Identifying Content Originator in Social Networks

Praboda Rajapaksha*‡, Reza Farahbakhsh*, Noël Crespi*
*Institut Mines-Télécom, Télécom SudParis, CNRS UMR 5157 SAMOVAR, France
{praboda.rajapaksha, reza.farahbakhsh, noel.crespi}@telecom-sudparis.eu
‡Uva Wellassa University of Sri Lanka, praboda@uwu.ac.lk

***Abstract*—Content originality detection is an interesting research topic in large-scale scenarios especially in social media where anyone has the ability to produce and disseminate content in different forms through their profiles and activities. What is missing in these communication sites is to be able to identify original content producers as some users spread information copied from other users without indicating its original producer, or where they found it. This paper provides a conceptualized approach for content originality detection and illustrates the efficiency of the model when applying it to a Twitter dataset. This approach amalgamates user's linguistic features and their online circadian behaviors to identify accurately the content originator for a given text. The proposed approach is evaluated using an F1-measure and the results indicate an accuracy of 95% or higher for all test scenarios. While achieving high accuracy in the test results, our approach, as a usecase, was applied in the context of news agencies popular worldwide to identify news producers and consumers by analyzing their Tweets. We investigated intra and inter news flows among several major news agencies considered in our dataset. Our results show that this proposed approach can distinguish *News Story Tellers* from *News Propagators* in the news agencies community as well as provide information that helps to understand the flow patterns between different news groups.**

***Index Terms*—Content propagation, originality, authorship, social media, Twitter.**

I. Introduction

In this era where social media is encouraging users to be active content producers instead of simple consumers, and where users have the ability to share almost everything from anybody, having the knowledge and a method with which to identify the main producer of a content is an important asset and a difficult challenge. Studies show the huge sharing activities of social media users [1] but a large portion of the content is simply copied/pasted from other accounts (can be called as *Cross Posts* [2]) or sources without referring to their original publishers.

Undoubtedly, plagiarism detection in Online Social Networks (OSNs) is important, especially when the content belong to popular users (e.g. celebrities, politicians etc.) or major news agencies. This study attempts to identify the content originator of textual content in social media and to detect information propagation patterns among users based on their linguistic features and temporal behaviors.

In particular, content originator detection is critical in the context of fraudulent news and social media hoaxes. Some fake news increases readers tension while providing dangerous irresponsible information. For instance, false news stories circulated on social media played a dubious role during 2016 US Presidential election campaigns [3]. Originality detection in OSNs based on unsupervised procedures are thus extremely important to identify the true identify of a user.

Authorship attribution is one key area that we can adapt to detect content originators. A large body of literature in author attribution has been proposed, utilizing users' writing styles [4] [5]. Towards that end, we implement a framework manipulating user's writing patterns using the SCAP method [6] (since earlier research has proven that SCAP is useful when identifying authorship [7], [8]) and users' online circadian behaviors. We then evaluate the framework using different test cases, with the goal of analyzing (as explained in Section IV) two research questions: (1) How efficient is the SCAP algorithm when applied to OSNs data (since the length of the text is limited) and what parameters do we need to consider in order to increase the accuracy of the system? And (2) Is the circadian typology behavior of users in OSNs useful for detecting content originators? Finally, in Section V, for a better understanding of the functionality of the proposed approach, we consider a use case with the top news agencies in the world, and utilized the proposed method to identify the flow of information (posts on specific news topics) between these agencies that publish on Twitter.

The main contributions of this study are summarized in the two following items: (i) the proposal of an advanced framework to identify cross-posts and the main originator of a post in a large dataset and subsequently to later on detect who is the original content publisher; (ii) the evaluation of the proposed method on a dataset including 145K tweets belonging to 8 popular international news agencies.

II. Literature Review

Author attribution in social media has become a major research area. Many recent studies on authorship attribution of short and noisy text in social media have used machine learning techniques, NLP and similarity based approaches such as topic identification, genre identification etc. However, similarity based approaches outperform other methods when considering a large number of authors, a limited text size, and a large training set [9]. Many social network authorship attribution studies have used different similarity based mechanisms such as word and character n-grams [10], Source Code Author Profile (SCAP) [7] [8], Latent Dirichlet Allocation (LDA) and Author Topic (AT) [11] [12] .

