Combinatorial Bounds for Broadcast Encryption

Michael Luby[1]* and Jessica Staddon[2]**

[1] International Computer Science Institute,
1947 Center St., Suite 600,
Berkeley, CA, 94704-1198.
E-mail: luby@icsi.berkeley.edu
[2] RSA Laboratories,
100 Marine Parkway, Suite 500,
Redwood City, CA, 94065-1031.
E-mail: jstaddon@rsa.com

Abstract. A broadcast encryption system allows a center to communicate securely over a broadcast channel with selected sets of users. Each time the set of privileged users changes, the center enacts a protocol to establish a new broadcast key that only the privileged users can obtain, and subsequent transmissions by the center are encrypted using the new broadcast key. We study the inherent trade-off between the number of establishment keys held by each user and the number of transmissions needed to establish a new broadcast key. For every given upper bound on the number of establishment keys held by each user, we prove a lower bound on the number of transmissions needed to establish a new broadcast key. We show that these bounds are essentially tight, by describing broadcast encryption systems that come close to these bounds.

1 Introduction

Broadcast encryption addresses the problem of the allocation of secret keys to users in order to enable a center to broadcast to selected subsets of users with security. This is an important problem in the larger area of network security, and it has increased in prominence with the growth of the pay-television industry.

Our model is a formalization of that of Fiat and Naor [7]. Each user initially holds a personalized subset of all possible establishment keys. Each time the center needs to establish a new broadcast key it enacts an establishment protocol. This protocol consists of a sequence of transmissions, each transmission is encrypted using a different establishment key. A transmission can only be decrypted by users who have the corresponding establishment key in their personalized set. The broadcast encryption system should be designed so that only privileged users are able to compute the new broadcast key when the protocol ends. Subsequent transmissions by the center are encrypted using the newly established broadcast key.

* The author's research was supported in part by the National Science Foundation operating grants CCR-9304722 and NCR-9416101.

** This author's research was done while a graduate student in the Mathematics Department, University of California at Berkeley.

As an example, consider the simple broadcast encryption system in which each user has a unique establishment key. To establish a new broadcast key, select a random key B and send one transmission for each of the privileged users, encrypting B with the establishment key of the user. This protocol only requires a small amount of storage as each user holds just one establishment key. However, it requires a large amount of communication because the number of transmissions is equal to the number of users in the privileged set. At the opposite end of the spectrum, consider the broadcast encryption system that assigns to each set of users a unique establishment key, and each user holds the keys for all sets in which it is a member. To establish a new broadcast key, select a random key B and send one transmission encrypting B using the establishment key associated with the set of privileged users. This system only requires one transmission (low communication). However, it requires each user to hold as many establishment keys as there are privileged sets in which it is a member (high storage). These two examples suggest there is a trade-off between the number of transmissions needed to establish a new broadcast key and the number of establishment keys held by each user. This trade-off is the subject of this paper.

We focus on the case in which the privileged sets consist of all sets of users of a certain fixed size. It is not unreasonable to focus on such a collection of privileged sets, since in practice the number of users requesting any given broadcast can be bounded accurately a priori. For example, the set of excluded users may just be those who have neglected to pay their pay-television bill that month. Over time, the number of delinquent users is likely to be relatively stable (and small).

We prove that for a given upper bound on the number of establishment keys held by each user there is an inherent lower bound on the number of transmissions needed to establish a new broadcast key. For different types of protocols, we then describe constructions that come within a constant factor of these bounds, thereby demonstrating that our trade-off bounds are close to optimal. We note that our bounds do not take into account how much information is sent with each transmission. Most of our constructions do not send much information (i.e. just a single key) with each transmission.

The organization of this paper is as follows: Section 2 describes previous work in this area, Section 3 contains all the definitions and notation, Section 4 describes the model from a set theoretic perspective and the mathematical tools that we use, Section 5 contains lower bounds on the number of keys per user in broadcast encryption systems (our main results), Section 6 describes constructions that are close to these bounds, and Section 7 is a brief conclusion.

2 Previous Work

Previously, bounds in the general broadcast encryption model have been given for various parameters. Fiat and Naor [7] introduce broadcast encryption and describe several constructions. They focus on a feature of broadcast encryption systems called *resiliency*. A broadcast encryption system is k-resilient if a center is able to broadcast to any set of privileged users with the assurance that no disjoint coalition of k excluded users can receive the broadcast even by sharing their establishment keys . They construct both unconditionally secure and computationally secure systems of various resiliencies. Our lower bounds apply even to 1-resilient protocols, and thus our bounds are as strong as possible. In addition, some of the establishment protocols we describe are resilient against arbitrary coalitions of excluded users and come close to the trade-off parameters of our lower bounds. We also describe less resilient establishment protocols that meet our lower bounds.

Blundo and Cresti study unconditionally secure broadcast encryption systems further in [2]. They prove information theoretic lower bounds for a model of unconditionally secure broadcast encryption focusing on zero-message broadcast encryption (no transmissions by the center) and interactive broadcast encryption. Here we present broadcast encryption systems in which the number of keys per user is much smaller than the zero-message schemes in [2] by allowing a positive number of transmissions. These transmissions take the form of one-way (i.e. noninteractive) broadcasts from the center to the users.

