# Score Fusion Based Authorship Attribution of Ancient Arabic Texts

**Halim Sayoud, Siham Ouamour**
USTHB University
*halim.sayoud@uni.de, siham.ouamour@uni.de*

**Abstract**

In this paper, we investigate the authorship of several short historical texts that are written by ten ancient Arabic travelers: this Arabic dataset, which was collected by the authors in 2011, and called AAAT (Authorship attribution of Ancient Arabic Texts) corpus, is considered as a reference dataset in Arabic.

Several experiments of authorship attribution are conducted by using different features namely: characters, character n-grams, and lexical features such as words, word n-grams, and rare words. On the other hand, different classifiers are employed, such as: statistical distances, Multi Layer Perceptron (MLP), Support Vector Machines (SVM) and Linear Regression (LR).

In this investigation, a new fusion technique is proposed to enhance the overall performances of the classifiers: it is called Score Based Fusion (SBF).

Results show good attribution performances with an optimal score between 80% and 90% of good authorship attribution. The proposed fusion technique raised this score to 100% of good authorship attribution. Moreover, this comparative survey has revealed interesting results concerning the Arabic language and more particularly with short texts.

## Introduction

Authorship Attribution (*AA*) is a research field of stylometry, which consists in identifying the author(s) of a piece of text by using some techniques of text mining and statistics. The longer is the text; the better is the identification accuracy.

In general, individuals have distinctive ways of speaking and writing (Sayoud, 2012), and there exists a long history of linguistic and stylistic investigation into authorship attribution.

Although several works are reported for the English (Küppers , 2012) (Juola , 2006) (Holmes, 1994) (Burrows , 1987) and Greek (Tambouratzis , 2003) (Tambouratzis , 2004) (Tambouratzis , 2004-bis) (Tambouratzis , 2005) (Tambouratzis , 2007) languages, the authors have not found a lot of serious research works made with Arabic texts. Most of authorship attribution researchers used lexical features to represent the author style. However other works used common words (*articles, prepositions, pronouns, etc.*) to discriminate between authors (Argamon , 2005) (Burrows , 1987). In other works, various sets of words have been used for English, we can quote the works of Abbasi and Chen in 2005 (Abbasi , 2005) who used a set of 150 function words; Argamon in 2003 (Argamon , 2003) who used a set of 303 words; Zhao and Zobel in 2005 (Zhao , 2005) who used a set of 365 function words; and Koppel and Schler in 2003 (Koppel , 2003) who proposed 480 function words. Similarly, in the works of Argamon in 2007 (Argamon , 2007), another set of 675 words was proposed. Koppel in 2007 (Koppel , 2007) used the 250 most frequent words, while Stamatatos in 2006 (Stamatatos , 2006) extracted the 1000 most frequent words. On a larger scale, Madigan (Madigan , 2005) used all the words that appear at least twice in the corpus.

In (Peng , 2004, (Sanderson , 2006) and (Coyotl-Morales , 2006), *word n-grams* have been proposed as textual features. Differently, Koppel and Schler in (Koppel , 2003) proposed various writing error measures to detect the idiosyncrasies of an author's style. So, a set of spelling and formatting errors has been defined. This type of information could be used for any natural language and give good results to quantify the writing style (Grieve , 2007).

Concerning the character *n*-grams, the application of this approach to authorship attribution has shown an interesting success. Character bigrams and trigrams have been used in the works of Kjell (Kjell , 1994) to discriminate the Federalist Papers. Forsyth and Holmes (Forsyth , 1996) found that bigrams and character *n*-grams of variable-length performed better than lexical features in several text classification tasks including authorship attribution. They have been also used in the works of Peng (Peng , 2003), Keselj (Keselj , 2003) and Stamatatos (Stamatatos , 2006-bis) giving interesting results. Moreover, in 2004, Juola (Juola , 2004) proposed one of the best performing algorithms based on a character *n*-gram representation for the task of authorship attribution. In another work, Grieve (Grieve , 2007) showed by a comparison of different lexical and character features that character *n*-grams were the most effective features.

However, it is not only the feature that is important; in fact, the choice of a suitable classifier is important too. So, in 2010, Jockers and Witten (Dasarathy , 1994) tested five different classifiers on the problem of the Federalist Papers. They reported that each of the five methods performed well, but nearest shrunken centroids and regularized discriminant analysis got the best performances.

