, we also prove a lower bound on the key storage for multi-sender BE schemes that support revocation of an arbitrary number of users, and that are constructed in a certain way. Specifically, we focus on so-called *unique predecessor* schemes [2], which are schemes that follow the OR approach and in which keys are derived from secret values by applying a hash function (possibly modeled as a random oracle), a pseudorandom generator, or a pseudorandom function to those values individually. To the best of our knowledge, this class includes all known computationally secure, single-sender schemes that improve on information-theoretic schemes, and lower bounds for unique predecessor schemes (in the single-sender case) were previously studied by Austrin and Kreitz [2]. Our bound shows that, in the multi-sender setting, any such scheme requires at least one user to store at least $\frac{n-1}{2}$ keys. Interestingly, we also show that this bound is tight, as there is a multi-sender BE scheme in which all users hold this many keys. (The bandwidth in this scheme is $n - r$, which is why we do not include it in Table 1.)

Austrin and Kreitz [2] also show that any (unique predecessor) single-sender BE scheme with polynomially many keys per user has bandwidth $\Omega(r)$ for small r, showing that our computationally secure scheme supporting unbounded revocation is asymptotically optimal in terms of both storage and bandwidth.

2 Definition of the Problem

We consider multi-sender broadcast encryption schemes, in which there is a set of users $[n] = \{1, \ldots, n\}$, each of whom is given some keying material by a trusted authority. Subsequently, each user $i \in [n]$ should be able to send a message to any desired subset of users $S \subseteq [n] \setminus \{i\}$ such that the revoked users $R = [n] \setminus (S \cup \{i\})$ cannot recover the message even if they all collude. We let $r = |R|$ denote the number of revoked users. Some schemes support revocation of an arbitrary number of users, whereas others impose an *a priori* bound $r_{\max}$ on the number of users who can be revoked. We consider two classes of schemes—information-theoretic and computational—both following the OR approach described in the previous section.

Information-theoretic schemes. In the information-theoretic schemes we consider, there is a set K of keys chosen uniformly and independently from some key space. Each user i is assigned a

set $K_i \subset K$, and the *per-user storage* of the scheme is defined as $\max_i |K_i|$. When user i wishes to send a message to a subset S, it finds the smallest $K_{i,S} \subseteq K_i$ such that (1) each user $j \in S$ holds at least one of the keys in $K_{i,S}$ (i.e., $K_j \cap K_{i,S} \neq \emptyset$ for $j \in S$) and (2) no user $j \in R$ holds any of the keys in $K_{i,S}$ (i.e., $K_j \cap K_{i,S} = \emptyset$ for $j \in R$). (For schemes supporting a bounded number $r_{\max}$ of revoked users, such a $K_{i,S}$ is only required to exist if $n - |S| - 1 \leq r_{\max}$.) User i can then encrypt its message using[4] each key in $K_{i,S}$. The *bandwidth* of the scheme for a given number of revoked users r is defined to be $\max_{i,S:n-|S|-1\leq r} |K_{i,S}|$.

Computationally secure schemes. We consider *unique predecessor* schemes, as defined by Austrin and Kreitz [2], that can offer reduced storage but only achieve computational security. In such schemes, we have sets K, K_i, and $K_{i,S}$ satisfying the same conditions as above, and the bandwidth is defined in the same way. Now, however, users need not store their keys explicitly. Instead, they may derive their keys from secret values they store, with the canonical example of this being the use of a single secret value v to derive keys $k_1 = F_v(1), \ldots, k_\ell = F_v(\ell)$ for F a pseudorandom function. Following [2], we model this by a set $V \supseteq K$ of "secret values" along with a directed graph G (a *key-derivation graph*) whose nodes are in one-to-one correspondence with the elements of V and such that each node has in-degree 0 or 1 (hence the name "unique predecessor"). To instantiate such a scheme, a uniform value is chosen for each node with in-degree 0; then, for every $v' \in V$ that is the ℓth child of some node $v \in V$, we set the value of v' equal to $F_v(\ell)$.

Nodes labeled with elements of K are called "keys"; we say $k \in K$ can be derived from $v \in V$ if v is an ancestor of k in the graph G (this includes the case $v = k$). Each user i is now given a subset $V_i \subset V$ of secret values, and we define K_i to be the set of keys that can be derived from V_i. The per-user storage is now $\max_i |V_i|$.

