IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555
Vol. 2, No.6, December 2012

The Streaming Server's Data Processing Technique Using Packet Priority Algorithm

Mi-Ra Kim
Dept. of Computer Science & Engineering
Ewha Womans University
Seoul, Republic of Korea

Dong-sub Cho
Dept. of Computer Science and Engineering
Ewha Womans University
Seoul, Republic of Korea

***Abstract*— The number of users who use multimedia data increased rapidly due to the development of the Internet. The user was able to meet a variety of needs using multimedia services that can be utilized, such as games, movies and education. Multimedia Server operators are taking advantage of the streaming techniques for efficient server operating. There was a serious breakdown in communications when multiple users provide access to the streaming server at the same time. Therefore we would propose techniques that provide for efficient streaming service. When the user is connected to a streaming server, a variety of information is stored in the user database such as payment amount, payment frequency, web page access time zones, web pages, and number of clicks. After the data preprocessing the user information, we would group users using the SVM algorithm. After we determine user's priority, based on this priority, we analyze the packet information. We would propose the way that first access to the high-priority IP. This would be known as Packet Priority Algorithm. Using the PPA algorithm, it can increase the efficiency of the server operating when the streaming server operating first highest first by connecting users. It would determine user access ranking faster because than other access ranking in the manner prescribed by the packet information.**

Keywords-SVM(Support Vector Machine); IP Splitting rule; packet analysis; Packet Priority Algorithm

I. INTRODUCTION

Recently we would use a variety of multimedia applications such as education, movies and games through the Internet. Because the increases of multimedia users using the Internet, several techniques have been proposed to manage them efficiently. Provide multimedia services via a streaming server, users can easily access the image data and can be downloaded. However, as an increasing number of users to connect to the streaming server, we would interfere when a large number of users connect to the streaming server.

In this paper, we would propose the way that we determine user's priority, and according to the priority access when users connect to the streaming server. We would user SVM (Support Vector Machine) algorithm to determine the priority of the user's way. Users would grouped by priority through a multi-user data preprocessing using the SVM algorithm. When users connect to the streaming server, by analyzing the portion of the IP packet, and the users first appointed to the rank of a server connection, based on the priority grouping. In chapter 2, we would examine the SVM algorithm and packet processing technique for IP Splitting rule. In chapter 3, we would design packet processing techniques from a streaming server environment using SVM. In chapter 4, we would implement the network packet processing technique using the SVM. In chapter 5, we would propose conclusions and future research directions.

II. RELATED WORKS

A. Support Vector Machines

First, Vapnik invented support vector machines[1]. In its simplest, linear form, an SVM is a hyperplane that separates a set of positive examples from a set of negative examples with maximum margin. Margin maximization can be expressed as given in [2] as

$$\min_{w,\xi} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{C}{n}\sum_i^n \xi_i, \quad \text{s.t. } \forall i,\ \xi_i \ge 0, \quad \forall i,\ y_i(\mathbf{w}\cdot\mathbf{x}_i + b) \ge 1-\xi_i. \tag{1}$$

Using a Lagrangian multiplier, this optimization problem can be converted into a dual form which is a QP problem, where the objective function L1 is solely dependent on a set of Lagrangian multipliers α :

$$\max_\alpha L_1(\alpha) = \sum_i^n \alpha_i - \frac{1}{2}\sum_i^n\sum_j^n \alpha_i\alpha_j \mathbf{x}_i\mathbf{x}_j, \tag{2}$$

subject to the inequality constraints,

$$\forall i,\ 0 \le \alpha_i \le \frac{C}{n}, \tag{3}$$

and one linear equality constraint,

$$\sum_i^n y_i\alpha_i = 0. \tag{4}$$

There is a one-to-one relationship between each Lagrangian multiplier and each training example. Once the Lagrangian multipliers are determined, the normal vector *w* and the threshold *b* can be derived from the Lagrangian multipliers:

$$\mathbf{w} = \sum_{i}^{n} y_i \alpha_i \mathbf{x}_i, \; b = -\mathbf{w} \cdot \mathbf{x}_k + y_k \quad \text{for some } \alpha_k > 0. \qquad (5)$$

B. Multi-class Support Vector Machines

We give introductions of multi-class SVM methods[3], one-against-all [4], one-against-one [5], and DAGSVM [6] which are based on binary classifiers.

Support vector machines are originally designed for binary pattern classification. Multi-class pattern recognition problems are commonly solved using a combination of binary SVMs and a decision strategy to decide the class of the input pattern. Each SVM is independently trained. The training data set (*xi, ci*) consists of N examples belonging to M classes. We assume that the number of examples for each class is the same, N/M.

