IJCSN International Journal of Computer Science and Network, Volume 4, Issue 5, October 2015
ISSN (Online) : 2277-5420 www.IJCSN.org
Impact Factor: 0.417

An Experimental Study on Authorship Identification for Cyber Forensics

[1] Smita Nirkhi, [2] Dr. R. V. Dharaskar , [3] Dr. V. M. Thakare

[1] Research Scholar, G.H. Raisoni, Department of Computer Science
Nagpur, Maharashtra, India

[2] Former Director DES(Disha-DIMAT) Group of Institutes
Raipur, Chhattisgarh, India

[3] Department of Computer Science & Engineering
S.G.B Amravati University, Amravati, Maharashtra, India

Abstract - Authorship Identification is subfield of authorship analysis deals with finding the plausible author of anonymous messages. The Authorship identification problem of online messages is challenging task because cyber predators make use of obscurity of Cyberspace and conceal the identity. By performing the forensic analysis of online messages, empirical evidence can be collected. These evidences can be used to prosecute the cybercriminal in a court and punish the guilty. This way cybercrimes can be minimized up to certain extent by detecting the true indentities.Therefore it is required to build up innovative tools & techniques to appropriately analyze large volumes of suspicious online messages. This paper compares the Performance of various classifiers in terms of accuracy for authorship identification task of online messages. Support Vector Machines, KNN, and Naïve Bayes classifiers are used for performing experimentation .This paper also investigate the appropriate classifier for solving authorship of anonymous online messages in the context of cyber forensics.

Keywords - *Authorship Identification, Cybercrime, Cyber Forensics, Support Vector Machine, K-NN, Naïve Bayes.*

1. Introduction

Internet provides us the convenient and efficient platform for sharing and exchanging information across the world. It has connected the world through a collective area called cyberspace. This connectivity gives many opportunities and advantages to internet users at large. However, this connectivity gives opportunity and possibility to different criminal elements for committing crimes. It is a very crucial task to identify who is behind the cyber crime. To punish the guilty, tremendous forensics and investigation capabilities are required. Cyber crime and cyber attacks are increasing all over the world. In India during last 17 years conviction rate is 0.5% and cyber crime growth rate has increased to 107%.Every year there is rise in cyber crime cases. The main cause behind these criminal activities is anonymous and decentralized nature of internet. Besides, there are many methods to hide the identity and pretending like defender [10].

Many criminals hide themselves among legitimate internet user to commit the crime and launch cyber attack. They compromise the computer and make it a part of botnet to perform all illegal activities. Due to available techniques of hiding the true identity it has become quite difficult to prove that a particular crime has been committed by him or her [3].

In this context, authorship attribution also called as Authorship identification is an important technique to detect the culprit. Data mining techniques along with the profile of accused to attribute the author of written text is the only emerging area of cybercrime investigation [5].
In India, there is lack of regulation and guidelines for effective investigation of cybercrimes. Therefore techno legal skill development is necessary. Authorship Attribution of anonymous online messages is an upcoming research area in recent years. Earlier it has been studied in many fields like Literary and poetic work, social psychology [2].

The first section of this paper explains the problem definition and various approaches used for authorship identification, secondly the methodology used for solving the attribution of author. The next section deals with

performance evaluation of various classifiers followed by conclusion.

2. Problem Description

There are three types of authorship analysis scenarios [1]:

Scenario1: There are many suspects and we have to attribute the text to one of them.

Scenario2: There is one suspect and we must determine if the suspect is the author of the text or not.

Scenario3: There are no suspects. The task is to provide as much psychological or demographic information about the author as possible.

Scenario1: When there are many suspects and we have to attribute the text to one of them. As shown in Figure 1, the authorship identification is to determine the probability of a candidate author who wrote a text, i.e. which candidate is the most likely author of the disputed text. Given a text *D* and a set of candidate authors *C*= {*A*1, *A*2... *An*}, matching the author from set *C* with the text in dispute *d* is the problem to solve in this scenario. If the author of text *d* is definitely in the set *C*, this is called a closed set problem; if the author of text *d* can also be someone who is not included in the set *C*, this is well-known as an open set problem. Different authors have different writing styles, based on which one author can be distinguished from the others.

