Unsupervised authorship analysis of phishing webpages.

Published version available at
http://dx.doi.org/10.1109/ISCIT.2012.6380857

Unsupervised Authorship Analysis of Phishing Webpages

Robert Layton
Internet Commerce Security Laboratory
University of Ballarat
r.layton@icsl.ballarat.edu.au

Paul Watters
Internet Commerce Security Laboratory
University of Ballarat
p.watters@icsl.ballarat.edu.au

Richard Dazeley
Data Mining and Informatics Research Group
University of Ballarat
r.dazeley@ballarat.edu.au

***Abstract*—Authorship analysis on phishing websites enables the investigation of phishing attacks, beyond basic analysis. In authorship analysis, salient features from documents are used to determine properties about the author, such as which of a set of candidate authors wrote a given document. In *unsupervised* authorship analysis, the aim is to group documents such that all documents by one author are grouped together. Applying this to cyber-attacks shows the size and scope of attacks from specific groups. This in turn allows investigators to focus their attention on specific attacking groups rather than trying to profile multiple independent attackers. In this paper, we analyse phishing websites using the current state of the art unsupervised authorship analysis method, called NUANCE. The results indicate that the application produces clusters which correlate strongly to authorship, evaluated using expert knowledge and external information as well as showing an improvement over a previous approach with known flaws.**

I. Introduction

Phishing is a type of attack in which a target institution is mimicked to defraud a victim, falling under the category of 'information gathering through apparently authentic appeals' [25]. In an online service phishing attack, information such as usernames and passwords can be obtained by creating a website that mimics the trusted party's legitimate website, called a *phishing website*. The Anti-Phishing Working Group reports that more than thirty thousand phishing websites were detected each month in the second quarter of 2010 [3]. Identifying information, such as online banking passwords or driver's license numbers, can then be used to access money, obtain financial loans or register for services (such as a mobile phone or car hire) in a victims name. An identity theft such as this has been shown to have a large negative impact on the victim, with the victim recovery process proving difficult, costly and traumatic [27], [23].

Despite advances in anti-phishing technologies and increasing public awareness, phishing attacks continue to cause extensive damage. The French Ministry of Finance was infiltrated by a targeted phishing attack, called *spear* phishing, in March of 2011 [4] leaking information relating to a G-20 economic group to attackers who remain unidentified. In February of 2011, the Canadian Government was forced to restrict web access to their online services due to a large number of targeted phishing attacks [5]. Email accounts originally compromised by the phishing attack were then used to spread a virus to other parts of the compromised network, leading to a greater level of penetration. Personal and private information can also be lost or leaked on a large scale through unintentional mistakes or ignorance [30] however the focus of this paper is on intentional acts, rather than these accidental losses.

The objective of this research is to analyse a set of phishing websites to determine the size and scope of the operations creating them. In order to perform this, a methodology able to perform automated and unsupervised authorship analysis was used, called NUANCE (n-gram Unsupervised Automated Natural Cluster Ensemble). This research publishes results that are currently in use in industry to assist with phishing response. This analysis could be applied to other cybercrimes involving written documents such as fraud detection and plagiarism investigation.

II. Authorship Analysis in Cybercrime

Cybercrime is difficult to investigate for a number of reasons, with one of those being the lack of appropriate models for determining the provenance of attacks. In order for this to be achieved, attacks need to be linked together in a reliable and robust method. One method for linking documents by provenance is **Authorship Analysis**, described as 'the process of examining the characteristics of a piece of work in order to draw conclusions on its authorship' [31, p.60], normally concerned with discovering *who* wrote a particular document. For cyberattacks consisting of written documents, such as phishing, authorship analysis could lead to insights into the attack's provenance.