The one-against-all (also known as 1-v-r or one-versus-rest) is the probably earliest implementation for multi-class SVM classification. In this approach, an SVM is constructed for each class by discriminating that class against the remaining (M - 1) classes. The number of SVMs used in this approach is M. A test pattern *x* is classified by using the *winner-takes-all* decision strategy, the class with the maximum value of the discriminant function f(x) is assigned to it. All the N training examples are used in constructing an SVM for a class. The SVM for class k is constructed using the set of training examples and their desired outputs, (*xi, yi*). The desired output *yi* for a training example *xi* is denied as follows:

$$y_i = \begin{cases} +1 & \text{if } c_i = k \\ -1 & \text{if } c_i \neq k \end{cases}$$

The examples with the desired output *yi* = +1 are called positive examples and the examples with the desired output *yi* = -1 are called negative examples. An optimal hyperplane is constructed to separate N=M positive examples from *N(M - 1)/M* negative examples. The one-against-all algorithm was implemented in MSVM.m with option '1vr' which extends a binary SVMs implementation SVM.m.

The one-against-one method is also known as 1-v-1 or one-versus-one method, and rst introduced on SVM as pairwise SVM. In this approach, an SVM is constructed for every pair of classes by training it to discriminate the two classes. Thus, the number of SVMs used in this approach is M(M - 1)/2. An SVM for a pair of classes (*k, m*) is constructed using training examples belonging to the two classes only. The desired output *yi* for a training example *xi* is denied as follows:

$$y_i = \begin{cases} +1 & \text{if } c_i = k \\ -1 & \text{if } c_i = m \end{cases}$$

The maxwins [7] strategy is commonly used to determine the class of a test pattern x in this approach. In this strategy, a majority voting scheme is used. Otherwise, class *m* wins a vote. Outputs of SVMs are used to determine the number of votes won by each class. The class with maximum number of votes is assigned to the test pattern. When there are multiple classes with the maximum number of votes, the class with maximum value of the total magnitude of discriminant functions (TMDF) is assigned. The total magnitude of discriminant functions for class *k* is denied as follows:

$$\text{TMDF}_k = \sum_{m} |f_{km}(\mathrm{x})|$$

Where the summation is over all m with which class *k* is paired. The one-against-one algorithm was implemented in MSVM.m with option '1v1' which extends a binary SVMs implementation SVM.m.

C. The Packet Splitting Rule

In the network the information consists of a number of packets, and using a packet of information, we would find out a lot of information [8]. We would be analyzed the information of packet statistically, and through this, prevent the leakage of information and outside intrusion. To identify the problems of the network, the system can be safely managed. And we would analysis of the various protocols, and figure out the analysis of network load, user behavior patterns and requirements of the various protocols. IP Splitting rule would apply Shift operations as much as N-bit which entered into by the user's IP address to accept the connection from the server. After applying the operation, the generated value applies m bit Mask operation. Packets are separated by IP and stored depending on the value of the generated h of. IP Splitting Rule steps are as follows Fig.1.

Step 1	We would input the IP of the user
Step 2	We would do shift operations the IP as long as n
Step 3	We would do mask operations S m bit (m = 1) the value generated in step 2
Step 4	We would generate the value of h
Step 5	Depending on the value of h, the IP protocol packets would be grouped by store

Figure 1. IP Splitting Rule Steps

IP Splitting rule generalization process is as follows Fig.2.

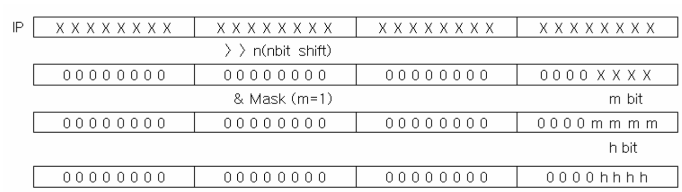

Figure 2. The generalization of IP Splitting Rule

The correlation of the field values of the proposed algorithm with the IP information would be determined. As seen earlier, the IP class Group is organized into five. The IP information would be applied to the IP Splitting Rule, and We would store the packets between the same IP Class Group. If you want to save the packet's information of the same IP Class, IP Splitting Rule would put the value of the desired field of IP Class Group.

