Personalized E-mail Filtering System Based on Usage Control

Changda Wang
School of Computer Science and Telecommunication Engineering, Jiangsu University, Zhenjiang, China
Email: changda@ujs.edu.cn

Tingting Gong and Patricia Ghann
School of Computer Science and Telecommunication Engineering, Jiangsu University, Zhenjiang, China
Email: 309367807@qq.com, pghann@gmail.com

***Abstract*—In order to cope with the problem of spam soaring, a personalized e-mail filtering method based on UCON is proposed. E-mails from different senders were classified as junk e-mail, suspicious e-mail and normal e-mail by trust third-party according to the maintained blacklist and embedded machine learning technology online. Suspicious e-mails will be classified further from users' point of view manually. Then the incoming e-mails would be sifted and processed differently according to their classification. Experiments results illustrate the method of the paper not only provide a personalization filtering but also more accurate and effective than the popular statistical spam filtering system.**

***Index Terms*—spam filtering, personalization filtering, usage control**

I. Introduction

With the rapid development of computer networks, information and communication technology has been integrated into various dimensions of our social life. Nowadays, e-mail has become one of the most popular network applications. After three decades of development, it has evolved from simply transmit a text message to an application being used to send images, sound, video clips and other types of multimedia information. However it associated with the challenges of spam proliferation. According to the "20 anti-spam investigation report" issued jointly by the 12321 Report Center with the Internet Society of China Anti-Spam Information Center[1], in the first quarter of 2010, Chinese citizens received an average of 12.5 spam per week, which account for 38.3% in overall volume of e-mails.

With spam growing, solutions were proposed at juristic and technical levels to alleviate it. Anti-spam legislation of one country may not approve by the others. Thereafter, anti-spam solutions at technical level are feasible all over the world.

In reference to the popular technologies and try to avoid their defects, this paper focuses on a filtering scheme to improve the performance through a personalized e-mail filtering system based on UCON (Usage control) model. At the beginning, from the user's point of view to set the categories of incoming e-mails, and then by refer to user's address list, challenge sender verification questions and/or query trust third-party of non spam authorization online, the usage control model for personalized e-mail filtering is built. Our scheme combines a variety of filtering techniques to do spam identification, moreover, instead of decided by the filtering system in respect to the characteristic of e-mail only, e-mail receivers are invited to join the decision of the spam identification.

The rest of the paper is organized as follows: Section 2 discusses related work. In section 3, we show some drawbacks of traditional access control and the new generation of access control, i.e., UCON_{ABC} model. Section 4 summarizes popular anti-spam techniques and highlights the design philosophy of the e-mail filtering scheme. Section 5 build the e-mail filtering formal model of our scheme and then illustrate the merits of it by compare with the others. The conclusions were given based on validation experimental results in section 6.

II. Related Work

E-mail filtering and spam block technologies are the new research hotspot after E-mail classification. The development of e-mail filtering technology has evolved of three stages [2,3]: the first generation of e-mail filtering technology utilize IP address filtering, keyword filtering, e-mail (attachment) size control, SMTP connection time and frequency control, etc., to do identification and then block the spam. The second generation of filtering technology has intelligence characters embedded in when compared with the first generation[4]. Furthermore, manual intervention is introduced to improve the accuracy. In order to deal with the active "noise" added by spam to disturb filtering, the third generation adopts the methods based on the behavior analysis, which judges legitimacy of e-mail in respect to sending behavioral characteristics.

The essence of e-mail filtering is the accuracy of e-mails classification. Current anti-spam strategies can be categorized into two types: Automatic filtering based on machine learning and semi-automatic with human intervention. Most of semi-automatic technologies with

doi:10.4304/jsw.8.4.860-867

human intervention were developed a few years earlier. At that time, human intervention is a complement since lack of intelligence character. For example, heuristic filtering[5] was widely used before “A Plan for Spam” published by Graham in 2002, an algorithm based on a set of given rules to classify and bound e-mail, brought on nearly 100% accuracy rate but also had unacceptable false positive rate. Damiani et al. designed the signature filter of message digest to identify spam[6]. Jung and Sit studied seven popular blacklists and found that 80% of spam sources are included in certain DNS blacklist[7]. By analyzing the set of parameters, Gomes et al. proposed that the characteristics of spam can be clearly distinguished from ham traffic in the filtering of traffic analysis [8]. The method dwelled on by Wang and Chen utilizing head message to filter out spam possibly made normal e-mail get a low false alarm [9]. These techniques gradually become the complement of machine learning techniques to do anti-spam in recent years.