In [4] constructions and lower bounds for a model of unconditionally secure broadcast encryption are presented. The authors of [4] are also interested in the communication-storage trade-off. In their model each user is given some secret information and the users use the information to compute common keys via a key predistribution scheme such as in [1] or [3]. The efficiency of the systems in [4] is measured by considering the amount of secret information held by each user as compared to the information content of the broadcast made to establish the broadcast key; and the size of the broadcast as compared to its information content (i.e. it is an information theoretic model). In this paper we assume that the users are actually given the keys (for example, in an integrated circuit (IC) card) rather than the information with which to compute them, and communication is measured in terms of the number of keys needed to establish the broadcast key (the number of transmissions). The efficiency of our systems is measured by comparing the number of keys per user to the number of transmissions. These are both important practical parameters. In an implementation of a broadcast encryption system, a user's keys may be contained in an IC card with only limited memory, and the broadcasting center may want to limit the number of transmissions due to cost-efficiency concerns. Because of the differences between our measurements of efficiency and those in [4], the optimal systems in [4] are generally not optimal in our model. For example, they present an optimal scheme using resolvable designs, with $\binom{x}{\frac{x}{2}}$ transmissions to broadcast to a privileged set of size x, out of a universe of n users, that requires each user to generate $\binom{n-1}{\frac{x}{2}-1}$ keys. In this paper we present a system in which each user has $\binom{n-1}{\frac{x}{2}-1}$ keys and only 2 transmissions are needed. Another difference between [4] and this paper is in the mathematical tools used to prove lower bounds on the trade-off between communication and storage. A study of the keys per user versus transmissions trade-off is well suited to a combinatorial analysis. Results from extremal set theory lead to tight bounds on the number of keys per user in terms of the number of transmissions.

Stinson and Trung [12] continue the analysis of the trade-off studied in [4] by presenting new constructions of key predistribution schemes and broadcast encryption systems. They also prove new lower bounds on information rates of the aforementioned trade-offs.

A survey of broadcast encryption systems (constructed prior to [4]) can be found in [11].

3 Definitions and Notation

We are largely motivated by the scenario of pay-TV in which there is a set of users who have paid to watch a particular TV station. We call the users who have paid for this service the set of *privileged users*, and the collection of users who are not in the set, *excluded users.* We want to allocate establishment keys to users in such a way that the center can establish a new broadcast key with which to encode the TV station for any particular set of privileged users. Any excluded user should be unable to decipher the broadcast key. We'll denote the collection of privileged sets of users by $\mathcal{P}$.

In this paper, we let $\mathcal{P}$ be the collection of all subsets of users of size $n - m$, where n is the total number of users, and m is the size of an excluded set of users. We show that the number of keys per user can be reasonably small when either m is much smaller than n ($m << n/2$) or m is very large ($m >> n/2$). It is likely that one of these scenarios will be the case in practice. For example, to a pay-television station m is the number of users who do not pay their bill in a given month; usually this is a small number. On the other hand, to a pay-per-view provider m is the (usually large) number of users who do not request to view a particular film.

Let S denote the set of all establishment keys. The set of establishment keys known by user u, is denoted by $U \subseteq S$. Let $K = |S|$ be the total number of establishment keys, and let $|U|_{max} = \max_u |U|$. For a set of privileged users, $P \in \mathcal{P}$, the set of establishment keys which the center uses to establish the new broadcast key will be denoted by $S_P \subseteq S$. The *number of transmissions* is defined to be $t = \max_{P \in \mathcal{P}} |S_P|$.

The focus of this paper is the trade-off between the number of transmissions, t, and the maximum number of keys per user, $|U|_{max}$.

The broadcast key used to encrypt the TV station for the users in P is denoted by B_P. To each privileged set P there is an associated establishment protocol. The establishment protocol defines which subsets of keys in S_P are sufficient to recover B_P. Two natural establishment protocols are what we call the *OR* and *AND* protocols. If the center is broadcasting to P with an *OR* protocol then a user needs at least one key in S_P to be able to decrypt B_P. With an *AND* protocol a user needs all the keys in S_P to be able to decrypt B_P. We will discuss specific examples of broadcast encryption systems that use these protocols later.

We'll often refer to establishment keys and establishment protocols as, simply, keys and protocols. However, we will always distinguish between (establishment) keys and broadcast keys.

Suppose the center wants to establish a broadcast key B_P to broadcast to the users in the privileged set P. The center first generates random binary strings B_P and T_P. For each key $k \in S_P$ the center then generates a string $c_k^P(B_P,T_P)$ based on the establishment protocol associated with P. The string $c_k^P(B_P,T_P)$ is then encrypted, in a computationally secure way, so that key k is necessary to decrypt it. Each user u is able to recover $\{c_k^P(B_P,T_P) : k \in U\}$. We assume that a user w for which $k \notin W$ gains no information about $c_k^P(B_P,T_P)$ from the encryption of $c_k^P(B_P,T_P)$. For each privileged set P there is a function d_P which on input all the information user $u \in P$ is able to decrypt, outputs B_P. The following conditions must be met by any establishment protocol:

I. Any privileged user is able to recover enough information to construct the broadcast key, i.e. $\forall u \in P$, $d_P(\{c_k^P(B_P,T_P) : k \in U\}) = B_P$.
II. For any possible decoding algorithm d'_P, and for any possible output string β of the decoding algorithm, each string $\alpha \in \{0,1\}^{|B_P|}$ is equally likely to be the broadcast key, i.e. $\forall w \notin P$, $\forall d'_P$, $\forall \alpha \in \{0,1\}^{|B_P|}$, $\forall \beta \in \{0,1\}^{|B_P|}$, $\Pr[B_P = \alpha | d'_P(\{c_k^P(B_P,T_P) : k \in W\}) = \beta] = \frac{1}{2^{|B_P|}}$.