Cybercrime is an important application of authorship analysis due to its increasing impact on today's lifestyle. The nature of the Internet provides an easy way to maintain anonymity while still allowing Internet based crime to gain large results, creating a need for indirect methods of attack attribution. The types of cybercrimes that authorship analysis has been applied to vary wildly. Examples include webpage spam [29], [28], malware [11], [9], pornography [21], online terrorism postings [1],web forum postings [24] and malware code [17], [16], [12].

A. Local n-grams Profiling

Local n-grams (LNG) methodologies are effective forms of authorship analysis that are language independent, automatable

and highly accurate. LNG is based on a concept of an 'author profile', which is 'the set of the L most frequent n-grams with their normalised frequencies' [15], for a given author. From these author profiles, several methods in the literature exist for determining the distance between two profiles for authors A_1 and A_2.

The original method of this type is the Common n-grams (CNG) method [15]. The CNG method uses the Relative Distance (RD) between two document profiles or author profiles with lower distances considered to infer that two profiles are from the same author. The frequencies for the L most frequently occurring n-grams are compared using equation 1, to determine a distance between the two profiles [15].

$$K = \sum_{x \in X_{P_1} \cup X_{P_2}} \left(\frac{2 \cdot (P_1(x) - P_2(x))}{P_1(x) + P_2(x)} \right)^2 \quad (1)$$

Where $P_i(x)$ is the frequency of term x in profile P_i and X_{P_i} is the set of all n-grams occurring in profile P_i.

Other Local n-gram methodologies for authorship analysis include the Source Code Author Profiling (SCAP) method [9], [10], the Recentred Local Profiles (RLP) method [20] and the Weighted Profile Intersection (WPI) method [7]. Each of these other methods employ a similar methodology to CNG with variations in the way in which profiles are generated and distances between profiles are calculated.

B. Unsupervised Authorship Analysis

The above listed methods are all supervised methods for authorship analysis. Without labelled corpora for discovering models, these methods cannot be applied to discover the authorship of phishing webpages and other cybercrimes. For this reason, *unsupervised* methods of authorship analysis are needed. There has been a traditional lack of research in the field of unsupervised authorship analysis, although recent research is working to resolve this gap in the literature. Juola surveyed the field of authorship attribution and listed only visualisation techniques for unsupervised authorship attribution [14]. Other survey papers of the field also failed to list any unsupervised authorship analysis methods that are not visualisation methods [26].

Since those listed survey papers, there has been some research on the field. Research by [13] performs a clustering of emails by authorship. Once this clustering is performed, the resulting clusters are then analysed using the Writeprints [2] technique used earlier to discover patterns that lead to the creation of the cluster. The clustering algorithms chosen require an estimate of the number of clusters, which was chosen in these experiments as the number of authors in the dataset - the 'correct' value of k. This is not practical for a real world application, where the correct value cannot be known *a-priori*.

[29], [28] use similarity detection focusing on mainly the non-alphanumeric characters in HTML source code to detect clusters of spam campaigns that were generated by the same automated process. The methods do not attempt to cluster by authorship, instead focusing on campaigns of spam - those generated by the same generation software but not necessarily all spam by the same author.

The USCAP methodology by [19] is an automated methodology for clustering documents and was performed on a set of phishing websites. This technique uses the SCAP methodology, a form of LNG using just the occurrence rather than the frequencies of n-grams, on document profiles and then clusters using SCAP's distance metric. The results showed a high precision but low recall; while there was evidence that the websites within each cluster 'belonged' together, there was substantial evidence found that the discovered clusters should be linked. This led to the conclusion that the discovered clusters represented campaigns of attacks, rather than authorship itself. In this research, the results of USCAP will be used as the baseline research for improvement.

The n-gram Unsupervised Automated Natural Cluster Ensemble (NUANCE) methodology proposed in [18] enables the clustering of documents by authorship and was shown to produce clusters with a high correlation to true authorship. The NUANCE methodology takes as input a set of documents and creates local n-gram profiles of each document to form an array of representations of the data. Each representation is then clustered multiple times using the k-means algorithm with randomised initialisation parameters. The resulting clusters are used to form a co-association matrix C such that $C_{i,j}$ is the number of times that instances i and j are clustered together. The resulting matrix C is then used to form a dendrogram, which is then cut using the Iterative Positive Silhoeutte (IPS) procedure, which cuts the dendrogram into k clusters where $k-1$ clusters results in a negative median Silhouette Coefficient.

