Genetic Optimization of Access Control Schemes in Virtual Local Area Networks

Igor Saenko and Igor Kotenko

St. Petersburg Institute for Informatics and Automation (SPIIRAS)
39, 14 Linija, St. Petersburg, Russia
{saenko,ivkote}@comsec.spb.ru

Abstract. The paper presents the formulation of the problem of access control to information resources located in virtual local area networks. We define the initial data, the objective function and constraints of the problem. To solve the proposed problem we suggest the method of genetic optimization of access control scheme based on the poly-chromosomal representation of intermediate points. The results of computer simulation and evaluation of the proposed method are discussed.

Keywords: access control, virtual local area networks, genetic optimization.

1 Introduction

Joint work of users in computer networks stipulates the need to restrict the access to information resources without the use of passwords. An example is the problem of protecting information from unauthorized access in computer classrooms of universities.

This problem has the following specificity. First, the student contingent has a strong heterogeneity, and all students can be considered as potential security infringers. In this case, the effectiveness of passwords and user accounts is low. Secondly, in classrooms the access control schemes require frequent retuning. This is due to the fact that in classrooms the lessons, having different composition of used information resources, are usually alternated.

The basic principles of information security in such integrated information systems are outlined, for instance, in the papers [1, 2]. These papers show that the discretionary model based on an access control matrix is most widely implemented in classrooms of universities. The first access control matrix as access control scheme was introduced in [3]. This model was considered in more details in many modern works, for example [4, 5]. Each cell of the matrix defines the subject authority to access a specific object or another access subject.

In practice, as a rule, the access control matrix is replaced by access control lists (ACL) [6] or "lists of capabilities" (C-lists) [7]. Switches, used for local area networks, also use ACL [8]. Such capability allows to implement virtual local area networks (VLANs) based on these solutions [9]. ACL lists ensure that certain traffic is sent to specific ports. This prevents the unauthorized access to confidential corporate

I. Kotenko and V. Skormin (Eds.): MMM-ACNS 2010, LNCS 6258, pp. 209–216, 2010.
© Springer-Verlag Berlin Heidelberg 2010

information and network congestion as a result of program attacks. As a result, on the one hand, VLANs provide an additional level of access control to network resources, and, on the other hand, the adjustment of VLANs is also determined by the access control matrix.

The generation of an access control matrix is a complex problem [10]. Under system operation the adjustment of access control schemes is repeated each time, when the equipment, software and users are changed. Nevertheless, usually the generation of access control scheme is still done manually, without the use of mathematical methods [11]. Creation of access control scheme can be automated, if we reduce it to an optimization problem and apply an effective way to solve it. One of these ways is to use genetic algorithms. Genetic algorithms allow to solve successfully the problems of structural and parametric optimization of various systems [12, 13].

The purpose of this paper is to test the idea of applying genetic algorithms to generate a correct access control matrix for a computer network on the base of constructing VLAN. We suggest the method of genetic optimization of access control scheme based on the poly-chromosomal representation of intermediate points.

The paper is structured as follows. *Section 2* considers mechanisms for access control to information in VLAN. *Section 3* outlines the proposed problem definition and analysis. In *section 4*, we suggest the method of genetic optimization of access control scheme. *Section 5* discusses the results of computer simulation and evaluation of the proposed method. *Conclusion* surveys the main paper results.

2 Mechanisms for Access Control to Information in VLAN

The efficient mechanisms for access control and protection of information against unauthorized access in VLAN are (1) the rational distribution of information resources and users on network nodes and (2) the organization of virtual subnets using network switches or routers. Let us consider these mechanisms.

2.1 Distribution of Information Resources and Users on Network Nodes

Information resources (files or directories) distributed on network nodes are called *access objects*. Network computers are *network nodes*. Users, working at computers at any given time, are called *access subjects.*

Several access objects can be situated on one network node at one time. In other words, there is a mapping $\mathbf{D}^{OU}$ of degree 1 : M among the set of access objects and the set of nodes. The same access subject can work only on one node. Consequently, the mapping $\mathbf{D}^{SU}$ among the set of subjects and the set of nodes has also the degree 1 : M.