Note that an excluded user w may be able to obtain some information if he has some of the keys used to encrypt the transmissions. The broadcast encryption system must be designed so that the broadcast key is uniformly distributed even with this information.

For an *OR* protocol the center sets $T_P = \emptyset$, the empty string. Since any key in S_P must be sufficient to decode B_P, the center defines $c_k^P(B_P,\emptyset)$ to be B_P for every $k \in S_P$. In other words, the center replicates B_P, $|S_P|$ number of times. Then $d_P(\{c_k^P(B_P,\emptyset) : k \in U\}) = c_k^P(B_P,\emptyset) = B_P$ for all $k \in S_P$. It is important to note here that *OR* protocols are secure against arbitrary coalitions of excluded users since any excluded user has none of the keys in S_P and it is necessary to have at least one of the keys in S_P to decode B_P.

The *OR* Protocol

- Any one key in S_P is sufficient to recover the broadcast key, B_P.
- Secure against arbitrary coalitions of excluded users, since any excluded user has *none* of the keys in S_P.
- Implementation:
  1. Set $T_P = \emptyset$.
  2. For all $k \in S_P$, $c_k^P(B_P,T_P) = B_P$.
  3. For all $u \in \mathcal{P}$, $\exists k \in S_P \cap U$ such that $d_P(c_k^P(B_P,T_P)) = B_P$.

For an *AND* protocol with $S_P = \{k_1,\ldots,k_r\}$, $r \leq t$, the center generates $r-1$ random strings $T_P^{k_1},\ldots,T_P^{k_{r-1}}$ and defines $T_P^{k_r}$ to be $B_P \oplus T_P^{k_1} \oplus \ldots \oplus$

$T_P^{k_{r-1}}$. The string T_P is the concatenation of $T_P^{k_1},\dots,T_P^{k_r}$. For each $k_i \in S_P$, $c_{k_i}^P(B_P,T_P) = T_P^{k_i}$, and for every user u in P, $d_P(\{c_k^P(B_P,T_P) : k \in U\}) = \oplus_{i=1}^r c_{k_i}^P(B_P,T_P) = B_P$. If a user is missing $k_i \in S_P$ then the user will be unable to decode $c_{k_i}^P(B_P,T_P) = T_P^{k_i}$, and hence will not be able to decode B_P.

The *AND* Protocol

- It is necessary to have *all* of the keys in S_P to recover B_P.
- Secure against a coalition of one excluded user; two excluded users may be able to recover B_P by pooling their keys.
- Implementation:
  1. Let $S_P = \{k_1,...,k_r\}$, $r \leq t$. For all $i < r$, $T_P^{k_i}$ is a randomly chosen string in $\{0,1\}^{|B_P|}$, $T_P^{k_r} = B_P \oplus T_P^{k_1} \oplus ... \oplus T_P^{k_{r-1}}$ and $T_P = T_P^{k_1} \parallel T_P^{k_2} \parallel ... \parallel T_P^{k_r}$.
  2. For all $k_i \in S_P$, $c_{k_i}^P(B_P,T_P) = T_P^{k_i}$.
  3. For all $u \in \mathcal{P}$, $\forall i = 1,...,r$ $k_i \in U$, and $d_P(\{c_{k_i}^P(B_P,T_P) : i = 1,...,r\}) = \oplus_{i=1}^r c_{k_i}^P(B_P,T_P) = B_P$.

We can implement other establishment protocols by using these same ideas of replication (as in the *OR* protocol) and exclusive-or (as in the *AND* protocol). We will discuss other establishment protocols more in later sections.

Finally, it will be helpful in our later discussion of establishment protocols to have a function associated with each privileged set P that on input a subset of S_P returns a value of 1 if the subset is sufficient to decode B_P, and 0 otherwise. This function is referred to as a *characteristic function for the establishment protocol associated with P*. This is formalized below.

Let $\chi_U \in \{0,1\}^K$ be the characteristic string of the keys held by user u. Let $\chi_{U_i} \cap \chi_{U_j}$ be the characteristic string of the intersection of sets U_i and U_j. Let D denote the inclusion poset on $\{0,1\}^K$ (see Section 4.2 for definition). For every subset $P \in \mathcal{P}$ we have a monotonically increasing function $f_P : D \longrightarrow \{0,1\}$. Let $\chi_{S_P} \in \{0,1\}^K$ be the characteristic string of S_P, then χ_{S_P}has at most t ones. The following hold:

I. $\forall P \in \mathcal{P}, \forall u \in P, f_P(\chi_U \cap \chi_{S_P}) = 1$
II. $\forall P \in \mathcal{P}, \forall w \notin P, f_P(\chi_W \cap \chi_{S_P}) = 0$

For example, let the number of users be $n = 3$ and let $\mathcal{P}$ be the collection of all subsets of users of size 2. Then with $K = 3$, $t = 2$, and *OR* protocols for each P_i $i = 1,2,\dots,6$, the following characteristic functions (monotonically extended) satisfy the above properties:

$$f_{\{u_1,u_2\}}(1,0,0) = 1, f_{\{u_1,u_2\}}(0,1,0) = 1, f_{\{u_1,u_2\}}(0,0,1) = 0$$
$$f_{\{u_1,u_3\}}(1,0,0) = 1, f_{\{u_1,u_3\}}(0,1,0) = 0, f_{\{u_1,u_3\}}(0,0,1) = 1$$

$f_{\{u_2,u_3\}}(1,0,0) = 0, f_{\{u_2,u_3\}}(0,1,0) = 1, f_{\{u_2,u_3\}}(0,0,1) = 1$

In this example, $\forall i, j$, $1 \leq i, j \leq 3$, $S_{\{u_i,u_j\}} = \{k_i, k_j\}$.