We remark that the information-theoretic setting is a special case of the above, where $V = K$ and all nodes have in-degree 0.

3 Constructions

We first consider two transformations of single-sender BE schemes to multi-sender BE schemes that are applicable for any number of revoked users. Then, we look at the special case where there is an *a priori* bound $r_{\max}$ on the number of users to be revoked.

3.1 A Trivial Construction

Let Π be a single-sender BE scheme for n users. The construction described here applies regardless of how Π works, but for simplicity we assume Π is a unique predecessor scheme as defined in Section 2 (adapted appropriately for the single-sender case). Thus, we let $\bar{V}$ denote the set of secret values in Π, let $\bar{V}_i \subset \bar{V}$ denote the values given to user i, and let $\bar{V}_0 \subset \bar{V}$ denote the values with in-degree 0 (these are the only values the sender needs to store). We construct a multi-sender scheme for n users by simply running Π in parallel n times, with each user acting as the sender in an instance of Π. Our set of secret values will be $V = [n] \times \bar{V}$, and user i will be given

$$V_i = \left\{(i, v) \mid v \in \bar{V}_0\right\} \cup \left\{(j, v) \mid j \neq i,\ v \in \bar{V}_i\right\};$$

that is, user i will be given the values that the sender would store in the ith instance of Π, and the values that user i would store (as a receiver) in all other instances of Π. For a user i to send a

[4]For long messages, user i can encrypt the message using a fresh key k and encrypt k using each key in $K_{i,S}$.

message to a designated subset S, that user will simply act as the sender would in the ith instance of Π when sending to S.

It is easy to see that this multi-sender scheme is secure if Π is. Consider any sender i and designated subset of receivers S. Since only the ith instance of Π will be used, we can focus our attention on values of the form $\{(i, v)\}_{v \in \bar{V}}$. But then security of Π implies that even if all the users in R collude, they will not be able to decrypt the message sent by user i. We thus have:

Theorem 1. *Let Π be a single-sender BE scheme with s^* sender storage, per-user storage s, and bandwidth b. Then the multi-sender BE scheme described above supports the same number of revoked users as Π does, and has per-user storage $(n-1) \cdot s + s^*$ and the same bandwidth as Π. Moreover, if Π is information-theoretic then so is the derived scheme.*

3.2 An Improved Construction

We now give an improved construction that uses key derivation applied to an information-theoretic, single-sender scheme Π. Let $\bar{K}$ denote the set of keys used by Π, and let $\bar{K}_i \subset \bar{K}$ denote the keys stored by user i in that scheme. Conceptually, in our multi-sender scheme we will again have n instances of Π, with each user acting as a sender in one of the schemes. Now, however, the keys in the various schemes will be *correlated*. Specifically, the keys used in the ith instance of Π will be $K^{(i)} = \{F_{\bar{k}}(i) \mid \bar{k} \in \bar{K}\}$. In our multi-sender scheme, each user i is given all the keys that the sender would store in the ith instance of Π (namely, $K^{(i)}$), as well as the values $\bar{K}_i \subset \bar{K}$ that can be used to derive the keys that user i would store (as a receiver) in all other instances of Π. Note that user i need not store $F_{\bar{k}}(i)$ for $\bar{k} \in \bar{K}_i$; hence the storage of user i is exactly $|\bar{K}|$.

More formally, we now have a set of keys $K = \{k_{i,j} \mid i \in [n],\ j \in \bar{K}\}$ and additional values $V_0 = \{k_{0,j} \mid j \in \bar{K}\}$; define $V = K \cup V_0$. The keys satisfy $k_{i,j} = F_{k_{0,j}}(i)$; in terms of the underlying key-derivation graph, all nodes corresponding to V_0 have in-degree 0, and node $k_{0,j}$ is a parent of all nodes of the form $k_{i,j}$. User i is given

$$K_i = \left\{k_{0,j} \mid j \in \bar{K}_i\right\} \cup \left\{k_{i,j} \mid j \in \bar{K}\right\}.$$

