

# Development of Aggression Detection Technique in Social Media

**Shah Zaib**

Department of Computer Science and IT, University of Lahore, Pakpattan Campus, Pakistan  
E-mail: shah.zaib@cs.uol.edu.pk

**Muhammad Asif and Maha Arooj**

Department of Computer Science and IT, University of Lahore, Pakpattan Campus, Pakistan  
E-mail: muhammad.asif@cs.uol.edu.pk, maha.arooj@cs.uol.edu.pk

Received: 06 March 2019; Accepted: 22 March 2019; Published: 08 May 2019

**Abstract**—Due to the enormous growth of social media the potential of social media mining has increased exponentially. Individual users are producing data at unprecedented rate by sharing and interacting through social media. This user generated data provides opportunities to explore what people think and express on social media. Users exhibit different behaviors on social media towards individuals, a group, a topic or an activity. In this paper, we present a social media mining approach to perform behavior analytics. In this research study, we performed a descriptive analysis of user generated data such as users' status, comments and replies to identify individual users or groups which can be a potential threat. Tokenization technique is used to estimate the polarity of the behavior of different users by considering their comments or feedbacks against different posts on Facebook. The proposed approach can help to identify possible threats reflected by the user's behavior towards a specific event. To evaluate the approach, a data set was developed containing comments on the Facebook from different users in different groups. The dataset was divided into different groups such as political, religious and sports. Most negative users' in different groups were identified successfully. In our research, we focused only on English content; however, it can be evaluated with other languages.

**Index Terms**—Social Media Mining, threat mining, text mining, crime mining, pattern analysis, community detection.

## I. INTRODUCTION

Social Media has connected billions of people around the world and enables them to communicate anytime, anywhere. People can share their moods, everyday activities and opinions on social media. It is being used for information sharing, online procuring, and surveys of products, picture sharing and advertisement. There is a variety of social media are available such as Facebook, Twitter, WhatsApp, WeChat etc. Facebook is now the most common and largest OSN in the world that have the

more than a billion of users[1]. It has become the most widely used social network that allow the users to create service-specific profiles, post videos and photos, do chatting and share with friends, family and other social groups.

The ubiquitous use of social media has created extraordinary volumes of social data. Social media data are immense, intense, formless, and vibrant in nature, so therefore unusual challenges emerge. Mining the social media has its potential to extract patterns that can be useful for commercial, users, and consumers. At this time, Facebook rules the digital marketing area, trailed diligently by Twitter. Blogs and YouTube platforms that offer very obvious benefits but still consider less ideal[2]. Social Media Mining[3] provides methods of discerning valuable or actionable information that generates significant amount of data. It states that the distinctive procedure of mining suitable raw information contain the knowledge findings from data (KDD)[4, 5]. Opinion oriented knowledge discovery is being widely studied in different domains such as social media marketing and others.

Social media mining can increase the researchers' ability of understanding the new marvels and advanced business aptitude to offer better facilities and develop new creative opportunities[6, 7]. It is becoming popular to identify and predict the criminal activities which can help to avoid the crimes[8]. This mining requires the human data experts and programmed software to examine through enormous volumes of raw social media files such as the records use on social media, information/data sharing, online activities, online procurement behavior, networks between individuals, etc. with the purpose to distinguish patterns and developments. The effective technique can identify crime patterns by analyzing Facebook posts. These reports related to past delinquencies[9] in spite of the fact that manual analysis can be curiously time-consuming[10].

In recent years, Facebook have recorded a large increase in user numbers over 1 billion daily active users[11]. Due to a large number of users, information quickly spreads in social media networks that support

similar minded people to easily link with others[12, 13]. In 2013, Facebook highlighted some cases of unpleasant use of language, hateful words or comments in online groups and community pages. [14, 15]. The opinions or sentiments analysis can be of interest of various audiences such as manufacturing companies, political parties, social organizations, and others.

We developed a dataset from users' comments, replies in different groups or pages. The comments collected may contain personal opinions which can exhibit different behaviors. This research study focuses on identifying potential threats by mining posts and comments from users on the Facebook. We used Facebook comments for sentiment analysis for the following reasons:

1. Facebook is used by different users to express their opinions about different posts which can be a valuable source to estimate the opinions of people.
2. The number of users is increasing day by day and diversity of opinions is also increasing.
3. Facebook provides different ways to create a page, group or community which may include education, political, business and others.