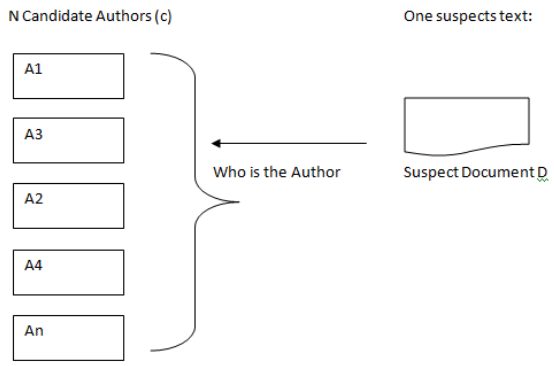

Fig. 1 Scenario1.

Scenario2: when there is one suspect and we must determine if the suspect is the author of the text or not. Given a suspicious text *d* and a set of known texts from one author *A*= {*D*1, *D*2 ...*Dn*}. The problem is to validate whether the text *D* in dispute was written by the same author who wrote the known texts from the set *A*. The same author has some writing styles that are believed to be difficult to hide in a short period of time, and therefore based on these writing styles a document claimed from the same author can be verified.

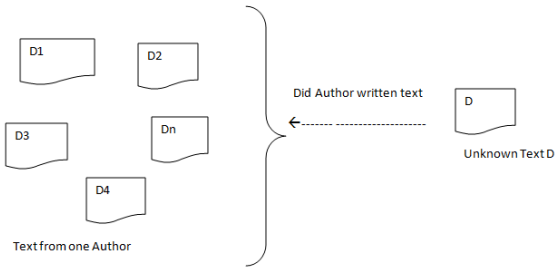

Fig. 2 Scenario2

Scenario 3: There are no suspects. The task is to provide as much psychological or demographic information about the author as possible like age, gender etc.

In the Scenerio3 there is one author and the study is to generate a demographic profile (for instance gender, age, native language) of the author based on the given documents [2].

The underlying rationale of author profiling is: Authors' written documents can reveal some of their personal characteristics without their consciousness, based on which a demographic author profile can be created.

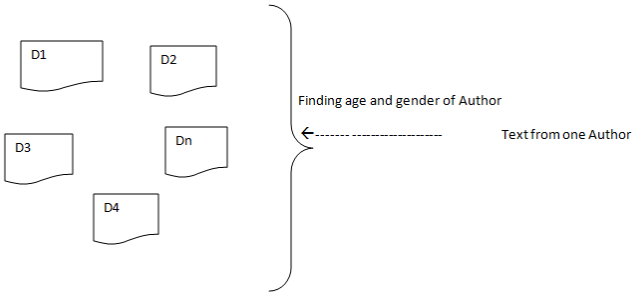

Fig. 3 Scenario3

2.1 Approaches for Authorship Identification

There are various authorship attribution methods, and according to Stamatatos [4],all the methods can be classified into two groups: profile-based approach and instance-based approach.

2.1.1 Profile Based Approach

As shown in Figure 4 ,the profile-based approach, which is a process of concatenating all the training texts of one author and generating an author profile. The features of each author are extracted from the concatenated text. Extracted features are used in the attribution model to determine the most likely author of the dispute text. However, a profile-based approach is criticized for losing much information because of the generating profile-based feature process which is required to remove all the dissimilar contents from the same author.

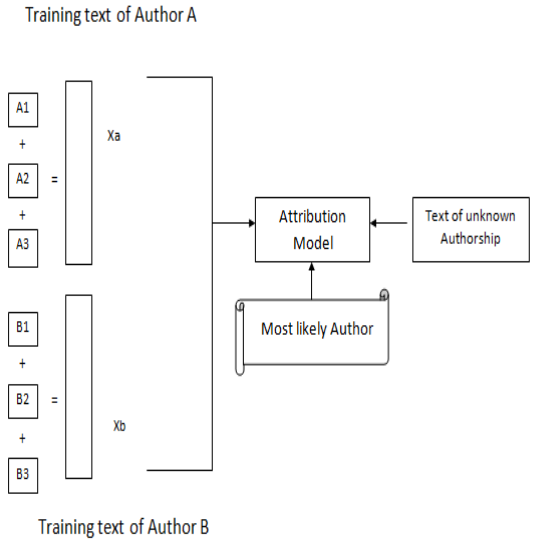

Fig. 4 Profile Based Approach

As shown in figure-4, feature vector of text is denoted by x. To represent the author A in term of his profile is denoted by xA and similarly to represent unknown author of the text, xU is used. The model then computes the distance between profile of unknown author and each known author.