There are many different anti-spam machine learning techniques. For example, Pelletier et al. adopted spam filtering as an increased spam filtering layer [10]; DMTP authorizes recipients more rights to get e-mails transportation paths [11]; the Leiba B, Fleizach C treat digital signature as a verification on the protocols of DomainKeys Identified Mail (DKIM) [12] and Occam[13], respectively. These anti-spam techniques are usually used in client-side for the convenience of adjusting the classification of e-mail filter.

The complete anti-spam technology, e.g. Bayes, a probability-based statistical algorithm, has been applied in many different aspects to block spam [14,15,16]. Biju Issac et al. utilize keywords in context to improve the Bayesian filtering algorithm[17]; Ming et al. built another model based on Bayesian for spam behavior recognition [18]. Fan Jieting et al. completed the research that the SVM maintains a rapid classification speed in the case of a very large size of samples[19]. ZhaoWenqing and Zhu Yongli categorized e-mail as spam, no-spam & suspicious ones based on rough set theory[20]. Chiu Y combined rough set theory with genetic algorithm and eXtended Classifier System (XCS) to set up spam filter[21]. Some scholars also tried to integrate Neural Network (NNet) into e-mail filter[22], for example, multi-layer perception classifier is used for detecting the type of ham and spam[23].

Currently, statistical based e-mail filtering technology is one of the most popular techniques to protect users from spam harassing. Encounter with the tactics constantly changed by spammers, machine learning technologies over-dependent on training samples and then cannot catch up spam changing in time, so e-mail filtering system integrated with multiple anti-spam methods based on UCON is proposed in the paper.

III. Usage Control

Access control is a kind of key technology of information security. Traditional access control models, e.g. access control matrix, BLP, RABC, etc., are not adapt for the modern network based information system. The reasons are as follows:

(1) Subjects’ rights are static. Before authorization, the rights of subject have no relationships with the execution environment, which is inadequate for dynamic, heterogeneous and distributed systems. Once subject had been granted a right, it will be kept by the subject until a clear process emerged to revoke it. This makes a security flaw since manually arrange revoke processes are easier to be ignored by the administrator. (2) Traditional access control can only be authorized before execution, which dissatisfies modern access control requirements of sustaining management for resources being in use. (3) Traditional access control focuses on the protection of digital resources in a closed system, and then it would be disabled after digital resources being disseminated outside the system. (4) Traditional access control only manages the users’ accessing behaviors within limited range, e.g. single operating system, which inadequate for network and distributed environments.

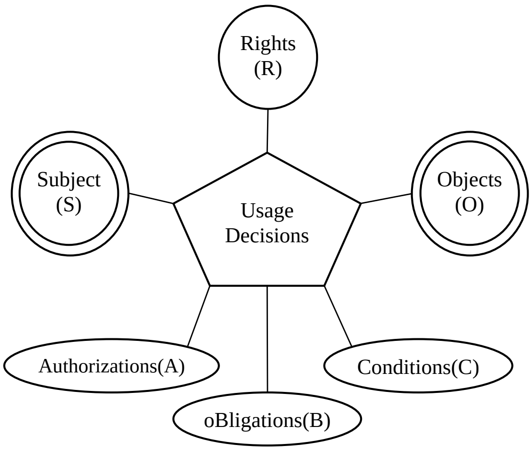

Fig.1. UCON$_{ABC}$ Model Components

UCON, a new access control model focused on usage control, proposed by Jaehong Park and Ravi Sandhu[24]. Since UCON extends traditional access control in many aspects[25], it is believed as the foundation of next generation of access control. UCON defined three decision factors including *authorizations*, *obligations* and *conditions*, as shown in Figure.1, and proposed two important attributes as *continuity* and *mutability*. UCON integrates Traditional Access Control (TAC), Trust Management (TM) and Digital Right Management (DRM) together. It provides a unified framework for protecting digital resources and prevents non-secure operation in modern information systems, as a new type of access control conceptual model.

Authorizations are functional predicates that have to be evaluated for usage decision and return whether the subject is allowed to perform the requested rights on the object. *Obligations* are another type functional predicates that verify mandatory requirements a subject has to perform before or during a usage exercise. *Conditions* are environmental or system-oriented decision factors. Condition predicates evaluate current environmental or system status to check whether relevant requirements are satisfied or not and return either true or false.