(We can also use the optimization mentioned above to reduce the storage slightly.) If we let $K^{(i)} = \{k_{i,j} \mid j \in \bar{K}\}$ and $K^{(i)}_\ell = \{k_{i,j} \mid j \in \bar{K}_\ell\} \subset K^{(i)}$, then the key observations are: (1) for each i, the sets $K^{(i)}, K^{(i)}_1, \ldots, K^{(i)}_n$ correspond to $\bar{K}, \bar{K}_1, \ldots, \bar{K}_n$, and we thus have n instances of Π; moreover, (2) user i can derive both $K^{(i)}$ as well as $K^{(j)}_i$ for all j. Put differently, user i can act as a sender in the ith instance of Π, and as a receiver in any other instance of Π. Thus, for a user i to send a message to some designated subset S, that user simply acts as the sender using keys $K^{(i)}$; each receiver $j \in S$ derives the keys $K^{(i)}_j$ and uses those to decrypt.

Security follows in a straightforward manner based on security of Π and the assumption that F is a pseudorandom function. We thus have:

Theorem 2. *Let Π be an information-theoretic, single-sender BE scheme with s^* total keys, per-user storage s, and bandwidth b. Then the multi-sender BE scheme described above is computationally secure, supports the same number of revoked users as Π does, and has per-user storage s^* and the same bandwidth as Π.*

3.3 A Construction Supporting Bounded Revocation

In this section we explore an approach for constructing multi-sender BE schemes supporting a bounded number of revoked users. Our construction uses the notion of r-cover-free families [23, 15]:

Definition 3. *Fix a universe K, and $1 \le r < n$. A family of sets $\mathcal{F} = \{K_1, \ldots, K_n\}$ with $K_i \subset K$ is* **r-cover free** *if $K_j \not\subseteq K_{i_1} \cup \cdots \cup K_{i_r}$ for any distinct $j, i_1, \ldots, i_r \in [n]$.*

Kumar et al. [24] show, for any r, n, an explicit construct of an r-cover-free family of size n with $|K| \le 16r^2 \log n$ and $|K_i| \le 4r \log n$ for all i. We remark that $r_{\max}$-cover-free families immediately imply single-sender broadcast encryption schemes supporting up to $r_{\max}$ revoked users. In general, however, the bandwidth of the resulting construction may be high.

We now show how to use an $r_{\max}$-cover-free family in conjunction with any single-sender broadcast encryption scheme Π supporting up to $r_{\max}$ revoked users to construct a *multi-sender* scheme supporting up to $r_{\max}$ revoked users.

Fix some $r_{\max}$, and let $\{T_1, \ldots, T_n\}$ be an $r_{\max}$-cover-free family over a set T of size t. The construction described below applies regardless of how Π works, but for simplicity we assume Π is a unique predecessor scheme as defined in Section 2 (adapted appropriately for the single-sender case). Thus, we let $\bar{V}$ denote the set of secret values in Π, let $\bar{V}_i \subset \bar{V}$ denote the values given to user i, and let $\bar{V}_0 \subset \bar{V}$ denote the values with in-degree 0 (these are the values the sender stores).

Our construction of a multi-sender scheme works by generating t independent instances of Π, and giving each user i (1) the values that the sender would store in the jth instance of Π, for all $j \in T_i$, and (2) the values that user i would store in all instances of Π. That is, our set of values is now $V = T \times \bar{V}$, and user i is given

$$V_i = \left\{(j, v) \mid j \in T_i,\ v \in \bar{V}_0\right\} \cup \left\{(i, v) \mid i \in T,\ v \in \bar{V}_i\right\}.$$

Say user i wants to send a message to some designated subset S, where $R = [n] \setminus (S \cup \{i\})$ has size at most $r_{\max}$. User i first finds an $i^* \in T_i$ such that $i^* \notin \bigcup_{j \in R} T_j$; such an i^* exists by the properties of the cover-free family. It then acts as the sender in instance i^* of Π, revoking the users in R. Security follows since Π is secure for at most $r_{\max}$ revoked users. Using [24], we thus have:

Theorem 4. *Let Π be a single-sender BE scheme supporting up to $r_{\max}$ revoked users, and having s^* sender storage, per-user storage s, and bandwidth $b(r)$ when revoking $r \le r_{\max}$ users. Then the multi-sender BE scheme described above supports up to $r_{\max}$ revoked users, and has per-user storage $O(s^*\, r_{\max}\, \log n + s\, r_{\max}^2\, \log n)$ and the same bandwidth as Π. If Π is information-theoretic then so is the derived scheme.*

