❑   1394

A Conceptual Model of Role Based Access Control Using Role Mining Algorithm

Nazirah Abd Hamid, Rabiah Ahmad, Siti Rahayu Selamat
Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Malaysia
Faculty of Information & Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia

Article Info	ABSTRACT
	Numerous studies have shown that currently, role-based access control has becoming one of the successful access control model because of its principle that could simplifies the work of security administrators. However, to construct a concise, role-based access control system, a good role mining algorithm structure is needed therefore the objectives of this paper are firstly, to provide a general overview on phases that involved in designing and developing the algorithm and secondly, to introduce a conceptual model that constructed based on the analysis and this model represents a general process in role mining model. This model involved series of phases that begin with the input of data, pre-processing stage, candidate role generation phase, role selection and role assignment process and lastly number of roles as generated output.

Corresponding Author:

Nazirah Abd Hamid,
Faculty of Informatics and Computing,
Universiti Sultan Zainal Abidin, Malaysia.
Email: owenira@gmail.com

1.   INTRODUCTION

Role-Based Access Control (RBAC) has demonstrated to be a proper access control model to manage authorizations aspect especially for the security administration due to its flexibility and ability to capture an organization's structure and objectives. According to NIST [1], RBAC is accepted as a security standard in numerous domains, such as Industrial, Military and Healthcare. Role engineering has been introduced by [2] and has been applied to define a requisite and correct set of roles and permissions and role mining is a concept in the role engineering that popular among the researchers due to the nature of applying computing-intensive approaches that could decrease the cost of maintaining the security features and also simplifies the work of security administrators.

The objectives of this paper are firstly, to analyze and classify on some of the present role mining algorithms from 2013 to 2017 and then provide a general overview on phases or stages that involved in designing and developing them and secondly, to propose a conceptual model that constructed based on the analysis of the aforementioned phases.

The remainder of the paper is structured as follows. The analysis of the background study is shown in Section 2. Section 3 presents a mathematical background of this study while Section 4 introduces the general process in role mining model. Lastly, Section 5 discusses the conclusions and the future work that can lead to further enhancement of this field.

2.   LITERATURE REVIEW

Numerous studies have reported that presently, role-based access control (RBAC) has becoming the predominant access control model because of it principle that could significantly simplifies the work of

security administrators [3-6]. The abovementioned principle of RBAC could be defined as every role is a group of permissions, and each user obtains the permissions only through the roles.

According to Ye et al. [4], RBAC system could be implemented through two approaches specifically the top-down and the bottom-up method. The authors have explained that the top-down approach builds a RBAC system with the involvement of experts' analysis on the business processes yet, this approach consumes a lot of time because of human participation [7]. The bottom-up approach, according to Hu et al. [8] can uncover roles from the existing user-permission assignments (UPA) automatically that is known as role mining and because of its nature that based on computing-intensive approach, it is widely applied to build a RBAC model.

However, to build and sustain a RBAC model, role mining is becoming a great interest [9-10] and the authors have identified the need of role mining to design and develop an algorithm to determine roles based on data mining methods because it could reduce the cost of allocating roles manually thus able to construct a concise RBAC system. The next section would provide in-depth analysis on methodology to build a RBAC model using role mining algorithm.

2.1. Role Mining Model

In general, Fuchs and Meier [11] has introduced a general Role Mining Process Model as in Figure 1 and the author has described the phases in Figure 1 as the following:
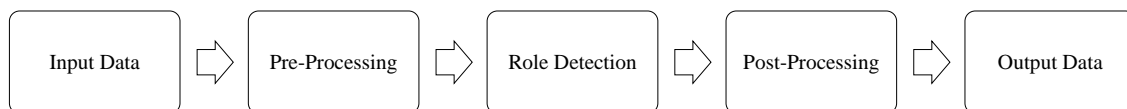

Figure 1. Role mining process model

2.1.1 Input Data

In most of the role mining algorithms, user-permission assignment (UPA) matrix can be considered as input data, correspondingly, an active research should be done to explore the possibility of other types of data to be used in role mining process.

2.1.2 Pre-processing

Many researchers have emphasized on the importance of this stage [12-13] particularly to generate a clean and quality data and usually pre-processing stage involve the process to clean the noises that might affect the results.

2.1.3 Role Detection

This stage is significant in the role mining process model because it involves the discovery of appropriate candidate roles from obtainable set of inputs preferably a clean data using data mining techniques or heuristics algorithms called as role mining algorithm.

2.1.4 Post-processing

In this stage, the acquired candidate roles from the previous stage are being selected and assigned optimally by using yet other suitable algorithms.