Our classification of incoming e-mails based on the following characteristics of UCON: decision factors, continuity of decision being either pre or ongoing with respect to the access in question and mutability that can allow updates on subject or object attributes at different times. With mutability properties on mutable attributes, occurrence of updates is possible before (pre), during (ongoing) or after (post) the right exercised, which may affect usage decision on this or next time. Among UCON model, subjects, subject attributes, objects, object attributes and rights continued the concept of traditional access control. However, rights are no longer static and independent of subject activities, whereas dynamically determines access rights based on the relationships of subject attributes and object attributes and the three decision factors mentioned above.

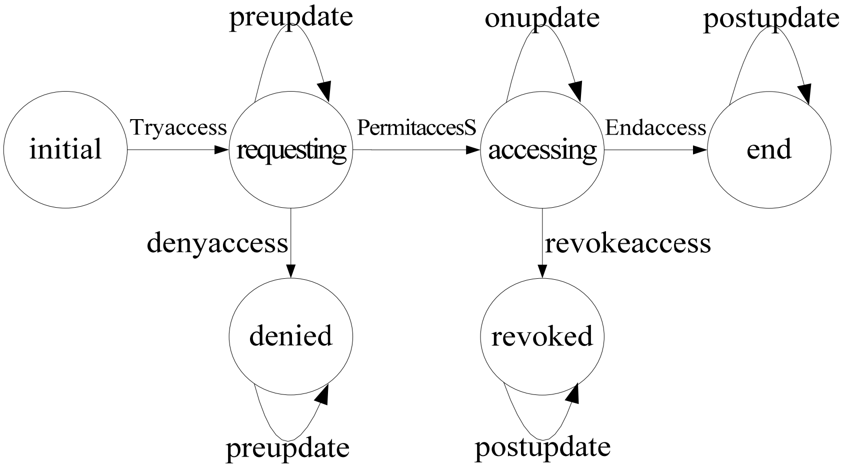

Fig.2. State transition of a single access

UCON can be classified into 16 basic sub-models for such e-mails classification. Each sub-model is produced only according to a single decision factor, the enumerated model spaces are listed in Table I. If a sub-model not reachable, it marked "N", otherwise marked "Y". In practice, different combinations of sub-models are made for different demands.

TABLE I.
THE 16 BASIC UCONABC MODELS

	immutable	pre-update	ongoing-update	post-update
preA	Y	Y	N	Y
onA	Y	Y	Y	Y
preB	Y	Y	N	Y
onB	Y	Y	Y	Y
preC	Y	N	N	N
onC	Y	N	N	N

IV. E-MAIL FILTERING SCHEME BASED ON UCON

A. Related techniques

The personalized scheme proposed in this paper combines various techniques which can be replaced or complement by other related techniques to meet users' needs. The following techniques are the reference of our scheme:

(1) Black-and-white list technology

All incoming e-mails, the senders were not in the white list, are set as spam, which may contain trust e-mail address, server domain name or IP address information. Only senders in the white list, the incoming e-mails were ready for read. Similarly, blacklist is an anti-spam technique on the contrary. Black and while list techniques are simple and reliable, by which can save bandwidth, storage capacity and processing time, and be applied to any level of the system. However, this kind of techniques cannot update and maintain the contents of the white and/or black lists in real-time; thereafter boring human intervention is needed. With spam growth, use black-and-white list technology alone to do spam filtering isn't feasible.

(2) Head message analysis

Some spam senders often forge 'From' address to cheat mail filtering system or users. Head message analysis check senders' addresses by comparing with the 'Received' domain, especially the 'Received' domain of the first mail server to avoid swallow the bait. The advantages of this method are simple, convenient, and easy to be implemented, while its disadvantages are high false positive rate and poorer processing performance.

(3) Challenge-response

The challenge-response mode is used to deal with those procedures of sending e-mail automatically. System temporarily keeps the incoming e-mail and then challenges the sender a question. In a given period, the e-mail would be put into the inbox of the receiver if and only if the sender response the challenge correctly. This method effectively frustrate spam from strangers, unfortunately, it also brings extra burden to legitimate senders.