4 Lower Bounds on Per-User Storage

In this section we consider bounds on the per-user storage s for multi-sender broadcast encryption schemes. We first observe a storage/communication tradeoff for information-theoretic schemes. Say there is a scheme with per-user storage s and bandwidth b when revoking r users. Consider some sender i storing keys K_i with $|K_i| = s$. There are $\binom{n-1}{r}$ different authorized subsets $S \subset [n]/\{i\}$ that exclude r users, and for each one the set of keys $K_{i,S} \subseteq K_i$ used by user i to encrypt must be different and non-empty. Moreover, $|K_{i,S}| \le b$ for all S. Thus, we must have

$$\sum_{j=1}^{b} \binom{s}{j} \ge \binom{n-1}{r}.$$

Simplifying, this gives $s \geq \binom{n-1}{r}^{1/b}$. If r is small, the above gives (asymptotically) $b \geq r\frac{\log(n-1)}{\log s}$. The most relevant consequence is that if r is constant, and the per-user storage is polylogarithmic in n, then $b = \omega(r)$. It is interesting to note (cf. Table 1) that when r is constant there is a computationally secure scheme with polylogarithmic storage and $b = O(r)$.

Can the storage be improved in computationally secure schemes? Unfortunately, the following theorem shows that any unique predecessor scheme supporting arbitrarily many revoked users must have per-user storage $\Omega(n)$.

Theorem 5. *Any unique predecessor scheme for n-user, multi-sender broadcast encryption supporting arbitrarily many revoked users has per-user storage at least $\lceil \frac{n-1}{2} \rceil$.*

Proof. The ability to revoke $r = n - 2$ users implies that each pair of distinct users i, j must be able to derive a shared key $k_{\{i,j\}}$ that cannot be derived by any other user. Call this the *pairwise key* for i and j. We claim that for each such i, j, either user i or user j (or possibly both) explicitly stores a value $v_{\{i,j\}}$ such that the *only* pairwise key that can be derived from $v_{\{i,j\}}$ is $k_{\{i,j\}}$. This implies that there are at least $\binom{n}{2}$ values $v_{\{i,j\}}$ that are stored overall, and hence some user must store at least $\binom{n}{2}/n = \frac{n-1}{2}$ values.

To prove the claim, let v_i (resp., v_j) denote the stored value used by user i (resp., user j) to derive $k_{\{i,j\}}$. Assume toward a contradiction that user i derives some other pairwise key (say, $k_{\{i,j'\}}$ with $j' \neq j$) from v_i, and that user j derives some other pairwise key (say, $k_{\{i',j\}}$ with $i' \neq i$) from v_j. (Security implies that user i cannot derive the pairwise key $k_{\{i',j'\}}$ if $i \notin \{i', j'\}$, and similarly for user j.) The unique predecessor property implies that v_i and v_j must lie on the same path in the underlying graph, and hence one must be an ancestor of the other. Without loss of generality, say v_i is an ancestor of v_j. But then user i can derive v_j and hence $k_{\{i',j\}}$, violating security. □

This lower bound is essentially tight, as we now show an n-user scheme in which each user stores exactly $\lceil \frac{n-1}{2} \rceil + 1$ values. For notational convenience, define $H(x) = F_x(0)$ where F is a pseudorandom function, and let $H^{(i)}(\cdot)$ denote the i-fold iteration of H. Number the users from 0 to $n - 1$. Each user i stores:

- A value v_i. Define $k_{\{i,j\}} = H^{(j)}(v_i)$ for $j = i + 1, \ldots, i + \lfloor \frac{n-1}{2} \rfloor$ (taken modulo n).
- Keys $k_{i,j}$ for $j = i + \lfloor \frac{n+1}{2} \rfloor, \ldots, i + n - 1$ (taken modulo n).

Each user stores $1 + \left(n - \lfloor \frac{n+1}{2} \rfloor\right) = \lceil \frac{n-1}{2} \rceil + 1$ values. Note that each key $k_{\{i,j\}}$ can be derived only by users i and j. Any user i can thus securely send a message to any designated subset S by using the set of keys $\{k_{\{i,j\}} \mid j \in S\}$.