Another important problem is to find the minimal amount of data in order to get reliable authorship attribution results. Hence, in 2013 Eder tried to respond to that question (Eder , 2010) by looking for the minimal size of text samples for authorship attribution that would provide stable results for different types of text genres and languages. He reported that there is no conclusive answer to this question, but it seems that for corpora of modern novels, irrespective of the language tested, the minimal sample size is some 5,000 words (tokens), while Latin prose required only 2,500 words, and Ancient Greek prose just a little more to display their optimal performance.

Concerning the Arabic language, there are not a lot of serious works that are reported, but some of the most recent research ones are perhaps the works of Sayoud in 2012 (Sayoud, 2012) and Shaker (Shaker , 2012). Sayoud conducted an investigation on authorship discrimination between two old Arabic religious books: the Quran (*The holy words and statements of God in the Islamic religion*) and Hadith (*statements said by the prophet Muhammad*) (Sayoud, 2012) by employing several classifiers and several types of features. Shaker investigated the Authorship Attribution problem in the Arabic language (Shaker , 2012), and compared it to Authorship Attribution in English Language using Function Words. A hybrid algorithm (*Evolutionary Algorithm /Linear Discriminant Analysis*) was used to reach the best least number of features to classify the problem. Finally, in this investigation, we propose some techniques of fusion to enhance the performances of authorship attribution in standard Arabic. For that reason, a special Arabic corpus has been built by the authors of this paper in order to assess several features and classifiers that are usually employed in stylometry, in a comparative way.

## Description of the text dataset

The text dataset, called AAAT corpus (*i.e. Authorship attribution of Ancient Arabic Texts*), is built by the authors of this paper for a purpose of authorship attribution. It contains 10 groups of old Arabic texts that are extracted from 10 different Arabic books and which belong respectively to 10 different ancient authors. Each group contains different texts that are written by the same author, which means that each group belongs to only one ancient author. This set of texts (30 different documents) has been collected in 2011 from "Alwaraq library", and is now freely downloadable from our website. Moreover this corpus represents a sort of reference dataset for authorship attribution in Arabic, which has been used by some researchers working in this field. In fact, several research works have been recently conducted on AAAT for a purpose of evaluation.

## Brief description of the different books

The different historical texts are summarized in table 1.

*Table 1. Books specifications.*

| Author name | Date AD | Book title | Topic |
|---|---|---|---|
| Ibn Batuta | 1325-1352 | Travels of Ibn Batuta | Travels |
| ابن بطوطة | | رحلة ابن بطوطة | |
| Ibn Jubayr | 1182-1185 | Travels of Ibn Jubayr | Travels |
| ابن جبير | | رحلة ابن جبير | |
| Nasser Khasru | 1045 | Book of the Travels | Travels |
| ناصر خسرو | | سفر نامه | |
| Ibn Fathlan | 921 | Travels of Ibn Fathlan | Travels |
| ابن فضلان | | رحلة ابن فضلان | |
| Ibn Al Mujawer | 1233 | History of the Mustabsir | Travels |
| ابن المجاور | | تاريخ المستبصر | |
| Al Yussee | 1684 | Conferences in language and literature | Travels |
| اليوسي | | المحاضرات في اللغة و الأدب | |
| Lessan Addin | 1684 | Khatrat Al Tife during the tra vel of the winter and summer | Travels |
| لسان الدين بن الخطيب | | خطرة الطيف في رحلة الشتاء و الصيف | |
| Al Alussi | 1852 | Strangeness of travels | Travels |
| الآلوسي | | غرائب الاغتراب | |
| Al Hamawi | 1542-1608 | Hady Alathaan Annajdia to the Egypt houses | Travels |
| محب الدين الحموي | | حادي الأظعان النجدية إلى الديار المصرية | |
| Al Balwi | Before 1364 | Taj Almafraq Fi Tahlyet oriental scientist | Travels |
| البلوي | | تاج المفرق في تحلية علماء المشرق | |

The texts are quite short: the average text length is about 550 words and some texts have less than 300 words. This situation involves severe experimental conditions, since it has been shown in previous research works (Eder , 2010) (Signoriello , 2005) that the minimum number of words per text should be at least 2500 words to get good attribution performances. We have chosen to use short texts in order to evaluate the different classifiers with small documents in Arabic. In fact, when short texts are used, the AA performances decrease and it becomes difficult to make an efficient identification (*i.e. severe conditions*).

