

# **I Workshop en Ciencia de Datos en Redes Sociales (CidReS 2018)**

DAVID CAMACHO  
MARÍA VICTORIA LUZÓN

**Secretaría Técnica:**  
EUGENIO MARTÍNEZ CÁMARA





---

# **I Workshop en Ciencia de Datos en Redes Sociales (CidReS 2018)**

## **SESIÓN 1**







# Using Facebook Reactions to Recognize Emotion in Political Domain

Flor Miriam Plaza-del-Arco, Salud María Jiménez-Zafra,  
 María-Teresa Martín-Valdivia, L. Alfonso Ureña-López  
*Department of Computer Science, Advanced Studies Center in ICT (CEATIC)*  
*Universidad de Jaén*  
 Campus las Lagunillas, 23071, Jaén, Spain  
 fmplaza@ujaen.es, sjzafra@ujaen.es, maite@ujaen.es, laurena@ujaen.es

**Abstract**—Opinion Mining and Emotion Mining are part of the Sentiment Analysis area, but they have different objectives. Opinion Mining is concerned with the study of opinions expressed in texts and its basic task is polarity detection, whereas Emotion Mining is related to the study of emotions and its basic task is emotion recognition. Polarity detection is usually a binary classification task with outputs such as *positive* vs. *negative* or *like* vs. *dislike*, while emotion recognition aims to enable computers recognize and express emotions. In this paper we focus on Spanish emotion classification. We first compile a corpus from Facebook using the reactions in comments and posts in order to label different emotions. Then we apply a basic machine-learning approach and two lexicon-based approaches, one using a Spanish version of the NRC Emotion Lexicon (Emolex) and another adapting WordNet-Affect to Spanish. The results demonstrate the difficulty of the task and show some interesting features in the lexicon approaches.

**Index Terms**—Emotion Mining, Natural Language Processing, Social media, Facebook reactions, lexicon, machine learning, Emolex, WordNet-Affect

## I. INTRODUCTION

Emotion classification is a new task that combines several disciplines including Artificial Intelligence and Psychology, although Natural Language Processing is perhaps the most challenging area [1].

Recognize emotions in texts is becoming more and more important due to the fact that it can provide substantial benefits for different sectors [2], for instance detecting signs of depression [3], identifying cases of cyber-bullying [4] or contributing to improve student motivation and performance [5].

On the other hand, social media have changed the way people interact, as they allow the creation and exchange of user-generated content. Social networks are one of the main social media sites in which subjective information is published. Not only this social emotions are expressed in social media networks like Twitter or Facebook but also are said to have a high impact on public discourse and communication in society [6].

In this paper we focus on emotion recognition in Spanish over Facebook posts and comments for several reasons:

- (i) Emotion mining is a difficult task and the current results are not as accurate as those of polarity classification due to its multi-label nature.
- (ii) Most of the existing studies are focused on English, but the presence of other languages on the web is greater every day. Spanish is the second most spoken language in the world and in the two main social networks: Facebook and Twitter.<sup>1</sup>
- (iii) There are few resources and corpora for emotion recognition in Spanish.
- (iv) Facebook is currently the most popular social network. It has approximately 2,167 million of users according to a study of statista<sup>2</sup>.

Therefore, in order to advance in emotion recognition in Spanish, we present a corpus of posts and comments in the political domain which was compiled from the Facebook page *DignidadResponsabilidad*<sup>3</sup>, a Spanish popular page about politics. We first propose a basic Machine Learning (ML) approach to classify emotions in Spanish using the Support Vector Machine (SVM). Then we compare the ML method with two lexicon-based approaches. The first one uses the Emolex Spanish lexicon [7] and the second one proposes adapting the well-known WordNet-Affect (WNA) resource [8] to Spanish.

The rest of the paper is organized as follows: Section 2 describes some related studies; Corpora, ML and lexicon-based approaches are presented in Section 3; Section 4 shows the results and discussion, and finally, our conclusions are presented in Section 5.

## II. BACKGROUND

Emotion recognition is becoming very popular, and some of the main conferences dealing with data and text mining and evaluation are currently including workshops and share tasks

<sup>1</sup>[http://www.cervantes.es/sobre\\_instituto\\_cervantes/prensa/2017/noticias/Presentaci%C3%B3n-Anuario-2017.htm](http://www.cervantes.es/sobre_instituto_cervantes/prensa/2017/noticias/Presentaci%C3%B3n-Anuario-2017.htm)

<sup>2</sup><https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users>

<sup>3</sup>[https://www.facebook.com/DignidadResponsabilidad/?ref=br\\_rs/](https://www.facebook.com/DignidadResponsabilidad/?ref=br_rs/)

related to it. These include Semantic Evaluation (SemEval) [9], Computational Approaches to Subjectivity and Sentiment Analysis (WASSA)<sup>4</sup> and workshops on Computational Modeling of People's opinions, personality and emotions in Social Media (PEOPLE)<sup>5</sup>.

Scientific studies on classification of human emotions date back to 1960s. From a psychological perspective, an emotion is basic only if it does not contain another emotion; that is, if it represents an atomic, irreducible psychological construct [10]. Moreover, emotions can be recognized by humans all over the world regardless of their race, culture, and language. Many theorists have proposed sets of emotions that tend to be basic ones. While psychologists do not agree on what model describes more accurately the set of basic emotions, the most widely used in computer research is the one proposed by Ekman [11], with 6 emotions (anger, disgust, fear, joy, sadness and surprise) [12]. There are different works that use this model to create labeled emotions corpora or to identify emotions in the text. For example, Mohammad [13] created a corpus of emotional tweets from Twitter (TEC)<sup>6</sup>. He targeted the six basic emotions proposed by Ekman and chose six hashtags addressing these emotions to search for appropriate tweets. Another work is the one proposed by Strapparava and Mihalcea [14]. They describe the construction of a large data set annotated for six basic emotions and propose and evaluate several knowledge-based and corpus based methods for the automatic identification of these emotions in text.

Emotion mining techniques can be classified into two categories: lexicon based approaches and machine learning approaches [15]. The first one is based on lexical resources such as lexicons, bags of words or ontologies. The second approach applies ML algorithms based on linguistic features.

Several interesting studies have explored emotion recognition and most of them deal with English texts [16]. Many of these studies focus on evaluating information from social networks since they are appropriate places to share one's feelings easily and widely. Recognize automatically emotions in social media texts can provide the tools to researchers, and citizens in general, to monitor the pulse of the society towards specific topics of interest, a task traditionally accomplished only through opinion polls, which are costly and time consuming to conduct, and therefore frequently limited to small sample sizes. Most of the works have explored emotions from Twitter and different machine learning techniques [17]. For example, Purver and Battersby [18] exploit both emoticons and Twitter hashtags for emotion recognition, Mohammad and Kiritchenko [19] use hashtags to capture fine emotion categories from tweets and Bollen et al. [20] analyze emotions of all tweets in a specific time frame using a psychometric test, names "Profile of Mood States" (POMS).

<sup>4</sup><https://wt-public.emm4u.eu/wassa2018/>

<sup>5</sup><https://peopleswksh.github.io/index.html>

<sup>6</sup><http://saifmohammad.com/WebPages/lexicons.html>

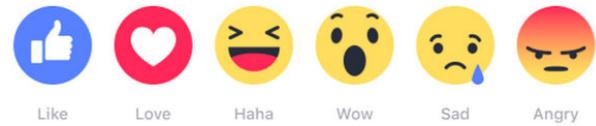


Fig. 1. Facebook reactions.

Regarding Facebook, there are few papers centered on emotion recognition, partly because it is difficult to get a labeled dataset for such a purpose. However, since February 2016 Facebook users have been able to express specific emotions in response to a given post or comment due to the newly introduced reaction feature. It has been observed that this new feature helps Facebook to know much more about its users and exploit this information for targeted advertising [21]. Krebs et al. [22] collected posts and their reactions from Facebook pages and constructed a dataset. They combine sentiment analysis and emotion mining techniques with neural network architectures in order to predict the distribution of reactions on Facebook posts. Moreover, they present a simple visualization environment. Pool and Nissim [23] take advantage of the Facebook reaction feature in a distant supervised fashion in order to train an SVM classifier for emotion detection, using several feature combinations and combining different Facebook pages. In our paper we follow the same idea to exploit the Facebook reaction feature, but applied to Spanish posts and comments.

### III. METHODOLOGY

In this section, we describe the process for generating the corpora and the approaches applied.

#### A. Corpus generation

In February 2016, Facebook made a more explicit reaction feature available world-wide. *Reactions* is an extension of the *Like* button which gives people more ways to express themselves and share their reaction to a post or comment. The collection of reactions include *Like*, *Love*, *Haha*, *Wow*, *Sad* and *Angry* (Figure 1).

We collected Facebook comments and posts with their corresponding reactions from a public page using the Facebook graph API<sup>7</sup>. We chose the Facebook page DignidadResponsabilidad<sup>8</sup>. It is a page that contains Spanish post and comments in the political domain. Currently it has 140,223 followers and 156,891 people like it. We downloaded the available posts and comments of this page from February 2016 to December 2017, recovering also the counts of reactions for each post and comment. The posts and comments are saved following a JSON structure. In the Figure 2 it can be seen an example.

<sup>7</sup><https://developers.facebook.com/docs/graph-api/>

<sup>8</sup>[https://www.facebook.com/DignidadResponsabilidad/?ref=br\\_rs/](https://www.facebook.com/DignidadResponsabilidad/?ref=br_rs/)



```
{
  {
    "status_id": "120857194649240_1259461114122170"
    "message": "REAL COMO LA VIDA MISMA Los Morancos y
    su particular versión sobre el juicio de Urmangarín."
    "status_published": "2017-03-12 08:31:59"
    "num_reactions": 88
    "num_comments": 3
    "num_shares": 32
    "num_likes": 74
    "num_loves": 4
    "num_wows": 0
    "num_hahas": 7
    "num_sads": 3
    "num_angrys": 0
  }
}
```

Fig. 2. Sample of resulting JSON file.

In order to assign the general emotion for the post or comment, we chose the majority reaction. In addition, we did not consider the reaction *Like* because it is the most generic that users tend to use and it would be biased. Finally, we obtained three different corpus: Post corpus (PC) with 1071 posts, Comment corpus (CC) that contains 1036 responses to the posts, and Post and comment corpus (PCC) with the union of the PC and CC, that is, 2107 samples in total. Statistics of Facebook reactions in these corpora can be observed in Table I. For the experimentation, we randomly partitioned the corpora into equally sized for training and testing (50% train and 50% test).

TABLE I  
NUMBER OF DIFFERENT FACEBOOK REACTIONS IN PC AND CC

Facebook reaction	post	comment
Angry	747	171
Haha	225	359
Love	68	317
Wow	13	85
Sad	18	104
Total	1071	1036

### B. Machine learning approach

In order to evaluate the ML method, we chose scikit learn package<sup>9</sup> of python and we applied as baseline the SVM algorithm. We selected the SVM formulation, known as C-SVC, the value of the C parameter was 1.0 and the kernel chosen was the linear. As has been mentioned before, 50% of the corpora was used for training and the remaining 50% for testing. Each document was represented as a vector of unigrams using the TF-IDF weighting scheme, but previously the following preprocessing step was carried out. The documents were tokenized using NLTK TweetTokenizer<sup>10</sup>, stopwords were removed, stemming was performed using NLTK

<sup>9</sup><http://scikit-learn.org/stable/>

<sup>10</sup><http://www.nltk.org/api/nltk.tokenize.html>

Snowball stemmer for Spanish<sup>11</sup> and all letters were converted to lower-case.

### C. Lexicon-based approach

In order to evaluate the lexicon-based approach for emotion classification we followed two different methods. Firstly, we applied the NRC Spanish Emotion Lexicon [7] and secondly, we adapted the English WNA Lexicon [8] to Spanish.

The Emolex Spanish Emotion lexicon is a version of the NRC Emotion Lexicon (Emolex)<sup>12</sup> that was built by translating English emotional terms into Spanish using Google Translator. In order to classify emotions the documents were tokenized using NLTK TweetTokenizer, stopwords were removed, and all letters were converted to lower-case. The emotion labels (EL) of the terms present in the text are obtained by identifying the presence of these terms in the Emolex lexicon. We map these labels into Facebook reactions (Table II) and then assign 1 as confidence value (CV). Additionally, we identify the emojis present in the text using the faces of an emoji lexicon<sup>13</sup>, we map them into Facebook reactions and we assign 1 as CV. Once the EL and CV of each word/emoji have been obtained, the general emotion is calculated in the following way: For each EL identified, we add the CV of the tokens that belong to it and we assign the EL with the highest sum to the text. In the case of two or more emotions having the same sum of CV or no EL being detected, we assign the most frequent EL, previously computed in the corresponding training corpus.

TABLE II  
MAPPING OF THE GENERAL EMOTION OF EMOLEX TO FACEBOOK REACTIONS

Facebook reaction	Emolex emotion
Sad	Sadness
Angry	Anger
Wow	Surprise
Love	
Haha	Joy

On the other hand, WNA is a linguistic resource that was built starting from WordNet Domains [24] through the selection and labeling of the synsets representing affective concepts. In particular, one or more affective labels are assigned to a number of WordNet synsets. This resource is focused on English, but we tackle Spanish texts. Therefore, we need to develop a method for obtaining affect labels for Spanish terms and obtaining their corresponding synsets with the aim of discovering the associated emotion in WNA. For this, we use the lexical disambiguator Babelfy<sup>14</sup> [25] to obtain the corresponding *BabelNet synset id* for each term in the

<sup>11</sup><http://www.nltk.org/api/nltk.stem.html>

<sup>12</sup><http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

<sup>13</sup><https://github.com/erunion/emoji-lexicon>

<sup>14</sup>Babelfy is based on the BabelNet multilingual semantic network and performs disambiguation and entity linking.

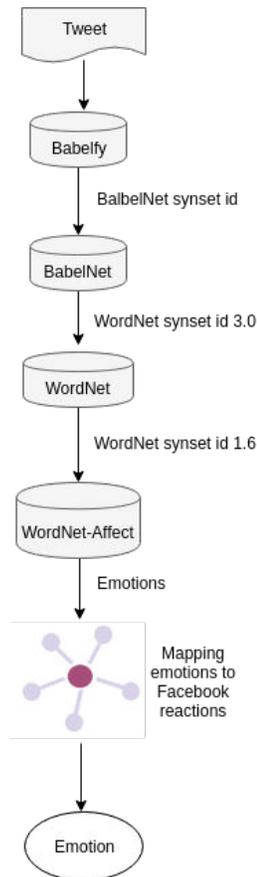


Fig. 3. Pipeline for adapting WNA to Spanish.

Spanish text. BalbelNet [26]<sup>15</sup> is a semantic network which is connected through synsets with WordNet. Using the BabelNet API we can obtain a correspondence between the *BalbelNet synset id* and the *WordNet synset id*. WNA includes a subset of appropriate synsets of WordNet 1.6 to represent affective concepts. However, the *WordNet synsets id* obtained with BabelNet API corresponds to the 3.0 version of WordNet. Therefore, we obtain the equivalent synset to the 3.0 version in the 1.6 version, and using the synset of the 1.6 version of WordNet we take the associated emotion and confident value from WNA. In Figure 3 it can be seen a diagram of the process described above.

WNA provides a set of emotional words organized in a tree. The leaf nodes represent specific emotions that are grouped into general categories (parent nodes). For example, *anger*, *hate* and *dislike* belong to the overall emotion *general-dislike*. However, WNA emotions are not the same as Facebook reactions. For this reason, each overall emotion of WNA has been mapped with Facebook reactions (Table III). The EL of the terms present in the text and their CV are obtained

<sup>15</sup>BabelNet is a multilingual encyclopedic dictionary, with lexicographic and encyclopedic coverage of terms in 271 languages, and a semantic network which connects concepts and named entities, made up of more than 13 million entries

using WNAffect python package<sup>16</sup>. These EL are mapped into Facebook reactions following Table III. Moreover, we identify the emojis present in the text using the faces of an emoji lexicon, we map them into Facebook reactions and we assign 1 as CV. Once the EL and CV of each word/emoji have been obtained, the overall emotion is computed as follows: For each EL identified, we add the CV of the tokens that belong to it and we assign the EL with the highest sum to the whole text. In the case of two or more emotions having the same sum of CV or no EL being detected, we assign the most frequent EL, previously computed in the corresponding training corpus.

TABLE III  
MAPPING OF THE GENERAL EMOTION OF WORDNET-AFFECT TO FACEBOOK REACTIONS

Facebook reaction	WNA emotion
Sad	apathy, neutral-unconcern, pensiveness, gravity, humility, compassion, despair, sadness
Angry	ambiguous-fear, ambiguous-expectation, ingratitude, shame, general-dislike
Wow	thing, ambiguous-agitation, surprise, positive-fear, positive-expectation, daze, anxiety, negative-fear
Love	gratitude, fearlessness, affection, self-pride, enthusiasm, positive-hope, calmness, love, liking
Haha	levity, joy

#### IV. RESULTS AND DISCUSSION

Table 2, 3 and 4 show the results obtained over the three datasets on the test sets (PC, CC and PCC). We can observe that the ML approach obtains better results as usual and expected. However, the results among the emotions are quite biased mainly due to the unbalance nature of the different corpora. Thus, it seems clear that we need to collect a corpus including enough samples for each emotion in order to our system can learn all of them.

Regarding the lexicon approaches, both methods achieve very low results. In addition, the majority emotion for each corpus determines the best result in this emotion because when our system does not recognize any emotion, we assign the majority one. For example, the *angry* emotion is the majority class in the PC and, thus, the results in this corpus are also biased to this emotion. For this reason, we performed other experiments, taking into account only the texts where our method recognizes one emotion, and neglecting the posts and comments with No-Emotion detected (WNA-NE and Emolex-NE). The results are very similar to the previous ones, but in this case it is interesting to note that WNA finds very few emotion words and, thus, the recall is very low, while with Emolex the texts detected with emotion almost reaches 80%. Specifically, WNA recognizes 67, 67 and 134 texts with emotions and Emolex

<sup>16</sup><https://github.com/clemtoy/WNAffect>



500, 327 and 827 for the PC, CC and PCC, respectively. Taking into account that the total number of texts in each test corpus is 534, 516 and 1050, the texts classified with emotion represent 13%, 13% and 13% for WNA and 94%, 63% and 79% for Emolex, respectively.

## V. CONCLUSION

In conclusion, emotion classification is a hard task that needs not only a deeper study but also specific linguistic resources in order to tackle the problem. Our next study will focus on collecting a larger and more balanced emotion corpus including others domains and studying the combination of different resources in order to generate a quality lexicon. In addition, we will compare the results obtained with English and Spanish corpora. Also, we plan to continue working on emotion recognition in Spanish because we have observed that the work in this language is very scarce, although it is the second most spoken language in the world and in the two main social networks: Facebook and Twitter.

It could be interesting to explore more affect lexicons because they provide prior information about the type and strength of emotion carried by each word of the text. Actually, in WASSA-2017 Shared Task on Emotion Intensity it was demonstrated that using features from affect lexicons is beneficial for emotion mining tasks [17].

Finally, this system could be used to measure the satisfaction of citizens with politicians that could be very useful, for example, to predict results in a political campaign based on the emotions transmitted by users in posts and comments of the Facebook page *Dignidad y Responsabilidad*<sup>17</sup>.

## ACKNOWLEDGEMENTS

This work has been partially supported by a grant from the Ministerio de Educación Cultura y Deporte (MECD - scholarship FPU014/00983), Fondo Europeo de Desarrollo Regional (FEDER) and REDES project (TIN2015-65136-C2-1-R) from the Spanish Government.

## REFERENCES

- [1] R. W. Picard *et al.*, “Affective computing,” 1995.
- [2] L. Canales, C. Strapparava, E. Boldrini, and P. Martinez-Barco, “Innovative semi-automatic methodology to annotate emotional corpora,” *PEOPLES 2016*, p. 91, 2016.
- [3] C. Cherry, S. M. Mohammad, and B. De Bruijn, “Binary classifiers and latent sequence models for emotion detection in suicide notes,” *Biomedical informatics insights*, vol. 5, no. Suppl 1, p. 147, 2012.
- [4] M. Dadvar, D. Trieschnigg, R. Ordelman, and F. de Jong, “Improving cyberbullying detection with user context.” in *ECIR*. Springer, 2013, pp. 693–696.
- [5] C. Suero Montero and J. Suhonen, “Emotion analysis meets learning analytics: online learner profiling beyond numerical data,” in *Proceedings of the 14th Koli calling international conference on computing education research*. ACM, 2014, pp. 165–169.
- [6] S. Stieglitz and L. Dang-Xuan, “Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior,” *Journal of management information systems*, vol. 29, no. 4, pp. 217–248, 2013.
- [7] S. M. Mohammad and P. D. Turney, “Crowdsourcing a word-emotion association lexicon,” vol. 29, no. 3, pp. 436–465, 2013.
- [8] C. Strapparava and A. Valitutti, “Wordnet affect: an affective extension of wordnet.” in *LREC*, vol. 4, 2004, pp. 1083–1086.
- [9] S. M. Mohammad, F. Bravo-Marquez, M. Salameh, and S. Kiritchenko, “Semeval-2018 Task 1: Affect in tweets,” in *Proceedings of International Workshop on Semantic Evaluation (SemEval-2018)*, New Orleans, LA, USA, 2018.
- [10] A. Celeghin, M. Diano, A. Bagnis, M. Viola, and M. Tamietto, “Basic emotions in human neuroscience: neuroimaging and beyond,” *Frontiers in Psychology*, vol. 8, p. 1432, 2017.
- [11] P. Ekman, “An argument for basic emotions,” *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.
- [12] A. Gholipour Shahraki, “Emotion mining from text,” 2015.
- [13] S. M. Mohammad, “# emotional tweets,” in *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2012, pp. 246–255.
- [14] C. Strapparava and R. Mihalcea, “Learning to identify emotions in text,” in *Proceedings of the 2008 ACM symposium on Applied computing*. ACM, 2008, pp. 1556–1560.
- [15] E. Cambria, “Affective computing and sentiment analysis,” *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102–107, 2016.
- [16] A. Yadollahi, A. G. Shahraki, and O. R. Zaiane, “Current state of text sentiment analysis from opinion to emotion mining,” *ACM Comput. Surv.*, vol. 50, no. 2, pp. 25:1–25:33, May 2017. [Online]. Available: <http://doi.acm.org/10.1145/3057270>
- [17] S. M. Mohammad and F. Bravo-Marquez, “Wassa-2017 shared task on emotion intensity,” *arXiv preprint arXiv:1708.03700*, 2017.

<sup>17</sup>[https://www.facebook.com/Dignidad y Responsabilidad/?ref=br\\_rs/](https://www.facebook.com/Dignidad y Responsabilidad/?ref=br_rs/)

TABLE IV  
RESULTS ON THE TEST SET OF PC

Approach	angry			haha			love			wow			sad		
	prec	rec	f	prec	rec	f	prec	rec	f	prec	rec	f	prec	rec	f
SVM	0.71	0.97	0.81	0.35	0.05	0.09	0	0	0	0	0	0	0	0	0
WNA	0.7	0.9	0.79	0.25	0.02	0.03	0.1	0.06	0.07	0.07	0.17	0.1	0	0	0
Emolex	0.67	0.67	0.67	0.16	0.07	0.1	0	0	0	0	0	0	0	0	0
WNA-NE	0.6	0.2	0.3	0.25	0.17	0.2	0.1	0.33	0.15	0.07	0.05	0.12	0	0	0
Emolex-NE	0.7	0.66	0.68	0.16	0.08	0.11	0	0	0	0	0	0	0	0	0

TABLE V  
RESULTS ON THE TEST SET OF CC

Approach	angry			haha			love			wow			sad		
	prec	rec	f	prec	rec	f	prec	rec	f	prec	rec	f	prec	rec	f
SVM	0.3	0.21	0.25	0.38	0.46	0.42	0.38	0.51	0.43	0	0	0	0.08	0.02	0.03
WNA	0.24	0.06	0.1	0.36	0.91	0.51	0.37	0.03	0.05	0	0	0	0.13	0.04	0.06
Emolex	0.16	0.21	0.18	0.4	0.71	0.51	0	0	0	0	0	0	0.16	0.25	0.2
WNA-NE	0.24	0.45	0.31	0.29	0.11	0.15	0.36	0.15	0.21	0	0	0	0.13	0.29	0.18
Emolex-NE	0.16	0.35	0.22	0.36	0.46	0.40	0	0	0	0	0	0	0.16	0.34	0.22

TABLE VI  
RESULTS ON THE TEST SET OF PCC

Approach	angry			haha			love			wow			sad		
	prec	rec	f	prec	rec	f	prec	rec	f	prec	rec	f	prec	rec	f
SVM	0.55	0.78	0.65	0.4	0.35	0.37	0.31	0.21	0.25	0	0	0	0	0	0
WNA	0.43	0.91	0.59	0.27	0.01	0.03	0.17	0.03	0.05	0.04	0.02	0.03	0.09	0.03	0.05
Emolex	0.43	0.66	0.52	0.27	0.13	0.18	0	0	0	0	0	0	0.07	0.21	0.10
WNA-NE	0.37	0.25	0.3	0.27	0.13	0.17	0.17	0.15	0.16	0.04	0.2	0.06	0.09	0.22	0.13
Emolex-NE	0.54	0.63	0.58	0.27	0.2	0.23	0	0	0	0	0	0	0.07	0.28	0.11

- [18] M. Purver and S. Battersby, "Experimenting with distant supervision for emotion classification," in *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 2012, pp. 482–491.
- [19] S. M. Mohammad and S. Kiritchenko, "Using hashtags to capture fine emotion categories from tweets," *Computational Intelligence*, vol. 31, no. 2, pp. 301–326, 2015.
- [20] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of computational science*, vol. 2, no. 1, pp. 1–8, 2011.
- [21] L. Stinson, "Facebook reactions, the totally redesigned like button, is here," *Wired*, February 24th, 2016.
- [22] F. Krebs, B. Lubascher, T. Moers, P. Schaap, and G. Spanakis, "Social emotion mining techniques for facebook posts reaction prediction," *arXiv preprint arXiv:1712.03249*, 2017.
- [23] C. Pool and M. Nissim, "Distant supervision for emotion detection using facebook reactions," *arXiv preprint arXiv:1611.02988*, 2016.
- [24] B. Magnini and G. Cavaglia, "Integrating subject field codes into wordnet," in *LREC*, 2000, pp. 1413–1418.
- [25] A. Moro, A. Raganato, and R. Navigli, "Entity linking meets word sense disambiguation: a unified approach," *Transactions of the Association for Computational Linguistics*, vol. 2, pp. 231–244, 2014.
- [26] R. Navigli and S. P. Ponzetto, "Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network," *Artificial Intelligence*, vol. 193, pp. 217–250, 2012.



# Visualización avanzada de datos de Twitter para su análisis como canal de comunicación

## Caso del mercado maduro de las bodegas con Denominación de Origen Calificada Rioja

Carmen Zarco

Departamento de Investigación de Mercados  
Universidad Internacional de La Rioja  
Logroño, ESPAÑA  
e-mail: [carmen.zarco@unir.net](mailto:carmen.zarco@unir.net)

Oscar Cordon

Instituto Andaluz Interuniversitario de Ciencia de Datos e  
Inteligencia Computacional (DaSCI)  
Universidad de Granada. Granada, ESPAÑA  
e-mail: [ocordon@decsai.ugr.es](mailto:ocordon@decsai.ugr.es)

**Resumen—** Las organizaciones se enfrentan a nuevas oportunidades y desafíos planteados por las redes sociales. Las empresas están evaluando el valor potencial de este fenómeno social para justificar la inversión de recursos en su adopción. Los Medios Sociales facilitan la distribución de información, referencias y reseñas electrónicas de boca en boca a través de redes sociales virtuales, foros de discusión, blogs y microblogs, incluido Twitter. El objetivo de este trabajo es identificar el modelo de comunicación desarrollado por las bodegas con Denominación de Origen Calificada Rioja en Twitter. Pretendemos conocer dicho modelo de comunicación desde una visión global y relacional, estudiando las características comunes y diferenciales de las distintas bodegas en su estrategia de comunicación en esta red social. Para ello, usaremos datos de presencia e impacto de las cuentas de Twitter de las bodegas y herramientas de visualización de datos avanzadas basadas en técnicas de Análisis de Redes Sociales.

**Palabras clave—** Twitter; Comunicación; Visualización de Datos; Análisis de Redes Sociales; Bodegas DO Ca Rioja.

### I. INTRODUCCION

Los Medios Sociales se han convertido en una importante plataforma de medios de comunicación que conecta a un tercio de la población [1]. Se ha producido una redistribución de los presupuestos destinados a la comunicación en las organizaciones a medida que los anunciantes se han alejado de los medios tradicionales invirtiendo mayores recursos en publicidad digital y redes sociales. Se estima el 58% de las marcas registradas en Twitter tienen más de 100.000 seguidores cada una [2]. El surgimiento de estas plataformas de redes sociales y su creciente adopción por parte de los clientes han precipitado un cambio de paradigma, alterando significativamente la forma en que los clientes se comunican e interactúan entre sí y con las empresas.

El uso de Twitter como plataforma digital global de interacción social no admite discusión alguna. Cuenta actualmente con 328 millones de usuarios activos al mes, registra 1.000 millones de visitas únicas mensuales a sitios web con acceso a la plataforma, genera más de 500 millones de tuits diarios, se encuentra disponible en más de 40 idiomas [3], [4].

Dentro de este contexto tecnológico, no es sorprendente que muchas marcas busquen nuevas e innovadoras formas de comunicarse con los consumidores. Las redes sociales como Twitter se han convertido en puntos de contacto con el cliente cada vez más populares, creando un nuevo foro para la interacción de los clientes con las marcas de vino como consecuencia de sus capacidades sociales y virales [5], [6].

En este trabajo analizaremos cómo empresas con una filosofía de tradición muy arraigada como las bodegas con Denominación de Origen Calificada (DOCa) Rioja utilizan Twitter como medio de comunicación. Pretendemos identificar el modelo desarrollado para comunicarse con sus seguidores y su impacto a través de su grado de presencia y actividad. Estudiaremos este modelo de comunicación desde una perspectiva global y relacional, descubriendo qué características comparten y diferencian a aquellas bodegas que hacen un mayor o menor uso del mismo. Para ello haremos uso de datos de presencia e impacto obtenidos de las cuentas de Twitter y de herramientas avanzadas de visualización de datos basadas en técnicas de Análisis de Redes Sociales (SNA) [7], que permitirán realizar un análisis detallado de las relaciones existentes entre las distintas organizaciones desde la perspectiva de su uso de Twitter como canal de comunicación. Nuestro objetivo es obtener representaciones visuales (*mapas*) del impacto y de las relaciones de similitud con respecto al posicionamiento de las distintas bodegas en Twitter que sean fácilmente interpretables por el analista considerando la distribución y la localización espacial de cada bodega en el mapa.

### II. MARCO TEÓRICO

El vino es un producto experiencial e implica inherentemente la socialización y construye comunidades alrededor del placer de compartir experiencias [8]. Las marcas españolas compiten para atraer y retener a los consumidores y muchas de ellas están adoptando las redes sociales para llegar a sus consumidores y comunicar su experiencia de marca, calidad y personalidad [9]. Los medios sociales y la comunicación a través de ellos mejoran ese elemento común de la apreciación y el consumo de vino creando comunidades por lo que es muy importante para aquellos involucrados en la industria del vino tener una presencia activa en las redes sociales [10].

La industria del vino está reconociendo progresivamente el papel cada vez más relevante que tienen las redes sociales como una herramienta apropiada y de gran valor para llegar a los consumidores. Las actividades de comunicación se encuentran en un estado de desarrollo evolutivo en el que continuamente surgen nuevas tendencias y las bodegas tienen la obligación de adaptarse al mismo si quieren tener éxito en sus campañas [11].

#### A. Twitter y sus usuarios

Twitter, fundada en 2006, se puede describir como una plataforma de microblogging y como una red social [12]. Tiene dos características especiales que lo diferencian de otro tipo de

plataformas similares: la limitación en el número de caracteres que se puede utilizar en cada mensaje (ahora a 280) y la forma en que se establecen las relaciones entre los usuarios. Al no requerirse una relación de amistad para interactuar, se pueden establecer relaciones con usuarios desconocidos facilitando el contacto con especialistas en temas específicos y el acceso a información actualizada relacionada con el área de interés [13].

### B. Perfil del consumidor de vino

El consumo de vino en España aumentó en 2016 después de varias décadas disminuyendo. Según el Observatorio Español del Mercado del Vino [14], la demanda de caldos registró un comportamiento positivo con una tasa de crecimiento del 2,5% al 4% hasta alcanzar un total de 9,8 millones de hectolitros. Aun así, y según los expertos, son datos muy por debajo de lo previsible para un país productor y de gran tradición vitivinícola.

Según el informe realizado por la consultora Nielsen para el Observatorio Español del Mercado del Vino [15], el perfil sociodemográfico del consumidor español de vino es fundamentalmente masculino, aunque en este momento el 43% de los consumidores son mujeres. En términos de edad, se pone de manifiesto el aumento de la proporción de consumidores aumenta con la edad. La penetración de consumidores de vino entre los más jóvenes (18 y 34 años) es del 52%, la cual aumenta en los grupos de edad entre 35 y 54 años y los superiores a 54 años hasta el 59% y 67%, respectivamente.

El vino como producto implica la socialización y construye comunidades alrededor el placer de compartir experiencias. Este intercambio y la construcción de la comunidad son procesos arraigados en redes sociales online que comparten principios similares a las relaciones humanas presenciales [16].

### C. Las bodegas y las políticas de comunicación

Investigadores y especialistas en marketing insisten en la importancia que tienen los medios sociales como un servicio de comunicación fácil y de bajo coste que proporciona una conexión inmediata con una gran cantidad de consumidores [17]. Sin embargo, estos expertos también sostienen que todavía queda un largo camino por recorrer por parte de la industria del vino para que se conviertan en una herramienta de comunicación y de marketing realmente eficiente [18]. Los expertos del sector vinícola han manifestado su inquietud por la ineficaz política de comunicación de la mayoría de bodegas españolas, pese a la notable presencia mediática que tienen los vinos [19].

Los medios tradicionales canalizan sus mensajes sobre vinos a través de la publicidad, las noticias que elaboran sus redactores y el trabajo de cada bodega para asegurarse espacios de difusión gratuita. Promueven así sus vinos, sus ideas y construyen un relato periodístico atractivo y explicativo de su actividad. Los consumidores desean participar ahora más que nunca en los procesos de comunicación. La pregunta ya no es sólo cómo llegar a ellos sino también cómo llegan ellos a las bodegas y cómo interactúan entre sí [20].

## III. METODOLOGÍA

Para llevar a cabo este análisis del uso de Twitter como canal de comunicación de las empresas con DOCa Rioja se han recopilado todos los datos relativos a su presencia en esta red de

microblogging. De las más de 590 bodegas que disponen de esta denominación, hemos comprobado que sólo 191 tienen presencia y actividad en Twitter. Uno de los hándicaps más importantes que se nos presentó fue la duplicidad de nombres de algunas bodegas que convergían en el mismo perfil dentro de la red social. De estas 191 bodegas se eliminaron tres que tenían dicha duplicidad de perfil en Twitter: Bodegas Berceo, Faustino Rivero y Bodegas Pujanza (es decir, dos bodegas distintas, generalmente de gran tamaño, empleaban la misma cuenta de Twitter en los tres casos), quedándonos finalmente con 188 bodegas/cuentas de Twitter. De cada cuenta se recopilaron datos de presencia reflejados en cuatro variables principales: *Tweets*, *Seguidores*, *Siguiendo* y *Me gusta*. La recogida de información se realizó entre los días 20 y 30 de diciembre de 2017, correspondientes a la campaña de Navidad, un periodo en el que las empresas vinícolas y los consumidores son muy activos en los medios sociales.

La base de nuestro estudio serán las técnicas de SNA para diseñar un mapa visual de las características de la comunicación en Twitter de estas bodegas con DOCa Rioja. El uso de técnicas de SNA [7], [21] ha demostrado su capacidad para generar visualizaciones esquemáticas de alta calidad de representaciones basadas en red en varios campos del conocimiento: psicología (para representar la estructura cognitiva de un tema) [22], modelado de sistemas (para analizar sistemas difusos) [23], depuración de software (para detección de bugs en sistemas multiagente) [24], optimización multiobjetivo (para visualizar la composición de las soluciones no dominadas, ayudando al decisor) [24] y cienciometría (para el análisis de grandes dominios científicos) [25], [26], entre otros.

En nuestro caso, en primer lugar diseñaremos una medida de similitud basada en los distintos indicadores de Twitter asociados a cada bodega. De esta forma, obtendremos una matriz relacional simétrica que determinará la similitud entre la estrategia de comunicación aplicada por cada par de bodegas, definiendo la estructura de la red social de bodegas correspondiente. Aplicaremos un algoritmo de poda de redes, una variante eficiente del algoritmo Pathfinder [27], sobre la matriz de similitud resultante con objeto de podar la red asociada, reduciendo las relaciones entre los modelos de comunicación definidos por las cuentas de Twitter de las bodegas para mantener solamente aquellas más significativas a nivel global. Para obtener el mapa visual, aplicaremos un algoritmo de dibujado (*layout*) de redes sociales basado en fuerzas, el algoritmo Kamada-Kawai [28], para situar los nodos de la red de modo que sus posiciones estén determinadas por las relaciones existentes.

### A. Medida de Similitud y Construcción de la Red Social

En este primer paso se consideran los datos de las cuatro variables: *Tweets*, *Seguidores*, *Siguiendo* y *Me Gusta*. Estos datos se han normalizado, con lo que cada bodega pasa a tener asociado un punto 4-dimensional en el espacio real  $[0,1]^4$ .

La medida de la similitud existente entre cada una de las bodegas de acuerdo a su política de comunicación en Twitter se define en función de la distancia Euclídea:

$$D(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^4 (x_i - y_i)^2} ,$$

lo que nos deja un cálculo como el siguiente:



$$D(\vec{x}, \vec{y}) = \sqrt{\begin{matrix} (Tweets_x - Tweets_y)^2 + \dots \\ (Siguiendo_x - Siguiendo_y)^2 + \dots \\ (Seguidores_x - Seguidores_y)^2 + \dots \\ (MeGusta_x - MeGusta_y)^2 + \dots \end{matrix}}$$

Volvemos a normalizar estos resultados localizando la distancia máxima y dividiendo las demás distancias por esta última. Después los convertimos en una medida de similitud invirtiendo esa distancia normalizada:

$$S(\vec{x}, \vec{y}) = 1 - D(\vec{x}, \vec{y}).$$

A continuación obtenemos la matriz que relaciona las distintas bodegas dependiendo del grado de similitud resultante de los datos gestionados. Es una matriz real cuadrada de dimensión 188×188, en la que no se considera la diagonal (la similitud entre una cuenta de Twitter y esa misma cuenta es máxima pero obviamente no es significativa) y que lógicamente es simétrica ( $S(\vec{x}, \vec{y}) = S(\vec{y}, \vec{x})$ ). Esa matriz define la red social asociada, una red ponderada no dirigida compuesta por 188 nodos, correspondientes a las bodegas analizadas, y  $\frac{188 \cdot 187}{2} = 17.578$  enlaces ponderados cuyo peso indica el valor de similitud entre las cuentas de Twitter de las dos bodegas que conecta.

### B. Poda de la Red Social

La visualización de redes sociales tiene varios problemas [29]: i) calidad (cuanto mayor es la red, más probabilidad hay de que existan errores en los datos; ii) complejidad: más variables, más detalle, más categorías; iii) velocidad (en muchos casos, queremos obtener resultados de nuestra red lo bastante rápido como para que pueda considerarse un proceso interactivo); y iv) análisis (¿qué orden de complejidad se requiere para los algoritmos?). En concreto, la gran dimensión que encontramos habitualmente en las redes sociales genera dificultades para obtener representaciones gráficas útiles para el análisis ya que puede provocar una sobrecarga de información para el analista, reduciendo la interpretabilidad de las representaciones gráficas que se pretende obtener.”

En la literatura existen tres alternativas predominantes en SNA para llevar a cabo esta tarea en redes ponderadas [30]:

- El primer método descarta los enlaces con pesos por debajo de un determinado umbral [31]. Este enfoque, aunque fácil de implementar, no considera la estructura intrínseca de la red subyacente. Por lo tanto, la red transformada puede no mostrar la naturaleza de la original, pudiendo obtener componentes conexas desconectadas que no se corresponden con la realidad.
- El segundo método se basa en la obtención de un árbol generador minimal a partir de la red [32]. Esto garantiza un número fijo y mínimo de enlaces (el número de nodos menos uno) pero no siempre refleja la información subyacente de un modo adecuado debido a la posibilidad de una poda excesiva.
- El tercer y último método establece restricciones en los caminos de la red y aquellos enlaces que no las satisfacen. El algoritmo Pathfinder [22] es el método usado habitualmente,

siendo conocido por sus propiedades matemáticas asociadas a la preservación de la desigualdad triangular en caminos de la red de longitud  $q$  (parámetro del algoritmo). El resultado de aplicar el método Pathfinder con parámetros  $q$  y  $r$  a una red social ponderada es una nueva red social ponderada podada denominada PFNET( $r, q$ ). Las distancias de los caminos se miden con una métrica paramétrica en la que  $r=1$  corresponde a la distancia Euclídea,  $r=2$  a la distancia cuadrática y  $r=\infty$  al peso del enlace de mayor peso. El algoritmo permite construir una secuencia de redes de complejidad decreciente aumentando el valor de  $q \in \{2, \dots, n-1\}$ . La PFNET( $r, q=1$ ) corresponde a la red original y la PFNET( $r, q=n-1$ ) está formada por el menor número posible de enlaces.

### C. Dibujado (Layout) de la Red Social

Una vez que se dispone de una PFNET o cualquier otro tipo de red social, existe una amplia batería de métodos que permiten su visualización de forma automática. Los algoritmos basados en fuerzas son los métodos más utilizados para dibujar estructuras de red en el área de las Ciencias de la Información [34]. Su propósito es ubicar los nodos de una red en un espacio bidimensional o tridimensional de forma que bien todos los enlaces o bien sean aproximadamente de igual longitud o bien se correspondan globalmente con las longitudes teóricas entre los nodos del grafo y que además se produzcan pocos cruces entre enlaces, tratando de obtener una visión lo más estética y agradable posible. Esta familia tiene a los algoritmos clásicos de Kamada-Kawai [28] y Fruchterman-Reingold [33] como sus métodos más representativos aunque existen propuestas recientes capaces de escalar hasta un millón de nodos [34].

Consideraremos el algoritmo Kamada-Kawai, que funciona bien con redes Pathfinder. Para enriquecer la visualización, usamos un código de colores para los nodos en el que el color azul refleja el valor más bajo y el rojo el valor más alto del indicador, pasando por el verde, el amarillo y el naranja que indican valores medios bajos, medios y medios altos.

## IV. VISUALIZACIÓN OBTENIDA Y ANÁLISIS DE RESULTADOS

### A. Generación del Mapa

La Figura 1 muestra la visualización resultante de aplicar la metodología de la sección III a los datos disponibles de las 188 bodegas. El parámetro de la distancia de Minkowski toma valor  $r=\infty$  ya que en estudios previos se ha comprobado que posibilita un mayor nivel de poda y por tanto una mejor visualización.