In this study, tidy-text approach[16] is applied. In word-wise tokenization technique, two scenarios are used a) identification of most negative comments of users b) measuring the polarity of the behavior of users towards a particular social group. Similarly, in sentence-wise technique, most negative comments are identified. This research also compares and analyzes sentence-wise technique which provides better results as compared to word-wise technique. Sentence-wise technique considers every word in line and possibly detects the actual true meaning of user behavior.

The contributions of our research study are as follows:

1. We present an approach to estimate aggression shown by different users in different Facebook groups or community pages.
2. We conduct an experimental evaluation on set of real data to prove our method is efficient to identify the intensity of the aggression shown by the users.

The rest of the article is organized as follows: Section 2 presents the related work of social media mining research. Section 3 elaborates the methods and material and section 4 contain the results and discussion with graphical representation of different categories. And last section 5 concludes the paper.

## II. RELATED WORK

In a study [17], authors proposed a prediction method to identify the users having the negative or positive behavior towards the law enforcement authority. The research study used machine learning and dictionary-based techniques to identify the negative attitude towards

law authority. The authors divided the data in two different categories: group P, and group N, that contains the user-generated data having negative behavior against the law enforcement and the data having neutral attitude respectively. The authors identified the individual user's behavior of the user in each group.

In another research study, [18] authors identified a crime patterns by analyzing Twitter dataset based on sentiment analysis. The objective was to find the crime rate in different cities and analyze the severity of criminal activity. Similarly, in another research work[19], authors proposed an approach called "Counter-Terror Social Network Analysis and Intent Recognition (CT-SNAIR)" technique to identify the background of criminal activities. According to the authors, it is possible to generate social graph by studying the unstructured text which can be helpful to identify criminal networks.

Another approach was proposed by [20] to identify the criminal activities and improve the process of solving the crime. Authors utilize semi-supervised learning technique to identify the crime patterns. This proposed system is implemented by using the geo-spatial plot to help the analysts to identify crimes by weighting different attributes. In another study [10], author proposed crime pattern analysis technique based on similarity measures.

## III. METHODS AND MATERIAL

In this study, data is collected manually from the Facebook. This data set consists of more than 250 comments and categorized into four different domains (Politics, Religious, Terrorism and Sports). A tidy text [16] approach was utilized in this study. The overall process is as illustrated in Fig.1.

The dataset collected consists of three attributes user id, group and comment. User id is used for the identification of a user, group attribute is used to identify group (Politics, Religious, Terrorism and Sports) and comment describes the user comment on different status/group.

The dataset was categorized and tagged into four different domains (Politics, Religious, Terrorism and Sports) depending on the nature of the post/comments collected. To find the possible threats in these four categories, descriptive analysis is performed by using following techniques:

1. Word-wise tokenizing
2. Sentence wise tokenizing

In this paper we have implemented these two techniques for three different scenarios. These scenarios are used to perform behavior analysis based on the users' comments.