III. Testing Methodology

The NUANCE methodology was applied to a dataset of phishing websites described in subsection III-A. The resulting clusters were then evaluated using a number of methods based on external information validating the clustering results. The evaluations are described in section IV.

The methodology in [18] used a large number of parameters and each of the three local n-gram methods described earlier (CNG, SCAP and RLP). However, the results of a leave-one-out ensemble consistently chose the CNG method with the following parameters for each of the different authorship problems in [18], which are $n = 3$, $L = 3000$; $n = 4$, $L = 5000$; $n = 4$, $L = 7500$; $n = 5$, $L = 7500$; and $n = 5$, $L = 10000$. Due to the stability of the choice of these parameters, they were chosen for the ensemble in this application. The application area of this research is of a different nature than the corpora used to select these parameters, as the HTML is much more formalised that the written documents used in the initial NUANCE training and contains many more punctuation characters. However both sets are in English and previous work using Local n-grams suggests similar n values, at least for classification performance. As an example, the work of [8] on programming source code (specifically C++

and Java) shows high accuracy for $n \geq 3$ with high L values, overlapping significantly with the parameter choices used here.

A. Testing Dataset

The dataset used for this experiment is a set of over 800 phishing websites targeting a major Australian financial institution. The phishing websites were collected between 2007 and 2011 by a monitoring system. The monitoring system used automated methods to collect phishing emails and discovery phishing websites from them. The URL is accessed and the resulting webpage verified as phishing by a human expert, ensuring little noise in the dataset.

IV. Evaluation Methods

The evaluation of the application of unsupervised methods is always a difficult task, as ground truth does not exist to provide fully independent feedback on the effectiveness of the result. To compensate for this, we use four independent methods of evaluation to assert quality. The first evaluation criteria is the site validation score, ensuring that likely matches by URL are within the same cluster. The second is the use of URL based rules to determine the purity of each cluster, called the URL pattern score. The third asserts that the clusters found through an application of USCAP, which was found to find phishing campaigns, form a subset of the clusters found by NUANCE. The fourth is an evaluation based on an expert's labelling of a sample of the websites. The results of the first, second and fourth evaluation methods will be compared against the results obtained from applying USCAP, under the expectation that NUANCE performs better in all cases.

A. URL Domain Evaluation

The URL validation criteria makes an assumption that if two phishing websites are hosted on the same URL, then they are likely to be created by the same author. This assumption is not a perfect one; a trivial counter-example would be the use of a URL as a free hosting platform, which is leveraged by different phishing groups. However it is considered likely that the assumption would hold true *in most cases*. This assumption has been used in previous phishing clustering research [19].

The **site validation score** is measured as the percentage of times that two phishing websites hosted on the same URL are clustered together. A value close to 1.0 is considered indicative of clusters with a high precision, while a low value would indicate incorrect clusters.

B. URL Based Rules

URL patterns are used by industry experts to determine which phishing group attacks come from, and are considered to be fairly accurate for some groups. A URL pattern is constructed based on observations by an expert to indicate that a phishing attack probably belongs to a known group. Different groups use different URL patterns themselves - a group could use more than one pattern. For this reason, it is considered that an authorship cluster should *contain* all phishing attacks hosted on URLs following a pattern. However it is not expected that all phishing attacks in an authorship cluster follow a URL pattern. The **URL pattern score** is therefore defined as the purity of the attacks matching a known URL pattern.