Hemos usado el algoritmo Force Atlas 2, una de las implementaciones del Kamada-Kawai de Gephi, una herramienta gráfica de libre distribución para análisis de redes muy extendida (<https://gephi.org/>). Se han usado los parámetros por defecto, cambiando el Escalado a 3.0 y considerando las opciones *Disuadir Hubs* y *Evitar el solapamiento*. Una vez obtenida la visualización de la red, se ha aplicado también el algoritmo *Ajuste de Etiquetas* de Gephi. La gradación de colores corresponde a una variable teórica *Posicionamiento en Twitter*, cuyo valor es equivalente a la media de los cuatro indicadores considerados para cada bodega.

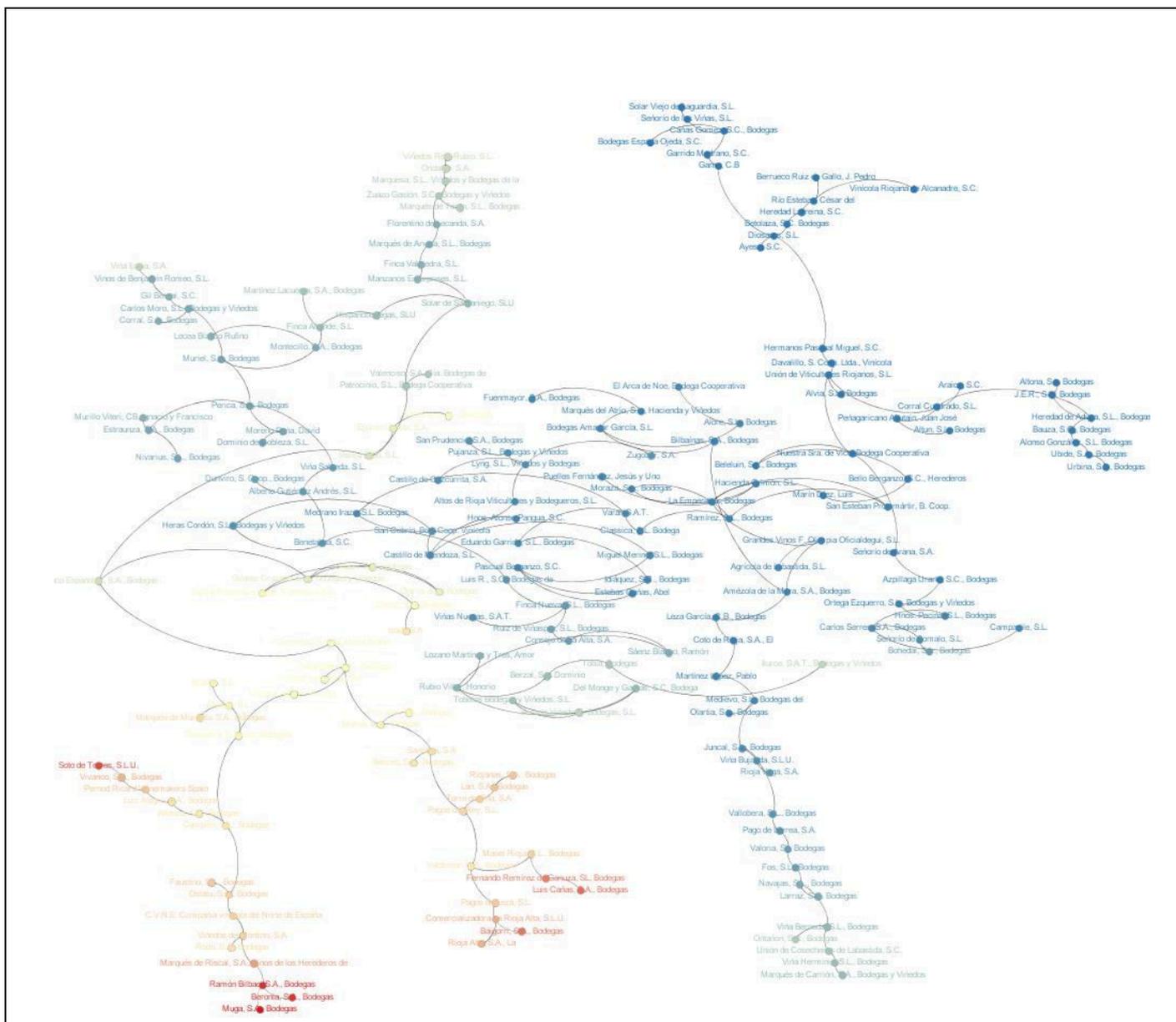


Fig. 1. Red PFNET( $r=\infty, q=n-1$ ) del modelo de comunicación de las bodegas con DOCa Rioja en Twitter (la escala de color de los nodos se basa en el posicionamiento global de las cuentas)

Esta representación visual obtenida es muy clara y permite identificar aspectos importantes de los datos analizados como las relaciones locales. Las distancias existentes entre los nodos, determinadas por el algoritmo de *layout* para hacerlas coincidir con las distancias globales de la red, representan claramente las similitudes y diferencias entre las estrategias de comunicación de las distintas bodegas. Se produce un efecto centro-periferia, habitual en este tipo de representaciones, en el que las bodegas situadas en la parte central tienen un comportamiento más similar entre sí y las posicionadas en los distintos extremos del mapa muestran políticas de comunicación muy diferenciadas con las del centro (es decir, valores significativamente distintos en los cuatro indicadores y, por tanto, en la medida de similitud) y, consecuentemente, diferentes también entre sí.

La gradación de colores nos permite observar claramente qué bodegas son las que tienen una estrategia de comunicación más activa en Twitter desde una perspectiva global de los cuatro indicadores (es decir, aquellas que presentan mayores valores en el agregado de los mismos y que por tanto tienen un tono más rojo en el mapa) y cuáles son las que la presentan menor actividad (menor valor agregado, tono más azul). Se observa fácilmente como las bodegas más cercanas entre sí presentan tonos de color (y por tanto estrategias de comunicación) similares, justificando el posicionamiento de los nodos en el mapa. Observamos como las bodegas con un modelo de comunicación más activo se sitúan en la parte inferior izquierda del mapa, siendo *Muga*, *Beronia*, *Ramón Bilbao*, *Soto de Torres*, *Baigorri*, *Luis Cañas* y *Fernando Ramírez de Ganuza* las más destacadas, por ese orden.



### B. Análisis del Mapa de Etiquetas de Posicionamiento Global

En esta sección pretendemos extraer conclusiones adicionales relacionadas con la distribución espacial de las bodegas en el mapa de modelo de comunicación en Twitter. En primer lugar se observa que en la parte superior del mapa se concentran principalmente las bodegas que apenas tienen actividad dentro de Twitter. Las bodegas que se encuentran en este grupo, como son, por ejemplo, *Solar Viejo de La Guardia*, *Señorío de las Viñas*, *Bodegas Espada Ojeda* y *Vinícola Riojana de Alcanadre*, reúnen características muy parecidas en lo que se refiere a la baja actividad en las cuatro variables analizadas y a la presencia de uno o varios ceros en las mismas. Como ejemplo, todas las bodegas citadas tienen valor cero en la variable *Me Gusta* junto con unos datos realmente bajos en las otras tres variables. Además, las dos primeras, situadas más en la periferia, presentan un cero también en *Tweets*. Nótese que el algoritmo de visualización separa claramente esta zona del resto del mapa a través de un enlace de gran longitud entre las bodegas *Hermanos Pascual Miguel* y *Diosares*, formándose un clúster con un comportamiento diferencial con respecto al centro del mapa que analizaremos a continuación.

Vemos cómo el mayor número de bodegas se concentran en el centro del mapa. Esto se debe a que la mayoría de las bodegas que utilizan Twitter tienen unos datos muy similares entre sí y con un comportamiento muy homogéneo. Como ejemplo tenemos las bodegas *Ramírez* y *La Emperatriz*, situadas en la parte más central del mapa y con valores relativamente bajos en todas las variables (*Tweets*=145, *Siguiendo*=258, *Seguidores*=309, *Me Gusta*=16 y *Tweets*=47, *Siguiendo*=176, *Seguidores*=365, *Me Gusta*=2, respectivamente), por lo que su política de comunicación en la red social se puede acabar diluyendo con el resto de empresas del sector vinícola. Este comportamiento nos confirma la inquietud que tienen las organizaciones con respecto a sus políticas de comunicación y el escaso uso que hacen de las nuevas tecnologías para realizar una política de comunicación potente, real e innovadora. Según nos acercamos tanto a la zona inferior como sobre todo a la zona izquierda del mapa, la tonalidad de color va aumentando, pasando al verde que indica un mayor valor en los indicadores. La utilidad de la información proporcionada por los enlaces a nivel local puede observarse en la bodega *Ilurce*, que aunque está situada en el centro de la parte inferior está enlazada con la rama situada a la izquierda formada por el grupo de bodegas con tonalidad verde (*Tobia*, *Berzal* y *Rubio Villar* en una rama, *Del Monge* y *Garbati* en otra, etc.). En concreto, los valores asociados a esta bodega son *Tweets*=1.418, *Siguiendo*=1.514, *Seguidores*=2.998 y *Me Gusta*=180, valores claramente superiores a los de la parte central.

Por último, vemos cómo se agrupan las bodegas en la parte baja en tres ramificaciones bien diferenciadas, cada una con una naturaleza distinta. En el grupo de la izquierda se aglutinan las bodegas que tienen un alto número de seguidores como *Ramón Bilbao* con 24.800 seguidores, *Muga* con 21.700 seguidores y *Faustino* con 10.400. Las tonalidades rojas y amarillas de esta zona nos indican que los valores del resto de variables también son altos, encontrándonos con el conjunto de bodegas con un uso más activo de la comunicación en Twitter.

En cambio, el grupo del centro incluye bodegas con los valores más altos en *Tweets*, donde destacan las bodegas

*Fernando Remírez de Ganuza*, *Luis Cañas* y *Baigorri*, con 13.000, 16.000 y 16.900 tweets respectivamente. De nuevo, las tonalidades rojas y amarillas vuelven a señalar una fuerte actividad, aunque con un modelo de comunicación distinto.

Finalmente, el tercer grupo situado a la derecha de los tres aquí analizados agrupa bodegas que tienen unos datos relativamente altos en *Tweets* y *Seguidores* pero que presentan una menor actividad desde un punto de vista global (tonos verdes, correspondientes a valores medios bajos). Dentro de este grupo destacan bodegas como *Marqués de Carrión*, *Unión de Cosecheros de Labástida*, *Ontañón* y *Viña Herminia*. Todas ellas superan holgadamente los 1.500 Tweets y cuentan con un total de *Seguidores* que oscila entre los 2.500 y 4.000, lo que da a entender que se trata de bodegas que tienen una actividad continua en comparación a la mayoría localizada en la parte central, aunque no tan intensa como pueden tener las bodegas que se sitúan más a la izquierda (en las otras dos variables, *Siguiendo* y *Me gusta*, los valores medios están en torno a 400).

### CONCLUSIONES

En este trabajo hemos analizado datos cuantitativos recopilados de la red social Twitter pertenecientes a las bodegas con DOCa Rioja. Con ello se pretendía conocer el modelo de comunicación que dichas empresas aplican en esta plataforma mediante el dibujo de un mapa que proporcionase una visualización estética de la distribución de las bodegas dependiendo del valor de las variables consideradas. El estudio realizado ha permitido desarrollar una aproximación para conocer de cerca la presencia de las bodegas y valorar el estado en el que se encuentran sus políticas de comunicación digitales, concretamente dentro de Twitter.

Una de las principales conclusiones es la escasa presencia que tienen estas organizaciones en la red social, ya que solamente 188 bodegas de un total de 591 tienen perfil en Twitter (aproximadamente un 32%) y no todas ellas tienen una actividad de comunicación activa y directa. Este hecho colisiona con el perfil del consumidor de vino y el perfil del usuario de Twitter cuyas características principales son coincidentes (hombres con edades comprendidas entre los 35 y 54 años), lo que sería razón de más para llevar a cabo una sólida política de comunicación en la red ya que es aquí donde se concentra el público objetivo al que se dirige el mercado vinícola español.

A través del mapa obtenido, hemos podido comprobar que las bodegas que acumulan un número mayor de seguidores son aquellas que tienen un nombre de marca consolidado y reconocido. No obstante ese dato no quiere decir que dichas bodegas mantengan una comunicación constante con sus seguidores a través de la red social pues en número de tweets las bodegas destacadas no poseían un alto número de notificaciones, a excepción de *Muga*, que ha demostrado tener un alto número de seguidores además de mantener una constante actividad dentro de la red de microblogging.

Por otro lado, las bodegas que sí han propiciado una acción por parte de sus seguidores para que marquen un “*Me Gusta*” son las que aglutinan más seguidores pero no son las que emiten más tweets. Esa característica se debe a la fidelidad que mantienen

estos seguidores que no permanecen en estado pasivo ante las notificaciones de las bodegas.

Todas estas interpretaciones nos dan a entender que el mercado vinícola español, y concretamente las bodegas con DOCa Rioja, siguen presentando determinadas carencias en sus comunicaciones en el mundo digital. La labor de las bodegas es hacer un buen vino y asignarle un precio atractivo pero también comunicarlo eficazmente a sus grupos de interés tanto presentes como futuros, a sus consumidores y al público en general, tanto a través de los medios de comunicación tradicionales como en los canales y plataformas de redes sociales. Las bodegas deben plantearse cuánto afecta una buena comunicación a la movilización del consumidor de vinos, cómo la comunicación crea y realza la imagen de las marcas de las bodegas, y cómo ese proceso induce a la compra. No hay que olvidar que el vino, pese a la bajada de ventas, tiene la ventaja de que es el producto alimenticio que más protagonismo tiene en los medios y es objeto de un caudal abundante de comentarios e informaciones que despiertan un notable interés entre usuarios y consumidores.

#### Agradecimientos

Este trabajo está soportado por el Ministerio de Economía y Competitividad de España bajo el proyecto NEWSOCO (ref. TIN2015-67661-P), incluyendo fondos FEDER.

#### BIBLIOGRAFÍA

- [1] K. Nelson-Field, E. Riebe y B. Sharp, «What's Not to "Like"?: Can a Facebook Fan Base Give a Brand The Advertising Reach It Needs?», *Journal of Advertising Research*, vol. 2, nº 52, pp. 262-269, 2012.
- [2] K. Smith, «Brandwatch», 2017. [En línea]. Available: <https://www.brandwatch.com>. [Último acceso: 4 Enero 2018].
- [3] S. Percastre-Mendizábal, C. Pont-Sorribes y L. Codina, «Propuesta de diseño muestral para el análisis de Twitter en comunicación política», *El profesional de la información*, vol. 26, nº 4, pp. 579-588, 2017.
- [4] Twitter, «Twitter», 2017. [En línea]. Available: <http://www.twitter.com>. [Último acceso: 2 Enero 2018].
- [5] N. Barber, T. Dodd y R. Ghiselli, «Capturing the younger wine consumer», *Journal of wine research*, vol. 19, nº 2, pp. 123-141, 2008.
- [6] M. Buelarca y S. Buelarca, «Twitter: a viable marketing tool for SMEs?», *Global Business & Management Research*, vol. 2, nº 4, pp. 296-309, 2010.
- [7] S. Wasserman y K. Faust, *Social Network Analysis: Methods And Applications (Structural Analysis in the Social Sciences)*, Cambridge: Cambridge University Press, 1994.
- [8] J. Bruwer y K. Alant, «The hedonic nature of wine tourism consumption: an experiential view», *International Journal of Wine Business Research*, vol. 21, nº 3, pp. 235-257, 2009.
- [9] Vinography, «Social Media and the Wine Industry: A New Era», 2014 Septiembre. [En línea]. Available: [http://www.vinography.com/archives/2012/02/social\\_media\\_and\\_the\\_wine\\_indu.html](http://www.vinography.com/archives/2012/02/social_media_and_the_wine_indu.html). [Último acceso: Enero 2018].
- [10] D. Wilson y S. Quinton, «Let's talk about wine: does Twitter have value?», *International Journal of Wine Business Research*, vol. 24, nº 4, pp. 271-286, 2012.
- [11] M. Stelzner, «Social Media Marketing Industry Report», *Social media examiner*, pp. 1-52, 2014.
- [12] M. Thelwall, K. Buckley y G. Paltoglou, «Sentiment in Twitter events», *Journal of the Association for Information Science and Technology*, vol. 62, nº 2, pp. 406-418, 2011.
- [13] H. Rodríguez, B. Restrepo y L. Fernando, «Conocimientos y uso del twitter por parte de estudiantes de educación superior», *Sophia*, vol. 11, nº 1, pp. 44-52, 2015.
- [14] Observatorio Español del Mercado del Vino, «OEMV», 2017. [En línea]. Available: <http://www.oemv.es/esp/-oemv.php>. [Último acceso: 15 Enero 2018].
- [15] Nielsen, «Caracterización del consumidor español de vino» 2017.
- [16] C. Cochran, «Wineries embracing social networking», *San Francisco Chronicle*, p. 7, 2010.
- [17] R. Dolan, J. Conduit, J. Fahy y S. Goodman, «Facebook for wine brands: an analysis of strategies for Facebook posts and user engagement actions», de *9th Academy of Wine Business*, Adelaide, 2016.
- [18] R. Fuentes Fernández, R. Vriesekoop y B. Urbano, «Social media as a means to access millennial wine consumers», *International Journal of Wine Business Research*, vol. 29, nº 3, pp. 269-284, 2017.
- [19] R. Castro Galiana, 27 Mayo 2014. [En línea]. Available: <http://castrogaliana.com/las-bodegas-los-vinos-y-las-politicas-de-comunicacion/>. [Último acceso: 2 Febrero 2018].
- [20] J. Bruwer y G. Wood, «The Australian online wine buying consumer: motivation and behaviour perspectives», *Journal of Wine Research*, vol. 6, nº 3, pp. 193-211, 2005.
- [21] J. Scott, *Social Network Analysis: A Handbook*, Sage Publications, 2000.
- [22] D. Dearholt y R. Schvaneveldt, «Properties of pathfinder networks», de *Pathfinder associative networks: Studies in knowledge organization*, 1990, p. 1-30.
- [23] D. Pancho, J. Alonso, O. Cerdón, A. Quirin y L. Magdalena, «FINGRAMS: visual representations of fuzzy rule-based inference for expert analysis of comprehensibility», *IEEE Transactions on Fuzzy Systems*, vol. 21, nº 6, pp. 1133-1149, 2013.
- [24] K. Trawinski, M. Chica, D. Pancho, S. Damas y O. Cerdón, «moGrams: A Network-based Methodology for Visualizing the Set of Non-dominated Solutions in Multiobjective Optimization», *IEEE Transactions on Cybernetics*, vol. 48, nº 2, pp. 474-485, 2018.
- [25] B. Vargas-Quesada y F. Moya-Anegón, *Visualizing the structure of science*, Springer Science & Business Media, 2007.
- [26] F. Moya-Anegón, B. Vargas-Quesada, Z. Chinchilla-Rodríguez, E. Corera-Álvarez, F. Muñoz-Fernández y V. Herrero-Solana, «Visualizing the marrow of science», *Journal of the Association for Information Science and Technology*, vol. 58, nº 14, pp. 2167-2179, 2007.
- [27] A. Quirin, O. Cerdón, J. Santamaría, B. Vargas-Quesada y F. Moya-Anegón, «A New Variant of the Pathfinder Algorithm to Generate Large Visual Science Maps in Cubic Time», *Information Processing and Management*, vol. 44, nº 4, p. 1611-1623, 2008.
- [28] T. Kamada y S. Kawai, «An algorithm for drawing general undirected graphs», *Information processing letters*, vol. 31, nº 1, pp. 7-15, 1989.
- [29] A. Unwin, M. Theus y H. Hofmann, *Graphics of large datasets: visualizing a million*, Springer Science & Business Media., 2008.
- [30] C. Chen y S. Morris, «Visualizing evolving networks: Minimum spanning trees versus pathfinder networks», de *Information Visualization*, 2003.
- [31] M. Zizi y M. Beaudouin-Lafon, «Accessing hyperdocuments through interactive dynamic maps», de *Proceedings of the 1994 ACM European conference on Hypermedia technology*, 1994.
- [32] S. Noel, C. Chu y V. Raghavan, «Visualization of document co-citation counts», de *IEEE Symposium on Information Visualisation*, 2002.
- [33] T. Fruchterman y E. Reingold, «Graph drawing by force-directed placement», *Software: Practice and experience*, vol. 21, nº 11, pp. 1129-1164, 1991.
- [34] S. Kobourov, «Force-directed drawing algorithms», de *Handbook of Graph Drawing and Visualization*, R. Tamassia, Ed., 2012.



# Identifying conflicts of interest of reviewers using authors bibliometric networks

V. Duarte-Martínez

Fac. de Ing. en Electricidad y Computación  
Escuela Superior Politécnica del Litoral  
Guayaquil, Ecuador  
vealduar@espol.edu.ec

A.G. López Herrera

Dept. Computer Science and  
Artificial Intelligence  
University of Granada, Granada, Spain  
lopez-herrera@decsai.ugr.es

M.J. Cobo

Dept. Computer Science and Engineering  
University of Cádiz, Cádiz, Spain  
manueljesus.cobo@uca.es

**Abstract**—The process of peer review is a very important task that must be followed by a scientific document to be approved and then published in a scientific journal. Being such a delicate process, it is required that the elected reviewers guarantee the validity and quality of the submitted work, without influencing in its decision any type of link that may exist between the authors of a scientific document and the reviewers. However, many times this goes unnoticed and the reviewers are selected only according to their experience in a given field.

This contribution proposes an approach to do recommendations based on the design of bibliometric networks such as: co-authorship networks, author co-citation networks and direct link-citation networks. In this way, we seek to discover relationships between reviewers and authors of scientific documents in order to suggest appropriate reviewers, thus avoiding the recurring problem of conflicts of interest.

**Index Terms**—Reviewers selection, bibliographic networks, co-author, author co-citation, direct link-citation

## I. INTRODUCTION

Much of the scientific job is based on the search for answers. To accomplish this, it must be performed many researches and experiments and their results are analyzed, discussed and disseminated, generally made through publications. However, in order that scientific documents become published must go through an assessment process, which includes peer review, to ensure the quality of the document and above all the validity of what has been investigated and is exposed in the document.

Scholarly peer review can be viewed as the central part of the publishing workflow, being considered necessary to ensure the quality of research journals. In fact, it must determine whether a manuscript should be accepted or rejected. This method leaves the work open to scrutiny, and often to the annotation or modification, by authors of similar or superior to the author's range. Since peer review is based in the criteria of experts, it needs of a set of experts in a particular research field (i.e. the research field of the manuscript to be reviewed) who are qualified and able to carry out an impartial review.

During this process, the role of referees is advisory. Moreover, in the scientific literature, the referees do not act as a group; not communicate with each other. Generally, they are not aware of the identity or the reports of their colleagues. In general, it is not necessary to achieve consensus. Therefore, the group dynamics are very different to a jury. Sometimes

their opinion is not unanimous. In such cases can be applied different options to make a decision.

Usually, it is desired to choose reviewers who are not close to the authors and also have no links with them. These are expected to inform publishers about potential conflicts of interest for evaluation. Some publishers or publications ask authors a list of potential arbitrators, as well as people they deem inappropriate to arbitrate their work. This is necessary, especially, when the subject of a work is so specialized that publishers cannot locate specialists in the field by themselves.

Therefore, the selection of impartial reviewers is a difficult and daunting task, since reviewers could have some conflicts of interest with the authors of the manuscript. For example, a reviewer could be a recent co-author, could be cited by the manuscript or could be usually cited by the author, which would mean, in a way that the reviewer might be interested in the publication of this manuscript. This and other issues could bias the reviewer criteria.

Usually, the selection of reviewers is based on the field of expertise of the reviewers, and the inherent and hidden social network of them is not taken into account in the process. So, in this contribution a novel approach based on bibliometric networks [1], to uncover hidden relationships between reviewers and authors, is presented. In particular, we focus on bibliometric networks based on the social aspects, i.e., authors. To do that, co-author, author co-citation and author direct link citation networks are used to build a global social bibliometric network.

## II. BIBLIOGRAPHIC NETWORKS

Networks provide an interesting abstraction of a variety of complex systems [2], such as, social networks or biological networks. Moreover, scientific knowledge could be understood as a complex system, where the network structure is frequently used to model the interaction of scientific actors (authors, journals, keywords, references, etc.).

On the other hand, bibliometric is dedicated to the analysis of the scientific output. Formally, Bibliometrics are a set of methods, which can be used to analyze academic literature quantitatively and its changes over time [3]. It is an important tool for assessing and analyzing the academic research output contributing to the progress of science in many different ways.

There are two main bibliometric methods for exploring a research field: performance analysis and science mapping [4]. While performance analysis aims to evaluate the citation impact of the scientific production of different scientific actors [5], science mapping aims at displaying the conceptual, social or intellectual structure of scientific research and its evolution and dynamical aspects [6] [7].

As above mentioned, the scientific knowledge can be represented as a graph. In this sense, science mapping analysis uses bibliographic networks [1] in order to represent the different relationships among scientific actors. Depending on the kind of aspects that will be represented, three kind of networks could be identified:

- Collaboration networks are used to show how authors or institutions relate to others in the field of scientific research. The most common kind of collaboration network are co-author networks. With this type of network can be discovered, for example, groups of regular authors, influence authors, hidden communities of authors, relevant institutions in a specific research field, etc. [8]
- Conceptual networks [9] represent relations between concepts or words in a set of publications. That means that, for instance, words which appear together in a document, will be related in a network. It is also known as co-word network. This type could be used to understand the topics covered by a research field [10], to define which are the most important and the most recent issues. It could also help in the study of the evolution of subjects over time and it could give good impression of cognitive relations between different research groups.
- Publication Citation networks [11] [12] show relationships between nodes which represent publications, while the edges can have different interpretations depending on the network type (co-citation, bibliographic coupling or direct link).

In this contribution we focus on the social bibliographic network which have become a very important topic of study that has attracted the attention of researchers because these kinds of networks represent a good prototype of complex evolving system to explore, where the network is constantly expanding with the occurrence of new authors of scientific documents, or new links between existing authors.

These networks are studied with two main aims. Mainly, to evaluate the collaboration status of scientific disciplines, that is related with bibliometric mapping, assisting in discovering of structural and dynamic aspects of scientific research, to better understand the organization of scientific fields and their evolution [6]. And also, to examine the performance of social structure of collaboration networks, using techniques of social networks analysis [13].

In Bibliometrics, social bibliographic networks are analyzed using indicators to measure and extract the inherent structure to a set of publications; but also these are studied as social networks to explore interactions between actors who may be researchers, groups, institutions, etc. that collaborate in the development of scientific papers on research topics. These are

usually weighted to represent the strength of different relations between actors.

Social bibliographic networks might reveal substantial knowledge such as: researcher communities, professional interactions between scientists, central nodes that act as hubs, leaders or gatekeepers; highly connected groups; and patterns of interactions between groups [14]. Very large networks can be assembled in this way.

### III. METHODOLOGY

As above mentioned, different bibliographic networks have been proposed based on co-occurrence, coupling and direct link. Moreover, depending of the selected unit of analysis, different aspects could be uncovered. The approach presented in this contribution focuses on the social aspects.

Particularly, in a research document the unit of analysis author can be found in two different parts. The former is the authors list, being the authors who write the research document. The latter is the author-references list, which are the authors cited by the research document (the intellectual social base). Using these two types of author lists, three kinds of bibliometric networks can be built: co-author, author co-citation and author direct link citation.

- Co-author networks, are those containing all authors of a document and related to each other, i.e. two scientists are considered connected if they have coauthored one or more papers together. Such networks reflect groups of researchers in scientific fields.
- Author co-citation networks, are those which relate all authors referenced in a document to each other. This kind of networks shows author groups which are referenced together in the same papers.
- Author direct citation networks, are those that express the direct relationship between the author of a document and the cited author, that is, it shows the authors cited by other authors. This information shows a very important social factor that may allow discover which are the inspirations of a creator author, and the bases that allows the author to develop his/her research.

These networks show us the connections that may arise between two researchers, according to the nature of their link, and represent the main study object in this document, for example, let suppose a set of three research documents where each document is composed of a set of authors, and a set of authors in the references list as it is shown in Table I. Thus, the three social bibliometric networks built from this example corpus are shown in Figure 1.

TABLE I  
EXAMPLE CORPUS.

Document	Authors	Authors in reference
Document 1	a1, a2, a3	a2, a4
Document 2	a2, a3, a4	a1, a3, a5
Document 3	a1, a3, a5	a2, a4, a6

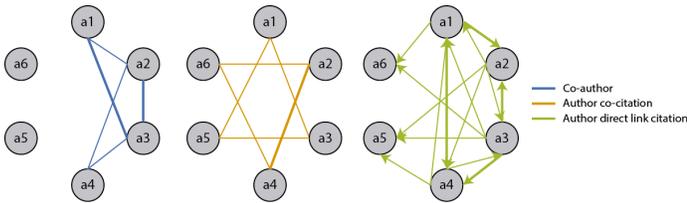


Fig. 1. Example of bibliometric network.

We should point out that an arc between two nodes represents a bibliometric relation between these authors. Also, the thickness of the arcs represents the raw co-occurrence or number of direct links depending on the kind of relationships.

Once the three kinds of bibliometric networks have been built, three levels of social relationships could be assessed:

- Level 1: authors who have only one kind of social relationship. At this level we just focus in one kind of relationship, and the remaining ones are not taken into account.
- Level 2: authors who have two kinds of social relationships. In this level, two authors could 1) collaborate together in a document and be co-cited, or 2) collaborate together in a document and have a direct link citation relationship between them, or 3) be co-cited and have a direct link citation between them.
- Level 3: authors who have the three kinds of social relationships. In this case, two authors collaborate together in (at least) a paper, are co-cited and also one of them cites to the other.

The above mentioned levels of relationships can be used to achieve the grade of conflict of interests between researchers. Therefore, from a set of bibliographic records downloaded from a bibliographic database (e.g. WoS or Scopus), a whole network containing the three types of relationship can be built.

Thus, from a set of documents related with a research field, a set of reviewers candidates could be achieved based on their bibliographic social relationships in order to avoid possible biases. To do that, the journal office should look at the social bibliometric relationships between the authors of the submitted manuscript and the reviewer community / panel in order to avoid biased reviewers. Furthermore, based on the three kinds of levels of social relationships, we propose an algorithm to identify different levels of bias prevention that could be configured. It is described as follows:

Having a graph  $G$  with a set of nodes  $V$  and edges between nodes  $V_i, V_j \in G$ . It is assumed that  $V_i$  is a document author while  $V_j$  is a reviewer candidate. Therefore, we could assess levels using the following inclusion criterion: **For each edge**  $(V_i, V_j)$  where  $V_i \neq V_j$ , level of bias prevention will be **Level 1** if  $V_i$  co-cited  $V_j$  or  $V_i$  co-authored  $V_j$  or  $V_i$  direct link cited  $V_j$ . Otherwise, level of bias prevention will be **Level 2** if  $(V_i$  !co-cited  $V_j$  and  $V_i$  !co-authored  $V_j)$  or  $(V_i$  !co-authored  $V_j$  and  $V_i$  !direct link cited  $V_j)$  or  $(V_i$  !co-cited  $V_j$  and  $V_i$  !direct link cited  $V_j)$ . In any another way, level of bias prevention will

be **Level 3** if  $V_i$  !co-cited  $V_j$  and  $V_i$  !co-authored  $V_j$  and  $V_i$  !direct link cited  $V_j$ .

The Level 3 is the most restrictive one. That means that the authors of the submitted manuscript and the reviewer candidate never appear together (neither in the author list, neither in the author-reference list) in any paper, in spite of, they investigate in the same research area. The Level 3 suggests no bias or conflicts of interest among authors and reviewers candidates, those last are supposed impartial. In Figure 2, it is shown a Venn diagram to explain the intersection of these three levels where each type of relationship is represented as a set and the levels are: The level 1, the less least restrictive, that covers every relationship; the Level 2, which allows the existence of two types of relationships at the same time; and the Level 3, which does not allow any kind of relationship.



Fig. 2. Different levels of bias prevention.

Thus, the level of rigor to be established depends on the editor's criteria; however, here is proposed this methodology that aims to deal with the problem of conflict of interest in scientific collaboration.

#### IV. CASE STUDY

With the aim of showing the applicability of this approach, a case of study have been built. It is described with details as follows.

Initially, we have collected data from Scopus database. For this, we have chosen 3891 research papers about Cloud Computing thematic only because this is a hot topic in Computer Sciences field.

Then, we have had used preprocessing techniques to decompose bibliographic data in order to extract specific data about principal authors and referenced authors. Afterwards, we have been able to build different bibliometric networks mentioned previously, that is, co-author, author co-citation and direct link citation networks. In these networks, we can explore

through their relationships and we can realize, as is logical, that an author related to another directly would not be a good candidate to be selected as a reviewer of a scientific document of him, but through these relationships you can navigate to find other researchers who have worked in the same subject but not related to the main author. The more levels of relationship are excluded, the more reliable it can be assured that there will be no conflicts of interest, obviously this can be assured from the point of view that encompasses scientific work.

In this analysis, the objective has been to find relationships of the type: *Friend of my friend who is not my friend*, taken to the field of research, Understanding that, if two researchers A and B have worked together, as collaborators of a scientific document, then from A and from B you can find possible reviewers candidates that are not related to A and/or B but with other researchers who have worked with authors A and B. In this way it is sought to avoid choosing reviewers of scientific articles that could have conflicts of interest with the authors of an article sent to a specific journal. The journal office, after receive the original manuscript, could use this analysis in order to select and recommend the best reviewers within the research area of the article, along with avoiding of conflict of interest, that is, authors which have no direct relationship with the authors of the submitted document (not cited author, or not collaborator).

To demonstrate this approach, we study each type of network that is generated from an author. For example, we have chosen, at random, an author of scientific documents dealing with the subject of this case study (Cloud Computing), the researcher Zheng Q. (Zheng, Qinghua). From this, co-author network, co-citation network and direct link citation network have been generated.

In Figure 3, we can see the co-author network built from one author (Zheng Q.) where the principal researcher is in purple node. In addition, the authors with whom he has shared co-authorship in scientific documents (Li J., Zhang J., Li X. and Li R.) are represented with orange nodes. From each author that have worked with Zheng Q. we have gotten all the recommended authors which are represented with green nodes. Recommended researchers are those who have worked with others who have co-authored scientific documents with Zheng Q., but have not worked directly with him. This would be the minimum level of connection that should be considered, although we could explore more in networks until we can find even more indirect relationships. In Table II, the list of recommended authors can be seen. We have pruned the networks built to show only 15 recommended authors thinking of simplifying the visualization of the results in this example.

We have also generated the co-citation network for the author, this means that the network has been built with all the relationships defined by the authors referenced in the scientific documents written by Zheng Q. This network can be seen in Figure 4, where node in purple color represents the principal author, orange nodes represents authors referenced in papers written by Zheng Q. and green nodes represents recommended authors. Summarized data can be seen in Table III.

TABLE II  
TABLE OF RECOMMENDATIONS FROM CO-AUTHOR NETWORKS.

Recommended Author
Cao Y.
Chen L.-X.
He T.-Q.
Lee S.-Y.
Li B.
Li C.
Liao S.
Su D.
Su G.
Wong D.W.S.
Xu P.
Xue W.
Yang Y.
Yu N.
Zhou F.-F.

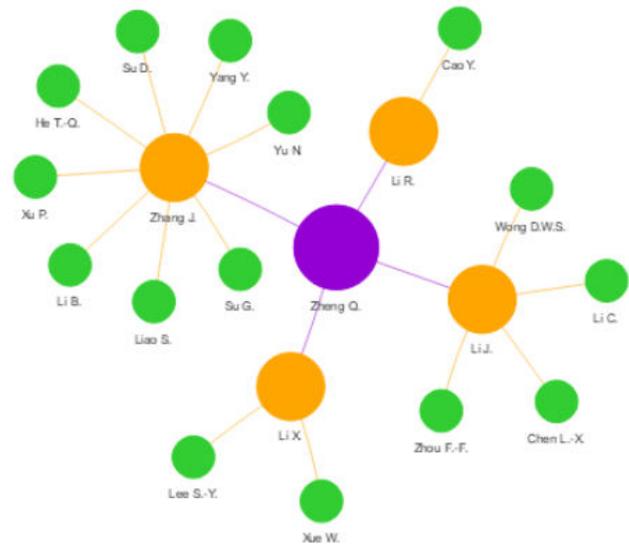


Fig. 3. Recommendations from co-author graph.

And finally, we have built an author direct link citation network. It models the relationships between the author and all cited authors in reference section from a scientific document. The Figure 5, shows the built graph where principal author, Zheng Q. is shown in purple node, referenced authors are shown in orange nodes and recommended authors are shown in green color. The list of recommended author is shown in the Table IV.

With all these recommended authors, we can identify who are those that fit in each level of bias prevention, according to what is established in section III. However, since the complete networks are very extensive (especially the co-citation networks that could contain thousands of relationships) we have summarized the results obtained. In Table V, it is shown some scientific authors that belong to level 1, level 2 and level 3 respectively. This means that, for example. the author Cheng



TABLE III  
TABLE OF RECOMMENDATIONS BASED ON AUTHOR CO-CITATION NETWORKS.

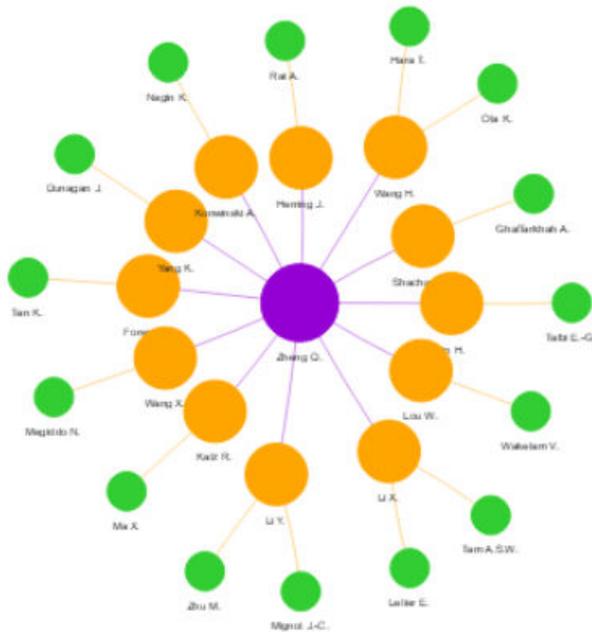


Fig. 4. Recommendations from author co-citation graph.

Recommended Author
Dunagan J.
Ghaffarkhah A.
Hara T.
Letier E.
Ma X.
Megiddo N.
Mignot J.-C.
Nagin K.
Ota K.
Rai A.
Talbi E.-G.
Tam A.S.W.
Tan K.
Wakelam V.
Zhu M.

TABLE IV  
TABLE OF RECOMMENDATIONS BASED ON DIRECT LINK CITATION NETWORKS.



Fig. 5. Recommendations from direct link citation graph.

Recommended Author
Campbell R.H.
Cardellini V.
Chao K.-M.
Chen C.
Jararweh Y.
Paraiso F.
Pernici B.
Rong H.
Safieddine I.
Sharma A.K.
Sivic J.
ValiKardan S.
Williams P. Yan Z.
Yang K.

C. who belongs to level 3, has not collaborated with Zheng Q. as co-author, nor has he quoted by him, nor does there exist direct link citation between them. Therefore, Cheng C. would be a good reviewer candidate because no relationship has been identified that could incur a problem of conflicts of interest when reviewing the scientific document written by Zheng Q. Depending on the level of restriction that you wish to apply, the authors that are pigeonholed in levels 1 and 2, can also be considered.

## V. CONCLUSIONS

Per review plays a central role in the research life, since scientific publications must pass the judge of anonymous colleagues in order to determine the quality of the manuscript. Also, per review is not only restricted to publications in journals, conferences or books. It is commonly used to determine research project proposals, grants and scholarship positions.

TABLE V  
TABLE OF BIAS PREVENTION LEVELS.

Level 1	Level 2	Level 3
Chao K.-M.	Campbell R.H.	Chen C.
Chen L.-X.	Cao Y.	Lee S.-Y.
Dunagan J.	Cardellini V.	Su G.
Ghaffarkhah A.	Hara T.	Xue W.
He T.-Q.	Jararweh Y.	Yan Z.
Letier E.	Li C.	
Ma X.	Liao S.	
Megiddo N.	Pernici B.	
Mignot J.-C.	Rong H.	
Nagin K.	Su D.	
Ota K.	Wong D.W.S.	
Paraiso F.	Yu N.	
Rai A.	Zhou F.-F.	
Safieddine I.		
Sharma A.K.		
Sivic J.		
Talbi E.-G.		
Tam A.S.W.		
Tan K.		
ValiKardan S.		
Wakelam V.		
Williams P. Xu P.		
Yang K.		
Yang Y.		
Zhu M.		

Reviewers must act following an ethic rules and conflict of interests must be avoided. The criteria of the researchers could be biased by several aspects: i) collaboration (e.g. co-authorship), ii) social relationship (e.g. belonging to the same department), and iii) intellectual (e.g. reviewer could be cited several time). In fact, the biases could be positive or negative. That is, a reviewer could favor or work against the author. Some of these effects could be avoided using a blind review, where the reviewers do not know the identity of the authors.

In this sense, the selection of reviewer without conflict of interests with the author could be a difficult task, since a reviewer should not have interacted socially with the author.

As above mentioned, researchers have many kinds of relationships with their colleagues in their academic life. These interactions could be modeled as a bibliographic network, where the nodes are the researchers and the edges are the interactions among them.

In this contribution, a novel approach to identify potential reviewers without conflict of interest based on social bibliographic networks is presented. Particularly, three kinds of social network are used: co-author, author co-citation and author direct link citation.

Using this variety of scientific social networks, three levels of social relationships can be determined. Therefore, three levels of bias prevention could be used, being the level 3, the most restrictive, where a researcher could be a reviewer if he/she never appear as co-author, neither in the author reference list.

For this, the transitivity property of networks have been studied in order to find possible candidates for reviewers using an undirected relationship between authors that we have called: *Friend of my friend who is not my friend*. This is an application of the transitivity property that indicates the fact that a node  $A$  knows a node  $B$  and a node  $B$  knows a node  $C$  does not guarantee that  $A$  knows  $C$ . Therefore, with these candidates we can apply level 3 of bias prevention in order to find real good possible reviewers.

The presented approach could be used by journal officers or panel experts in order to select the most suitable reviewers to evaluate different kind of research outputs.

As future work, we propose to include in the preprocessing process a disambiguation task in order to avoid the problem of finding two or more nodes that actually represent the same person. For instance, we have found {DÍAZ BRUGERA, JAVIER} and {DÍAZ BRUGUERA, JAVIER} which obviously refers to the same person but his/her name is misspelled.

#### ACKNOWLEDGMENT

This study was funded by FEDER funds in the Spanish Department for Economy and Competitiveness project (TIN2016-75850-R).