**Scenario 1:** User Behavior in a Social Group

**Scenario 2:** Identification of a user with most negative comments

**Scenario 3:** User Behavior in Different Social Groups is the overall approach is illustrated in Fig.1.

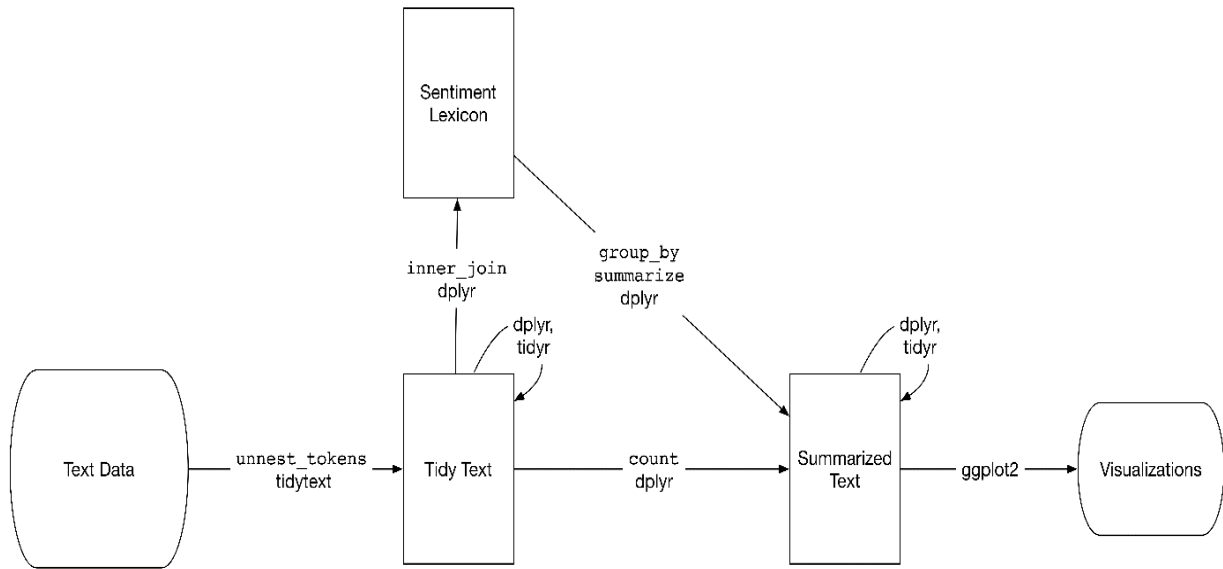


Fig.1. Overall process adapted from [16]

The overall mining process consists of following steps: Eq. 3.

**Step 1:** First, transform the data according to data frame called “tibble” which is a part of “dplyr” package.

**Step 2:** Implement the tokenizing method to restructure the data in the one token per row format.

**Step 3:** Remove unimportant words such as “the”, “and”, “of” etc.

**Step 4:** After this, finding the most common words in the whole data that the people used in their comments.

**Step 5:** There is a variety of methods and dictionaries available that can be used for sentiment analysis.

Following are three general type of lexicons exists utilized for tidy-text approach[16, 21].

**AFINN:** Calculate the score of positive and negative words between 5 to -5.

**Bing:** Identify the positive and negativity in the words

**Nrc:** calculate the emotions of different type like joy, anger etc.

**Step 6:** Now inner join of “Bing” sentiment implemented to find out the positive and negative sentiment in each word, then find that how many positive and negative comments in each group as shown in Eq1.

$$T \bowtie L = \{t U l | t \in T \wedge l \in L \wedge Sen(t U l)\} \quad (1)$$

Whereas  $T$  represents the tibble,  $L$  represents the lexicon and  $Sen$  represents the function of inner join with the condition.

After this sentiment score is calculated based on the result of inner join from (1).

$$Score(x) = \sum_{i=1}^n t_i \quad (2)$$

$$Polarity(x) = \begin{cases} -x, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (3)$$

**Step 7:** Moreover, by implementing the “Bing” sentiment in the data, we identify the persons who uses more abusive language towards the certain group as shown in

$$Person(y) = \begin{cases} -y, & |x| < |-x| \\ y, & |x| > |-x| \end{cases} \quad (4)$$

Here is the Algorithm to find the polarity of the user id

**Begin**

{

**Step 1:** Pre-process the data ( $d$ )

**Step 2:** Divide the processed data into each comment ( $c$ )

**Step 3:** Word Tokenization of each comment ( $tok$ )

**Step 4:** Repeat while ( $i \leq d$ )

Repeat ( $j \leq tok$ )

**If** ( $word == positive$ )

$countPos = countPos++$

**else**

$countNeg = countNeg++$

**end**

$Polarity = countPos - countNeg$

**end (step 4)**

}

## IV. RESULTS AND DISCUSSION

The following sections present the results of the scenarios.

### A. Scenario1: User Behavior in a Social Group

In this scenario, the major task was to find a group with the most negative comments. First, we calculated the word frequency in the data set and later calculated the sentiments for each group. Spread() function was used to calculate these sentiments. Now, inner join with the “bing” dictionary was done to analyze the polarity of each word whether it falls in a negative group or positive. Fig. 2 presents an overview of the results obtained. The y-axis in the graph shows the polarity of the comments and x-axis represents the respective user.