(4) Probability statistics

Probability statistics approaches utilize classified e-mails as training samples at beginning to extract the characterization from various types of e-mails. Based on that, an incoming e-mail would be identified as a spam or not according to the results of its spam probability calculation. Generally, this kind of method is more accurate and occupies less storage space, but it over relies on training samples with higher computational complexity. Training is both the time and resources consuming processes.

B. Personalized e-mail filtering

From the user point of view, e-mail can be categorized into legitimate e-mail (ham) and illegitimate e-mail (spam). Usually an e-mail belong to one of the following types:

I. Address List Contacts. The contacts in e-mail address list makes the social networking of a user. They are possible in person, but mostly referred to connections in the workplace, universities, and high school, as well as neighborhood. Generally, e-mails of this type account for majority of total normal communication except some public service e-mail address;

II. Temporary Address Contacts. Contacts from a temporary address list is the prospect Class I contacts of an e-mail user, moreover, Class I contacts may degrade to this type under some circumstance, generally refers to people familiar with the user or used to communicate with the user, but that used e-mail account unavailable now. The contacts of Class II includes: (1) the prospect Class I contacts at the beginning stage and the e-mail receivers are not sure whether these contemporary users can be evolve to an Class I contacts one day thereafter refuse to add them into their contact list; (2) a Class I contact uses a temporary e-mail address in the case of the previous e-mail addresses are not available. For example, a person may not use the e-mail account of his previous worked company when he employed by a new company.

III. Stranger Address. Received an e-mail, not a junk e-mail, from stranger address has low probability. We roughly categorized them as followings: (1) e-mails from fixed address occasionally, such as a statement of business or government officials; (2) large number of messages received within a period, for example, recruitment unit receives the job applications from candidates; (3) strange mails often received, for example, complaint e-mails from customers.

By classify the possible relationships between user and senders, the different filtering techniques are adopted to achieve optimal filtering effect. The main idea of our personalized e-mail filtering scheme is as follows:

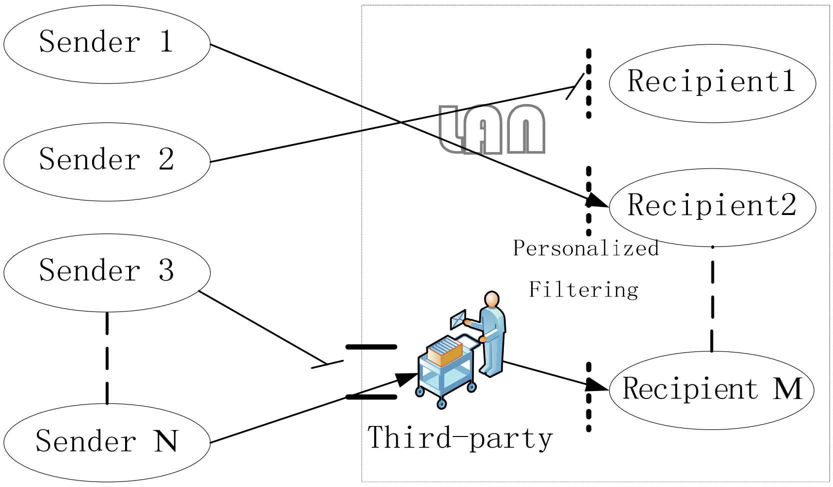

Fig.3. Personalized e-mail filtering model

a) Senders of Class *I* have the highest communication frequency with users, and it is also one type of the trust senders. Class *I* therefore can be adopt white list verification technique, a simple, rapid method and occupy less system resources.

b) Senders of Class *II*, whose current e-mail address were unknown by receiver but not a stranger. Receiver can adopt challenge-response to do verification, i.e. by correctly answer a private question of receiver to get through the spam filtering.

c) Senders of Class *III*, because of the uncertainty of the identity, are the most difficult type to judge an incoming e-mail is a ham or spam. Like our daily life we may consult a trust and knowledgeable person to give advices for uncertain situations. A trust third-party of e-mail filtering is built online. The trust third-party utilizes relatively accurate method based on machine learning which can be supplemented by a method about head message analysis, to do spam filtering online. A certification would be issued with those considered as a ham to confirm the receiver.

d) In the view of receivers, most of them prefer to receive spam instead of block all the suspicious e-mails. Therefore, trust third-party online should take on a technology with well identification rate at the cost of letting a small number of spam get through. To improve the performance, third-party needs to set up a blacklist in aid of filtering and to do spam audit.