As described, key derivation requires $O(n)$ invocations of H. Using a tree-based construction, however, this can be improved to $O(\log n)$.

5 Conclusion

We have introduced the problem of *multi-sender* broadcast encryption, a natural generalization of symmetric-key broadcast encryption, and explored upper- and lower bounds on such schemes.

The most pressing question is whether or not there exists a computationally secure scheme with storage $o(n)$ using an altogether different paradigm. It would also be interesting to find an information-theoretic scheme with storage $O(n)$ and bandwidth better than the trivial $n - r - 1$, or to show that to do asymptotically better is not possible.

6 Acknowledgments

We thank Bill Gasarch for organizing the University of Maryland Combinatorial Algorithms Applied Research (CAAR) REU program. We thank Daniel Apon, Seung Geol Choi, Jordan Schneider, and Arkady Yerukhimovich for discussing various aspects of this problem with us.

References

[1] William Aiello, Sachin Lodha, and Rafail Ostrovsky. Fast digital identity revocation. In *Advances in Cryptology—Crypto '98*, volume 1462 of *LNCS*, pages 137–152. Springer, 1998.

[2] Per Austrin and Gunnar Kreitz. Lower bounds for subset cover based broadcast encryption. In *Progress in Cryptology—Africacrypt 2008*, volume 5023 of *LNCS*, pages 343–356. Springer, 2008.

[3] David Balenson, David McGrew, and Alan Sherman. Key management for large dynamic groups: One-way function trees and amortized initialization. Internet Draft, 1999.

[4] Sanjat Bhattacharjee and Palash Sarkar. Reducing communication overhead of the subset difference scheme. To appear.

[5] Sanjay Bhattacherjee and Palash Sarkar. Concrete analysis and trade-offs for the (complete tree) layered subset difference broadcast encryption scheme. *IEEE Trans. on Computers*, 63(7):1709–1722, 2014.

[6] Rolf Blom. An optimal class of symmetric key generation systems. In *Advances in Cryptology—Eurocrypt '84*, volume 209 of *LNCS*, pages 335–338. Springer, 1985.

[7] Carlo Blundo, Alfredo De Santis, Amir Herzberg, Shay Kutten, Ugo Vaccaro, and Moti Yung. Perfectly secure key distribution for dynamic conferences. *Information and Computation*, 146(1):1–23, 1998.

[8] Carlo Blundo, Luiz A. Frota Mattos, and Douglas R. Stinson. Trade-offs between communication and storage in unconditionally secure schemes for broadcast encryption and interactive key distribution. In *Advances in Cryptology—Crypto '96*, volume 1109 of *LNCS*, pages 387–400. Springer, 1996.

[9] Dan Boneh, Craig Gentry, and Brent Waters. Collusion resistant broadcast encryption with short ciphertexts and private keys. In *Advances in Cryptology—Crypto 2005*, volume 3621 of *LNCS*, pages 258–275. Springer, 2005.

[10] Ran Canetti, Juan A. Garay, Gene Itkis, Daniele Micciancio, Moni Naor, and Benny Pinkas. Multicast security: A taxonomy and some efficient constructions. In *IEEE INFOCOM*, pages 708–716, 1999.

[11] Ran Canetti, Tal Malkin, and Kobbi Nissim. Efficient communication-storage tradeoffs for multicast encryption. In *Advances in Cryptology—Eurocrypt '99*, volume 1592 of *LNCS*, pages 459–474. Springer, 1999.

[12] Jung Hee Cheon, Nam-Su Jho, Myung-Hwan Kim, and Eun Sun Yoo. Skipping, cascade, and combined chain schemes for broadcast encryption. *IEEE Trans. Information Theory*, 54(11):5155–5171, 2008.

[13] Yevgeniy Dodis and Nelly Fazio. Public-key broadcast encryption for stateless receivers. In *Security and Privacy in Digital Rights Management (ACM CCS Workshop)*, pages 61–80. ACM, 2002.

[14] Martin Dyer, Trevor Fenner, Alan Frieze, and Andrew Thomason. On key storage in secure networks. *Journal of Cryptology*, 8(4):189–200, 1995.

