Anonymous Distribution of Broadcast Keys in Ad Hoc Systems

Jacek Cichoń, Łukasz Krzywiecki, Mirosław Kutyłowski,
(Wrocław University of Technology)
and Paweł Wlaź
(Technical University Lublin)

MADNES'2005, Singapore

Encoded broadcast

Application areas:

- pay TV
- services in 3G telecommunication networks

Features:

- pay for the access time only
- single broadcast channel, all subscribers get the same data

Solutions

- broadcast encrypted with a symmetric key K (session key)
- a subscriber that is logged in obtains K
- without K it is impossible to decode the transmission

A new subscriber Alice logs in

1. Alice contacts broadcasting system (request for a key + authorisation through a private channel)
2. the system responds with a message containing the current key K

A subscriber logs off

1. the session key K is changed and distributed to the users that remain in the system
2. transmission channel:

   option 1: private channel to each user (costly!)

   option 2: key update through appropriate messages in the broadcast channel (cheap!)

Update scenarios

scenario 1: only a few users leave the system at a time (most literature)

scenario 2: rapid changes

Our scenario

- the set of active users changes rapidly (mobility, consumers behavior...)
- it is unpredictable who requests the service and when
- the number of potential users is moderate

Communication model

cellular broadcast system:

- the service area divided into cells
- in each cell a base station broadcasts through a channel accessible by all mobile users in this cell
- a single broadcast channel of limited capacity

Privacy goals

1. the encryption key should not be decodable by unauthorized users
2. Alice should not be able to derive what Bob is doing - regardless whether or not Alice is logged in
3. a competition company should not be able to derive any information on the system usage

Privacy goals

1. the encryption key should not be decodable by unauthorized users
2. Alice should not be able to derive what Bob is doing - regardless whether or not Alice is logged in
3. a competition company should not be able to derive any information on the system usage

Privacy goals

1. the encryption key should not be decodable by unauthorized users
2. Alice should not be able to derive what Bob is doing
   - regardless whether or not Alice is logged in
3. a competition company should not be able to derive any information on the system usage

Users distribution

- $N =$ the total number of subscribers
  N is large (e.g. $N \approx 10^8$)
- $n =$ the maximal number of users requesting data in a cell,
  n is moderate (e.g. $n \approx 10^4$)

Private secrets

- a user A has a secret $s(A)$ shared with the broadcasting system
- some symmetric cryptography for authorization

Simple Solution 1

Goal: Alice, Bob, and Paul should get key K

1. transmission encoding a new key K:
   $E_{s(Alice)}(K), E_{s(Bob)}(K), E_{s(Paul)}(K)$ + test sequence: $E_K(date)$
2. Paul decrypts the first three ciphertexts with $s(Paul)$; Paul obtains K and two junk keys
3. Paul decrypts $E_K(date)$ with all keys
   K identified easily!

Simple Solution 1

Goal: Alice, Bob, and Paul should get key K

1. transmission encoding a new key K:
   $E_{s(Alice)}(K), E_{s(Bob)}(K), E_{s(Paul)}(K)$ + test sequence:
   $E_K(date)$
2. Paul decrypts the first three ciphertexts with $s(Paul)$;
   Paul obtains K and two junk keys
3. Paul decrypts $E_K(date)$ with all keys
   K identified easily!

Simple Solution 1

Goal: Alice, Bob, and Paul should get key K

1. transmission encoding a new key K:
   $E_{s(Alice)}(K), E_{s(Bob)}(K), E_{s(Paul)}(K)$ + test sequence: $E_K(date)$
2. Paul decrypts the first three ciphertexts with $s(Paul)$;
   Paul obtains K and two junk keys
3. Paul decrypts $E_K(date)$ with all keys
   K identified easily!