#### REFERENCES

- [1] V. Batagelj and M. Cerinšek, "On bibliographic networks," *Scientometrics*, vol. 96, no. 3, pp. 845–864, 2013.
- [2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, "Complex networks: Structure and dynamics," pp. 175–308, 2006.
- [3] M. Cobo, A. López-Herrera, E. Herrera-Viedma, and F. Herrera, "An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field," *Journal of Informetrics*, vol. 5, no. 1, pp. 146–166, 2011.
- [4] E. Noyons, H. Moed, and M. Luwel, "Combining mapping and citation analysis for evaluative bibliometric purposes: A bibliometric study," *Journal of the American Society for Information Science*, vol. 50, no. 2, pp. 115–131, 1999.
- [5] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proceedings of the National Academy of Sciences*, vol. 102, no. 46, pp. 16 569–16 572, 11 2005.
- [6] M. J. Cobo, A. G. López-Herrera, E. Herrera-Viedma, and F. Herrera, "Science mapping software tools: Review, analysis, and cooperative study among tools," *Journal of the American Society for Information Science and Technology*, vol. 62, no. 7, pp. 1382–1402, 2011.
- [7] M. J. Cobo, A. G. López-Herrera, E. Herrera-Viedma, and F. Herrera, "SciMAT: A new science mapping analysis software tool," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 8, pp. 1609–1630, 2012.
- [8] W. Glanzel, "National characteristics in international scientific co-authorship relations," *Scientometrics*, vol. 51, no. 1, pp. 69–115, 2001.
- [9] M. Callon, J.-P. Courtial, W. A. Turner, and S. Bauin, "From translations to problematic networks: An introduction to co-word analysis," *Social Science Information*, vol. 22, no. 2, pp. 191–235, 3 1983.
- [10] M. Cobo, M. Martínez, M. Gutiérrez-Salcedo, H. Fujita, and E. Herrera-Viedma, "25years at Knowledge-Based Systems: A bibliometric analysis," *Knowledge-Based Systems*, vol. 80, pp. 3–13, 5 2015.
- [11] M. M. Kessler, "Bibliographic coupling between scientific papers," *American Documentation*, vol. 14, no. 1, pp. 10–25, 1963.
- [12] H. Small, "Co-citation in the scientific literature: A new measure of the relationship between two documents," *Journal of the American Society for Information Science*, vol. 24, no. 4, pp. 265–269, 7 1973.
- [13] C. C. Aggarwal, Ed., *Social Network Data Analytics*. Boston, MA: Springer US, 2011.
- [14] S. Wasserman and K. Faust, *Social network analysis : methods and applications*. Cambridge University Press, 1994.

# Science Mapping Artificial Intelligence

Benjamín Vargas-Quesada  
Departamento de Información y  
Comunicación  
Universidad de Granada.  
Granada, Spain  
e-mail: [benjamin@ugr.es](mailto:benjamin@ugr.es)

Zaida Chinchilla Rodríguez  
Instituto de Políticas y Bienes  
Públicos (IPP). Consejo Superior de  
Investigaciones Científicas (CSIC).  
Madrid, Spain  
e-mail: [zaida.chinchilla@cchs.csic.es](mailto:zaida.chinchilla@cchs.csic.es)

Antonio Perianes-Rodríguez  
Departamento de Biblioteconomía y  
Documentación.  
Universidad Carlos III  
Madrid, Spain  
e-mail: [antonio.perianes@uc3m.es](mailto:antonio.perianes@uc3m.es)

**Abstract**— Mediante *science mapping*, se muestra la visualización de la estructura intelectual de la inteligencia artificial, a nivel mundial. Se identifican sus principales líneas de investigación a partir de las citas de sus trabajos, como interactúan entre sí, mediante la detección de los nodos con un alto grado de centralidad de *betweenness* o intermediación, así como el nivel de actividad, de cada una de ellas, por medio del uso de un indicador denominado *citation burst*. Se finaliza con un estudio longitudinal y evolutivo de las líneas detectadas.

**Keywords**— *Science Mapping; Inteligencia Artificial; Líneas de Investigación.*

## I. INTRODUCCIÓN

Hace más de 50 años, que Moreno [1] propuso el uso de las redes sociales para analizar el comportamiento humano y detectar patrones e individuos relevantes dentro de las mismas. Hoy día, el concepto de red social ha cambiado, trasladándose la representación gráfica de las interacciones entre los individuos o entidades (grafos), a denominar a las compañías que ofrecen servicios de redes y medios sociales.

La aparición del *big data* ha permitido que grandes conjuntos de datos, que hasta hace poco, eran imposibles de procesar mediante el uso de herramientas tradicionales debido a su tamaño y complejidad, puedan ahora ser tratados y utilizados con fines analíticos y de predicción. La combinación del *big data*, *datamining*, bibliometría y redes sociales, ha dado lugar a lo que se ha denominado como *science mapping* [2], cuyo objetivo es visualizar la estructura y relaciones de la ciencia en general, y de las áreas científicas, disciplinas, líneas de investigación, autorías, etc., en particular. *Science mapping* ha sido utilizado para mostrar la columna vertebral de la ciencia a nivel mundial [3], representar y analizar la estructura de grandes bases de datos científicas [4] y, visualizar la estructura intelectual del grafeno, sus principales líneas de investigación y evolución mundial [5], entre otras importantes iniciativas. Hasta donde llega nuestro conocimiento, no hay ningún trabajo que visualice la estructura intelectual de la IA, ni que detecte e

identifique sus principales líneas de investigación, a partir de ella. Este trabajo persigue los siguientes objetivos:

- Visualizar y analizar la estructura de la red de la Inteligencia Artificial (IA).
- Identificar las principales líneas de investigación.
- Localizar los trabajos que actúan como conectores de las líneas de investigación.
- Detectar tendencias emergentes de investigación.
- Realizar un breve estudio longitudinal de las principales líneas de investigación.

## II. MATERIAL Y MÉTODOS

Por medio de la consulta:  $WC=(Computer\ Science, Artificial\ Intelligence) AND TS=(Artificial\ Intelligence)$ , filtrando por *papers, proceedings paper o reviews*, el 28 de mayo de 2018, descargamos de Web of Science 8.635 documentos publicados entre 1971 y 2017, conteniendo un total de 172.109 referencias. Para la visualización y análisis de los datos, hemos utilizado *CiteSpace*, que es un *software* gratuito diseñado específicamente para este tipo de tareas. En aras a nuestro objetivo principal, utilizamos las relaciones de citación de los documentos descargados, para establecer la estructura intelectual e identificar los clusters o líneas de investigación. Cada línea es etiquetada en base a los *keywords* de los artículos que las citan, como se indica en la Figura 1.

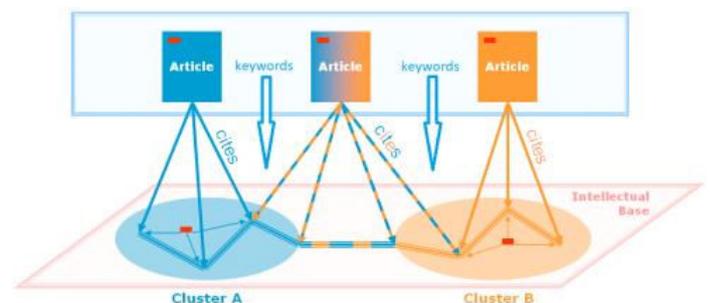


Fig. 1. Proceso para la detección y etiquetado de líneas de investigación

El uso de las citas, al igual que otras unidades de análisis de corte bibliométrico como son los autores, términos

extraídos de títulos y/o resúmenes, keywords, etc., podría ser considerado como una limitación, por las distorsiones que algunas de sus características inherentes como la intencionalidad, cobertura, popularidad, etc., pueden causar. Para salvar esta limitación, en cada visualización, se han establecido periodos de 5 años, y sólo se han representado los 50 documentos más citados de cada periodo. De esta forma se reduce en gran medida las distorsiones que se puedan producir, al representar la opinión consensuada de los autores, mediante sus citas.

### III. RESULTADOS

Las visualizaciones y resultados aquí mostrados, pueden ser replicados mediante el programa *Citespace*, disponible en: <http://cluster.cis.drexel.edu/~cchen/citespace>, en combinación con los registros descargados de Web of Science, accesibles

en:

<https://www.dropbox.com/sh/5i0nb0npcsavg4/AAA4uGb05fsCLN-L-18jekZca?dl=0>.

#### A. Estructura de la Red de IA

La red de IA a nivel mundial, cuenta con 382 nodos, 805 enlaces y una densidad de 0.011. El componente de mayor tamaño, Figura 2, está compuesto por un total de 287 nodos (75% de la red total). Mediante el método de Ward [6], se han identificado 8 clusters o líneas principales de investigación. La modularidad Q [7] de la red, es de 0.8066, indicando que está dividida en líneas de investigación altamente enlazadas y relacionadas entre sí. Su media Silhouette [8] es de 0.3387, sugiriendo, que la homogeneidad de dichas líneas es más bien baja.



Fig. 2. Science map mundial de IA y principales líneas de investigación

#### B. Principales Líneas de Investigación

Para su denominación, hemos recurrido a los *keywords* de los artículos fuente de cada línea, seleccionando el más representativo mediante el algoritmo likelihood ratio, por sus buenos resultados en cobertura y unicidad. Por orden de tamaño, las principales líneas de investigación son las siguientes: *Natural intelligence*, *Computational and artificial intelligence*, *Function aproximacion*, *Case-based reasoning*, *Distributed artificial intelligence*, *Support vector machines*, *Heuristics*, y *Firefly algorithms*, como se puede ver en la Figura 2, precedidas por el símbolo #.

#### C. Trabajos Responsables de Conectar las Líneas de Investigación

La centralidad de intermediación (*betweenness*) de un nodo en una red, mide el grado en que el nodo es parte de *paths* que conectan otros nodos [9]. En este caso, con valores comprendidos entre 0 y 1, la utilizamos para detectar las referencias responsables de conectar distintas líneas de investigación. Las referencias con un *betweenness* alto, se muestran con un borde morado. Cuanto más grueso es el borde, mayor es su valor de intermediación, ver Figura 2. Por ejemplo, el trabajo de Russell SJ, 2003, *Artificial Intelligence*, tiene una centralidad de 0.37, la más alta de toda la red, y es el responsable de la conexión de las líneas de investigación *Natural intelligence*, *Computational and artificial intelligence*, *Case-base reasoning* y *Support vector machines*. Este análisis es extensible al resto de nodos de color rojo, en función de su tamaño.

D. Detectar tendencias emergentes de investigación

Citation burst (explosión de citas) [10], es un indicador utilizado para la detección de referencias y líneas de investigación activas, que reciben un elevado número de citas durante uno o varios años. La explosión de citas, proporciona la evidencia de que un trabajo concreto ha atraído un alto grado de atención por parte de sus colegas, reflejándolo

mediante sus citas, Figura 3. Si una línea de investigación contiene numerosas referencias de este tipo, se puede considerar que es muy activa y con una tendencia emergente. Si un nodo presenta un círculo rojo en su interior, Figura 4, indicará que atrae un gran número de citas. A mayor citation burst, mayor tamaño del círculo.

Top 25 References with the Strongest Citation Bursts

References	Year	Strength	Begin	End
KOLOUDNER J, 1993, CASE BASED REASONING, V0, P0	1993	11.5028	1996	2001
PEARL J, 1988, PROBABILISTIC REASON, V0, P0	1988	11.2276	1991	1996
GOLDBERG DE, 1989, GENETIC ALGORITHMS S, V0, P0	1989	10.5589	1991	1997
BROOKS RA, 1991, ARTIF INTELL, V47, P139, DOI	1991	8.8768	1995	1998
WOOLDRIDGE M, 1995, KNOWL ENG REV, V10, P115, DOI	1995	8.3438	1997	2003
BOND AH, 1988, READINGS DISTRIBUTED, V0, P0	1988	7.6561	1991	1995
WINOGRAD T, 1986, UNDERSTANDING COMPUT, V0, P0	1986	5.7848	1991	1994
WATERMAN D, 1986, GUIDE EXPERT SYSTEMS, V0, P0	1986	5.7848	1991	1994
RUMELHART DE, 1986, PARALLEL DISTRIBUTED, V1, P0	1986	5.2037	1991	1994
MINTON S, 1992, ARTIF INTELL, V58, P161, DOI	1992	4.9822	1997	2000
HAYESROTH B, 1985, ARTIF INTELL, V26, P251, DOI	1985	4.7946	1991	1993
HARNAD S, 1990, PHYSICA D, V42, P335, DOI	1990	4.6658	1993	1998
REITER R, 1987, ARTIF INTELL, V32, P57, DOI	1987	4.5838	1992	1995
NORVIG P, 1995, ARTIFICIAL INTELLIGE, V0, P0	1995	4.4763	1996	2003
GENESERETH MR, 1987, LOGICAL F ARTIFICIAL, V0, P0	1987	4.3618	1991	1995
RICH E, 1991, ARTIFICIAL INTELLIGE, V0, P0	1991	4.3304	1993	1999
KUIPERS B, 1986, ARTIF INTELL, V29, P289, DOI	1986	4.2375	1992	1994
RIESBECK CK, 1989, INSIDE CASE BASED RE, V0, P0	1989	4.2184	1992	1997
NEWELL A, 1990, UNIFIED THEORIES COG, V0, P0	1990	4.2055	1993	1995
GASSER L, 1991, ARTIF INTELL, V47, P107, DOI	1991	4.0773	1992	1999
DECHTER R, 1991, ARTIF INTELL, V49, P61, DOI	1991	4.0773	1992	1999
DEKLEER J, 1986, ARTIF INTELL, V28, P127, DOI	1986	4.0433	1991	1994
WELLMAN MP, 1993, J ARTIF INTELL RES, V1, P1	1993	4.0285	1996	1999
GASSER L, 1989, DISTRIBUTED ARTIFICI, V2, P0	1989	3.8146	1991	1995
SUCHMAN LA, 1987, PLANS SITUATED ACTIO, V0, P0	1987	3.8146	1991	1995

Fig. 3. Top 25 trabajos con mayor citation burst

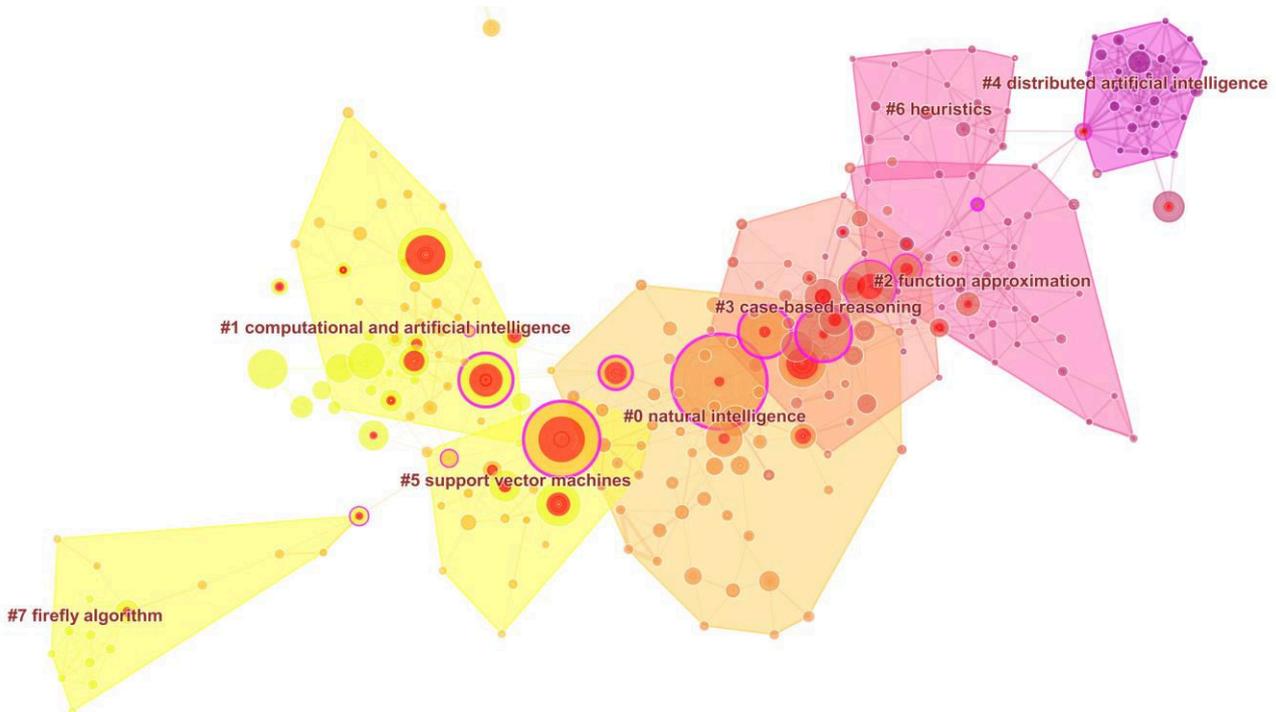


Fig. 4. Representación de los Top 25 trabajos con mayor citation burst

Por ejemplo, la línea de investigación 3 *Case-based reasoning*, contiene bastantes trabajos con un nivel de *citation burst* elevado, mostrando su alto nivel de actividad y emergencia. Así, el trabajo de Kolodner J, 1993, Case Based Reasoning, con un valor de 11,5028, seguido muy de cerca del de Pearl J, 1988, Probabilistic Reason, con 11,2276 de *citation burst*, respectivamente, ponen de manifiesto el amplio uso de estos trabajos y la relevancia de esa la línea de investigación.

E. Breve estudio longitudinal de las principales líneas de investigación etectar tendencias emergentes de investigación

*Timeline view*, Figura 5, es un tipo de representación que proporciona una visualización de la red por medio de líneas de tiempo horizontales. En este caso, se ha dividido en periodos

de cuatro años y la etiqueta de cada línea de investigación, aparece al final de la línea temporal, a la derecha. Las referencias citadas, se representan con círculos con un tamaño proporcional a la cantidad de citas recibidas en ese segmento de tiempo. Por lo tanto, un gran tamaño denota una referencia altamente citada. El color de cada referencia se corresponde con el segmento de tiempo en el que fueron citadas, pero se posicionan en base al año en que fueron publicadas. Además, se utilizan otro tipo de colores, explicados anteriormente, para resaltar atributos especiales como el *betweenness* o el *citation burst*. En este caso, las líneas que conectan dos referencias en la visualización, representan enlaces de cocitación, e indican la relación entre ellas, así como entre las líneas de investigación a las que pertenecen.

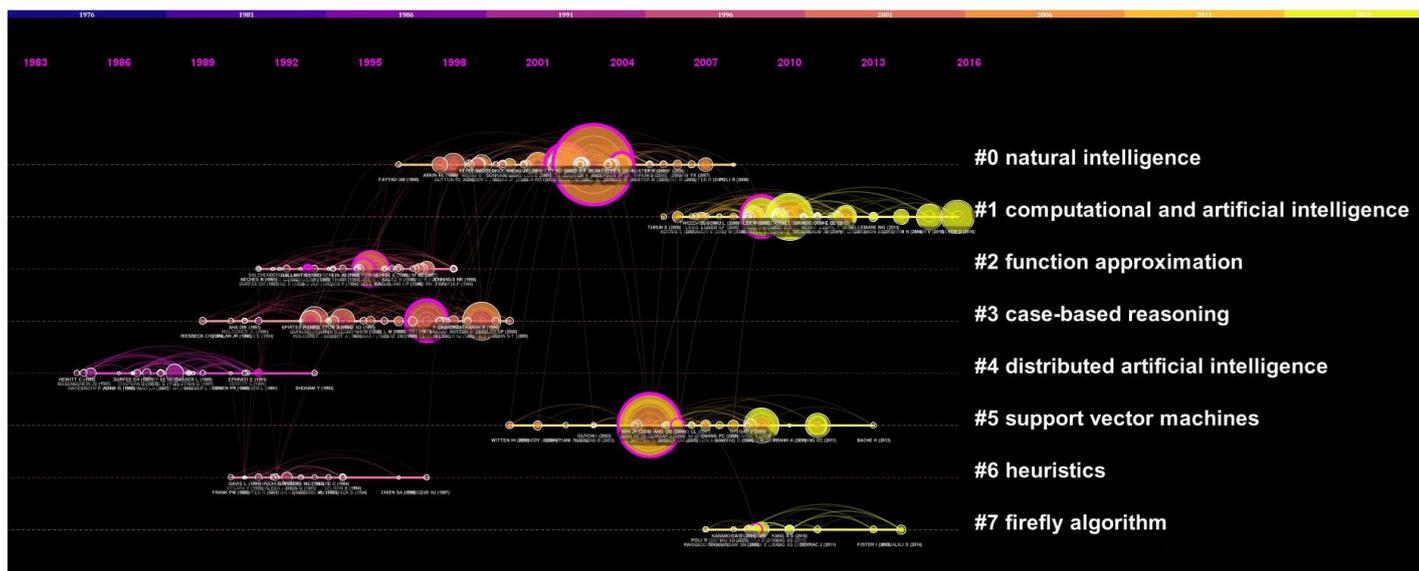


Fig. 5. *Timeline view* de las principales líneas de investigación

A nivel general y de forma breve, podemos decir que las líneas de investigación que se nutren de conocimientos más lejanos en el tiempo son *Distributed artificial intelligence* y *Heuristic*, pero sobre todo, la primera. Son líneas muy consolidadas, pero que no reciben muchas citas, como indica el tamaño de sus nodos, también son bastante lejanas en el tiempo, como muestra el color de sus nodos. Esto revela poca actividad, y una base de conocimiento con poca evolución.

*Function approximation* y *Case-base reasoning*, también son líneas consolidadas que se alimentan mucho de trabajos pretéritos. No obstante, sus citas son de fechas más recientes, y son puntos importantes de interconexión con otras líneas, como muestran las aureolas moradas de muchos de sus nodos. Son también bastante activas y atraen la atención de los investigadores, como podemos ver por su elevado número de nodos rojos.

*Natural intelligence* es la línea de investigación de mayor tamaño, la que más interrelaciona con el resto, como indica el contar con la referencia de mayor *betweenness*, así como otras con valores similares, a la vez que permanece bastante activa. Su número de citas no es muy elevado, pero sí es superior a las anteriores líneas de investigación.

*Computational and artificial intelligence*, *Support vector machines* y *Firefly Algorithm*, son las líneas con un mayor número de citas recientes. Sobre todo la primera.

*Computational and artificial intelligence* y *Support vector machines*, reciben un número elevado de citas a lo largo del tiempo, tienen trabajos con un elevado índice de *citation burst*, y cuentan con referencias con un alto grado de *betweenness*, permitiéndoles relacionarse entre sí y sobre todo, con *Natural intelligence*.



*Firefly Algorithm* es una línea joven, como se observa en su momento de aparición y las fechas de las citas que recibe, pero ya cuenta con varias referencias que le permiten ser una línea activa e interconectada con el resto.

#### IV. CONCLUSIONES

Mostramos por primera vez la red de la estructura intelectual de la IA a nivel mundial. Se han identificado ocho líneas principales de investigación altamente relacionadas entre sí, a la vez que poco homogéneas a nivel individual. Cinco de ellas están muy consolidadas y tres, son de más reciente aparición. Prácticamente, todas las líneas interactúan entre sí, pero sobre todo lo hace *Natural intelligence*, que es a su vez la de mayor tamaño. Son especialmente activas, por el número de citas que atraen: *Case-based reasoning*, *Computational and artificial intelligence* y *Support vector machines*.

#### BIBLIOGRAFIA

- [1] J.L. Moreno. Who shall survive? Beacon House, New York, 1953.
- [2] M.J. Cobo, A.G. López-Herrera, E. Herrera-Viedma and F. Herrera. "Science mapping software tools: Review, analysis, and cooperative study among tools". *Journal of the American Society for Information Science and Technology*, vol. 62, nº 7, pp. 1382-1402, 2011.
- [3] F. Moya-Anegón, B. Vargas-Quesada, Z. Chinchilla-Rodríguez, E. Corera-Álvarez, F.J. Muñoz-Fernández, and V. Herrero-Solana, "Visualizing the marrow of science". *Journal of the American Society for Information Science and Technology*, vol. 58, pp. 2167-2179, 2007.
- [4] A.J. Gómez-Núñez, B. Vargas-Quesada, Z. Chinchilla-Rodríguez, V. Batagelj, and F. Moya-Anegón, "Visualization and analysis of SCImago Journal & Country Rank structure via journal clustering", *Aslib Journal of Information Management*, vol. 68, nº 5, pp. 607-627, 2016.
- [5] B. Vargas-Quesada, Z. Chinchilla-Rodríguez, and N. Rodríguez, "Identification and Visualization of the Intellectual Structure in Graphene Research", *Frontiers in Research Metrics and Analytics*, vol. 2, pp. 1-22, 2017.
- [6] J.H.Jr. Ward, "Hierarchical Grouping to Optimize an Objective Function", *Journal of the American Statistical Association*, vol. 58, pp. 236-244, 1963.
- [7] C. Chen, "Predictive effects of structural variation on citation counts", *Journal of the American Society for Information Science and Technology*, vol. 63, nº 3, pp. 431-449, 2012.
- [8] C. Chen, F. Ibekwe-SanJuan, and J. Hou, "The structure and dynamics of co-citation clusters: a multiple-perspective co-citation analysis". *Journal of the American Society for Information Science and Technology*, vol. 61, nº 7, pp. 1386-1409, 2010.
- [9] L.C. Freeman, "A set of measuring centrality based on betweenness", *Sociometry*, vol. 40, pp. 35-41, 1977.
- [10] J. Kleinber, "Bursty and hierarchical structure in streams", 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Edmonton, Alberta, Canada, 2002, <http://www.cs.cornell.edu/home/kleinber/bhs.pdf>



---

# **I Workshop en Ciencia de Datos en Redes Sociales (CidReS 2018)**

SESIÓN 2







# On the analysis of the influence of the evaluation metric in community detection using GRASP

Sergio Pérez-Peló

Dept. of Computer Sciences

Universidad Rey Juan Carlos

C/Tulipán, S/N, 28933, Móstoles, Spain

Email: sergio.perez.pelo@urjc.es

Jesús Sánchez-Oro

Dept. of Computer Sciences

Universidad Rey Juan Carlos

C/Tulipán, S/N, 28933, Móstoles, Spain

Email: jesus.sanchezoro@urjc.es

Abraham Duarte

Dept. of Computer Sciences

Universidad Rey Juan Carlos

C/Tulipán, S/N, 28933, Móstoles, Spain

Email: abraham.duarte@urjc.es

**Abstract**—Community detection in social networks is becoming one of the key tasks in social network analysis, since it helps analyze groups of users with similar interests, detect radicalisms, or reduce the size of the data to be analyzed, among other applications. This paper presents a metaheuristic approach based on Greedy Randomized Adaptive Search Procedure methodology for detecting communities in social networks. The community detection is modeled as an optimization problem where the objective function to be optimized is the modularity of the network, a well known metric in community detection. The results obtained outperforms traditional methods of community detection as Edge Betweenness, Fast Greedy and Infomap over a set of real-life instances derived from Twitter.

## I. INTRODUCTION

The evolution of social networks in the last decades has aroused the interest of scientist from different and diverse areas, from psychology to computer sciences. Millions of people constantly share all their personal and professional information in several social networks. Furthermore, social networks have become one of the most used information sources, mainly due to their ability to provide the user with real-time content. Social networks are not only a new way of communication, but also a powerful tool that can be used to gather information related to several issues: which political party is the favourite for the next elections, what are the most commented movies in the last year, which is the best rated restaurant in a certain area, etc.

The analysis of social networks has become one of the most popular and challenging tasks in data science [1]. One of the most tackled problems in social networks is the analysis of the relevance of the users in a given social network. The relevance of a user is usually related to the number of followers or friends that the user has in a certain social network. However, this concept can be extended since a user may be relevant not only if he/she is connected with a large number of users, but with users that are also relevant. Several metrics have been proposed for analyzing the relevance of a user in a social network, emerging the PageRank [2] as one of the most used.

The problem of evaluating the relevance of a user has evolved in a more complex problem which consists of detecting specific users, often named influencers, with certain personal attributes that can be personal (credibility or enthusiasm) or related to their social networks (connectivity or centrality).

These attributes allow them to influence a large number of users either directly or indirectly [3].

Another important problem regarding the influence of people in other users is the analysis of sentiments in social networks. It is focused on finding out what do people think about a certain topic by analyzing the information they post in social networks. We refer the reader to [4] to find a complete survey on sentiment analysis techniques.

The previously described problems are related to individual users. However, there also exists some problems related to the structure of the network, devoted to find specific attributes and properties that can help to infer additional information of the social network. Community detection emerges as one of the most studied problems related to the structure of the network.

Most of the social networks present a common feature named community structure. Networks that have this property has the capacity to be divided into groups in such a way that the connections among users in the same group are dense, while connections among users in different groups are sparse. A connection can represent different features depending on the social network and the user profile, from professional relationships to friendships or hobbies in common. Community detection tasks are devoted to find and analyze these groups in order to better understand and visualize the structure of network and the relationships among their users.

Performing community detection algorithms over current social networks requires from a huge computationally effort mainly due to the continuous growth of social networks. Furthermore, since social networks are constantly changing (new friendships, mentions to users, viral information, etc.), it is interesting to perform the community detection in the shortest possible computing time, producing real-time information. These features make traditional exact methods not suitable for the current size of social networks, requiring from approximation algorithms in order to accelerate the process without losing quality. Recent works have tackled the community detection algorithm from a non-exact perspective in order to generate high quality solutions in short computing time [5].

The growth of social networks complicates their representation and understanding. The communities of a social network usually summarizes the whole network but reducing

its size and, therefore, making it easier to analyze. Furthermore, detecting communities in social networks has several practical applications. Recommender systems leverage the data of similar users in order to suggest new contents. In order to find similar users in a network we can simply perform a community detection over the network [6], improving the results of the recommender system. Communities in social networks also identifies people with similar interests, allowing us to evaluate the popularity of a political party [7], or even to detect radicalisms in social networks [8].

The remaining of the paper is structured as follows: Section II formally defines the problem considered as well as the metrics proposed for the evaluation of solutions; Section III describes the traditional algorithms proposed for detecting communities in social networks; Section IV presents the new algorithm proposed for detecting communities; Section V shows the computational experiments performed to test the quality of the proposal; and finally Section VI draws some conclusions on the research.

## II. PROBLEM STATEMENT

A social network is represented as a graph  $G = (V, E)$ , where the set of vertices  $V$ , with  $|V| = n$ , represents the users of the network and the set of edges  $E$ , with  $|E| = m$ , represents relations between users belonging to the network. An edge  $(v_1, v_2) \in E$ , with  $v_1, v_2 \in V$  can represent different types of relations depending on the social network under consideration. For example, in Twitter a relation represents that a user follows / is followed by an other user, while in LinkedIn it represents a professional relationship.

This work is focused on the Community Detection Problem (CDP), which involves grouping users of a social network into clusters. A desirable community in a social network is densely connected to the nodes in the same community and sparsely connected (or even unconnected) to nodes in other communities. Therefore, the main objective is to obtain groups or clusters of users that are similar among them and, at the same time, different to the users in other clusters with respect to a certain criterion.

A solution for the CDP is represented by a set of decision variables  $S$ , with  $|S| = n$ , where  $S_v = j$  indicates that vertex  $v$  is assigned to cluster  $j$  in solution  $S$ . Figure 1(a) shows an example graph with 19 vertices and 31 edges derived from a social network. In this example, an edge represents a friendship relationship between two users; for instance, users A and B are friends, while users A and C are not friends but they have a friend in common, which is vertex D.

Figure 1(b) shows a possible solution  $S$  for the community detection problem, where each cluster is represented with a different color. Regarding the solution representation previously defined, we can check the cluster for each vertex. Table I shows the community in which each vertex has been inserted (for example, vertex A belongs to community 1, vertex G to community 2, and so on).

The CDP then can be modeled as an optimization problem which consists of finding a solution  $S^*$  that maximizes a certain objective function value. In mathematical terms,

$$S^* \leftarrow \arg \max_{S \in \mathbb{S}} f(S)$$

where  $\mathbb{S}$  is the set of all possible solutions for a given social network.

There exists a large variety of quality metrics that can be used as objective function for finding high quality solutions. Several metrics considers that the optimal partition (ground truth) is known beforehand, and tries to minimize the distance of the current partition with respect to the optimal one (e.g., Omega-Index [5]). However, this work considers networks where the optimal partition is not known. In this case, most of the metrics are focused on maximizing the density of intra-cluster edges (those connecting vertices of the same cluster) while minimizing inter-cluster edges (those connecting vertices in different clusters).

We consider two metrics that has been traditionally considered for optimizing the quality of a solution for the CDP: conductance and modularity [9]. For the sake of simplicity, all metrics are normalized in the range 0–1, where 1 indicates the value for the optimal partition and 0 is the expected score for a random assignment of users to clusters. Notice that in some of the metrics the optimal score of 1 is not possible for some networks due to their internal structure.

The first metric considered is the conductance [10]. Given a cluster  $k$ , its conductance,  $C_n(k, G)$ , is defined as the number of edges that connect vertices of different clusters divided by the minimum between the number of edges with an endpoint in the cluster and the number of edges with no endpoint in the cluster. A large value in the conductance indicates that there are several edges connecting vertices in different clusters and, therefore, the cluster does not represent a community. More formally,

$$\begin{aligned} C_n(k, G) &= \frac{|(v, u) \in E : S_v = k \wedge S_u \neq k|}{\min\{E_k, \overline{E}_k\}} \\ E_k &= |(v, u) \in E : S_v = k \vee S_u = k| \\ \overline{E}_k &= |(v, u) \in E : S_v \neq k \wedge S_u \neq k| \end{aligned}$$

Then, the conductance of a complete solution  $C_n(S, G)$  is evaluated as the average conductance for all the clusters in the graph. In order to have a direct comparison with other metrics, we subtract that value from 1, so the objective is again to maximize the opposite of the conductance  $\overline{C}_n(S, G)$  to produce high quality solutions (i.e.,  $\overline{C}_n(S, G) = 1 - C_n(S, G)$ ). Then, the opposite of the conductance value for the example depicted in Figure 1(b) is  $\overline{C}_n(S, G) = 0.63$ .

The third metric studied is the modularity [11], which evaluates, for each edge connecting vertices in the same cluster, the probability of the existence of that edge in a random graph. The modularity is evaluated as:

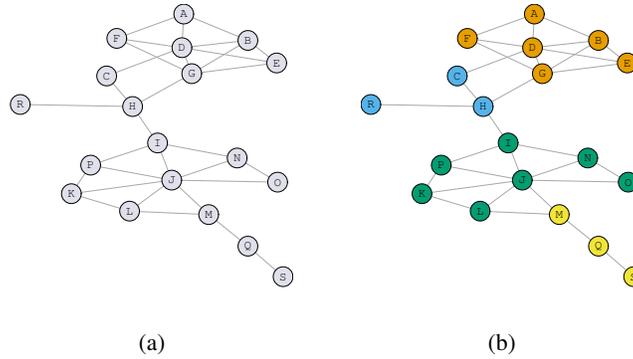


Fig. 1: 1(a) Example of a graph derived from a social network and 1(b) a possible solution for the community detection (each community is represented with a different color).

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	1	1	1	1	1	2	2	3	2	3	3	3	3	3	4	3	4	4

TABLE I: Cluster assigned to each vertex in the solution depicted in Figure 1(b).

$$Md(S, G) = \sum_{j=1}^{\max(S)} (e_{jj} - a_j^2)$$

$$e_{jj} = \frac{|\{(v, u) \in E : S_v = S_u = j\}|}{|E|}$$

$$a_j = \frac{|\{(v, u) \in E : S_v = j\}|}{|E|}$$

where  $\max(S)$  is the maximum value for the  $S_i$  variables, which corresponds to the number of clusters in the solution. The majority of the traditional algorithms for community detection considers this metric as the one to be optimized in order to find high quality communities. The modularity value for the graph depicted in Figure 1(b) is  $Md(S, G) = 0.50$ .

### III. ALGORITHMS FOR COMMUNITY DETECTION

Several algorithms has been proposed for detecting communities in social networks. Community detection algorithms can be classified in two different classes: agglomerative or divisive clustering. On the one hand, agglomerative methods starts from a solution where each vertex is located in a different cluster and tries to optimize a given objective function by joining two or more communities at each step. On the other hand, divisive methods starts from a solution with all the vertices located in a single cluster, and the objective function is optimized by dividing one or more clusters in each step. Most of the algorithms are not exact procedures, since in most of the networks it is not feasible to find the optimal solution in a reasonable time, mainly due to the number of users in the network. This Section is devoted to describe the most used algorithms in the state of the art for the CDP, in order to have a framework of comparison for the algorithm presented in this work.

#### A. Edge-Betweenness

The idea of the Edge-Betweenness algorithm [12] relies on identifying those vertices that appears in the majority of the paths in the graph. Specifically, authors define the edge betweenness of an edge as the number of shortest paths between pairs of vertices that contains the edge under evaluation. Therefore, groups or communities are generated by removing the edge with the largest edge betweenness value in each step. This algorithm presents a complexity of  $O(m^2n)$ .

#### B. Fast-Greedy

The Fast-Greedy algorithm [13] is focused on optimizing the modularity of the solutions generated. This agglomerative method starts from a solution where each vertex is located in a different cluster and iteratively join the two clusters that produce the solution with maximum modularity value. The optimization and data structures presented in the original work reduces the complexity of the algorithm to  $O(n \cdot m \cdot \log n)$ .

#### C. Infomap

The Infomap algorithm [14] proposed a fast stochastic and recursive search method which is based on joining neighbor vertices into the same community. The method starts with each vertex located in a different community. Then, it randomly selects a vertex and assigns it to the community that minimizes the map equation. The map equation is presented in this work and it is an efficient estimation of the optimality of a certain partition. Then, the method creates a new network where the new vertices are the communities detected until now. The algorithm stops when no changes are produced in the communities.

#### D. Evaluation of the previous methods

This Section is devoted to evaluate the results obtained by the different methods over an example graph that presents community structure [15]. Figure 2 illustrates the graphical

results over the community detection in the graph, where each community is represented with a different color.

As it can be seen, the results are different for each algorithm. Additionally, Table II presents the results obtained by each considered algorithm over the example graph depicted in Figure 2, considering the three metrics described in Section II and the number of communities found.

First of all, we will analyze the modularity metric, since it is the most used metric in community detection optimization. The best results in modularity corresponds to the Fast Greedy (0.5284) algorithm, closely followed by Edge Betweenness (0.5245) and InfoMap (0.5231). All the values are close since the considered algorithms are focused on optimizing the modularity.

Regarding the conductance, we can see that best results are obtained again with the Fast Greedy approach, closely followed by the Edge Betweenness algorithm. In this case, both Label Propagation and Infomap present worst results in term of conductance. These results suggest that optimizing one of the metrics does not guarantee a good result in the other considered ones.

Finally, analyzing the number of communities detected, most of the algorithms detect 6 communities, which seems to be the actual number of communities in the social network. The Label Propagation algorithms prematurely stops the search, resulting in only 4 communities, while the largest number of communities, 8, is found by Infomap algorithm.

#### IV. GREEDY RANDOMIZED ADAPTIVE SEARCH PROCEDURE

Greedy Randomized Adaptive Search Procedure (GRASP) is a metaheuristic originally presented in [16] and formally defined in [17]. We refer the reader to [18] for a recent survey on this methodology. This metaheuristic can be divided into two main phases: solution generation and local improvement.

The solution generation phase iteratively adds elements to an initially empty solution until it becomes feasible. The first element is usually selected at random, acting as a seed for the procedure. The algorithm then constructs a candidate list ( $CL$ ) with all the elements that must be included in the solution. After that, a Restricted Candidate List ( $RCL$ ) is created with the most promising elements of the  $CL$  according to a predefined greedy function. Then, in each iteration, an element is selected at random from the  $RCL$  and added to the solution under construction, updating the  $CL$  and  $RCL$  in each step until reaching a feasible solution.

The construction phase of the GRASP algorithm presents a random part devoted to increase the diversity of the solutions generated. In particular, in the previous description, the random part relies on the selection of the next element from the  $RCL$ . Therefore, most of the obtained solutions are not a local optimum and can be improved by means of a local optimizer. The second phase of the GRASP algorithm is intended to find a local optimum of the solution generated, usually applying a local search method, although it can be replaced with a more complex optimizer.

The algorithm presented in this section is able to optimize any of the metrics defined in Section I. However, since the algorithm considered for the comparison are focused on optimizing the modularity, the proposed algorithm is also focused on optimizing the modularity, which has been traditionally considered as a good optimization metric.

#### A. Constructive procedure

The constructive procedure designed for the community detection problem, named *GRASPAGG* follows an agglomerative approach, where each element is initially located in a different cluster. Then, *GRASPAGG* iteratively joins two of the most promising clusters with the objective of maximizing one of the aforementioned metrics. Algorithm 1 shows the pseudocode of the *GRASPAGG* constructive method.

---

#### Algorithm 1 *GRASPAGG*( $G, \alpha$ )

---

```

1:  $S_v \leftarrow v \forall v \in V$ 
2:  $CL \leftarrow \{1, 2, \dots, n\}$ 
3:  $continue \leftarrow \mathbf{True}$ 
4: while  $continue$  do
5:    $continue \leftarrow \mathbf{False}$ 
6:    $g_{\min} \leftarrow \min_{j \in CL} (e_{jj} - a_j^2)$ 
7:    $g_{\max} \leftarrow \max_{j \in CL} (e_{jj} - a_j^2)$ 
8:    $\mu \leftarrow g_{\min} + \alpha \cdot (g_{\max} - g_{\min})$ 
9:    $RCL \leftarrow \{j \in CL : (e_{jj} - a_j^2) \geq \mu\}$ 
10:   $j_1 \leftarrow \mathit{Random}(RCL)$ 
11:   $Md_{best} \leftarrow Md(S, G)$ 
12:   $j_2 \leftarrow -1$ 
13:  for  $j' \in 1 \dots CL$  do
14:     $S' \leftarrow S$ 
15:     $S'_v \leftarrow j_1 \forall S_v = j'$ 
16:    if  $Md(S', G) > Md_{best}$  then
17:       $Md_{best} \leftarrow Md(S', G)$ 
18:       $continue \leftarrow \mathbf{True}$ 
19:       $j_2 \leftarrow j'$ 
20:    end if
21:  end for
22:  if  $continue$  then
23:     $S_v \leftarrow j_1 \forall S_v = j_2$ 
24:     $CL \leftarrow CL \setminus \{j_2\}$ 
25:  end if
26: end while
27: return  $S$ 

```

---

The method starts by assigning a different cluster to each node in the graph  $G$  (step 1). Then, the  $CL$  is constructed with every cluster in the solution  $S$  under construction (step 2). Then, the minimum ( $g_{\min}$ ) and maximum ( $g_{\max}$ ) values for the greedy function under evaluation are calculated (steps 6-7). The proposed greedy function is the modularity value of each cluster  $j$ , which is  $e_{jj} - a_j^2$ , as stated in Section II. Then, a threshold  $\mu$  is evaluated (step 8) to construct the  $RCL$  with the most promising candidates in  $CL$  (step 9). The next steps selects the two clusters that will be merged in the current iteration, being the first cluster  $j_1$  to be merged selected at

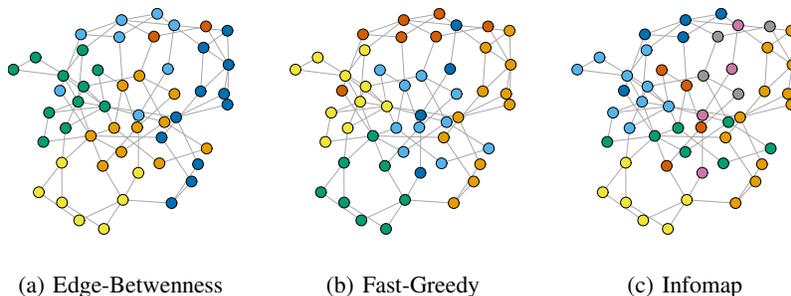


Fig. 2: Comparison of the community detection of the described algorithms over a example graph with 50 nodes that presents community structure.

