Detection of Satiric News on Social Media: Analysis of the Phenomenon with a French Dataset

Zhan Liu
Institute of Information Systems
University of Applied Sciences and Arts Western Switzerland
(HES-SO Valais-Wallis)
Sierre, Switzerland
zhan.liu@hevs.ch

Shaban Shabani
University of Applied Sciences Western Switzerland
HEG Arc HES-SO
Neuchâtel, Switzerland
shaban.shabani@he-arc.ch

Nicole Glassey Balet
Institute of Information Systems
University of Applied Sciences and Arts Western Switzerland
(HES-SO Valais-Wallis)
Sierre, Switzerland
nicole.glassey@hevs.ch

Maria Sokhn
University of Applied Sciences Western Switzerland
HEG Arc HES-SO
Neuchâtel, Switzerland
maria.sokhn@he-arc.ch

Abstract—The topic of deceptive and satiric news has drawn attention from both the public and the academic community, as such misinformation has the potential to have extremely adverse effects on individuals and society. Detecting false and satiric news automatically is a challenging problem in deception detection, and it has tremendous real-word political and social influences. In this paper, we contribute a useful French satiric dataset to the research community and provide a satiric news detection system using machine learning to automate classifications significantly. In addition, we present the preliminary results of our research designed to discriminate real news from satiric stories, and thus ultimately reduce false and satiric news distribution.

Index Terms—news satire, French dataset, fake news, machine learning, classification, social media

I. Introduction

Deceptive and satiric news is defined as that written and published with the intent to mislead its readers to gain financially or politically, often with sensationalist, exaggerated, or patently false headlines to attract readers' attention. Recently, this has been a popular topic throughout the world, with a 365% increase in the term's use in 2019 according to Digital marketing Ramblings [1]. In a recent report Statista released in 2017 [5], 31% of survey respondents believed that mainstream media sometimes reported false news. The same situation occurred in Switzerland. According to a qualitative study of "fake news" Zurich-based consultancy Publicom conducted that was published in the 2017 edition of the annual MediaBrands [4], Swiss Media's average credibility score was 4.5 of a maximum 6 possible, while the score for online social media in Switzerland was only 3. Deceptive and satiric news has become an unfortunate element of the media landscape that spreads wildly across computer screens and mobile phones and is shared on social media networks. A poll Morning Consult conducted

for the Alliance in September 2017 found that 41 percent of people have turned to their social networks for news. Online social networks today have become a strong communication channel that engages people in political, environmental, civic, and social events. However, huge amounts of data and multiple sources of unverified information derived from untraceable sources are a critical problem in the context of data processing, source verification, and information quality, among others. This indicates that social network platforms that have, in some ways, democratized news more than ever before, are allowing the public to be deceived every day by false news that scrolls across their screens. As a distribution channel, the Internet has replaced traditional media for most people, and breaks down barriers to content creation and dissemination, such that the line between content users generate and traditional media content has become blurred increasingly. Therefore, before we think about the way we can combat such false and satiric news, we need firstly to understand when and whom it affects.

Obviously, there are different types of false news that has been classified usually in different ways. Some of these are false deliberately (i.e., disinformation—the willful creation and dissemination of information known to be false), others are attributable to human error or biases (i.e., misinformation—the inadvertent sharing of false information). In either case, they are connected very loosely with the truth and basically lay along a continuum of the intent to deceive. According to [21], false news may be defined according to 4 types: news parody, photo manipulation, advertising and public relations, and propaganda. Another report from [3] categorized fabricated news into 5 types: 1) satire or parody; 2) misleading news that is somewhat true, but used in the wrong context; 3) sloppy reporting that fits an agenda; 4) misleading news that is not

based on facts, but supports an on-going narrative, and 5) intentionally deceptive. Moreover, First Draft News [2] made an interesting effort to classify 7 types of misinformation and disinformation, based on the type of content, the motivations of those who create it, and the ways that content is disseminated:

- Satire or parody - no intention to cause harm, but has potential to deceive
- False connection - when headlines, visuals, or captions do not support the content
- Misleading content - misleading use of information to frame an issue or individual
- False context - when genuine content is combined with false contextual information
- Imposter content - when genuine sources are impersonated
- Manipulated content - when genuine information or imagery is manipulated to deceive
- Fabricated content - new content that is 100% false, and designed to deceive and do harm