Training samples and the characteristics of spam is extracted from mass e-mails header analysis, by which to feed into the core filtering system to pick out spam as intelligence as possible. The incoming e-mails get through the online third-party filtering and client-side personalized e-mail filtering were put into the inbox of the user and ready for read. A suspicious e-mail pick out by the filter system will be marked as a real spam after receiver's confirmation, thereafter spam sender address may blacklisted optionally. The spam, include spam from blacklist, will be used as a training data for filter system to update and adjust filtering rules to deal with new spam emerging.

From the above design ideas, the filter program is designed to improve the speed and accuracy of the LAN classification of spam. This system consists of two modules: Online third-party filtering based on machine learning and client-side personalized filtering. Online third-party filtering mechanism deployed at the entrance of the LAN, by which checks the incoming e-mails before them reach the destination mailbox. Furthermore, client-side personalized filtering (including white list verification and challenge-response method) is located on the terminal to block the spam. Personalized e-mail filtering scheme and the usage processes are as follows, fig.4.

Usage Rules

```
START
  an e-mail incoming;
  IF incoming e-mail apply for certification
    IF the mailing address is reliable proved by head message
      IF the trust third-party confirms the legitimacy of the e-mail
        issue a certification to authenticate the incoming e-mail and mark it as a ham;
      ELSE junk the e-mail;
    ELSE junk the e-mail ;
  ELSE send to client-side for checking;
        IF the sender's address among white list
            mark e-mail as a ham;
        ELSE IF challenge-response mode authenticate the sender
            mark e-mail as a ham;
        ELSE junk the e-mail;
  IF receiver finds it is a spam get through the authentication of the third-party
        reject and send the certification back to third-party, thereafter the third-party would
        add sender's address in blacklist;
END
```

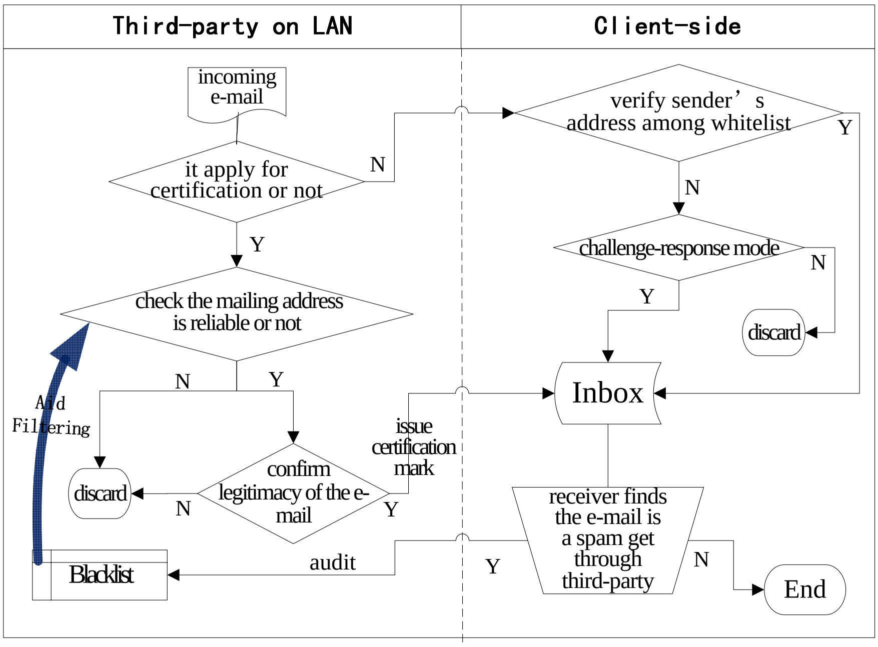
Third-party on LAN
Client-side
incoming e-mail
it apply for certification or not
N
Y
check the mailing address is reliable or not
N
Y
Aid Filtering
discard
N
confirm legitimacy of the e-mail
Y
issue certification mark
Blacklist
audit
verify sender's address among whitelist
Y
N
challenge-response mode
N
Y
discard
Inbox
receiver finds the e-mail is a spam get through third-party
Y
N
End