A number of Internet-scale author detection studies focused on using a user's stylometric features [4] in various disciplines

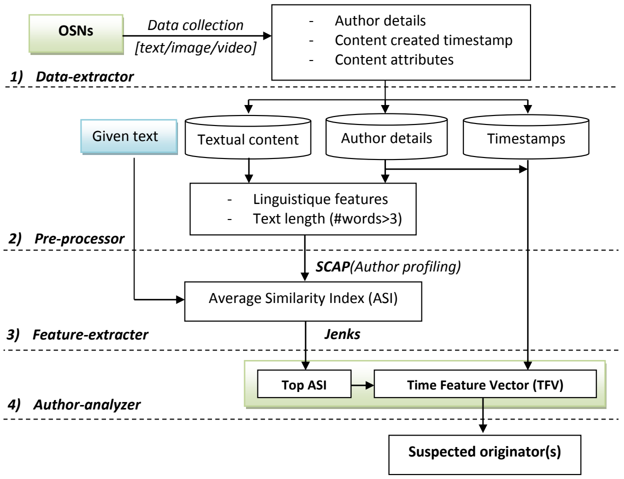

Fig. 1. Content Originality Detection Framework: ConOrigina.

such as, cyber-criminal detection [5], identifying the likability of tweets [10], and cross domain authorship attribution in social media [13]. Additionally, some other works are based on analyzing a user's writing style to detect fraud in digital forensics [14]. In addition, there are few patent-granted studies for the determination of content originality in the Web[1]. These mainly focused on electronic commerce and provided recommendations for an item based on the originality of similar items and identify the matching content objects.

In this study, we propose a novel framework to identify content originators in social media enhancing the SCAP approach by integrating circadian behavior of a user in online platforms.

III. Methodology and Proposed Solution

The methods described in Section II are successful author identification mechanisms for social media based on users' writing patterns. Moreover, a limited #studies have considered how temporal changes of user's writing style affects author attribution. Hosein et al. [8] identified that authors do change their writing styles at different time periods, and different authors change differently by using a time-aware feature sampling approach. Our originality detection mechanism proposed in this study aims to combine users' temporal behaviors and linguistic features to detect content originators in OSNs.

Compared to other web forums, OSNs contain more information on user-generated content (UGC), combined with a number of attributes such as timestamp, geo-location, content type (text/image/video), and user's profile. This work is designed to identify textual content originator in OSNs; the same idea can be applied to other content types as well. We propose a framework, namely *ConOrigina*, based on SCAP method using the four distinct phases depicted in Figure 1: (1) Data-crawler, (2) Pre-processor, (3) Feature-extractor and (4) Author-analyzer.

A. Data-crawler

A data-crawler extracts relevant information from OSNs to build a knowledge base. Our approach's knowledge base consists of each user's social structure and the documents containing each user's shared text. In this study, to evaluate our proposed method we have selected Twitter as an OSN source and therefore, we implemented a crawler using a Tweepy python library to access the entire Twitter RESTful API methods. The crawler's input is the user's Twitter name; the crawler invokes a Twitter API to extract tweets, timestamps, and associated profile details. Twitter API allows the retrieval of only 3,200 recent tweets, and so this present study is focused on a single API query per user.

Since this study is an attempt to detect content originality, the identified main attributes here are each user's writing profile and temporal behaviors in OSNs. We implement each users' writing profile based on her shared tweets. To discern these, we consider the user's tweets before and after a month has passed with respect to a timestamp of a given text. Apart from this step, we use the timestamps of all posts to identify users' temporal behaviors using time feature vectors (TFV). The TFV method is elaborated in detail in the author-analyzer phase of the framework. Twitter API allows timestamps to be retrieved formatted as UTC, hence TFV implementation is much easier and accurate in term of comparing posts with near timestamp from different geo-locations.

B. Pre-processor

In this phase, our dataset is classified into different groups including texts, author details (i.e, username) and timestamp of the posts. We can use a raw dataset directly without pre-processing, and with the pre-processed texts in the SCAP method. For instance, in the 1st scenario, a Twitter dataset may contain all the #tags, @username, and URLs. The pre-processor remove all these components from the tweets.