4 A Set Theoretic Approach to Broadcast Encryption

4.1 Establishment Protocols

In Section 3 we introduce the functions $\{f_P\}_{\mathcal{P}}$; the characteristic functions of the establishment protocols associated with the privileged sets. From each f_P a set theoretic description of each privileged set can be derived. This description suggests a natural construction with *OR* protocols. Also, this description may be helpful in proving other protocol specific lower bounds. We first describe how each f_P is equivalent to a certain logical formula involving the boolean operations $\vee$ and $\wedge$, and then we show how to translate this logical formula into a set formula for P.

Definition 1. Let Σ be a set containing the symbols $k_1, \ldots, k_K$ that is closed under the boolean operations $\wedge$ and $\vee$. A *formula* is any member of Σ.

To find a formula that corresponds to a function f_P, we simply consider all sets of keys $\{A_i\}$ that suffice to receive the broadcast key (i.e. $f_P(\chi_{A_i}) = 1$). An f_P function can then be expressed as a formula by taking the disjunction of all formulas of the form $\wedge_{k \in A_i} k$. To find an equivalent formula we only consider minimal sets A_i that suffice to receive the broadcast key. For example, let $S_P = \{k_1, k_2, k_3\}$ and let f_P be defined as follows:

$f_P(0,0,0) = 0, f_P(1,0,0) = 0, f_P(0,1,0) = 0, f_P(0,0,1) = 0,$
$f_P(1,1,0) = 1, f_P(1,0,1) = 1, f_P(0,1,1) = 1, f_P(1,1,1) = 1.$

Then we can represent f_P by the formula $(k_1 \wedge k_2) \vee (k_1 \wedge k_3) \vee (k_2 \wedge k_3) \vee (k_1 \wedge k_2 \wedge k_3)$ or equivalently, $\sigma_{f_P} = (k_1 \wedge k_2) \vee (k_1 \wedge k_3) \vee (k_2 \wedge k_3)$. We can translate this into a set theoretic formulation by letting κ_i denote the set of users who have key k_i.

To implement a protocol given a formula, simply use a separate *AND* protocol on each of the conjunctive subformulas as described in Section 3. To implement the previous example use three independently generated *AND* protocols, for the same broadcast key B_P, on the conjunctive subformulas.

Definition 2. Let Σ_S be a collection of sets containing the symbols $\kappa_1, \ldots, \kappa_K$ that is closed under the operations of intersection, $\cap$, and union, $\cup$. A *set formula* is any member of Σ_S.

We have the following theorem that holds for any set system $\mathcal{P}$.

Theorem 3. *A broadcast encryption system with characteristic functions $\{f_P\}_{P \in \mathcal{P}}$ and K keys total exists if and only if there are K sets $\kappa_1, \ldots, \kappa_K$, each contained in $\{u_1, \ldots, u_n\}$, such that $\forall P \in \mathcal{P}$ there exists a set $\{i_1, \ldots, i_{r_P}\} \subseteq \{1, ..., K\}$, $(r_P \leq t)$ and P is equal to a set formula σ_{f_P} with set symbols $\kappa_{i_1}, \ldots, \kappa_{i_{r_P}}$.*

Proof: Assume we have such a broadcast encryption system. Then for all $P \in \mathcal{P}$ there is a boolean function f_P and a set S_P of at most t keys, that returns one on input a characteristic string $\chi_U \cap S_P$ if $u \in P$, and returns zero if $u \notin P$. To construct a formula that describes f_P, first form the conjunction of the set of key symbols corresponding to a minimal set of keys in S_P that suffices to decrypt B_P for each privileged user. The formula consists of the disjunction of all the subformulas formed in this way (i.e. one for each privileged user). If we substitute κ_i for k_i, $\cap$ for $\wedge$, and $\cup$ for $\vee$ then we obtain a set formula for P.

Conversely, allocate to user i key k_j if and only if $u_i \in \kappa_j$. Translate the set formulas into formulas for monotonic encryption functions by reversing the above substitutions. Then we have a broadcast encryption system for $\mathcal{P}$ with at most t transmissions. The system can be implemented as described previously.□

The previous theorem proves broadcast encryption systems can also be defined in a set theoretic manner. This description doesn't capture all aspects of the implementation of the system; for example, the length of the transmissions is not explicitly defined.

The following Corollary gives a necessary and sufficient characterization of *OR* protocols.

Corollary 4. *There is a broadcast encryption system with* OR *protocols for $\mathcal{P}$, at most t transmissions, and K keys total if and only if there are K subsets $\kappa = \{\kappa_1, \dots, \kappa_K\}$ of $\{u_1, \dots, u_n\}$ such that for all $P \in \mathcal{P}$ there are $1 \le i_1, \dots, i_{r_P} \le K$, $r_P \le t$, such that $P = \cup_{j=1}^{r} \kappa_{i_j}$.*

Given any collection of key establishment protocols we can prove corollaries to Theorem 3, as we did above for a collection of *OR* protocols. A construction of a broadcast encryption system with *OR* protocols follows naturally from the set theoretic characterization given here.

