Approximation Algorithms for Key Management in Secure Multicast

A. Chan[1] R. Rajaraman[1] Z. Sun[1] F. Zhu[2]

[1]Department of Computer Science
Northeastern University

[2]Cisco Systems

COCOON, 2009

Outline

Outline

Motivation

- Publish-subscribe systems need to guarantee the privacy and authenticity of the participants.
  - Interactive gaming, stock data distribution, video conferencing, etc.
- Most systems rely on *symmetric key* cryptography to multicast messages.
  - We refer to key being used as *group key*.
- Any user should have access to the data only during the time periods that the user is a member of the group.
  - Need to update group key when set of group members changes.

Key update cost models

- Minimize the number of update messages sent.
  Motivation: consume minimum resources at the server.
- Minimize the total routing cost of update messages.
  Motivation: reduce network traffic.
- We consider both update models.

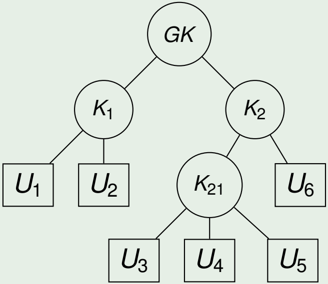
Key update approaches

- Naive approach: update one member at a time using his/her public key.
- Logical key hierarchy.
  - A single group key for data communication.
  - A group controller distribute *auxiliary subgroup key* to the group members according to a key hierarchy.
  - Each member stores auxiliary keys coresponding to all the nodes in the path to the root in the hierachy.

Example of a logical key hierarchy

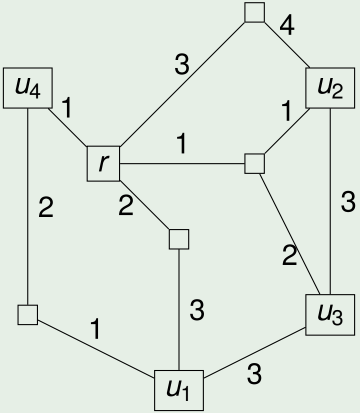
- GK is the group key.
- K's are auxiliary keys.
- Each user holds keys that lie along the path to the root.
  - U_3 has key GK, K_2, K_{21} and U_3's public key.
- When there is an update at a leaf, need to change group key.
  - View each leaf as a subgroup of users; whenever a user joins/leaves, an update occurs at the leaf.

LKH example

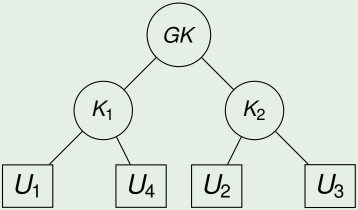


Example: routing cost of update messages

Routing network



Logical key hierarchy



If u_2 requests key update, the cost will be $2 + 3 + 4 + 4 = 13$.

Outline

Problem input

An instance of the Key Hierarchy Problem is given by the tuple (S, w, G, c).

- S is the set of group members.
- $w : S \to Z$ is the weight function (capturing the update probabilities).
- $G = (V, E)$ is the underlying communication network with $V \supseteq S \cup \{r\}$ where r is a distinguished node representing the group controller.
- $c : E \to Z$ gives the cost of the edges in G.

Cost of key hierarchy

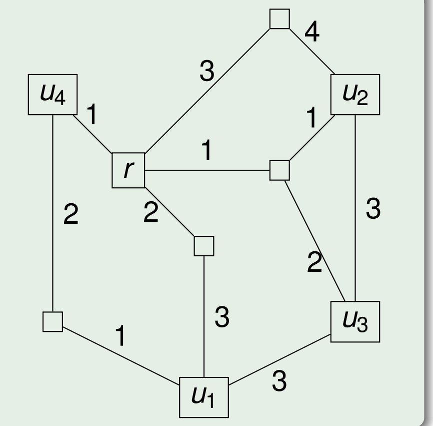
- A hierarchy on a set $X \subseteq S$ to be a rooted tree H whose leaves are the elements of X.
- Cost of a member x with respect to H is given by

$$\sum_{\text{ancestor } u \text{ of } x} \; \sum_{\text{child } v \text{ of } u} M(T_v)$$