Since our framework is using byte-level n-grams, the text language is especially important and hence, we categorize and filter texts into different datasets based on the language. The analysis of this study is performed on tweets in English.

C. Feature-extractor

The feature-extractor receives as input a dataset classified in the pre-processor phase, and its output is a set of attributes (n-gram profiles) that are manipulated based on users' writing styles. We then, execute SCAP (Source Code Author Profiling) [6] method with these attributes. The SCAP method, one of the character-level n-gram approaches, was designed to identify the author of a computer program by profiling an author based on his commonly-used n-grams. In the SCAP approach, n-gram frequencies are considered as an author's profile. These author profiles are used to examine the similar writing styles of different users based on the intersection of their n-grams using a Jaccard index. The higher the overlap measurements, the more similar those user profiles are.

Despite the fact that the SCAP method is used to measure the overlap similarity of author profiles, applying it directly to analyze a short text with large textual content is not very efficient [13]. Therefore, we examine the outcome of the

[1] https://www.google.com/patents/US8707459 & .../patents/US8412718

Algorithm 1 Originility Detection Algorithm

```
procedure ORIGINALITY($AUT, AKT, N, X$)    ▷
N: #text used in 1 SCAP execution
X: Set of top similarity index authors
AUT: author unknown text
AKT: author known texts
    if #words in AUT & AKT $> 3$ then
        while $i = \#authors$ do
            for k=0,j select N text of $AKT_i$ do
                out_$AKT_k$ = SCAP(AUT,$AKT_i$)
                result.$append$(out_$AKT_k$)
            end for
            Similarity-Index.$append$($avg$(result))
        end while
    end if
    X = $Jenks$(Similarity-Index)
    for each user i in X do
        if TT in TFV(AKT) then
            originators.append(i)
        end if
    end for
    return $sort$(originators) based on timestamp    ▷ top user in the list is the originator
end procedure
```

Algorithm 2 Time Feature Vector (TFV)

```
function TFV($AKT$)    ▷ n: #posts in AKT
    for i=0, j=6, i+=6, j+=6 do
        if $i < T_{up} < j$ then
            $T_{up} \in F_{i,j}$
        ▷ $T_{up}$:timestamp of post-p of user-u and $F_{i,j}$ :set of posts belongs to time period $i - j$hrs
        end if
    end for
    $TFV =< \sum F_{0,6}/n, \sum F_{6,12}/n, \sum F_{12,18}/n, \sum F_{18,24}/n >$
    return $TFV$[time duration of max index]
end function
```

SCAP approach by manipulating n-gram author-sub-profiles, generated for different time periods and #texts.

In this proposed framework, initially, the SCAP method is executed for each user's document by reading a chunk (N lines) where each document contains posts crawled 1 month before the date of the given post and 1 month after. Next, the SCAP method is executed between the author-sub-profiles generated for each chunk (1 sub-profile per chunk and n in range(4,6)) and n-grams of the provided text to produce a similarity score. Finally, the *feature extractor returns an Average Similarity Index (ASI) per user* considering the similarity scores obtained for all author-sub-profiles versus the given text.

D. Author-analyzer

The author-analyzer phase is where most adequate author(s) for a given text are predominantly identified by utilizing the results from the feature-extractor phase.

The Jenks unsupervised classification algorithm is used to classify X number of users into the best cluster. Since Jenks dynamically chooses each cluster, we do not set any specific threshold to cluster X users with maximum ASI identified in the Feature-extractor phase. Furthermore, in this study, the value of the Goodness of Variance Fit (GVF) used in the Jenks algorithm is set to 0.8. Jenks algorithm returns one or more users that have the same writing styles with reference to the provided text. This study aims to assess the best matching author(s) for a given text based on the results obtained from the Jenks algorithm, as elaborated in algorithm 1.

