International Journal of Research
ISSN: 2348-6848 Vol-3, Special Issue-6
National Conference on Recent Engineering Trends
NCRET- 2016
Held on 16th March 2016 organized by Institute of Knowledge College of Engineering, Pune, India.

Spam filtering

Kanade Priyanka Arjun
(*BE-Dept. of CSE*) SP"s institute of knowledge college of engineering
Email Id-kanadepriya1403@gmail.com
Dhokale Vijay Nanabhau
(*BE-Dept. of CSE*) SP"s institute of knowledge college of engineering
Email Id-vijaydhokale15@gmail.com
Bhakare Mahesh Mahadev
(*BE-Dept. of CSE*) SP"s institute of knowledge college of engineering
Email Id-maheshbhakare@gmail.com
Ajay K. Gupta
(*M.Tech*) SP"s institute of knowledge college of engineering
Email Id-ajay2006-07@yahoo.co.in

Abstract- *Extending pattern classification theory and design methods to adversarial settings is thus a novel and very relevant research direction, which has not yet been pursued in a systematic way. The system evaluates at design phase the security of pattern classifiers, namely, the performance degradation under potential attacks they may incur during operation. A framework is used for evaluation of classifier security that formalizes and generalizes the training and testing datasets. Pattern classification systems are commonly used in adversarial applications, like biometric authentication, network intrusion detection, and spam filtering, in which data can be purposely manipulated by humans to undermine their operation.*

Keywords-Pattern classification; adversarial classification; performance evaluation security evaluation; robustness evaluation.

1. INTRODUCTION

Pattern classification systems are commonly used in adversarial applications, like biometric authentication, network intrusion detection, and spam filtering, in which data can be purposely manipulated by humans to undermine their operation. As this adversarial scenario is not taken into account by classical design methods, pattern classification systems may exhibit vulnerabilities, whose exploitation may severely affect their performance, and consequently limit their practical utility. Extending pattern classification theory and design methods to adversarial settings is thus a novel and very relevant research direction, which has not yet been pursued in a systematic way. In this paper, we address one of the main open issues: evaluating at design phase the security of pattern classifiers, namely the performance degradation under potential attacks they may incur during operation. We propose a framework for empirical evaluation of classifier security that formalizes and generalizes the main ideas proposed in the literature, and give examples of its use in three real applications. Reported results show that security evaluation can provide a more complete understanding of the classifier's behavior in adversarial environments, and lead to better design choices Adversarial scenarios can also occur in intelligent data analysis and information retrieval .

It is now acknowledged that, since pattern classification systems based on classical theory and design methods do not take into account adversarial settings, they exhibit vulnerabilities to several potential attacks, allowing adversaries to undermine their effectiveness. A systematic and unified treatment of this issue is thus needed to allow the trusted adoption of pattern classifiers in adversarial environments, starting from the theoretical foundations up to novel design methods, extending the classical design cycle of. In particular, three main open issues can be identified: (i) analyzing the vulnerabilities of classification algorithms, and the corresponding attacks (ii) developing novel methods to assess classifier security against these attacks, which is not possible using classical performance evaluation methods (iii) developing novel design methods to guarantee classifier security in adversarial environments.

Abbreviations and Acronyms

SMD : Spam Mail Detection
DFD : Data Flow Diagram
URL : Unified resource locator
LOC : Line of code
SRS : Software requirement specification
GUI : Graphical user interface
UML : Unified modelling language
STP : Software test plan
TH : Threshold

2. RELATED WORK

Existing System

Design assignment frameworks predicated on traditional hypothesis and conjuration systems don't consider antagonistic settings, they display susceptibilities to a few potential assaults, authorizing foes to undermine their adequacy. A deliberate and cumulated treatment of this issue is subsequently expected to authorize the trusted selection of example classifiers in ill disposed situations, beginning from the hypothetical substructures up to novel outline strategies, extending the traditional conjuration cycle of Specifically, three fundamental open issues can be recognized: 1. break down the susceptibilities of assignment calculations, and the comparing assaults. 2. Developing novel techniques to survey classifier security against these assailment, which are impractical using traditional execution assessment routines. 3. Developing novel conjuration techniques to guarantee classifier security in ill disposed situations. In the Year 2009 A. Kolcz and Teo developed method for Feature weighting for improved classifier robustness, in 6th Conf. on Email and Anti-Spam and in the year 2010 Abernethy, Chapelle and Castillo developed prototype Graph regularization methods for Web spam detection.

Disadvantages of existing system

Poor analyzing the vulnerabilities of classification algorithms, and the corresponding attacks. A malicious webmaster may manipulate search engine rankings to artificially promote their website.