If you want to save your information packet IP Class A Group, we would do n=31 bit shift operation. And if we would do the mask of m=1, we would generate the two value of h=0 and h=1. Depending on the value of generated h, a packet of information is stored. In h=0, the packet's information of the IP Class A is stored, and in h=1, the other packet's information stored is. When we would want to save the information packet of IP Class B Group, we would do n=30 bit shift operations and m=11 to the Mask, and then we would generate the four values of h=00, h=01, h=10, h=11. Depending on the value of h, the packet's information is stored in four Group. And in h = 00 and h = 01 the packet's information of IP Class A Group will be stored, in h = 10 the packet's information of IP Class B Group is stored are. These characteristics of IP Class would apply to IP Splitting Rule, and we would save a packet of the IP Group.

III. THE DESIGN OF THE STREAMING SERVER'S DATA PROCESSING TECHNIQUE USING PACKET PRIORITY ALGORITHM

A. The Framework of the streaming server's data processing technique using Packet Priority Algorithm

In this paper, we would propose streaming packet analysis processing technique of using the multi-SVM to improve the efficiency of server connection for the configuration in Fig. 3 below.

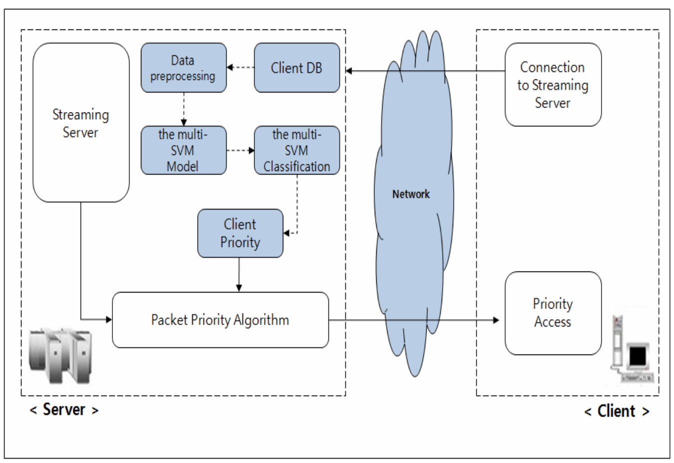

Figure 3. The Framework of the streaming server's user processing technique using Packet Priority Algorithm

Users would connect to the streaming server through the network. If you connect to the streaming server, the user is working on the server details and user information is stored in the user database. Data is stored in the user database through data preprocessing to build the multi-SVM model. We would determine the user's priority using the multi-SVM algorithm. The users are divided into groups based on priority of each. After we would determine user's priority set forth by the multi-SVM algorithm, then we would analyze the packet information when users access specifies the access priority. When users connect to the streaming server, Packet Priority algorithm by extracting the information of the IP packet is used to analyze the user's IP. And appointed users are classified by priority using the multi-SVM. We would set the user access ranking as a result of this classification.

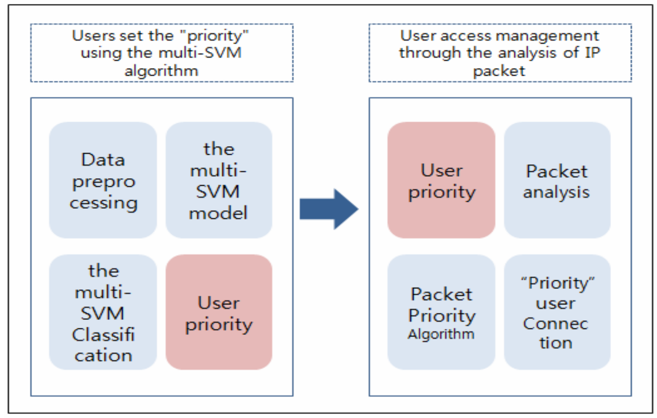

Figure 4. The phase of Packet Priority Algorithm

There would be divided into two phase that the user rank analysis and packet processing. Each step is shown in Fig.4. The first step is the priority setting phase that would determine using the SVM algorithm. It would be constructed SVM models through data preprocessing. It would be grouping user data by SVM classification through the resulting model using SVM. Through this process, the priority is determined. The second step is user access management through the analysis of IP packet. By analyzing packet when the user connect to the streaming server, it would be priority classification using the Packet Priority algorithm. With this classified results, it would be enable the user to access by user's priority.

B. Data Classification using the multi-SVM

The first step would be priority setting step using the multi-SVM algorithm. Through data preprocessing, the multi-SVM models are constructed. Using the resulting the multi-SVM model, it would be data grouping by the multi-SVM classification. Through this process, the priority would be determined. Each process would be in Fig.5. To apply the multi-SVM algorithm, it would be vectorization to the data of the user database.