TABLE II: Evaluation of the solution generated by each algorithm over the example graph using the three considered metrics

Algorithm	Modularity	Coverage	Conductance	Number of communities
Edge Betweenness	0.5245	0.7250	0.5248	6
Fast Greedy	0.5284	0.7125	0.5306	6
Infomap	0.5231	0.6750	0.4732	8

random from the *RCL* (step 10). The second cluster  $j_2$  is the one that maximizes the modularity of the resulting solution after merging clusters  $j_1$  and  $j_2$  (steps 11-21). If the method has found an improvement in the modularity after joining both clusters, a new iteration is performed, updating the incumbent solution (step 23) and the candidate list (step 24). *GRASPAGG* stops when it is not possible to join two clusters improving the modularity, returning the best solution found.

### B. Local optimization

This section presents a local search procedure designed to find a local optimum for every solution constructed in the previous phase. In order to define a local search method we firstly need to define the neighborhood in which the local optimum will be found. For this problem, we consider all the solutions that can be reached from a given solution  $S$  by moving a node from one cluster to another. Specifically, after performing the move  $Move(S, v, j)$ , the vertex  $v$  will be located at cluster  $j$  (i.e.,  $S_v \leftarrow j$ ). Notice that if  $v$  was the last vertex in its original cluster  $j'$ , then cluster  $j'$  will disappear after performing the move.

The next step for defining the local search method is the selection of the vertex to be moved to another community. For this purpose, we define a heuristic criteria based on the number of intra-cluster edges of the vertex under evaluation with respect to the total number of edges in the graph. Specifically, the local search selects the vertex  $v$  with the smallest ratio  $r$  between number of edges in the same cluster and the total number of incident edges to  $v$ . More formally,

$$r(v, S) \leftarrow \frac{|(v, u) \in E : S_v = S_u|}{|(v, w) \in E|} \quad \forall u, w \in V$$

The local search method selects, for each community, the node with the smallest value of this selection criteria among all nodes in the graph. Then, the node is moved to the

community that maximizes the modularity among all the existing communities in the incumbent solution.

The proposed local search procedure follows a first improvement approach. In particular, the first improvement move found is performed, restarting the search again, opposite to performing the best available move, which is often rather time consuming. The method stops when no improvement is found after evaluating the move of a node in every community.

### V. COMPUTATIONAL RESULTS

This Section is devoted to analyze the quality of the proposed algorithm when compared with the most popular community detection algorithms presented in Section III. Since most of the algorithms are focused on optimizing the modularity, the evaluation of the quality must be performed over a different metric. In this work we consider conductance as the evaluation metric, for testing the robustness of the methods. We additionally include the modularity value obtained by each algorithm, although it should not be considered in the evaluation of the quality of the community detection. However, we consider that it is interesting to analyze how far an algorithm is able to optimize the detection considering the modularity value. The proposed algorithm have been implemented in Java 8 and the experiments have been conducted in an Intel Core 2 Duo E7300 2.66 GHz with 4 GB RAM.

The instances used for the experiment have been extracted from the Twitter SNAP dataset<sup>1</sup>. Specifically, we have selected 100 instances with vertices ranging from 50 to 250, that represents the ego-network of several Twitter users (data is anonymized in the dataset).

The first experiment is devoted to tune the  $\alpha$  parameter of the *GRASPAPP* procedure. This parameter controls the degree of randomness of the method: on the one hand,  $\alpha = 0$  results in a totally random method, while  $\alpha = 1$  considers

<sup>1</sup><https://snap.stanford.edu/data/egonets-Twitter.html>

a completely greedy method. Therefore, it is interesting to test values distributed in the range 0–1 to analyze whether the best results for the CDP are obtained with a small or large percentage of randomness in the construction. In this experiment we have considered  $\alpha = \{0.25, 0.50, 0.75, RND\}$ , where *RND* indicates that a random value of  $\alpha$  is selected for each construction. This experiment has been conducted over a subset of 20 representative instances in order to avoid overfitting.

Table III reports the results obtained with the different values of  $\alpha$ . Specifically, three statistics are considered: Modularity, the average of the best modularity value obtained for each instance; Dev (%) the average deviation with respect to the best solution found in the experiment; and #Best, the number of times that an algorithm reaches that best solution.

TABLE III: Results obtained by the GRASP algorithm considering different values for  $\alpha$  parameter

$\alpha$	Modularity	Dev (%)	#Best
0.25	0.31961	1.60	9
0.50	0.32019	1.30	5
0.75	0.32063	1.08	4
RND	0.32080	1.08	6

As it can be derived from Table III, the best results are obtained with  $\alpha = 0.75$ . In particular, is able to obtain the best modularity values and reaches the best solution in 9 out of 20 instances. The average deviation value of 0.65% indicates that, in those instances in which it is not able to reach the best value, it remains rather close to it. Therefore, we select  $\alpha = 0.75$  for the final experiment.

The final experiment is intended to compare the quality of the solutions provided by our proposal with respect to the traditional methods described in Section III.

TABLE IV: Comparison of the considered metrics over all the algorithms presented in Section III and the proposed GRASP method.

	Modularity	Conductance
EB	0.14272	0.11319
FG	0.25064	0.34494
IM	0.14216	0.33014
GRASP	0.26029	0.38206

These results show the superiority of the proposal when considering both modularity and conductance, supported by a *p*-value lower than 0.0001 when applying the Friedman statistical test.

## VI. CONCLUSIONS

This paper has proposed a new metaheuristic method for community detection in social network based on Greedy Randomized Adaptive Search Procedure methodology. The problem is addressed by optimizing the modularity metric, which is a robust metric to evaluate the quality of a partition in a social network. The algorithm is compared with several well-known traditional algorithms for community detection using

conductance as evaluation metric. The computational results show how GRASP is able to obtain better results in both metrics than the previous methods, emerging as a competitive algorithm for detecting communities in social networks.

## ACKNOWLEDGMENT

This work has been partially funded by Ministerio de Economía y Competitividad with grant ref. TIN2015-65460-C2-2-P.

## REFERENCES

- [1] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford University Press, Oxford, 2003.
- [2] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: Bringing order to the web." Stanford InfoLab, Tech. Rep., 1999.
- [3] M. Gladwell, *The Tipping Point How Little Things Can Make a Big Difference*. Boston: Little Brown and Company, 2000.
- [4] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1–135, 2008.
- [5] A. González-Pardo, J. J. Jung, and D. Camacho, "ACO-based clustering for Ego Network analysis." *Future Generation Comp. Syst.*, vol. 66, pp. 160–170, 2017.
- [6] C. Cao, Q. Ni, and Y. Zhai, "An Improved Collaborative Filtering Recommendation Algorithm Based on Community Detection in Social Networks." in *GECCO*, S. Silva and A. I. Esparcia-Alcázar, Eds. ACM, 2015, pp. 1–8.
- [7] N. Zalmout and M. Ghanem, "Multidimensional community detection in Twitter." in *ICITST*. IEEE, 2013, pp. 83–88.
- [8] D. Camacho, A. González-Pardo, A. Ortigosa, I. Gilpérez-López, and C. Urruela, "RiskTrack: A New Approach for Risk Assessment on Radicalisation Based on Social Media Data." in *AfCAI*, ser. CEUR Workshop Proceedings, M. T. H. Ezquerro, G. J. Nalepa, and J. T. P. Mendez, Eds., vol. 1794. CEUR-WS.org, 2016.
- [9] S. Emmons, S. Kobourov, M. Gallant, and K. Börner, "Analysis of Network Clustering Algorithms and Cluster Quality Metrics at Scale," *PLoS ONE*, vol. 11, no. 7, 2016.
- [10] H. Almeida, D. O. G. Neto, W. M. Jr., and M. J. Zaki, "Is There a Best Quality Metric for Graph Clusters?" in *ECML/PKDD (1)*, ser. Lecture Notes in Computer Science, D. Gunopulos, T. Hofmann, D. Malerba, and M. Vazirgiannis, Eds., vol. 6911. Springer, 2011, pp. 44–59.
- [11] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review*, vol. E 69, no. 026113, 2004.
- [12] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *PNAS*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [13] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," *Physical Review E*, vol. 70, p. 066111, 2004.
- [14] M. Rosvall, D. Axelsson, and C. T. Bergstrom, "The map equation," *The European Physical Journal Special Topics*, vol. 178, no. 1, pp. 13–23, 2009.
- [15] A. Lancichinetti, S. Fortunato, and F. Radicchi, "Benchmark graphs for testing community detection algorithms," *Phys. Rev. E*, vol. 78, no. 4, p. 046110, 2008.
- [16] T. A. Feo and M. G. C. Resende, "A probabilistic heuristic for a computationally difficult set covering problem," *Operations Research Letters*, vol. 8, no. 2, pp. 67 – 71, 1989.
- [17] T. A. Feo, M. G. C. Resende, and S. H. Smith, "A Greedy Randomized Adaptive Search Procedure for Maximum Independent Set." *Operations Research*, vol. 42, no. 5, pp. 860–878, 1994.
- [18] M. G. C. Resende and C. C. Ribeiro, *GRASP: Greedy Randomized Adaptive Search Procedures*. Boston, MA: Springer US, 2014, pp. 287–312.



# Analizando la influencia de los atentados del 11-M en las elecciones generales de 2004: un enfoque ABM

Ignacio Moya\*, Manuel Chica\*, José L. Sáez-Lozano †, Oscar Cordón\*

\*Instituto Andaluz Interuniversitario DaSCI (Data Science and Computational Intelligence), Universidad de Granada

† Departamento de Economía Internacional y de España y GIADE-UGR, Universidad de Granada

Emails: imoya@ugr.es, manuelchica@ugr.es, josaez@ugr.es, ocordova@decsai.ugr.es

El 11 de marzo de 2004 (11-M), tres días antes de las elecciones generales del 14 de marzo, un grupo de terroristas detonó varios explosivos en trenes que se dirigían a la estación de Atocha en Madrid. 193 personas perdieron la vida y cerca de 2000 resultaron heridas. Los atentados alteraron el proceso electoral: la mañana del 11-M se suspendió la campaña; el 12 de marzo hubo manifestaciones contra el terrorismo en las principales ciudades españolas; el 13 de marzo hubo una manifestación frente a la sede del Partido Popular (PP), el partido en el gobierno en aquel momento. Finalmente las elecciones trajeron un inesperado cambio de gobierno.

Tras los atentados se generó una cantidad de información por parte del gobierno, políticos, y medios de información, forzando a los candidatos a posicionarse con respecto a los atentados. Una posición que los votantes incorporarían a su decisión de voto [1]. Este evento tuvo un marco comunicativo que se estructuró entorno a dos framings: la autoría de los atentados por parte de ETA o Al'Qaeda. La primera era la posición defendida por el gobierno del PP, mientras que el principal partido de la oposición, el Partido Socialista Obrero Español (PSOE), y otras fuerzas políticas defendían la segunda [2]. Muchos votantes evaluaban positivamente la gestión del PP en la lucha contra ETA, por lo que el presidente Aznar declaró a ETA responsable de los atentados del 11-M como estrategia electoral [3]. Por contra, la decisión del gobierno de participar en la invasión de Irak iba contra gran parte de la opinión pública y el resto de partidos políticos.

Distintos estudios postelectorales muestran que los atentados del 11-M influyeron en los votantes [3], [4], [5]. La interpretación más común es que el resultado de las elecciones fue un castigo al partido del gobierno tanto por su gestión de los atentados como por su política exterior. Aun así, no existen estudios anteriores que expliquen el efecto framing en la decisión electoral. Chong y Druckman [6] definen el efecto framing como el proceso psicológico que permite a las personas desarrollar una conceptualización *ad hoc* de un evento, ajustando su opinión en consecuencia. Según el CIS, el 11 % de los votantes cambiaron su voto tras los atentados [7].

Dada la importancia socio-económica y política de estos hechos en la historia reciente de España, nuestro objetivo es analizar el efecto framing generado tras los atentados y su influencia en la decisión de quienes tenían la intención de

votar al PP, al PSOE, o abstenerse tras los atentados. Ello nos permitirá conocer la influencia del tratamiento mediático de la responsabilidad de los atentados y la difusión por boca a boca (WOM) [8], [9]. Modelamos este escenario empleando *agent-based modeling* (ABM) [10], [11], dado que ya ha sido aplicada con éxito en otros escenarios políticos [12], [13]. La metodología ABM utiliza una población de entidades autónomas (agentes) que siguen reglas simples y tienen la capacidad de interactuar entre ellos. La agregación de ambos permiten representar dinámicas emergentes y definir escenarios *what-if*.

De este modo simulamos las 72 horas siguientes a los atentados y estudiamos como la información que se difunde durante este periodo de tiempo afecta a la población hasta llegar el momento de la votación. Segmentamos la población de agentes usando datos pre-electorales para replicar las principales opciones políticas: PP, PSOE, y abstención. Nuestro *framework* de simulación también reproduce la información de los medios y del WOM usando redes sociales artificiales [14]. Modelamos WOM difundiendo las percepciones de los votantes [8] a través de una red libre de escalas [14]. Respecto a los medios de información elegidos, tomamos los de mayor audiencia en esa época (i.e., televisión, radio, y prensa).

Usando datos pre-electorales reales como entrada, validamos nuestro diseño para ajustar su comportamiento a los resultados de las elecciones, calibrando algunos de sus parámetros tomando como objetivo el resultado de las elecciones. Aunque los valores de algunos parámetros se ajustan directamente usando datos reales de confianza, otros son más difíciles de estimar y deben ser ajustados usando calibración automática (como los parámetros que regulan la difusión de WOM y la influencia de los medios). En concreto, nuestra estrategia de calibración usa un algoritmo memético [15] basado en un algoritmo genético [16] y una búsqueda local, obteniendo un ajuste del 99.13 %. En la Figura 1 mostramos la evolución de votos de nuestro modelo para cada día de este periodo.

Definimos distintos experimentos donde observamos cómo los medios y el WOM influyen a los votantes a través de sus agentes correspondientes en el modelo calibrado. Estos experimentos son escenarios *what-if* donde analizamos comunicaciones alternativas de los medios y del WOM. Las alternativas para los medios suponen alterar el mensaje original para favorecer uno u otro marco. En el caso del WOM,

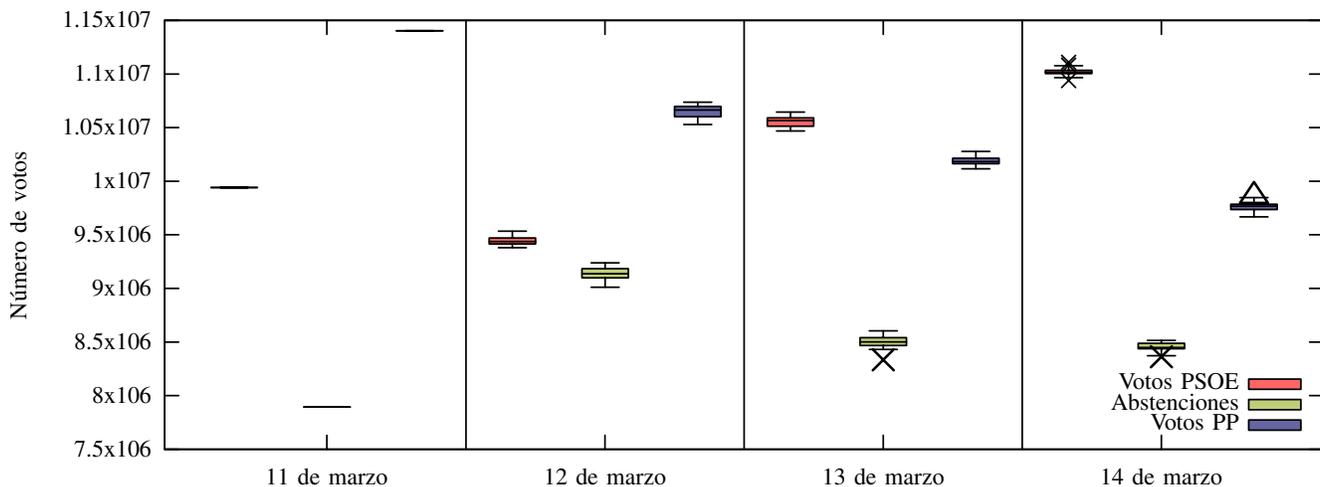


Figura 1: Promedio de votos por opción política para cada día entre los atentados y las elecciones.

los escenarios consideran comportamientos alternativos que modifican cómo los votantes reaccionan a la difusión de percepciones políticas. De este modo, monitorizamos el resultado de las elecciones para cada escenario.

Los resultados de los experimentos sugieren que el efecto marco pudo efectivamente influir en las elecciones movilizándolo abstencionistas y desactivando votantes del PP. En este punto podemos resaltar que los mecanismos de difusión tienen un papel muy importante durante este periodo, dado que aparece un gran intercambio de votos al modificar la densidad de la red social o la probabilidad de difusión. Esto a su vez es consistente con otros trabajos que tratan la difusión de percepciones políticas [17], [18]. También podemos observar que la red social tiene un papel importante al exponer a los agentes a mensajes altamente polarizados.

Los resultados también reflejan que la agregación de los medios de comunicación consigue un efecto mayor que la suma individual de los medios. Incluso si la televisión se distingue claramente como el medio más influyente, radio y prensa incrementan el efecto agregado de manera notable. Esto no solo resalta que la televisión tuvo un papel importante en la difusión de información durante este periodo, sino que la presencia de otros medios fue decisiva para los resultados de las elecciones.

#### AGRADECIMIENTOS

Este trabajo está financiado por el Ministerio de Economía y Competitividad bajo el proyecto NEWSOCO (ref. TIN2015-67661-P), incluyendo Fondos Europeos de Desarrollo Regional (ERDF).

#### REFERENCIAS

- [1] T. Holbrook, *Do campaigns matter?* Contemporary American Politics. SAGE publications, 1996.
- [2] J. A. Olmeda, "Fear or falsehood? Framing the 3/11 terrorist attacks in Madrid and electoral accountability," *Boletín Elcano*, no. 70, 2005.

- [3] I. Lago and J. R. Montero, *The 2004 election in Spain: Terrorism, accountability, and voting*. Barcelona: Institut de Ciències Polítiques i Socials, 2006.
- [4] I. Lago Peñas, J. R. Montero, and M. Torcal, "Del 11-M al 14-M: Los mecanismos del cambio electoral," *Claves de la razón práctica*, vol. 149, pp. 36–45, 2005.
- [5] V. A. Bali, "Terror and elections: Lessons from Spain," *Electoral Studies*, vol. 26, no. 3, pp. 669–687, 2007.
- [6] D. Chong and J. N. Druckman, "Framing theory," *Annual Review of Political Science*, vol. 10, pp. 103–126, 2007.
- [7] Centro de Investigaciones Sociológicas, "Estudio 2555," CIS Data bank, 2004.
- [8] D. Duzevik, A. Anev, P. Funes, and P. Gaudiano, "The effects of word-of-mouth: An agent-based simulation of interpersonal influence in social networks," *Word of Mouth Research Symposium*, 2007.
- [9] B. Libai, E. Muller, and R. Peres, "Decomposing the value of word-of-mouth seeding programs: Acceleration versus expansion," *Journal of Marketing Research*, vol. 50, no. 2, pp. 161–176, 2013.
- [10] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proceedings of the National Academy of Sciences*, vol. 99, no. suppl 3, pp. 7280–7287, 2002.
- [11] J. M. Epstein, *Generative social science: Studies in agent-based computational modeling*. Princeton University Press, 2006.
- [12] M. Laver, "Policy and the dynamics of political competition," *American Political Science Review*, vol. 99, no. 02, pp. 263–281, 2005.
- [13] J. Muis, "Simulating political stability and change in the netherlands (1998-2002): an agent-based model of party competition with media effects empirically tested," *Journal of Artificial Societies and Social Simulation*, vol. 13, no. 2, p. 4, 2010.
- [14] A. L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [15] P. Moscato, "On evolution, search, optimization, genetic algorithms and martial arts: towards memetic algorithms," Caltech Concurrent Computation Program, Pasadena, USA, Tech. Rep. 826, 1989.
- [16] T. Back, D. B. Fogel, and Z. Michalewicz, *Handbook of evolutionary computation*. Bristol (UK): IOP Publishing Ltd., 1997.
- [17] V. K. Singh, S. Basak, and N. Modanwal, "Agent based modeling of individual voting preferences with social influence," in *Trends in Computer Science, Engineering and Information Technology*. Springer, 2011, pp. 542–552.
- [18] Y. Sudo, S. Kato, and A. Mutoh, "The impact of exchanging opinions in political decision-making on voting by using multi-agent simulation," in *PRIMA 2013: Principles and Practice of Multi-Agent Systems*. Springer, 2013, pp. 340–354.



# Análisis de la influencia de los sentimientos en el proceso de detección de tópicos en redes sociales

Karel Gutiérrez-Batista  
Departamento de Ciencias de la  
Computación e Inteligencia Artificial  
Universidad de Granada  
Granada, España  
karel@decsai.ugr.es

Jesús R. Campaña  
Departamento de Ciencias de la  
Computación e Inteligencia Artificial  
Universidad de Granada  
Granada, España  
jesuscg@decsai.ugr.es

Maria-Amparo Vila  
Departamento de Ciencias de la  
Computación e Inteligencia Artificial  
Universidad de Granada  
Granada, España  
vila@decsai.ugr.es

Maria J. Martín-Bautista  
Departamento de Ciencias de la  
Computación e Inteligencia Artificial  
Universidad de Granada  
Granada, España  
mbautis@decsai.ugr.es

**Resumen**—En el presente trabajo se propone realizar un estudio para analizar la influencia que tienen los términos que expresan sentimientos en la detección automática de tópicos en redes sociales. Esta propuesta utiliza una metodología basada en una ontología, a la cual se le incorpora la capacidad de identificar y eliminar aquellos términos que presenten una orientación sentimental en textos de redes sociales, los cuales pueden influir de forma negativa en la detección de tópicos. Para ello se han utilizado dos recursos orientados al análisis de sentimientos con el fin de detectar dichos términos. El sistema propuesto ha sido evaluado con conjuntos de datos reales de las redes sociales Twitter y DreamChatchers en inglés y español respectivamente, demostrando en ambos casos la influencia de los términos con orientación sentimental en la detección de tópicos en textos de redes sociales.

**Palabras Claves**—Detección tópicos, Agrupamiento jerárquico, Etiquetado grupos, Términos sentimientos, Multilingüe

## I. INTRODUCCIÓN

Hoy día es un hecho reconocido el crecimiento y la popularidad alcanzada por las redes sociales, y como consecuencia de ello, el aumento del número de usuarios interactuando en dichas redes, lo que provoca la acumulación de grandes volúmenes de datos textuales no estructurados. Por tal motivo, las redes sociales constituyen una fuente de información de gran importancia, por lo que es de esperar que organizaciones, investigadores, etc., empleen tiempo y recursos en el estudio de estas. Sin embargo, el gran cúmulo y la falta de estructura de los textos, hace que sea prácticamente imposible su procesamiento y análisis automático de forma masiva, motivo por el cual resulta conveniente tener los textos previamente organizados teniendo en cuenta la temática abordada.

La detección de tópicos a partir de textos no estructurados, permite organizar dichos textos por temáticas, lo cual facilita su posterior análisis integrado con datos convencionales. En

[1] se ha propuesto una metodología multilingüe para la detección automática de tópicos en datos textuales. Mediante la experimentación, se demostró la viabilidad de la propuesta, aunque se debe resaltar que los resultados no son lo suficientemente buenos cuando los textos provienen de redes sociales.

Esto se debe a que la detección de tópicos en textos más elaborados (librerías digitales, sitios web de noticias, etc.) es diferente cuando los textos pertenecen a redes sociales. En dichos sistemas los usuarios expresan ideas, hechos y sentimientos sobre cualquier tema utilizando un lenguaje coloquial, por lo que es de esperar que en los textos aparezcan con alta frecuencia términos que permiten expresar sentimientos relacionados con determinados productos, servicios, etc.

Motivado por la problemática anterior, en este artículo se realiza un estudio para analizar la influencia de los términos que expresan sentimientos en la detección automática de tópicos en redes sociales. Para ello se propone un nuevo enfoque para mejorar la metodología para la detección automática de los principales tópicos presentes en datos textuales propuesta en [1], la cual utiliza técnicas de minería de datos, recursos relacionados con el análisis de sentimientos, y una base de conocimiento multilingüe. La nueva propuesta permite identificar y eliminar los términos con orientación sentimental, con el objetivo de mejorar los resultados del sistema sobre textos de redes sociales.

La idea básica consiste en realizar un filtrado que elimine las palabras que expresen sentimientos durante la etapa de preprocesamiento semántico presente en la metodología. Para ello, se utilizan los recursos léxicos SentiWordNet [2] y WordNet Affect [3] por separado y juntos, para luego comparar los resultados obtenidos. Para la experimentación se han utilizado cuatro conjuntos de datos, los cuales pertenecen a las redes sociales Twitter y Dreamcatchers, en inglés y español

respectivamente.

El resto de este artículo está estructurado de la siguiente forma. Se presenta una revisión de los trabajos previos relacionados con el presente tema en la Sección 2. La Sección 3 describe brevemente la propuesta seleccionada para la detección automática de tópicos. La Sección 4, brinda una descripción detallada del proceso principal que permite analizar la influencia de los términos que expresan sentimientos en la detección automática de tópicos en datos textuales de redes sociales. Seguidamente la Sección 5 presenta los resultados experimentales. Finalmente se presentan las conclusiones, así como los trabajos futuros derivados de la presente investigación en la Sección 6.

## II. ANTECEDENTES

La detección de tópicos a partir de grandes volúmenes de textos, ha sido un tema ampliamente analizado en la literatura desde varios puntos de vista. Entre ellos destaca el uso de métodos tales como algoritmos de clasificación, Latent Dirichlet Allocation (LDA) y algoritmos de agrupamiento, entre otros. Para el caso de los algoritmos de clasificación es necesario contar con un conjunto de datos de entrenamiento que permita entrenar el clasificador, mientras tanto LDA y los algoritmos de agrupamiento no resulta necesario contar con un corpus previamente clasificado.

Son muchos los trabajos que podemos encontrar relacionados con la detección de tópicos mediante el uso de algoritmos de agrupamiento jerárquico supervisados y semi-supervisados, no así para los no supervisados. Tales son los casos propuestos en [4]–[6], donde los autores proponen enfoques basados en el uso de información experta, para de esta forma mejorar los resultados en la detección de los principales tópicos.

Desde el punto de vista no supervisado, en [1] se presenta una propuesta para la detección automática de tópicos en datos textuales basada en ontologías. Se utilizó el recurso léxico WordNet Domains [7] con el fin de homogeneizar la representación sintáctica de los conceptos presentes en los textos y así reducir considerablemente la dimensionalidad del problema. Para la experimentación se empleó el conjunto de datos Reuters-21578 el cual contiene textos relacionados con publicaciones de noticias. Los resultados muestran la viabilidad de la propuesta, donde los valores del Coeficiente de Silueta son mejores cuando se aplica la metodología propuesta.

Se debe destacar que aunque los textos de Reuters-21578 constituyen datos reales, son textos largos, están bien elaborados y además pertenecen a un dominio restringido. Si tenemos en cuenta que los textos presentes en redes sociales son textos cortos y los usuarios principalmente expresan sus sentimientos sobre un tema determinado, dichos textos deberían ser preprocesados de forma diferente con el fin de extraer los tópicos presentes.

En redes sociales, la detección de tópicos ha sido extensamente utilizada para el análisis de datos textuales. Muchas han sido las soluciones que han aparecido para el análisis textual en redes sociales, tales como el análisis de sentimientos [8], el filtrado de contenidos [9], [10], el modelado de los

intereses del usuario [11], así como el seguimiento de eventos de interés [4], [12]. En [13] se realiza una comparación entre el contenido de los textos de Twitter con un medio de comunicación tradicional, el New York Times. Para ello se utiliza el modelado de tópicos sin supervisión utilizando el modelo Twitter-LDA, para descubrir dichos tópicos en mensajes cortos.

Por otra parte, son varios los trabajos que presentan un modelo donde se fusionan la detección y análisis de tópicos con el análisis de sentimientos [14]–[18]. En todos ellos la detección de tópicos se lleva a cabo sin tener en cuenta la influencia que tienen los términos con determinada orientación sentimental en dicha tarea.

En el presente trabajo se analiza la influencia de los términos con orientación sentimental en la detección de tópicos en redes sociales. Para ello se aplica un filtro durante el preprocesamiento semántico con el objetivo de eliminar los términos que expresan sentimientos, los cuales pueden introducir ruido en la detección de tópicos. Esta propuesta es totalmente novedosa ya que a diferencia de los trabajos mencionados donde se fusiona la detección de tópicos con el análisis de sentimientos, en este caso lo que se hace es descartar los términos de sentimientos, permaneciendo sólo los términos que aportan información útil para la detección automática de los principales tópicos.

## III. DESCRIPCIÓN DE LA PROPUESTA PARA LA DETECCIÓN DE TÓPICOS

Como se mencionó anteriormente, en este artículo se analiza la influencia de los términos con orientación sentimental en la detección automática de tópicos en redes sociales. Teniendo en cuenta que los datos textuales de redes sociales se encuentran escritos de una manera más coloquial, y los usuarios tienden a expresar sus sentimientos y opiniones sobre determinados productos, servicios, entidades, atributos de estos, etc., resulta útil detectar y eliminar aquellos términos con una determinada orientación sentimental, ya que dichos términos no aportan información útil para la detección de tópicos.

A continuación, se presenta un resumen de cada una de las fases del sistema propuesto, para más detalles ver [1]. Para el caso específico de la fase de Preprocesamiento semántico, se ha resaltado el filtro que permite identificar y descartar los términos de sentimientos, el cual constituye el principal aporte del presente trabajo, y será explicado detalladamente en la Sección IV.

### *Preprocesamiento sintáctico*

Una de las fases fundamentales en la detección de tópicos es el preprocesamiento sintáctico, el cual consiste en una limpieza sintáctica donde se aplican filtros a los datos textuales para facilitar su procesamiento automático. Primero, son ejecutados los procesos de etiquetado de categoría gramatical y de reconocimiento de entidades, dichos procesos se realizan con las herramientas Stanford POS [19] y Stanford NER respectivamente [20]. Luego se aplica el filtro de tokenización y los filtros necesarios para eliminar los términos que pertenecen al conjunto de palabras vacías, los que no son identificados



como sustantivos por el etiquetador gramatical, los que son identificados como sustantivos propios por el identificador de entidades, así como los aquellos términos que no se encuentren en la base de conocimiento Multilingual Central Repository (MCR) [21], ya que todos ellos no aportan información útil para la detección de los tópicos.

#### *Preprocesamiento semántico*

Una vez que los textos han sido preprocesados sintácticamente, se procede con el análisis semántico. En nuestro caso, el objetivo del preprocesamiento semántico, es el de homogeneizar la representación sintáctica de los conceptos presentes en el texto. Lo que se hace es sustituir las etiquetas de WordNet Domains [7] con las que han sido etiquetados los sentidos de WordNet por los términos presentes en los textos originales. Como ya se ha mencionado, en este trabajo se desea analizar la influencia de los términos con orientación sentimental en detección de tópicos, por tal motivo en la Sección IV, se explica en profundidad el proceso relacionado con la identificación y eliminación de los términos con orientación sentimental.

#### *Agrupamiento Jerárquico*

Una vez homogeneizados los textos, se procede a realizar el agrupamiento jerárquico de los textos a partir de las etiquetas de WordNet Domain. Para representar las características se ha utilizado el enfoque propuesto en [1]. En este artículo sólo se analizará el algoritmo de agrupamiento jerárquico Complete Link utilizando como medida de similitud la distancia del coseno.

#### *Etiquetado de grupos*

Cuando termina la fase de agrupamiento, se realiza el proceso de selección de etiquetas de los grupos, la cual constituye una tarea de gran importancia, sobre todo en aplicaciones relacionadas con el análisis de datos, donde el usuario final necesita conocer de qué trata determinado grupo [22]. En el presente trabajo se ha utilizado la Media Aritmética para determinar las etiquetas más relevantes de cada grupo de textos.

### IV. INFLUENCIA DE LOS SENTIMIENTOS EN LA DETECCIÓN DE TÓPICOS

En esta sección se describe el principal aporte del presente trabajo, que consiste en detectar y descartar los términos con orientación sentimental (positiva o negativa) con el fin de mejorar la detección de tópicos.

#### *IV-A. Análisis de los términos de sentimientos en la detección de tópicos*

Para detectar los términos con orientación sentimental, se han utilizado los recursos SentiWordNet [2] y WordNet Affect [3], los cuales están basados en WordNet y permiten determinar si un término en un contexto determinado expresa algún tipo de sentimiento. El primer paso sería desambiguar el término, así de esta forma se conoce el verdadero significado

del término en cuestión y finalmente determinar si tiene o no orientación sentimental.

Se debe resaltar que la idea del filtro aplicado para detectar los términos que expresan sentimientos, es totalmente novedosa, ya que permite separar los términos con información relevante para la detección de tópicos, de aquellos términos vinculados con algún tipo de sentimiento.

*SentiWordNet* : Es un recurso léxico creado especialmente para tareas relacionadas con la clasificación de sentimientos, así como en aplicaciones basadas en la minería de opinión [2]. Constituye una versión mejorada de SentiWordNet 1.0 [23] y se encuentra disponible públicamente para propósitos de investigación. SentiWordNet es el resultado de asignar a todos los sentidos de WordNet tres valores numéricos que indican el valor de polaridad (positivo, negativo y neutro), y dichos valores están en el rango [0,1] [2].

En nuestro caso, una vez desambiguado cada término, determinamos el valor positivo y negativo asignado en SentiWordNet al sentido correspondiente para cada uno de los términos analizados. En caso de que el sentido presente un valor positivo mayor que cero o un valor negativo mayor que cero, este término queda totalmente descartado y no se tiene en cuenta para el posterior análisis mediante el cual se detectan los principales tópicos abordados en los textos.

*WordNet Affect*: Es un recurso lingüístico para la representación léxica del conocimiento afectivo. Al igual que SentiWordNet, está basado en WordNet. Fue creado mediante la selección y etiquetado de sentidos de WordNet, que representan conceptos afectivos, y luego fue extendido mediante el uso de las reacciones entre términos y conceptos presentes en WordNet [3].

A diferencia de SentiWordNet, WordNet Affect no presenta la polaridad de los distintos conceptos, en su lugar etiqueta los sentidos con un conjunto de 1,903 categorías que constituyen estados afectivos. Actualmente consta de 2,874 sentidos y 4,787 términos [3]. La forma en la que ha sido utilizado es similar al anterior, cuando los términos son desambiguados, son descartados si su significado en el contexto actual ha sido etiquetado con algún estado mental en WordNet Affect.

De forma general, cuando los términos que expresan sentimientos son descartados, independientemente del recurso léxico utilizado, los términos más frecuentes que pertenecen a un tópico determinado son más afines al mismo. Además se debe señalar que son muchos los textos que tras aplicar el filtrado de sentimientos no pueden ser procesados, ya que son descartados todos sus términos. Esto se debe también en gran medida a que la gran mayoría de los textos de redes sociales solamente expresan opiniones sobre cierta temática.

### V. EXPERIMENTOS

A continuación demostraremos de forma experimental la validez de nuestra propuesta. Como se ha mencionado, no contamos con información previa de los tópicos presentes en los textos (categorías o etiquetas), por tal motivo debemos utilizar una medida no supervisada. En este caso se ha seleccionado el

Coefficiente de Silueta [24], que permite determinar la cantidad de grupos para la que el algoritmo brinda un mejor resultado.

Los conjuntos de datos para evaluar el sistema pertenecen a dos redes sociales Twitter y Dreamcatchers. Lo primera es una de las redes sociales más populares y de las más utilizadas en investigaciones relacionadas con el tema. La segunda ha sido desarrollada bajo un enfoque colaborativo entre sus miembros, y se cuenta con la base de datos que le da soporte. Los datos seleccionados de Twitter y Dreamcatchers se encuentran en inglés y español respectivamente, demostrando que la propuesta es independiente del idioma.

#### V-A. Conjuntos de datos

Se han seleccionado cuatro conjuntos de datos Tabla I, pertenecientes a Twitter y a la Red Social Dreamcatchers. Los datos de Twitter se obtuvieron del conjunto de entrenamiento de Sentiment140<sup>1</sup>, se encuentran en formato CSV y consta de seis campos, entre ellos el texto de los *tweets* el cual será utilizado en el presente trabajo. Se debe mencionar que se seleccionaron estos datos por estar orientados al Análisis de Sentimientos, y constituyen una fuente de gran importancia para la experimentación.

Por otra parte, se cuenta con la base de datos de Dreamcatchers, con un total de 61 tablas. La información recogida es toda la relacionada con los datos personales y de afiliación del usuario, así como, las interacciones que realiza en su perfil y con otros usuarios. En este caso utilizaremos los *Comentarios* de los usuarios.

Se debe mencionar que además del idioma, los textos de Twitter difieren con los de Dreamcatchers en los siguientes aspectos:

- Los textos de Dreamcatchers aunque abordan diversos tópicos, principalmente constituyen temas relacionados con el ámbito universitario, ya que los usuarios de dicha red social son estudiantes de la Universidad de Camagüey.
- La longitud de los textos de Twitter que hemos utilizado está restringida a 140 caracteres, mientras los textos de Dreamcatchers no presentan ninguna restricción.

#### V-B. Evaluación

En esta sección, se explica el procedimiento para evaluar el funcionamiento de la metodología para la detección automática de tópicos aplicando el filtro para eliminar los términos que expresan sentimientos con los distintos recursos y sin aplicarlo.

Para ello se ha utilizado como medida el Coeficiente de Silueta [24] 1. Esta medida permite analizar la calidad de los grupos creados por los algoritmos de agrupamiento jerárquico. Los valores de este coeficiente están en el intervalo de [-1;1], siendo 1 el valor ideal que deben alcanzar los distintos algoritmos de agrupamiento jerárquico.

$$S(i) = \frac{a(i) - b(i)}{\max\{a(i), b(i)\}} \quad (1)$$

<sup>1</sup><http://www.sentiment140.com/>

La experimentación se ha llevado a cabo utilizando el método Complete Link y realizando cortes para las siguientes cantidades de grupos (17, 25, 40, 60, 80, 100 y 120), y para cada caso se ha determinado el valor del Coeficiente de Silueta cuando no se aplica el filtro para eliminar los sentimientos, así como cuando es aplicado haciendo uso de los distintos recursos. En todos los casos se ha utilizado como medida de similitud la distancia del coseno.

#### V-C. Resultados y discusión

En las Tablas II-V se muestran los valores del Coeficiente de Silueta de cada conjunto experimental y las distintas cantidades de grupos, tanto cuando no se aplica el filtro para eliminar los términos que expresan sentimientos (**NF**) y cuando es aplicado utilizando SentiWordNet (**SWN**), WordNet Affect (**WA**) and la combinación de ambos (**SWN-WA**). Se puede observar a simple vista que cuando se aplica el filtrado de sentimientos, se obtienen mejores resultados que cuando no es aplicado (valores en negrita).

Con el fin de realizar un análisis más detallado de los resultados, se ha realizado un análisis estadístico para así poder determinar si existen diferencias significativas entre los valores obtenidos para los distintos número de grupos, los recursos utilizados y la fuente de los datos textuales. Las Figuras 1 y 2 muestran las gráficas del Coeficiente de Silueta con respecto a la cantidad de grupos y el recurso utilizado para detectar los términos con orientación sentimental respectivamente, utilizando en ambos casos la prueba de Kruskal-Wallis [25]. A continuación se resumen las conclusiones del análisis:

- A partir de 60 grupos en adelante los valores del Coeficiente de Silueta se estabilizan, mostrando diferencias significativas con las cantidades anteriores.
- Existe una notable diferencia entre los resultados obtenidos con los recursos **SWN** y **SWN-WA**, y los obtenidos con el recurso **WA** y cuando no se aplica ningún filtro para detectar los términos con orientación sentimental, aunque no existen diferencias significativas entre **SWN** y **SWN-WA** si tenemos en cuenta que la utilización de **SWN-WA** conlleva el uso de otro recurso léxico.

Por otra parte, la Figura 3 muestra la gráfica del Coeficiente de Silueta con respecto a la red social de la cual provienen los textos. Se realizó la prueba de Wilcoxon [26] y se puede concluir que existe una gran diferencia para las redes sociales utilizadas, pues Dreamcatchers brinda los mejores resultados. Esto se debe en gran medida a que gran parte de los usuarios de esta red pertenecen a un contexto universitario, por lo que el dominio de conversación es más restringido.



Tabla I  
DESCRIPCIÓN DE LOS CONJUNTOS DE DATOS

Conjunto	Cantidad de documentos	Fuente	Idioma	Intervención	Cantidad de términos diferentes	Cantidad total de términos
Conjunto 1	5000	Twitter	Inglés	Tweet	3189	12915
Conjunto 2	10000	Twitter	Inglés	Tweet	4597	25634
Conjunto 3	5000	Dreamcatchers	Español	Comentario	1661	8851
Conjunto 4	10000	Dreamcatchers	Español	Comentario	2218	17141

Tabla II  
COEFICIENTE DE SILUETA DEL CONJUNTO 1

Recurso/ Grupos	17	25	40	60	80	100	120
NF	0.0	0.09	0.19	0.23	0.24	0.24	0.22
SWN	0.03	0.1	0.2	0.26	<b>0.28</b>	0.27	0.27
WA	0.0	0.09	0.18	0.22	0.23	0.24	0.22
SWN-WA	0.02	0.09	0.21	0.26	0.27	0.26	0.27

Tabla III  
COEFICIENTE DE SILUETA DEL CONJUNTO 2

Recurso/ Grupos	17	25	40	60	80	100	120
NF	0.04	0.09	0.14	0.21	0.22	0.23	0.22
SWN	0.06	0.09	0.15	0.25	0.26	<b>0.27</b>	0.26
WA	0.05	0.08	0.12	0.21	0.23	0.23	0.22
SWN-WA	0.06	0.08	0.15	0.24	<b>0.27</b>	<b>0.27</b>	0.25

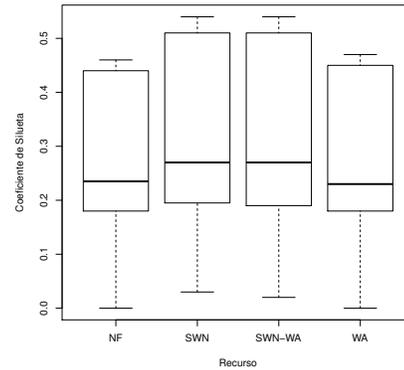


Figura 2. Gráfica entre el Coeficiente de Silueta y los recursos utilizados

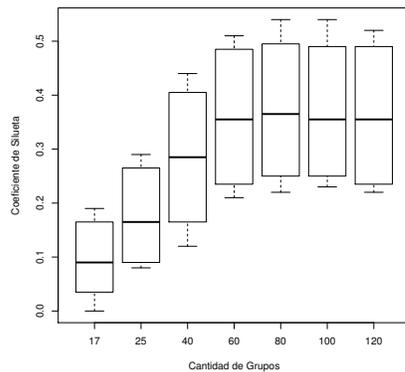


Figura 1. Gráfica entre el Coeficiente de Silueta y el número de grupos

Tabla IV  
COEFICIENTE DE SILUETA DEL CONJUNTO 3

Recurso/ Grupos	17	25	40	60	80	100	120
NF	0.13	0.24	0.37	0.45	0.45	0.44	0.44
SWN	0.19	0.29	0.44	0.51	0.52	0.52	0.52
WA	0.12	0.23	0.37	0.45	0.46	0.45	0.45
SWN-WA	0.16	0.27	0.44	0.51	<b>0.53</b>	0.51	0.52

Tabla V  
COEFICIENTE DE SILUETA DEL CONJUNTO 4

Recurso/ Grupos	17	25	40	60	80	100	120
NF	0.17	0.24	0.36	0.45	0.46	0.45	0.45
SWN	0.15	0.28	0.43	0.51	<b>0.54</b>	<b>0.54</b>	0.52
WA	0.18	0.26	0.38	0.46	0.47	0.47	0.46
SWN-WA	0.17	0.28	0.44	0.51	<b>0.54</b>	0.53	0.52

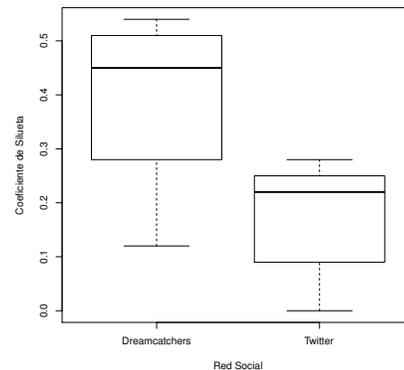


Figura 3. Gráfica entre el Coeficiente de Silueta y las redes sociales utilizadas

## VI. CONCLUSIONES Y TRABAJOS FUTUROS

En este artículo se presenta una nueva propuesta para la detección automática de tópicos en textos de redes sociales. Para ello se ha incorporado un filtro durante el preprocesamiento semántico de los textos, permitiendo detectar y eliminar los términos que presenten una orientación sentimental, ya que estos no constituyen información relevante para la detección de tópicos. Con este fin, fueron utilizados los recursos SentiWordNet y WordNet Affect por separado y combinándolos.