**B. Scenario 2: Identification of a user with most negative comments**

Similar to scenario1 when the preprocessing done, the dataset was divided into tokens and removed unnecessary words from it, such as “the, am, that, is” etc. Later, we classify those users who gave most negative comments by using “bing” library which includes more than 27, 00 words. It separates the negative words from the library which count 4,772 negative words.

First, a list of negative words from the Bing lexicon are assembled than made a data frame of how many words are in each comment to normalize the length of comments. Afterwards, identify the number of negative words in each comment and divide by the total number of words in each comment. The comments with the highest proportion of negative words are considered as most negative.

For example, in Fig. 4 user id 169 falls in sports group which uses most negative comments in all categories.

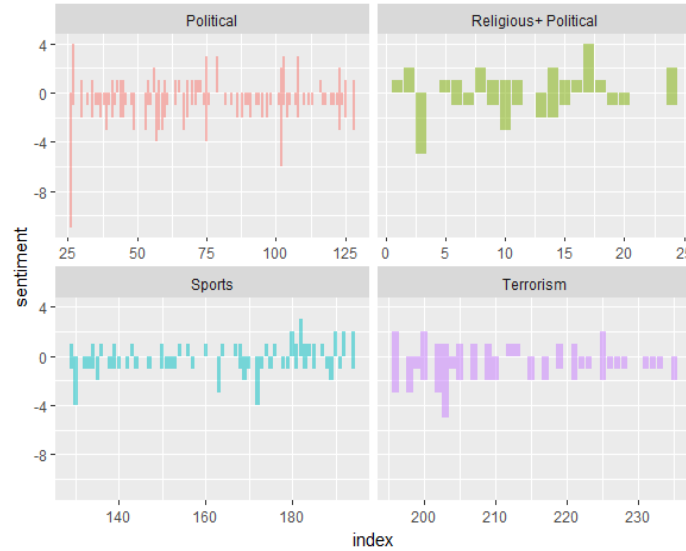


Fig.2. User Behavior

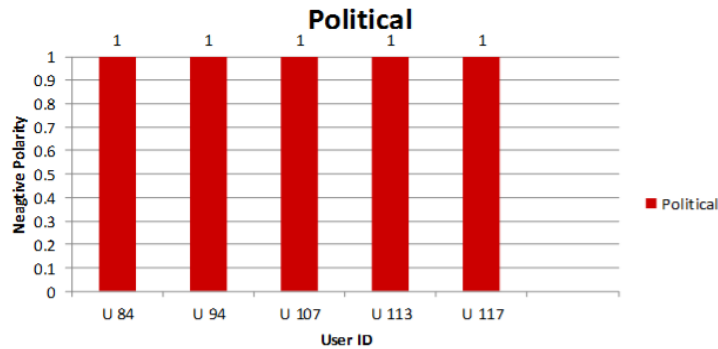


Fig.3. Political User Behavior

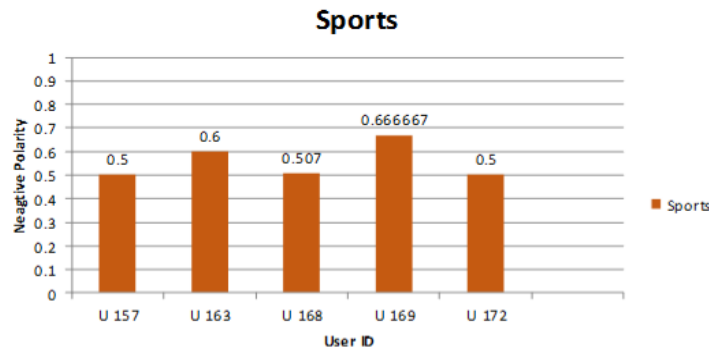


Fig.4. Sports User Behavior

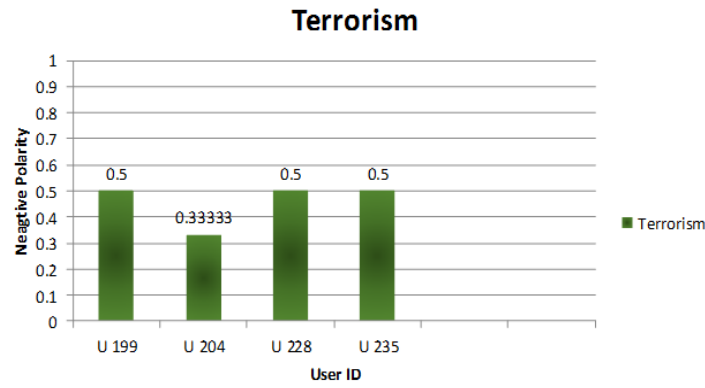


Fig.5. Terrorism User Behavior

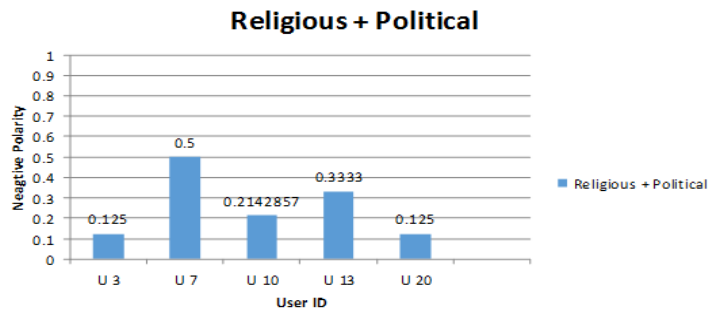


Fig.6. Religious+ Political Group

C. Scenario 3: User Behavior in Different Social Groups

For this scenario, sentence level tokenization was performed to analyze that produces more accurate results

as compared to word-wise technique. To apply sentence-wise technique, we used “sentimentr” library. In this approach,