Proposed System

In this work we address issues above by developing a framework for the empirical evaluation of classifier security at design phase that elongates the model cull and performance evaluation steps of the classical design cycle .We summarize anterior work, and point out three main conceptions that emerge from it. We then formalize and generalize them in our framework (Section 3). First, to pursue security in the context

of an arms race it is not sufficient to react to observed attacks, but it is additionally compulsory to proactively anticipate the adversary by prognosticating the most pertinent, potential attacks through a what-if analysis; this sanctions one to develop congruous countermeasures afore the assailment authentically occurs, according to the principle of security by design. Second, to provide practical guidelines for simulating authentic attack scenarios, we define a general model of the adversary, in terms of her goal, cognizance, and capability, which encompass and generalize models proposed in anterior work. Third, since the presence of conscientiously targeted attacks may affect the distribution of training and testing data discretely, we propose a model of the data distribution that can formally characterize this comportment, and that sanctions us to take into account an immensely colossal number of potential attacks; we withal propose an algorithm for the generation of training and testing sets to be utilized for security evaluation, which can naturally accommodate application-concrete and heuristic techniques for simulating attacks.

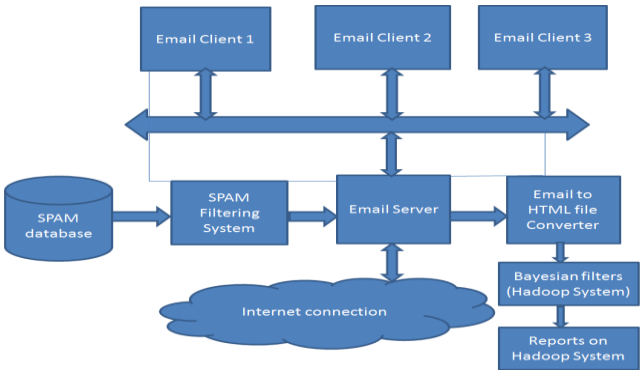

Fig: System Architecture Diagram

Modules

1. Spam Database:

   Table is maintain for-

   - Collections of users.
   - Collections of domains.
   - Collections of spam mails.
   - Collections of ham mails.

2. Spam Filtering System: Detects an attack and gives a notification of spam mails.
3. Email Server And Email Client: Used for request and response.
4. Bayesian System (Hadoop System): Analyze a number of ham and spam mails.

Advantages of proposed system

Prevents developing novel methods to assess classier security against this attack. The presence of an intelligent and adaptive adversary makes the classification problem highly non-stationary .Reduce chances of Attack by detecting it in early stage. Saves cost as this prototype can be used in multiple applications.

3. SYSTEM REQUIREMENT

a. Software requirements

1. Operating System : Windows XP/7.
2. Coding Language : JAVA/J2EE.
3. IDE : Eclipse .
4. Database : Hadoop , MYSQL.

b. Hardware requirements

1. System : Pentium IV 2.4 GHz.
2. Hard Disk : 40 GB.
3. RAM : 4 GB.
4. Monitor : 15 VGA Colour.
5. Mouse : Logitech.

4. IMPLEMENTATION

Screenshots

International Journal of Research
ISSN: 2348-6848 Vol-3, Special Issue-6
National Conference on Recent Engineering Trends
NCRET- 2016
Held on 16th March 2016 organized by Institute of Knowledge College of Engineering, Pune, India.