2.1.2 Instance Based Approach

On the contrary, instance-based approach, which is used in most of the contemporary authorship attribution research, can keep most of the information from the given texts. Instance based approach, as shown in figure 5, includes every instance of training text to conclude the author of unseen text. This is discriminative approach where xa1 indicates the vector for first training sample of author A. These vectors are given as input to train the classifier and tested against the unknown authorship to find out most likely author.

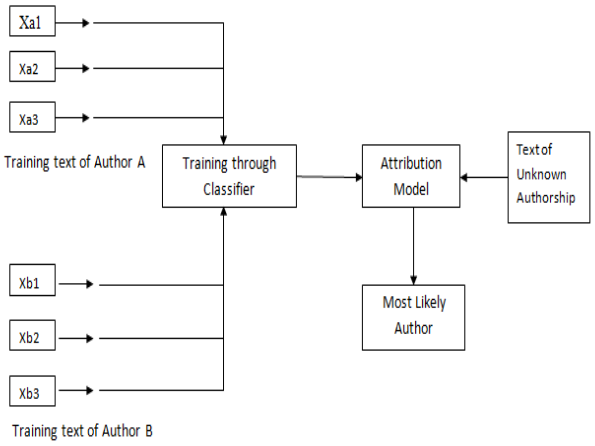

Fig. 5 Instance Based Approach

3. Methodology for Solving the Problem

Two critical research issues which influence the performance of authorship analysis is finding out the effective discriminators and approach to discriminating texts by authors based on the selected features[6]. Figure 6 shows the methodology for solving the problem. The corpus is divided into training and testing samples. Instance based approach is used. Training samples are used for Creation of Feature vector and given as input to the classifier.

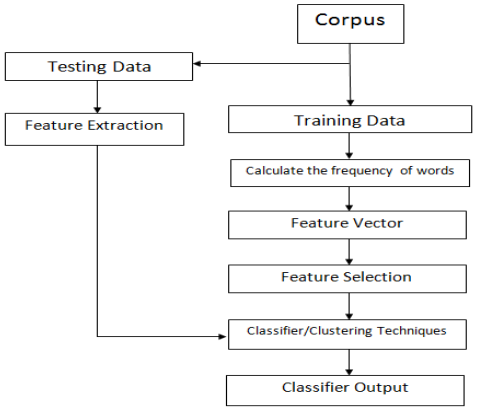

Fig. 6 Proposed Approach

3.1 Features Used

In the proposed approach, features used are Lexical Features, Character Features, unique word-print, co-occurrence of words.

1. Lexical Features includes word frequencies, word n-gram
2. Character Features are letters, digits and character n-grams.
3. Unique word-print means the list of unique words used by particular author and not by any other author.
4. Finding text style with co-occurrence of words

3.2 Classifiers Used

Authorship identification is a single label and multiclass text classification problem[4]. Therefore selection of classifier used for performing attribution should be done carefully. This experimentation is carried out to evaluate the performance of classifiers with the selected feature sets. Classifiers used for experimentation are Support Vector Machine, Naive Bayes, and K-NN.

3.3 Corpus Used

Brennan-Greenstadt Corpus [8][9] is used for experimentation pupose.It was prepared at Drexel University. Two datasets are prepared by them. One dataset contain 12 authors and another contains 45 different authors. Authors were asked to submit 5000 words of their writing for demographic survey. The participated authors were colleagues and friends.

4. Experimental Results

To carry out experimentation, setup was run for different number authors. Numbers of authors are 2, 5,10,20,30 for all the three classifier and same set of features as mentioned in section 3.1 .Different training and testing samples are used for conducting experiments. Accuracy is calculated by dividing the total number of messages to total number of messages whose author is correctly identified. N-gram approach is used with word n-gram=1 .Results are shown in Table1, Table2, Table3 respectively for support vector machine, naive bayes, K-NN classifier.

