ORIENTAL JOURNAL OF
COMPUTER SCIENCE & TECHNOLOGY
An International Open Free Access, Peer Reviewed Research Journal
Published By: **Oriental Scientific Publishing Co., India.**
www.computerscijournal.org

ISSN: 0974-6471
March 2013,
Vol. 6, No. (1):
Pgs. 103-110

Visual Elicitation of Roles: using A Hybrid Approach

A. AROCKIA EUCHARISTA and K. HARIBASKAR

[1]Mount Zion College of Engineering and Technology,
Computer Science and Engineering, Pudukottai, India

(Received: February 15, 2013; Accepted: March 01, 2013)

ABSTRACT

Access control is the process of mediating requests to data and services maintained a system, determining which requests should be granted or denied. Significant research has focused on providing formal representation of access control models. Role Based Access Control (RBAC) has become the norm in most organizations. This success is greatly due to its simplicity: a role identifies a set of Permissions; users in turn are assigned to roles based on their responsibilities. To implement a RBAC system, it is important to devise a complete set of roles. This design task, known as role engineering, has been recognized as the costliest part of a RBAC – oriented project. We propose a new role engineering approach to Role – Based Access Control (RBAC) referred to as visual role mining. The main aim is to graphically represent user – permission assignments to enable quick analysis and elicitation of meaningful roles. We propose two algorithms: VISRODE (VISualize Roles using DicE) and EXTRACT (Exception Tolerant Role ACTualizer). A heuristic algorithm VISRODE is used to sort the users and permissions matrix to avoid the large gaps between items using DicE coefficient. EXTRACT is a probabilistic algorithm and it generates a list of pseudo roles. This paper offers a graphical way to effectively navigate the result so that it reduces the time complexity in visualizing the roles.

Key words: Datasets, Matrix Sorting, Role Engineering, Role mining Algorithms, Pseudo Roles, Visual Mining.

INTRODUCTION

Access Control is the process of mediating requests to data and services maintained by a system, determining which requests should be granted or denied. Significant research has focused on providing formal representation of access control models. it is an important component of IAM (Identity and Access Management). Role Based Access Control (RBAC) is not the perfect solution, it enables greater shared responsibility and more effective and efficient permissions management for IT and business operations. RBAC facilities and relative to other approaches, reduces costs associated with governance, risk, and compliance (GRC) activities and through greater visibility of permissions assigned to users and easier verification of internal controls; access control policy maintenance, attestation of access control policies in place, certification of regulated information systems, and access control policy audits conducted by internal and external auditors. The proposals and adopted standard largely

eliminated the uncertainty and confusion about RBAC's utility and definition; it has served as a foundation for software product development, evaluation, and procurement specifications. Permissions sometimes referred to as privileges or entitlements; specify what operations a user may perform on a specific object. Typical operations include read, write, delete and execute, or complex transactions such as a money transfer. Users can create new tables, add new information to existing tables, or modify information that exists in the tables. For many applications, there is only a single permission allowing a user to execute the application. The task of assigning, terminating and modifying user's permissions is referred to as provisioning. When a new user joins an organization, he must be given all of the permissions necessary to perform his job. The process by which access permissions are removed from user is referred to as "de provisioning". When a user leaves the organization all permissions must be terminated. To understand what access control policy is in place, an organization must be able to review user's permissions. Acquiring this information requires time by IT administrators, human resources (HR), and management. Many systems provide a means of placing users in to one or more groups, with permissions attached to both a group and individual users within the group.