Fig.4. Filtering processes

Figure.3 shows our scheme of personalized e-mail filtering model is a collaboration of senders, recipients and a trust third-party online. With a certification from third-party, an e-mail of class *III* can be received by users under the decision of preB sub-model of UCON, which also get through head message checking and machine learning filtering. Due to existence of personalized filtering in client-side, a large number of spam turn to the authentication of the online trust third-party. Undoubtedly, that will increase the workload of the filtering system and be likely to bring false alarm. Trust third-party online equipped with a blacklist used for post auditing, assistant filtering and machine learning, which act as a globalB sub-model of UCON. Post auditing operated by blacklist will reduce some of the potential spam. For convenience, personalized authentication is built in client side for the e-mails of Class *I* and Class *II*. When users received such an e-mail, they validate it by white list or by challenge-response mode, and this process served as the sub decision model of preA and preB of UCON. Finally, user manually check e-mails served as the sub decision model of onA of UCON.

V. Advantages

A. Formal Usage Control Model

Among spam filtering process, trust third-party online, client-side and incoming e-mails are all participants. Sub-model preA, onA, preB and globalB of UCON dynamical control the access rights in the usage process. *S*, *O*, *r* represent the subject set, the object set and the requested rights, respectively. The strategies listed in figure.2 and figure.4 are formally present as the follows:

(1) Certification of the trust third-party online

In this stage, e-mail is going to be proved legitimate by certification of the third-party online. This stage can use UCONpreApreB Model, in which *s*, *o*, *obs*, *obo*, *ob* represent subjects, objects, obligation of subjects, obligation of objects and obligation of operations, respectively.

permitaccess(s,o,r)→◆(Θ(requesting(s,o,r)∧ (o.application=TRUE)∧ addressIntegrity(Third-party,o) filter(Third-party,o)) ∧send(Third-party,authenMark) ∧send(Third-party,o))

(2) Validation in client-side

Uzilize UCONpreApreB model, in which *s.att*, *o.att* represent subject attributes and object attributes, respectively.

permitaccess(s,o,r)→◆(◆requesting(s,o,r)∧ Θ((o.address∈s.addressBook) || ((o.address□s.addressBook)∧question(system,sender))) ∧accept(inbox,o))

(3) Validation by user

Utilize the decision model of UCONonA.

□(¬(s.judgement=TRUE)∧(state(s,o,r)=accessing)→ revokeaccess(s,o,r))

(4) audit stage

Utilize the sub-model of UCONglobal.

□(exist(o,authenMark)∧revokeaccess(s,o,r)→ enterup(blacklist,o.address)

B. Model Checking

In general, e-mail filtering is separated into training phase and application phase. The training phase is to calculate priori probability of a given spam by its characteristic values, while the application stage is to pick up spam from incoming e-mails according to their characteristics in the spam database.

In order to verify the effectiveness of the proposed method, some popular anti-spam methods, e.g. simple Bayesian filtering method and adaptive spam filtering method [26], are introduced as the reference model. The aim of our experiments is to compare the anti-spam effect with single probability statistics filtering method. In order to accurately describe the results of the comparison experiment, the same test set was selected as in[26], the sizes of them are 100, 200, 300, 400, 500, 600, 700, 800, respectively, and the amount of spam accounted for 30% of the totality of the e-mails. We assume that the challenge question set by user is impossible to be leaked or be guessed by a spam sender, and the senders of Class *I* and *II* never send spam.

Spam filtering performance evaluation usually adopts text classification indicators. In this paper, the following two important evaluation indicators are *recall* and *accuracy*. In order to define a few variables, we assume the total of e-mails is N in the test set, Table II, where N=A+B+C+D.

TABLE II.
THE SITUATION OF JUDGMENT ON SPAM

	the actual ham	**the actual spam**
system judgment is ham	A	C
system judgment is spam	B	D

(1)Recall: the proportion is that identified spam accounted for the totality of actual spam, R=D/(C+D)*100%, i.e. spam recall reflects the system detecting ability.

(2)Accuracy: the proportion is that the totalities of e-mails are accounted for by the correctly classified mails, Acc=(A+D)/N*100%, which reflects the correctly classified ability of system.

After using the filtration system, the blacklist formed by each LAN is different. Here we adopt model checking to illustrate the effectiveness of personalized e-mail filtering scheme.

Experimental Software Environment: Windows XP Professional, 2002, NuSMV 2.5.2;

Experiments Hardware Environment: Intel (R) Core (TM) 2 Duo CPUs 2.93 GHz, G-300 hard disk, 1.96G memory.