Los experimentos realizados tanto con Twitter como con Dreamcatchers, permiten mostrar la viabilidad del sistema. Se ha experimentado sin aplicar el filtro para eliminar los sentimientos y aplicándolo con dos recursos léxicos relacionados

con el análisis de sentimientos. En cada caso se realizaron cortes en las cantidades de grupos ya mencionadas y se calculó el Coeficiente de Silueta. Los resultados alcanzados cuando se aplicó el filtro mejoran los resultados obtenidos cuando no se aplica el filtro. El recurso con el que se obtuvo un mejor rendimiento fue SentiWordNet, pues aunque la combinación de ambos recursos mejora a SentiWordNet en determinados casos, las diferencias no son significativas teniendo en cuenta que se incorporan todos los conceptos de WordNet Affect.

El uso de algoritmos de agrupamiento jerárquico, brinda la posibilidad de agrupar los documentos en tópicos o temas y crear una jerarquía de tópicos, la cual puede ser utilizada como jerarquía de una dimensión en un modelo multidimensional y de esta forma facilitar el análisis de los datos de las redes sociales. La metodología propuesta puede ser aplicada a distintas redes sociales independientemente del idioma y de los temas tratados en dichas redes, gracias a la base de conocimiento utilizada (MCR). Se debe mencionar que MCR está basado en el recurso WordNet, integra recursos como WordNet Domains, Base Concepts, Top Ontology y la ontología AdimenSUMO y además mediante un índice (Inter-Lingual-Index ILI) integra wordnets de seis idiomas diferentes: Inglés, Español, Catalán, Euskera, Gallego y Portugués [21].

Como se mencionó anteriormente el tema tratado en la presente investigación deja abierta la posibilidad para realizar extensiones tales como incluir la jerarquía de tópicos obtenida como una nueva dimensión en un modelo multidimensional, permitiendo un mejor análisis de los datos principalmente en redes sociales, y luego de tener creados los tópicos estudiar el proceso relacionado con la de detección automática de tópicos para un nuevo texto.

#### REFERENCIAS

- [1] K. Gutiérrez-Batista, J. R. Campaña, M.-A. Vila, y M. J. Martín-Bautista, "An ontology-based framework for automatic topic detection in multilingual environments," *International Journal of Intelligent Systems*, vol. 33, no. 7, pp. 1459–1475, 2018. [Online]. Disponible: <https://onlinelibrary.wiley.com/doi/abs/10.1002/int.21986>
- [2] S. Baccianella, A. Esuli, y F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining." in *LREC*, vol. 10, 2010, pp. 2200–2204.
- [3] R. Valitutti, "Wordnet-affect: an affective extension of wordnet," in *Proceedings of the 4th International Conference on Language Resources and Evaluation*, 2004, pp. 1083–1086.
- [4] L. Chung-Hong, "Unsupervised and supervised learning to evaluate event relatedness based on content mining from social-media streams," *Expert Systems with Applications*, vol. 39, no. 18, pp. 13 338 – 13 356, 2012.
- [5] A. G. Skarmeta, A. Bensaid, y N. Tazi, "Data mining for text categorization with semi-supervised agglomerative hierarchical clustering," *International Journal of Intelligent Systems*, vol. 15, no. 7, pp. 633–646, 2000. [Online]. Disponible: [http://dx.doi.org/10.1002/\(SICI\)1098-111X\(200007\)15:7<633::AID-INT4>3.0.CO;2-8](http://dx.doi.org/10.1002/(SICI)1098-111X(200007)15:7<633::AID-INT4>3.0.CO;2-8)
- [6] L. Zheng y T. Li, "Semi-supervised hierarchical clustering," in *Proceedings of the 2011 IEEE 11th International Conference on Data Mining*, ser. ICDM '11. Washington, DC, USA: IEEE Computer Society, 2011, pp. 982–991. [Online]. Disponible: <http://dx.doi.org/10.1109/ICDM.2011.130>
- [7] B. Magnini y G. Cavaglia, "Integrating subject field codes into wordnet," in *LREC*. European Language Resources Association, 2000.
- [8] C. Lin y Y. He, "Joint sentiment/topic model for sentiment analysis," in *18th ACM Conference on Information and Knowledge Management (CIKM09)*. New York, NY, USA: ACM, 2009, pp. 375–384.
- [9] J. Duan y J. Zeng, "Web objectionable text content detection using topic modeling technique," *Expert Systems with Applications*, vol. 40, pp. 6094–6104., 2013.
- [10] J. Martínez-Romo y L. Araujo, "Detecting malicious tweets in trending topics using a statistical analysis of language," *Expert Systems with Applications*, vol. 40, pp. 2992–3000, 2013.
- [11] M. Pennacchiotti y S. Gurumurthy, "Investigating topic models for social media user recommendation," in *20th International Conference Companion on World Wide Web*. New York, NY, USA: ACM, 2011, pp. 101–102.
- [12] J. Wu, W. Gao, B. Zhang, J. Liu, y C. Li, "Cluster based detection and analysis of internet topics," in *4th International Symposium on Computational Intelligence and Design, ISCID 2011*, vol. 2, 2011, pp. 371–374.
- [13] W. X. Zhao, J. Weng, J. He, E.-P. Lim, y H. Yan, "Comparing twitter and traditional media using topic models," in *33rd European conference on advances in information retrieval (ECIR11)*. Berlin, Heidelberg: Springer-Verlag., 2011, pp. 338–349.
- [14] C. Lin, Y. He, R. Everson, y S. Ruger, "Weakly supervised joint sentiment-topic detection from text," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 24, no. 6, pp. 1134–1145, June 2012.
- [15] J. Sowmiya y S. Chandrakala, "Joint sentiment/topic extraction from text," 2015, pp. 611–615, cited By 0.
- [16] Y. Rao, Q. Li, X. Mao, y L. Wenyin, "Sentiment topic models for social emotion mining," *Information Sciences*, vol. 266, pp. 90 – 100, 2014. [Online]. Disponible: <http://www.sciencedirect.com/science/article/pii/S002002551400019X>
- [17] K. Cai, S. Spangler, Y. Chen, y L. Zhang, "Leveraging sentiment analysis for topic detection," in *Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT '08. IEEE/WIC/ACM International Conference on*, vol. 1, Dec 2008, pp. 265–271.
- [18] X. Ding, L. Zhang, Y. Tian, X. Gong, y W. Wang, "Dynamic topic detection model by fusing sentiment polarity," vol. 159, 2015, pp. 65–71, cited By 0.
- [19] K. Toutanova, D. Klein, C. D. Manning, y Y. Singer, "Feature-rich part-of-speech tagging with a cyclic dependency network," in *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1*, ser. NAACL '03. Stroudsburg, PA, USA: Association for Computational Linguistics, 2003, pp. 173–180. [Online]. Disponible: <http://dx.doi.org/10.3115/1073445.1073478>
- [20] J. R. Finkel, T. Grenager, y C. Manning, "Incorporating non-local information into information extraction systems by gibbs sampling," in *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, ser. ACL '05. Stroudsburg, PA, USA: Association for Computational Linguistics, 2005, pp. 363–370. [Online]. Disponible: <http://dx.doi.org/10.3115/1219840.1219885>
- [21] A. Gonzalez-Agíre, E. Laparra, y G. Laparra, "Multilingual central repository version 3.0," in *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. Istanbul, Turkey: European Language Resources Association (ELRA), may 2012.
- [22] C. D. Manning, P. Raghavan, y H. Schütze, *Introduction to Information Retrieval*. New York, NY, USA: Cambridge University Press, 2008.
- [23] A. Esuli y F. Sebastiani, "Sentiwordnet: A publicly available lexical resource for opinion mining," in *In Proceedings of the 5th Conference on Language Resources and Evaluation (LREC06)*, 2006, pp. 417–422.
- [24] P. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, no. 1, pp. 53–65, Nov. 1987. [Online]. Disponible: [http://dx.doi.org/10.1016/0377-0427\(87\)90125-7](http://dx.doi.org/10.1016/0377-0427(87)90125-7)
- [25] W. Kruskal y W. Wallis, "Use of ranks in one-criterion variance analysis," *Journal of the American Statistical Association*, pp. 583–621, 1952.
- [26] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.



# Un Algoritmo Memético, con búsqueda local basada en Label Propagation, para detectar comunidades en redes dinámicas

1<sup>st</sup> Ángel Panizo

*Departamento de ciencias de la computación  
Universidad Autónoma de Madrid  
Madrid, España  
angel.panizo@uam.es*

2<sup>nd</sup> Gema Bello-Orgaz

*Departamento de ciencias de la computación  
Universidad Autónoma de Madrid  
Madrid, España  
gema.bello@uam.es*

3<sup>rd</sup> Alfonso Ortega

*Departamento de ciencias de la computación  
Universidad Autónoma de Madrid  
Madrid, España  
alfonso.ortega@uam.es*

4<sup>th</sup> David Camacho

*Departamento de ciencias de la computación  
Universidad Autónoma de Madrid  
Madrid, España  
david.camacho@uam.es*

**Abstract**—El análisis y la detección de comunidades en redes complejas es actualmente un área de estudio en auge, ya que muchos sistemas se pueden representar como redes de nodos interconectados. Tradicionalmente el esfuerzo se ha aplicado en estudiar métodos para analizar redes estáticas, es decir, redes que no cambian en el tiempo. En el mundo real, a menudo, las redes son dinámicas, es decir, evolucionan según pasa el tiempo, debido a este fenómeno la comunidad científica se ha comenzado a interesar por el análisis de este tipo de redes. En la literatura ya se han usado con éxito tanto algoritmos genéticos como otros algoritmos bio-inspirados para detectar comunidades en redes dinámicas, pero pocos artículos se han centrado en buscar mecanismos para mejorar estos métodos explotando la temporalidad de estas redes. En este artículo presentamos un Algoritmo Memético, con búsqueda local basada en *Label Propagation*, para identificar comunidades en redes dinámicas.

**Index Terms**—detección dinámica de comunidades, algoritmos genéticos, análisis de redes complejas, algoritmos meméticos

## I. INTRODUCCIÓN

El análisis y la detección de comunidades en redes complejas es actualmente un área de estudio en auge, ya que muchos sistemas complejos se pueden representar como redes de nodos interconectados. Esta idea se ha aplicado de forma satisfactoria a una gran cantidad áreas de estudio: como el marketing [1], la salud pública [2], el cibercrimen [3]–[5] o el análisis de redes sociales [6]–[8]. La detección de comunidades es un problema *mal definido*, ya que existen varias definiciones de qué es una comunidad y no se ha llegado a un consenso entre la comunidad científica para decidir cuál de ellas es la correcta. Nosotros definiremos una comunidad

Este trabajo ha sido cofinanciado por los siguientes proyectos de investigación: EphemCH (TIN2014-56494-C4-4-P) y DeepBio (TIN2017-85727-C4-3-P). Ministerio de Economía y Competitividad de España, bajo los fondos Europeos de desarrollo regional (FEDER).

como un conjunto de nodos que interactúan más entre ellos que entre el resto de nodos de la red. Además, podemos definir dos tipos de comunidades, *solapadas* y *no-solapadas*, dependiendo del número de comunidades a las que un nodo puede pertenecer a la vez. Si un nodo puede pertenecer a varias comunidades a la vez, entonces estaremos hablando de comunidades *solapadas*. Mientras que si un nodo sólo puede pertenecer a una única comunidad, entonces hablaremos de comunidades *no-solapadas*.

Tradicionalmente el esfuerzo se ha aplicado en estudiar métodos para analizar redes estáticas, es decir, redes que no cambian en el tiempo. En el mundo real, a menudo, las redes son dinámicas, es decir, evolucionan según pasa el tiempo. Debido a este fenómeno la comunidad científica se ha comenzado a interesar por el análisis de este tipo de redes. Una red dinámica se puede modelar usando una secuencia de instancias, dónde cada instancia representa el estado de la red en un momento concreto. Las comunidades de una red dinámica son conocidas como *comunidades dinámicas*. Detectar comunidades dinámicas tiene una dificultad añadida si lo comparamos con sus homologas estáticas. No sólo es necesario agrupar cada instancia de la red en comunidades, sino que también hay que rastrear estas comunidades a lo largo de las diferentes instancias de la red.

Los métodos de detección de comunidades dinámicas se pueden dividir en dos grupos. El primer grupo incluye a aquellos métodos que hacen el paso de agrupación y el paso de rastreo de manera independiente, por ejemplo, primero se agrupa cada instancia de la red en comunidades y después se analizan los cambios que ha sufrido estas comunidades a lo largo de las distintas instancias. El segundo grupo, incluye los métodos que siguen la filosofía del *evolutionary clustering* [9], que consiste en agrupar cada instancia de la

red de tal manera que las comunidades encontradas en una instancia se parezcan, en la mayor medida de lo posible, a las comunidades encontradas en la instancia anterior. Esta filosofía asume que no se dan cambios abruptos en la red en un periodo corto de tiempo. Por motivos de claridad, a partir de ahora, a los métodos del primer grupo los llamaremos *métodos temporalmente suaves* y a los métodos del segundo grupo *métodos temporalmente no-suaves*.

Los Algoritmos Genéticos (GAs) son una meta-heurística inspirada en los principios de la selección natural de Charles Darwin y en la genética mendeliana. Los GAs son un método efectivo de resolver problemas de optimización combinatoria y se han usado anteriormente para detectar comunidades tanto estáticas [10], [11] como dinámicas [12]. También se han usado Algoritmos Meméticos para detectar comunidades estáticas. Un Algoritmo Memético es un Algoritmo genético que se ha emparejado con un método de *búsqueda local*. Los métodos de *búsqueda local* se han aplicado principalmente durante la creación de la población inicial [13] o durante el proceso de mutación [14]. En este artículo presentamos un Algoritmo Memético que usa un proceso de *búsqueda local* basado en *Label Propagation* [15] para detectar comunidades dinámicas, *no solapadas* y siguiendo un *método temporalmente no-suave*. En el método de *Label Propagation* cada nodo, de manera iterativa, cambia su comunidad para pertenecer a la misma comunidad que la mayoría de sus vecinos. El algoritmo termina cuando en una iteración ningún nodo ha cambiado su comunidad. La idea principal de nuestro algoritmo es utilizar las comunidades detectadas en una instancia de la red para guiar la búsqueda de la instancia siguiente. Nuestro método ejecuta un GAs para cada instancia de la red dinámica. La población de cada GAs se construye utilizando la población obtenida como solución para la instancia anterior. Debido a los cambios en la red, es posible que algunos de los individuos dejen de ser soluciones válidas cuando se pasan de una instancia a otra. Para solucionar este problema proponemos usar una búsqueda local basada en *Label Propagation* para *reparar* estos individuos intentando mantener la mayor cantidad de información posible. Este trabajo se centra en la agrupación de las diferentes instancias de tiempo en comunidades, dejando el posterior rastreo fuera del ámbito de este.

La principal contribución de este trabajo es la presentación y el análisis de la viabilidad de un método de búsqueda local para mejorar la eficiencia de un GAs capaz de detectar comunidades dinámicas.

El resto del artículo está organizado de la siguiente manera: en la sección II se describe la detección dinámica de comunidades de una manera formal y se presenta el Algoritmo Memético propuesto. En la sección III se presentan los procedimientos seguidos para comprobar la viabilidad y la eficiencia del método usando un *dataset* real. Por último, se muestran las conclusiones obtenidas y se proponen futuras líneas de investigación.

## II. ALGORITMO

### A. Formulación del problema

Sea  $\aleph = \{G^0..G^n\}$  una red dinámica con  $n$  instancias ordenadas en orden cronológico. Sea  $G^t$  una instancia de dicha red modelada como un grafo  $G^t(V^t, E^t)$ . Sea  $V^t$  el conjunto de vértices o nodos presentes en el instante  $t$ . Sea  $E^t$  un conjunto de conexiones, llamadas aristas, que conectan dos elementos de  $V^t$  en un instante concreto de tiempo  $t$ . Nuestro objetivo es agrupar los vértices de cada instancia de la red dinámica  $G^t$  de tal manera que los vértices que pertenezcan a un mismo grupo tengan más conexiones entre ellos que con el resto de vértices de  $G^t$  y además cada vértice pertenezca a un único grupo en cada instancia de tiempo.

### B. Algoritmo Memético

El pseudocódigo del Algoritmo Memético propuesto está disponible en el *Algoritmo 1*. Usamos un Algoritmo Genético con elitismo para detectar las comunidades en cada instancia,  $G^t$ , de la red dinámica  $\aleph$  (líneas 1-13). Para la primera instancia de la red,  $G^0$ , el método evoluciona una población generada aleatoriamente. Sin embargo, para el resto de las instancias,  $G^t | t > 0$ , en vez de evolucionar una población aleatoria, se evoluciona una población generada a partir de la población obtenida como solución para la instancia anterior. Una vez procesadas todas las instancias  $G^t$  el resultado devuelto por el algoritmo ( $C_{dinamica}$ ) es la solución codificada por el mejor individuo de cada una de las poblaciones evolucionadas (línea 13).

Cada individuo de una población codifica una posible agrupación de nodos en comunidades para una instancia  $G^t$  de la red. Los individuos codifican las soluciones usando la codificación *locus-based adjacency* propuesta en [16]. Tal como comentamos en el párrafo anterior, para la primera instancia de la red,  $G^0$ , generamos una población aleatoria (línea 7). Para generar cada individuo de la población aleatoria rellenamos cada uno de sus genes con un vecino aleatorio del nodo que codifica dicho gen. Por el contrario, para el resto de instancias de la red ( $\{G^t | t > 0\}$ ), los individuos de la población inicial se generan aplicando la función *labelBL* a cada uno de los individuos devueltos por el Algoritmo Genético ejecutado para la instancia anterior (líneas 9-11). La función *labelBL* es la función de búsqueda local basada en *Label Propagation* que *traspasa* la información de un individuo de una generación a otra. Se encarga de cambiar los genes de un individuo que remiten a nodos que han dejado de ser vecinos de el nodo que el gen codifica (líneas 16-23). Para seleccionar un nuevo valor para cada uno de estos genes: primero buscamos todos los vecinos del nodo, que codifica el gen afectado, en la instancia actual  $G^t$  (línea 18). A continuación, utilizamos la mejor solución ( $C_{last}$ ) de la instancia anterior  $G^{t-1}$  para encontrar la comunidad a la que pertenecen el mayor número de vecinos (línea 20). Por último, el valor del gen se cambia por un vecino aleatorio que pertenezca a la comunidad seleccionada en el paso anterior (línea 22).



Una vez generada la población inicial, se evoluciona siguiendo un Algoritmo Genético Elistista (línea 12) con las siguientes funciones: *Two Points Crossover*, *Tournament Selection* y *Modularity* [17] como *función de Fitness*. Además, como mutación se ha implementado una función que selecciona una serie de genes al azar y cambia sus valores por vecinos aleatorios del nodo que codifica el gen afectado. Estas funciones se seleccionaron después de revisar los operadores más utilizados en la literatura.

---

**Algorithm 1** Algoritmo Memético para detectar comunidades dinámicas

---

```

1: function ALGORITMOMEMETICO( $\mathbb{N}$ )
2:    $Conf \leftarrow ELITISM, CRXRATE$ 
3:    $Conf \leftarrow Conf \cup MUTR, MAXGEN, NCONV$ 
4:    $C_{dinamica} \leftarrow \emptyset$ 
5:    $P_{ant} \leftarrow null$ 
6:    $C_{ant} \leftarrow null$ 
7:   for all  $G^t \in \mathbb{N}$  do
8:     if  $t = 0$  then
9:        $P_{ini} \leftarrow poblacionAleatoria(G^t, POPSIZE)$ 
10:    else
11:       $P_{ini} \leftarrow \emptyset$ 
12:      for all  $ind \in P_{ant}$  do
13:         $P_{ini} \leftarrow P_{ini} \cup labelBL(ind, G^t, C_{ant})$ 
14:       $P_{ant} \leftarrow elitistaGA(P_{ini}, G^t, Conf)$ 
15:       $C_{ant} \leftarrow mejorSolucion(P_{ant})$ 
16:       $C_{dinamica} \leftarrow C_{dinamica} \cup C_{ant}$ 
17:    return  $C_{dinamica}$ 
18: function LABELBL( $ind, G^t, C_{ant}$ )
19:   for all  $i \in [1, size(ind)]$  do
20:      $vecinos \leftarrow getVecinos(i, G^t)$ 
21:     if  $ind[i] \notin vecinos$  then
22:        $com \leftarrow getComunidadComun(vecinos, C_{ant})$ 
23:        $vecinosOk \leftarrow nodosEnCom(vecinos, com)$ 
24:        $ind[i] \leftarrow seleccionAzar(vecinosOk)$ 
25:   return  $individuo$ 

```

---

### III. EXPERIMENTACIÓN

#### A. Descripción del dataset

En esta sección estudiaremos la efectividad del algoritmo propuesto probándolo contra el *dataset* de *Enron* [18]. Este dataset consta de los correos enviados por los empleados de la corporación Enron entre los años de 1999 y 2003. Cada vértice en el dataset representa a un trabajador de Enron y una arista uniendo dos nodos significa que esos dos empleados mantuvieron una comunicación por e-mail. Cada arista tiene asociada la fecha de cuando se produjo dicha comunicación. Usando estas fechas se ha creado una red dinámica compuesta por ocho instancias. Cada instancia tiene los correos enviados entre empleados de Enron durante quince días, los fines de semana se han excluido debido a la poca actividad que ocurre en ellos. Se han seleccionado quince días como la resolución de cada instancia de la red porque una resolución menor daba

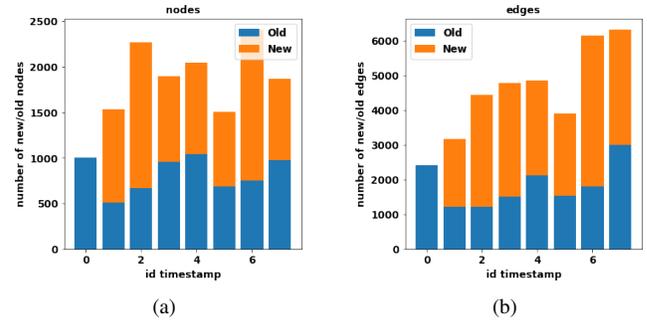


Fig. 1: Número de vértices y aristas en cada instancia. La parte azul de la barra muestra el número de elementos que estaban presentes en la instancia anterior además de en la actual. La parte naranja muestra el número de elementos que están presentes en la estancia actual pero no en la anterior.

TABLE I: Parámetros utilizados en los Algoritmos Genético y Memético

Parámetro	Descripción	Valor
<i>POPSIZE</i>	Tamaño de la población	300
<i>ELITISM</i>	Número de individuos que forman la élite	10%
<i>MAXGEN</i>	Número de máximo de generaciones	500
<i>NCONV</i>	Número de generaciones con el mismo <i>fitness</i> (+/-0.01)	10
<i>CRXRATE</i>	Probabilidad de cruce	1.0
<i>MUTR</i>	Probabilidad de Mutación	0.1

lugar a redes demasiado dinámicas para ser analizadas usando este método. Las ocho instancias que se han seleccionado son consecutivas y empiezan el 1 de Enero del 2000 y terminan el 30 de Abril del mismo año. En la figura 1 se muestran el número de nodos y aristas de cada instancia.

#### B. Resultados Experimentales

Para poder validar nuestra propuesta vamos a comparar nuestro Algoritmo Memético contra un Algoritmo Genético que no usa búsqueda local (para cada instancia de la red evoluciona una población aleatoria). Para evaluar ambos algoritmos compararemos el *fitness* del mejor individuo de cada generación obtenidos por los dos métodos. De esta manera no sólo podremos comparar cuál de los dos algoritmos obtiene un resultado mejor, sino que también podremos valorar cuál de ellos lo obtiene en menos generaciones. Tanto el Algoritmo Memético como el Genético usan la misma configuración que se encuentra disponible en la Tabla I. Los diferentes parámetros se han calibrado de manera manual siguiendo un paradigma de ensayo y error.

Cada experimento se ha ejecutado 30 veces, la mediana de todas las ejecuciones está disponible en las Figuras 2 y 3. Analizando los resultados podemos concluir que nuestro Algoritmo Memético consigue mejores resultados en todas las instancias de la red a excepción de la *instancia 3* (Figura 3a). La primera instancia, la cero (Figura 2a), es un caso especial dado que al no existir una solución previa no se puede

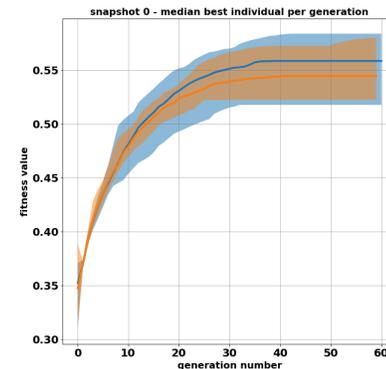
llevar a cabo la búsqueda local y se comporta igual que el Algoritmo Genético, obteniendo ambos resultados similares. Para el resto de instancias, a excepción de la tres, la población obtenida por el Algoritmo Memético es mejor que la población obtenida por el Genético durante toda la ejecución. En la *instancia 3* la población generada por el Mémetico comienza siendo mejor que la población del Genético, pero en la generación 15 este le sobrepasa y se mantiene así hasta el final. En nuestra opinión aquí yace el principal problema de este método: aunque las poblaciones del Memético siempre empiezan siendo mejores que la del genético, la velocidad de convergencia de las primeras son menores que la velocidad de las segundas. Esto se debe a que el método de búsqueda local, usado por el Algoritmo Memético, genera poblaciones con una diversidad menor que el método aleatorio, usado por el Genético, reduciendo las capacidades exploratorias del primero. Este efecto hace que el Algoritmo Memético sea más propenso a quedarse atascado en un mínimo local, justo lo que está pasando en la *instancia 3*.

#### IV. CONCLUSIONES

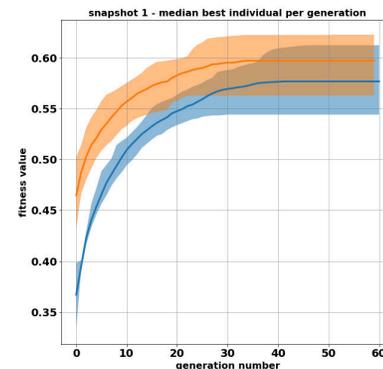
En este artículo hemos presentado un nuevo Algoritmo Memético para detectar comunidades en redes dinámicas. El método ejecuta un Algoritmo Genético elitista para cada una de las instancias que conforman la red dinámica. Entre cada ejecución del Algoritmo Genético se aplica un operador de búsqueda local basado en *Label propagation* que aprovecha la solución anterior para guiar la ejecución actual y obtener mejores resultados. Para probar la validez del método se ha comparado nuestro algoritmo contra un Algoritmo Genético sin búsqueda local usando el *dataset* de Enron. Los experimentos que se han llevado a cabo muestran que el Algoritmo Memético encuentra soluciones de más calidad que el Genético, sin embargo, también muestran que el Algoritmo Memético tiene menos capacidad exploratoria que el Genético y tiene más posibilidades de quedar estancado en un mínimo local, aunque en la mayoría de los casos estudiados esto no ocurrió. Por esta razón, nuestras futuras líneas de investigación se centrarán en estudiar otros métodos de búsqueda local o combinaciones de estos que generen poblaciones que sean a su vez diversas y de calidad. Además, también compararemos nuestro método con otros métodos del estado del arte y otros *dataset*, tanto sintéticos como reales, para determinar con más certeza si el método propuesto es o no prometedor.

#### REFERENCES

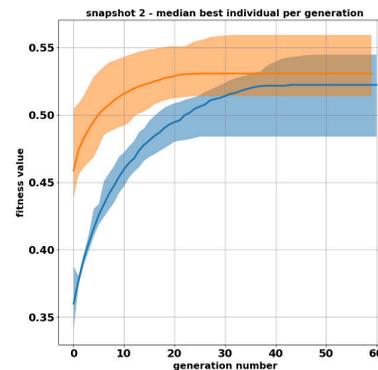
- [1] G. Bello-Orgaz, H. Menéndez, S. Okazaki, and D. Camacho, "Combining social-based data mining techniques to extract collective trends from twitter," *Malaysian Journal of Computer Science*, vol. 27, no. 2, 2014.
- [2] G. Bello-Orgaz, J. Hernandez-Castro, and D. Camacho, "Detecting discussion communities on vaccination in twitter," *Future Generation Computer Systems*, vol. 66, pp. 125–136, 2016.
- [3] R. Lara-Cabrera, A. Gonzalez-Pardo, M. Barhamgi, and D. Camacho, "Extracting radicalisation behavioural patterns from social network data," in *2017 28th International Workshop on Database and Expert Systems Applications (DEXA)*, Aug 2017, pp. 6–10.
- [4] R. Lara-Cabrera, A. Gonzalez-Pardo, and D. Camacho, "Statistical analysis of risk assessment factors and metrics to evaluate radicalisation in twitter," *Future Generation Computer Systems*, 2017.



(a)

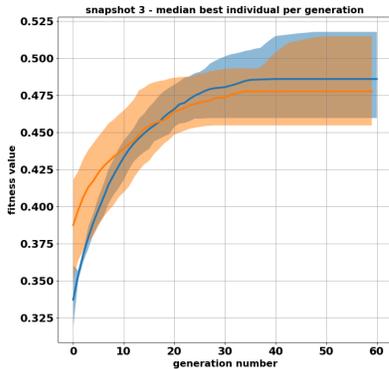


(b)

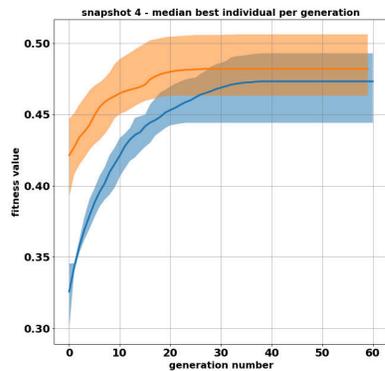


(c)

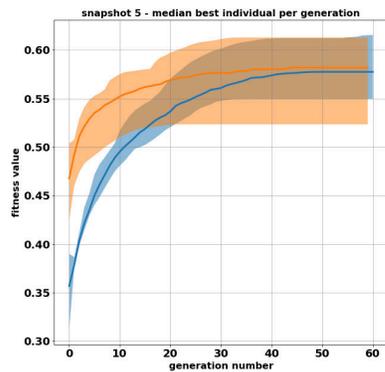
Fig. 2: Cada gráfico corresponde con una instancia de la red. En el eje X se muestra el número de generación. En eje Y el valor de la modularidad de el mejor individuo de la población. La curva naranja muestra la evolución de las poblaciones generadas pos el Algoritmo Memético y la azul la evolución del Algoritmo Genético.



(a)

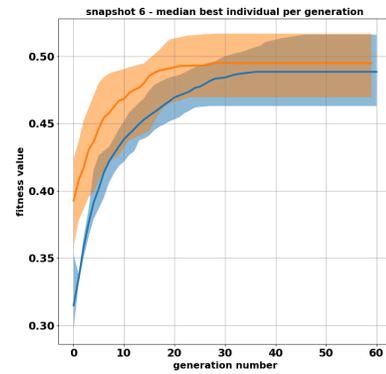


(b)

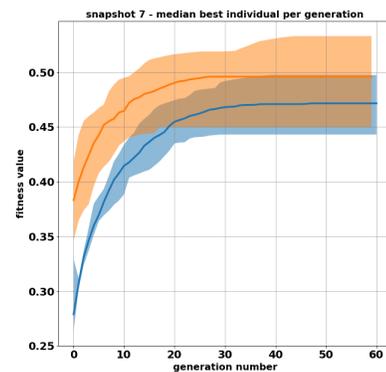


(c)

Fig. 3: Cada gráfico corresponde con una instancia de la red. En el eje X se muestra el número de generación. En eje Y el valor de la modularidad de el mejor individuo de la población. La curva naranja muestra la evolución de las poblaciones generadas pos el Algoritmo Memético y la azul la evolución del Algoritmo Genético.



(a)



(b)

Fig. 4: Cada gráfico corresponde con una instancia de la red. En el eje X se muestra el número de generación. En eje Y el valor de la modularidad de el mejor individuo de la población. La curva naranja muestra la evolución de las poblaciones generadas pos el Algoritmo Memético y la azul la evolución del Algoritmo Genético.

- [5] A. Malm and G. Bichler, "Networks of collaborating criminals: Assessing the structural vulnerability of drug markets," *Journal of Research in Crime and Delinquency*, vol. 48, no. 2, pp. 271–297, 2011.
- [6] A. Gonzalez-Pardo, J. J. Jung, and D. Camacho, "Aco-based clustering for ego network analysis," *Future Generation Computer Systems*, vol. 66, pp. 160–170, 2017.
- [7] R. Lara-Cabrera, A. G. Pardo, K. Benouaret, N. Faci, D. Benslimane, and D. Camacho, "Measuring the radicalisation risk in social networks," *IEEE Access*, vol. 5, pp. 10 892–10 900, 2017.
- [8] J. Del Ser, J. L. Lobo, E. Villar-Rodríguez, M. N. Bilbao, and C. Perfecto, "Community detection in graphs based on surprise maximization using firefly heuristics," in *Evolutionary Computation (CEC), 2016 IEEE Congress on*. IEEE, 2016, pp. 2233–2239.
- [9] D. Chakrabarti, R. Kumar, and A. Tomkins, "Evolutionary clustering," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2006, pp. 554–560.
- [10] G. Bello-Orgaz, H. D. Menéndez, and D. Camacho, "Adaptive k-means algorithm for overlapped graph clustering," *International journal of neural systems*, vol. 22, no. 05, p. 1250018, 2012.
- [11] R. Francisquini, V. Rosset, and M. C. Nascimento, "Ga-lp: A genetic algorithm based on label propagation to detect communities in directed networks," *Expert Systems with Applications*, vol. 74, pp. 127–138, 2017.
- [12] F. Folino and C. Pizzuti, "An evolutionary multiobjective approach for community discovery in dynamic networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 8, pp. 1838–1852, 2014.



- [13] S. B. Mathias, V. Rosset, and M. C. Nascimento, "Community detection by consensus genetic-based algorithm for directed networks," *Procedia Computer Science*, vol. 96, pp. 90–99, 2016.
- [14] S. Li, Y. Chen, H. Du, and M. W. Feldman, "A genetic algorithm with local search strategy for improved detection of community structure," *Complexity*, vol. 15, no. 4, pp. 53–60, 2010.
- [15] U. N. Raghavan, R. Albert, and S. Kumara, "Near linear time algorithm to detect community structures in large-scale networks," *Physical review E*, vol. 76, no. 3, p. 036106, 2007.
- [16] Y. Park and M. Song, "A genetic algorithm for clustering problems," in *Proceedings of the third annual conference on genetic programming*, vol. 1998, 1998, pp. 568–575.
- [17] M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical review E*, vol. 69, no. 2, p. 026113, 2004.
- [18] B. Klimt and Y. Yang, "Introducing the enron corpus." in *CEAS*, 2004.



# Design of an ACO algorithm for Solving Community Finding Problems

Antonio Gonzalez-Pardo  
Computer Science Department  
Universidad Autónoma de Madrid  
Email: antonio.gonzalez@uam.es

David Camacho  
Computer Science Department  
Universidad Autónoma de Madrid  
Email: david.camacho@uam.es

**Abstract**—The amount of data generated by social media users is increasing exponentially mainly produced by the high number of users connected everyday that interacts with each other through the Social Network (SN). As a result, SNs has become an interesting domain for research due to the wide variety of problems that can be solved. Among these problems, this work is focused on Community Finding Problems (CFP) whose goal is to group the different users in several clusters in such a way users belonging to the same cluster are similar (according to a specific metric) whereas they are different from the users of the other clusters. In this work, we describe the algorithm proposed in [1]. This algorithm for CFP is based on Ant Colony Optimization (ACO) algorithm, and it uses the information regarding the topology of the network, i.e. the connections of the users in the SN. For the experimental phase, we have compared the performance of the described algorithm against the performance of some well-known algorithms extracted from the State-of-the-Art. The results reveal that the proposed Topology-based ACO algorithm is a good approach to solve the community finding problem and it provides competitive results against the analyzed algorithms.

## I. INTRODUCTION

Nowadays, Social Networks (SNs) has become a powerful tool for a wide variety of purposes. The main reason lies in the number of users that connects to these SNs everyday. Initially, the goal of any SN is to put people in touch with each other and also to publish different media contents. In this sense, for example Instagram is mainly used for sharing photos, whereas Twitter is used for sharing small text messages called tweets.

Nevertheless, the number of users connected to these SNs generates such quantity of information that its analysis has attracted the attention of both, the research community and the industry. The main reason is the wide variety of applications that can be developed using these data.

In this sense, the most classical problem faced when working with Twitter is to analyze the content of the tweets and perform sentiment analysis over the tweets published by the users [2]. Other works try to identify the key players in a social network [3]. This last application domain is motivated by the usage that radical groups, as terrorist supporters, are using the SNs, such as Islamic State of Iraq and Syria (ISIS), to disseminate their propaganda. Joining both applications

appear works like [4], [5], [6] where authors try to determine the radicalization level of different users based on the content of their tweets. Finally, other works tries to determine the structure hidden by the relations of users in the SNs, like [7], what is called Community Finding Problems (CFPs).

This work provides a summary of our current results presented in [1] on the detection of communities within a given SN. Although there are several valid approaches to perform this task such as Iterated Greedy [7], or Genetic Algorithms [8], in this work the algorithm selected to perform this task has been Ant Colony Optimization (ACO) [9].

## II. THE ACO MODEL FOR CFP

This section provides a short description of the ACO model designed to find the different communities in the SN [1].

Any CFP can be modeled as a Constraint Satisfaction Problem (CSP) where the goal is to assign a cluster to each user of the SN whereas some constraints are satisfied, such as: each cluster must contain at least one user, or users belonging to the same are similar. Traditionally, solving any CSP using an ACO algorithm implies the creation of a decision graph where the ants are executed. This graph is composed by  $X \cdot k$  nodes, where  $X$  is the number of elements of the dataset and  $k$  is the number of clusters. As it can be understood, the main disadvantage of this representation is related to the size of the resulting graph. If the number of nodes increases, it could be difficult to work with the corresponding graph. And also, this representation implies the definition of the number of clusters before execution.

The model described in this paper is focused on the reduction of the resulting graph. In such a way, the resulting graph does not depend on the number of clusters, but on the number of elements in the dataset (i.e. the number of users in the SN). The task of assigning a cluster to the different users is carried out by the ants, while they are moving through the graph. In this case, the ants travel through the graph visiting the different nodes that represents the users of the dataset. The goal of each ant is to visit all the nodes, because the need to analyze all the users of the dataset. Each time, any ant arrives to a node, the ant has to decide the cluster that

will contain this user using the standard equation of the ACO algorithm.

In order to do that, the ants use a modified heuristic function  $\tau_{ij}$  that measures how good is to incorporate the user  $e_i$  to the cluster  $C_j$  by analyzing the different connections between  $e_i$  and the other users that compose the cluster  $C_j$ . This heuristic function is shown in Eq. 1.

$$\tau_{ij} = \text{Topology}(e_i, C_j) = \frac{|\mathcal{N}_i \cap C_j|}{|\mathcal{N}_i|} \quad (1)$$

Where  $\mathcal{N}_i$  represents the set of elements connected to  $e_i$  and  $|\cdot|$  is a function that compute the number of elements contained in a specific set. Therefore, this function takes into account how many connections of  $e_i$  belongs to the cluster  $C_j$ .

### III. EXPERIMENTAL EVALUATION

For testing the designed algorithm, we have used the dataset published by SNAP<sup>1</sup>. This dataset is composed by 10 different Ego Networks, and it has been selected because the dataset contains the groundtruth with the communities that compose the networks.

For the experimental phase, we have compared the performance of the ACO model against some of the well-known algorithms extracted from the State-of-the-Art. Table I shows for each algorithm how many networks each algorithm provides the best solution. Note that the best solution is the one with the highest Omega Index.

	Facebook	
	Total	Percentage
Topology-based ACO	5	50
Walktrap	1	10
Infomap	3	30
Clauset Newman Moore	1	10
Label Propagation	0	0

TABLE I

THIS TABLE SHOWS FOR EACH ALGORITHM (FIRST COLUMN), HOW MANY EGO NETWORKS EACH ALGORITHM HAS PROVIDED THE BEST SOLUTION AND THE RELATED PERCENTAGE (%) FOR EACH DATASET ANALYZED. THE BEST SOLUTION IS THE ONE THAT PROVIDES THE HIGHEST OMEGA INDEX SCORE BECAUSE IT IS THE CLOSEST TO THE GROUNDTRUTH.

### IV. CONCLUSIONS AND FUTURE WORK

This work provides a summary of the work published in [1]. In that work, we described an implementation of a new Ant Colony Optimization (ACO) algorithm to perform the community finding tasks in Social Networks (SNs), based on the topology information of the network.

In the experimental phase, we compared the performance of the described algorithm against some of the state-of-the-art

algorithms. The results reveal that the proposed Topology-based ACO algorithm is a good approach to solve the community finding problem because it provides competitive results against the analyzed algorithms.

As future work, we are working on the adaptation of novel bio-inspired algorithm to perform community finding tasks. These algorithms are Artificial Bee Colony [3] and a new algorithm based on the calling behavior of male Japanese tree frogs [10].