According to the authors' best knowledge, very few studies have considered the temporal changes of users' writing patterns, and none of them have considered user's behaviors in social networks for author identification. Therefore, we can express our proposed methodology as a novel framework that characterizes users based upon their OSNs behaviors. We named this method the *Time Feature Vector (TFV)*, as it utilizes the timestamps of all posts to identify the text distribution in different time periods. There seems to be no compelling reason to argue that the *circadian typology* [15], the physiological and behavioral measures, of a user are notably important when implementing TFV, particularly with regards to a user's social media routines. Generally, the circadian typology classifies individuals into three different types: morning, evening and neither. Within the social interaction framework it is probable that a morning type individual prefers to post during the morning hours while an evening type person posts during the evening. The foregoing discussion implies how we adapted the circadian behavior of an individual in the social network to identify content originators.

The TFV is applied on users who are classified as the top X users in the Jenks algorithm. In order to generate the TFV, first inspect the frequency of #posts shared in different time periods. We consider four time periods to categorize users according to their circadian behaviors (0-6hrs, 6-12hrs, 12-18hrs, and 18-24hrs). Then, for each time period, the relative frequencies of the #posts are calculated and we denote the TFV as a vector of four elements. The social circadian behavior of a user is the time duration that belongs to the largest element in the TFV vector. The TFV method is shown in algorithm 2. At the end, the timestamp of a given text is used to map with one of the circadian typology periods to identify the potential author(s) as presented in algorithm 1.

IV. Evaluation of the Framework

To carry out an evaluation of the proposed method, this section describes; (1) the most effective parameters in the SCAP method when applying it to identify the content originator of very short texts (tweets), and (2) the efficiency of the originality detection approach when we consider a user's circadian typology. The first step is to have a test dataset that includes similar posts in different user accounts. One of the communities that share many identical posts in social media is news agencies. In this respect, we considered 8 popular news agencies and collected about 145K tweets (Table I). All of the Twitter accounts used in the dataset are legitimate publishers of the respective news agencies. We follow our methodology described in section II, starting with our first question mentioned above.

TABLE I
DATASET DESCRIPTION

News Agency	#users	list of the considered users in the news agency	#followers	Total #tweets
Reuters	12	Reuters, ReutersBiz, ReutersChina, ReutersIndia, ReutersLive, ReutersOpinion, ReutersPakistan, ReutersPolitics, ReutersTV, ReutersUK, ReutersWorld, LukeReuters	144.5K	38,756
BBC	8	BBCBreaking, BBCBusiness, BBCNews, BBCNewsAsia, BBC-NorthAmerica, BBCSport, BBCWalesSport, BBCWorld	8.4M	25,747
CNN	7	CNN, cnnbrk, CNNent, cnntech, CNNMoney, CNNPolitics, cnni	1.9M	22,571
NYT	6	nytimes, nytimesworld, NYTNational, nytopinion, nytpolitics, NYT-Sports	1.1M	19,366
WSJ	6	WSJ, WSJOpinion, WSJPolitics, WSJSports, WSJTech, WSJusnews	391.3K	19,382
Fox	4	FoxBusiness, FoxNews, foxnewspolitics, FoxNewsTech	990.7K	12,933
ABC	2	ABC, ABCPolitics	2.6M	6,463
SC	1	SportsCenter	31.6M	3,225

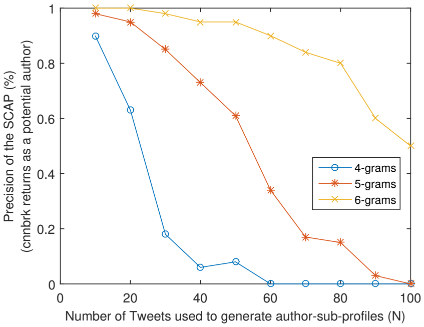

Fig. 2. Precision of SCAP for short texts.

TABLE II
TEST TWITTER DATASET

Twitter account - #Tweets	Precision%	Recall%	F1 Score
BBCSport - 40 ***Injected Tweets: CNN - 10, Reuters - 10***	97.50	100.00	98.73
ABC - 40 ***Injected Tweets: CNN - 10, WSJ - 10***	97.50	95.12	96.30
nytimes - 40 ***Injected Tweets: Reuters - 10, WSJ - 10***	95.00	95.00	95.00

Precision - TP/(TP+FP): the proportion of retrieved Twitter accounts that are relevant where retrieved originator and the considered Twitter account are identical. Recall - TP/(TP+FN): the proportion of relevant Twitter accounts that are retrieved. F1 score: harmonic average of precision and recall.