Access subjects have full access to those objects which are located on their own network node. At the same time, sometimes, the subject has to access one or more objects on other nodes (for example, on a network server). This capability is achieved by assigning to access object a special *shared access flag*. It should be noted that there is the opportunity of a password based shared access. However, this type of shared access is not taken into account in the statement of the problem due the specificity considered in the Introduction. The password based shared access control is assigned to VLAN.

It is supposed that the access of a specific subject to a particular object is determined by the following rules: (1) access is possible, if the object does not have a shared access flag, but is located on the node at which the subject operates; (2) access is possible, if the object has that flag (in this case it is not important on which node the object is located); (3) access is denied in all other cases.

2.2 Organization of Virtual Subnets

Virtual subnets are realized by using managed network switches. These network devices have a memory that stores information on banning (permitting) the exchange of information between certain pairs of computers connected to switches. As a result, it is possible instead of a fully connected exchange scheme between the ports to organize a selective scheme with segregation of virtual local subnets.

The following access rule is used in VLAN for all computers: if two computers are not in the same subnet, then the information exchange between them is impossible.

VLAN implementation requires a change of the second rule outlined above. Now this rule is as follows: access is possible, if the object has a shared access flag and the computer, on which the object is located, and the computer, on which the subject works, are in the same virtual subnet.

The simultaneous use of two considered access control mechanisms makes up a "real" access control scheme. At the same time the usage the specific software determine a "required" access control scheme. In the general case a "real" and a "required" access control schemes may be different.

Thus, an informal statement of the problem of access control with usage of VLAN can be formulated as follows: using mentioned access control mechanisms it is needed to ensure that the “real” access control scheme has minimal differences with the “required” scheme, and coincides with it in ideal case.

3 Formal Statement of the Problem

Let us specify the formal statement of the problem of on-line optimization of access control schemes in VLANs.

The initial data for the formal statement of the problem are as follows:

$\mathbf{OD} = \{od_i\}$, $i = 1 \ldots I$ – set of access objects (files, directories);

$\mathbf{SD} = \{sd_j\}$, $j = 1 \ldots J$ – set of access subjects (for example, learners and teachers working in a computer network);

$\mathbf{U} = \{u_k\}$, $k = 1 \ldots K$ – set of network nodes;

$\mathbf{R}^{\text{req}} = \|r^{\text{req}}{}_{ij}\|$ – requirements for different levels of access (required access control scheme), where $r^{\text{req}}{}_{ij} = 1$, if sd_j should have access od_i, and $r^{\text{req}}{}_{ij} = 0$ otherwise.

Since the problem variables should fully determine the decisions on the distribution of objects and subjects and the structure of VLAN, we assume that these decisions are as follows:

$\mathbf{D}^{\text{OU}} = \|d^{\text{OU}}{}_{ik}\|$ – matrix of distribution of objects on network nodes, where $d^{\text{OU}}{}_{ik} = 1$, if od_i is located on the node u_k, and $d^{\text{OU}}{}_{ik} = 0$ otherwise;

$\mathbf{D}^{\text{SU}} = \|d^{\text{SU}}{}_{jk}\|$ – matrix of distribution of subjects on network nodes, where $d^{\text{SU}}{}_{jk} = 1$, if sd_j is located on the node u_k, and $d^{\text{SU}}{}_{jk} = 0$ otherwise;

$\mathbf{V} = \{v_i\}$ – vector of shared access flags of network resources, where $v_i = 1$, if od_i is given in the share, and $v_i = 0$ otherwise;

$\mathbf{X} = \|x_{mn}\|$, $m,n = 1 \ldots K$ – matrix of VLAN structure, where $x_{mn} = 1$, if nodes u_m and u_n belong to one virtual subnet, and $x_{mn} = 0$ otherwise.

As an objective function should be used the function, evaluating the difference between the real control access scheme $\mathbf{R}^{\text{real}}$, stipulated by values $\mathbf{D}^{\text{OU}}$, $\mathbf{D}^{\text{SU}}$, $\mathbf{V}$ and $\mathbf{X}$, and the required access scheme $\mathbf{R}^{\text{req}}$.

Let us show how to obtain the functional form of scheme $\mathbf{R}^{\text{real}}$.