People form opinions based on media news coverage. An analysis [7] presented in 2006 has shown the crisis in the predominance of news in the following order: attribution of responsibility, economic, conflict, and human interest. News can influence consumers' perceptions, and to a certain degree, their opinions. Although today's communication channels are improving worldwide interactions, the dissemination of uncorroborated information has affected the content's credibility negatively. The Oxford Dictionary listed "post-truth" as its 2016 international word of the year. Post-truth denotes "…circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief" (Dictionary, O. E. Oed online). Thus, false and satiric news rapidly has become an umbrella term that discredits all forms of stories with real-world consequences.

Overall, the tone with which unverified and misleading information is delivered affects both producers and consumers strongly. Problems misleading information causes in some sectors may vary from discrediting a person's image to those that have a more critical global influence. For example, in tourism, the proliferation of fabricated and satiric reviews have a dramatic effect on a business's online reputation and consequently, its performance [16]. Politicians have expressed the need to access verified information and analyze the critical issue of the way an individual user with no track record or reputation can reach so many readers rapidly, as is the case of well-known news chains [6]. As a result, misleading stories might have a large influence even on presidential elections [6] [10].

The problem of detecting false and satiric news is more challenging than detecting deceptive reviews, as the posts on social media networks, or political language in TV interviews largely are short statements. Automatic satiric news detection is conducted primarily with two different approaches: machine learning and natural language processing. Machine learning technique systems examine the statement and score it based on how likely it is to be untrue. This solution generates a message to explain why the score was assigned, and thus, users understand the way the program flagged the statement as fabricated or factual news. Unlike the machine learning approach, natural language processing systems organize and process unstructured information to derive insights from large datasets using linguistic cues, which is useful in categorizing vast volumes of online statements. However, people are not satisfied with the current perceived accuracy of automatic satiric news detection, as sometimes these solutions cross different topics and domains that still are not understood well. In addition, none of them focuses on verifying the statements contained in news and opinion articles.

In this study, we focused on contributing a useful French dataset to the research community and provide a satiric news detection system by applying machine learning models to classify news articles as satiric or accurate automatically. To accomplish this goal, the paper provides answers to the following research questions:

- What methods can be used to identify French satirical articles on social media that is labelled accurately?
- What features are important in detecting satiric news automatically using machine learning models?

The paper is organized as follows: Section 2 presents the background and related works. In section 3, we introduce our dataset and the data collection methods. Then, we explain the approaches we used to detect satiric stories automatically and provide some preliminary results in section 4. Finally, we conclude with a summary of the current work, and offer suggestions for future research.

II. Related Work

Fabricated and satiric news spreads faster and deeper online than does true news. A recent study from MIT [22] examined a decade of tweets that included 126,000 news stories 3 million people shared on Twitter 4.5 million times, and verified that false information outperformed true information—false and satiric news spreads faster and deeper than true stories, and people, not bots, primarily were responsible for the spread of disinformation. Interestingly, they found that satiric information was significantly more novel and negative than the truth. Clearly, satiric news's novelty makes it more appealing to share, and people are more likely to share novel and negative stories on social media that are emotionally or morally charged to attract people's attention, even if they are unverified. Thus, a negative story has a relatively higher risk of containing false information that will be disseminated more widely.

A dataset of false and satirical stories was published recently [11], and consists of news published only in English and related to American politics. The authors extracted, labelled, and verified 483 news stories from different sources manually. Work [19] conducted showed that automatic classification can be 81% accurate in identifying satiric stories and fabricated news. Moreover, their hybrid confidence approach based on machine learning and crowdsourcing achieved a final accuracy of 87%.

Among several subject areas dissemination of satiric news affects, recent efforts' goal has been to verify the facts in journalism articles from multimodal online sources. False and satiric news providers use automated agents, called social bots, to create false accounts on social media to disseminate manipulated content and misinformation. BotOrNot [9] is a web service available publicly that determines automatically whether a social media account on Twitter is a real account or a social bot. The system takes advantage of more than one thousand features to evaluate the degree to which a Twitter account exhibits a similarity to social bots' known characteristics.