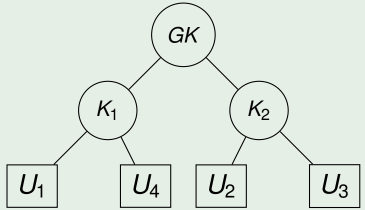
  - T_v is the set of leaves in the subtree of T rooted at v.
  - $M(Y)$ is the cost of multicasting from the root r to Y in G.
- Cost of a hierarchy H over X is the sum of the weighted costs of all the members of X with respect to H.

Illustrating routing cost

Routing structure



Logical key hierarchy

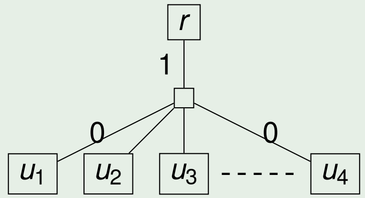


If u_2 requests key update, the cost will be $2 + 3 + 4 + 4 = 13$.

Uniform and non-uniform multicast model

- Minimizing the number of update messages is a special case of minimizing the routing cost of update messages.
- Refer minimizing the number of update messages as uniform multicast model.
- Refer minimizing the routing cost of update messages as nonuniform multicast model.

Uniform multicast routing graph



Outline

Results for uniform multicast model

- Identical update probabilities: We compute the optimal key hierarchy in polynomial time.
- General update probabilities: We give a PTAS (polynomial time approximation scheme).
  - Cost of this key hierarchy is within $1 + \epsilon$ times the cost of the optimal key hierarchy, where $\epsilon > 0$ and can be arbitrarily small.

Outline

Results for nonuniform multicast model

Hardness results:

- The Key Hierarchy Problem is NP-complete when group members have different weights and the routing network is a tree.
- The Key Hierarchy Problem is NP-complete when group members have the same weights and the routing netwrok is a general graph.

Approximation results:

- An 11-approximation algorithm when the routing network is a tree.
- A 75-approximation algorithm when the routing network is a general graph.

Outline

Divide and conquer

Lemma

For any instance, there exists a 3-approximate binary hierarchy.

So we can focus on finding a good binary key hierarchy.

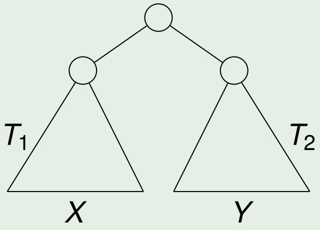
- Firstly, *partition* the member set into 2 subsets.
- Then find a "good" binary key hierarchy for each subset recursively.
- Lastly, *combine* these 2 binary key hierarchies.

Keys of partitioning:

- Make close users "close" in the hierarchy.
- Balance the weight of binary hierarchy.

Combine logical key hierarchies

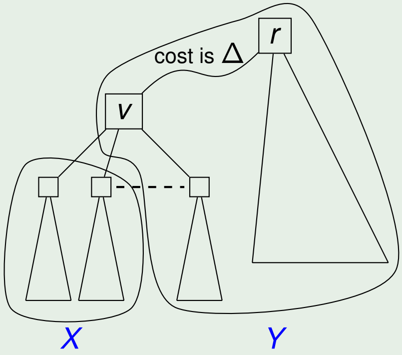
Binary key hierarchy T



- Let T_1 be a "good" binary hierarchy for member set X.
- Let T_2 be a "good" binary hierarchy for member set Y.
- Define combine(T_1, T_2) to be the following. Add a new root r, and make T_1 the left subtree, T_2 the right subtree.

Partition member set

Partition procedure



- Assume the routing network is a tree, controller is the root, and members are the leaves.
- $W(S)/3 \leq W(X), W(Y) \leq 2W(S)/3$, where $S = X \cup Y$ and $W(\cdot)$ is the total weight of the members in the set.