### ACKNOWLEDGMENT

This work has been co-funded by the following research projects: Spanish Ministry of Science and Education and Competitiveness (MINECO) and European Regional Development Fund (FEDER) under grant TIN2017-85727-C4-3-P (DeepBio); and the Justice Programme of the European Union (2014-2020) 723180 – RiskTrack – JUST-2015-JCOO-AG/JUST-2015-JCOO-AG-1. The contents of this publication are the sole responsibility of their authors and can in no way be taken to reflect the views of the European Commission.

### REFERENCES

- [1] A. Gonzalez-Pardo, J. J. Jung, and D. Camacho, "Aco-based clustering for ego network analysis," *Future Generation Computer Systems*, vol. 66, pp. 160 – 170, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167739X16302278>
- [2] A. Giachanou and F. Crestani, "Like it or not: A survey of twitter sentiment analysis methods," *ACM Comput. Surv.*, vol. 49, no. 2, pp. 28:1–28:41, Jun. 2016. [Online]. Available: <http://doi.acm.org/10.1145/2938640>
- [3] M. Lozano, C. Garca-Martinez, F. J. Rodriguez, and H. M. Trujillo, "Optimizing network attacks by artificial bee colony," *Information Sciences*, vol. 377, pp. 30 – 50, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0020025516312075>
- [4] M. Fernandez, M. Asif, and H. Alani, "Understanding the roots of radicalisation on twitter," in *In WebSci '18: 10th ACM Conference on Web Science*. ACM (Association for Computing Machinery), May 2018. [Online]. Available: <http://oro.open.ac.uk/54344/>
- [5] R. Lara-Cabrera, A. Gonzalez-Pardo, and D. Camacho, "Statistical analysis of risk assessment factors and metrics to evaluate radicalisation in twitter," *Future Generation Computer Systems*, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167739X17308348>
- [6] R. Lara-Cabrera, A. G. Pardo, K. Benouaret, N. Faci, D. Benslimane, and D. Camacho, "Measuring the radicalisation risk in social networks," *IEEE Access*, vol. 5, pp. 10 892 – 10 900, 2017.
- [7] J. Sánchez-Oro and A. Duarte, "Iterated greedy algorithm for performing community detection in social networks," *Future Generation Computer Systems*, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167739X17323932>
- [8] C. Pizzuti, "Ga-net: A genetic algorithm for community detection in social networks," in *Parallel Problem Solving from Nature – PPSN X*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 1081–1090.
- [9] M. Dorigo, "Optimization, learning and natural algorithms (in Italian)," Ph.D. dissertation, Dipartimento di Elettronica, Politecnico di Milano, Milan, Italy, 1992.
- [10] C. Blum, B. Calvo, and M. J. Blesa, "Frogcol and frogmis: new decentralized algorithms for finding large independent sets in graphs," *Swarm Intelligence*, vol. 9, no. 2, pp. 205–227, Sep 2015. [Online]. Available: <https://doi.org/10.1007/s11721-015-0110-1>

<sup>1</sup><http://snap.stanford.edu/data/egonets-Facebook.html>

**I Workshop en  
Ciencia de Datos  
en Redes Sociales  
(CidReS 2018)**

SESIÓN 3







# Red Neural Recurrente para la Desambiguación de Entidades en Datos de Medios Sociales

Cristina Zuheros, Siham Tabik, Ana Valdivia, Eugenio Martínez-Cámara y Francisco Herrera  
Instituto Andaluz de Investigación en Ciencias de Datos e Inteligencia Computacional,  
Universidad de Granada, 18071 Granada, España  
Email: {czuheros, siham, avaldivia}@ugr.es, {emcamara, fherrera}@decsai.ugr.es

**Resumen**—Un reto particular en el ámbito del Procesamiento del Lenguaje Natural es la desambiguación de palabras polisémicas. La gran disponibilidad, diversidad y rapidez cambiante de los datos en línea fuerzan el desarrollo de sistemas de desambiguación con una dependencia reducida de los recursos lingüísticos. Se parte de la base que la codificación del contexto de una entidad específica no requiere del uso de recursos lingüísticos externos como bases de conocimiento, por lo que se propone una arquitectura de red neuronal basada en el uso de las Redes Neuronales Recurrentes *Long Short-Term Memory* para codificar el contexto de una entidad, concretamente el modelo *Two k-Contextual Windows*. Se considera un problema real que requiere la desambiguación de la entidad *Granada*. Se genera y proporciona un corpus etiquetado de comentarios extraídos de medios sociales escritos en inglés, los cuales se utilizan para evaluar nuestra propuesta. Los resultados demuestran la validez de nuestra hipótesis.

## I. INTRODUCCIÓN

La mayoría de las palabras tienen más de un significado posible ya que el lenguaje humano es inherentemente ambiguo. Los humanos somos capaces de identificar el sentido de una palabra analizando el contexto en el que esta aparece. Actualmente, una enorme cantidad de información, incluyendo textos, imágenes y vídeos, está disponible en distintas plataformas de Internet como son *Twitter*, *Instagram*, *Tripadvisor* y *Booking*, entre otros [1] y [2]. La gran cantidad de información que puede obtenerse de esas fuentes es muy valiosa en distintos campos. Por ejemplo, saber lo que se piensa sobre una entidad, un producto, un evento, una persona o un lugar, nos permite usar dicha información para mejorar un servicio y aumentar el número de clientes o turistas. Extraer esta información sobre una entidad con un significado específico de una enorme cantidad de datos requiere filtrar aquellas menciones que hacen referencia a otros posibles significados de la entidad en cuestión.

La desambiguación del sentido de la palabra (*Word Sense Disambiguation* - WSD) es la tarea del Procesamiento del Lenguaje Natural que se encarga de la asignación computacional del sentido o significado correcto de una palabra dependiendo de su contexto [3] y [4]. Esta tarea requiere una gran cantidad de información y conocimiento. Actualmente, se vuelve cada vez más desafiante, ya que los sentidos de las palabras son muy diversos e incluso cambiantes, especialmente en escenarios donde la vinculación de un nuevo sentido a una entidad es impredecible. Además, los sistemas supervisados de WSD dependen altamente de las características anotadas a mano, las

cuales también se ven afectadas por el problema anteriormente mencionado. Por tanto, la tarea de desambiguación se ve limitada por el cuello de botella de la adquisición de conocimiento y la naturaleza siempre cambiante del lenguaje.

El estado del arte en WSD queda determinado principalmente por modelos supervisados y actualmente por métodos de aprendizaje profundo basados en un tipo específico de redes neuronales recurrentes conocido como modelo *Long Short-Term Memory* (LSTM), como el modelo *Targeted Two k-Contextual Windows* [5]. Sin embargo, esos modelos neuronales no han sido evaluados en *corpora* obtenidos de medios sociales. Este tipo de datos presentan unas características propias como el uso informal del lenguaje y el uso de *emojis* para la expresión de emociones [6]. Estas características específicas hacen más desafiante la tarea.

Las contribuciones que se hacen en este trabajo se pueden resumir en: (1) nueva arquitectura neuronal basada en LSTM para la desambiguación de entidades, llamada *Two k-Contextual Windows*. Consideramos como caso de estudio la desambiguación de la entidad *Granada*. Y (2) nuevo corpus etiquetado manualmente cada uno de los sentidos de la entidad objeto de estudio. El corpus está compuesto por textos publicados en *Twitter* e *Instagram*.

Dicho documento está estructurado de la siguiente manera: la Sección II abarca la definición de WSD como tarea de clasificación y un breve análisis sobre los trabajos más cercanos. La descripción del modelo propuesto se muestra en la Sección III, y la generación del corpus utilizado para la evaluación se presenta en la Sección IV. El marco experimental se muestra en la Sección V, mientras que los resultados experimentales se presentan en la Sección VI. Por último, en la Sección VII se presentan las conclusiones y los trabajos futuros.

## II. CONTEXTO

En esta Sección se exponen las bases para entender nuestra propuesta. En la Sección II-A se define la tarea de clasificación de WSD junto con los *word embeddings*. En la Sección II-B se presenta la LSTM RNN y en la Sección II-C se revisan los trabajos más relacionados.

### II-A. Aprendizaje de secuencias para WSD y embeddings

WSD puede definirse como la identificación automática del sentido más adecuado de una palabra a la misma en función de su contexto. Puede verse como una tarea de clasificación.

Sea  $T$  una secuencia de  $n$  palabras  $\{w_1, \dots, w_n\}$ , la tarea de desambiguación consiste en encontrar una función  $A$  de palabras a sentidos, tal que  $A(w_j) \subseteq \text{sentido}_D(w_j)$  siendo  $\text{sentido}_D(w_j)$  el conjunto de sentidos posibles de la palabra  $(w_j)$  en un diccionario de sentidos  $D$ .

En este trabajo,  $T$  es el conjunto de comentarios extraídos de *Twitter* e *Instagram*, la palabra a desambiguar  $w_t$  es *Granada* y  $D$  es el conjunto de posibles significados que se detallan en la Sección IV-A. Se propone representar mediante vectores de *word embeddings* la secuencia de palabras  $\{w_1, \dots, w_n\}$  de cada secuencia de entrada  $T$  y codificarlas posteriormente con una LSTM RNN.

Un *embedding* es una representación de un objeto topológico en un espacio determinado, de forma que se conservan sus propiedades de conectividad o algebraicas [7] y [8]. Un vector *word embedding* es la representación del espacio semántico ideal de palabras en un espacio vectorial continuo. Matemáticamente, se define un *word embedding*  $w_j$ , como un vector  $d$ -dimensional  $\mathbf{we}_j^T = (we_1, \dots, we_d) \in \mathbb{R}^d$ . Por tanto, se define  $WE_{1:n} \in \mathbb{R}^{d \times n}$  como la matriz de *word embeddings* del conjunto de palabras  $\{w_1, \dots, w_n\}$ . Algunos de los modelos más destacados para realizar *word embeddings* son *word2vec* [9] y *GLoVe* [10].

### II-B. Red neuronal de memoria de corto y largo plazo

LSTM es un tipo de RNN con la capacidad de representar entradas secuenciales de tamaño arbitrario en un vector de tamaño fijo y prestar atención a las propiedades estructuradas de las entradas [11]. Dado que se pretende desambiguar una palabra en una entrada secuencial de tamaño arbitrario, LSTM cumple con los requisitos de nuestro problema.

A continuación, se definen las bases de RNN que también son la base de LSTM. Una RNN se define como una función  $R$  que se aplica recursivamente a una secuencia de palabras, concretamente en nuestro caso una secuencia de *word embeddings*  $(\mathbf{we}_1, \dots, \mathbf{we}_n)$ . La función  $R$  toma como entrada un vector de estado  $\mathbf{s}_{j-1}$  y un vector  $\mathbf{we}_j$ , y genera un nuevo vector estado  $\mathbf{s}_j$ . Este vector estado es proyectado en un vector salida  $\mathbf{y}_j$  mediante una función determinística  $O(\cdot)$ . La Ecuación 1 resume dicha definición.

$$\begin{aligned} RNN(WE_{1:n}; \mathbf{s}_0) &= \mathbf{y} \\ \mathbf{y}_j &= O(\mathbf{s}_j) \\ \mathbf{s}_j &= R(\mathbf{s}_{j-1}, \mathbf{we}_j) \end{aligned} \quad (1)$$

Las RNNs presentan el problema del desvanecimiento del gradiente [12] y [13], lo que significa que la multiplicación repetida de parámetros en la función  $R$  puede hacer que los valores de esos parámetros se desvanezcan o exploten, lo que hace más difícil el entrenamiento de la red neuronal. LSTM fue la primera RNN con una arquitectura diseñada para resolver dicho problema [12] y [13].

LSTM [14] separa el vector estado ( $\mathbf{s}_j$ ) en dos mitades, siendo una mitad considerada como “celdas de memoria” y la otra como memoria de trabajo. El mecanismo de LSTM decide en cada entrada de la secuencia qué cantidad de la nueva entrada ( $we_j$ ) debe escribirse en la celda de memoria

y qué cantidad del contenido de la celda de memoria debe olvidarse [15].

Desde un punto de vista lingüístico, la operación recursiva de LSTM significa que cada palabra ( $w_j$ ) está codificada con el significado de las palabras anteriores ( $R_{LSTM}(\mathbf{s}_{j-1}, \mathbf{we}_j)$ ). Sin embargo, el contexto de una palabra no sólo depende de las palabras localizadas antes de esa palabra, sino también en las palabras que se encuentran después de ella en la oración. Las LSTMs son versátiles, pueden funcionar en direcciones diferentes y sus salidas pueden combinarse de maneras diferentes. Formalmente, dado un vector de *word embeddings*  $(\mathbf{we}_1, \dots, \mathbf{we}_n)$ , y una palabra a desambiguar,  $\mathbf{we}_t$ , una LSTM puede codificar las palabras de la secuencia de entrada tanto de izquierda a derecha ( $\overrightarrow{\text{LSTM}}$ ) como de derecha a izquierda ( $\overleftarrow{\text{LSTM}}$ ), como indican las flechas. Las salidas de ambas LSTMs pueden ser combinadas y procesadas por una o más capas. Por tanto, codificar el significado de una secuencia de palabras puede ser expresado mediante una función  $G(\cdot)$  que combina la salida de un número  $m$  de redes LSTMs (ver Ecuación 2).

$$\text{meaning} = G(\text{LSTM}_{1:m}) \quad (2)$$

### II-C. Trabajos afines

Los principales enfoques que se encuentran en la amplia literatura [3] y [4] sobre WSD pueden clasificarse en cuatro grupos: (1) Métodos basados en el conocimiento, (2) no supervisados, (3) semisupervisados, y (4) supervisados. Nos centramos en la categoría supervisada.

Aunque los modelos supervisados requieren que el sentido de los datos sea manualmente etiquetado por humanos expertos, superan a los sistemas basados en conocimiento en los *benchmarks* estándar [16]. Se basan en el hecho de que palabras semánticamente similares tienden a tener distribuciones contextuales similares [17], [18] y [19]. Gran cantidad de modelos utilizan características manuales y métodos tradicionales de aprendizaje automático, como el sistema IMS [20] y [21].

Recientemente, los métodos de aprendizaje profundo están mostrando resultados muy prometedores [22], [23], [24], [25] y [26]. Melamud et al. [27] presentaron un modelo no supervisado basado en LSTM bidireccional. En el trabajo [28] utilizaron una arquitectura similar de red neuronal pero bajo un enfoque semi-supervisado para codificar el contexto de la palabra a desambiguar. En el trabajo [5] utilizaron una arquitectura similar combinada con capas ocultas bajo un enfoque supervisado, logrando una alta precisión en la tarea de desambiguación léxica de las primeras ediciones de Senseval [29] y [30].

## III. MODELO NEURONAL PARA WSD

En la Sección III-A, se detalla el modelo LSTM basado en una arquitectura similar a la utilizada en [5]. En la Sección III-B, se presenta nuestra propuesta.

Ambos modelos<sup>1</sup> se basan en la codificación de una ventana contextual, que definimos como un número fijo de  $k$  palabras

<sup>1</sup>Sentencia mostrada en las Figuras 1 y 2: “La ciudad de Granada es muy bonita en invierno.”



consecutivas de la secuencia de entrada. Además, comparten la capa de entrada y representación. La capa de entrada es definida como una secuencia de  $n$  palabras,  $\{w_1, \dots, w_n\}$ . La capa de representación es una matriz de *word embeddings*, definida como  $WE_{1:n}$ . Se usa el conjunto de pre-entrenado de vectores de *word embeddings* 6B *GloVe* [10]. Se aplica *Dropword* como método de regularización.

Los modelos difieren en: (1) la dirección en la que el modelo procesa las palabras de cada ventana contextual, y (2) la posición de la palabra en la sentencia, a partir de la cual el modelo inicia el procesamiento de cada ventana de contexto.

### III-A. Targeted Two $k$ -Contextual Windows

El modelo *Targeted Two  $k$ -Contextual Windows*, al que nos referiremos como  $CW^{t-k}CW^{t+k}$  tiene una arquitectura similar al modelo presentado en [5]. Determina las  $k$ -palabras de la ventana de contexto basándose en la posición de la palabra a desambiguar  $w_t$ . La primera ventana de contexto incluye las  $k$ -palabras situadas a la izquierda de la palabra a desambiguar mientras que la segunda ventana considera las  $k$ -palabras situadas a la derecha de  $w_t$  (ver Figura 1).

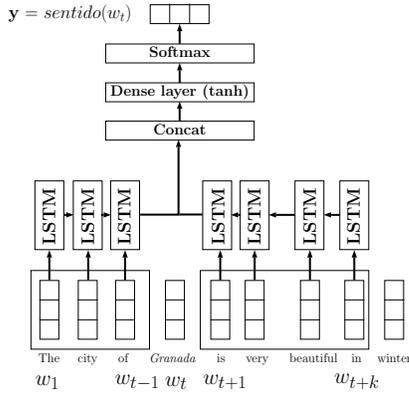


Figura 1. Modelo Targeted Two  $k$ -Contextual Windows considerando  $k = 4$ .

Las dos ventanas de contexto son procesadas por dos LSTM con  $d_{LSTM}$  unidades ocultas. El último vector de estado de las dos LSTMs son concatenadas y procesadas por una capa completamente conectada con  $h_{dense}$  unidades ocultas activadas mediante una función tangente hiperbólica ( $\tanh$ ). Finalmente, una capa *softmax* calcula la distribución de probabilidad correspondiente al conjunto de sentidos posibles de  $w_t$ . La ecuación 3 muestra las operaciones subyacentes.

$$\begin{aligned}
 sentido(w_t) &= \arg \max_{s \in S} (\mathbf{y}), \mathbf{y} \in \mathbb{R}^s \\
 \mathbf{y} &= \text{softmax}(\mathbf{c}), \mathbf{c} \in \mathbb{R}^{h_{dense}} \\
 \mathbf{c} &= \text{dense}(\text{concat}), \text{concat} \in \mathbb{R}^{2 \cdot h_{LSTM}} \\
 \text{concat} &= [\mathbf{m}^1; \mathbf{m}^2], \mathbf{m}^1, \mathbf{m}^2 \in \mathbb{R}^{h_{LSTM}} \\
 \mathbf{m}^2 &= \underline{\text{LSTM}}(WE_{t+1:t+k}), WE_{t+1:t+k} \in \mathbb{R}^{d \times k} \\
 \mathbf{m}^1 &= \underline{\text{LSTM}}(WE_{t-k:t-1}), WE_{t-k:t-1} \in \mathbb{R}^{d \times k}
 \end{aligned} \tag{3}$$

### III-B. Two $k$ -Contextual Windows

Se propone un modelo que, a diferencia del modelo previo (ver Sección III-A), permite codificar las primeras y últimas  $k$ -palabras de la secuencia de entrada sin tener en cuenta la

posición de  $w_t$  y permite solapar las dos ventanas de contexto. El modelo *Two  $k$ -Contextual Windows* ( $CW^kCW^{n-(k-1)}$ ) procesa dos ventanas de contexto de  $k$ -palabras en direcciones opuestas. La primera analiza la oración hasta la palabra localizada en la posición  $k$  y la segunda analiza desde la palabra localizada en la posición  $n-(k-1)$ , siguiendo las direcciones indicadas por las flechas, como muestra la Figura 2.

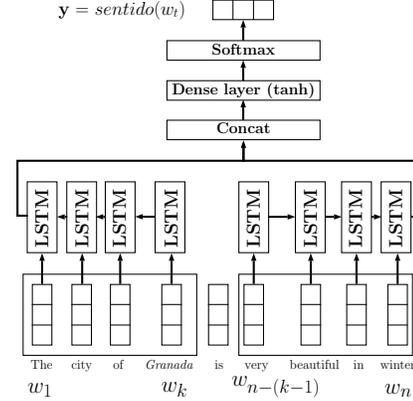


Figura 2. Modelo Two  $k$ -Contextual Windows considerando  $k = 4$ .

Ambas ventanas de contexto son individualmente procesadas usando una LSTM con  $h_{LSTM}$  unidades ocultas. Las salidas son concatenadas y procesadas por una capa completamente conectada con  $h_{dense}$  unidades ocultas activadas mediante una función tangente hiperbólica. Finalmente, *softmax* proporciona el sentido de la palabra a desambiguar  $w_t$ . La ecuación 4 muestra las operaciones realizadas por cada capa.

$$\begin{aligned}
 sentido(w_t) &= \arg \max_{s \in S} (\mathbf{y}), \mathbf{y} \in \mathbb{R}^s \\
 \mathbf{y} &= \text{softmax}(\mathbf{c}), \mathbf{c} \in \mathbb{R}^{h_{dense}} \\
 \mathbf{c} &= \tanh(\text{concat}), \text{concat} \in \mathbb{R}^{2 \cdot h_{LSTM}} \\
 \text{concat} &= [\mathbf{m}^1; \mathbf{m}^2], \mathbf{m}^1, \mathbf{m}^2 \in \mathbb{R}^{h_{LSTM}} \\
 \mathbf{m}^2 &= \underline{\text{LSTM}}(WE_{1:k}), WE_{1:k} \in \mathbb{R}^{d \times k} \\
 \mathbf{m}^1 &= \underline{\text{LSTM}}(WE_{n-(k-1):n}), WE_{n-(k-1):n} \in \mathbb{R}^{d \times k}
 \end{aligned} \tag{4}$$

## IV. GENERACIÓN DEL CORPUS

En la siguiente Sección se describen los diferentes sentidos de la entidad *Granada* (Sección IV-A), se expone el pre-procesamiento de los datos (Sección IV-B) y la anotación del corpus (Sección IV-C).

### IV-A. Palabra a desambiguar

Un análisis profundo realizado en diferentes fuentes de Internet ha demostrado que la palabra *Granada* tiene al menos 54 sentidos. Los sentidos más frecuentes pueden agruparse en: Sentido 1: Provincia/Ciudad/Calle localizada en el sureste de España, evento (no relacionado con música o deporte) o monumento/establecimiento de dicha localización.

Sentido 2: País/Provincia/Ciudad/Distrito/Calle/Río no localizado en España.

Sentido 3: Fruto, granada.

Sentido 4: Referencias a música o deporte.

Sentido 5: Otros significados menos relevantes como usuarios de redes sociales, canales de televisión, vehículos, periódicos, entre otros.

#### IV-B. Pre-procesamiento del corpus

El conjunto de datos contiene un total de 876.969 comentarios y opiniones que incluyen la palabra *Granada* publicados por usuarios de *Twitter* e *Instagram*. Cada instancia del dataset presenta los campos siguientes:

- *uid*: identificador de la instancia.
- *source*: fuente de la instancia.
- *body*: texto, comentario de la instancia.

La mayoría de las instancias que provienen de medios sociales contienen *emojis*, ya que son útiles para expresar los sentimientos de las personas [31]. Estos sentimientos pueden sugerir si una persona está hablando sobre un tema u otro. Por ejemplo, si una usuaria usa algunos *emojis* mostrando frutas, probablemente no esté hablando sobre deporte. Como los *emojis* no pueden ser procesados directamente por los modelos, se utiliza una descripción con palabras para representar cada *emoji*. Se elabora un diccionario de *emojis* público que contiene un total de 837 *emojis* expresados en inglés<sup>2</sup>.

Se eliminan todas URLs e imágenes, dado que el procesamiento de imágenes no forma parte del flujo de trabajo de nuestra propuesta. Se eliminan los comentarios carentes de contenido y los caracteres repetitivos ya que no son útiles para nuestra tarea. Finalmente, se reemplazan caracteres como  $<3, =)$  por “love”/“amor”, “I am happy”/“Estoy feliz”, entre otros.

#### IV-C. Etiquetado del corpus

La primera autora, que tiene un alto dominio del inglés, etiquetó manualmente las instancias del conjunto de datos. Cada instancia se anota con un solo sentido. La Tabla I muestra el número de comentarios por cada sentido. El conjunto de instancias etiquetadas se almacena en un archivos XML.

Sentido 1	Sentido 2	Sentido 3	Sentido 4	Sentido 5	Total
11.985	3.381	452	7.438	1.445	24.701

Tabla I  
NÚMERO DE INSTANCIAS ETIQUETADAS PARA CADA SENTIDO.

### V. MARCO EXPERIMENTAL

Se realiza una experimentación profunda con el fin de evaluar nuestra propuesta. En primer lugar, se establecen dos subconjuntos de entrenamiento y test con el objetivo de evaluar nuestro modelo neuronal mediante diferentes distribuciones de los datos. También se evalúa la relevancia del uso de *emojis* como características para la tarea de desambiguación.

Además, se considera XGBoost como caso base, ya que presenta buenos resultados en tareas de clasificación. Dado que se considera WSD como una tarea de clasificación, se utilizan las medidas de evaluación propias de tareas de clasificación para evaluar la conducta de nuestra propuesta. Los detalles de la experimentación son los siguientes.

**Partición de los datos** Se crean dos particiones diferentes de conjuntos de entrenamiento y de test para evaluar nuestra propuesta: (1) P1, el 80 % de las instancias se encuentran en

el conjunto de entrenamiento y el conjunto de test contiene el 20 % restante, y (2) P2, el 70 % de las instancias están en el conjunto de entrenamiento mientras que el 30 % restante están en el test.. El número de elementos de cada conjunto se muestra en el Tabla II.

**Emojis** Puesto que se considera que los *emojis* pueden ser una característica relevante para la tarea de desambiguación, comparamos resultados de: (1) eliminar los *emojis* de las instancias, y (2) sustituir cada *emoji* por una expresión textual.

**Entrenamiento de los modelos** Los valores de los hiperparámetros utilizados en los modelos neuronales se muestran en la Tabla III. El entrenamiento se llevó a cabo utilizando la función de pérdida de entropía cruzada y el algoritmo de optimización *Momentum* [32].

**Baseline** Con el fin de comparar, se aplica el modelo de clasificación *XGBoost* para bolsa de palabras [33]. *XGBoost* es un sistema de clasificación basado en árboles de decisión cuyo entrenamiento se basa en el método de aumento de gradientes. Los textos se *tokenizan* y las palabras se representan mediante su valor *tf-idf*.

**Métricas de evaluación** Dado que se re-formula la tarea de WSD como un problema de clasificación multiclase, se usan las versiones macro-promediadas de Precisión, Recall y F1.

### VI. RESULTADOS Y ANÁLISIS

En esta sección se describen y analizan los resultados alcanzados para el caso base, XGBoost, y los dos modelos presentados en la Sección III,  $\langle CW^{t-k}CW^{t+k} \rangle$  y  $\langle CW^kCW^{n-(k-1)} \rangle$ . En primer lugar se analiza y compara el rendimiento global de todos los modelos, y posteriormente se estudia el rendimiento por clase del mejor modelo. Finalmente, se analiza el impacto de eliminar o incluir *emojis* y el uso de *GloVe*.

La Tabla IV muestra los resultados alcanzados para los tres modelos indicados para cada partición de los datos incluyendo y sin incluir *emojis*. Los dos modelos neuronales superan al modelo XGBoost y logran resultados muy beneficiosos. Los mejores resultados para todos los modelos se obtienen en la partición P2 incluyendo *emojis*. Las mejores medidas para todas las particiones consideradas son proporcionadas por el modelo  $\langle CW^kCW^{n-(k-1)} \rangle$ , seguido por  $\langle CW^{t-k}CW^{t+k} \rangle$ .  $\langle CW^{t-k}CW^{t+k} \rangle$  alcanza peores resultados ya que considera menos información que  $\langle CW^kCW^{n-(k-1)} \rangle$ , debido a que

P1 entrenamiento	P1 test	P2 entrenamiento	P2 test
19.760	4.941	17.290	7.411

Tabla II  
TAMAÑO DEL CONJUNTO DE ENTRENAMIENTO Y TEST PARA P1 Y P2.

Hiperparámetro	Valor
Tamaño ventana de contexto ( <i>k</i> )	70
Dimensión <i>word embedding</i> ( <i>d</i> )	100
Unidades capa LSTM ( $h_{LSTM}$ )	74
Unidades capa Dense ( $h_{dense}$ )	200

Tabla III  
HIPERPARÁMETROS DE AMBOS MODELOS NEURONALES.

<sup>2</sup>[https://github.com/cristinazuhe/Disambiguation\\_SocialMedia\\_LSTM](https://github.com/cristinazuhe/Disambiguation_SocialMedia_LSTM).



$CW^{t-k}CW^{t+k}$  analiza dos ventadas contextuales disjuntas. Por ejemplo, cuando la palabra a desambiguar se encuentra al principio o al final de la sentencia, solamente algunas de las palabras que están situadas a la derecha o a la izquierda de dicha palabra, respectivamente, se tienen en cuenta en el análisis de  $CW^{t-k}CW^{t+k}$ .

La evaluación en términos de Recall, Precisión y F1 del modelo  $CW^kCW^{n-(k-1)}$  para cada sentido se muestra en la Tabla V. Los mejores resultados se obtienen para la partición P2. Incluir expresiones para los *emojis* mejora el rendimiento del modelo en todas las clases excepto para el sentido 5. El valor más alto para la medida F1 es alcanzado en los sentidos 1 y 5 mientras que los resultados más bajos se obtienen en los sentidos 3 y 4. Se puede concluir que sustituir los *emojis* por sus correspondientes expresiones es beneficioso para el rendimiento del sistema. En base a la medida F1, los *emojis* son buenas características para los sentidos 3 y 4, por ejemplo la medida F1 para el sentido 3 aumenta en un 3,14 % al usar *emojis* como características.

**Emojis** Los *emojis* que afectan positivamente a cada sentido después de reemplazarlos con sus correspondientes descripciones mediante palabras se muestran en la Tabla VI. Si se eliminan los *emojis* mostrados en la primera fila, el modelo clasifica las correspondientes instancias con el sentido 1. Sin embargo, estas instancias se clasifican con el sentido correcto, 4, si las instancias sí incluyen las descripciones de los *emojis*. Intuitivamente, es evidente que todos los *emojis* mostrados tienen una fuerte relación con las clases a las que se asocian después de entrenar los modelos, por lo que son características informativas. Por ejemplo, el melón y la fresa están relacionados con la fruta, los bíceps y las notas musicales están relacionadas con el deporte/música. Sin embargo, algunos *emojis* no están asociados trivialmente a un sentido particular. Por ejemplo, los *emojis* de la última fila, el coche de policía y la luz giratoria del coche de policía, se asocian durante el entrenamiento al sentido 2, que hace referencia a las localizaciones que no están situados en España. Por tanto, se concluye con nuestro estudio que en estas áreas hay demasiados disturbios policiales.

**Word embedding** Se analiza la importancia de usar *word embeddings* pre-entrenados. Se comparan los resultados al-

Partición	Emojis	Sentido	%Prec.	%Recall	%F1
P1	No	S1	96,44	98,91	97,66
		S2	96,61	96,74	96,67
		S3	100,0	92,31	96,00
		S4	96,67	93,74	95,18
		S5	99,35	95,60	97,44
	Sí	S1	96,40	98,70	97,53
		S2	96,49	97,31	96,90
		S3	97,37	94,87	96,10
		S4	93,47	96,66	95,03
		S5	99,67	95,60	97,59
P2	No	S1	96,57	98,62	97,58
		S2	96,63	96,72	96,68
		S3	97,46	89,84	93,50
		S4	96,52	93,81	95,15
		S5	98,90	98,25	98,57
	Sí	S1	97,22	98,79	98,00
		S2	96,28	97,21	96,74
		S3	97,60	95,31	96,44
		S4	97,25	94,57	95,89
		S5	98,46	98,03	98,24

Tabla V  
RESULTADOS DEL MEJOR MODELO ANALIZADOS PARA CADA SENTIDO PARA P1 Y P2 CON Y SIN *emojis*.

Sin emojis	Emojis	Con emojis
1		4
4		1
1		3
1		2

Tabla VI  
EL MODELO CLASIFICA CORRECTAMENTE EL SENTIDO DE LA PALABRA USANDO LOS *emojis* DE LA COLUMNA *Emojis*.

canzados usando *GloVe* con los resultados obtenidos con *word embeddings* inicializados mediante una distribución gaussiana. Se realiza el análisis sobre la partición P2 incluyendo la descripción de los *emojis*. Los resultados se muestran en la Figura 3. La clasificación a partir de los *embeddings* pre-entrenados usando *GloVe* alcanzan resultados ligeramente más altos que la clasificación a partir de los *embeddings* gaussianos como se esperaba, ya que los *embeddings* pre-entrenados incorporan conocimiento previo semántico a los modelos neuronales.

Modelo	Partición	Emojis	Prec.	Recall	F1
XGBoost	P1	No	18,41	17,26	17,82
		Sí	18,63	20,22	19,39
	P2	No	28,20	24,20	26,05
		Sí	55,19	49,14	51,99
$CW^{t-k}CW^{t+k}$	P1	No	95,27	93,63	94,44
		Sí	95,78	93,89	94,72
	P2	No	96,23	93,76	94,98
		Sí	95,67	95,35	95,51
$CW^kCW^{n-(k-1)}$	P1	No	<b>97,81</b>	<b>95,46</b>	<b>96,63</b>
		Sí	<b>97,32</b>	<b>95,99</b>	<b>96,65</b>
	P2	No	<b>97,22</b>	<b>95,45</b>	<b>96,32</b>
		Sí	<b>97,36</b>	<b>96,78</b>	<b>97,07</b>

Tabla IV  
RESULTADOS DE LOS MODELOS PARA P1 Y P2 CON Y SIN *emojis*.

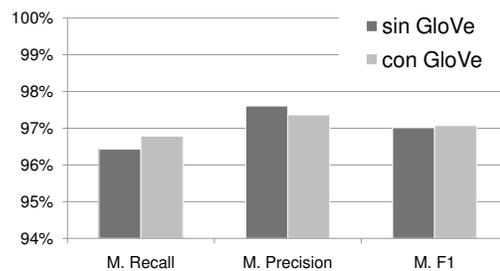


Figura 3. Resultados del mejor modelo usando *word embeddings* aleatorios (sin *GloVe*) o pre-entrenados (usando *GloVe*).

## VII. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se propone una nueva arquitectura neuronal basada en el uso de las redes LSTM para la tarea de desambiguación de entidades, específicamente el modelo *Two k-Contextual Windows*. Además, se proporciona un nuevo corpus etiquetado para la desambiguación de la entidad *Granada* en textos extraídos de medios sociales, lo cuál es una aportación relevante para la comunidad investigadora.

En nuestra experimentación se pone en manifiesto que nuestro modelo propuesto logra resultados muy buenos en textos escritos en inglés extraídos de medios sociales. En contraste con [5], la codificación de la palabra a desambiguar contribuye a mejorar el rendimiento del sistema. Finalmente, tanto el uso de *embeddings* como la inclusión de expresiones para representar a los *emojis* permite alcanzar mejores resultados.

Como trabajo futuro, se desarrollarán nuevas técnicas para abordar el problema del desbalanceo de clases. Se trabajará en la creación de nuevos corpus de datos en otros idiomas con el objetivo de realizar experimentos en un entorno multilingüe.

## AGRADECIMIENTOS

Agradecemos a la empresa Mabrian por compartir los datos que han sido utilizados en la evaluación de nuestra propuesta. Este trabajo contó con el apoyo parcial del Ministerio de Economía y Competitividad de España en el marco del proyecto TIN2017-89517-P, y con una subvención del FEDER. Siham Tabik contó con el apoyo del Programa Ramón y Cajal (RYC-2015-18136) del Gobierno de España, y Eugenio Martínez Cámara fue apoyado por el Programa Juan de la Cierva Formación (FJCI-2016-28353) del Gobierno de España. En esta investigación usamos Titan X Pascal donada por NVIDIA Corporation.

## REFERENCIAS

- [1] M. T. Thai, W. Wu, and H. Xiong, *Big Data in Complex and Social Networks*. Chapman and Hall/CRC, 11 2016.
- [2] A. Farzindar and D. Inkpen, *Natural Language Processing for Social Media*, 2nd ed., ser. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers, 12 2017, vol. 10.
- [3] E. Agirre and P. Edmonds, *Word Sense Disambiguation: Algorithms and Applications (Text, Speech and Language Technology)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [4] R. Navigli, "Word sense disambiguation: A survey," *ACM Comput. Surv.*, vol. 41, no. 2, pp. 10:1–10:69, Feb. 2009.
- [5] M. Kägebäck and H. Salomonsson, "Word sense disambiguation using a bidirectional LSTM," in *Proceedings of the 5th Workshop on Cognitive Aspects of the Lexicon (CogALex - V)*. Osaka, Japan: The COLING 2016 Organizing Committee, December 2016, pp. 51–56.
- [6] E. Martínez-Cámara, M. T. Martín-Valdivia, L. A. Ureña López, and A. Montejo-Ráez, "Sentiment analysis in Twitter," *Natural Language Engineering*, vol. 20, no. 1, p. 1–28, 2014.
- [7] I. Iacobacci, M. T. Pilehvar, and R. Navigli, "Embeddings for word sense disambiguation: An evaluation study," in *Proceedings of the 54th Annual Meeting of the ACL (Vol 1: Long Papers)*, 2016, pp. 897–907.
- [8] M. Insall, T. Rowland, and E. W. Weistein, "Embedding."from mathworld—a wolfram web resource (access March 12, 2018), 2015.
- [9] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint:1301.3781*, 2013.
- [10] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. Doha, Qatar: ACL, October 2014, pp. 1532–1543.
- [11] Y. Goldberg, *Neural Network Methods for Natural Language Processing*. Morgan & Claypool Publishers, 2017.
- [12] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proceedings of the 30th International Conference on Machine Learning - Volume 28*, ser. ICML'13. JMLR.org, 2013, pp. III–1310–III–1318.
- [13] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, ch. 10.
- [14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
- [15] A. Graves, "Supervised sequence labelling with recurrent neural networks," Ph.D. dissertation, Technische Universität München, 2008.
- [16] A. Raganato, J. Camacho-Collados, and R. Navigli, "Word sense disambiguation: A unified evaluation framework and empirical comparison," in *Proceedings of the 15th Conference of the European Chapter of the ACL: Volume 1, Long Papers*. ACL, 2017, pp. 99–110.
- [17] Z. S. Harris, "Distributional structure," *WORD*, vol. 10, no. 2-3, pp. 146–162, 1954.
- [18] J. R. Firth, "A synopsis of linguistic theory 1930-55," vol. 1952-59, pp. 1–32, 1957.
- [19] G. A. Miller and W. G. Charles, "Contextual correlates of semantic similarity," *LCP*, vol. 6, no. 1, pp. 1–28, 1991.
- [20] Z. Zhong and H. T. Ng, "It makes sense: A wide-coverage word sense disambiguation system for free text," in *Proceedings of the ACL 2010 System Demonstrations*. Uppsala, Sweden: ACL, July 2010, pp. 78–83.
- [21] H. Shen, R. Bunescu, and R. Mihalcea, "Coarse to fine grained sense disambiguation in wikipedia," in *Second Joint Conference on Lexical and Computational Semantics (\*SEM), Volume 1: Proceedings of the Main Conference and the Shared Task: Semantic Textual Similarity*. Atlanta, Georgia, USA: ACL, June 2013, pp. 22–31.
- [22] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *Journal of Machine Learning Research*, vol. 12, pp. 2493–2537, Nov. 2011.
- [23] C. D. Santos and B. Zadrozny, "Learning character-level representations for part-of-speech tagging," in *Proceedings of the 31st International Conference on Machine Learning*, E. P. Xing and T. Jebara, Eds., vol. 32, no. 2, 2014, pp. 1818–1826.
- [24] E. Kiperwasser and Y. Goldberg, "Simple and accurate dependency parsing using bidirectional lstm feature representations," *Transactions of the ACL*, vol. 4, pp. 313–327, 2016.
- [25] D. F. Wong, Y. Lu, and L. S. Chao, "Bilingual recursive neural network based data selection for statistical machine translation," *Knowledge-Based Systems*, vol. 108, pp. 15 – 24, 2016, new Avenues in Knowledge Bases for Natural Language Processing.
- [26] K. Lin, D. Li, X. He, M.-t. Sun, and Z. Zhang, "Adversarial ranking for language generation," in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Curran Associates, Inc., 2017, pp. 3158–3168.
- [27] O. Melamud, J. Goldberger, and I. Dagan, "context2vec: Learning generic context embedding with bidirectional lstm," in *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning*. ACL, 2016, pp. 51–61.
- [28] K. Taghipour and H. T. Ng, "Semi-supervised word sense disambiguation using word embeddings in general and specific domains," in *Proceedings of the 2015 Conference of the North American Chapter of the ACL: Human Language Technologies*. Denver, Colorado: ACL, May–June 2015, pp. 314–323.
- [29] A. Kilgarriff, "English lexical sample task description," in *The Proceedings of the Second International Workshop on Evaluating Word Sense Disambiguation Systems*. ACL, 2001, pp. 17–20.
- [30] R. Mihalcea, T. Chklovski, and A. Kilgarriff, "The senseval-3 english lexical sample task," in *Proceedings of senseval-3, the third international workshop on the evaluation of systems for the semantic analysis of text*, 2004.
- [31] J. Read, "Using emoticons to reduce dependency in machine learning techniques for sentiment classification," in *Proceedings of the ACL Student Research Workshop*. ACL, 2005, pp. 43–48.
- [32] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural Networks*, vol. 12, no. 1, pp. 145 – 151, 1999.
- [33] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 785–794.



# Detección de nodos tramposos en procesos de consenso en redes

M. Rebollo\*<sup>†</sup>

\*Grupo de Tec. Informática – Inteligencia Artificial  
Universitat Politècnica de València  
Valencia, Spain  
mrebollo@upv.es

R.M. Benito<sup>†</sup>, J.C. Losada<sup>†</sup>, J. Galeano<sup>†</sup>

<sup>†</sup>Grupo de Sistemas Complejos  
Universidad Politécnica de Madrid  
Madrid, Spain  
{rosa.benito, juancarlos.losada, javier.galeano}@upm.es

**Resumen**—Los procesos de consenso en redes son procesos de difusión de información que permiten realizar cálculos de forma distribuida en una red, intercambiando información únicamente con los vecinos directos y sin tener conocimiento global de la estructura, tamaño, valores ni otras características de la red. Sin embargo, este proceso requiere que todos los nodos sigan el mismo algoritmo y cualquier cambio provoca una alteración en el valor de consenso alcanzado. Este trabajo propone un modelo para la detección de nodos que de forma deliberada o accidental no siguen el proceso de consenso. La red podrá, de forma autónoma y no supervisada, detectar estas desviaciones y corregirlas. Se muestra su aplicación a tres escenarios: un sistema de voto distribuido, la manipulación de imágenes por *adversarial examples* y el problema de los generales bizantinos, que es la base de los protocolos usados para la manipulación de cadenas de bloques.

**Index Terms**—complex networks, consenso, votaciones, blockchain, adversarial examples, tolerancia a fallos, resiliencia

## I. INTRODUCCIÓN

Los procesos de consenso en redes permiten calcular de forma distribuida el valor de una función común. Cada nodo emplea únicamente sus datos y la información de sus vecinos directos para recalcular de forma iterativa el valor de la función y propagarlo a través de sus vecinos. Este proceso converge a un valor único final para la función que se está calculando. Los nodos no tienen ningún tipo de conocimiento sobre el tamaño, la topología de la red ni ninguna otra característica.