4.2 Mathematical Tools

In this section we describe a couple of concepts and theorems that we use to establish our main results; lower bounds on the number of keys per user in broadcast encryption systems. The previous section indicates that it's helpful to think of broadcast encryption systems in a set theoretic way, and the mathematics we discuss here is from the area of extremal set theory.

Definition 5. A *poset* (partially ordered set) is a set A with a binary relation $\le$ such that:

(i) $a \le a$ for all $a \in A$ (reflexivity)

(ii) if $a \le b$ and $b \le c$ then $a \le c$ (transitivity)

(iii) if $a \le b$ and $b \le a$ then $a = b$ (antisymmetry).

Example 1. The *inclusion poset* on $\{1, \dots, K\}$ consists of the subsets of $\{1, \dots, K\}$ ordered by inclusion.

Definition 6. An *antichain* is a set of elements of a poset that are pairwise incomparable.

Sperner [10] proved a famous result on the size of an antichain in the inclusion poset (often called a *Sperner family*). The following is a strengthening of this result. It was discovered independently by Lubell [8], Meshalkin [9] and Yamamoto [14]. Although it is also a special case of a result of Bollobás [5], it is usually referred to as the *LYM* inequality.

Lemma 7 LYM Inequality, Bollobás, Lubell, Meshalkin and Yamamoto. *Let $S_1, \ldots, S_r$ be subsets of $\{1, \ldots, K\}$ such that $\{S_i\}_{i=1}^r$ is an antichain in the inclusion poset, and let f_ℓ denote the number of sets of size ℓ, $0 \leq \ell \leq K$. Then*

$$\sum_{\ell=0}^{K} f_\ell \binom{K}{\ell}^{-1} \leq 1.$$

To prove lower bounds, we rely heavily on the combinatorial concept of a *sunflower.*

Definition 8. A set system $\mathcal{F} = \{F_1, \ldots, F_M\}$ is a *sunflower* with M *petals* if $\forall i \neq j, 1 \leq i, j \leq M$

$$F_i \cap F_j = \bigcap_{r=1}^{M} F_r$$

$\bigcap_{r=1}^{M} F_r = C_{\mathcal{F}}$ is called the *center* of the sunflower.
A *petal* in the sunflower $\mathcal{F}$ is a set of the form $F_i - (\cap_{r=1}^{M} F_r) = F_i - C_{\mathcal{F}}$.

The following famous results gives a lower bound on the size of a sunflower in a set system.

Lemma 9 Sunflower Lemma, Erdös and Rado. *Let t, n be positive integers. Let $\mathcal{F}$ be a collection of n sets, each of size at most t. Then $\mathcal{F}$ contains a sunflower of size at least $\frac{n^{1/t}}{t}$.*

5 Lower Bounds

In this section we prove lower bounds on the number of keys per user for a variety of protocols. We begin with *OR* protocols, as these are both simple and very secure. In Section 5.2, we show that the ideas behind the proofs in Section 5.1 can be extended without much difficulty to a much larger class of protocols that we call *consistent* protocols. Consistent protocols are interesting because they are more general, but still easy to implement. In Section 5.3 we prove lower bounds for a broadcast encryption system with an arbitrary collection of protocols. For the case of all *OR* protocols and consistent protocols, the bounds are the same. For an arbitrary collection of protocols and small t, we prove a lower bound on $|U|_{max}$ that is on the same order as in the previous two cases.

5.1 *OR* Protocols

In this section all the protocols are *OR* protocols. We are particularly interested in *OR* protocols because, as mentioned in Section 3, any *OR* protocol is resilient against arbitrary coalitions of excluded users.

To motivate our lower bounds on the number of keys per user we first consider the relationship between K the total number of keys, and $|\mathcal{P}|$, the number of privileged sets. To do this, we prove the following simple corollary of the *LYM* inequality.

Corollary 10. *If $S_1, \ldots, S_r$ are subsets of $\{1, \ldots, K\}$ such that $\forall i$, $1 \leq |S_i| \leq t \leq K/2$ and $\{S_i\}_{i=1}^{r}$ is an antichain in the inclusion poset, then $r \leq \binom{K}{t}$.*

Proof: If $t \leq K/2$ then for every ℓ, $1 \leq \ell \leq t$, then $\binom{K}{\ell} \leq \binom{K}{t}$. The result follows from the *LYM* inequality. □

Lemma 11. *Let $t \leq K/2$. Then in any broadcast encryption system with* OR *protocols, K is $\Omega(\binom{n}{m}^{1/t})$.*

Proof: Since $\{S_P\}_{P \in \mathcal{P}}$ is an antichain, we can apply Corollary 10 to get $\binom{K}{t} \geq \binom{n}{m}$. □

It follows from the previous lemma that $|U|_{max}$ is $\Omega(\frac{1}{n}\binom{n}{m}^{1/t})$. However a larger lower bound can be proven when m is much smaller than n. We show that for arbitrary but fixed (with respect to n) values of m and t, the maximum number of keys per user and the average number of keys per user are both $\Omega(\binom{n}{m}^{1/t})$. To prove these two results we rely on the Sunflower lemma of Erdös and Rado.