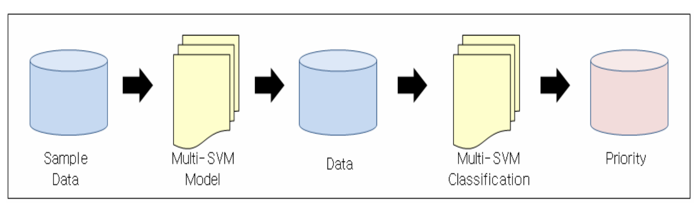

Figure 5. The first phase of Data Classification using the multi-SVM

To apply the multi-SVM algorithm, it would be vectorization to the data of the user database. The vectorization to go through the process of data preprocessing. Data extracted from the database of users is divided into seven fields. User access time, payment amount, frequency of access, connection time, etc. Each of these would be the name of the field [16].

C. Packet analysis with Packet Priority Algorithm

We would be separated and handle the packets by user-specific priorities using Packet Priority Algorithm. In the previous step, it would be grouped the user data using the multi-SVM. After the Packet Priority Algorithm priority by grouping, it would be grouped by user's priority. When user would connect to the streaming server, access each group differentiation. Each process would be in Fig.6.

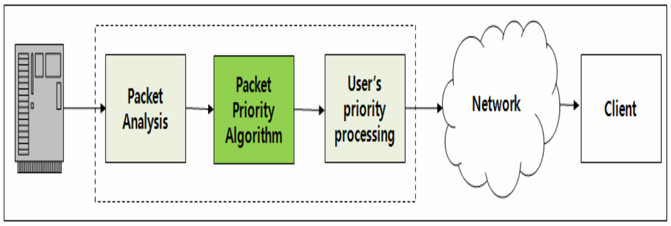

Figure 6. The second phase of Packet Priority Algorithm

Based on the IP Splitting rule, we would propose Packet Priority Algorithm. The IP Splitting rule would be applied shift operations as much as the N-bit, enough to accept the connection from the server to the user's IP address. Generated after applying the operation on the value of m bit Mask operation would be apply. Packet would be separated and stored by IP depending on the value of the generated h of. Packet Priority Algorithm would be applied to according to the grouping value for SVM user-specific rankings. In this paper, we would classify into four groups. By a 30-bit hexadecimal packet analysis, the shift operation is applied. And [0,0], [0,1], [1,0], [1,1] are classified into four groups by applying a 2-bit mask operations. We would set each group through user's priority using multi-SVM, and would allow to connect to server.

IV. THE IMPLEMENTED AND EXPERIMENTATION

The user packet processing technique would be implemented the following work flow in Fig.7. After the data preprocessing, the multi-SVM models would be constructed. We would be the multi-SVM classification after data preprocessing. Through this process, the user specifies the priority. We would analysis of user access packet IP. Using Packet Priority Algorithm, we would process with user's priority processing.

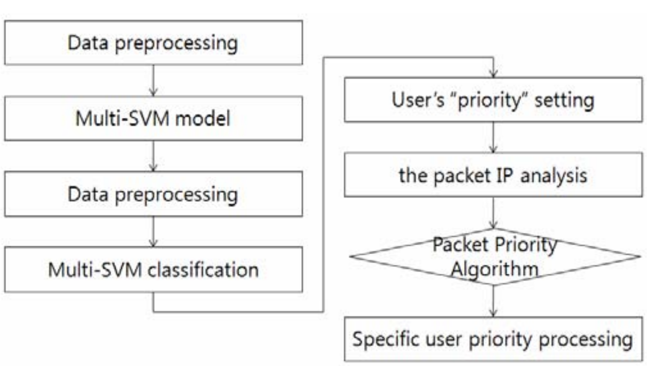

Figure 7. The work flow of the streaming server's data processing technique

First, through a process of data preprocessing, we should establish the multi-SVM model. Experimental data for constructing the multi-SVM model is in Table I. We would extract the data that corresponds to the field, such as class, the number of accesses, duration, payment, number of payments, time zone, downloads, page views from the user database. The data for each would be [1,10,10,100,000,10,24,21,45],[1,12,11,120,000,8,11,17,38],[1,9,9,90,000,12,5,15,42],[2,7,7,50,000,7,10,10,30],[2,8,6,45,000,8,13,12,33],[2,7,7,37,000,1,11,11,29],[3,5,5,10,000,5,5,8,27],[3,5,7,150,000,6,8,7,18],[4,1,1,0,1,6,0,12],[4,2,2,5,000,2,15,3,9].