Rather than assigning permissions directly to users, under RBAC, permissions are assigned to roles engineered in software systems and users are assigned the roles necessary to do their jobs. Permissions can be grouped in to roles based on location, business function, department, or other attributes of users. Already in the 1st ACM workshop on RBAC in 1995, Edward Coyne[11] keen out that "Role Engineering entails defining the roles that will determine which employees have access to which data and to which applications, as well as roles' relationships to one another, role hierarchy, and role constraints". There are top down engineering, bottom up engineering, and by example. The design of role mining algorithms, existing techniques deal with three main practical issues: Meaning: Organizations are unwilling to deploy roles they cannot fully understand. Noise: Exceptionally or accidentally granted permissions can hinder the role mining task proposed noise models do not always fit real cases, especially when exceptions are legitimate and cannot be avoided. Correlations: The identification of relationships among roles (e.g., similarities, permission set inclusion, etc.) can further ease the identification of roles.

To overcome all issues we propose a new approach, referred to as Visual Role Mining. The following figure 1 shows the concepts of extracting the role from the user permission relations.

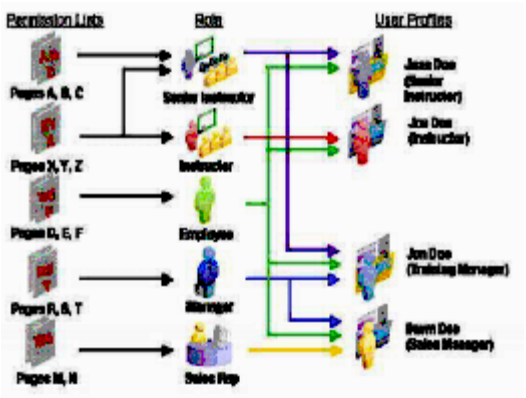

Fig. 1: User inherits permissions through roles

Finally, correlations among roles are shown as overlapping patterns, hence providing an intuitive way to discover and utilize these relations. Our work shows that a proper representation of user – permission assignments allows role designers to gain insight, draw conclusions, and design meaningful roles from both IT and business perspectives. We propose two algorithms called VISRODE DICE coefficient is a similarity measure used to avoid large gaps between items. Finally, a probabilistic algorithm EXTRACT[1] is used to generate a list of pseudo roles. We offer a graphical way to effectively navigate the result.

Related work

Coyne[11] was the first to propose the role engineering problem and the top – down approach to role engineering. Several subsequent papers[12-14] focused on the top down approach is generally very expensive as it is human intensive and requires the collaboration of security experts and domain experts to extract the knowledge of

business process descriptions. The role mining approach was first proposed by Kuhlmann et al.[15] in 2003. Vaidya et al.,[16] randomly generate RBAC states and mine the flattened UP relationship two variants of RMP were later introduced. In[6], Zhang et al. suggest minimizing the size of each relation (W=(0,1,1,1, “)), and the size of the RBAC representation Molloy et al.[18] propose the notion of weighted Structural Complexity (WSC) that subsumes many the above metrics and is used in this work. Colantanio et al.,[3] describe a measure similar to WSC with an additional abstract cost function c: R → R, where R is the set of all possible roles, which allows an administrator to increase or decrease the cost associated with a role based on its desirability, such as underlying business processes or semantic meaning. Without domain knowledge, c maps all values to a constant. Several works[16,17] allow for solutions that grant users permissions not granted in UP. These are typically modeled as δ – consistent solutions for δ > 0. We see several problems with δ– approximate role – mining, primarily stemming from its treatment of over and under assignments identically. To address[1] all the issues, this paper proposes a methodology that helps role engineers to leverage business information during the role mining process.

In[18] this work we allow under assignment only by allowing direct user – permission assignments. These are additional permission assignments [4] that are required in addition to the mined RBAC state to maintain 0-consistency with the input. Our approach is partially inspired by [19]. Our approach is greatly differs from[20], first, we adopt a different visualization cost metric that is more suitable for role engineering[21] incompatible with the core of their theory second, we show hoe to obtain a matrix representation without resorting to any existing mining algorithm.

Visual elicitation

This section addresses the following problem[1]: given a set already discovered roles of interest, we want to identify the best graphical representation for them. In particular, we want the representation for user-permission assignments[1] that allows for both an intuitive role validation and a visual identification of the relationships among roles.