## Description of the classification process

The choice of the optimal classifier is crucial before any application of pattern recognition, that is why we have decided to use several types of classifiers and evaluate them in the same experimental conditions. These experiments are conducted by employing the following distances and classifiers: Manhattan distance, Cosine Distance, Stamatatos distance, Canberra distance, Multi Layer Perceptron (MLP), Support Vector Machines (SVM), and Linear regression. Moreover, several features were employed namely: characters, character n-grams, words, word n-grams and rare words. This diversity in the feature type is quite interesting for discovering the most reliable characteristics for the Arabic language. At the end, a Vote Based Fusion (*VBF*) has been proposed to enhance the overall classification performances. In the following subsections, some brief definitions of the different classifiers are given and commented. The general classification process is divided into two methods: Training Model based Classification and Nearest Neighbor based Classification. In the first type, a training step is required to build the model or the centroid (*in case of similarity measures*); afterward, the testing step could be performed by using the resulting model. In the second type, the training is not required, since a simple similarity distance is computed between the unknown document and each referential text: the smallest distance gives an indication on the most probable class. Furthermore two types of measures are employed: a simple distance and a centroid based distance. The first type is known to be inaccurate, while the second one (*i.e. centroid*) is more accurate and robust against noises.

The first classification type includes the following classifiers: Centroid based Similarity measures, Multi-Layer Perceptron, Support Vector Machines and Linear Regression; whiles the second classification type includes only the nearest neighbor similarity measures (*table 2)*.
After every identification test, a score of good authorship attribution is computed in order to get estimation on the overall classification performances.

## Experiments of authorship attribution

In this section, we present the different experiments of authorship attribution, which are conducted on the historical Arabic texts.

### Experimental results

Several features are tested such as: characters, character bigrams, character trigrams, character tetragrams, words, word bigrams, word trigrams, word tetragrams and rare words.

On the other hand, different types of classifiers (MLP, SVM and LR) and distances are employed to ensure the automatic authorship classification.

Note that the Authorship Attribution Score (*AAS*) is calculated, in our investigation, by using the Rand Accuracy formula, as follows:

$$AAS\ score = Rand\ Accuracy = \frac{number\ of\ texts\ that\ are\ well\ attributed}{total\ number\ of\ texts} \quad (1)$$

*Table 2. Authorship attribution accuracy for the different classifiers (* means 600 most frequent features only).*

| AAS Accuracy | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Feature / Classifier | Characters | Charac. bigram | Charac. trigram | Char. tetragram | Word | Word bigram | Word trigram | Word tetragram | Rare word |
| Canberra distance | 0.3 | 0.6 | 0.6 | 0.6 | 0.2 | 0.1 | 0.1 | 0.1 | 0.2 |
| Cosine distance | 0.5 | 0.6 | 0.6 | 0.6 | 0.6 | 0.2 | 0.1 | 0.2 | 0.6 |
| Manhattan distance | 0.4 | 0.5 | 0.7 | 0.5 | 0.6 | 0.3 | 0 | 0.1 | 0.6 |
| Stamatatos distance | 0.4 | 0.5 | 0.6 | 0.6 | 0.2 | 0.1 | 0.1 | 0.1 | 0.2 |
| Canberra Centroid dist | 0.5 | 0.6 | 0.6 | 0.6 | 0.1 | 0.1 | 0.2 | 0.1 | 0.1 |
| Cosine Centroid dist | 0.3 | 0.4 | 0.4 | 0.6 | 0.5 | 0.2 | 0.2 | 0.1 | 0.6 |
| Manhattan Centroid dist | 0.3 | 0.4 | 0.7 | 0.9 | 0.5 | 0.4 | 0.2 | 0 | 0.7 |
| Stamatatos Centroid dist | 0.1 | 0.4 | 0.6 | 0.6 | 0.1 | 0.1 | 0.2 | 0.1 | 0.1 |
| MLP | 0.6 | 0.8 | 0.8* | 0.6* | 0.6* | 0.4* | 0.2* | 0.2* | 0.7* |
| SVM | 0.6 | 0.8 | 0.8* | 0.7* | 0.7* | 0.2* | 0.1* | 0.1* | 0.8* |
| Linear Regression | 0.5 | 0.7 | 0.6* | 0.6* | 0.6* | 0.5* | 0.2* | 0.1* | 0.6* |