[15] P. Erdös, P. Frankl, and Z. Füredi. Families of finite sets in which no set is covered by the union of r others. *Israeli Journal of Mathematics*, 51(1–2):79–89, 1985.

[16] Amos Fiat and Moni Naor. Broadcast encryption. In *Advances in Cryptology—Crypto '93*, volume 773 of *LNCS*, pages 480–491. Springer, 1994.

[17] Eli Gafni, Jessica Staddon, and Yiqun Lisa Yin. Efficient methods for integrating traceability and broadcast encryption. In *Advances in Cryptology—Crypto '99*, volume 1666 of *LNCS*, pages 372–387. Springer, 1999.

[18] Craig Gentry, Zulfikar Ramzan, and David P. Woodruff. Explicit exclusive set systems with applications to broadcast encryption. In *47th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 27–38. IEEE, 2006.

[19] Michael T. Goodrich, Jonathan Z. Sun, and Roberto Tamassia. Efficient tree-based revocation in groups of low-state devices. In *Advances in Cryptology—Crypto 2004*, volume 3152 of *LNCS*, pages 511–527. Springer, 2004.

[20] Dani Halevy and Adi Shamir. The LSD broadcast encryption scheme. In *Advances in Cryptology—Crypto 2002*, volume 2442 of *LNCS*, pages 47–60. Springer, 2002.

[21] Jung Yeon Hwang, Dong Hoon Lee, and Jongin Lim. Generic transformation for scalable broadcast encryption schemes. In *Advances in Cryptology—Crypto 2005*, volume 3621 of *LNCS*, pages 276–292. Springer, 2005.

[22] Nam-Su Jho, Jung Yeon Hwang, Jung Hee Cheon, Myung-Hwan Kim, Dong Hoon Lee, and Eun Sun Yoo. One-way chain based broadcast encryption schemes. In *Advances in Cryptology—Eurocrypt 2005*, volume 3494 of *LNCS*, pages 559–574. Springer, 2005.

[23] W.H. Kautz and R.C. Singleton. Nonrandom binary superimposed codes. *IEEE Trans. Information Theory*, 10(4):363–377, 1964.

[24] Ravi Kumar, Sridhar Rajagopalan, and Amit Sahai. Coding constructions for blacklisting problems without computational assumptions. In *Advances in Cryptology—Crypto '99*, volume 1666 of *LNCS*, pages 609–623. Springer, 1999.

[25] Ravi Kumar and Alex Russell. A note on the set systems used for broadcast encryption. In *14th Annual Symposium on Discrete Algorithms (SODA)*, pages 470–471. ACM-SIAM, 2003.

[26] Michael Luby and Jessica Staddon. Combinatorial bounds for broadcast encryption. In *Advances in Cryptology—Eurocrypt '98*, volume 1403 of *LNCS*, pages 512–526. Springer, 1998.

[27] Daniele Micciancio and Saurabh Panjwani. Optimal communication complexity of generic multicast key distribution. In *Advances in Cryptology—Eurocrypt 2004*, volume 3027 of *LNCS*, pages 153–170. Springer, 2004.

[28] Chris J. Mitchell and Fred C. Piper. Key storage in secure networks. *Discrete Applied Mathematics*, 21(3):215–228, 1988.

[29] Dalit Naor, Moni Naor, and Jeffery Lotspiech. Revocation and tracing schemes for stateless receivers. In *Advances in Cryptology—Crypto 2001*, volume 2139 of *LNCS*, pages 41–62. Springer, 2001.

[30] Debby M. Wallner, Eric J. Harder, and Ryan C. Agee. Key management for multicast: Issues and architectures. Internet Draft, RFC 2627, 1999.

[31] Shyh-Yih Wang, Wu-Chuan Yang, and Ying-Jen Lin. Balanced double subset difference broadcast encryption scheme. *Security and Communication Networks*, 8(8):1447–1460, 2015.

[32] Chung Kei Wong, Mohamed Gouda, and Simon S. Lam. Secure group communications using key graphs. In *Proceedings of ACM SIGCOMM*, pages 68–79, 1998.

A Information-Theoretic Single-Sender Schemes

In this section we describe various single-sender schemes that, to the best of our knowledge, have not appeared previously in the literature. The parameters of the schemes presented here do not beat the parameters of the best known single-sender schemes, but they have the advantage of having information-theoretic security.