Assume that the access control scheme is determined only by the decisions $\mathbf{D}^{\text{OU}}$ and $\mathbf{D}^{\text{SU}}$ (in other words, all the elements of $\mathbf{V}$ and $\mathbf{X}$ are equal to 1). We call this scheme unconditional and denote $\mathbf{R}^{\text{uc}}$. In this case we have

$$\mathbf{R}^{\text{uc}} = \mathbf{D}^{\text{SU}} \cdot (\mathbf{D}^{\text{OU}})^{\text{T}}, \tag{1}$$

where the elements of the matrix $\mathbf{R}^{\text{uc}}$ are determined by the expression

$$r^{\text{uc}}{}_{ij} = \sum_{k=1}^{K} \left(d^{\text{SU}}{}_{ik} \cdot d^{\text{OU}}{}_{kj} \right). \tag{2}$$

Note that in (2), as in all subsequent expressions, summation and product are the logical operators OR and AND respectively.

Suppose that the decision $\mathbf{V}$ takes effect (that is, there are v_i=0). We call this access control scheme as "conditional on $\mathbf{V}$" and denote $\mathbf{R}^{\text{V}}$.

If $v_i = 1$, then the resource od_i is available for all subjects. In this case, for any j, $r^{\text{V}}{}_{ij} = 1$. If $v_i = 0$, then the availability of the resource od_i is defined by a matrix $\mathbf{R}^{\text{uc}}$. Consequently, the element of the matrix $\mathbf{R}^{\text{V}}$ is defined by the following expression

$$r^{\text{V}}{}_{ij} = r^{\text{uc}}{}_{ij} + v_i (1 - r^{\text{uc}}{}_{ij}). \tag{3}$$

Now suppose that in addition to $\mathbf{V}$, the decision $\mathbf{X}$ enters into force. In this case, the access control scheme is a "real" control access scheme $\mathbf{R}^{\text{real}}$.

The actual availability of the resource od_i to subject sd_j occurs when there is a virtual subnet joining this resource and this subject together. In other words, the following expression is true:

$$r^{\text{real}}{}_{ij} = \sum_{k=1}^{K} x_{ik} \cdot r^{\text{V}}{}_{kj}. \tag{4}$$

It is easy to see that expressions (2)–(4) completely determine $\mathbf{R}^{\text{real}}$ as a function of variables $\mathbf{D}^{\text{OU}}$, $\mathbf{D}^{\text{SU}}$, $\mathbf{V}$ and $\mathbf{X}$.

Objective function of optimization problem statement is defined as a measure of divergence between $\mathbf{R}^{\text{real}}$ and $\mathbf{R}^{\text{req}}$

$$\Delta\mathbf{R} = \sum_{i=1}^{I} \sum_{j=1}^{J} \left| r^{\text{real}}{}_{ij} - r^{\text{req}}{}_{ij} \right|. \tag{5}$$

The discrepancy between $\mathbf{R}^{real}$ and $\mathbf{R}^{req}$ should be minimal. Therefore, the synthesis criterion formulated in the problem statement has the form

$$\Delta\mathbf{R}\left(\mathbf{D}^{OU}, \mathbf{D}^{SU}, \mathbf{V}, \mathbf{X}, \mathbf{R}^{req}\right) \Rightarrow \min . \tag{6}$$

The constraints of the problem statement are as follows:

1) on a single node there can not be more than one subject, and therefore the following condition is true:

$$\sum_{k=1}^{K}\left(d^{SU}{}_{ik}\right) \le 1 ; \tag{7}$$

2) one file can be only on one node, so the following expression is valid

$$\sum_{k=1}^{K}\left(d^{OU}{}_{ik}\right) \le 1 . \tag{8}$$

4 Method of Solving the Problem

The problem defined by expressions (2) – (8) belongs to a class of non-linear Boolean programming problems, when the variables are given in the vector and matrix form. The exact solution of this problem is possible only by an exhaustive search of variables that can not be acceptable for practical purposes.

We offer for its solution a method which implements genetic optimization algorithms (GOA), successfully used in many synthesis problems [12, 13].