The full list of URL patterns used in this research is considered a trade secret and cannot be published, however it was created with the help of a cybercrime expert and is considered accurate based on their experience. Three examples of the URL patterns are given below.:

- Matches the pattern: `http://%/%<BANKING_DOMAIN>/%` [1]
- Contains the phrase: `%.user%` [2]
- Contains the phrase: `%_email=%`[3]

The evaluation score for the URL patterns is calculated as the mean purity of the labels for each URL pattern. The concept of purity asserts that members of one class appear predominately within another class [6] and is related to the concept of precision. It is often used as a supervised metric, in which the predicted classes are asserted to be subsets of actual classes, or vice versa. The purity for a class O_i^1 for class O_j^2 is the proportion of instances in O_i^1 that are within the comparison cluster O_j^2. The classes (or clusters) O_i^1 and O_j^2 (for any value of j) are often from a differing clustering or partitioning of a dataset. The purity for a class O_i^1 for class O_j^2 is given the notation $purity(O_i^1, O_j^2)$ and calculated using equation 2, where n is the number of instances in the dataset.

$$purity(O_i^1, O_j^2) = \frac{1}{n}\left(|O_i^1 \cap O_j^2|\right) \tag{2}$$

The purity for a class O_i^1 is the maximum value of the above metric for all possible other classes O_j^2. Formally, the purity for class O_i^1 is given as $purity(O_i^1)$ and calculated using equation 3.

$$purity(O_i^1) = \max_j purity(O_i^1, O_j^2) \tag{3}$$

The purity for a set of classes $O^1 = \{O_1^1, ...O_i^1, ...O_k^1\}$ of a dataset is given as the mean of the purity for each class (or cluster) in O^1.

C. USCAP Sub-cluster Purity

The clusters obtained by the USCAP methodology were shown to have a high precision but a low recall [19]. The instances in each cluster were likely to have belonged to a single author but there was evidence that clusters should be joined. The USCAP clusters can then be used as a further purity based evaluation of the NUANCE clusters.

The **USCAP purity** is therefore measured using a method similar to the URL pattern score. For each USCAP cluster K^U, the NUANCE cluster K^N containing the most instances from K^U is noted as the 'expected cluster'. The percentage of instances from K^U in K^N is given as the USCAP purity score.

[1] Indicating that the phishing kit targets many websites, or that it is trying to fool the user by including the domain in the folders list. The percentage sign (%) is a wildcard matching none or many characters.

[2] A pattern previously used by *rock-phish*.

[3] This phrase is usually followed by an email identifier, which can be used to verify and track email addresses caught by the phishing attacks.

Values closer to 1 indicate that the entire USCAP cluster is contained within a NUANCE cluster, suggesting further that the NUANCE clusters are of a high precision.

D. Expert Evaluation

A team of cybercrime experts were used to evaluate the resulting clusters, by providing their own labelling on a sample of the phishing webpages. The experts are part of a commercial security team with a large amount of experience in cybercrime investigation, including phishing attacks. The labels provided by the experts were their evaluation of which *phishing kit* was behind the attack. Different phishing kits likely share the same author. As a result, the labels provided by the experts should form subsets of the clusters found by NUANCE. A sample was chosen of 30.6% of the websites, of which the expert team was able to identify the phishing kit behind 52% of these attacks. This sample size corresponded to a 95% confidence of 5% error rate on the websites in the population, however not all of the attacks within the sample were able to be identified. The purity measure was calculated for the labels with values close to 1 indicating that the phishing kit labels belong within the NUANCE clusters, which was the expected scenario.

Stratified sampling was used, taking a proportional sample from each cluster, which was chosen to ensure that most clusters were represented. Any cluster with three or fewer documents was ignored in the sample (as the expected sample size is less than 1), while all larger cluster have a proportionate amount.

V. Results and Discussion

The application of the automatic and unsupervised authorship analysis methodology to the phishing webpage dataset resulted in 21 clusters being found. Seven clusters had just one member, with another eight clusters having less than ten members. The largest cluster contained 75.4% of all of the phishing webpages while the next largest cluster contained 10.6% of entries.