Table 3 presents the AAS scores that are got by the different classifiers and by using the nine different features. As we can observe, the authorship attribution performances, which seem to be diversified and very different, depend closely to the type of feature and classifier that are employed. So, for instance, a score of good attribution of 90% (i.e. 0.9) has been obtained by using Manhattan centroid distance and character tetragrams. This score represents the best score that is obtained during all the experiments of this investigation.

## Comparative performances

In the overall, concerning the classifiers, we remark that the best scores for the distances (*Canberra, Cosine, Stamatatos and Manhattan*) are obtained with tetragram feature, whereas for the machine learning types (*MLP, SVM and linear regression*) the best scores are given using character bigrams as feature. We also notice that by using Manhattan distance, the score increases with the length of character n-grams.

We remark that Manhattan centroid distance seems to be very accurate, with a score of 90% (i.e. 0.9), followed by the learning machines (MLP and SVM), with a score of 80% (i.e. 0.8), after that, we retrieve Manhattan nearest neighbor distance and the linear regression classifier, which provide a score of 70% (i.e. 0.7). Finally, the remaining distances: Canberra, Cosine and Stamatatos distances, give the worst performances with a score of only 60% (i.e. 0.6).

In another figure (*not presented in this paper*), we have presented the average authorship attribution performances for every feature. Those performances scores are obtained by calculating the mean of all the scores of a specific feature. From that figure, we were able to deduce that the best feature is character trigrams, followed by character tetragrams, character bigrams and rare words. The performances of authorship attribution continue to decrease respectively by using words, characters, word bigrams, word trigrams and finally, word tetragrams, which represents the worst features in our experiments. In overall, we notice two important points: On one hand, the attribution score increases with the character n-gram size (ie. the size n) and decreases with the word n-gram size. On the other hand, character n-grams seem to be more accurate than word n-grams (and rare words).

Once again, we can observe that character n-grams are better than word n-grams and we can also notice that the system presents a failure when using word n-grams. These last ones seem to be not suitable for the authorship attribution of short texts: this result is logical because short texts do not contain enough words or enough word n-grams either to make a fair statistical representation of the features.

Concerning the best score given by each used feature. As we can see, a score of 90% has been given by character tetragrams, followed by a score of 80% obtained by using character bigrams, character trigrams and rare words, thereafter, a score of 70% is given by words, a score of 60% is given by characters, a score of 50% is given by using word bigrams and finally, a score of 20% is obtained by using word trigrams and tetragrams.

## Score Based Fusion using the COST parameter

In order to enhance the attribution performance, we thought to use a special combination in order to get a lower discrimination error: this combination is called Fusion.

The fusion in the broad sense can be performed at different hierarchical levels or processing stages. A very commonly encountered taxonomy of data fusion is given by the following three-stage hierarchy (Dasarathy , 1994) (Verlinde , 1999) (Jain , 2004): Feature level, Score (matching) and Decision. In our case, we are interested in the second fusion type.

In fact, we noticed that, usually, when ancient Arabic authors wrote a series of poems, they made a termination similarity between the neighboring sentences of the text (like poems), such as a same final syllable or letter. To evaluate that termination similarity, a new parameter estimating the degree of text chain (*in a text of several sentences*) has been proposed: the COST parameter (Sayoud, 2012).

For instance, the COST parameter for sentence “j” is computed by adding all the occurrence marks (*values*) between sentence “j” and its neighboring sentences (*sentence “j-1” and sentence “j+1”*). In our case, the occurrence marks concern only the two last letters of the sentence (Sayoud, 2012).

The COST parameter, in this case, can give some information on the structure of the text (ending structure). In this investigation, it has been employed to see if the documents respect certain regularities in the text structure or not, and if so, to assess the corresponding regularity ratio.