During the past decade, several works have been published in the context of verifying facts and identifying false and satiric news [12], [17], [23]. Several recent works are related specifically to false news detection, and [20] provided a comprehensive review of this technique. In 2012, Google launched its Digital News Initiative (DNI) to approach the future of media and promote innovation in digital journalism. Moreover, [8] reviewed the increasing interest in digital journalism topics since 2014. A community both of professionals and nonprofessionals has undertaken other initiatives designed to fight deceptive and satiric journalism with fact verification, as is the case of WikiTribune, launched in 2017. Atlantic Media recently has exploited artificial intelligence resources to promote "conversational journalism" freely. By using embedded chatbots, its Quartz initiative seeks to provide the public with verified publications.

TABLE I
ENGLISH LANGUAGE PUBLIC DATASETS RELATED TO FAKE AND SATIRE NEWS

Dataset	Records	Type of data	Classes
KaggleBS	12,999	posts	10 (fakenews, satire, etc)
LIAR	12,836	short statements	6 (true, false, pants-fire, etc)
Emergent	2,595	news articles: headline	3 (for, against, observing)
fnc-1	49,972	news articles: headline and body	4 (agree, disagree, discuss, unrelated)

Today, automatic false and satiric news classification has been reported to be accurate under specific scenarios. Table I presents the public datasets available most widely that have been used for research on automatic classification. However, most existing studies have focused on mining social data in the English language, and very little research has been conducted using satiric news detection methods with French sources. Moreover, the French public satiric news dataset for scientific research is still in its infancy. Therefore, we contributed a French dataset of satirical stories in this paper to fill the research gap.

III. DATASET DESCRIPTION

This study analyzed French satirical stories on Twitter. As we wished to build our machine learning model using a trusted dataset with rich information sources, we defined a set of guidelines for data collection at the beginning of the project, as follows:

- Define "satiric news". There are several definitions for satiric news. In this project, we used the following definition: "...a type of parody presented in a format typical of mainstream journalism, and called a satire because of its content. News satire has been around almost as long as journalism itself, but it is particularly popular on the web, with websites like The Onion, where it is relatively easy to mimic a legitimate news source. News satire relies heavily on irony and deadpan humor." Two slightly different types of news satire exist. One uses satirical commentary and sketch comedy to comment on real-world news events, while the other presents wholly fictionalized news stories. Based on this definition, we selected several trustworthy Twitter accounts to analyze their tweets that include both real and satirical content.
- Enrich content. Although Twitter doubled the length of individual tweets allowed from 140 to 280 characters, this number still is not sufficient to use for machine-learning training and obtain a good detection accuracy from automatic classifications. However, the URL that is provided with each tweet allowed us to crawl for more useful contents from their official websites, such as the article's title, snippet, and body of the text. These additional contents can be used for training in machine learning models.
- Diversify sources. Because there are more real French news sources than satirical news websites, we created a diverse set with articles from different sources to avoid common writing style issues. Moreover, we balanced the number of real and satirical articles to avoid classification bias and fitting to the sources.
- Diversify topics. To date, most existing English datasets include news related to politics only. However, satiric news is not limited to one topic, and can cover any domain. To this end, we collected satirical articles on a variety of topics, including economics, politics, sports, culture, people, and technology. In addition, we balanced the number of articles on each topic to ensure consistent classification among all topics.
- Current articles. We attempted to use current data, so we collected factual articles from different sources that were posted since January 2018, and satirical articles posted since January 2016. The time frame difference was attributable to the number of data available to ensure the subjects the articles discussed were similar.

We began by identifying the factual and satiric news in newspapers' Twitter accounts. Our goal was to select true accounts according to specific criteria, such as the number of followers, retweets, likes, and activity levels. Finally, four French daily newspapers were selected to collect accurate articles: Le Monde and Le Figaro from France, and 20 Minutes and Le Matin from Switzerland. Moreover, two French websites for satiric news were used to collect satirical articles:

Le Gorafi and Nordpresse.