2.1.5 Ouput Data

The output of these processes is normally a set of roles and a RBAC state such as hierarchy or involving constraints.

2.2. Role Mining Algorithm Phases

Numerous studies have proposed role mining algorithms that could solve Role Mining Problem (RMP) in Role Based Access Control (RBAC) system and this following section would analyze and classify on some of the present role mining algorithms from 2013 to 2017 according to the phases that stated in 2.1 section and the detail is presented in Table 1.

Table 1. Role mining algorithm phases

No	Authors (Year) & Title	Candidate Roles (CR) Phase and Role Selection (RS) & Assignment Phase (AP)	Input & Output Data
1.	14] An Efficiency Approach for RBAC Reconfiguration with Minimal Roles and Perturbation	CR: generate the candidate roles using FastMiner algorithm [15] RS & AP: select a subset from CR that minimizes the perturbation	Input: access history log & UPA matrix Output: GR as the generated roles and QR as the qualified roles
2.	[16] An Approach for Hierarchical RBAC Reconfiguration with Minimal Perturbation	CR: generate the candidate roles using FastMiner algorithm [15] RS & AP: select the similar roles which is the most similar the set of qualified roles	Input: access history log & UPA matrix Output: hierarchical RBAC state (RH)
3.	[10] The RBAC System Based on Role Risk and User Trust	CR: clustering roles based on risk RS & AP: for each role identify the permission (P) > assigned trust threshold	Phase 1: clustering roles based on risk Input: UPA matrix Output: RBAC roles & RH stable Phase 2: trust Input: RBAC roles Output: roles
4.	[17] Mutual Exclusion Role Constraint Mining Based on Weight in Role-Based Access Control System	CR: generate the candidate permission sets based on weight RS & AP: generate all combinations of permission sets whose weighted support is greater than the user specified minimum weighted	Input: UPA matrix Output: roles
5.	[18] Scott D. Stoller and Thang Bui (2016) Mining Hierarchical Temporal Roles with Multiple Metrics	CR: generates initial roles and then creates additional candidate roles by intersecting sets of initial roles using FastMiner [15] RS & AP: Construct role hierarchy	Input: ACL policy Output: role hierarchy
6.	[19] Mining Approximate Roles under Important Assignment	CR: UPA is decomposed into two assignments, NUPA & IUPA RS & AP: (1) NUPA is processed by the δ -Approx RM algorithm and generates NRoles (2) IUPA is processed by any algorithm to generate IRoles	Input: UPA matrix Output: roles
7.	[20] Mining Temporal Roles using Many-Valued Concepts	CR: construct the concepts only for a pre-determined value of θ (many-value concept). RS & AP: maximum area of coverage is selected and is added to the final set of concepts	Input: TUPA matrix Output: UA, PA & REB
8.	[21] Towards User-oriented RBAC Model	CR: three different ways of generating candidate roles (1) itself (2) intersection (3) association RS & AP: enforce maximal role assignment constraints (1) greedy (2) fewest (3) most (4) rand	Input: UPA matrix Output: UA & PA
9.	[22] The Generalized Temporal Role Mining Problem	CR: creation of candidate role set by taking union of the sets of units, initial and generated roles RS & AP: iterative selection of a minimal cardinality subset of the candidate role set using any one of the four greedy heuristics	Input: TUPA matrix Output: R, UA, PA & REB
10.	[23] Role Mining Based on Permission Cardinality Constraint and User Cardinality Constraint	CR: initial role: (1) one is from the prerequisite role set (2) the initial role set generation algorithm RS & AP: role selection algorithm & role state generation algorithm	Input: UPA matrix Output: UA & PA
11.	[7] Role Mining based on Cardinality Constraints	CR: generating the initial role set based cardinality constraints of roles and permissions RS & AP: selecting role pair for role update algorithm (hierarchical relationships) & updating the initial role state (graph optimization algorithm)	Input: UPA matrix Output: roles