A. Site Validation Score

The site validation score was significantly higher for NUANCE than was achieved for USCAP when applied to the phishing websites. The site validation score was 0.979 indicating that when two phishing webpages were hosted on the same domain, 97.9% of the time they were clustered together. There were 285 joint uses of URLs with 6 URLs appearing in multiple clusters. Three of the errors occurred due to URLs appearing in both clusters 7 and 15. The other three appeared in K_0^N (NUANCE cluster numbered 0), with each URL also appearing in clusters K_1^N, K_5^N and K_{20}^N.

In contrast, the USCAP methodology scored just 0.902. The USCAP methodology made 26 errors from 266 decisions a significant increase on a smaller decision set. In the original USCAP results, 12 errors came from overlap between clusters K_8^U and K_9^U (USCAP clusters 8 and 9). Both of these clusters are subsets of cluster K_0^N, strongly supporting the conclusion that these two clusters be merged. The increase in the site validation score also suggests strong evidence showing that NUANCE provides clusters with a higher precision than those obtained by USCAP.

B. URL Pattern Score

The URL pattern score was also high with a score of 0.970, indicating that the attacks corresponding to each of the known cluster patterns were in the same cluster for an average of 97% of the time. Of the 21 URL patterns used in the evaluation, 14 had a perfect purity of 1.0 which indicated that all phishing attacks matching a URL pattern appeared in the same NUANCE cluster. Four of the patterns had a purity above 0.9, indicating that 90% of the attacks fell within the same cluster, while the remaining three patterns had purity scores above 0.8.

Three URL patterns had significant overlap between clusters with more than 3 attacks appearing in the non-dominate cluster. All three of the URL patterns were variants of `BANK_DOMAIN`, suggesting that these rules may not be accurate. One of these patterns linked K_0^N and K_1^N using the main banking domain. Another URL pattern linked clusters K_0^N and K_{15}^N using the full online banking domain path (including subdomain). The final pattern with errors linked K_0^N and K_{15}^N again using the banking domain as a folder name.

The USCAP results scored 0.861 for the URL pattern scores, representing a marked improvement by NUANCE. Again this shows a higher precision of the clusters obtained by NUANCE compared to USCAP. In the next section, the clusters resulting from both methodologies are compared.

C. USCAP Purity Score

The USCAP purity score was 0.945, indicating strongly that the clusters from the application of USCAP are typically subsets of the clusters from NUANCE. Only two USCAP clusters were not subsets of a NUANCE cluster. The first was USCAP cluster 7 (K_7^U), which had attacks appearing in fourteen NUANCE clusters overall. It is worth noting that K_7^U had a poor intra-cluster distance and was similar to other clusters, K_8^U and K_9^U. This suggests that the initial clustering by USCAP was poor for this cluster. The second was K_{12}^U, appearing in two NUANCE clusters. Together, these two clusters account for 41 phishing websites. Figure 1 in the Appendix illustrates the relationship between clusters from NUANCE (the 'N' clusters) and those from USCAP (the 'U' clusters). In this diagram, two clusters are linked if they jointly contain the same attacks. Clusters with full purity map completely into another cluster, as is the case with many of the clusters shown.

D. Expert Evaluation

The expert evaluation resulted in 16% of all phishing attacks being labelled. There were 74 different labels, with many attacks being the only instance of that phishing kit within the sample. In total there were 25 labels on two or more attacks and just six of those labels on five or more attacks. The purity for these labels would be artificially high due to the large