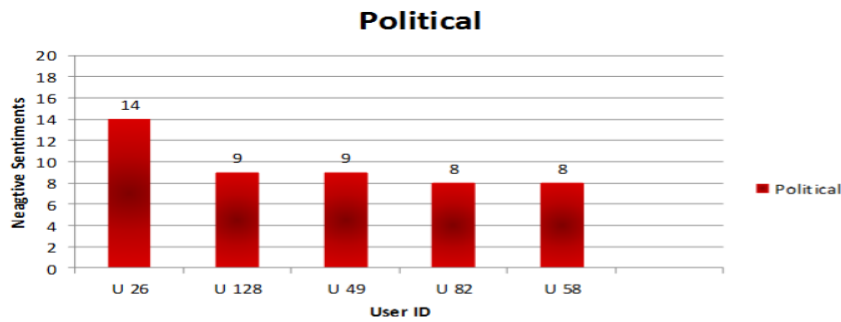


Fig.7. Political User Behavior

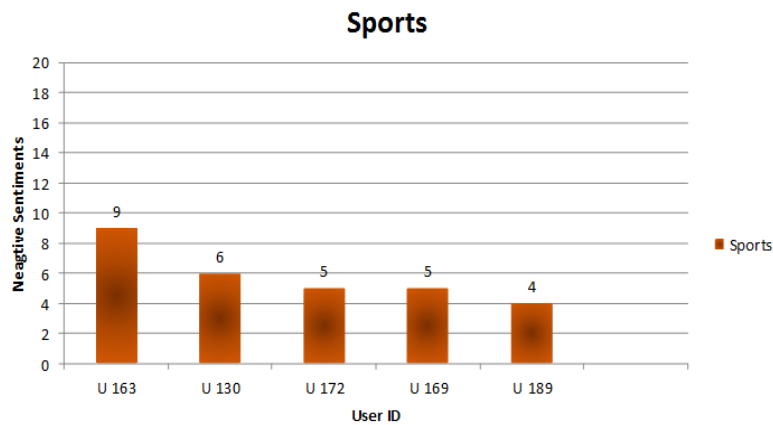


Fig.8. Sports User Behavior

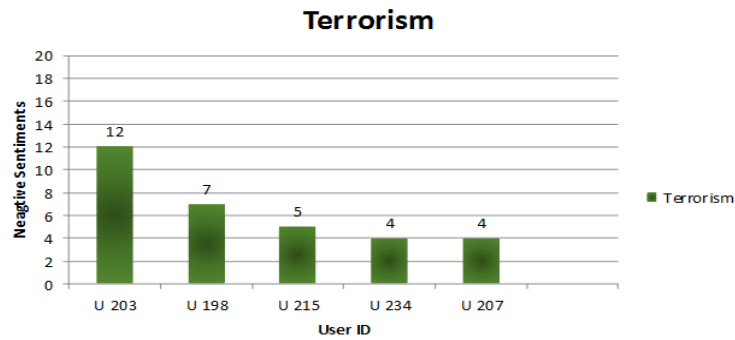


Fig.9. Terrorism User Behavior

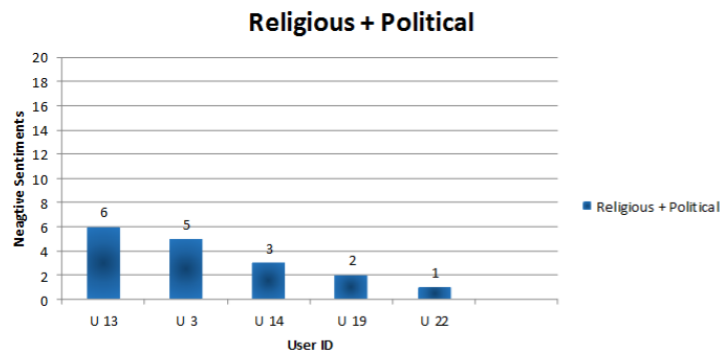


Fig.10. Religious+ Political User Behavior

The sentence numbers are, tracked and create an index and then “unnest” the sentences by words. This generates a “tibble” that has individual words by sentence within each user id. Now, with the help of “AFINN” lexicon a net sentiment score for each user id is calculated. The graphical representation of this scenario is shown below in Fig.7 -10.

#### V. CONCLUSION

This research study focused on estimating the users’ behaviors by using tidy-text approach. A dataset consists of comments and replies by users on the Facebook were collected to perform descriptive analysis. Three different scenarios were used to identify the most negative comments by a user in a particular group. The experiment was performed based on word-wise and sentence-wise tokenization approach with the help of existing sentiment lexicons. It was found that sentence-wise approach performed better. The research work can be extended to analyze different Facebook groups or community pages to identify any unusual or offensive posts by the people against government agencies or others.

#### REFERENCES

- [1] Kansara, K.B. and N.M. Shekokar, A framework for cyberbullying detection in social network. *International Journal of Current Engineering and Technology*, 2015. 5(1): p. 494-498. E-ISSN 2277 – 4106, P-ISSN 2347 – 5161