Experimental results of analysis

TABLE III.
COMPARE WITH SIMPLE BAYESIAN FILTERING

test	Simple Bayesian filtering		UCON-based filtering	
set	recall/%	accuracy/%	recall/%	accuracy/%
100	62.4	69.6	91.1	88.5
200	70.3	75.1	92.9	89.7
300	73.5	77.4	93.7	90.6
400	76.1	79.1	94.4	91.3
500	79.9	82.2	95.0	92.6
600	82.5	84.6	95.6	93.3
700	84.2	85.1	96.3	93.8
800	86.3	86.9	96.5	94.6

TABLE IV.
COMPARE WITH ADAPTIVE BAYESIAN FILTERING

test	Adaptive Bayesian filtering		UCON-based filtering	
set	recall/%	accuracy/%	recall/%	accuracy/%
100	80.5	83.2	96.2	93.2
200	82.7	84.1	95.7	93.2
300	84.0	84.8	96.1	93.3
400	85.1	86.0	96.1	94.0
500	85.9	87.2	96.3	94.3
600	87.0	87.9	96.5	94.8
700	88.2	88.8	97.0	95.3
800	89.4	89.8	97.3	95.6

As can be seen from Table III and Table IV, the merits of the filter system based on UCON is obviously. According to the comparison on recall and precision, it can be seen that spam pick out in the personalized e-mail filtering system based on UCON is obviously better than those in Bayesian models. Especially, when the training samples were small, the advantages are more apparent, because it is difficult for users to collect enough spam as training samples when they start to use filters. Combination with blacklist technique can not only speed up the filtering rate of the third-party, but also provide the samples of spam training.

Experimental results illustrate the system has achieved our desired effect with higher accuracy and speed.

VI. CONCLUSIONS

The proposed spam filtering system based on UCON, combined various filtering methods, is utilized to classify and filter different types of e-mails. Due to the effect of the filtering system is better than single filtering method based on probability and statistics and easier to be implemented, furthermore, the online trust third party can embedded various e-mail filtering techniques to enrich anti-spam system. Our anti-spam scheme is a flexible and effective e-mail filtering system.

ACKNOWLEDGMENT

This work was supported in part by the National Nature Science Foundation of China under Grant 61003288, Nature Science Foundation of Jiangsu Province under Grant BK.2010192 and the Ph.d Foundation of Ministry of Education under Grant 20093227110005.

REFERENCES

[1] 12321 "*Network adverse Spam Report Reception Center*", 2010-06-13. http://www.12321.cn/viewnews.php?id=12858.

[2] Qian Gao. *The invasion and defense of spam*, Computer Security, 2008, No.6, pages: 52-53.

[3] Zhian Yi, Yan Mao. *Overview of spam filtering technologies*, Journal of Yangtze University (Nat Sci Edit), 2010, Vol.7, No.1, pages: 256-258.

[4] Jingtao Sun, Qiuyu Zhang and Zhanting Yuan. *Application of refined LSA and MD5 algorithms in spam filtering*, Journal of Computers, 2009, Vol.4, No.3, pages: 245-250.

[5] David Anderson. *Statistical Spam Filtering*, EECS595, Fall 2006.

[6] Damiani E, De Capitni di Vimercati S, Paraboschi S, Samarati P. *P2p-based collaborative spam detection and filtering. Proceedings of the fourth international conference on peer-to-peer computing* (P2P' 04), IEEE Computer Society, 2004, pages: 176–183.

[7] Jung J, Sit E. *An empirical study of spam traffic and the use of DNS black lists*, Proceedings of the 2004 ACM SIGCOMM Internet Measurement Conference, IMC 2004. Taormina, Sicily, Italy, October 2004, pages: 370-378.

[8] Gomes L H, Cazita C, Almeida J, Almeida V, Meira Jr W. *Characterizing a spam traffic*, Proceedings of the 2004 ACM SIGCOMM Internet Measurement Conference, IMC, 2004, pages: 356-369.

[9] Wang C-C, Chen S-Y. *Using header session messages to anti-spamming*, computers & security 2007, Vol.26, No.5, pages: 381-390.

[10] Pelletier L, Almhana J, Choulakian V. *Adaptive filtering of spam*, Proceedings of the Second Annual Conference on Communication Networks and Services Research. Fredericton, NB, Canada 19-21, May 2004, pages: 218-224.