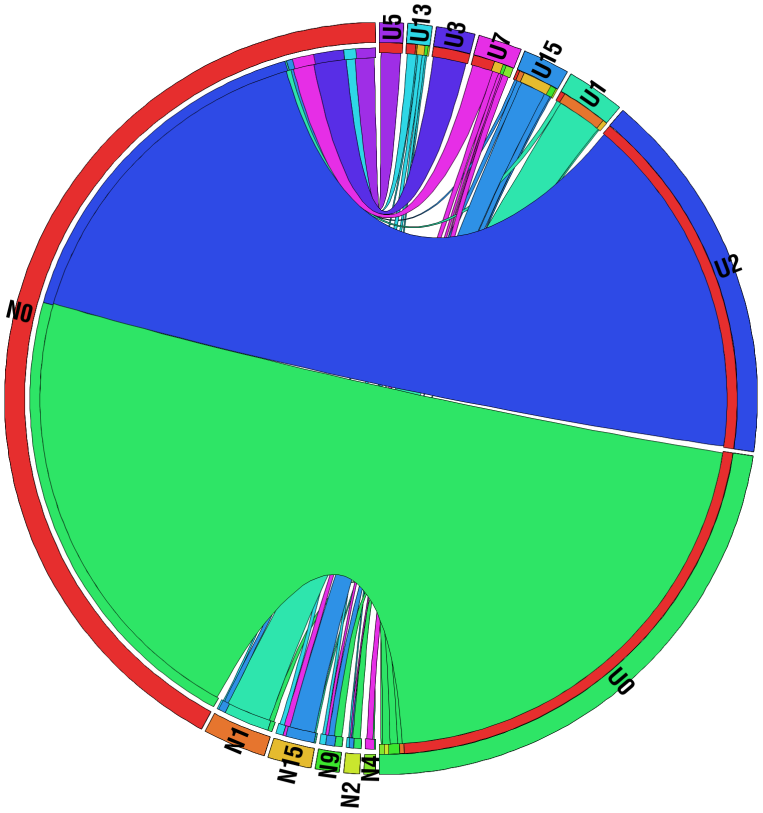

Fig. 1. Circos diagram of the relationship between the USCAP (U) and NUANCE (N) clusters with more than 10 members. Coloured bands show the size of overlap between the clusters labelled at either end.

number of single label instances (as each of these would have a purity of 1.0). As a result, all clusters with just one attack were excluded from the calculation. The resulting purity was 0.980, with just one label producing an error. The USCAP clusters also had the same and only error.

The label that produced an error was on just two attacks, with one attack in cluster K_2^N and one attack in cluster K_{15}^N. Examination of the source code indicates that this is indeed an error with the NUANCE clustering and not with the expert labelling. Both of these attacks have a similar structure, with the structure of the HTML and CSS code (contained within the HTML) very similar. The differences are semantic differences within the HTML, the most notable is the change between class names, which are of different forms. The HTML in one attack is also formatted with more whitespace, with all sub-elements tabulated an extra level than their parent elements. The strictness of the formatting suggests that this may be the result of using an automatic formatter rather than manually performed. This indicates a potential strategy for phishers to evade future authorship analysis. Automatically post-processing the HTML can remove stylistic choices which may dampen the quality of future attribution. This problem is inherent in any non-structural form of authorship analysis and must be addressed in future research.

Metric	Result
Site Validation	0.979
URL Pattern	0.970
USCAP Purity	0.945
Expert Evaluation	0.980

TABLE I
SUMMARY OF THE RESULTS USING DIFFERENT EVALUATION METRICS.

VI. CONCLUSIONS

In this research, the NUANCE methodology was applied to a set of phishing attacks targeting a major Australian banking institution. The results are compared against four types of evaluation based on expert knowledge, URL patterns, site validation and the previously used USCAP method. These evaluations indicate strongly that the clusters found by NU-

ANCE do in fact correspond to actual authorship clusters. In particular, there is a strong relationship between expert knowledge of phishing campaigns and distinct authorship classes. Table I summarises the results, showing consistently high scores for each evaluation metric used.

These results also serve to help investigators in their efforts to properly understand these attacks. In this application, the clusters were derived automatically and then analysed to investigate their correlation to known information about the phishing landscape. These discoveries have already led to benefits in monitoring, investigating and protecting against future attacks for anti-phishing responders in industry. Knowing the creator of a phishing attack has allowed responders to target their response to the specific attack before manual investigation of the attacks takes place.