However, we note that, as shown by expression (3) and (4), the set of variables in the objective function (6) can be reduced by replacing the two matrices $\mathbf{D}^{OU}$ and $\mathbf{D}^{SU}$ on a single matrix $\mathbf{R}^{uc}$.

The method based on GOA is as follows. On initialization stage, an initial set of solutions (or *population*) is randomly formed. Each solution (or *individual*) is characterized by a string isomorphically related to the vectors and matrices of variables that determine this solution. This string is called a *chromosome* and a single character in it – a *gene.*

At each subsequent stage the following steps are fulfilled.

Pairs from the population of individuals are randomly selected. They are called *parents*. Between them the process of *crossing-over* occurs. As a result of this process, a couple of new individuals appear. These individuals are called *descendants*. The chromosome of each of the descendants is formed from two parts: one part is taken from the chromosomes of the "father", and the second – from the "mother"'s chromosomes. The descendants are added to the general population.

The population has quantitative restrictions, so individuals with the lowest suitability function are removed from the population ("die"). The role of suitability function is played by the function (5).

In addition, at each stage a part of the individuals is subjected to *mutation*. During mutation the genes in the chromosome are changed randomly.

An essential feature of proposed GOA is his poly-chromosomal character, i.e. individuals have not one, but three chromosomes $\mathbf{R}^{uc}$, $\mathbf{V}$ и $\mathbf{X}$.

Let us offer the forms of these chromosomes.

Since $\mathbf{R}^{uc}$ is a matrix of dimension I x J, it is not symmetric. Therefore, the only way to build a chromosome mapping this matrix is a serial concatenation of rows of $\mathbf{R}^{uc}$ into one big string:

$$[\mathbf{R}]_{chr} = [r_{11}, \ldots, r_{1J}; x_{21}, \ldots, x_{2J}; \ldots; x_{i1}, \ldots, x_{iJ}; \ldots; x_{I1}, \ldots, x_{IJ}]. \quad (9)$$

Vector $\mathbf{V}$ by its very nature is a chromosome, in which an element v_i carries the role of individual gene:

$$[\mathbf{V}]_{chr} = [v_1, v_2, \ldots, v_i, \ldots, v_I]. \quad (10)$$

Matrix $\mathbf{X}$ is a symmetric matrix. Each element of its main diagonal is 1. Therefore, to construct the chromosome which maps $\mathbf{X}$, the following string is used:

$$[\mathbf{X}]_{chr} = [x_{12}, \ldots, x_{1K}; x_{23}, \ldots, x_{2K}; \ldots; x_{i,i+1}, \ldots, x_{iK}; \ldots; x_{K-1,K}]. \quad (11)$$

As a result of poly-chromosomal crossing-over, not two, as in the traditional case, but eight descendants ($2^3 = 8$) will appear.

The GOA is completed, when the population goes to a stable state, in which the individual with the maximum value of efficiency is taken as the final solution of the problem.

5 Evaluation of the Method

The evaluation was conducted in two phases. On the first phase, we estimated computational complexity and performance. On the second phase, we estimated the LAN security based on the method developed.

Analysis shows that GOA has a polynomial computational complexity $O(N_{pop} \cdot N_{ind} \cdot K)$, where N_{pop} – number of populations needed to obtain a solution, N_{ind} – number of individuals in the population, K – number of network nodes. In the experiments, the value of N_{po} was in the interval [25; 100], K had values {5; 10; 15} and $N_{ind} = 200$.

Evaluation of GOA performance demonstrated that a full coincidence of the resulting access scheme with the required one is observed only at small values of I, in particular, when $I = 6$. Moreover, the coincidence is reached at population number in the range from 25 to 30.

Data on security evaluation are given in Table 1.