Sea  $G = (V, E)$  un grafo no dirigido formado por un conjunto de vértices  $V$  y un conjunto de enlaces  $E \subseteq V \times V$  donde  $(i, j) \in E$  si existe un enlace entre los nodos  $i$  y  $j$ . Un vector  $x = (x_1, \dots, x_n)^T$  contiene los valores iniciales de las variables asociadas a cada nodo de la red. Olfati–Saber and Murray [1] proponen un algoritmo cuya aplicación iterativa converge al valor medio de  $x$ . La convergencia de este método está asegurada.

$$x_i(t+1) = x_i(t) + \varepsilon \sum_{j \in N_i} [x_j(t) - x_i(t)] \quad (1)$$

Este proceso converge al valor medio de los valores iniciales

$$\lim_{t \rightarrow \infty} x(t) = \frac{1}{n} \sum_i x_i(0) = \bar{x} \quad (2)$$

This work is supported by the PROMETEOII/2013/019, TIN2015-65515-C4-1-R and MTM2015-63914-P projects of the Spanish government

y en todas las iteraciones se conserva la suma de  $x$ .

$$\sum_i x_i(0) = \sum_i x_i(t) \quad \forall t \leq 0 \quad (3)$$

Sin embargo, algunas funciones no pueden calcularse por este proceso. Particularmente, aquellas que requieren la consideración de valores repetidos (como la moda) o valores acumulados (por ejemplo, obtener la suma de los valores iniciales). Por otra parte, todos los nodos deben seguir exactamente el proceso de forma obligatoria para que el resultado sea correcto. Para alterar el resultado, basta con que uno de los nodos altere el valor que intercambia en una de las iteraciones para que el valor de consenso se desvíe del valor esperado. Existen mecanismos para detectar cuándo se produce una variación, pero requieren guardar el estado completo de la red en cada iteración y además es necesario mantener una visión global de la red y conocer su estructura, pues se basan en matrices de observación que se calculan a partir de la matriz de adyacencia de la red completa [2]. Otras alternativas requieren la existencia de un nodo que hace las funciones de coordinador [3]. También están relacionados otros trabajos enfocados a reducir el efecto del ruido en redes de sensores, pero sus soluciones están centradas en obtener la topología de red óptima para minimizar el efecto del ruido [4], [5]

En el presente trabajo se propone un método completamente descentralizado que no requiere información global de la red y que es independiente de la topología, si bien es cierto que existen determinadas restricciones que se deben cumplir para poder anular por completo las desviaciones producidas.

Para mostrar su utilidad, se aplica a tres casos: un mecanismo de voto distribuido, la difusión de imágenes a través de una red y el problema de los generales bizantinos, que es la base de las cadenas de bloques. En todos ellos nos centraremos en cómo detectar y corregir las posibles manipulaciones sobre el proceso de consenso para alterar el resultado de la votación, modificar la imagen para que no se clasifique correctamente (usando *adversarial example*) o evitar la difusión de mensajes falseados en el caso de los generales.

## II. DETECCIÓN DE DESVIACIONES EN EL PROCESO DE CONSENSO

El proceso de consenso definido en (1) asume que todos los participantes ejecutan el mismo algoritmo. Si un nodo desea

alterar el resultado basta con enviar a los vecinos cualquier otro valor. Por ejemplo, si un nodo  $i$  envía un valor fijo  $x_i(t) = x_i(0) \forall t$ , la red completa convergerá a dicho valor  $x_i(0)$ . En general, podemos modelar el proceso de consenso con desviaciones como

$$x_i(t+1) = x_i(t) + \varepsilon \sum_{j \in N_i} [x_j(t) - x_i(t)] + u_i(t) \quad (4)$$

donde  $u_i(t)$  es el error introducido por el nodo  $i$  en la iteración  $t$ . La pregunta que se plantea es si los vecinos de  $i$  pueden detectar cuándo  $u_i(t) \neq 0$ .

### II-A. Detección de fallos en el protocolo

Despejando (1) podemos obtener la siguiente expresión:

$$dv_i(t+1) = x_i(t+1) + (\varepsilon d_i - 1)x_i(t) - \varepsilon \sum_{j \in N_i} x_j(t) \quad (5)$$

siendo  $d_i$  el grado del nodo  $i$ . Si los nodos de la red siguen el algoritmo, cada iteración debe cumplir

$$dv_i(t) = 0 \quad \forall t \geq 0 \quad (6)$$

y cuando  $dv_i(t) > 0$  el nodo puede cuestionar la validez de los datos recibidos. La primera parte de la expresión depende solo de los valores de  $i$  y el último término es la suma de los valores recibidos en la iteración anterior, así que basta con guardar esta suma para controlar el proceso. La única restricción del método es que el primer intercambio debe ser el real, es decir, el valor inicial no contendrá errores.

De esta manera se detecta que el resultado obtenido en el proceso de consenso no es correcto. La siguiente pregunta que se plantea es ¿puede corregirse esta desviación?

### II-B. Corrección de las desviaciones

Una primera aproximación es compensar la desviación cuando se detecta en los nodos de la red. Vamos a asumir, sin perder generalidad y para facilitar la explicación, que la desviación se produce una única vez y en uno de los nodos  $i$ . Todos sus vecinos  $j \in N_i$  lo detectan según (5), con  $dv_j(t) > 0$ . En la siguiente iteración, estos nodos descuentan la desviación de su valor calculado, haciendo

$$z_j(t+1) = x_j(t+1) - dv_j(t+1) \quad (7)$$

siendo  $z_j(t+1)$  el valor a intercambiar en la siguiente iteración. Si todos los nodos se comportan de esta forma, la suma se conserva y además se puede demostrar que  $u(t) = \sum_i dv_i(t+1)$ . El principal problema de esta solución es que el nodo que introduce la desviación debe colaborar para compensar el resultado, lo que no es probable.

De (5) se puede deducir cómo se reparte el exceso entre los vecinos. El nodo tramposo  $i$  se guarda  $(1 - \varepsilon d_i)u_i(t)$  para él mismo y reparte  $\varepsilon u_i(t)$  para cada uno de sus vecinos, luego

$$dv_j(t+1) = \varepsilon \sum_{k \in N_j} u_k(t), \quad \frac{dv_j(t+1)}{\varepsilon} = \sum_{k \in N_j} u_k(t) \quad (8)$$

Dividiendo la desviación obtenida en el nodo por el factor  $\varepsilon$ , cualquier nodo conoce cuál es el error total que se ha introducido en el consenso. Sin embargo, dado que los valores de  $\sum_{k \in N_j} u_k(t)$  están agregados, lo que no es posible es determinar qué nodo ha introducido el error.

Esta solución permite a los vecinos del nodo con la desviación saber que algo está mal en el proceso de consenso y corregirlo. Para que esta corrección llegue a todos los nodos se debe propagar también por la red, por lo que se efectúa un proceso de consenso paralelo

Es importante destacar que este método no impide la manipulación ni lo corrige cuando lo detecta. El proceso converge al valor manipulado. Sin embargo, cuando finaliza el proceso, los nodos disponen de la información necesaria para saber si se ha producido alguna desviación de los datos y cómo corregirla.

## III. CASOS DE APLICACIÓN

A continuación se van a presentar tres casos de aplicación de la detección y corrección de fallos a problemas para los que disponer de sistemas robustos y seguros es de vital importancia: (i) un sistema de votación electrónica, (ii) la detección de imágenes manipuladas mediante *adversarial examples* y (iii) el problema de los generales bizantinos, base de las cadenas de bloques.

### III-A. Votación distribuida

Supongamos que se desea obtener el resultado de una votación que se realiza de forma distribuida en una red. Supongamos que se dispone de  $m$  opciones y cada nodo solo puede votar a una de ellas. El proceso de consenso consistirá en intercambiar con los vecinos un vector  $x_i = (x_i^1, \dots, x_i^m)$ , donde  $x_i^k = 1$  si  $k$  es la opción seleccionada por  $i$  and  $x_i^j = 0 \forall j \neq k$ . Es decir, el vector contiene un 1 en la posición de la opción votada y ceros en todas las demás. Cuando se realiza el proceso de consenso, el resultado converge al valor medio de la votación. Para obtener el resultado general, basta con multiplicar el resultado por el número de nodos, que puede obtenerse de forma sencilla con un proceso de consenso en paralelo inicializado de una forma particular.

El proceso de consenso converge a la media de los valores iniciales. Si  $\sum_i y_i(0) = 1$ , el proceso de consenso convergerá para algún instante de tiempo  $t$  a  $y_i(t) = \frac{1}{n}$  por lo que el tamaño de la red  $n = \frac{1}{y_i(t)}$ . Ambos procesos de consenso son independientes y pueden ejecutarse en paralelo.

Extendemos el vector de votos con una columna adicional

$$(x_i | y_i) = (x_i^1, \dots, x_i^m | y_i) \quad (9)$$

e, inicialmente,  $y_i = 0 \forall i$ . Sin perder generalidad, podemos introducir un nodo adicional en la red cuyos valores iniciales sean

$$(x_0 | y_0) = (\underbrace{0, \dots, 0}_m | 1) \quad (10)$$

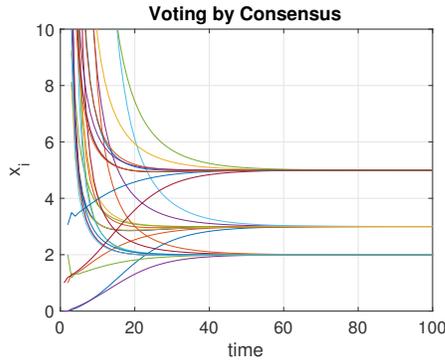


Figura 1. Evolución del proceso de consenso para calcular el resultado de una votación de forma distribuida. red aleatoria con  $n = 10$  nodos y  $m = 3$  opciones. Los votos son  $x(0) = \{3, 3, 1, 3, 2, 1, 1, 2, 3, 3\}$ , lo que da un resultado de  $\{3, 2, 5\}$ .

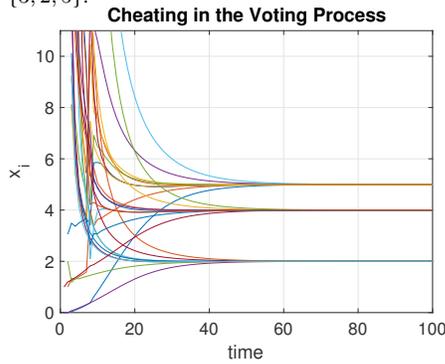


Figura 2. El nodo  $i = 7$  añade 2 votos a la primera opción y elimina un voto de la tercera en la iteración  $t = 6$ . El resultado alterado converge a  $\{5, 2, 4\}$ .

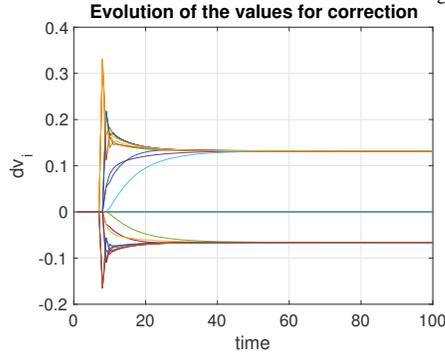


Figura 3. El proceso de consenso sobre las desviaciones detectadas converge a  $dv(t) = (0, 1320, 0, -0, 0660)$ . Aplicando (8) adaptada a votaciones, con  $\varepsilon = 0, 1650$  y  $w = 0, 4$  se obtiene la corrección deseada  $\frac{dv/\varepsilon}{w} = (2, 0, -1)$

Este nodo no afecta al resultado de la votación puesto que no vota por ninguna opción. Cuando el proceso de consenso converge, basta con que cada nodo divida los valores finales  $x_i$  por el valor obtenido en la última columna  $y_i$  para determinar el resultado general de la votación tal y como muestra (11)

$$\begin{aligned} \frac{x_i(t)}{y_i(t)} &= \frac{(x^1, \dots, x^m)}{1/n} = (nx^1, \dots, nx^m) \\ &= \left( \sum_j x_j^1, \dots, \sum_j x_j^m \right) \end{aligned} \quad (11)$$

La Figura 1 muestra un ejemplo de la aplicación de este

proceso sobre una red aleatoria de diez nodos. Los nodos eligen entre 3 opciones y la elección de cada uno es  $x(0) = \{3, 3, 1, 3, 2, 1, 1, 2, 3, 3\}$ . Por lo tanto,  $(x_1|y_1) = (0, 0, 1|0)$ ,  $(x_2|y_2) = (0, 0, 1|0)$ ,  $(x_3|y_3) = (1, 0, 0|0)$ , y así sucesivamente. Las líneas muestran la evolución del cociente  $\frac{x_i^k(t)}{y_i(t)}$ . Puede observarse que la red converge al resultado de la votación  $\{3, 2, 5\}$ .

Supongamos que el nodo  $i = 7$  desea añadir dos votos a la primera opción y eliminar un voto de la tercera:  $u_7 = (2, 0, -1)$ . El resultado del consenso convergerá a  $\{5, 2, 4\}$  y la opción ganadora sería la primera, tal y como deseaba el nodo 7 (ver Figura 2). Aplicando (5), la red puede detectar esta manipulación. La Figura 3 muestra el proceso de consenso sobre los valores de  $dv_i(t)$ . Para corregir las desviación en el proceso de voto, el consenso de las desviaciones también se extiende con un valor  $w_i$  adicional que en este caso se emplea para determinar cuántos nodos han detectado la desviación. El resultado obtenido al aplicar (8) debe dividirse por este valor  $w_i$ . En el caso de la Figura 3,  $\varepsilon = 0, 1650$  y la manipulación es detectada por cuatro nodos, lo que arroja un valor de  $w = 0, 4$ , con lo que la desviación detectada a partir de los valores de  $dv(t)$  obtenidos será  $(2, 0, -1)$  en todos los nodos.

De esta forma, al terminar el proceso de consenso, aunque el resultado esté manipulado, todos los nodos de la red tienen la información que les permite corregirlo.

### III-B. Adversarial examples

Un segundo caso de aplicación es la transmisión de imágenes a través de una red. Asumiremos que hay un nodo que tiene la imagen (por ejemplo, una imagen capturada en una cámara, como la matrícula de un coche) y quiere propagarla a todos los demás. Sin perder generalidad, será  $i = 1$ .

La representación de la imagen será a través de un vector que contiene el mapa de bits, con tantos elementos como resolución tiene la imagen y con el valor de RGB de cada pixel. El proceso de consenso se debe modificar de forma semejante al apartado anterior. A dicho vector, se le añade una columna  $y_i$  que valdrá cero para todos los nodos excepto para el que contenga la imagen, para el que  $y_1 = 0$ . Si la imagen contiene  $p$  píxeles,  $(x_1|y_1) = (x_1^1, \dots, x_1^p | 1)$  y para el resto de nodos  $(x_i|y_i) = (0, \dots, 0 | 0)$ . Cuando el proceso de consenso converge, basta con dividir el valor obtenido para cada pixel entre  $y_i(t)$  para recuperar la imagen original.

Sin embargo, un nodo malicioso puede tratar de modificar la imagen para que el sistema no la reconozca o la identifica erróneamente. Una de las técnicas empleadas se conoce como *adversarial example* [6] y consiste en añadir a la imagen una ligera distorsión de manera que los algoritmos de *deep learning* fallen y clasifiquen la imagen de forma incorrecta (ver Figura 4).

En este caso, el nodo  $i = 5$  aplica una perturbación a la imagen en la iteración  $t = 20$ , tal y como muestra la Figura 4. A la perturbación se le aplica un factor de 0.007, que es suficiente para hacer que los clasificadores que no contemplan *adversarial example* fallen, aunque para el ojo humano parece que la imagen no ha sido alterada. La



Figura 4. Adversarial example. Si a la imagen de la derecha (panda) se le añade la perturbación (centro), da como resultado una imagen que es erróneamente clasificada como un gibón (derecha)

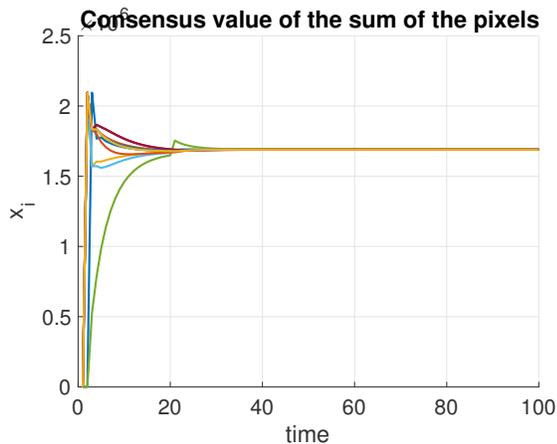


Figura 5. Alteración del consenso en la iteración 20 para incluir la imagen modificada para *adversarial example*. El valor de  $x_i$  es la suma del mapa de bits asociado a la imagen

Figura 5 muestra el proceso de consenso, donde el valor es la suma del mapa de bits asociado a la imagen. Se aprecia la perturbación introducida. Cuando converge, todos los nodos tienen la imagen alterada.

Pero de la misma forma que en el caso de la votación, la alteración se detecta inmediatamente por los vecinos utilizando (8) y esta desviación respecto a la imagen original se puede corregir cuando el proceso ha finalizado.

En esta ocasión, vamos a relajar la corrección, de manera que no deseamos recuperar la imagen original, si no simplemente tratar de paliar el efecto de la perturbación para que la imagen vuelva a ser clasificada correctamente. El resultado que muestra la Figura 6 se ha obtenido haciendo un consenso sobre las desviaciones  $dv_i$  obtenidas. De esta forma, la parte de la perturbación que se queda el nodo que modifica la imagen  $\varepsilon(1 - d_i)u_i(t)$  no se puede recuperar.

Pero como se puede ver de los resultados de la Tabla I, aunque no se recupere la imagen original, la corrección es lo suficientemente buena como para que la imagen vuelva a ser correctamente clasificada.

Los métodos actuales suelen incorporar bibliotecas que incorporan el *adversarial example*, por lo que frecuentemente



Figura 6. *Adversarial example*. Imagen corregida después del proceso de consenso. Se clasifica como panda con una precisión del 0.9986 (googleNet)

el único efecto que tiene es disminuir la precisión del clasificador en la imagen seleccionada, pero sigue clasificándose de forma correcta. La imagen corregida por consenso muestra un pixelado mayor, perceptible a simple vista (Figura 6) Sin embargo, la precisión que obtiene por el clasificador es equivalente a la imagen original. La implementación de los distintos algoritmos corresponden a las versiones de Matlab para la *Neural Network Toolbox*. En todos los casos excepto VGG-16, las imágenes son clasificadas como panda gigante, la imagen modificada obtiene una menor precisión y la imagen corregida, sin ser una corrección total, logra resultados comparables a la imagen original. En el caso de VGG-16, se observa el comportamiento detectado en el trabajo de Goodfellow [6] y la imagen manipulada se clasifica como un gibón. La corrección vuelve a clasificarse correctamente, y esta vez con una precisión notablemente superior.

### III-C. Problema de los generales bizantinos

Un tercer caso de aplicación de particular interés: el problema de los generales bizantinos. La primera solución la plantean Lamport et al. [7] y propone una solución basada en prueba de trabajo (*proof of work -POW-*) como las que incorporan en la actualidad los algoritmos para cadenas de bloques y algunas criptomonedas, como bitcoin.



Cuadro I

PRECISIÓN DE LA CLASIFICACIÓN DE IMÁGENES POR DISTINTOS ALGORITMOS. LA TABLA MUESTRA LA PRECISIÓN DE LA IMAGEN ORIGINAL, LA IMAGEN MODIFICADA POR *adversarial example*, Y LA IMAGEN CORREGIDA DURANTE EL PROCESO DE CONSENSO.

	Google Net	Res-Net 50	Alex Net	VGG-16
original	0.9873	0.9844	0.7479	0.5715 (panda)
alterada	0.9313	0.9429	0.6887	0.4439 (gibón)
corregida	0.9986	0.9863	0.9220	0.8302 (panda)

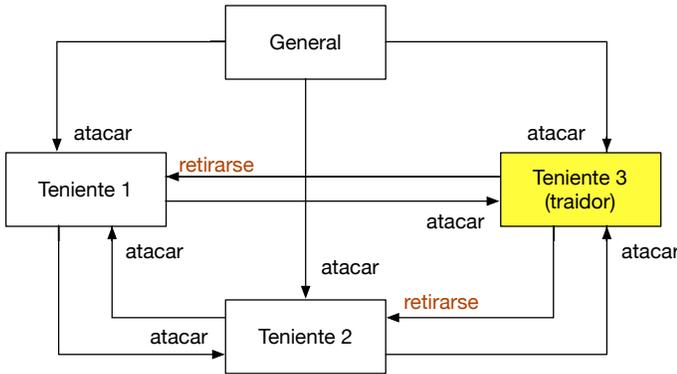


Figura 7. Problema de los generales bizantinos con 4 generales y con un mensaje simple de atacar o retirarse. Los generales solo atacarán si están seguros de que van a hacerlo todos.

El problema de los generales bizantinos consiste en lo siguiente. Supongamos que un conjunto de generales están sitiando una ciudad y tienen que coordinarse para atacar. Si no lo hacen todos al mismo tiempo, fracasarán. Dentro del grupo hay algunos generales traidores que tratarán de prevenir el acuerdo. Para poder atacar, uno de los generales mandará un mensaje al resto (los llamaremos tenientes para diferenciarlos) con la hora a la que van a atacar o indicando si se van a retirar. Cada teniente leal transmitirá el mensaje tal y como lo ha recibido, Pero los traidores alterarán el mensaje indicando una hora distinta. Por ejemplo, si el ataque es a las 9:00, un general traidor mandará un mensaje para que el ataque sea a las 8:00. Cuando lleguen las 8:00, los generales que hayan sido engañados atacarán y, al no ser suficientes, perderán. Luego, los generales restantes atacarán a las 9:00 y de nuevo perderán por no ser suficientes.

¿Cómo puede el algoritmo de consenso resolver este problema? En primer lugar, veamos cómo se pueden sincronizar mediante consenso. Supongamos que no hay traidores. La inicialización será la siguiente: el general (nodo  $i = 1$ ) la hora de ataque como valor inicial  $x_1(0) = 8$  y la columna adicional con  $y_1 = 1$ , con un vector extendido  $(x_1|y_1) = (8 | 1)$ . El resto de nodos se inicializan a cero:  $(x_i|y_i) = (0 | 0), \forall i \neq 1$ . Una vez alcanzada la convergencia, todos los generales sabrán la hora de ataque, que será  $\frac{x_i(t)}{y_i(t)} = 8$ .

Si hay algún traidor, como muestra la Figura 8, la red converge al valor modificado. En este caso, el nodo  $i = 15$  a incrementado en una la hora de ataque y la red converge a 9.

En este caso necesitamos obtener la hora precisa, por lo que es necesario hacer una corrección completa siguiendo (8). De

esta forma, los generales determinan que la hora recibida de ataque tiene una desviación de 1 hora adicional sobre la hora que contenía el mensaje original, luego la hora de ataque serán las 8:00 (Figura 9).

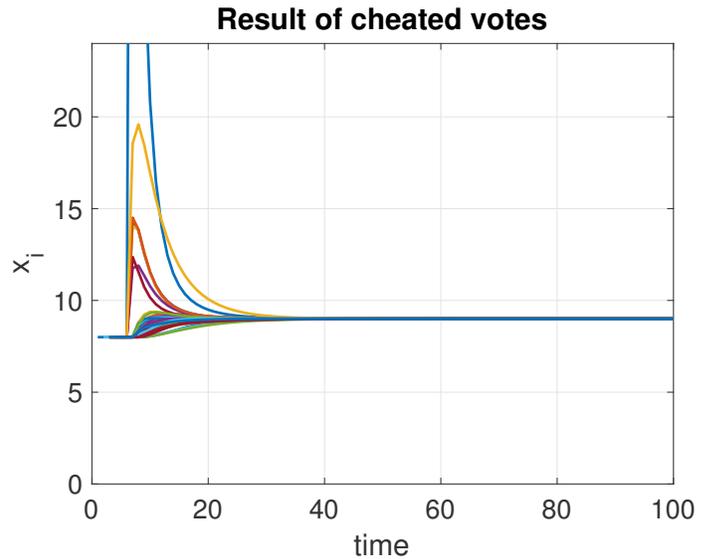


Figura 8. Proceso de consenso para determinar la hora de ataque en una red aleatoria de 50 nodos. El nodo  $i = 15$  es un traidor y cambia la hora de ataque de las 8:00 a las 9:00.

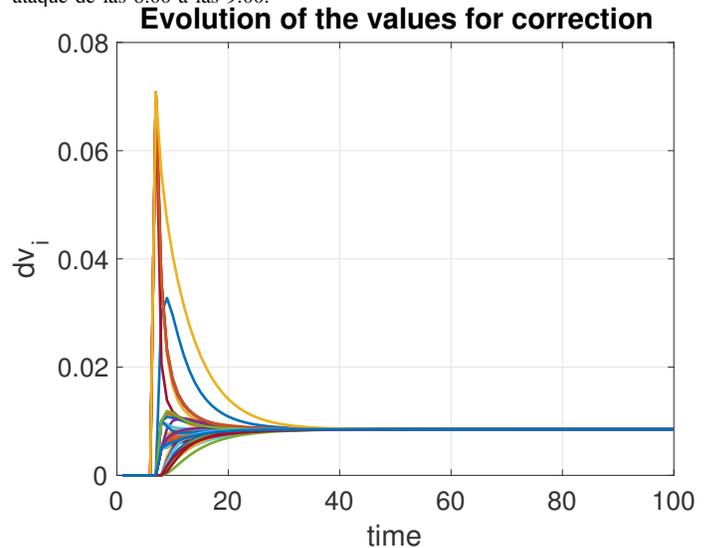


Figura 9. Evolución del factor de corrección  $dv_i(t)$ . El valor de convergencia es de 0.0085 y aplicando (8) se obtiene que la diferencia con la hora de ataque tiene un exceso de 1 hora

### III-D. Validación en redes sintéticas

Finalmente, se han realizado un conjunto de experimentos para validar cómo se degrada la solución a medida que aumentan los nodos que no siguen el algoritmo de consenso. Las pruebas se han realizado sobre redes aleatorias de 100 nodos. Se ha variado el número de nodos que no siguen el proceso desde 5 hasta 50 (representa el 50% de la red). Para cada caso, se han generado 20 redes diferentes y en cada una de ellas se han ejecutado 50 experimentos variando qué nodos

hacen trampas, con un total de 1000 ejecuciones para cada caso. En la Figura 10 se muestra el valor medio (la línea central) y la desviación típica (el área sombreada). Puede verse que siempre hay una desviación mínima, debido a la ubicación de los nodos tramosos.

Para que la solución de Lamport para problema de los generales bizantinos o la corrección planteada por Sundaram funcionen, se requiere que

1. en la red haya al menos  $3m + 1$  nodos si hay  $m$  nodos tramosos
2. al menos la mitad de los vecinos de cada nodo deben comportarse correctamente

Puesto que para estas pruebas se han generado redes aleatorias sin ningún control sobre el tipo de nodos y su conectividad, y que los ejemplos llegan a una distribución de nodos de confianza y tramosos de 50–50, es de esperar que haya casos en los que la red no sea capaz de compensar la desviación. Resulta interesante ver que el proceso de consenso sigue funcionando aunque se supere el límite de un tercio de nodos tramosos, si bien es algo que debe estudiarse en más profundidad y queda para trabajos futuros.

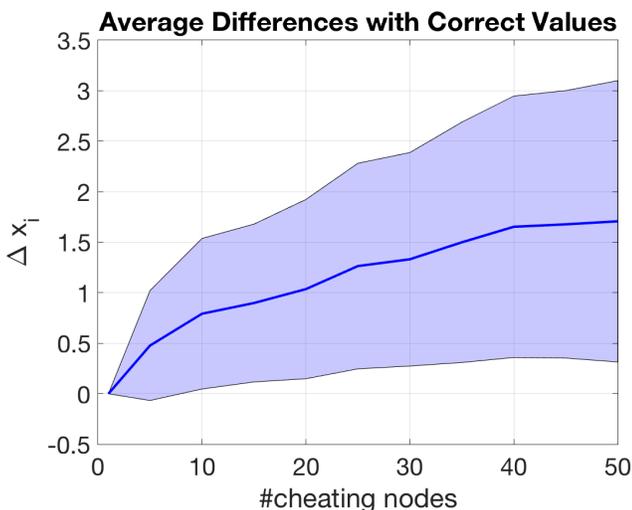


Figura 10. Desviaciones sin corregir en una red de 100 nodos variando el número de nodos tramosos. La línea central es el valor medio y el área sombreada la desviación típica.

#### IV. CONCLUSIONES

Los procesos de consenso en redes nos permiten realizar cálculos distribuidos sobre el valor de cierta función sobre los valores iniciales de la red, como la media, máximos, mínimos o valores agregados, utilizando tan solo información local y sin conocer la topología ni ninguna otra característica de la red. Sin embargo, para que el proceso funcione, todos los nodos están obligados a seguir el mismo algoritmo. Si tan solo uno de ellos realiza un cambio en uno de los pasos, el valor al que

converge la red se ve alterado. Para evitarlo, en el presente trabajo ha desarrollado un mecanismo que permite (i) detectar cuándo se produce una desviación en el proceso de consenso y (ii) corregir la desviación sobre el valor final obtenido.

Se ha mostrado su utilidad en tres problemas de interés: un sistema de votación distribuido, la alteración de imágenes usando *adversarial example* y la resolución del problema de los generales bizantinos. En todos ellos, se ha realizado una adaptación del proceso de consenso para poder hacer cálculos sobre valores agregados, aprovechando la asignación concreta de los valores iniciales y realizando un consenso doble en paralelo sobre variables independientes. En los tres casos, se ha aplicado con éxito el algoritmo de consenso con detección de desviaciones. En el caso de las votaciones, se evita que un nodo intente añadir o eliminar votos a alguna de las opciones. Para la difusión de imágenes, se pueden anular los efectos en los errores de clasificación de imágenes a las que se les ha introducido una perturbación mediante *adversarial example*. Finalmente, en el problema de los generales bizantinos se corrigen los cambios que los traidores hacen en los mensajes y se recupera la hora original. De esta manera, se consigue un método robusto que permite seguir aplicando el consenso en situaciones de incertidumbre, en las que se pueden producir fallos de forma accidental o deliberada y los valores que se transmiten no son los correctos.

Una limitación del método expuesto es que los nodos necesitan tener un número mínimo de vecinos de confianza que propague la información siguiendo el algoritmo. La cantidad de nodos tramosos no debe superar un tercio del total de nodos de la red. Los resultados obtenidos experimentalmente indican que utilizando procesos de consenso se pueden reducir estos límites, pero es necesario un estudio teórico en profundidad sobre el comportamiento del algoritmo presentado. Por otra parte, no se contempla la posibilidad de que los nodos tramosos coordinen sus ataques. Se asume que cada nodo intenta afectar el resultado de forma individual frente a toda la red. Un planteamiento que resuelva estas situaciones se contempla como trabajo futuro.

#### REFERENCIAS

- [1] R. Olfati-Saber and R. M. Murray, *Consensus problems in networks of agents with switching topology and time-delays* IEEE TAC. **49**(9), 1520–1533 (2004)
- [2] S. Sundaram, S., and C.N. Hadjicostis, *Distributed function calculation via linear iterative strategies in the presence of malicious agents*. IEEE TAC, **56**(7), 1495–1508 (2011)
- [3] M. Rafailescu, *Fault Tolerant Leader Election In Distributed Systems*. IJCSIT, **9**(1), 13–20 (2017)
- [4] Ma, C., Li, T. and Zhang, J., *Consensus control for leader-following multi-agent systems with measurement noises*. J Syst Sci Complex **23**(1), 35–49 (2010)
- [5] Wang, J.Z., Mareels, I. and Tan, Y. *Robustness of Distributed Multi-Agent Consensus*, IFAC Proceedings Volumes **41**(2), 1510–1515 (2016)
- [6] Goodfellow, Ian, Shlens, Jonathon and Szegedy, Christian. *Explaining and Harnessing Adversarial Examples*. arXiv 1412.6572. (2014)
- [7] Lamport, L. and Shostak, R. and Pease, M. *The Byzantine Generals Problem*. ACM TPLS, **4**(3), 382–401 (1982)



# Resumiendo Opiniones Negativas con *Deep Learning* y Reglas Descriptivas

Ana Valdivia, Eugenio Martínez-Cámara, M. Victoria Luzón, Francisco Herrera  
Instituto Andaluz de Investigación en Ciencia de Datos e Inteligencia Computacional  
Universidad de Granada, 18014 Granada, España  
Email: avaldivia@ugr.es, emcamara@decsai.ugr.es, luzon@ugr.es, herrera@decsai.ugr.es

**Resumen**—La tarea del Análisis de Opiniones a Nivel de Aspecto se basa en realizar un análisis sobre las opiniones con el mayor detalle posible, pues detecta aspectos específicos que aparecen en el texto de la opinión. No obstante, debido a que la información que se extrae de estos algoritmos es muy precisa, aspecto por aspecto, ésta no contribuye a obtener un conocimiento general de todo el corpus. Por ello, en este trabajo se propone una metodología que, a partir de la extracción de aspectos, se resume todo el contenido de una manera descriptiva, visualizando las conexiones más relevantes entre aspectos y una cierta polaridad, que en nuestro caso es la negativa. En este estudio, demostramos que este nuevo método puede emplearse para obtener una descripción global de un corpus entero a partir de las relaciones más significativas entre un aspecto y su polaridad.

## I. INTRODUCCIÓN

El Análisis de Opiniones (AO), que tiene como objetivo extraer sentimientos de textos escritos [1], ha alcanzado un alto interés en la sociedad debido a la gran cantidad de textos generados por usuarios que existen en blogs, foros, redes sociales, etc. Dentro de este campo, existen diferentes niveles de análisis dependiendo de su granularidad. El Análisis de Opiniones a Nivel de Aspecto (AONA)<sup>1</sup> es considerado el nivel más profundo de análisis [2]. Todos los métodos basados en AONA tienen como objetivo detectar polaridades en entidades y aspectos mencionados en el texto. Esta subtarea del AO se ha viralizado en los últimos años debido a su capacidad de extraer información muy fina y detallada.

No obstante, los métodos de AONA trabajan a nivel de entidad, lo que implica que la información que extraen es específica de cada opinión que analiza. Permiten obtener un análisis muy preciso de cada opinión, por lo que la información que se extrae a partir de estas técnicas es muy local.

Debido a ello, en este trabajo se propone una nueva metodología que extrae información general de las opiniones a través de algoritmos basados en AONA. La idea es mejorar la capacidad de generalización mediante Reglas Descriptivas (RD) que muestran las relaciones más relevantes entre los aspectos y polaridades. El proceso se basa en tres fases: (1) Extracción de aspectos de las opiniones, (2) Agrupación de aspectos que se refieren a conceptos similares y (3) Extracción de reglas descriptivas de aspectos y polaridades.

Los experimentos están centrados en las opiniones en inglés de TripAdvisor de tres monumentos culturales: la Alhambra,

la Sagrada Familia y la Mezquita de Córdoba. El propósito es el de caracterizar el conjunto de todas las opiniones negativas de estos monumentos, encontrando los aspectos más relevantes de esta clase, pues son los que ayudan a entender qué entidades se podrían mejorar de la visita. Los resultados muestran claramente que la metodología es eficaz para abordar la falta de generalización de los algoritmos AONA, y más específicamente, podría ser utilizado por el personal gerente de monumentos culturales para mejorar la experiencia de la visita cultural.

La estructura de este trabajo sigue así: la Sección II presenta una breve introducción de los conceptos básicos para una mejor comprensión de este estudio. La Sección III presenta la metodología propuesta para abordar el problema identificado. La Sección IV muestra los experimentos realizados. Finalmente, la Sección V resume las conclusiones del trabajo y propone futuras líneas de investigación.

## II. CONTEXTO

En esta sección se describen los conceptos generales en los que se basa esta metodología. Primero, se define AONA (Sección II-A), luego, se describen los algoritmos que existen basados en *Deep Learning* (DL) para AONA. Finalmente, se presentan las reglas descriptivas (Sección II-C).

### II-A. Análisis de Opiniones a Nivel de Aspecto (AONA)

AONA es una subtarea de AO que se basa en extraer aspectos y polaridades de entidades que son nombradas en opiniones. Es una tarea que aporta información granular del propósito de la opinión [3], y por lo tanto también es más precisa, ya que los usuarios no suelen mostrar el mismo sentimiento en todo un mismo texto. Por ejemplo, si analizamos el siguiente enunciado:

*“Los jardines de la Alhambra son espectaculares, pero las colas que se forman antes de entrar son horribles (deberían mejorar el sistema de entradas).”*

En el ejemplo se observa que la usuaria primero expresa un sentimiento positivo hacia los jardines, pero luego critica el sistema de entradas por las colas que se forman para acceder al monumento. Por lo tanto, la opinión global de la oración no está clara, pero si se realiza el análisis a nivel de entidad, se identifica un aspecto positivo (la Alhambra) y otro negativo (sistema de entradas).

<sup>1</sup>En inglés: *Aspect Based Sentiment Analysis* (ABSA).

Existen diferentes métodos basados en la extracción de aspectos. Los primeros que se desarrollaron, conocidos como Métodos Basados en Frecuencias, están basados en detectar los nombres simples y compuestos más frecuentes. Por ejemplo, Hu et al. identificaron características de productos a través de algoritmos con reglas de asociación [2]. Este método fue también aplicado al dominio del turismo, donde Marrese et al. extrajeron aspectos de restaurantes y hoteles en [4]. No obstante, estos métodos no detectan aspectos con una frecuencia muy baja en el texto, los cuáles también pueden ser útiles para comprender la opinión. Por otro lado, los Métodos Basados en Sintaxis se centran en analizar las relaciones sintácticas en el texto y así extraer los aspectos [5]. Para que estos métodos tengan un alto rendimiento, se necesitan definir previamente reglas sintácticas, tantas como sean posibles, lo cuál complica considerablemente el desarrollo de estos algoritmos. Por último, durante los últimos años se ha experimentado una explosión de técnicas basadas en *deep learning*. Este tipo de métodos han conseguido mejorar el rendimiento de muchos problemas dentro de las Ciencias de la Computación, consiguiendo resultados que baten a la mayoría de metodologías pasadas. AONA es un ejemplo de cómo estos algoritmos han mejorado los resultados de ciertas tareas, el cual explicaremos en la siguiente sección.

### II-B. Deep Learning (DL)

Durante los últimos años, el DL ha revolucionado gran parte de la comunidad científica. Este tipo de estructuras algorítmicas proporciona la versatilidad del aprendizaje supervisado y sin necesitar una selección previa de características para optimizar el rendimiento final. Los modelos de *deep learning* son clasificadores no lineales que pueden ajustarse a los datos de una manera más precisa. Durante la última década, este tipo de métodos han mejorado la mayoría de *baselines* de problemas del Procesamiento del Lenguaje Natural (PLN). Por ejemplo, Collobert et al. presentaron el primer estudio que introdujo el uso de Redes Neuronales Convolucionales (CNN) en tareas de PLN [6]. Poria et al. presentaron un enfoque de AONA basado en el aprendizaje profundo, que se basa en dos capas de CNN combinadas con un conjunto de patrones lingüísticos [7]. En este trabajo, se utilizará este sistema para extraer los aspectos debido al buen rendimiento que presenta en el estudio.

### II-C. Reglas Descriptivas (RD)

El aprendizaje supervisado comprende todos aquellos métodos que aprenden una función que asigna instancias a una clase previamente etiquetada. Se utilizan cuando el objetivo es predecir la clase de nuevas instancias. En cambio, el aprendizaje no supervisado son todos aquellos métodos que infieren estructuras ocultas a partir de datos no etiquetados. En este caso, se conciben como técnicas de descriptivas de datos y analizan su estructura inherente.

Una de las técnicas más populares de aprendizaje no supervisado son las RD. Se definen como el conjunto de técnicas que apuntan a descubrir un conocimiento descriptivo guiado

por una variable de clase supervisada. El objetivo principal de las RD es entender los patrones que se transmiten en los datos, en lugar de clasificar las instancias con respecto a una variable de clase mediante reglas. Recordamos que una regla es una expresión que relaciona un cierto atributo (consecuente) con un conjunto de características (antecedente). Una de las técnicas más conocidas de RD es el *Subgroup Discover* (SD). SD es una técnica de minería de datos no supervisada que describe una conjunción de características que son propia de una clase, dónde el antecedente es el conjunto característico de valores propios de la clase.

$$R: \{\text{Subgroup Conditions}\} \longrightarrow \{\text{Class}\},$$

Esta tarea no se centra en encontrar relaciones complejas en los datos, sino que intenta cubrir las instancias a partir de los datos de una manera exhaustiva.

Existen diferentes métricas que tratan de medir la calidad y relevancia de las reglas extraídas. Las más populares son [8] y [2]:

- **Cobertura:** Número de instancias cubiertas por la regla.

$$Cov(R) = \frac{|\text{Subgroup Conditions}|}{N},$$

donde  $N$  es el total de instancias del dataset.

- **Soporte:** Número de instancias en el dataset que cumplen las condiciones del antecedente (*Subgroup Conditions*) y el valor de la clase (*Class*).

$$Sup(R) = \frac{|\text{Subgroup Conditions} \cap \text{Class}|}{N}.$$

- **Confianza:** Mide el grado de cobertura de una regla en relación al conjunto de ejemplos que satisfacen el antecedente.

$$Conf(R) = \frac{|\text{Subgroup Conditions} \cap \text{Class}|}{|\text{Subgroup Conditions}|}.$$

### III. UNA NUEVA METODOLOGÍA PARA DESCRIBIR OPINIONES

El objetivo de las tareas basadas en AONA es el de obtener información detallada sobre la opinión. Estas tareas relacionan los aspectos mencionados en un texto con una polaridad específica. Sin embargo, en este trabajo detectamos la falta de generalización de estas técnicas, pues no sirven para tener una idea global del contenido de las opiniones. Por ello, proponemos una nueva metodología que combina la extracción de aspectos con técnicas de RD para obtener una descripción global de las relaciones más relevantes entre aspectos y polaridades dentro de un corpus. En este sentido, se combinan dos poderosas herramientas: una red neuronal basada en el uso de varias capas convolucionales CNN para la extracción de aspectos y un método de SD para la agregación de información, es decir, reglas que asocien un conjunto de aspectos con una cierta polaridad, en este caso, la negativa:

$$R: \{\text{aspect}_a = 1, \text{aspect}_b = 1\}$$

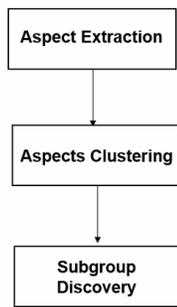


Figura 1. Flujo de la metodología propuesta.

→ {sentiment = negative}.

La metodología propuesta está basada en tres fases (ver Figura 1). La primera es extraer aspectos usando una arquitectura neuronal (Section III-A). Luego, se agrupan aspectos similares para representar el mismo concepto en una sola entidad (Sección III-B). Finalmente, se utiliza un algoritmo basado en SD para extraer la información agregada y así describir el conjunto de opiniones mediante reglas entre aspectos y polaridades (Sección III-C).

### III-A. Deep Learning para extraer aspectos

Los modelos de DL son clasificadores no lineales, lo cuáles se han convertido en el estado del arte de la gran mayoría de las tareas de PLN, debido al buen rendimiento que se ha observado. Por esta razón, en este trabajo se propone utilizar el método DL presentado en [7] para la extracción de aspectos. Este método se basa en una arquitectura convolucional, la cual utiliza como entrada las palabras que rodean a un aspecto. La arquitectura de esta red está formada por una capa de entrada, dos capas de convolución, dos capas de *pool* máximo y una capa totalmente conectada con una salida *softmax*. La salida de cada capa de convolución está calculada usando *tanh* como función de activación. Por último, se utiliza un conjunto de patrones lingüísticos que se ejecutan en la salida, lo cuales mejoran el rendimiento del método de extracción de aspectos.