12.	[24] Migrating from DAC to RBAC	CR: each iteration uses the DEMiner algorithm to generates a candidate role RS & AP: at each iteration, the new candidate role is intersected with the roles in R and UPA are performed at each iteration to reflect the updates in R.	Input: UPA matrix Output: UA & roles
13.	[5] Handling Least Privilege Problem and Role Mining in RBAC	CR: (1) The first kind of candidate roles set (FCR) (2) The second kind of candidate roles set (SCR) RS & AP: least privileges principle	Input: a set of users, a set of privileges, and a set of user-privilege assignment relation Output: roles
14.	[25] Performance of AI Algorithms for Mining Meaningful Roles	CR: generating roles using elimination algorithm RS & AP: in each generation, the elitist selection scheme is applied to guarantee that the fittest member of each generation is copied directly into the next generation (GA)	Input: RH Output: roles
15.	[26] An Optimization Framework for Role Mining	CR: three different ways of generating candidate roles (1) itself (2) intersection (3) association RS & AP: greedy algorithm	Input: UPA Output: UA & PA
16.	[27] Visual Elicitation of Roles: using A Hybrid Approach	CR: Random Data Generator (RDG) RS & AP: Matrix sorting algorithm	Input: UPA Output: role sets
17.	[28] Toward Mining of Temporal Roles	CR: enumerates the set of candidate roles from an input TUPA matrix Role Selection and assignment: elects the least possible number of roles from the candidate roles using a greedy heuristic	Input: TUPA matrix Output: UA, PA & REB
18.	[3] Role Mining Using Boolean Matrix Decomposition with Hierarchy	CR: the candidate roles through formal concept analysis RS & AP: redundant roles can be removed according to cost-utility analysis	Input: UPA matrix Output: UA, PA, RH, UA′ & PA′ matrix
19.	[29] Mining Parameterized Role-based Policies	CR: use CompleteMiner [15] to generate candidate roles. RS & AP: (1) It selects roles from highest quality to lowest (2) Compute Role Hierarchy (3)	Input: UPA matrix Output: UA, PA & RH
20.	[30] Evolving role definitions through permission invocation patterns	CR: which are selected to optimize an objective function that balances distance from the original roles with behaviorial similarity in the form of permission RS & AP: assigned to roles according to a criterion that mitigates redundancy	Input: access history log Output: roles

3.   MATHEMATICAL BACKGROUND

This section presents some of the formal definitions that related to the Role Based Access Control (RBAC) as well as Role Mining Problem (RMP) and its variants and some of terms are associated to the conceptual model.

Definition 1. (RBAC Model)

The RBAC model has the following basic elements [14]:
a)   U, R and P are signifying the set of users, roles and permissions.
b) $UA \subseteq U \times R$ is representing the user-role assignments.
c) $PA \subseteq P \times R$ is defining the role-permission assignments.
d) $UPA \subseteq U \times P$ is the user-permission assignments.
e) $RH \subseteq R \times R$, a partial order on roles described the inheritance relationships.

Definition 2. (RBAC state)

RBAC state could be expressed as <R, UA, PA, RH> that is consistent with an access control configuration $\rho = <U, P, UPA>$, where U defines a set of all users, P is a set of all permissions and $UPA \subseteq U \times P$ as the user-permission relation [10].

Definition 3. (Basic RMP)

Assumed a set of users (U), a set of permissions (P) and a user-permission assignment (UPA) is given, acquire a set of roles (R), a user-role assignment (UA) and a role-permission assignment (PA) while reducing the number of roles |R| = k. In matrix notation, it can be formulated as [12]:

$$\|UA \otimes PA - UPA\|1 = 0 \tag{1}$$

Definition 4. (δ-Approx RMP)

δ-approx RMP is a part of Basic-RMP that allows a partial match between the user-permission assignment (UPA) and the generated user-role assignment (UA) and a role-permission assignment (PA) and can occasionally decrease the total number of roles, k, substantially. It can be formulated in matrix representation, such that [31]:

$$\|UA \otimes PA - UPA\|1 \leq \delta \tag{2}$$

Definition 5. (MinNoise RMP)

For MinNoise RMP, the number of roles (k) is bounded so that the number of mismatches between the UPA and the generated UPA is minimized. So, presumed a set of users (U), a set of permissions (P), a user-permission assignment (UPA) and a number of roles (k) is given, discover a set of k roles (R), a user-role assignment UA and a role-permission assignment PA by minimizing [31]:

$$\| UA \otimes PA - UPA \|1 \tag{3}$$

Definition 6. (User-Permission Assignment)

The user-permission assignment (UPA) matrix is an m × n binary matrix UPA, m is representing the number of users, while n can be defined as the number of permissions. The element UPA (i, j) = 1 indicates the assignment of permission j to user i [14].

Definition 7. (Access History Log)

Access history log is a series of quaternion (U, P, R, t) and this series indicates an access event in the system and represents the user (U) invoke the permission (P) by activating the role (R) at the time t [16].