Table 1. Security evaluation

I	P_0	p_{pw}	N_1	N_2	P_1	P_2	k_{UUD}
6	10^{-4}	10^{-4}	6	0	0,00070	10^{-4}	7,00
6	10^{-5}	10^{-4}	6	0	0,00061	10^{-5}	60,98
6	10^{-4}	10^{-5}	6	0	0,00016	10^{-4}	1,60
12	10^{-4}	10^{-4}	12	5	0,00130	0,00060	2,17
12	10^{-5}	10^{-4}	12	5	0,00121	0,00051	2,37
12	10^{-4}	10^{-5}	12	5	0,00022	0,00015	1,47
20	10^{-4}	10^{-4}	20	18	0,00210	0,00190	1,11
20	10^{-5}	10^{-4}	20	18	0,00201	0,00180	1,11
20	10^{-4}	10^{-5}	20	18	0,00030	0,00028	1,07

The table 1 uses the following parameters: P_0 – the probability of unauthorized access the information caused by other reasons other than the compromise of shared passwords; p_{pw} – the probability of password compromising; $N1$ and $N2$ – the number of objects which require access password protection in the traditional case and in the case of using the proposed method, respectively; P_1 и P_2 – the probability of unauthorized access in the traditional case and in the case of using the proposed method, respectively; $k_{UUD} = P_1 / P_2$ – degree of security increase.

Table 1 shows that for various configurations of the simulated system the gain in security increase varies from 7 to 600 percentages. The greatest gain in 60 times takes place only when p_{pw} is greater than P_0 in 10 times, and the simulated system has a low dimension, when the resulting access scheme, organized by means of VLANs, is the same as required one. In all other cases, when the probability of compromising the password is much more than the probability of unauthorized access by other reasons, the gain is also significant.

6 Conclusion

The paper shows that combining the technologies of VLAN and GOA can be an effective means of protecting information against unauthorized access to the information stored in local networks. On the one hand, the proposed method of protecting information from unauthorized access takes into account the requirements of security policy. On the other hand, it provides multi-level use of organizational and technical measures of protection. The method has high efficiency and improves the security on 7–11 percentages for large-scaled systems or in 7–60 times for small systems. In this case the network performance was not reduced significantly, and the cost of routine work of security administrators was greatly decreased.

Software implementation of proposed method may be included in the arsenal of information security means available to network security administrators. It may be actively used to create dynamically configurable schemes of custom access to network resources.

Acknowledgments

This research is partly funded by the EU under SecFutur project, the grant of the Russian Foundation of Basic Research (Project No.10-01-00826) and Program of fundamental research of the Department for Nanotechnologies and Informational Technologies of the Russian Academy of Sciences (Contract No.3.2).

References

1. Bubnov, R.V., Chernikov, A.S.: Basic principles of security in an integrated information system to support the management of the University. Vestnik MSTU Bauman, No.2 (2004) (in Russian)
2. Simonenko, S.N.: Review of discretionary access control mechanisms in relation to information systems, http://www.philippovich.ru/Library/Books/ITS/wwwbook/IST7/simonenko/Simonenko.htm#_ftn1 (in Russian)
3. Lampson, B.W.: Protection. In: Proceedings of the 5th Princeton Conference on Information Sciences and Systems (1971)
4. Bishop, M.: Computer Security: art and science. Pearson Education, Inc., Boston (2002)
5. Kizza, J.M.: Computer Network Security. Springer Science+Business Media, Inc., New York (2005)
6. Galatenko, V.A.: Identification and authentication, access control, http://www.citforum.ru/security/articles/galatenko (in Russian)
7. Gyikovich, V.Y.: Fundamentals of Information Technology Security, http://bezpeka.ladimir.kiev.ua/pq/show/zi.htm (in Russian)
8. Hildebrandt, W.: Security at all levels. LAN 11 (2004) (in Russian)
9. Research Report: Secure Use of VLANs: An @stake Security Assessment (2002)
10. The main protective mechanisms used in systems to protect information, http://asher.ru/security/book/its/07 (in Russian)
11. How To Design an Access Control Matrix for Your Organization, http://www.howtodothings.com/business/how-to-design-an-access-control-matrix-for-your-organization
12. Wang, G., Dexter, T.W., Punch, V.F., Goodman, E.D.: Optimization of GA and Within a GA for a 2-Dimensional Layout Problem. In: First International Conference on Evolutionary Computation and Its Application (1996)
13. Shaffer, J.D., Eshelman, L.J.: Combinatorial Optimization by Genetic Algorithms: The Value of the Genotype/Phenotype Distinction. In: First International Conference on Evolutionary Computation and Its Application (1996)