Source	Class	Columns	Records
Le Figaro	real	title, snippet, body, url	1,000
Le Monde	real	title, snippet, body, url	1,393
Le Matin	real	title, snippet, body, url	198
20 Minutes	real	title, snippet, body, url	250
Le Gorafi	satire	title, snippet, body, url	1,745
Nordpresse	satire	title, snippet, body, url	1,096

The most efficient method to collect data from Twitter is to use the Twitter public API (Application Programming Interface), because the data are returned in a data-interchange format, JSON, which is simple for machines and humans to read and analyze. Once the tweets were collected with the corresponding URLs that link with the original news contents, we developed one web crawler for each source's website to obtain the additional information. As shown in Table II, our benchmark dataset includes a total of 5,682 French articles, half of which half are accurate and half of which are satirical. Each record includes a tweet message, title, snippet, full article text, the URL to the source website, and the classification. Table III presents two exemplar records from the dataset of real and satirical articles.

Title	Il traverse la rue et se fait percuter par un emploi
Snippet	Lucas n'a pas pris au sérieux les paroles du président Emmanuel Macron et a failli le payer de sa vie.
Body	Hier, Lucas, 33 ans, a traversé la rue sans regarder et s'est fait percuter par un emploi. "Je suis bête, je sais qu'il y en a partout, les emplois pullulent dans la rue. J'étais dans mes pensees et j'ai traverse en dehors du passage piéton et paf! Je me suis pris un CDD de 3 mois au smic horaire dans la restauration en pleine face" nous confie ce professeur d'EPS du 12ème arrondissement de Paris...
URL	http://legorafi.fr/u/7o9
Classified	satire

Title	Emmanuel Macron explique la crise des "gilets jaunes" aux enfants
Snippet	Le président a rencontré des éluspuis des enfants pour tenter de rassurer sur l'issue du grand débat national.
Body	Question compliquée, réponse simple. C'est parfois l'avantage avec les enfants que de pouvoir esquiver la complexité. Pour sa treizième participation à un «grand débat», Emmanuel Macron a choisi de s'adresser à une quarantaine d'écoliers de CM2 et de 6e dont certains font partie de conseils municipaux de Beaupréau-en-Mauges et de villages alentours. Au bout d'une heure d'échanges...
URL	https://t.co/OllWidKljm
Classified	real

We analyzed the dataset to check the text available in the title, snippet, and the text body individually. There were approximately 12 words in the title section, 26 in the snippet, and approximately 343 words in the text body. Overall, considering the three columns jointly, there were 380 words on average, in which the shortest articles had only 4 words and the longest 4,406 words. Figure 1 shows the ratio of words per article.
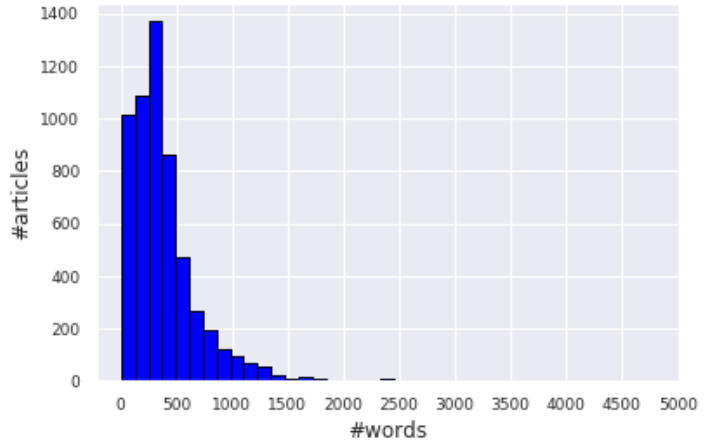

Fig. 1. Words per articles distribution

IV. AUTOMATIC DETECTION OF SATIRIC STORIES

After completing and labeling the dataset, our next step was to analyze the application of machine learning to classify the news articles collected automatically. The task is defined as a binary classification problem with two classes: factual and satire. For this text classification task, we used only the text data available from the news: title, snippet, and body. Our goal was to inspect and evaluate whether text-based data are sufficient to build models that detect satiric news stories with high accuracy.