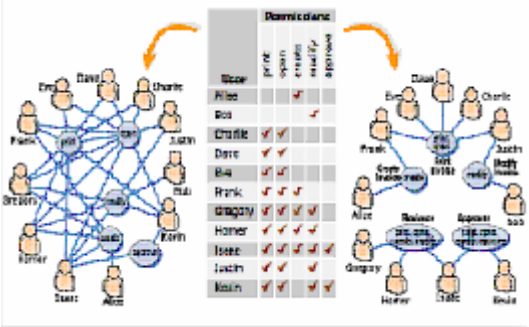

Fig. 2. Representation of Extracting the Roles

User – Permission Assignments
{(u1, p2), (u1, p3), (u1, p5), (u2, p1), (u2, p2), (u2, p4), (u2, p6), (u3, p2), (u3, p4), (u3,p6) (u4, p1), (u4, p2), (u4, p3), (u4, p5), (u5, p1), (u5, p2), (u5, p3), (u5, p4), (u5,p5), (u5, p6)}

Fig. 3: Input Datasets

	P1	P2	P3	P4	P5	P6
U1		■	■		■	
U2	■	■		■		■
U3		■		■		■
U4	■	■	■		■	
U5	■	■	■	■		■

Fig. 4: Matrix Representation

	P2	P3	P1	P5	P4	P6
U1	■	■	■	■		
U4	■	■	■	■		
U5	■	■	■	■	■	■
U2	■				■	■
U3	■				■	■

Fig. 5: Sorted Matrix

Roles	Permissions	Users
r1	p2	u1,u2,u3,u4,u5,
r2	p3,p5,p1	u1,u4,u5
r3	p4,p6	u2,u3,u5

Fig. 6: Role Generation

	P2	P3	P1	P5	P4	P6
U1	■	■	■	■		
U4	■	■	■	■		
U5	■	■	■	■	■	■
U2	■				■	■
U3	■				■	■

Fig. 7: Extracting Role R1- [$U_{(1, 4, 5, 2, 3)} \rightarrow P_2$]

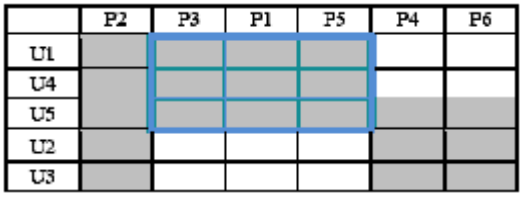

Fig. 8: Extracting Role R2- [$U_{(1, 4, 5)} \rightarrow P_{(3, 1, 5)}$]

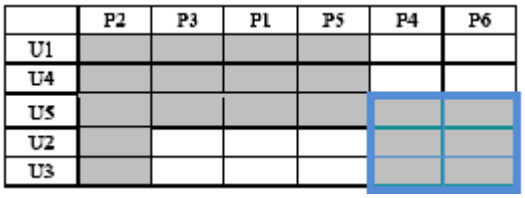

Fig. 9: Extracting Role R3- [$U_{(5, 2, 3)} \rightarrow P_{(4, 6)}$]

We will show that roles are easier to recogonize than describe via a binary matrix representation. The proposed representation can answer questions that classical statistical or mining such a tool an ideal companion for any existing role mining algorithm. The following figure 2 shown as the datasets for elicits roles. In the above Fig. 3 the admin should create the input data sets by using users and permissions as input.

User – Permission Assignment

First we have to create the input datasets in which the permissions are assigned to the users. Permission Assignment is to generate users and permissions store in to database. The initial step is in Fig. 3 is to create the users of the system and assign designation to the users to persist the data in to the database. The next step would be to create permissions and then to persist the data in to the database, permissions must be assigned to users and the dataset is to be formed.

Matrix Representation

A natural representation for these permission assignments is binary matrix which is represented in Fig. 4. Matrix contains rows and columns. Each value of the rows and columns correspond to users and permissions, and each cell represents '1' means a certain user has a certain permission is not granted and '0' is not granted.

