An Efficient Similarity-Based Approach for Optimal Mining of Role Hierarchy

Hassan Takabi and James Joshi {hatakabi, jjoshi}@sis.pitt.edu
Laboratory for Education and Research on Security Assured Information Systems (LERSAIS), School of Information Sciences, University of Pittsburgh

Motivation and Background

Motivation

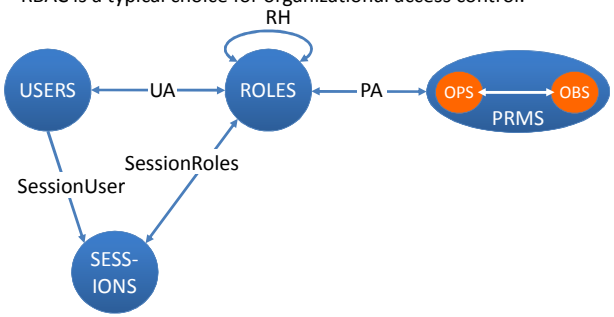
- Deploy of RBAC requires *identification* of a complete set of roles.
- ***Role engineering*** - one of the costliest tasks in migrating to RBAC.
- Most of the existing approaches do not consider roles that already exist.

Contributions

- Formally define the problem of mining role hierarchy with minimal perturbation
- Present a heuristic algorithm to find an RBAC state as similar as possible to the existing state and the optimal state

Background

Role Based Access control (RBAC) - a high level authorization model in which access decisions are based on the roles that users hold within an organization.
RBAC is a typical choice for organizational access control.



- RBAC State: $\gamma = \langle R, UA, PA, RH \rangle$
- System Configuration: $\rho = \langle U, P, UP \rangle$

•Role Engineering:

- The top-down approach: analysis of business processes
- The bottom-up approach: data mining techniques
- The hybrid approach

•Formal Concept Analysis

A formal context: triple *(G,M,I)*

- *objects, attributes* $gIm : (g, m) \in I$

Concept: *(X, Y)*

- *Y* : the set of all attributes shared by all objects in *X*
- *X* : the set of all objects that share all attributes in *Y*

concept lattice : the concepts complies mathematical axioms defining a lattice.
reduced concept lattice : the result of removing redundant attributes and objects.

The Problem of Mining Role Hierarchy with Minimal Perturbation

The ***reduced concept lattice*** defines a complete RBAC state.

- Each concept represents a role
- The lattice can be viewed as the role hierarchy
- The sub-concept relation corresponds to the role inheritance relation.

We need a measure to compare the different role hierarchies and identify which one is more desirable.
We define two different measure:

A measure for Goodness of an RBAC State

Given $W = \langle w_r, w_u, w_p, w_h \rangle$ where $w_r, w_u, w_p, w_h \in Q^+ \cup \{\infty\}$, the weighted structural complexity of an RBAC state is defined as follows:

$$wsc(\gamma, W) = w_r * |R| + w_u * |UA| + w_p * |PA + w_h * |t\_r(RH)|$$

Where
Q^+: is the set of all non-negative rational numbers,
$|.|$: the size of the set or relation
t_r(RH): the transitive reduction of role-hierarchy.

A Measure for Minimal Perturbation

The similarity between two roles:

Three different similarity measures

- *Permission centric*
- *User centric*
- *Hierarchy relation centric*

The Role-Role Similarity $sim(r_1, r_2)$ is defined by combining all these measures with adjustable weights.

The similarity between two role sets rs_1 and rs_2 :
$\forall r_i \in rs_1$, find $Max_{r_j \in rs_2} sim(r_i, r_j)$ such that for all selected pairs (r_i, r_j) and (r_x, r_y), if $r_x \neq r_y$ then $r_i \neq r_j$.

- In this step every role in rs_1 is matched with exactly one distinct role in rs_2
- There are some roles in rs_2 that have not been matched with any role from rs_1 .
- Define a threshold *t* and consider only roles that have a similarity measure above the threshold.
- Take average over all of chosen similarities.

The Problem of Mining Role Hierarchy with Minimal Perturbation

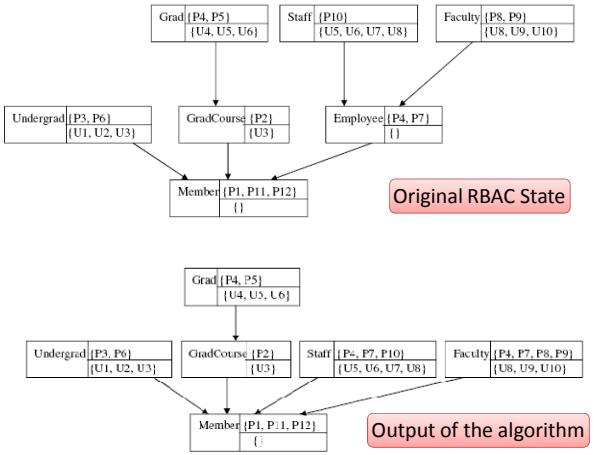
The goal is to minimize the global optimization function of the predefined optimality measure (i.e., weighted structural complexity) and the predefined perturbation measure (i.e., dissimilarity measure).

The Proposed Algorithm

The Proposed Algorithm

- First phase- generate the reduced concept lattice
- Second phase- prune this lattice and select the final RBAC state.
- The greedy algorithm prunes the reduced concept lattice based on combination function that defined above.

Example



Results

- Generates significantly fewer roles than the original state.
- Smaller *wsc* than the HierarchicalMiner and closer to the optimal solution.
- Provides better results compared to VAG algorithm.

Challenges and Directions

- **Parameterized roles**
- **Separation of Duty Constraints**
- **Roles with semantic meanings**
- **Role mining in multi-domain environments**

Future Work

- Evaluate the proposed algorithm using real data, and comparing it with the existing approaches.
- Propose an approach that considers separation of duty constrains and its effects on process of migrating to RBAC.