The studied documents do not correspond to poems but some of the old Arabic authors were used to employing a certain ending similarity, which make them quite differentiable in terms of stylistic structure.

In our experiments, we fused the Character-4grams features with the COST parameter by using the Manhattan centroid distance as classifier.

In this second investigation, we have proposed a new fusion technique based on the combination of the two different scores (distances). We called it SBF or Score Based Fusion.

$$SBF_{Fusion(i,k)} = Char4gram_{distance(i,k)} + \alpha . COST_{distance(i,k)} \quad (2)$$

where i represents the $i^{th}$ testing sample and k represents the $k^{th}$ reference sample. The constant $\alpha$ is tuned experimentally and should be smaller than 1.

That is, as we can see in the following table (table 3), the fusion permits us to correct some confusion errors during the identification process, leading to a correct identification of the text documents.

*Table 3: Confusion matrix* using the Score Fusion.*

| Score Fusion (Char-4grams & COST) inter-distances | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $Test_1$ | $Test_2$ | $Test_3$ | $Test_4$ | $Test_5$ | $Test_6$ | $Test_7$ | $Test_8$ | $Test_9$ | $Test_{10}$ |
| $Ref_1$ | 0.62 | 0.67 | 0.72 | 0.87 | 0.75 | 0.75 | 0.83 | 0.77 | 0.76 | 0.88 |
| $Ref_2$ | 0.74 | 0.53 | 0.85 | 0.75 | 0.72 | 0.70 | 0.83 | 0.96 | 0.89 | 0.91 |
| $Ref_3$ | 0.84 | 0.64 | 0.71 | 0.73 | 0.66 | 0.72 | 0.90 | 1.04 | 0.95 | 0.98 |
| $Ref_4$ | 0.85 | 0.66 | 0.84 | 0.61 | 0.67 | 0.75 | 0.91 | 1.05 | 0.94 | 1.01 |
| $Ref_5$ | 0.86 | 0.68 | 0.91 | 0.77 | 0.61 | 0.72 | 0.98 | 0.95 | 0.96 | 0.94 |
| $Ref_6$ | 0.76 | 0.65 | 0.83 | 0.80 | 0.66 | 0.62 | 0.91 | 0.91 | 0.88 | 0.94 |
| $Ref_7$ | 1.15 | 1.19 | 1.25 | 1.25 | 1.31 | 1.24 | 0.50 | 0.90 | 0.83 | 0.76 |
| $Ref_8$ | 1.04 | 1.06 | 1.13 | 1.20 | 1.14 | 1.07 | 0.71 | 0.72 | 0.76 | 0.76 |
| $Ref_9$ | 0.78 | 0.80 | 0.96 | 0.97 | 0.93 | 0.91 | 0.77 | 0.71 | 0.64 | 0.87 |
| $Ref_{10}$ | 1.03 | 1.00 | 1.10 | 1.18 | 1.15 | 1.16 | 0.62 | 0.75 | 0.67 | 0.76 |

Concerning the overall authorship attribution experiments using the proposed Score Fusion between the COST parameter and Char-4grams, we got an authorship attribution accuracy of 100% on the AAAT corpus.

We recall that the best previous scores without fusion were between 80% and 90%, which shows that the proposed Score Fusion has further enhanced the classification performances.

## Discussion

In this research work a new authorship attribution investigation has been conducted on an old Arabic set of documents that were written by ten ancient Arabic travelers. Several features have been experimented for the Arabic language and particularly for short texts. These two particularities (*-ancient Arabic language and -small text size*) represent another originality for this research work, since only few works were reported for such cases. Hence, eleven different classifiers have been used for the attribution task, by using nine different features as described in section 4. Moreover a new fusion techniques (i.e. SBF) was proposed and a new COST parameter was employed.