Complexity measures

transmission size for key update - broadcast channel capacity is limited

energy usage: receiving time of a user the receiver consumes energy from batteries of a mobile device, the receiver should be switched off as long as possible

Drawbacks of Solution 1

- high energy usage - all ciphertexts must be received (in the worst case)
- a large number of decryptions

Simple Solution 2

1. instead of $E_{s(A)}(K)$ transmission contains $A, E_{s(A)}(K)$ or an indexing data determining the location of $E_{s(A)}(K)$
2. for privacy: A can be replaced by $H(A, E_{s(A)}(t), t)$ for a hash function H and $t =$ current time

Features of Solution 2

- the number of decryptions $= 1$
- size of transmission data for keysize 64
  example: $N = 10.000.000$, $n = 1000$
  - indexing data: $\geq 1000 \cdot \log N \geq 1000 \cdot 23$ bits
  - ciphertexts of the key: $1000 \cdot 64$ bits
  - overhead: increase of transmission size by 36%
- privacy - OK

Lower bound

- can we transmit k-bit key to n users with a message of length $\ll n \cdot k$?
- lower bound: it is impossible

Lower bound transmission size is at least

$$n \cdot (k - \log n) - (0.5 \log n + 3) - k$$

Proof idea of lower bound

- a transmission and a session key K determine a unique subset of users (which retreive K)
- average transmission length + length of $K \geq$ $\log(\text{number of subsets})$

Design goals

- transmission size $\approx kn$
- small energy cost for mobile users
- full privacy

Tools: solution based on Shamir's secret sharing

- users $A_{j_1}, A_{j_2}, \ldots, A_{j_m}$
- q random but known
- let $u_i := H(q, s(A_{j_i})), \quad x_i := H'(q, s(A_{j_i}))$
  for $i = 1 \ldots, m$,

  where H, H' are different hash functions

Solution based on Shamir's secret sharing ...

- build a polynomial f of degree m
  such that $f(0) = K$, and $f(x_i) = u_i$ for $i \leq m$
- message transmitting K:

$$f(1), f(2), \ldots, f(m)$$

Reconstruction of K

- $m+1$ points are necessary for reconstruction of f,
- a value of f for one more point needed, apart from $f(1), \ldots, f(m)$.
    - otherwise **no** information on K,
- A_{j_i} uses (x_i, u_i) and $(1, f(1)), \ldots, (m, f(m))$: and Lagrange interpolation for reconstructing f and $f(0)$

Features of the scheme

- perfect anonymity
- not practical for a large m — due to computational effort

Main idea

- keys are transmitted in buckets corresponding to *bins*
- each bin is responsible for up to c users
- transmission in a bin is fully anonymous
- in each bin use the Shamir's scheme

Problems to solve

1. how to assign the users evenly to the bins?
2. how the user determines its own bin?
3. how to preserve anonymity?

Assignment to bins

Parameters:

- n – the number of users
- B – the number of bins (for instance $B = n/100$)
- d – a solution parameter
- F – a pseudorandom cryptographic function with the range $\{1, \dots, B/d\}$

Assignment to bins - naive solution

Parameters:

- the bin of A determined by $H(A, E_{s(A)}(t))$ (or any other pseudorandom function)
- problem: with high probability there is a bin that will contain many users above the average number

Assignment to bins - naive solution

Parameters:

- the bin of A determined by $H(A, E_{s(A)}(t))$ (or any other pseudorandom function)
- problem: with high probability there is a bin that will contain many users above the average number

Assignment to bins – *left*[d] procedure

Choice based on *left*[d] procedure by Berthold Vöcking:

- the sender chooses and broadcasts a random number ρ
- d groups of bins:
  $\{1, \ldots, B/d\}, \{B/d + 1, \ldots, 2B/d\}, \ldots$
- preliminary choice: user A assigned to d bins
- the ith bin chosen for A has index:
  $(i - 1) \cdot B/d + F(\rho, A, s(A), i)$
  (a "random" bin in group i)
  these bins can be determined by the sender and by A only

Assignment to bins – *left*[*d*] procedure

Choice based on *left*[*d*] procedure by Berthold Vöcking:

- the sender chooses and broadcasts a random number ρ
- d groups of bins:
  $\{1, \ldots, B/d\}, \{B/d + 1, \ldots, 2B/d\}, \ldots$
- preliminary choice: user A assigned to d bins
- the ith bin chosen for A has index:
  $(i - 1) \cdot B/d + F(\rho, A, s(A), i)$
  (a "random" bin in group i)
  these bins can be determined by the sender and by A only

Assignment to bins – *left*[*d*] procedure

Choice based on *left*[*d*] procedure by Berthold Vöcking:

- the sender chooses and broadcasts a random number ρ
- d groups of bins:
  $\{1, \ldots, B/d\}, \{B/d + 1, \ldots, 2B/d\}, \ldots$
- preliminary choice: user A assigned to d bins
- the ith bin chosen for A has index:
  $(i - 1) \cdot B/d + F(\rho, A, s(A), i)$
  (a "random" bin in group i)
  these bins can be determined by the sender and by A only

Assignment to bins – *left*[*d*] procedure

Choice based on *left*[*d*] procedure by Berthold Vöcking:

- the sender chooses and broadcasts a random number ρ
- d groups of bins:
  $\{1, \ldots, B/d\}, \{B/d + 1, \ldots, 2B/d\}, \ldots$
- preliminary choice: user A assigned to d bins
- the ith bin chosen for A has index:
  $(i - 1) \cdot B/d + F(\rho, A, s(A), i)$
  (a "random" bin in group i)
  these bins can be determined by the sender and by A only

left[*d*] procedure

- for $i = 1, 2, \ldots$ the sender uses **one bin** among the bins given by preliminary choice
- the bin chosen for the ith user: the bin with the smallest load after assigning bins for users 1 through $i - 1$

left[*d*] procedure

- for $i = 1, 2, \ldots$ the sender uses **one bin** among the bins given by preliminary choice
- the bin chosen for the ith user: the bin with the smallest load after assigning bins for users 1 through $i - 1$

Properties

- The number of users assigned to some bin exceeds

$$\frac{n}{B} + i + \gamma$$

(γ is some constant) with a probability that can be bounded by a function of i

- for $i = O(\frac{\log \log n}{d})$ this probability is $O(1/n)$.

Assignment to bins on sender side - summary

- preliminary d bins for a user – chosen in a pseudorandom way
- fixing one out of d bins for a user – a sequential process

Assignment to bins - user's point of view

- preliminary d bins for A – computed easily with the secret key $s(A)$,
- determining the bin used by the sender for encoding the key for A – only by testing the keys derived

Complexity measures

Energy cost

- each user has to receive $n/B \cdot d$ ciphertexts
- the choice of ciphertexts – off-line

Transmission length

- theoretical value:

$$nk \cdot \left(1 + O\left(\frac{B \log \log B}{nd \ln \Phi_d}\right)\right) .$$

- the parameters B, d can be chosen freely except that $d \geq 2$.

Experimental values for practical parameter choice

- a sequence of 100 experiments
- number of users in a cell 10^6
- $B = 10^4$

d	n/B	max load	$\#$ of bins with load $> n/B$	B
1	100	145	4.764	10.000
2	100	103	3.109	10.000
4	100	**101**	1.322	10.000
10	100	**101**	539	10.000

Experimental values for practical parameter choice

- a sequence of 100 experiments
- number of users in a cell 10^4
- $B = 10^2$

d	n/B	max load	$\#$ of bins with load $> n/B$	B
1	100	142	54	100
2	100	102	38	100
4	100	**101**	17	100
10	100	**101**	8	100

Practical values - conclusion

- for $d = 4$ transmission size is practically $1.01 \cdot nk$
- even if something bad happens – the sender may change random parameter ρ

Conclusion

- substantial savings regarding energy usage with almost nk transmission size
- full anonymity

Open problems

- how to expel few users with short transmission, small energy use, and anonymity?
- previous tree based methods provide no privacy

thanks for your attention

http://kutylowski.im.pwr.wroc.pl