1) SCAP approach: To give a brief example of the efficiency of the SCAP approach, we considered 100 sample tweets from *cnnbrk* (cnnbrk shows the highest #active followers (47.3M) in our dataset). We assume that, these tweets are originated from *cnnbrk* and therefore we excluded using RT-tweets in the sample dataset. The SCAP method was executed dynamically between each test tweet (timestamp of the tweet:TT) and N tweets published by all news agencies in 1 month before and after TT. The value of N varied from 10 to 100 and in each execution we incremented N by 10. For each set of N tweets, the SCAP method generates n-gram (n in 4,5,6) profiles; in this study we name them author-sub-profiles. We use these author-sub-profiles to measure ASI. Figure 2 depicts, on average, the precision of identifying *cnnbrk* as a potential author of the test Tweet for different N and n values.

The results show that the precision of the SCAP is higher if 6-grams are used to build author-profiles and also, if user's author-sub-profiles are generated using 10 or 20 tweets at a time. In view of that, the accuracy of identifying *cnnbrk* as a potential author among the list of users classified by the SCAP is 100%. We achieved this result after using pre-processed texts, while, raw dataset achieved only 86% accuracy for the same test dataset. This result indicates that the, SCAP method is very efficient for content originality identification. Therefore, we use the SCAP method in our study to identify the content producer, with evaluations based on 6-grams profiles, and use 10 tweets (N=10) to build author-sub-profiles.

2) TFV approach (n=6 & N=10): On average, the SCAP+Jenks algorithm returns 11 users for the considered dataset in Figure 2, where we consider them as the potential originators and they have the same writing styles. Therefore, TFVapproach is applied to reduce #users classified in the SCAP. In this respect, at the end, originator is identified by sorting (based on timestamps) the list of potential authors detected in the TFV method.

As described in section II, the TFV method uses circadian behavior of user's online activities. The circadian behaviors of many of the news agencies (49.01%) considered in this work belong to 12-18hrs durations, including BBCWorld, BBCBreaking, Reuters, ReutersWorld etc. Among all the Twitter accounts we considered, 33.33% of them publish their content between 18h00-24h00. Further, 11.76% of the Twitter accounts publish during 06h00-12h00, and the circadian behaviors of the remaining Twitter accounts fall within 0-6hrs durations. The outcome of the TFV method is determined by these publishing patterns and thus, we identified only 6 users out of 11 users in the example considered in Figure 2. Hence, by using TFV classification on the SCAP result, we can filter for the best potential authors for a given text.

Another evaluation was performed using classic information retrieval metrics: Precision, Recall, and F1 scores, as shown in Table II, based on 3 different test cases that are manipulated using 60 tweets. For each test case, we manually identified 40 tweets from BBCSport, ABC, and nytimes separately. In the test dataset, 2/3 of the tweets are from these 3 main Twitter accounts and 1/3 are injected from other Twitter accounts. We manually checked the whole dataset to verify that the considered 60 tweets originated from the respective Twitter account. We performed the methodology described in this work and present the results in Table II. The results indicate high precision and high recall parameters and therefore we can conclude that the model we used to detect content originality is very accurate. Similarly, the F1 score results also show that the considered approach is very accurate as its values are more than 95% for 3 test scenarios. In particular, among 20 test tweets injected in 3 different cases, only a small portion were classified as False Negatives, in which the TT of the tweet was not within the circadian typology period of the originated account. Therefore, classification shows the injected Twitter account as the originator in those cases.

V. NEWS AGENCIES ANALYSIS - A USECASE

As mentioned earlier, the proposed framework has the potential to be used in many different scenarios. In this section, we use it to analyze the news agency community, as the users in this community (e.g BBC, CNN, etc.) are frequently

TABLE III
DISTRIBUTION OF THE IDENTIFIED ORIGINAL PUBLISHERS OF TWEETS PER NEWS AGENCIES.