Fig 1: Home page

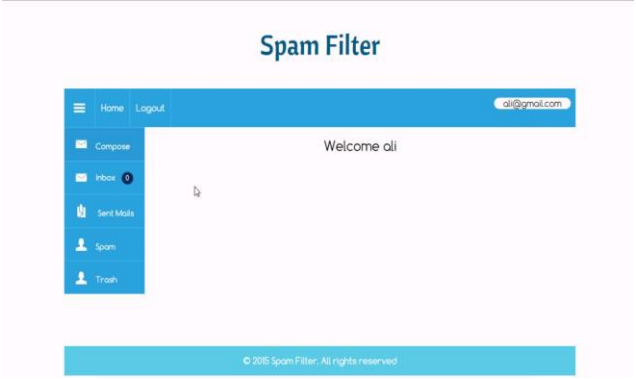

Fig 2: Account page

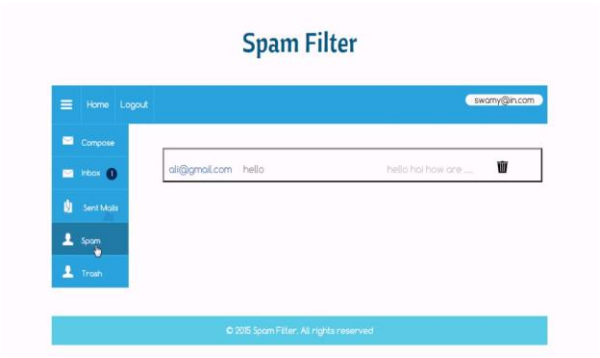

Fig 3: Spam mail

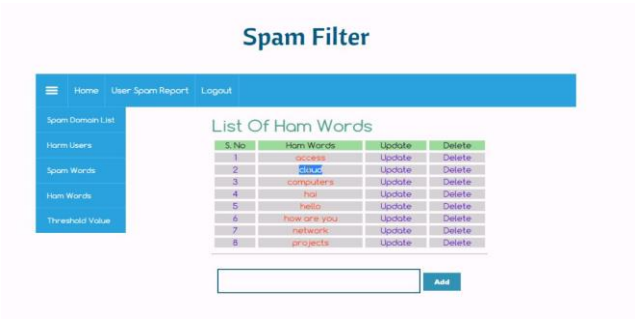

Fig 4: Ham words

Performance

the performance is usually measured in terms of genuine acceptance rate (GAR) and false acceptance rate (FAR), respectively the fraction of genuine and impostor attempts that are accepted as genuine by the system. We use here the complete ROC curve, which shows the GAR as under the above model selection setting (two classifiers, and four feature subsets) eight different classifier mode is must be evaluated. Each model is trained on TR. SVMs are implemented with the Lib SV Ms Software The C parameter of their learning algorithm is chosen by maximizing theAUC10 percent through a 5-fold cross-validation on TR. An online gradient descent algorithm is used for LR.

Fig 5- a conceptual representation in arm race in adversarial classification

5. EXPERIMENTAL RESULTS

Attacks	pattern	classifier	Potential
0.0992	2	6	10
0.0995	5	5	20
0.0996	5	5	30
0.0997	7	8	50
1	5	10	60

Table 1- classification of pattern classifier potential

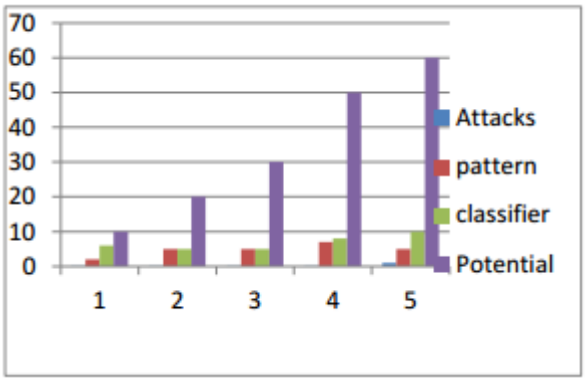

Fig 6- Function of classifier values

Each model decreases that is it drops to zero[8] for values between 3and 5 (depending on the classifier). This means that all testing spam emails gotmis classified as legitimate, after adding or obfuscating from3 to 5words.The pattern and attack classifiers perform very similarly when they are not under attack, regardless of the feature set size; therefore, according to the viewpoint of classical performance evaluation, the designer could choose any of the eight models. However, security evaluation

6. CONCLUSION:

Here we have developed the Prototype for SPAM Filtering which provides security to users against SPAM emails. Here we also analysed that with the help of HADOOP framework we could provide advance Analysis to the Administrator so that Admin can take the appropriate action.

7. REFERENCES

[1] R.N. Rodrigues, L.L. Ling, and V. Govindaraju, "Robustness of Multimodal Biometric Fusion Methods against Spoof Attacks," J. Visual Languages and Computing, vol. 20, no. 3, pp. 169-179, 2009.

[2] P. Johnson, B. Tan, and S. Schuckers, "Multimodal Fusion Vulnerability to Non-Zero Effort (Spoof) Imposters," Proc. IEEE Int'l Workshop Information Forensics and Security, pp. 1-5, 2010.

[3] P. Fogla, M. Sharif, R. Perdisci, O. Kolesnikov, and W. Lee, "Polymorphic blending attacks," in Proc. 15th Conf. on USENIX Security Symp. CA, USA: USENIX Association, 2006.

[4] D. Lowd and C. Meek, "Good word attacks on statistical spam filters," in 2nd Conf. on Email and Anti-Spam, CA, USA, 2005.

[5] A. Kolcz and C. H. Teo, "Feature weighting for improved classifier robustness," in 6th Conf. on Email and Anti-Spam, CA, USA, 2009.

[6] D. Fetterly, "Adversarial information retrieval: The manipulation of web content," ACM Computing Reviews, 2007.

[7] M. Barreno, B. Nelson, A. Joseph, and J. Tygar, "The Security of Machine Learning," Machine Learning, vol. 81, pp. 121-148, 2010.

[8] Abernethy, J., O. Chapelle, and C. Castillo: 2010, 'Graph regularization methods for Web spam detection'. Machine Learning Journal 81(2). DOI: 10.1007/s10994-009-5154-2.