Light approximate shortest-path tree (LAST)

Our approximation algorithm uses the elegant algorithm of Khuller-Raghavachari-Young for finding spanning trees that simultaneously approximates both the minimum spanning tree and the shortest path tree. An (α, β)-LAST of a given weighted graph G is a spanning tree T of G, such that

- shortest path in T from root to any vertex is at most α times the shortest path from the root to the vertex in G,
- total weight of T is at most β times the minimum spanning tree of G.

Approximating the multicast cost

- If the routing network is a graph, the optimum multicast to a member set is obtained by a minimum Steiner tree, computing which is NP-hard.
- There is an easy 2-approximation algorithm using a minimum spanning tree (MST) in the metric space defined by the routing graph.
- So we approximate $M(Y)$ by the cost of MST connecting the root r to Y in the complete graph $G(Y)$ whose vertex set is $S \cup \{r\}$ and the weight of edge (u, v) is the shortest path distance between u and v in the routing graph G.

Outline

ApproxGraph(S)

- If S is singleton, return trivial hierarchy with one node.
- Compute complete graph on $S \cup \{root\}$; weight of (u, v) is the length of shortest path between u and v in the original routing graph.
- Compute minimum spanning tree on this complete graph.
- Compute an (α, β)-LAST L of MST(S).
- $(X, v) =$ **partition**(L).
- Let Δ be the cost from root to partition node v. If $\Delta \leq M(S)/5$, $T_1 =$ ApproxGraph(X). Otherwise, $T_1 =$ PTAS(X). $T_2 =$ ApproxGraph(Y).
- $T_2 =$ ApproxGraph(Y).
- Return **combine**(T_1, T_2).

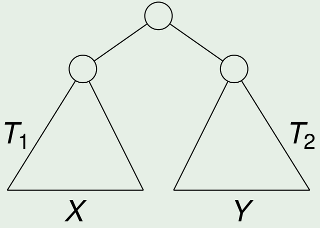
Proof sketch of constant approximation ratio

Theorem

*Algorithm **ApproxGraph** is a constant-factor approximation.*

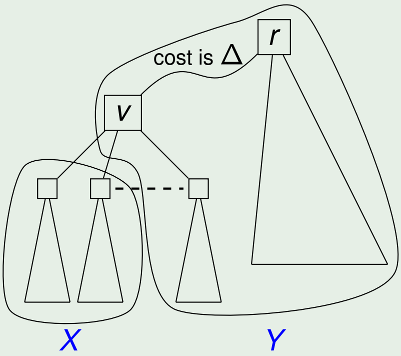
Proof uses induction on the number of members in S.

Binary key hierarchy T



- $ALG(S)$ cost of hierarchy produced by ApproxGraph.
- $OPT(S)$ cost of optimal hierarchy.
- $ALG(S) = ALG(X) + ALG(Y) + W(S)\,[M(X) + M(Y)]$.
- $OPT(S) \geq OPT(X) + OPT(Y)$.

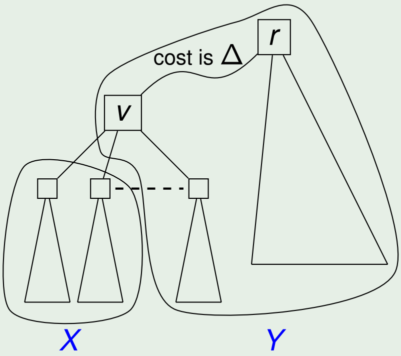
(α, β)-LAST of routing network



Case 1: $\Delta > M(S)/5$

- Distance from r to any elem in X is bigger than Δ.
- This distance is close to shortest path in the original graph.
- Multicast cost to any subset of X is “roughly” the same. Use PTAS to get better approx on $ALG(X)$.
- Apply induction hypothesis on Y.

(α, β)-LAST of routing network



Case 2: $\Delta \leq M(S)/5$

- Apply induction hypothesis on both X and Y.

Open problems

- Hardness result for uniform multicast cost but non-uniform key update probabilities.
- Dynamic maintenance of key hierarchies when members change update probabilities.
- Design key hierarchies where members have a bound on the number of auxiliary keys they store.