News Agency (#users)	**Twitter User** (#tweets)	**Portion of tweets (%) originated by the:**			
		same user	associated users from same agency in dataset (exact user)	users from other news agencies in the dataset (two major ones)	users outside of the dataset or not classified
Reuters (11)	Reuters (3251)	30.52	41.91 (ReutersWorld-29.07%)	7.48 (NYT-23.36%, CNN-20.09%)	21.75
BBC (8)	BBCBreaking (3239)	63.26	5.28 (BBCNewsAsia-44.44%)	4.14 (CNN-44.78%, Reuters-36.57%)	27.32
CNN (7)	cnnbrk (3212)	71.31	8.34 (CNNPolitics-37.1%)	7.13 (Reuters-26.64%, BBC-24.45%)	13.22
New York Times (6)	nytimes (3236)	56.47	11.68 (nytpolitics-42.86%)	6.92 (CNN-53.57%, Reuters-16.52%)	24.93
WSJ (6)	WSJ (3245)	23.94	5.21 (WSJPolitics-36.69%)	29.39 (NYT-34.78%, Reuters-8.71%)	41.46
Fox (4)	FoxNews (3230)	33.12	5.69 (FoxBusiness-77.17%)	8.60 (Reuters-29.49%, CNN-26.62%)	47.41
ABC (2)	ABC (3247)	56.13	6.88 (ABCPolitics-9.35%)	10.53 (CNN-39.88%, Reuters-26.30%)	26.46
SportsCenter (1)	SportsCenter (3225)	46.65	-	1.64 (NYT-60.38%, CNN-22.64)	51.71

reporting news on similar topics and have a great potential to use each other published posts as source of information. It is interesting to understand how the information flow among these users, and how news agencies disseminate their content in Twitter. Towards that end, major news agencies published posts are used in the proposed framework to identify cross posts in each user's profile and analyzed as described in the following paragraphs.

A. Dataset and Scenario Description

We used the collected dataset described in Table I, which includes more than 145K tweets across 46 different Twitter users of 8 major news agencies. The analysis of this section is based on one Twitter user selected from each news agency: the user that exhibits the highest #followers compared with other users from the same news agency.

The main objective is to find different publishing patterns in terms of originality of their content. Further to this, we apply the proposed approach in this study to analyze cross posts among different users with the aim of identifying intra information flow among users in each news agency and information flow patterns between inter news groups.

B. News Story Tellers vs. News Propagators

Similar to many other communities, we expect to have different groups of users in the news agency community in term of original content publication where, a groups of users who have fresh news and actually are somehow news originator (*Story Tellers*) and another group is composed of users who somehow use other users as a source for their posts (*News Propagators*). Many related previous work in economic and management domains are focused on manual data collection whereas, our approach is an automated process[2].

In this section, we apply our framework on the published tweets of a set of major news agencies to understand and quantify *Story Tellers* and *News Propagators* what portion of the published content have very similar text on same topics (Cross-posts). Table III presents the results for the above scenarios considering the most active (depends on #followers) user accounts in each news agency. In addition, Figure 3 visualize the distribution of the originated tweets in different categories we considered in Table III. It clearly indicates that, among the classified content, cnnbrk and BBCBreaking have the highest #tweets originating from them (71% and 63% respectively) compared to other users and they incorporate minimum #tweets originated by other news agencies (shown in column 5 of Table III), indicating that they perform as *Story Tellers* in our dataset. The least #originated tweets is from WSJ (23.94%) and this interprets WSJ as a *News Propagator*. We can see a main source of content for them is NYT where 34.78% of their posts are similar to the published posts in *nytimes*. Even-though a great #followers in Reuters belongs to *Reuters* Twitter account, its many information (41.91%) is obtained from other Reuters users.

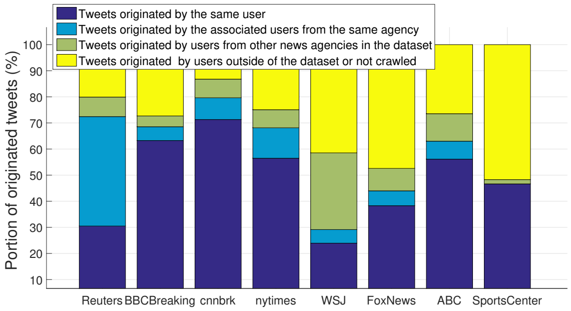

Fig. 3. Distribution of the tweets originated by different users.