### III-B. *k-nn* para agrupar aspectos

Debido a la riqueza del lenguaje natural que utilizan los humanos, cuando una persona escribe no suele utilizar la misma palabra o expresión para transmitir una misma idea. Por lo tanto, la variedad de aspectos extraídos por cualquier método de extracción de aspectos es muy amplio, y muchos de ellos pueden referirse a una misma entidad. Por ejemplo, cuando se escribe sobre la experiencia de una visita a un monumento cultural, existen muchas maneras de expresar lo mismo:

**ticket** → {onsite ticket office, senior ticket, ticket area, ticket check points, ticket office, entry ticket, service ticket, machine ticket, ticket staff, ticket price, ticket process, ...}.

La riqueza del lenguaje implica que: (1) se debe lidiar con la alta dimensionalidad de los aspectos, porque una misma entidad puede estar representada con una gran variedad de palabras idea; (2) aspectos similares semánticamente tienen representaciones diferentes. Para abordar estos problemas, en este trabajo proponemos agrupar estos aspectos en un mismo grupo o clúster, y así disminuir la dimensionalidad de aspectos extraídos.

**III-B1. De palabras a vectores:** Los *word embeddings* son representaciones de palabras en forma de vector numérico. Mikolov et al. presentaron en [9] el primer *word embedding* del estado del arte del PLN, conocido como *word2vec*. Levy et al. generalizaron este modelo teniendo en cuenta el contexto de las palabras que se representan [10]. Demostraron que los contextos sintácticos capturan otro tipo de información, por lo que su modelo recoge más similitud funcional entre las palabras representadas como vectores. Por esa razón, en este trabajo se han usado este tipo de *embeddings*<sup>2</sup>. Para esos aspectos que están compuestos por más de una palabra, su representación en *embeddings* es la media de todos los *embeddings* de las palabras que componen dicho aspecto.

**III-B2. De vectores a clústers:** El *clustering* es la tarea de agrupar un conjunto de objetos de tal manera que los que pertenecen a un mismo grupo o clúster son más similares entre ellos que entre los que pertenecen a otro clúster. En toda la literatura, existe una gran variedad de algoritmos de *clustering*. La principal diferencia entre ellos es la noción que tienen de similitud entre los objetos a la hora de formar los grupos y cómo de eficiente es la agrupación. Uno de los más populares es el *k-means*. Este es un algoritmo iterativo tiene como objetivo minimizar la varianza entre las instancias que forman un mismo grupo. Más formalmente se puede expresar como:

Dado un conjunto de elementos  $\{w_1, \dots, w_n\}$ , *k-means* agrupa las  $n$  observaciones en  $k$  clusters ( $\{C_1, \dots, C_k\}$ ), minimizando la función:

$$\arg \min_C \sum_{i=1}^k \sum_{w \in C_i} \|w - \mu_i\|^2, \quad (1)$$

donde  $\mu_i$  es la media de los puntos en  $C_i$ .

Una vez se han agrupado los aspectos similares en un mismo clúster, construimos la matriz *review-aspect* (ver Figura 1), la cuál tiene la misma estructura que una matriz *término-documento*: el elemento  $a_{ij}$  es igual a 1 si la opinión  $i$  contiene algún aspecto del cluster  $j$ , si no, es igual a 0. Por último, añadimos a esta matriz la columna de la polaridad de TripAdvisor, es decir, la puntuación que el usuario asigna a su experiencia. Como esta puntuación se basa en una escala del 1 al 5, configuramos las puntuaciones de 1 a 3 como opiniones negativas y de 4 a 5 como opiniones positivas.

<sup>2</sup><https://levyomer.wordpress.com/2014/04/25/dependency-based-word-embeddings/>

### III-C. SD para describir reglas

Los algoritmos de reglas de asociación tienen como objetivo obtener relaciones entre las variables del conjunto de datos. En este caso, las variables pueden aparecer tanto en el antecedente como en el consecuente. En los algoritmos SD, en el consecuente se asigna la variable clase, que en nuestro caso es la polaridad. Esto significa que los algoritmos de reglas de asociación se pueden adaptar perfectamente a las tareas SD.

El algoritmo Apriori [11] fue diseñado para operar en un conjunto de datos transaccional, en la que cada transacción es vista como un conjunto de ítems o características. Dado un valor de un umbral  $C$  de confianza, el algoritmo Apriori identifica todos los conjuntos de ítems que son subconjuntos de al menos  $C$  transacciones. Este algoritmo utiliza un enfoque *bottom up*, donde subconjuntos frecuentes son ampliados por un ítem a la vez (este paso se conoce como generación de candidatos). El algoritmo finaliza cuando no encuentra más ampliaciones exitosas de los conjuntos previos de ítems. Apriori-SD es la versión de este algoritmo para SD. Fue desarrollado añadiendo varias modificaciones del algoritmo Apriori C [12].

En nuestra metodología, se aplica el algoritmo Apriori-SD teniendo en cuenta:

- *ítems* son los aspectos,
- la *transacción* es la matriz review-aspects,
- el *antecedente* es el conjunto de aspectos que aparecen siempre juntos,
- el *consecuente* es la polaridad prefijada, negativa o positiva.

Por lo tanto, el objetivo es el de encontrar los aspectos más relevantes dentro de todas las opiniones negativas de TripAdvisor. Como medidas de evaluación de las reglas extraídas, proponemos emplear el Soporte y la Confianza.

## IV. EXPERIMENTOS

En esta sección se evalúa la eficacia de nuestra propuesta en diferentes bases de datos de opiniones. En esta sección, primero se describen los datasets empleados (Sección IV-A), luego se analizan cómo se han agrupado los aspectos (Sección IV-B), y los resultados de las reglas descriptivas (Sección IV-C).

### IV-A. El conjunto de datos

TripAdvisor es una compañía estadounidense basada en una página web de viajes que ofrece reseñas de experiencias de viajeros sobre hoteles, restaurantes y monumentos. Este sitio web ha formado la mayor comunidad de viajeros, alcanzando un total de 630 millones de visitantes únicos mensuales, y 350 millones de comentarios y opiniones que cubren más de 7.5 millones de alojamientos, restaurantes y atracciones en 49 mercados en todo el mundo<sup>3</sup>. Una de las características más relevantes de TripAdvisor es la gran cantidad de opiniones de millones de turistas cotidianos que contiene. De hecho,

<sup>3</sup>Fuente: <https://tripadvisor.mediaroom.com/uk-about-us>

Monumento	Opiniones	Opiniones con Aspectos	Aspectos
Alhambra	7.217	6.186	9.284
Mezquita	3.526	2.802	3.688
Sagrada Familia	34.558	26.386	18.553

Tabla I

NÚMERO DE OPINIONES, OPINIONES CON ASPECTOS Y ASPECTOS ÚNICOS.

Monumento	Positivo	Negativo
Alhambra	6.781 (93,96 %)	436 (6,04 %)
Mezquita	3.454 (97,96 %)	72 (2,04 %)
Sagrada Familia	32.664 (94,52 %)	1.894 (5,48 %)

Tabla II

DISTRIBUCIÓN DE LAS POLARIDAD POSITIVAS Y NEGATIVAS POR MONUMENTO.

sus opiniones han sido utilizadas como fuente de datos para muchos estudios de AO [4], [13]–[15].

Para evaluar nuestra propuesta, se propone aplicar la metodología en tres de los principales monumentos culturales a nivel nacional: la Alhambra (Granada), la Sagrada Familia (Barcelona) y la Mezquita de Córdoba (Córdoba). Para ello, se han recopilado 45,301 opiniones escritas en inglés, desde Julio de 2012 hasta Junio de 2016. La Tabla I muestra el número de revisiones por monumento, el número de revisiones con aspectos detectados, y el número de aspectos extraídos por el método descrito en la Sección III-A.

Antes de empezar con el análisis de los resultados, se propone analizar la distribución de los sentimientos en cada conjunto de datos. Se establecen las valoraciones de los usuarios de 1 a 3 como negativas, y de 4 a 5 como positivas. Como muestra en la Tabla II, las polaridades están altamente desbalanceadas. Las opiniones positivas son mucho más numerosas que las negativas, lo que significa que los usuarios de TripAdvisor tienden a evaluar positivamente su visita en estos monumentos.

### IV-B. Una idea, cien palabras

En esta sección se describen los resultados del agrupamiento de aspectos similares. Un aspecto puede estar formado por una palabra o más de una. En el segundo caso, la representación en *word embeddings* consiste en la media de todas las palabras que forman el aspecto. Además, también existen aspectos que no tienen representación en *word embeddings*, debido a faltas de ortografía, falta de espacios entre dos palabras, etc. En estos casos, no serán considerados para el agrupamiento, y formaran un clúster único. La Tabla III muestra la distribución de los aspectos con y sin representación en *word embeddings*.

Este análisis muestra la agrupación de los clústers, eligiendo  $k$  igual a 5, 20, 50, 100, 200, 500 and 1.000. Se observa que para valores muy pequeños de  $k$ , los clústers son formados

Monumento	Aspectos	Aspectos con embeddings	Aspects sin embeddings
Alhambra	9.284	5.430	3.854
Mezquita	3.688	2.291	1.397
Sagrada Familia	18.553	10.247	8.306

Tabla III

NÚMERO TOTAL DE ASPECTOS, CON Y SIN REPRESENTACIÓN EN VECTOR.



por gran número de aspectos, lo que hace que el clúster no sea una representación de una idea en común. En cambio, con  $k = 500$ , la mayoría de clústers corresponden a un mismo concepto, lo que indica que los resultados son coherentes con los de [16]. La Tabla IV muestra algunos ejemplos de clústers con  $k = 500$ .

Otra ventaja importante de la agrupación de aspectos es la reducción de la dimensionalidad. Como muestra la Tabla III, se obtienen 9.284, 3.688 y 18.553 aspectos por cada monumento, respectivamente. Después del proceso de agrupación con  $k = 500$ , estos aspectos se reducen a 4.353, 1.897 y 8.806 aspectos, respectivamente.

#### IV-C. Caracterizando las opiniones negativas de los monumentos culturales

Antes de aplicar algoritmos SD, se analizan frecuencias de los aspectos agrupados en los tres conjuntos de datos. En este análisis, la gran mayoría de los aspectos ocurren menos de 5 veces. La mayoría de estos aspectos corresponden a aquellas palabras sin representaciones vectoriales (ver Tabla III). Por otro lado, los aspectos agrupados obtienen valores con alta frecuencia, lo que tiene sentido porque representan varios aspectos, por ende se suma la frecuencia de cada aspecto que pertenece a ese clúster.

También se analizan los aspectos agrupados más frecuentes en los tres conjuntos de datos. Como se observa en la Tabla IV, en los tres monumentos las palabras más populares están relacionadas con temas arquitectónicos. Por consiguiente, se concluye que los usuarios tienden a describir el monumento mientras opinan sobre su visita en TripAdvisor.

Finalmente, se utiliza la versión del algoritmo Apriori-SD para identificar aspectos en opiniones negativa. Se fija el consecuente de las reglas con la clase *negativa* y se aplica el algoritmo SD para que descubra los aspectos en el lado del antecedente. Las Tablas V, VI y VII presentan las reglas más relevantes de las opiniones negativas de la Alhambra, Mezquita y Sagrada Familia, respectivamente. Se obtienen valores muy bajos tanto para las medidas de soporte y confianza. Los valores de soporte bajos se deben a la escasez de aspectos en todo el dataset. En cambio, los valores bajos de confianza se deben a la poca frecuencia de los aspectos respecto al conjunto total de opiniones.

Analizando el contenido de las reglas, se detectan algunos patrones interesantes en los datos. En las opiniones de la Alhambra, los clústers relacionados con *staff* (personal), *tickets* (tiquets) y *queue* (cola) son los más relevantes. De hecho, el clúster que representa todo lo relacionado con el personal es el que mayor confianza obtiene (0,3), lo que implica que es una característica muy distintiva de esta clase. En ese caso, se concluye que los usuarios de TripAdvisor que han visitado la Alhambra suelen quejar de ese aspecto. El mismo hecho se observa analizando las opiniones de la Sagrada Familia con el aspecto *queue*, a pesar que su confianza es relativamente más baja (0,06). También se observa que existen otro tipo de reglas descriptivas relacionadas con las características del monumento en sí: {*gardens*, *architecture*,

*building*, *church*, ...}. Esto es debido a que los usuarios de TripAdvisor suelen escribir información objetiva en sus opiniones, es decir, realizan descripciones arquitectónicas del monumento.

## V. CONCLUSIONES

Este trabajo presenta una metodología novedosa y eficaz para describir un conjunto de opiniones sobre alguna entidad según los aspectos nombrados. Los algoritmos de AONA extraen información muy detallada de opiniones a través de aspectos, pero no proporcionan una visión general de lo que contiene el texto. En consecuencia, se propuso combinar métodos de extracción de aspectos con técnicas de reglas, para así representar el contenido de un texto. Esta metodología se ha basado en tres fases: (1) Extracción de Aspectos, (2) Agrupamiento de Aspectos y (3) Reglas Descriptivas. Luego, esta metodología se ha aplicado sobre el conjunto de opiniones negativas de tres monumentos culturales, ya que aportan la información más valiosa para mejorar la experiencia del visitante a estos monumentos.

Los resultados muestran que la metodología propuesta es eficaz para resumir un conjunto de opiniones según sus aspectos. La principal ventaja es que ofrece una representación directa del contenido del texto. Se pudo describir una gran cantidad de reseñas culturales a través de aspectos y reglas descriptivas, lo que refuerza la comprensión de las opiniones. Por ejemplo, los visitantes de la Alhambra suelen quejarse del personal, del sistema de tickets y de las largas colas de espera. También se detectó que los usuarios tienden a describir los elementos del monumento visitado, lo que se considera como información objetiva. Este hecho es muy interesante porque no se observa cuando se analizan opiniones de restaurantes u hoteles. Sin embargo, se detectó que las métricas de las reglas son muy bajas, debido a dos razones: (1) el bajo porcentaje de opiniones negativas y (2) la poca frecuencia de los aspectos. También se detecta que en algunos casos, usar la polaridad de la opinión para todos los aspectos que aparecen en ella puede llevar a una mala interpretación del contenido.

Hay varias direcciones a seguir, guiadas por los resultados obtenidos. La primera está por la baja confianza de las reglas. Se propone aplicar la misma metodología pero extrayendo la polaridad del aspecto en concreto. Otra línea futura sería extender esta metodología a diferentes contextos como: restaurantes, hoteles o crítica de productos.

## AGRADECIMIENTOS

Este trabajo ha sido desarrollado gracias al proyecto TIN2017-89517-P del Gobierno de España. Eugenio Martínez-Cámara contó con el apoyo del Programa Juan de la Cierva Formación (FJCI-2016-28353).

## REFERENCIAS

- [1] B. Liu, *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press, 2015.
- [2] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004, pp. 168–177.

Monumento	Etiqueta del cluster	Contenido del Cluster
Alhambra	FAA	alhambra staff, member of staff, staff member, local staff, ticket staff, entry staff, palace staff, male staff, staffs
Alhambra	BAG	ticket price, ticket collectors, ticket seller, ticket prices, price child ticket, ticket buying, ticket printer, ticket stock
Alhambra	DBI	garden benches, architecture garden complexes, garden areas, garden tiles, manicured garden, botanical garden, garden mazes
Mezquita	B	ceilings, quoined ceilings, domed ceilings, pews, altars ceilings, mosques ceilings
Mezquita	DIE	guard, security guard, security guard berating
Mezquita	BFH	audio tour, audio, audio program, audio rental, multilingual audio, audio devices
Sagrada Familia	BAH	entrance map, gain entrance, entrance slot, entrance queues
Sagrada Familia	BCI	crowd waiting, waits, waiting costs, lines waiting, waiting queue, waiting times, waiting line
Sagrada Familia	DEB	shop, souvenir shop, merch shop, bookshop, adjoint shop

Tabla IV

EJEMPLO DE ASPECTOS AGRUPADOS EN UN MISMO CLUSTER.

Cluster Content	Rule	Cover.	Support	Conf.
staff	{ BBA = 1 } $\rightarrow$ { sentiment = negative }	0,02	< 0,01	0,30
tickets	{ DJG = 1 } $\rightarrow$ { sentiment = negative }	0,05	< 0,01	0,12
queue	{ BH = 1 } $\rightarrow$ { sentiment = negative }	0,05	< 0,01	0,13
alhambra	{ CAF = 1 } $\rightarrow$ { sentiment = negative }	0,27	0,02	0,07
gardens	{ CFE = 1 } $\rightarrow$ { sentiment = negative }	0,17	< 0,01	0,04

Tabla V

MOST RELEVANT RULES OF THE ALHAMBRA MONUMENT.

Cluster Content	Rule	Cover.	Support	Conf.
mosque	{ EHE = 1 } $\rightarrow$ { sentiment = negative }	0,18	< 0,01	0,02
architecture	{ EHF = 1 } $\rightarrow$ { sentiment = negative }	0,16	< 0,01	0,02
place	{ BDE = 1 } $\rightarrow$ { sentiment = negative }	0,06	< 0,01	0,02
arches	{ DJD = 1 } $\rightarrow$ { sentiment = negative }	0,05	< 0,01	0,02
building	{ DDB = 1 } $\rightarrow$ { sentiment = negative }	0,09	< 0,01	< 0,01

Tabla VI

MOST RELEVANT RULES OF THE MEZQUITA MONUMENT.

Cluster Content	Rule	Cover.	Support	Conf.
queue	{ ECA = 1 } $\rightarrow$ { sentiment = negative }	0,05	< 0,01	0,06
view	{ EJB = 1 } $\rightarrow$ { sentiment = negative }	0,06	< 0,01	0,06
construction	{ BGA = 1 } $\rightarrow$ { sentiment = negative }	0,08	< 0,01	0,05
church	{ CBG = 1 } $\rightarrow$ { sentiment = negative }	0,09	< 0,01	0,05
building	{ GB = 1 } $\rightarrow$ { sentiment = negative }	0,08	< 0,01	0,05

Tabla VII

MOST RELEVANT RULES OF THE SAGRADA FAMILIA MONUMENT.

- [3] K. Schouten and F. Frasincar, "Survey on aspect-level sentiment analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 3, pp. 813–830, 2016.
- [4] E. Marrese-Taylor, J. D. Velázquez, F. Bravo-Marquez, and Y. Matsuo, "Identifying customer preferences about tourism products using an aspect-based opinion mining approach," *Procedia Computer Science*, vol. 22, pp. 182–191, 2013.
- [5] Y. Zhao, B. Qin, S. Hu, and T. Liu, "Generalizing syntactic structures for product attribute candidate extraction," in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 2010, pp. 377–380.
- [6] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *Journal of Machine Learning Research*, vol. 12, no. Aug, pp. 2493–2537, 2011.
- [7] S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowledge-Based Systems*, vol. 108, pp. 42–49, 2016.
- [8] N. Lavrač, B. Kavšek, P. Flach, and L. Todorovski, "Subgroup discovery with cn2-sd," *Journal of Machine Learning Research*, vol. 5, no. Feb, pp. 153–188, 2004.
- [9] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.
- [10] O. Levy and Y. Goldberg, "Dependency-based word embeddings," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, vol. 2, 2014, pp. 302–308.
- [11] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, A. I. Verkamo *et al.*, "Fast discovery of association rules," *Advances in knowledge discovery and data mining*, vol. 12, no. 1, pp. 307–328, 1996.
- [12] V. Jovanoski and N. Lavrač, "Classification rule learning with apriori-c," in *Portuguese Conference on Artificial Intelligence*. Springer, 2001, pp. 44–51.
- [13] A. Valdivia, M. V. Luzón, and F. Herrera, "Sentiment analysis in tripadvisor," *IEEE Intelligent Systems*, vol. 32, no. 4, pp. 72–77, 2017.
- [14] B. Lu, M. Ott, C. Cardie, and B. K. Tsou, "Multi-aspect sentiment analysis with topic models," in *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on*. IEEE, 2011, pp. 81–88.
- [15] W. Kasper and M. Vela, "Sentiment analysis for hotel reviews," in *Computational linguistics-applications conference*, vol. 231527, 2011, pp. 45–52.
- [16] X. Zhang and Q. Yu, "Hotel reviews sentiment analysis based on word vector clustering," in *Computational Intelligence and Applications (ICCIA), 2017 2nd IEEE International Conference on*. IEEE, 2017, pp. 260–264.



# Minería de Opinión no supervisada en Twitter

J. Ángel Díaz-García  
Dpto. Ciencias de la Computación e I.A.  
Universidad de Granada

M. Dolores Ruiz  
Dpto. Ingeniería Informática  
Universidad de Cádiz

Maria J. Martín-Bautista  
Dpto. Ciencias de la Computación e I.A.  
Universidad de Granada

**Resumen**—La minería de opiniones ha sido ampliamente estudiada en la última década dado su gran interés en el ámbito de investigación e innumerables aplicaciones al mundo real. En este trabajo proponemos un sistema que combina reglas de asociación, generalización de reglas y análisis de sentimientos para catalogar y descubrir tendencias de opinión en la red social Twitter. A diferencia de lo extendido, se usa el análisis de sentimientos para favorecer la generalización de las reglas de asociación. Para ello, primeramente mediante minería de textos se resume un conjunto inicial de 1.7 millones de tuits captados de manera no dirigida en un conjunto de entrada para los algoritmos de reglas y análisis de sentimientos de 140718 tuits. Sobre este último conjunto se obtienen sets de reglas, estándar y generalizadas, fácilmente interpretables sobre personajes que el propio sistema revelará como interesantes.

**Index Terms**—Minería de opiniones, reglas de asociación, análisis de sentimientos, análisis de tendencias, aprendizaje no supervisado

## I. INTRODUCCIÓN

Las técnicas de minería de datos están presentes en casi todas las vertientes de estudio y desarrollo con las que los seres humanos actualmente trabajan. Hay ciertos problemas en los que estas técnicas destacan notablemente influenciados por los nuevos paradigmas económicos y sociales, donde las redes sociales han tomado un papel relevante. Es en este último punto de las redes sociales donde surge lo que conocemos como análisis de tendencias o minería de opiniones en la que se utilizan las técnicas de minería de datos para el análisis de opiniones. Objeto de estudio en el que se trata de comprender o analizar comportamientos, actividades y opiniones, por ejemplo, de consumidores de cierto producto o usuarios de cierta red social.

En este trabajo, se propone un novedoso enfoque no supervisado para el problema del análisis de tendencias y la minería de opiniones cuya base nace de la unión de dos técnicas bien diferenciadas dentro del ámbito de la minería de datos, las reglas de asociación y la minería de opiniones o análisis de sentimientos, aunque diferenciando la supervisada y no supervisada. Estas técnicas por separado, han sido ampliamente estudiadas en la literatura, donde se ha constatado la potencia de las reglas de asociación para resumir y descubrir conocimiento de un gran conjunto de datos, así como la gran utilidad del análisis de sentimientos para la realización de un análisis subjetivo de los problemas o dominios donde se aplican estas técnicas. El presente proyecto, propone por tanto combinar las tres vertientes utilizando el análisis de sentimientos para enriquecer el proceso posterior de obtención

de reglas de asociación, cuya finalidad será descubrir patrones de opinión, algo que difiere de la práctica totalidad de la literatura, donde se aplican las reglas de asociación para mejorar el paso posterior de análisis de sentimientos, polarizando entidades textuales, como por ejemplo tuits en buenos, neutros o malos sin obtención de patrones sobre aquellos factores que implicarán esos resultados, así como su interpretación.

La aportación del estudio al estado del arte en la materia de la minería de opiniones y el análisis de tendencias está por tanto en el diseño de un sistema que es capaz de trabajar con datos sin filtrado que puedan ser obtenidos de Twitter, por ejemplo de un país o lengua en concreto durante un período temporal que podrá variar. Sobre estos datos, el sistema será capaz de resumir cientos de miles de tuits, en un conjunto de reglas fácilmente interpretable sobre aquella temática o personaje que al analista interese en un determinado momento. Posteriormente se podrá generalizar y visualizar estas reglas en función de los sentimientos generados, de manera que pasaremos de un conjunto de datos desestructurado a conjuntos muy reducidos de reglas que representarán los sentimientos y las tendencias, con las que podremos categorizar las opiniones sobre personajes que el sistema obtendrá de manera no dirigida.

La metodología seguida por el sistema para conseguir este objetivo es la que podemos ver en la figura 1. Partiendo de un gran conjunto de tuits, el sistema propuesto limpia en primera instancia los datos para posteriormente realizar una selección de muestras basada en la localización de personas relevantes, tras esto el sistema polariza en paralelo los términos de los tuits y obtiene reglas de asociación sobre los mismos de manera que puede combinar la polarización de sentimientos y las reglas en un enfoque jerárquico, ofreciendo en la salida, para un mismo personaje, una nube de términos con los patrones y tendencias de opinión y un set de reglas generalizadas que nos ofrecen un nivel más de obtención de información sobre esas tendencias de opinión.

Para validar el buen funcionamiento del sistema propuesto, se han escogido dos conocidos políticos de EEUU, Donald Trump y Hillary Clinton. El motivo de escoger estos personajes, entre todos los que el sistema descubrió como relevantes en la red social Twitter, es que podemos validar acorde a los actos acaecidos en el último año las opiniones, reglas y reglas basadas en sentimientos que el modelo sea capaz de minar sobre los mismos. Cabe destacar, que podríamos haber escogido cualquier otro personaje y el sistema habría funcionado también correctamente.

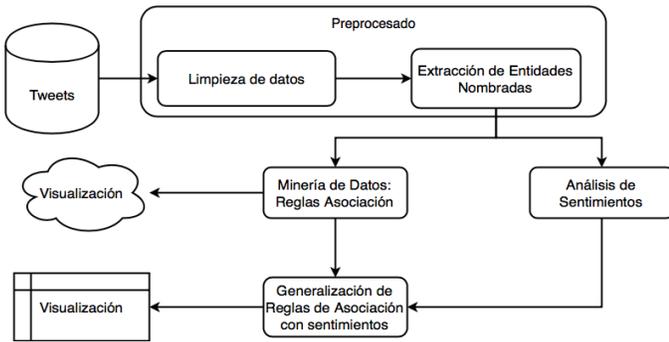


Figura 1. Flujo de la metodología.

## II. TRABAJOS RELACIONADOS

En esta sección, se estudian los trabajos relacionados con el presente estudio en función de las técnicas que se usan.

### *Reglas de asociación y minería de redes sociales.*

Uno de los principales trabajos que encontramos en el campo de las reglas de asociación es el propuesto en el 2000 por Silverstein et al. [1], donde se usan técnicas de reglas de asociación para el conocido problema de las cestas de la compra, que relaciona la compra de un determinado producto con la posibilidad de comprar otro distinto. Posteriormente, el campo ha sido ampliamente estudiado con artículos de gran interés aunque no será hasta 2010 cuando aparezca uno de los principales artículos que relacione el uso de reglas de asociación de causalidad en medios sociales. El artículo en cuestión, es propuesto por Oktay et al. [2], y en él, se estudia la relación entre la aparición de ciertos términos en preguntas de la web *Stack Overflow*, con la aparición de ciertos términos en las respuestas a estas preguntas, lo cual podría relacionarse con nuestro estudio, en el que tratamos de obtener e interpretar a relación de unos términos con otros, aunque las técnicas y el dominio difieren por completo.

Si nos centramos en nuestro dominio, Twitter, es a raíz de la publicación del trabajo de Pak and Paroubek [3], donde denotan la potencia de los datos obtenidos de Twitter para la minería de opiniones, cuando comienzan a aparecer multitud de estudios que se centran en esta red social como origen de datos sobre los cuales aplicar distintas técnicas de minería de datos, entre las que se encuentran, aunque en menor medida que otros enfoques, las reglas de asociación. En este área, encontramos trabajos como el propuesto por Cagliero y Fiori [4] o el propuesto por Erlandsson et al. [5]. En el primer estudio los autores usan reglas de asociación dinámicas, es decir, en el que las medidas de confianza y soporte cambian a lo largo del tiempo, con el fin de obtener datos sobre los hábitos y comportamientos de los usuarios en Twitter. En el segundo, se propone un análisis basado en reglas de asociación para encontrar *influencers* en Twitter, estudio que aborda el **estudio de influencias** dentro de las distintas vertientes de la minería de medios sociales. Nuestro trabajo podría verse por tanto como un híbrido entre estos dos estudios, donde se creará un modelo capaz de obtener tendencias de opinión o

relaciones acorde a personajes influyentes que a muy seguro aparecerán en el proceso. Esto se diferencia de otros trabajos como el de Meduru et al. [6] o el de Abascal-Mena et al. [7] entre otros, donde se realiza un filtrado previo de los tuits sobre los que se aplicará el procesado. En nuestro estudio, no se realiza ningún filtrado previo de los datos de Twitter.

### *Reglas de asociación y análisis de sentimientos.*

En cuanto al campo del análisis de sentimientos y las reglas de asociación, hay pocos trabajos relacionados debido al predominio de los métodos de clasificación en este área, pero, cabe mencionar el reciente interés en esta temática donde encontramos estudios como el de Hai et al. [8] donde se aplica un enfoque basado en reglas de asociación, co-ocurrencias de palabras y clustering, para obtener las características más comunes respecto a determinados grupos de palabras que puedan representar una opinión. El fin del estudio es ofrecer una vuelta de tuerca al proceso de análisis de sentimientos, que simplemente polariza una opinión, para poder refinarlo de manera que no solo se polarice ésta sino que se pueda ver acorde a qué palabras o características de opinión se ha llevado a cabo esta polarización. En este enfoque de análisis de sentimientos y aprendizaje no supervisado encontramos también el trabajo de Yuan et al. [9], en el que los autores proponen una nueva medida para la discriminación de términos frecuentes sin orientación aparente de las opiniones, lo que favorece el proceso de análisis de sentimientos posterior. El nexo de unión de estos estudios a grandes rasgos es el uso de las reglas de asociación e itemsets frecuentes para mejorar el proceso de análisis de sentimientos, esto difiere de nuestro estudio en que una vez minadas e interpretadas las reglas, se aplicará un enfoque jerárquico de las mismas basado en sentimientos, o lo que es lo mismo, usaremos el análisis de sentimientos para mejorar la interpretación de las reglas de asociación.

### *Reglas de asociación generalizadas*

Las aproximaciones jerárquicas al proceso de minado de reglas de asociación, están siendo estudiadas últimamente, debido en gran medida a la necesidad de condensar la información que estas representan de cara por ejemplo a mejorar los procesos de visualización. Un reciente ejemplo de este uso, es propuesto por Hahsler y Karpienko [10], donde se propone una visualización basada en matrices que hace uso de una simplificación jerárquica de los ítems que aparecen en las reglas de asociación. En el presente estudio, también se usa el enfoque jerárquico para simplificar las reglas, pero en lugar de hacer esto por categorías de ítems, lo haremos por sentimientos. Es en este punto, donde encontramos uno de los trabajos más íntimamente ligados al presente estudio, donde, como hemos introducido anteriormente, se usa el análisis de sentimientos para mejorar la información aportada por las reglas de asociación. Este trabajo realizado por Dehkharghani et al. [11], propone el uso de reglas de asociación para relacionar la co-ocurrencia de términos en tuits, a los que



posteriormente se clasifica en función de los sentimientos de estos términos que forman las reglas obtenidas.

Nuestro estudio, está ligado al anterior trabajo en la interpretación de correlación que hacemos de la co-ocurrencia de términos en un determinado tuit y en el uso de análisis de sentimientos para interpretar desde un enfoque subjetivo las reglas de asociación obtenidas. Nuestra propuesta se diferencia de la anterior en que realiza el análisis sin filtrar los datos de entrada, las palabras y términos son polarizados antes de la obtención de las reglas en un enfoque jerárquico basado en sentimientos tras lo que conseguimos un resumen de las reglas pudiendo, además de obtener opiniones y tendencias, categorizar los sentimientos suscitados por un personaje (también podría ser una marca, un producto o incluso un lugar) en un conjunto de tuits recopilados durante un período de tiempo.

### III. METODOLOGÍA PROPUESTA

En esta sección estudiaremos la metodología propuesta.

#### III-A. Preprocesado

Atendiendo a la naturaleza y cantidad de los datos con los que vamos a trabajar donde prácticamente cada uno de los de tuits contiene algún elemento que hace que sea totalmente distinto de los demás, se hace estrictamente necesaria una etapa previa de preprocesado de datos. Las técnicas usadas han sido las siguientes:

1. Eliminación de palabras vacías en inglés. A estas se le ha añadido la palabra *via*, que podemos considerar vacía en el ámbito que nos incumbe.
2. Eliminación de enlaces, eliminación de signos de puntuación, caracteres no alfanuméricos y valores perdidos (tuits vacíos).
3. Términos poco comunes. Se eliminan las palabras cuya frecuencia de aparición sea inferior a 30 ocurrencias, además de aquellas palabras que a pesar de tener más de 30 ocurrencias, tienen una longitud mayor de 13 letras. Esto último indicará que provienen de hashtags o unión de palabras que no tienen significado real o importancia en nuestro proceso.
4. Extracción de Entidades Nombradas: Se ha realizado un proceso de selección de instancias en la que mantendremos solo aquellas tuits que hablen de personas. Esto se ha realizado usando la técnica de *Name Entity Recognition* [12], de ahora en adelante NER, propuesta por la Universidad de Stanford y que se incluye en el *coreNLP* propuesto por Manning et al. [13] en la misma universidad. Tras la ejecución del proceso NER, obtenemos resultados bastante aceptables donde se localizan 140.718 tuits que hacen referencia a personas. Sobre este conjunto de tuits, vuelven a aplicarse las técnicas anteriores de preprocesado para refinar el proceso.
5. Paso a minúsculas: Se pasa todo el contenido a letras minúsculas para evitar discordancias. Pese a que el paso a minúsculas es uno de los principales pasos en minería de textos, en nuestro caso se aplica tras el proceso de

NER, ya que el uso de mayúsculas en los nombres propios facilita y mejora los resultados del proceso.

6. 2-gramas: Cabe esperar la posibilidad de obtener nombres compuestos por dos términos cuyo análisis en conjunto sea mucho más interesante y evite la aparición de reglas de asociación redundantes. La idea es relacionar o fusionar términos como **donald** seguido de **trump** en una sola palabra del tipo **donald-trump**. Para evitar trabajo innecesario, utilizamos un estudio que nos permita discernir si esta premisa de nombres compuestos es acertada y si sus frecuencias son tales como para ser útil esta unión. Mediante un estudio de los bigramas más comunes, damos por cierta la premisa de que hay ciertas palabras que cuya unión evitará redundancia en las reglas. Tras identificar estos términos se realiza mediante un proceso iterativo la unión de los mismos mediante un guión. Por otro lado con el análisis de comenzamos a obtener información sobre el dominio de los datos y las conversaciones en la red social que nos permitirán guiar los procesos de obtención de información posteriores.
7. Por último transformaremos los tuits en transacciones, donde cada palabra será un ítem.

Hemos obviado el proceso de *steaming* debido a que se podría perder interpretabilidad de cara a los procesos posteriores.

#### III-B. Obtención de reglas de asociación

La etapa de minería de datos ha sido llevada a cabo usando reglas de asociación mediante el algoritmo Apriori [14] sobre el que hemos aplicado valores de soporte de 0.001 y 0.0001 junto con un valor de confianza de 0.7. De esta manera obtendremos un número considerable de reglas pero cohesionadas y fuertes.

#### III-C. Polarización de términos basada en sentimientos

La motivación en este estudio para realizar análisis de sentimientos es el tener la posibilidad de polarizar las reglas de asociación que hemos obtenido en la anterior etapa de la metodología en función de los términos que aparezcan en el antecedente o consecuente de las mismas. Se ha propuesto un proceso iterativo en el que se recorre cada uno de los tuits obteniendo el sentimiento asociado a cada palabra en ese tuit, para esto usamos el paquete *syuzhet* [16] el cual hace uso del diccionario de sentimientos de mismo nombre creado por el laboratorio de literatura de Nebraska y que tiene en cuenta las 8 emociones básicas más extendidas propuestas por el psicólogo Plutchik [17]. Haciendo esto se genera una estructura de datos, en la que para cada palabra se obtiene cuantas ocurrencias tiene para cada sentimiento. Algunas de estas palabras y su sentimiento pueden verse en la figura 2.

#### III-D. Generalización de reglas basada en sentimientos

El último paso de la metodología pasa por combinar la etapa vista en el punto III-B de obtención de reglas y el punto III-C de polarización por sentimientos de los términos. Una de las diferencias más radicales del estudio frente a





Cuadro III  
REGLAS POR SENTIMIENTOS SOBRE DONALD TRUMP

Antecedente	Consecuente	Sop	Conf	Lift
{trust}	=>{donald-trump}	0.94592745	1	1
{anticipation}	=>{donald-trump}	0.59411362	1	1
{surprise}	=>{donald-trump}	0.42505133	1	1
{anger}	=>{donald-trump}	0.34565366	1	1
{fear}	=>{donald-trump}	0.29500342	1	1
{joy}	=>{donald-trump}	0.22655715	1	1
{disgust}	=>{donald-trump}	0.11293634	1	1
{sadness}	=>{donald-trump}	0.07460643	1	1

Cuadro IV  
REGLAS POR SENTIMIENTOS SOBRE HILLARY CLINTON

Antecedente	Consecuente	Sop	Conf	Lift
{trust}	=>{hillary-clinton}	0.93968872	1	1
{anger}	=>{hillary-clinton}	0.49221790	1	1
{anticipation}	=>{hillary-clinton}	0.48638132	1	1
{fear}	=>{hillary-clinton}	0.29961089	1	1
{surprise}	=>{hillary-clinton}	0.20038911	1	1
{joy}	=>{hillary-clinton}	0.14591440	1	1
{sadness}	=>{hillary-clinton}	0.07976654	1	1
{disgust}	=>{hillary-clinton}	0.07782101	1	1

Viendo la reglas se obtiene un ranking de los sentimientos que identifican a cada una de las personas estudiadas. Lo primero que sale a la vista y que podríamos concluir es que en Twitter se han emitido más tuits de apoyo y respaldo contra ambos candidatos que de otro tipo de sentimiento. Una interpretación muy interesante es la que podemos hacer del sentimiento *anger*, donde vemos cómo el 50% de los tuits que hablan de Hillary Clinton, tienen a su vez relacionados este sentimiento, por contra, Trump, tiene un 20% menos de este sentimiento, por lo que parece que la sociedad americana, a pesar de lo que parecía en Europa, estaba más en contra de Hillary Clinton que de Trump. Esto posteriormente se vería confirmado al vencer el republicano.

## V. CONCLUSIÓN Y TRABAJO FUTURO

Se ha desarrollado un modelo que es capaz de obtener patrones de comportamiento en la red social Twitter que podrían ser catalogados como tendencia y que llevados a un enfoque basado en *data streaming* podría ser incluso utilizado para categorizar las opiniones y sentimientos mayoritarios de un determinado lugar sobre un personaje en tiempo real. Se ha constatado la potencia de los métodos no dirigidos y su fácil interpretabilidad, muy cercana al lenguaje natural, en problemas similares pese a la actual hegemonía de los métodos dirigidos. Un resultado similar al obtenido con las reglas de asociación podría obtenerse mediante un estudio de frecuencias, pero estos serían computacionalmente costosos de obtener para la cantidad de datos barajada, eliminarían en cierta medida el componente no dirigido que ofrecen las reglas, al tener que decidir sobre qué palabras obtener las relaciones con otras palabras, y por último ofrecerían métricas distintas de las que pueden obtenerse de algoritmos como Apriori. De igual modo resulta interesante una comparativa entre ambos resultados y se deja la misma como trabajo futuro

que complementará al actual. Por último es necesario destacar la infinidad de temas distintos que pueden ser tratados en Twitter, haciendo que los dataset obtenidos de este sean muy interesantes y ruidosos a la par.

## AGRADECIMIENTOS

La investigación de este trabajo ha sido parcialmente apoyada por el Ministerio de Economía y Competitividad del Gobierno de España bajo el proyecto TIN2015-64776-C3-1-R.

## REFERENCIAS

- [1] C. Silverstein, S. Brin, R. Motwani, J. Ullman. Scalable techniques for mining causal structures. *Data Mining and Knowledge Discovery*, 4 (2-3) (2000), pp. 163-192
- [2] H. Oktay, B.J. Taylor, D.D. Jensen Causal discovery in social media using quasi-experimental designs. Proceedings of the first workshop on social media analytics, ACM (2010), pp. 1-9
- [3] Pak, A., & Parubek. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *LREC*
- [4] L. Cagliero and A. Fiori. Analyzing Twitter User Behaviors and Topic Trends by Exploiting Dynamic Rules. *Behavior Computing: Modeling, Analysis, Mining and Decision*. Springer, 2012 pp. 267-287.
- [5] F Erlandsson, P Bródka, A Borg, H Johnson: Finding Influential Users in Social Media Using Association Rule Learning. *Entropy* 18: 164 (2016).
- [6] Meduru, M., Mahimkar, A., Subramanian, K., Y. Padiya, P., & N. Gunjgur, P. (2017). Opinion Mining Using Twitter Feeds for Political Analysis. *International Journal Of Computer (IJC)*, 25(1), 116-123.
- [7] Abascal-Mena R., López-Ornelas E., Zepeda-Hernández J.S. User Generated Content: An Analysis of User Behavior by Mining Political Tweets. In: Ozok A.A., Zaphiris P. *Online Communities and Social Computing, OCSC 2013*. Lecture Notes in Computer Science, vol 8029. Springer, Berlin, Heidelberg
- [8] Hai Z., Chang K., Kim J. (2011) Implicit Feature Identification via Co-occurrence Association Rule Mining. In: Gelbukh A.F. (eds) *Computational Linguistics and Intelligent Text Processing, CICLing 2011*. Lecture Notes in Computer Science, vol 6608. Springer, Berlin, Heidelberg
- [9] Yuan M., Ouyang Y., Xiong Z., Sheng H. (2013) Sentiment Classification of Web Review Using Association Rules. In: Ozok A.A., Zaphiris P. (eds) *Online Communities and Social Computing, OCSC 2013*. Lecture Notes in Computer Science, vol 8029. Springer, Berlin, Heidelberg
- [10] Hahsler, M., Karpienko, R. Visualizing association rules in hierarchical groups. (2017) *Journal of Business Economics*, 87, pp. 317-335.
- [11] R. Dehkharghani, H. Mercan, A. Javeed, Y. Saygin: Sentimental causal rule discovery from Twitter. *Expert Syst. Appl.* 41(10): 4950-4958 (2014).
- [12] Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL 2005)*, pp. 363-370.
- [13] Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55-60.
- [14] R. Agrawal and R. Srikant Fast algorithms for mining association rules in large databases. 1994. *Proceedings of the 20th International Conference on Very Large Data Bases, VLDB*, pp. 487-499.
- [15] B Liu, L Zhang . A survey of opinion mining and sentiment analysis. *Mining text data*, 2012. Springer.
- [16] Jockers, M. (2017). Package 'syuzhet'. URL: <https://cran.r-project.org/web/packages/syuzhet>.
- [17] Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist*, 89(4), 344-350.
- [18] Matei Zaharia, Reynold S. Xin, Patrick Wendell, Tathagata Das, Michael Armbrust, Ankur Dave, Xiangrui Meng, Josh Rosen, Shivaram Venkataraman, Michael J. Franklin, Ali Ghodsi, Joseph Gonzalez, Scott Shenker, and Ion Stoica. 2016. Apache Spark: a unified engine for big data processing. *Commun. ACM* 59, 11 (2016), 56-65.