Theorem 12. *In any broadcast encryption system with* OR *protocols, $|U|_{max} \geq \left(\frac{\binom{n}{m}^{1/t}}{t} - 1\right)/m$.*

Proof: From Lemma 9 we know that the set system $\{S_P\}_{P \in \mathcal{P}}$ contains a sunflower, $\mathcal{F}$, of size at least $\frac{\binom{n}{m}^{1/t}}{t}$. Consider a set $S_P \in \mathcal{F}$. The users in P^c must, as a group, contain at least one key in each of the other petals of the sunflower; therefore they collectively have at least $\frac{\binom{n}{m}^{1/t}}{t} - 1$ keys, and so, some user in the group has at least $\left(\frac{\binom{n}{m}^{1/t}}{t} - 1\right)/m$ keys. □

We can also use Lemma 9 to get a lower bound on the average number of keys per user. For this it suffices to show that the number of sets in $\{S_P\}_{P \in \mathcal{P}}$ that aren't in a sufficiently large sunflower is exponentially small.

Theorem 13. *In any broadcast encryption system with* OR *protocols the average number of keys per user is at least $\frac{\binom{n}{m}^{1/t}}{8tm}$.*

Proof: Let $\mathcal{S} = \{S_P : P \in \mathcal{P}\}$. Find sunflowers $\mathcal{F}_i$ in $\mathcal{S} - \cup_1^{i-1}\mathcal{F}_j$ each of size $\ell = \frac{|\mathcal{P}|^{1/t}}{2t}$ until there are no more. Let s be the number of sunflowers found in this way. The number of sets in $\mathcal{S}$ that aren't in sunflowers is less than $\frac{|\mathcal{P}|}{2^t}$ by Lemma 9.

Let $\mathcal{F}_i = \{S_{P_{i_1}}, \ldots, S_{P_{i_\ell}}\}$, $1 \leq i \leq s$. Consider a set $S_{P_{i_j}}$ in sunflower $\mathcal{F}_i$. None of the users in $P_{i_j}^c$ have any of the keys in $S_{P_{i_j}}$, so as a group they must have a key in each of the petals $\{S_{P_{i_r}} - C_{\mathcal{F}_i}\}_{r \neq j}$. Therefore, $\sum_{u \in P_{i_j}^c} |U| \geq \frac{|\mathcal{P}|^{1/t}}{2t} - 1$. If we let T be the sum of $|U|$ for all users u who are excluded by some set P_{i_j} of some sunflower, $\mathcal{F}_i$, then since each user is excluded by $\frac{m|\mathcal{P}|}{n}$ privileged sets, we have:

$$|U_1| + |U_2| + \ldots + |U_n| \geq T \geq \left[\frac{|\mathcal{P}|^{1/t}}{2t} - 1\right] |\mathcal{P}| \left[1 - 1/2^t\right] \frac{1}{m/n|\mathcal{P}|}$$

Therefore, the average number of keys per user is at least $\frac{|\mathcal{P}|^{1/t}}{8tm}$. □

5.2 Consistent Establishment Protocols

Just as we've considered broadcast encryption systems with all *OR* protocols, we might consider broadcast encryption systems in which the protocols are all the same, though not necessarily of the *OR* type (e.g. all *AND* protocols). In fact, we can generalize this notion a bit to obtain what we call *consistent* protocols and show that the results from Section 5.1 hold when the protocols for the privileged sets are consistent. Such a collection of protocols will in general be simpler to implement than an arbitrary collection of protocols, but they will not generally have the high security of the *OR* protocols. Informally, the protocol for P is consistent with the protocol for P' if any subset $V \subseteq S_P \cap S'_P$ suffices to receive B_P if and only if it suffices to receive $B_{P'}$. We formalize the definition of consistent establishment protocols in terms of the characteristic functions below and prove lower bounds in this case.

Definition 14. The functions $\{f_P\}_{P \in \mathcal{P}}$ are *consistent* iff for all $P_1 \neq P_2$, and for all characteristic strings χ_V where $V \subseteq S_{P_1} \cap S_{P_2}$, $f_{P_1}(\chi_V) = f_{P_2}(\chi_V)$.

Theorem 15. *In any broadcast encryption system with consistent protocols,* $|U|_{max} \geq \frac{\binom{n}{m}^{1/t}}{2tm}$.

Proof: From Lemma 9 we know that the set system $\{S_P\}_{P \in \mathcal{P}}$ contains a sunflower, $\mathcal{F}$, of size at least $\frac{|\binom{n}{m}|^{1/t}}{t}$. Consider $S_{P_1} \in \mathcal{F}$. For every $S_{P_i} \in \mathcal{F}$ different from S_{P_1}, there is some user $u \in P_1^c \cap P_i$. Since the functions $\{f_P\}_P$ are consistent user u must have a key in S_{P_i} that's not in $S_{P_1} \cap S_{P_i}$, so user u has a key in the petal $S_{P_i} - C_{\mathcal{F}}$. By this argument, at least one of the m users in P_1^c must have a key in each of the petals $S_{P_i} - C_{\mathcal{F}}$, $i \neq 1$, so some user in P_1^c has at least $\left(\frac{\binom{n}{m}^{1/t}}{t} - 1\right)/m$ keys. □

Theorem 16. *In any broadcast encryption system with consistent protocols the average number of keys per user is at least* $\frac{\binom{n}{m}^{1/t}}{8tm}$.

Proof: Apply the same modification to Theorem 13 as was applied to Theorem 12 to prove Theorem 15. □

5.3 General Establishment Protocols

In this section we consider broadcast encryption systems in which the protocols corresponding to the individual privileged sets are not necessarily related. When t is small, we can extend the ideas of the previous section to get a lower bound on the maximum number of keys per user. Recall that, practically, a small value for t is a desirable feature of a broadcast encryption system, as it usually indicates that a broadcast key can be established quickly and inexpensively.