4.   CONCEPTUAL MODEL

A conceptual model determines a comprehensive understanding and scopes of a proposed solution using the organized concepts that are linked together [32] and for this paper, the conceptual model is constructed based on literature review in Section 2 and this model represents a general process in role mining model as shown in Figure 2.
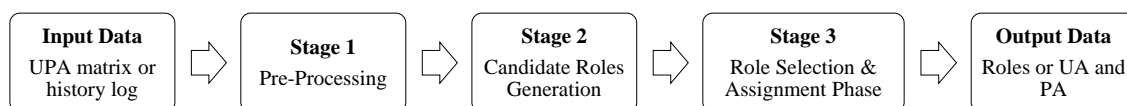

Figure 2. General process in role mining model

The components of the model and the relationships between them are described as follows. The process starts by inserting input data and as listed in Table 1, most of the role mining algorithms are utilizing variants of user-permission assignment (UPA) matrix as an input data and some of the researchers also using access history log. Then the data is transferred to pre-processing stage and the general activity in this stage

includes the data cleaning and data normalization specifically the acts of removing noise and handling missing data and the purpose of this stage is to create a set of correct data for the next process.

Next, the clean data is advanced to the next stage namely candidate roles generation phase. This stage is the most meaningful process because it involves the discovery of appropriate candidate roles by exploiting any suitable data mining techniques or heuristics algorithms or known as role mining algorithms. This stage usually produce a big pool of candidate roles, therefore in the next step, role selection and role assignment phase, more specific and smaller roles are produced according to the desired objectives and outputs by utilizing any appropriate data or role mining algorithms. Lastly, as mentioned before and based on information in Table 1, the outputs could be varied but most of the researchers prefer number of roles as the main output.

5. CONCLUSION AND FUTURE WORKS

We have proposed a conceptual model that is constructed based on the literature review and this model represents a general process in role mining model. This model involves series of phases that begin with the input of data, pre-processing stage, candidate role generation phase, role selection and role assignment process and lastly number of roles as generated output. For the future works, we intend to improve the conceptual model with a comprehensive and complete model.

REFERENCES

[1] "Role Based Access Control, "National Institute of Standards and Technology", 2016. [Online]. Available: https://csrc.nist.gov/projects/role-based-access-control#rbac-standard.
[2] E. J. Coyne, "Role Engineering," in *Proceedings of the first ACM Workshop on Role-based access control*, 1996, no. 4, pp. 15–16.
[3] W. Ye, R. Li, and H. Li, "Role Mining Using Boolean Matrix Decomposition with Hierarchy," in *12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, 2013, pp. 805–812.
[4] W. Ye, R. Li, X. Gu, Y. Li, and K. Wen, "Role Mining using Answer Set Programming," *Futur. Gener. Comput. Syst.*, vol. 55, pp. 336–343, Feb. 2016.
[5] H. Huang, F. Shang, J. Liu, and H. Du, "Handling Least Privilege Problem and Role Mining in RBAC," *J. Comb. Optim.*, vol. 30, no. 1, pp. 63–86, Jul. 2015.
[6] Y. R. More and S. V Gumaste, "Performance Evaluation of A Role Based Access Control Constraints in Role Mining Using Cardinality," *Int. J. Adv. Res. Sci. Manag. Technol.*, vol. 2, no. 7, pp. 1–7, 2016.
[7] R. Li, H. Li, X. Gu, Y. Li, W. Ye, and X. Ma, "Role Mining based on Cardinality Constraints," *Concurr. Comput. Pract. Exp.*, vol. 27, pp. 3126–3144, 2015.
[8] J. Hu, K. M. Khan, Y. Zhang, Y. Bai, and R. Li, "Role Updating in Information Systems using Model Checking," *Knowl. Inf. Syst.*, vol. 51, no. 1, pp. 187–234, Apr. 2017.
[9] M. Frank, J. M. Buhmann, and D. Basin, "On the Definition of Role Mining," in *Proceeding of the 15th ACM symposium on Access control models and technologies - SACMAT '10*, 2010, p. 35.
[10] C. Jin, A. Shen, and W. Yu, "The RBAC System Based on Role Risk and User Trust," *Int. J. Comput. Commun. Eng.*, vol. 5, no. 5, pp. 374–380, 2016.
[11] L. Fuchs and S. Meier, "The Role Mining Process Model - Underlining the Need for a Comprehensive Research Perspective," in *2011 Sixth International Conference on Availability, Reliability and Security*, 2011, pp. 35–42.
[12] S. Vavilis, A. I. Egner, M. Petkovic, and N. Zannone, "Role Mining with Missing Values," in *11th International Conference on Availability, Reliability and Security (ARES)*, 2016, pp. 167–176.
[13] J. Lu and Q. Zhu, "An Effective Algorithm Based on Density Clustering Framework," *IEEE Access*, vol. 5, pp. 4991–5000, 2017.
[14] N. Pan, Z. Zhu, L. He, and L. Sun, "An Efficiency Approach for RBAC Reconfiguration with Minimal Roles and Perturbation," *Concurr. Comput. Pract. Exp.*, no. December, pp. 1–15, Dec. 2017.
[15] J. Vaidya, V. Atluri, and J. Warner, "RoleMiner: Mining Roles using Subset Enumeration," in *Proceedings of the 13th ACM conference on Computer and communications security - CCS '06*, 2006, no. January, p. 144.
[16] N. Pan, L. Sun, L. He, and Z. Zhu, "An Approach for Hierarchical RBAC Reconfiguration with Minimal Perturbation," *IEEE Access*, vol. 2169–3536, no. c, pp. 1–11, 2017.
[17] X. Ma, J. Wang, L. Zhao, and R. Li, "Mutual Exclusion Role Constraint Mining based on Weight in Role-Based Access Control System," *Int. J. Innov. Comput. Inf. Control*, vol. 12, no. 1, pp. 91–101, 2016.