Initially, we split the dataset into an 80% training set (4,545 samples) and 20% test set (1,137 samples). We used the *training_samples* method from Scikit-Learn [14], which splits the data into training and test sets randomly while maintaining a balance in the source (newspapers articles).

A. Feature Extraction

With the training and test sets ready, the next step was to merge the text from the title, snippet, and body into a single vector of words. Thereafter, we performed a text cleaning process that removes special characters and stopwords. Then, we used the tf-idf [18] method to compute term-frequency and inverse document-frequency with n_grams (1,5), and ended with a feature vector length of 3000. Before conducting training, cross-validation, and testing, we ran the *min-max* feature normalization method to ensure that all data points had the same scale and were equally important.

B. Evaluation

For the evaluation, we chose different models for testing. Results are presented in Table IV. We chose different test models in the evaluation, the results of which are presented in Table IV. In the 10-fold cross-validation phase, the Logistic Regression algorithm performed best, with an accuracy of

95.25%, followed by Neural Networks with 94.68%, and SVM with 93.58%, while in the test phase, the Logistic model again achieved the highest accuracy of 92.17%. Another advantage of the Logistic model is that it is simpler and faster compared to SVM and Neural Networks. Moreover, the feature extraction and classification pipeline includes no external tools to extract additional features, but considers only the text available from news, which makes the process of classification simple overall.

TABLE IV
ACCURACY OF THE CLASSIFICATION MODELS

Model	Cross-Validation	Testing
Neural Networks	94.68%	89.88%
SVM	93.58%	89.01%
Random Forest	95.18%	86.10%
Logistic Regression	**95.25%**	**92.17%**
Naive Bayes	92.32%	85.14%

V. DISCUSSION AND CONCLUSION

In this paper, we built a French-language dataset of satirical stories to serve as a contribution to the false news research community. Our project is still in the development stage, and we are considering including more data from other sources, such as *secretnews.fr*; *jesuissceptique.com*. The final dataset will be available and published at the end of 2019. To date, 5,682 French-language articles from six different sources have been collected that include the articles' full text and topic, as well as labels indicating their type (real or satire). We hope this dataset will be useful for researchers to fill the research gap in the field of automatic false news detection for French-language articles.

Moreover, the preliminary results were promising, as the accuracy of automatic detection reached 92.17% with the Logistic Regression algorithm. Such an achievement also validates the quality of our dataset in this early stage, although the test and training datasets derive from the same source, which may cause some noise. In addition, the exploration of other features, such as sentiment-related features, the text similarity features extracted by querying Google [13], and the paralinguistic features using LIWC [15], could provide a better understanding of the detection of satirical articles based on a French-language dataset.

Future research should concentrate on adding more news sources, enriching the dataset with French-language news articles from diverse sources in order to avoid the issue of fitting to the source. In particular, the test and training datasets should differ. In addition, because of the language used and because false news generators enhance their writing style, which makes these articles more sophisticated, automatic false news detection using only machine learning algorithms is a challenging task. This problem requires fact-checking, and humans are known to perform better in these scenarios. We believe that using humans' cognitive and fact-checking skills will improve the accuracy in detecting false news. Therefore, in the next step of our project, we intend to design and develop a hybrid false news detection system using both machine learning to automate the procedure significantly, and active learning from crowdsourcing with human validation to improve the accuracy rate further.

REFERENCES

[1] Digital marketing ramblings, alarming fake news statistics and facts, https://expandedramblings.com/index.php/fake-news-statistics/, (Date last accessed 01-March-2019)

[2] First draft news, fake news. it's complicated., https://firstdraftnews.org/fake-news-complicated/, (Date last accessed 01-March-2019)

[3] Five types of 'fake news' and why they matter, https://www.ogilvy.com/, (Date last accessed 01-March-2019)

[4] Publicom, kein glaubwürdigkeitsverlust bei schweizer medienmarken, https://www.publicom.ch/2017/09/08/kein-glaubwuerdigkeitsverlust-bei-schweizer-medienmarken/, (Date last accessed 01-March-2019)

[5] Statista, most likely sources of fake news stories in the united states, https://www.statista.com/statistics/697774/fake-news-sources/, (Date last accessed 01-March-2019)