The different experiments conducted during this investigation have led to the following important points:

- Character n-grams appear to be quite interesting for the task of authorship attribution, since the score of good attribution reaches 90% with small texts.
- Particularly for Manhattan distance (*centroid technique*), we strangely notice that the attribution score increases with the size of the n-grams length (*i.e. accuracy of 0.3, 0.4, 0.7 and 0.9 for character unigrams, bigrams, trigrams and tetragrams respectively*). Hence, it is clear that the amount of authorship information is closely linked to the n-gram size: in other terms, the greater is this size, the higher will be the linguistic information (*theoretically speaking*).
- Concerning the average performance of the classifiers, we have noticed that the best classifier is the SVM. This result was expected as mentioned previously.
- In a purpose of comparison between the features, we can notice that character-based features are better than word-based features (*for almost all classifiers*). Although, in other works (*not cited in the present paper*) we got quite interesting results with words and word n-grams for Arabic language, we notice that in this investigation, we do not obtain such performances: probably, because we are using very short texts and then the amount of words or word n-grams is not sufficient to help training the classifier. Moreover, a failure has particularly been noted for the word feature, which presents an attribution score of about 35% only.
- Concerning the proposed SBF fusion technique, which employs the association of the classical features scores with the so-called COST parameter, measuring the text endings similarity, we notice that it provided the best performances at all, namely: an accuracy of 100%. However, the use of the COST parameter is not so easy since in most cases it was computed manually, which makes this last result quite difficult to reach.
- Concerning the overall performances of our identification approaches, although the text size was quite small (*i.e. the average text length is about 550 words*), the score of good AA remains quite interesting.

Finally, this investigation on the ancient Arabic language, which is considered as the real standard Arabic, shows a real motivation and interest for this type of language. It also shows that the fusion approach can really improve the AA results if it is judiciously performed. However, an important question would be: what are the actual performances for dialectical Arabic language (*eg. Egyptian or Saudi dialects*) and what kind of problems could we meet in practice (*ie. modern spoken Arabic*)?

## References


Sayoud H., "*Author Discrimination between the Holy Quran and Prophet's Statements*", Literary and Linguistic Computing, Volume 27 Issue 4, pp.427-444, 2012.

Küppers R., and Conrad S., "*A Set-Based Approach to Plagiarism Detection*", PAN 2012 Lab Uncovering Plagiarism, Authorship, and Social Software Misuse held in conjunction with the CLEF 2012, 17-20 September, Rome, Italy, 2012.

Juola P., "*Authorship Attribution, a survey and technical monograph on authorship attribution, the process of inferring the author or author's characteristics from the text of a document*", published through NOW Publishers, 2006.

Holmes D.I., "*Authorship attribution*", Computers and the Humanities, 28, pp. 87–106, 1994.

Burrows J.F., "*Word patterns and story shapes: The statistical analysis of narrative style*", Literary and Linguistic Computing, 2, pp. 61-70, 1987

Tambouratzis G., Hairetakis G., Markantonatou S., & Carayannis G. (2003). *Applying the SOM Model to Text Classification According to Register and Stylistic Content*. International Journal of Neural Systems, 13(1), 1-11.

Tambouratzis G., Markantonatou S., Hairetakis N., Vassiliou M., Carayannis G. & Tambouratzis D. (2004). *Discriminating the Registers and Styles in the Modern Greek Language-Part 1: Diglossia in Stylistic Analysis*. Liter. & Ling. Comp. 19(2), 197-220.

Tambouratzis G., Markantonatou S., Hairetakis N., Vassiliou M., Carayannis G. & Tambouratzis D. (2004). *Discriminating the Registers and Styles in the Modern Greek Language-Part 2: Extending the Feature Vector to Optimize Author Discrimination*. Literary and Linguistic Computing, 19(2), 221-242.

Tambouratzis G. (2005). *Assessing the Effectiveness of Feature Groups in Author Recognition Tasks with the SOM Model*. IEEE Transactions on Systems, Man & Cybernetics – Part C: Applications and Reviews, 36(2), 249-259.

Tambouratzis G. & Vassiliou M. (2007). *Employing Thematic Variables for Enhancing Classification Accuracy Within Author Discrimination Experiments*. Literary and Linguistic Computing, 22(2), 207-224.

Argamon S., and Levitan S., "*Measuring the usefulness of function words for authorship attribution*", In Proceedings of the Joint Conference of the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing, 2005.

Abbasi A., and Chen H., "*Applying authorship analysis to extremist-group web forum messages*", IEEE Intelligent Systems, Volume 20 Issue 5, pp. 67-75, 2005.