[11] Duan Z-H, Dong Y-F, Gopalan K. *DMTP: Controlling spam through message delivery differentiation*, COMPUTER NETWORKS, 2007, Vol.51, No.10, pages: 2616-2630.

[12] Leiba B, Fenton J. *Domain keys identified mail (DKIM): Using digital signatures for domain verification*, Proceedings of the 4th Conference on Email and Anti-Spam (CEAS 2007).

[13] Fleizach C, Voelker G M, Savage S. *Slicing spam with occam's razor*, Proceedings of the 4th Conference on Email and Anti-Spam, CEAS 2007.

[14] Deshpande V P, Erbacher R F. *An evaluation of Naive Bayesian anti-spam filtering techniques*, Proceedings of the 2007 IEEE Workshop on Information Assurance United States Military Academy, West Point, NY, 2007, pages: 333-340.

[15] Begriche Y, Labiod H. *A prior distribution for anti-spam statistical Bayesian model*, N2S'09: International Conference on Network and service Security. IEEE Piscataway, NJ, 2009, pages: 1-5.

[16] Peiyu Liu, Liwei Zhang and Zhenfang Zhu. *Research on e-mail filtering based on improved Bayesian*, Journal of Computers, 2009, Vol.4, No.3, pages: 271-275.

[17] Biju Issac, Wendy Japutra Jap, Jofry Hadi Sutanto. *Improved Bayesian Anti-Spam Filter—Implementation and Analysis on Independent Spam Corpuses*, International 2009 IEEE DOI 10.1109/ICCET.2009.170, pages: 326-330.

[18] Ming L, Yunchun L, Wei L. *Spam Filtering by Stages*, Proceedings of International Conference on Convergence Information Technology, 2007, pages: 2209-2213.

[19] Jieting Fan, Huicheng Lai. *A method of spam filtering based on SVM algorithm*, Computer Engineering and Applications, 2008, Vol.44, No.28, pages: 95-97.

[20] Wenqing Zhao and Zhu Yongli . *An email classification scheme based on decision-theoretic rough set theory and analysis of email security*. Proceedings of IEEE Region 10 Annual International Conference, Melbourne, Australia, 2007.

[21] Y. Chiu, C. Chen, B. Jeng, H. Lin. *An Alliance-based Anti-Spam Approach*, Proceedings of Third International Conference on Natural Computation, 2007, pages: 203-207.

[22] Carpinteiro OAS, Lima I, Assis JMC, de Souza ACZ, Moreira EM, Pinheiro CAM. *A neural model in anti-spam systems*, Proceedings of the 16th International Conference on Artificial Neural Networks, ICANN. 2006, Vol.4132, pages: 847-855.

[23] Zhan Chuan, Lu Xianlinag, Hou Mengshu,et al. *A LVQ-based neural network anti-spam email approach*, ACM SIGOPS Operating Systems Review, 2004, pages: 35-39.

[24] J.Park, R.Sandhu. *The UCONABC usage control Model*, ACM Transactions on Information and System Security, 2004,Vol.7, No.1, pages: 128-147.

[25] X.Zhang, F.Parisi-Presicce, R.Sandhu, J.Park. *Formal model and policy specification of usage control*, ACM Transactions on Information and System Security, 2005, Vol.8, No.4, pages: 351-387.

[26] Shaojun Ning, Hengming Zou. *Self-adaptable spam filtering based on Bayesian algorithm*, Computer Applications and Software, 2007, Vol.24, No.11, pages: 189-191.

Changda Wang received the B.S. degree from mathematic department of Soochow University in 1994, and the M.S. degree and the Ph.d degree from school of computer science of Jiangsu University in 1999 and 2006, respectively.

He is an associate professor of computer science, Jiangsu University. From April 2004 to April 2005, he was supported by the scholarship of the Jiangsu province government to conduct his research in school of computer, Carleton University, Ottawa, Canada.

Tingting Gong received the B.S. degree from computer science department of Hefei College in 2009, Now she is a Master student of school of computer science and telecommunications engineering, Jiangsu University.

Patricia Ghann received B.S. degree from natural resource management department of Kwame Nkrumah University of Science and Technology, Ghana, in 2002, and the M.S. degree from Computer Science department of De Montfort University, United Kingdom, in 2005. She was a lecturer at the Ghana Telecom University, Ghana. Now, she is a Ph.d candidate of computer science, Jiangsu University, China.