[6] Allcott, H., Gentzkow, M.: Social media and fake news in the 2016 election. Journal of economic perspectives **31**(2), 211–36 (2017)

[7] An, S.K., Gower, K.K.: How do the news media frame crises? a content analysis of crisis news coverage. Public Relations Review **35**(2), 107–112 (2009)

[8] Ausserhofer, J., Gutounig, R., Oppermann, M., Matiasek, S., Goldgruber, E.: The datafication of data journalism scholarship: Focal points, methods, and research propositions for the investigation of data-intensive newswork. Journalism (2017). https://doi.org/10.1177/1464884917700667

[9] Davis, C.A., Varol, O., Ferrara, E., Flammini, A., Menczer, F.: Botornot: A system to evaluate social bots. In: Proceedings of the 25th International Conference Companion on World Wide Web. pp. 273–274. International World Wide Web Conferences Steering Committee (2016)

[10] Dewey, C.: Facebook fake-news writer:'i think donald trump is in the white house because of me'. The Washington Post **17** (2016)

[11] Golbeck, J., Mauriello, M., Auxier, B., Bhanushali, K.H., Bonk, C., Bouzaghrane, M.A., Buntain, C., Chanduka, R., Cheakalos, P., Everett, J.B., Falak, W., Gieringer, C., Graney, J., Hoffman, K.M., Huth, L., Ma, Z., Jha, M., Khan, M., Kori, V., Lewis, E., Mirano, G., Mohn IV, W.T., Mussenden, S., Nelson, T.M., Mcwillie, S., Pant, A., Shetye, P., Shrestha, R., Steinheimer, A., Subramanian, A., Visnansky, G.: Fake news vs satire: A dataset and analysis. In: Proceedings of the 10th ACM Conference on Web Science. pp. 17–21. WebSci '18, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3201064.3201100, http://doi.acm.org/10.1145/3201064.3201100

[12] Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, B.J., Wong, K.F., Cha, M.: Detecting rumors from microblogs with recurrent neural networks. In: IJCAI. pp. 3818–3824 (2016)

[13] Olivieri, A., Shabani, S., Sokhn, M., Cudré-Mauroux, P.: Creating task-generic features for fake news detection. In: Proceedings of the 52nd Hawaii International Conference on System Sciences (2019)

[14] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011)

[15] Pennebaker, J.W., Francis, M.E., Booth, R.J.: Linguistic inquiry and word count: Liwc 2001. Mahway: Lawrence Erlbaum Associates **71**(2001), 2001 (2001)

[16] Phillips, P., Barnes, S., Zigan, K., Schegg, R.: Understanding the impact of online reviews on hotel performance: an empirical analysis. Journal of Travel Research **56**(2), 235–249 (2017)

[17] Ruchansky, N., Seo, S., Liu, Y.: Csi: A hybrid deep model for fake news detection. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. pp. 797–806. CIKM '17, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3132847.3132877, http://doi.acm.org/10.1145/3132847.3132877

[18] Salton, G., McGill, M.J.: Introduction to modern information retrieval (1986)

[19] Shabani, S., Sokhn, M.: Hybrid machine-crowd approach for fake news detection. In: 2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC). pp. 299–306. IEEE (2018)

[20] Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H.: Fake news detection on social media: A data mining perspective. SIGKDD Explor. Newsl. **19**(1), 22–36 (Sep 2017). https://doi.org/10.1145/3137597.3137600, http://doi.acm.org/10.1145/3137597.3137600

[21] Tandoc Jr, E.C., Lim, Z.W., Ling, R.: Defining "fake news" a typology of scholarly definitions. Digital Journalism **6**(2), 137–153 (2018)

[22] Vosoughi, S., Roy, D., Aral, S.: The spread of true and false news online. Science **359**(6380), 1146–1151 (2018)

[23] Wang, Yang, W.: 'liar, liar pants on fire': A new benchmark dataset for fake news detection. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). pp. 422–426. Association for Computational Linguistics (2017). https://doi.org/10.18653/v1/P17-2067, http://aclweb.org/anthology/P17-2067