Table1.Performance of Support Vector Machine

Experiment id	*Number of Authors*	*Accuracy*	*Training Samples*	*Testing Samples*
1	*2*	*89%*	*25*	*6*
2	*5*	*93%*	*71*	*15*
3	*10*	*91%*	*110*	*49*
4	*20*	*81%*	*244*	*61*
5	*30*	*77%*	*340*	*105*
6	*40*	*70%*	*445*	*142*
7	*45*	*69%*	*493*	*162*

Table2.Performance of Naïve Bayes Classifier

Experiment id	*Number of Authors*	*Accuracy*	*Training Samples*	*Testing Samples*
1	*2*	*97.2%*	*25*	*6*
2	*5*	*91%*	*71*	*15*
3	*10*	*85%*	*110*	*49*
4	*20*	*65%*	*244*	*61*
5	*30*	*48%*	*340*	*105*
6	*40*	*41%*	*445*	*142*
7	*45*	*43%*	*493*	*162*

Table3.Performance of Naïve Bayes Classifier

Experiment id	*Number of Authors*	*Accuracy*	*Training Samples*	*Testing Samples*
1	*2*	*95%*	*25*	*6*
2	*5*	*87%*	*71*	*15*
3	*10*	*79%*	*110*	*49*
4	*20*	*74%*	*244*	*61*
5	*30*	*66%*	*340*	*105*
6	*40*	*47%*	*445*	*142*
7	*45*	*58%*	*493*	*162*

5. Conclusions

We have designed and proposed a technique for performing forensics of online messages to help the investigators to collect practical evidence by automatically analyzing large collection of suspicious online messages. The analysis is performed on the textual contents of a message. The proposed technique used the frequency of common words from the training and testing data. Function word usage and unique word usage by each author can work as discriminator to uniquely identify the plausible author of disputed text. SVM outperforms Naïve Bays and K-NN classifiers. Different parameter settings of authorship identification had an impact on performance.

References

[1] S.Argamon, M. Koppel, J. Pennebaker and J.Schler, "Automatically Profiling the Author of an Anonymous Text ", Communications of the ACM, 52 (2): 119-123, 2009.

[2] Koppel, M., Schler, J., & Argamon, S. Computational Methods in Authorship Attribution .Journal of the American Society for Information Science and Technology,2009, ,(pp. 60(1):9–26).

[3] R. Zheng, J. Li, H. Chen, Z. Huang. "A framework for authorship identification of online messages: Writing-style features and classification techniques", Journal of the American Society for Information Science and Technology, 57(**3**), pp.378-393, 2006.

[4] Stamatatos, E. (2009). A Survey of Modern Authorship Attribution Methods. *Journal of the American Society for Information Science and Technology, 60*(3), 238-556.

[5] R. Hadjidj, M. Debbabi, H. Lounis, F. Iqbal, A. Szporer, and D. Benredjem, "Towards an integrated e-mail forensic analysis framework", Digital Investigation, 5(3-4):124 – 137, 2009.

[6] Smita Nirkhi and R.V.Dharaskar," Comparative study of authorship identification techniques for cyber forensics analysis", International journal of advanced computer science and application, vol. 4, no. 5, 2013

[7] Ahmed Abbasi and Hsinchun Chen," Writeprints: A stylometric approach to identity-level identification and similarity detection in cyberspace", ACM Transactions on Information Systems (TOIS), 26(2), 2008.

[8] Michael Brennan, Sadia Afroz, and Rachel Greenstadt,"Adversarial stylometry: Circumventing authorship recognition to preserve privacy and anonymity", ACM Trans. Inf. Syst. Security. 15, 3, Article 12 (November 2012).

[9] Michael Brennan and Rachel Greenstadt. Practical Attacks Against Authorship Recognition Techniques in Proceedings of the Twenty-First Conference on Innovative Applications of Artificial Intelligence (IAAI), Pasadena, California, July 2009.

[10] S.M.Nirkhi, R. V. Dharaskar, V.M.Thakre, "Analysis of online messages for identity tracing in cybercrime investigation", 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), pp. 300 - 305, 2012.

First Author

Asst. Prof. Ms. S. M. Nirkhi has completed M.Tech in Computer Science & Engineering & currently Pursuing PHD in computer science. She has received RPS grant of 8 lakhs from AICTE for her Research. She has attended 6 STTPworkshops along with other training programs. She has Published 20 papers in international conferences & 12 papers in international journals. She had presented paper at International Conference at Singapore. She has 13 years of professional experience. Her areas of interest include soft computing, Data mining, web mining, pattern recognition, MANET, Digital Forensics.