[18] S. D. Stoller and T. Bui, "Mining Hierarchical Temporal Roles with Multiple Metrics," in *IFIP Annual Conference on Data and Applications Security and Privacy*, 2016, pp. 79–95.

[19] N. Pan, Z. Zhu, L. He, L. Sun, and H. Su, "Mining Approximate Roles under Important Assignment," in *2nd IEEE International Conference on Computer and Communications*, 2016, pp. 1319–1324.

[20] B. Mitra, S. Sural, J. Vaidya, and V. Atluri, "Mining Temporal Roles using Many-Valued Concepts," *Comput. Secur.*, vol. 60, pp. 79–94, Jul. 2016.

[21] H. Lu, Y. Hong, Y. Yang, L. Duan, and N. Badar, "Towards User-oriented RBAC Model," *J. Comput. Secur.*, vol. 23, no. 1, pp. 107–129, Mar. 2015.

[22] B. Mitra, S. Sural, V. Atluri, and J. Vaidya, "The Generalized Temporal Role Mining Problem," *J. Comput. Secur.*, vol. 23, no. 1, pp. 31–58, 2015.

[23] X. Ma, R. Li, H. Wang, and H. Li, "Role Mining based on Permission Cardinality Constraint and User Cardinality Constraint," *Secur. Commun. Networks*, vol. 8, no. 13, pp. 2317–2328, 2015.

[24] E. Uzun, D. Lorenzi, V. Atluri, J. Vaidya, and S. Sural, "Migrating from DAC to RBAC," in *Lecture Notes in Computer Science*, vol. 9149, P. Samarati, Ed. Cham: Springer International Publishing, 2015, pp. 69–84.

[25] X. Du and X. Chang, "Performance of AI Algorithms for Mining Meaningful Roles," in *2014 IEEE Congress on Evolutionary Computation (CEC)*, 2014, pp. 2070–2076.

[26] H. Lu, J. Vaidya, and V. Atluri, "An Optimization Framework for Role Mining," *J. Comput. Secur.*, vol. 22, no. 1, pp. 1–31, Jan. 2014.

[27] A. A. Eucharista and K. Haribaskar, "Visual Elicitation of Roles : using A Hybrid Approach," *Orient. J. Comput. Sci. Technol.*, vol. 6, no. 1, pp. 103–110, 2013.

[28] B. Mitra, S. Sural, V. Atluri, and J. Vaidya, "Toward Mining of Temporal Roles," in *Lecture Notes in Computer Science*, vol. 7964 LNCS, IFIP International Federation for Information Processing, 2013, pp. 65–80.

[29] Z. Xu and S. D. Stoller, "Mining Parameterized Role-based Policies," in *Data and application security and privacy*, 2013, p. 255.

[30] W. Zhang, Y. Chen, C. Gunter, D. Liebovitz, and B. Malin, "Evolving Role Definitions Through Permission Invocation Patterns," in *Proceedings of the 18th ACM symposium on Access control models and technologies - SACMAT '13*, 2013, no. June, p. 37.

[31] B. Mitra, S. Sural, J. Vaidya, and V. Atluri, "A Survey of Role Mining," *ACM Comput. Surv.*, vol. 48, no. 4, pp. 1–37, 2016.

[32] Y. Jabareen, "Building a Conceptual Framework: Philosophy, Definitions, and Procedure," *Int. J. Qual. Methods*, vol. 8, no. 4, pp. 49–62, Mar. 2009.