TABLE I. THE EXPERIMENTAL DATA OF PACKET PRIORITY ALGORITHM

Class	Accesses	Duration	Pay	Num	Time zone	Down	Page
1	10	10	100,000	10	24:00	21	45
1	12	11	120,000	8	11:00	17	38
1	9	9	90,000	12	05:00	15	42
2	7	7	50,000	7	10:00	10	30
2	8	6	45,000	8	13:00	12	33
2	7	7	37,000	1	11:00	11	29
3	5	5	10,000	5	05:00	8	27
3	5	7	150,000	6	08:00	7	18
4	1	1	0	1	06:00	0	10
4	2	2	5,000	2	15:00	3	9

After building a SVM model, the value of the experimental data to go through the process of data preprocessing. Priority is determined after applying the data to the multi-SVM algorithm in table II. Using experimental data, and then build the multi-SVM model, applied to the multi-SVM. The value of the result of applying the multi-SVM prioritize data. The data for each would be [5,3,10000,3,15,8,7,3] in group 3, [7,7,30000,5,14,10,7] in group 3, [15,10,150000,9,8,5,1] in group 1, [10,5,50000,5,22,5,8] in group 2 .

TABLE II. THE DATA PREPROCESSING AND THE RESULT OF THE SVM

Num	1	2	3	4	5	6	7	Priority
1	5	3	10,000	3	15	8	7	3
2	7	7	30,000	5	14	10	7	3
3	15	10	150,000	9	08	5	1	1
4	10	5	50,000	5	22	5	8	2
5	7	3	40,000	1	18	6	4	2
6	12	5	100,000	8	17	5	12	1
7	10	14	80,000	10	23	7	10	1
8	1	1	0	0	10	0	5	4
9	3	1	0	0	18	0	8	4
10	7	3	30,000	1	18	5	17	3
11	8	2	20,000	3	16	3	10	3
12	3	1	10,000	1	20	5	8	3

Users are grouped by priority. Through multiple SVM, users are classified into four groups. Group 1 would be the experiment data such as [100.115.203.477 / 100.115.203.477 / 100.115.203.477], and IP range would be from 100.1.1.1 to 140.255.255.255. Group 2 would be the Experiment data such as [100.115.203.477 / 100.115.203.477 / 100.115.203.477], and IP range would be from 100.1.1.1 to 140. 255.255.255. Group 3 would be the Experiment data such as [100.115.203.477 / 100.115.203.477 / 100.115.203.477], and IP range would be from 100.1.1.1 to 140.255.255.255. Group 4 would be the Experiment data such as [100.115.203.477 / 100.115.203.477 / 100.115.203.477], and IP range would be from 100.1.1.1 to 140. 255.255.255. Each data would be in table III.

TABLE III. THE USER'S PRIORITY OF EXPERIMENT IP

Num	Priority	Experiment IP	IP range
1	Group 1	201.115.36.25 203.42.2.56 241.120.36.110	200.1.1.1 ~ 255.255.255.255
2	Group 2	122.20.46.12 112.132.63.111 123.25.23.14	112.1.1.1 ~ 129. 255.255.255
3	Group 3	155.120.14.13 138.142.130.26 143.115.23.4	130.1.1.1 ~ 200. 255.255.255
4	Group 4	58.115.10.156 65.132.15.14 49.175.41.123	10.1.1.1 ~ 111. 255.255.255

User access to the server, and users of the IP portion of the packet capture. Using Packet Priority Algorithm, of IP packets that are captured by analyzing a few second priority group, the group should identify. If you belong to one among a group of users of 1,2,3,4 will be processed first. Each data would be in table IV and table V.

TABLE IV. THE EXPERIMENTAL USER IP AND PACKET ANALYSIS

Num	User IP	Packet(32bit)
1	204.10.153.21	11001100 00001010 10011001 00010101
2	56.124.52.123	00111000 01111100 00110100 01111011
3	80.123.52.152	01010000 01111011 00110100 10011000
4	235.25.13.25	11101011 00011001 00001101 00011001
5	145.200.36.8	10010001 11001000 00100100 00001000
6	122.100.112.142	01111010 01100100 01110000 10001110
7	169.120.47.174	10101001 01111000 00101111 10101110
8	152.32.57.8	10011000 00100000 00111001 00001000
9	222.35.124.78	11011110 00100011 01111100 01001110
10	57.126.25.14	00111001 01111110 00011001 00001110
11	147.15.23.146	10010011 00001111 00010111 10010010
12	230.169.28.145	11100110 10101001 00011100 10010001