This observations show, a big part of the news agencies tend to write the content in a similar way to the content presented in CNN, Reuters, and NYT and importantly, CNN shows only 13.29% of their #tweets were originated by the outside users. The analysis also reveals that news agencies try to publish content mostly on political and business perspectives.

The graph of the presented flow of information is depicted in Figure 4. Let the graph pictured in Figure 4 be represented by G=(V,A) where V is a set of nodes (V=46) that represent 46 Twitter users of 8 news agencies in Table I and A is a set of weighted directed edges (A=252) from one node to another. Assume that, $X, Y \in V$ and edge $X \rightarrow Y \in A$. The weight of the edge, $W_{X \rightarrow Y}$, is calculated using the following equation.

$$\mathrm{W}_{X \rightarrow Y} = \frac{\text{\#tweets published by } Y \text{ but originated in } X}{\text{\#total classified tweets in } Y} \times 100$$

The size of a edge is proportional to the strength of the relationship between nodes. The nodes in the middle of graph G depict users with highest #followers in Twitter.When moving from the center towards the circumference in G, the size of nodes is decreasing, which means that they are less influence on the information flow patterns. The degree of a node is defined as the #propagative relationships from one node (user) to another and vice versa. The greater the degree the more interactions performed with other users and we can

[2] The framework will be public to the community for further research in GitHub in the camera ready version.

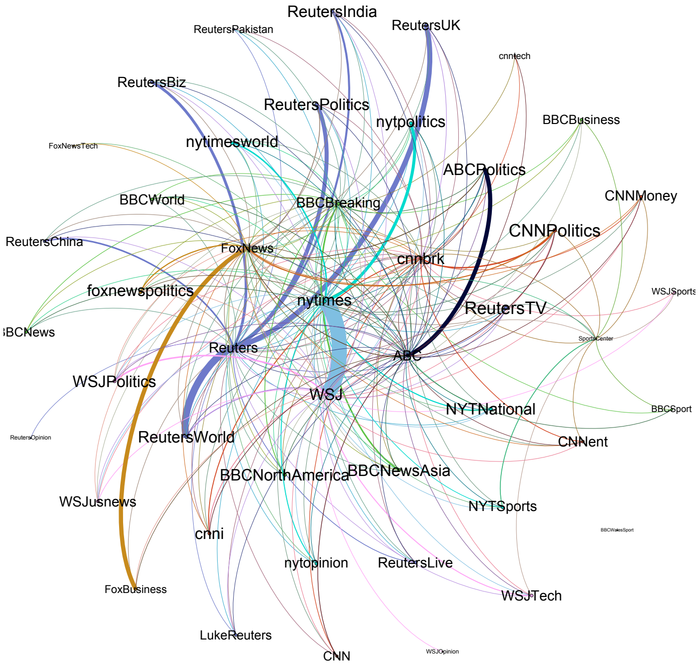

Fig. 4. News propagation flow in intra and inter news agencies. (The flow direction is shown by curve direction. A clockwise curve is an outgoing edge and a counterclockwise curve is an ingoing edge. Edge coloring represents different flow from news agencies).

use in-degree and out-degree of each node (in-degree of node X: from how many users X is duplicating/coping content, out-degree: how many other users are copying content from X) as this is a directed graph. Higher out-degree indicates that other users in the network attracts more information of this node. The highest in-degree (36) is presented by WSJ and ABC, but the weight of the edges are different in both Twitter users. This again proves that WSJ publish more similar information as users. Apart from that, the maximum value of out-degree is 7, identified for cnnbrk, nytimes, and WSJ.

Figure 4 clearly elaborates the findings in Table III. We can observe a big link between WSJ and nytimes showing a majority of WSJ content are similar to posts of nytimes. Also, Reuters share many information directly obtained from its sibling users than others. Additionally, in many news agencies, political information supplier is providing content to others within the same news group indicating that media tend to spread more political information than other categorical news (tech,sports). As shown, CNN acts as a content provider to all other news agencies considered in this study as it does not has any link towards from other news groups.

In nutshell, by applying our proposed framework to the news agencies usecase, we identified different patterns of users behavior in term of original content publication.