The analysis, using the NUANCE methodology, could be applied to other cybercrimes involving written documents such as fraud detection and plagiarism investigation. In particular, there is a strong need for higher levels of inference to be taken from profiling methodologies such as NUANCE. NUANCE clusters documents by authorship, but does not infer any extra information about them. Ethnographic investigative methods such as that employed by [22] may have a high level of synergy with NUANCE; NUANCE can investigate a large number of attacks while ethnographic investigations can create higher levels of inference, profiling those attacks and the attackers behind them. Combining these types of approaches could lead to in-depth focused investigations of cybercrime.

ACKNOWLEDGEMENT

This research was conducted at the Internet Commerce Security Laboratory and was funded by the State Government of Victoria, IBM, Westpac, the Australian Federal Police and the University of Ballarat. More information can be found at http://www.icsl.com.au

REFERENCES

[1] . C. H. Abbasi, Ahmed, "Applying authorship analysis to extremist-group web forum messages," *IEEE Intelligent Systems*, vol. 20, no. 5, pp. 67–75, 2005.

[2] A. Abbasi and H. Chen, "Writeprints: A stylometric approach to identity-level identification and similarity detection in cyberspace," *ACM Transactions on Information Systems*, vol. 26, no. 6, pp. 7:1–7:29, 2008.

[3] Anti-Phishing Working Group, "Phishing Activity Trends Report 2nd Quarter 2010," *APWG Industry Advisory*, 2010.

[4] P. Bright. (2011) Hackers spear-phish, infiltrate french ministry of finances. Retrieved July 4th, 2011. [Online]. Available: http://arstechnica.com/security/news/2011/03/hackers-spear-phish-infiltrate-french-ministry-of-finances.ars

[5] R. Charette. (2011) Canadian government restricts web access due to phishing attacks. Retrieved July 4th, 2011. [Online]. Available: http://spectrum.ieee.org/riskfactor/telecom/internet/phishing-attacks-makes-canadian-government-restrict-web-access

[6] H. Elghazel and K. Benabdeslem, "Towards b-coloring of som," in *Proceedings of the 6th International Conference on Machine Learning and Data Mining in Pattern Recognition*, ser. MLDM '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 322–336.

[7] H. Escalante, M. Montes-y Gómez, T. Solorio, I. Batyrshin, and G. Sidorov, "A Weighted Profile Intersection Measure for Profile-Based Authorship Attribution," in *10th Mexican International Conference on Artificial Intelligence, MICAI 2011*, ser. Lecture Notes in Computer Science, vol. 7094. Puebla, Mexico: Springer Berlin Heidelberg, 2011, pp. 232–243.

[8] G. Frantzeskou, S. Gritzalis, and S. G. Macdonell, "Source code authorship analysis for supporting the cybercrime investigation process," in *Proceedings of the first International Conference on e-business and Telecommunications Networks (ICETE04), Vol*, 2004, pp. 85–92.

[9] G. Frantzeskou, E. Stamatatos, S. Gritzalis, and C. E. Chaski, "Identifying authorship by byte-level n-grams: The source code author profile (SCAP) method," *Int. Journal of Digital Evidence*, vol. 6, 2007.

[10] G. Frantzeskou, E. Stamatatos, S. Gritzalis, and S. Katsikas, "Source code author identification based on n-gram author profiles," *Artificial Intelligence Applications and Innovations*, pp. 508–515, 2006.

[11] ——, "Source Code Author Identification Based on N-gram Author Profiles," *Artificial Intelligence Applications and Innovations*, vol. Volume 204/2006, pp. 508–515, 2006.