TABLE V. THE CLASSIFICATION BY PACKET PRIORITY ALGORITHM

Num	30bit shift	Mask(11)	Result	Priority
1	11	11	11	1
2	00	11	00	4
3	01	11	01	2
4	11	11	11	1
5	10	11	10	3
6	01	11	01	2
7	10	11	10	3
8	10	11	10	3
9	11	11	11	1
10	00	11	00	4
11	10	11	10	3
12	11	11	11	1

V. CONCLUSIONS

Using a variety of multimedia services over the Internet, there would be the exponential growth of the number of Internet users. The position of the companies that provide these multimedia service provider would require the way that provide the users a more stable and efficient service. When you run the server, we would distinguish between users more profit to the streaming server, and provide preferential service to them. In this paper, by analyzing the pattern of the user, we would determine access priority to fit into the user priority. The patterns of users would be grouped the user's priority using SVM. We would set priorities by using the user's fields affecting to connect to the streaming server database. We would propose the Packet Priority Algorithm that use user packet information in determining user access ranking. Through the Packet Priority algorithm, the user first access ranking by ranking packet classification, appointed to a streaming server, so that would connect.

In the future research projects, we would automate the implementation of each module, and implement the test environment that the actual users would connect to the streaming server. And by applying the method proposed to evaluate the performance of the proposed system, we would be trying to prove the superiority.

REFERENCES

[1] Changki Lee, Myung-Gil Jang, "Fast Training of Structured SVM using Fixed-Threshold sequential minimal optimization," ETRI Journal, Volume 31, Number 2, April 2009
[2] V. Vapnik, "Statistical Learning Theory," Wiley, New York, 1998.
[3] Naotoshi Seo, sonots "A Comparison of Multi-class Support Vector Machine Methods for Face Recognition," December 6, 2007
[4] V. Vapnik, "Statistical learning theory," Wiley, New York, 1998.
[5] S. Knerr, L. Personnaz, and G. Dreyfus, "Nurocosingle-layer learning revisited: A stepwise proce-dure for building and training a neural network," Springer, 1990.
[6] J. Shawe-Taylor J. Platt, N. Cristianini, "Large margin dags for multiclass classification, in Advances in Neural Information Processing Systems," 2000.

[7] J. H. Friedman, "Another approach to polychotomous classification, Technical report," Stanford, Department of Statistics, 1996.

[8] Ji-hye Ok,Dong-sub Cho, "Load Balancing in Distributed System for Packet Mining," ICEE, 2002

[9] E. Osuna, R. Freund, and F. Girosi, "Training Support Vector Machines: An Application to Face Detection," Proc. CVPR, pp. 130-136, 1997.

[10] J. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines," Microsoft Research Technical Report MSR-TR-98-14, 1998.

[11] I. Tsochantaridis et al., "Support Vector Machine Learning for Interdependent and Structured Output Spaces," Proc. ICML, p. 104, 2004.

[12] B. Scholkopf and A. Smola, "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond," MIT Press, 2001.

[13] Cristianini, N. and Shawe-Taylor, J. "An introduction to support vector machines and other kernel based learning methods," Cambridge University Press, 2000.

[14] A. Bestavros et al. "Distributed Packet Rewriting and its Application to Scalable Web Server Architectures," Proc.6th IEEE Int'l Conf. Network Protocols, IEEE Computer Soc. Press, Los Alamitos, 1988.

[15] Busch, Costas, "A Study on distributed structures," Brown University, 2000.

[16] [Online]. Available: http://svmlight.joachims.org/svm_multiclass.html

AUTHORS PROFILE

Mi-Ra Kim received her MS degree in Computer Science and Engineering at Ewha Womans University in 2003. She is currently a PhD candidate in Dept. of Computer Science and Engineering at Ewha Womans University. Her research interests include Automatic Inforamtion Classification, Personalization and Recommendation, Web Mining, Web Programming, Smart-grid.

Dong-sub Cho received his BS degree in Electrical Engineering, Seoul National University in 1979. He received his MS degree in Electrical Engineering at Seoul National University in 1981. He received his PhD in Computer Engineering at Seoul National University in 1986. He is currently a faculty member of the Department of Computer Science and Engineering, Ewha Womans University. His research interests include Ubiquitous Device and Pervasive Computing, e-Commerce and M-Commerce Design, High Performance Web Server Architecture, and High Available Web Mail Server.