We begin with a simple scheme that revokes exactly one user (i.e., $r_{\max} = 1$). Fix some b, and identify the n users with b-tuples whose coordinates range from 1 to $n^{1/b}$. The sender holds a set of keys $K = \{k_{i,w}\}_{i\in[b],\, w\in[n^{1/b}]}$ of size $b \cdot n^{1/b}$. The user associated with tuple $(w_1, \ldots, w_b)$ is given the set of keys $\{k_{i,w}\}_{i\in[b],\, w\neq w_i}$; in other words, key $k_{i,w}$ is held by all users whose ith coordinate is not w. To revoke the single user $(w_1, \ldots, w_b)$, the sender encrypts the message using the b keys $k_{1,w_1}, \ldots, k_{b,w_b}$ not held by that user. It follows that:

Theorem 6. *For any b, there is an information-theoretic, single-sender BE scheme with $r_{\max} = 1$ having per-user storage $b \cdot n^{1/b} - b$, bandwidth b, and $b \cdot n^{1/b}$ total keys.*

Gentry et al. [18] show that in any information-theoretic, single-sender scheme with $r_{\max} = 1$, storage s, and bandwidth b, it holds that $n \leq s^b$. The above scheme shows this bound is tight within a constant factor.

We now show how to build an information-theoretic scheme Π^* revoking any number of users based on any scheme Π revoking a single user. The high-level idea is to apply the SD approach [29] but to schemes rather than keys. In the SD approach, users are arranged at the leaves of a binary tree, and for each pair of nodes i, j in the tree with i a parent of j, we let $S_{i,j}$ denote the users who are descendants of i but not descendants of j. Naor et al. show that any set of users S can be partitioned into $O(r)$ such sets, where $r = n - |S|$ is the number of revoked users. In the SD

scheme, for all i, j as above there is a single key $k_{i,j}$ that is known exactly to those users in $S_{i,j}$; hence, the bandwidth of the scheme is $O(r)$. Here, we generalize the approach so that there is a *set* of keys allowing only those users in $S_{i,j}$ to decrypt.

We again arrange the users at the leaves of a binary tree. In this tree, let T_i denote the sub-tree rooted at some node i. For each such sub-tree T_i of height h, we associate the root node i of that sub-tree with h instances of Π (recall, Π is a single-sender scheme supporting revocation of a single user) corresponding to the h levels of T_i not including the root node itself. The "virtual users" of instance $\ell \in \{0, \ldots, h-1\}$ of Π correspond to the nodes at height ℓ in T_i, and we imagine giving each node the keys it would receive as a virtual user in all instances of Π in which it is involved. The real users, at the leaves, store the keys that would be given to its ancestors.

To send a message to a subset S of the users, the sender partitions S into a collection of subsets $S_{i,j}$ as in the SD scheme. To encrypt a message such that only the users in $S_{i,j}$ can read it, the sender uses the instance of Π in which node i is the sender and the nodes on the same level as j are the receivers, and revokes user j.

Rather than analyzing the above in the general case, we compute the bandwidth and storage when applied to the single-sender scheme Π from Theorem 6. Naor et al. showed that any set of S users can be partitioned into at most $2r-1$ subsets $S_{i,j}$, where $r = n - |S|$ is the number of revoked users. Since the scheme Π from Theorem 6 has fixed bandwidth b independent of the number of users, we conclude that the bandwidth of our scheme here is at most $b \cdot (2r-1)$. The storage per user is given by $\sum_{h=1}^{\log n} \sum_{\ell=0}^{h-1} (n/2^{h-\ell})^{1/b} = O(n^{1/b})$. Similarly, one can show that the total number of keys is $O(n)$. Summarizing:

Theorem 7 (Scheme 1)**.** *For any b, there is an information-theoretic, single-sender BE scheme supporting arbitrarily many revoked users having per-user storage $O(n^{1/b})$, bandwidth $O(b \cdot r)$, and $O(n)$ total keys.*

Specifically, there is an information-theoretic, single-sender BE scheme supporting arbitrarily many revoked users having per-user storage $O(\sqrt{n})$, bandwidth $O(r)$, and $O(n)$ total keys.