Theorem 17. *In any broadcast encryption system with at most* $t < \sqrt{\log n}$ *transmissions,* $|U|_{max}$ *is* $\Omega(\binom{n}{m}^{1/t})$.

To facilitate the proof of Theorem 17 we have the following definitions. Let $\mathcal{F}$ be a sunflower with center $C_{\mathcal{F}}$. Let $T \subseteq C_{\mathcal{F}}$.

Definition 18. A *block*, L_T, is the set of all users u_i such that $U_i \cap C_{\mathcal{F}} = T$.

Definition 19. A block is *split* by a petal $S_P - C_{\mathcal{F}}$ of the sunflower $\mathcal{F}$ if there exist $u_i, u_j \in L_T$ such that $u_i \in P$ and $u_j \in P^c$.

Proof: By Lemma 9, the set system $\{S_P\}_{P \in \mathcal{P}}$ contains a sunflower $\mathcal{F}$ of size $\frac{|\mathcal{P}|^{1/t}}{t}$, with center $C_{\mathcal{F}}$. There are at most 2^t subsets of $C_{\mathcal{F}}$. Let $L_1, \ldots, L_\ell$ ($\ell \leq 2^t$) be the blocks corresponding to those subsets.

Since for all $P_i \in \mathcal{P}$, $|P_i^c| = m$ there are at most 2^{tm} petals in $\mathcal{F}$ which don't split any of the blocks, L_i. Therefore, there are at least $\frac{|\mathcal{P}|^{1/t}}{t} - 2^{tm}$ petals that each split some block. Some block must be split by at least $\frac{\frac{|\mathcal{P}|^{1/t}}{t} - 2^{tm}}{2^t}$ petals. Let L_i be such a block. We have the following two cases:

(i) If $|L_i| \leq 2m$ then there are at most $4m^2$ ordered pairs of users that could be split. So, some user has at least $\frac{\frac{|\mathcal{P}|^{1/t}}{t} - 2^{tm}}{2^t 4m^2 2}$ keys.

(ii) If $|L_i| > 2m$ then if S_P is a petal that splits L_i, P must include at least $1/2$ of the users in L_i. Therefore the average number of keys amongst the users in L_i is at least $\frac{1}{2}\left(\frac{|\mathcal{P}|^{1/t}}{t2^t} - 2^{t(m-1)}\right)$ keys.

□

6 Constructions

In this section we demonstrate that the lower bounds from Section 5 are essentially tight by describing broadcast encryption systems that come close to the bounds. The first construction is for the most secure case of *OR* protocols, the second construction uses all *AND* protocols and the third construction uses consistent protocols. Also, in the first two constructions a relatively small amount of information is sent with each transmission. The last construction may have large transmission sizes.

Theorem 20. *There is a broadcast encryption system with* OR *protocols in which* $|U|_{max}$ *is* $\binom{n-1}{\lceil\frac{n-m}{t}\rceil-1}$. *This is close to optimal for large t.*

Proof: Note that with *OR* protocols we would never need more than $n-m$ transmissions. Let κ consist of all subsets of $\{u_1,\ldots,u_n\}$ of size $\lceil\frac{n-m}{t}\rceil$. This construction (construction *I*) has $\binom{n}{\lceil\frac{n-m}{t}\rceil}$ keys total.

We can use approximations to binomials to show that the ratio of $\binom{n}{\lceil\frac{n-m}{t}\rceil}$ to our bound from Theorem 12 is $O(\frac{(et)^{\frac{n-m}{t}}}{n^2})$. In particular,

$$(et)^{\frac{n-m}{t}} > \frac{\binom{n}{\frac{n-m}{t}}}{\binom{n}{n-m}^{1/t}} > \left(\frac{t}{e}\right)^{\frac{n-m}{t}}$$

Therefore, for large t, $\binom{n}{\frac{n-m}{t}}$ is close to $\binom{n}{n-m}^{1/t}$, and so the above construction is close to optimal. □

Although the above theorem shows that we cannot always reach our lower bound on the number of keys per user with *OR* protocols, we can construct optimal broadcast encryption systems for arbitrary t and m with other protocols. For both of the following simple broadcast encryption systems the number of keys per user and the total number of keys are on the order of $\binom{n}{m}^{1/t}$. Except in the case $m \geq n/2, m \geq t$ they are not as resilient against colluding users as *OR* protocols. Also, construction **III** may require that a large amount of information be sent in each transmission.

II. A broadcast encryption system for $t \leq m$:

Let $K = \binom{n}{\lceil\frac{m}{t}\rceil}$. Note that this implies that the number of keys per user is on the order of our proven lower bounds.

For every subset A of $\lceil m/t\rceil$ users create a key, k_A. Give k_A to every user except those in A. Given any set of m excluded users, P^c, choose $A_1,\ldots,A_t \subseteq P^c$ such that $\cup_{i=1}^{t} A_i = P^c$ and $|A_i| = \lceil m/t\rceil$. Let k_i be the key that all the users in A_i are missing. We'll decide to transmit information using these keys in such a way that a user must have all of $k_1,\ldots,k_t$ to receive the broadcast key. This system uses *AND* protocols with $S_M = \{k_1,\ldots,k_t\}$.

In the notation of the previous section, $\kappa = \{\kappa_1,\ldots,\kappa_K\}$ is the collection of all subsets of $\{u_1,\ldots,u_n\}$ of size $n-\lceil\frac{m}{t}\rceil$.