[12] A. Gray, P. Sallis, and S. MacDonell, "Identified (integrated dictionary-based extraction of non-language-dependent token information for forensic identification, examination, and discrimination): a dictionary-based system for extracting source code metrics for software forensics," in *Software Engineering: Education & Practice, 1998. Proceedings. 1998 International Conference*, Jan. 1998, pp. 252 –259.

[13] F. Iqbal, H. Binsalleeh, B. C. M. Fung, and M. Debbabi, "Mining writeprints from anonymous e-mails for forensic investigation," *Digital Investigation*, vol. 7, no. 1-2, pp. 56–64, 2010.

[14] P. Juola, *Authorship attribution*. Now Publishing, 2008.

[15] V. Kešelj, F. Peng, N. Cercone, and C. Thomas, "N-gram-based author profiles for authorship attribution," in *Proceedings of the Pacific Association for Computational Linguistics*, 2003.

[16] R. Kilgour, A. Gray, P. Sallis, and S. MacDonell, "A fuzzy logic approach to computer software source code authorship analysis," *International Conference on Neural Information Processing and Intelligent Information Systems*, pp. 865–868, 1997.

[17] I. Krsul and E. Spafford, "Authorship analysis: Identifying the author of a program," *Computers & Security*, vol. 16, pp. 233–257, 1997.

[18] R. Layton, P. Watters, and R. Dazeley, "Automated Unsupervised Authorship Analysis Using Evidence Accumulation Clustering," *submitted*, 2011.

[19] ——, "Automatically determining phishing campaigns using the uscap methodology," in *eCrime Researchers Summit (eCrime), 2010*. IEEE, 2011, pp. 1–8.

[20] ——, "Recentred Local Profiles for Authorship Attribution," *Journal of Natural Language Engineering*, 2011, available on CJO 2011.

[21] J. Ma, G. Teng, Y. Zhang, Y. Li, and Y. Li, "A Cybercrime Forensic Method for Chinese Web Information Authorship Analysis," *Intelligence and Security Informatics*, pp. 14–24, 2009.

[22] S. McCombie and J. Pieprzyk, "Winning the phishing war: a strategy for Australia," in *2010 Second Cybercrime and Trustworthy Computing Workshop*. IEEE, 2010, pp. 79–86.

[23] M. McNally, G. Newman, and C. Graham, "Perspectives on identity theft," Tech. Rep., 2008.

[24] S. Pillay and T. Solorio, "Authorship attribution of web forum posts," in *eCrime Researchers Summit (eCrime), 2010*. IEEE, 2011, pp. 1–7.

[25] A. Stabek, P. Watters, and R. Layton, "The seven scam types: Mapping the terrain of cybercrime," in *Cybercrime and Trustworthy Computing Workshop (CTC), 2010 Second*, july 2010, pp. 41 –51.

[26] E. Stamatatos, "A survey of modern authorship attribution methods," *Journal of the American Society for Information Science and Technology*, vol. 60, 2009.

[27] K. Turville, J. Yearwood, and C. Miller, "Understanding victims of identity theft: Preliminary insights," *Cybercrime and Trustworthy Computing, Workshop*, vol. 0, pp. 60–68, 2010.

[28] T. Urvoy, E. Chauveau, P. Filoche, and T. Lavergne, "Tracking web spam with html style similarities," *ACM Transactions of the Web*, vol. 2, no. 1, pp. 1–28, 2008.

[29] T. Urvoy, T. Lavergne, and P. Filoche, "Tracking web spam with hidden style similarity," *AIRWeb 2006 Program*, vol. 29, p. 25, 2006.

[30] P. Watters, "Data loss in the british government: A bounty of credentials for organised crime," in *Ubiquitous, Autonomic and Trusted Computing, 2009. UIC-ATC'09. Symposia and Workshops on*. IEEE, 2009, pp. 531–536.

[31] R. Zheng, Y. Qin, Z. Huang, and H. Chen, "Authorship analysis in Cybercrime investigation," *Lecture Notes in Computer Science*, pp. 59–73, 2003.