Candidate Roles

In Fig. 6, using input data values group the permissions by separating each user and name a role to the groups. Generate candidate roles mainly based on permissions. Random Data Generator (RDG): The RDG was used in[19]; it takes five parameters { nu, nr, np, mr, mp} where nu, nr, np, are the number of users, roles ad permissions, respectively and where mr, mp are the maximum number of roles and the maximum number of permissions a user can have. Finally, for each user, the user – permissions assignment are computed based on user – role assignments and role – Permission assignments. The data generated[21] by the random data generator does not contain any structure and treats each user, role and permissions as statistically independent. We present two data generation algorithms that consider different structures and role hierarchies.

Matrix sorting algorithm

Get the user permission matrix and apply the matrix sorting algorithm VISRODE (VISualize Roles using DicE coefficient). Given a set of roles, this algorithm is able to provide a compact representation of them. In this VISRODE using DICE coefficients to avoid or reduce large gaps between the item sets. VISRODE is a heuristic algorithm used to sort the user permission matrix. We measure the similarity between two item using DICE coefficient:

$$DICE(i,i') = \frac{2 * \sum_{r \in roles(i) \cap roles(i')} |m(r)|}{\sum_{r \in roles(i)} + \sum_{r \in roles(i')}} \quad ...(1)$$

The Algorithm description of VISRODE

- Rows and columns are sorted independently.
- If some items are assigned to the same set of roles, they are put together.
- Items set positions are decided one – by – one.
- To avoid large gaps by putting item sets close to each other when they share large roles.
- Each item set is preferentially positioned at the beginning or at the end of already sorted item sets.
- Item set sorting is converted to item sorting.

Generate Pseudo Role

Visual analysis of user permission assignments are used to identify the roles. An approach is to first compute all possible closed permission sets and later trying to best represent them. A permission set is "closed" when no proper supersets of permissions possessed by the same users exist. Closed permission sets provide a compressed representation of all possible permission combinations that can be found within users. Closed permissions sets are roles in RBAC terminology. The number of closed permission sets is often too large when compared to the number of users and permissions. Hence, closed permission set leading to long running time and huge memory footprint.

Visual Elicitation

All pseudo roles are generated using a probabilistic algorithm called EXTRACT. After generating pseudo roles, each role is represented separately and also visualize user – permission assignments. Note that if the visualization quality is poor due to the approximated frequency values, it is possible to improve the quality by performing just additional samples. Suppose to have the matrix representation generated by feeding the algorithm VISRODE with the output of the algorithm EXTRACT with k samples: if we are not satisfied by this matrix, we can use k0 additional samples (namely, we run the loop) in order to have a more accurate frequency estimation.

EXPERIMENTAL RESULTS

Visualization Fragmentation Cost

Since the goal of our algorithm is to minimize the visualization cost of a set of hyper rectangles, visual cost (σ_U, σ_P, σ_R) of sampled datasets using three different conversions. It depends on the number of role fragments (i.e., sub matrices made up of contiguous "on" cells). In this below equation (2) represented by the quantity

$$V\sigma up = u(r) \times P(r) \quad ...(2)$$

Where $V_{\sigma up}$ represents weighted by the number of cells "wasted" to represent the role with respect to its compact representation as (i.e.)

$$\omega_u(r) \times \omega P(r) - |ass\_users(r)| \times |ass\_perms(r)| \quad ...(3)$$

Notice that when all the cells of a role are contiguous, the corresponding cost is zero.