Argamon S., Saric M., and Stein S., "*Style mining of electronic messages for multiple authorship discrimination: First results*", In Proceedings of the 9th ACM SIGKDD, pp. 475-480, 2003.

Zhao Y., and Zobel J., "*Effective and scalable authorship attribution using function words*", In Proceedings of the 2nd Asia Information Retrieval Symposium, 2005

Koppel M., Schler J., "*Exploiting stylistic idiosyncrasies for authorship attribution*", Workshop on Computational Approaches to Style Analysis and Synthesis, pp. 69-72, 2003.

Argamon S., Whitelaw C., Chase P., Hota S.R., Garg N., and Levitan S., "*Stylistic text classification using functional lexical features*", Journal of the American Society for Information Science and Technology, Volume 58 Issue 6, pp. 802-822, 2007.

Koppel M., Schler J., and Bonchek-Dokow E., "*Measuring differentiability: Unmasking pseudonymous authors*", Journal of Machine Learning Research, 8, pp. 1261-1276, 2007.

Stamatatos E., "*Authorship attribution based on feature set subspacing ensembles*", International Journal on Artificial Intelligence Tools, Volume 15 Issue 5, pp. 823-838, 2006.

Madigan D., Genkin A., Lewis D., Argamon S., Fradkin D., and Ye L., "*Author identification on the large scale*", In Proceedings of CSNA-05, 2005.

Peng F., Shuurmans D., and Wang S., "*Augmenting naive Bayes classifiers with statistical language models*", Information Retrieval Journal, Volume 7 Issue 1, pp. 317-345, 2004.

Sanderson C., and Guenter S., "*Short text authorship attribution via sequence kernels, Markov chains and author unmasking: An investigation*", In Proceedings of the International Conference on Empirical Methods in Natural Language Eng., pp. 482-491, 2006.

Coyotl-Morales R.M., Villaseñor-Pineda L., Montes-y-Gómez M., and Rosso P., "*Authorship attribution using word sequences*", In Proceedings of the 11th Iberoamerican Congress on Pattern Recognition, pp. 844-85, Springer, 2006.

Grieve J., "*Quantitative authorship attribution: An evaluation of techniques*", Literary and Linguistic Computing, Volume 22 Issue 3, pp. 251-270, 2007.

Kjell B., "*Discrimination of authorship using visualization*", Information Proc. and Management, Vol. 30:1, pp. 141-150, 1994.

Forsyth R., and Holmes D., "*Feature-finding for text classification*", Liter. & Ling. Comput., 11:4, pp. 163-174, 1996.

Peng F., Shuurmans D., Keselj V., and Wang S., "*Language independent authorship attribution using character level language models*", In Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics, pp. 267-274, 2003.

Keselj V., Peng F., Cercone N., and Thomas C., "*N-gram-based author profiles for authorship attribution*", In Proceedings of the Pacific Assoc. for Computational Linguistics, pp. 255-264, 2003.

Stamatatos E., "*Ensemble-based author identification using character n-grams*", In Proceedings of the 3rd International Workshop on Text-based Information Retrieval, pp. 41-46, 2006.

Juola P., "*Ad-hoc authorship attribution competition*", In Proceedings of the Joint Conference of the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing, pp. 175-176, 2004.

Eder M., "*Does size matter? : autorship attribution, short samples, big problem*" In Digital humanities 2010 conference, pp. 132-135, London, 2010.

Shaker K., "*Investigating features and techniques for Arabic authoriship attribution*", PhD thesis Heriot-Watt Univ., 2012.

Signoriello D. J., Jain S., Berryman M. J., Abbott D., "*Advanced text authorship detection methods and their application to biblical texts*" SPIE'2005, Volume 6039, Spie, pp. 163–175, 2005.

Dasarathy B.V., "*Decision Fusion*", In proceedings of IEEE Computer Society Press, Los Alamitos, CA, 1994.

Verlinde P., "*Contribution à la vérification multimodale d'identité en utilisant la fusion de décisions*" PhD thesis, Ecole Nationale Supérieure des Télécom., Paris, France, September 17th, 1999.

Jain A. K., Ross A., and Prabhakar S., "*An Introduction to Biometric Recognition*", Journal of IEEE Transactions on Circuits and Systems for Video Technology, Volume 14 Issue 1, pp. 4-20, 2004.