Two broadcast encryption systems that use the same set of keys are complementary if the set of keys that user i holds in one system is the complement of the set of keys the same user holds in the other system. Therefore, if we can use *OR* protocols to broadcast to any set of $n - m$ users in a *BES*, then we can use *AND* protocols to broadcast to any set of m users in the complementary *BES*. When $m \geq n/2$ and $m \geq t$ we can use the method of construction **II** to find a broadcast encryption system for $\mathcal{P}^c = \{P^c : P \in \mathcal{P}\}$ (the complementary *BES*). This system uses *OR* protocols for $\mathcal{P}$ and the number of keys per user is close to our lower bounds.

We can increase the resiliency of this broadcast encryption system by increasing the size of the subsets A of the first paragraph. This will increase the likelihood that a subset of colluding users are all missing a particular key. It will also increase the number of keys per user.

III. A broadcast encryption system for $t > m$:

Let $\binom{K}{\lceil \frac{t}{m} \rceil}$ be the least integer greater than or equal to n, so the number of keys per user is on the order of the proven lower bounds.

Let U_i be the set of keys held by user u_i. Choose n subsets $U_1, \ldots, U_n$ of the key set $\{1, \ldots, K\}$ each of size $K - \lceil \frac{t}{m} \rceil$. Let P be a set of privileged users. We'll transmit to user $u_r \in P$ with the keys in $U_r \bigcap [\bigcup_{i \in P^c} U_i^c]$ using an *AND* protocol. The number of transmissions is $| \cup_{i \in P^c} U_i^c| \leq t$ (often, this is a strict inequality). This last inequality holds because each user is missing exactly $\lceil \frac{t}{m} \rceil$ keys.

Note that for u_s to be able to recover the broadcast key, u_s must have all the keys in $U_r \bigcap [\bigcup_{i \in P^c} U_i^c]$ for some $u_r \in P$. If $U_r \bigcap [\bigcup_{i \in P^c} U_i^c] \subseteq U_s$ and $u_s \in P^c$ then $U_r \cap U_s^c = \emptyset$. This implies that $U_r = U_s$, a contradiction, so the system is secure.

7 Conclusion

In this paper, we've studied the trade-off between the number of keys per user and the number of transmissions in broadcast encryption systems. These are important parameters to study because they measure quantities that effect the cost-effectiveness and speed of a broadcast encryption system. The number of keys per user has a positive correlation with the amount of memory per user, and the number of transmissions effects the speed of the system. These are the first proven lower bounds for these parameters, as far as we know. Some simple constructions demonstrate that these bounds are essentially tight.

An additional consideration, not fully addressed here, is that of the size of each transmission (or bandwidth). Our first two constructions are efficient in this respect, as they each require only that a binary string of the same size as the broadcast key be sent with each transmission. The third construction, however, requires that $O(n)$ binary strings of the same size as the broadcast key be sent with each transmission.

Acknowledgement

We would like to thank Benny Chor, Amos Fiat, Moni Naor and Rafail Ostrovsky for helpful discussions on broadcast encryption, and Matt Robshaw and Yiqun Lisa Yin for helpful comments on an earlier draft of this paper.

References

1. R. Blom, *An optimal class of symmetric key generation systems* , "Advances in Cryptology-EUROCRYPT '84", *Lecture Notes in Computer Science* **209** (1984), 335-338.
2. C. Blundo, A. Cresti, *Space requirements for broadcast encryption*, "Advances in Cryptology-EUROCRYPT '94", *Lecture Notes in Computer Science* **950** (1995), pp 287-298.
3. C. Blundo, A. De Santis, A. Herzberg, S. Kutten, U. Vaccaro and M. Yung, *Perfectly secure key distribution in dynamic conferences*, "Advances in Cryptology-CRYPTO '92", *Lecture Notes in Computer Science* **740** (1993), pp 471-486.
4. C. Blundo, L. A. Frota Mattos, D. R. Stinson, *Trade-offs between communication and storage in unconditionally secure schemes for broadcast encryption and interactive key distribution*, "Advances in Cryptology-CRYPTO '96", *Lecture Notes in Computer Science* **1109** (1996), pp 387-400.
5. B. Bollobás, *On generalized graphs*, Acta Math. Acad. Sci. Hungar., **16** (1965), pp 447-452.
6. P. Erdös, R. Rado, *Intersection theorems for systems of sets*, Journal London Math. Soc., **35** (1960), pp 85-90.
7. A. Fiat, M. Naor, *Broadcast encryption*, "Advances in Cryptology-CRYPTO '93", *Lecture Notes in Computer Science* **773** (1994), pp 480-491.
8. D. Lubell, *A short proof of Sperner's lemma*, J. Combinatorial Theory, **1** (1966), p 299.
9. L. D. Meshalkin, *A generalization of Sperner's lemma on the number of subsets of a finite set (English translation)*, Theory of Probab. and its Applns., **8** (1964), pp 204-205.
10. E. Sperner, *Ein Satz über Untermengen einer endlichen Menge*, Math. Zeitschrift, **27** (1928), pp 544-548.
11. D. R. Stinson, *On some methods for unconditionally secure key distribution and broadcast encryption*, Designs, Codes and Cryptography, **12** (1997), pp 215-243.
12. D. R. Stinson and T. van Trung, *Some new results on key distribution patterns and broadcast encryption*, submitted for publication.
13. J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*, Cambridge University Press, Cambridge, 1992.
14. K. Yamamoto, *Logarithmic order of free distributive lattices*, J. Math. Soc. Japan, **6** (1954), pp 343-353.