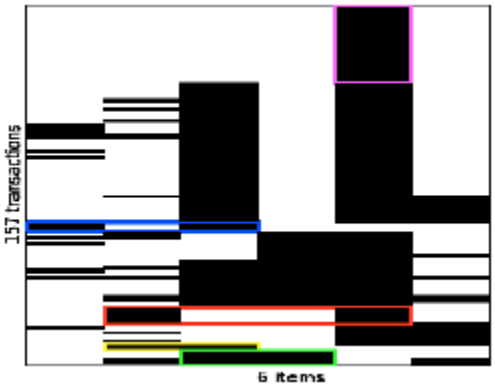

Fig. 10: Visualization Cost graph for [1]

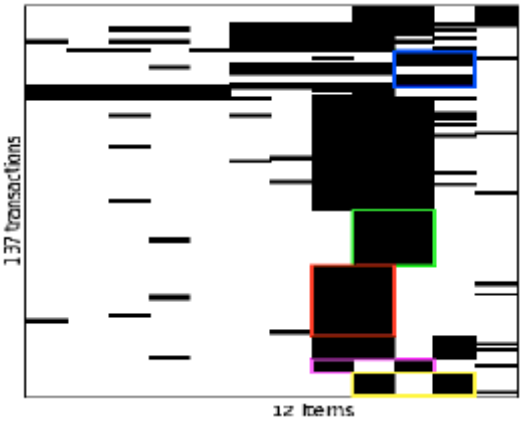

Fig. 11: Visualization Cost graph for VISRODE

The above Fig. 10 and Fig. 11 shown as the visualization cost for item sets in algorithms which we are described in the above sections. These graphs are to reduce the gap between the item sets by comparing VISRODE and[1].

Half Perimeter Cost

We would like to point out that an alternative visualization cost that we could have used here is the half perimeter [1], namely the sum of the height and width of roles in the given matrix representation. In our opinion, role visualization cost is more straightforward because a high role fragmentation greatly hinders the readability of the matrix, an aspect that does not catch. We will

support this statement by comparing the two measures in several real scenarios.

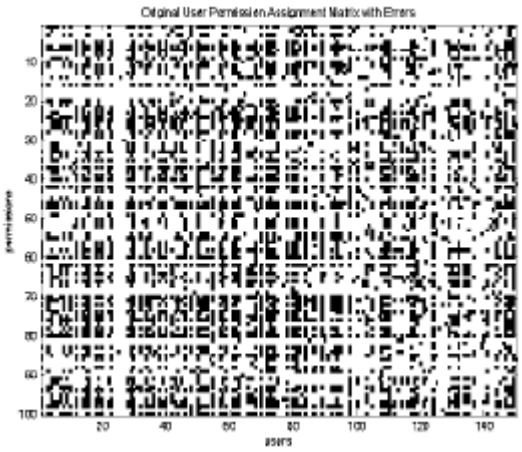

Fig. 12: Half Perimeter Cost graph for [1]

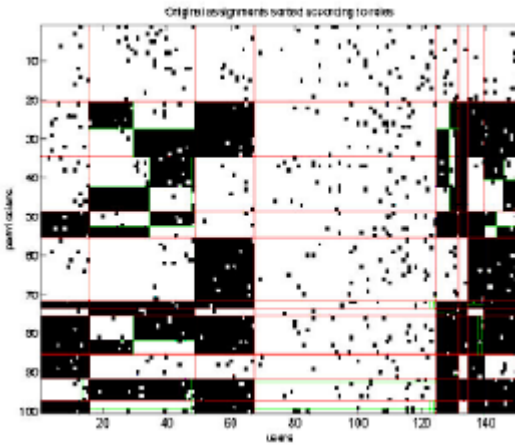

Fig. 13: Half Perimeter Cost graph for VISRODE

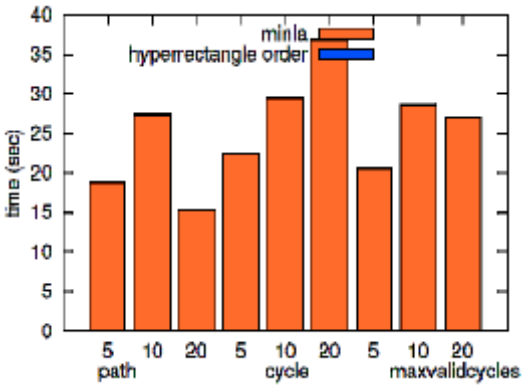

Fig. 14: Running Time complexity for [1]