VI. Conclusion

We present an advanced framework for detecting the originator of a published content in OSNs by identifying the best features for the SCAP in order to detect author of short texts, and exploring to what extent does user's circadian behaviors in OSNs help to distinguish content originators. The proposed approach is based on users' linguistic features and online circadian behaviors. The evaluation results obtained using a Twitter dataset give high F1 parameters in all the considered scenarios. We show how, by using a user's temporal changes in their writing patterns and their temporal behaviors, the originator of a given text can be recognized with high accuracy. Next, we applied the framework to identify information flow patterns in the context of major news agencies. The dataset used in this scenario consists of 145K tweets of 46 Twitter accounts belong to 8 different major news agencies. Our observation revealed two different categories (*News Story Tellers* and *News Propagators*) of new agencies based on their publication patterns. As future direction, we are going to apply the proposed method on a larger dataset to analyze more deeply the news agency behavior worldwide and expanding across different social media. In addition, spotting fake news in social media based on the originality detection framework is a target as the future direction of this research.

References

[1] R. Motamedi, R. Gonzalez, R. Farahbakhsh, A. Cuevas, R. Cuevas, and R. Rejaie, "Characterizing group-level user behavior in major online social networks," available at: https://goo.gl/UhF1dS, Tech. Rep., 2014.

[2] R. Farahbakhsh, A. Cuevas, and N. Crespi, "Characterization of cross-posting activity for professional users across facebook, twitter and google+," *Social Network Analysis and Mining*, vol. 6, no. 1, 2016.

[3] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," National Bureau of Economic Research, Tech. Rep., 2017.

[4] A. Narayanan, H. Paskov, N. Z. Gong, J. Bethencourt, E. Stefanov, E. C. R. Shin, and D. Song, "On the feasibility of internet-scale author identification," in *2012 IEEE Symposium on Security and Privacy*. IEEE, 2012, pp. 300–314.

[5] S. Afroz, A. C. Islam, A. Stolerman, R. Greenstadt, and D. McCoy, "Doppelgänger finder: Taking stylometry to the underground," in *2014 IEEE Symposium on Security and Privacy*. IEEE, 2014, pp. 212–226.

[6] G. Frantzeskou, E. Stamatatos, S. Gritzalis, C. E. Chaski, and B. S. Howald, "Identifying authorship by byte-level n-grams: The source code author profile (scap) method," *International Journal of Digital Evidence*, vol. 6, no. 1, pp. 1–18, 2007.

[7] R. Layton, P. Watters, and R. Dazeley, "Authorship attribution for twitter in 140 characters or less," in *Cybercrime and Trustworthy Computing Workshop (CTC), 2010 Second*. IEEE, 2010, pp. 1–8.

[8] H. Azarbonyad, M. Dehghani, M. Marx, and J. Kamps, "Time-aware authorship attribution for short text streams," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2015, pp. 727–730.

[9] E. Stamatatos, "A survey of modern authorship attribution methods," *Journal of the American Society for information Science and Technology*, vol. 60, no. 3, pp. 538–556, 2009.

[10] M. Almishari, D. Kaafar, E. Oguz, and G. Tsudik, "Stylometric linkability of tweets," in *Proceedings of the 13th Workshop on Privacy in the Electronic Society*. ACM, 2014, pp. 205–208.

[11] Y. Seroussi, F. Bohnert, and I. Zukerman, "Authorship attribution with author-aware topic models," in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2*. Association for Computational Linguistics, 2012, pp. 264–269.

[12] W. X. Zhao, J. Jiang, J. Weng, J. He, E.-P. Lim, H. Yan, and X. Li, "Comparing twitter and traditional media using topic models," in *European Conference on Information Retrieval*. Springer, 2011, pp. 338–349.

[13] R. Overdorf and R. Greenstadt, "Blogs, twitter feeds, and reddit comments: Cross-domain authorship attribution," *Proceedings on Privacy Enhancing Technologies*, vol. 2016, no. 3, pp. 155–171, 2016.

[14] S. Afroz, M. Brennan, and R. Greenstadt, "Detecting hoaxes, frauds, and deception in writing style online," in *2012 IEEE Symposium on Security and Privacy*. IEEE, 2012, pp. 461–475.

[15] J. A. Horne and O. Ostberg, "A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms." *International journal of chronobiology*, vol. 4, no. 2, pp. 97–110, 1975.