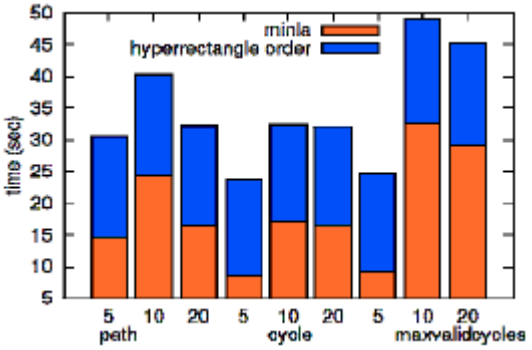

Fig. 15: Running Time complexity for VISRODE

Notice that Fig. 12 and Fig. 13 there is no particular strategy in positioning each role within the matrix: the algorithm only strives to reduce the number of fragments required to represent each role. We pointed out that a good matrix permutation can help role engineers to elicit candidate roles.

Running Time Complexity

In our algorithms we call the MinLA algorithm many times to get an improved graph order. We distinguish the running time of our algorithm with the running time of MinLA (in multiple times) of [1] in figure 14. We can see that in most cases the running time of our algorithms is almost negligible in comparison with MinLA of [20] with only one omission, the kosarak dataset.

We believe this is due to its unusually high number of unique items, which is much higher than other datasets. In addition, the running time of our algorithms remains almost constant when the number of hyper-rectangles increases.

CONCLUSIONS

Visualizing user-permission assignments in an intuitive graphical form that makes it possible to simplify the role engineering process. We offered a formal description of the visual role mining problem. We demonstrated that constructing the binary matrix representation of userpermission relations. Moreover, we proposed a novel heuristic algorithm called VISRODE to generate a matrix representation starting from the outcome of any role mining algorithm. We also described an efficient, tunable, and probabilistic tool referred to as EXTRACT. It produces approximate patterns

that can be used in conjunction with VISRODE to obtain high-quality visualization results the quality of the results produced by EXTRACT is formally proved.

Future enhancement

Data filtering, zooming algorithms, or approximated representation of data are just some examples of possible directions to investigate. Further, the problem of coping with very large data sets would deserve a deeper analysis.

ACKNOWLEDGMENTS

This is to acknowledge our sincere thanks to my advisor Mr. K. HariBaskar, M.E, (PhD) Assistant Professor, Mount Zion College of Engineering & Technology and all others who assisted me in bringing out this work successfully.

REFERENCES

1. Alessandro Colantonio, Roberto Di Pietro, Alberto Ocello, and Nino Vincenzo Verde (June 2012) "Visual Role Mining: A Picture Is Worth a Thousand Roles", *Proc. IEEE transactions on knowledge and data engineering* vol. 24 no. 6.
2. Chen, "Top 10 Unsolved Information Visualization Problems", *IEEE Trans. Computer Graphics and Applications,* vol. 25, no. 4, pp. 12-16, (2005).
3. Colantonio, R. Di Pietro, A. Ocello, and N.V. Verde, "A Formal Framework to Elicit Roles with Business Meaning in RBAC Systems", *Proc. 14th ACM Symp. Access Control Models and Technologies (SACMAT '09),*pp. 85-94 (2009).
4. Colantonio, R. Di Pietro, A. Ocello, and N.V. Verde, "Taming Role Mining Complexity in RBAC", *Computers Security*, **29**: 548-564, (2010).
5. D.A. Keim, G. Andrienko, J.D. Fekete, C. Gorg, J.Kohlhammer, and G. Melancon, "Visual Analytics: Definition, Process and Challenges" *Information* Visualization: Human – Centered Issues and Perspectives, **4950**: 154 -175 (2008).
6. Dana Zhang, Kotagiri Ramamohanarao and Tim Ebringer, "" Role engineering using graoh optimization", *Proc. ACM* Symposium on Access Control Models and Technologies (*SACMAT '07),*pp. 139-144 (2007).
7. Gustaf Neumann and Mark Strmbeck, "A scenario – driven role engineering process for functional RBAC roles", Proc.ACM Symposium on Access Control Models and *Technologies (SACMAT '02),* pp. 33, 42, New York, NY, USA (2002).
8. Haibing Lu, Jaideep Vaidya, and Vijayalakshmi Atluri, "Optimal Boolean matrix decomposition: Application to role engineering", *Proc. International Conference on Data Engiineering (ICDE '08),* pp. 297,306 (2008).
9. Haio Roeckle, Gerhard Schimpf, and Rupert Weidinger, "Process – Oriented approach for role – finding to implement role – based security administration in a large industrial organization", Proc. ACM Workshop on Role – Based Access Control (RBAC '00), pp. 103,110 (2000).
10. Ian Molloy, Hong Chen, Tiancheng Li, Qihua Wang, Ninghui Li, Elisa Bertino, Seraphin Calo and Jorge Lobo, "Mining roles with semantic meanings" Proc. ACM Symposium on Access Control Models and Technologies (SACMAT '08), pp. 21-30 (2008).
11. J. Edward Coyne, "Role Engineering", *Proc. ACM Workshop Role – Based Access Control (RBAC '95),* pp. 15-16 (1995).
12. Jaideep Vaidya, Vijayalakshmi Atluri, and Qi Guo, "The role mining problem: Finding a minimal descriptive set of roles" *Proc. ACM Symposium on Access Control Models and Technologies (SACMAT '07),* New York, NY, USA, 2007. ACM Press.
13. M. Frank, A.P. Streich, D. Basin, and J.M. Buhmann, "A Probabilistic Approach to Hybrid Role Mining", *Proc. 16th ACM Conf.*

Computer and Comm. Security (CCS '09), pp. 101-111 (2009).

14. M. Frank, D. Basin, and J.M. Buhmann, "A Class of Probabilistic Models for Role Engineering", *Proc. 15th ACM Conf. Computer and Comm. Security (CCS '08),* pp.299-310 (2008).
15. M. Kuhlmann, D. Shohat, and G. Schimpf, "Role Mining—Revealing Business Roles for Security Administration Using Data Mining Technology", *Proc.* Eighth ACM Symp. Access Control Models and *Technologies (SACMAT '03),* pp. 179-186 (2003).
16. M.J. Zaki and C.-J. Hsiao, "Efficient Algorithms for Mining Closed temsets and Their Lattice Structure", *IEEE Trans. Knowledge and Data Eng.,* **17**(4): pp. 462-478 (2005).
17. M.R. Garey, D.S. Johnson, and L. Stockmeyer, "Some Simplified NP-complete Problems," *Proc. Sixth Ann. ACM Symp. Theory of Computing (STOC '74),* pp. 47-63 (1974).
18. Molloy, N. Li, T. Li, Z. Mao, Q. Wang, and J. Lobo, "Evaluating Role Mining Algorithms", *Proc. 14th ACM* Symp. Access Control Models and Technologies (SACMAT *'09),* pp. 95-104 (2009).
19. R. Gupta, G. Fang, B. Field, M. Steinbach, and V. Kumar, "Quantitative Evaluation of Approximate Frequent Pattern Mining Algorithms," *Proc. 14th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD '08),* pp. 301-309 (2008).
20. R. Jin, Y. Xiang, D. Fuhry, and F.F. Dragan, "Overlapping Matrix Pattern Visualization: A Hypergraph Approach" , *Proc. IEEE Int'l Conf. Data Mining (ICDM '08),* pp. 313-322 (2008).
21. R. Santamaria, R. Theron, and L. Quintales, "BicOverlapper: A Tool for Bicluster Visualization", *Bioinformatics*, **24**(9):. 1212-1213, (2008).