- [2] Neri, F., et al. Sentiment analysis on social media. in *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. 2012. IEEE. DOI: 10.1109/ASONAM.2012.164
- [3] Tan, P.-N., M. Steinbach, and V. Kumar, *Introduction to data mining*: Pearson addison wesley. Boston, 2005.
- [4] Han, J., J. Pei, and M. Kamber, *Data mining: concepts and techniques*. 2011: Elsevier.
- [5] Kantardzic, M., *Data mining: concepts, models, methods, and algorithms*. 2011: John Wiley & Sons.
- [6] Gundecha, P. and H. Liu, *Mining social media: a brief introduction*, in *New Directions in Informatics, Optimization, Logistics, and Production*. 2012, Informis. p. 1-17. <http://dx.doi.org/10.1287/educ.1120.0105>.
- [7] Kumar, D. and D. Bhardwaj, Rise of data mining: current and future application areas. *International Journal of Computer Science Issues (IJCSI)*, 2011. 8(5): p. 256.
- [8] Berk, R., et al., Forecasting murder within a population of probationers and parolees: a high stakes application of statistical learning. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2009. 172(1): p. 191-211. 0964–1998/09/172000.
- [9] Gwinn, S.L., et al., *Exploring crime analysis: Readings on essential skills*. 2008: International Association of Crime Analysts.
- [10] Wang, T., et al. Learning to detect patterns of crime. in *Joint European conference on machine learning and knowledge discovery in databases*. 2013. Springer 515–530, 2013.
- [11] News Room: Available from: URL: <http://newsroom.fb.com/company-info/>. 06/02/2016.
- [12] Bretschneider, U. and R. Peters. Detecting offensive statements towards foreigners in social media. in *Proceedings of the 50th Hawaii International Conference on System Sciences*. 2017, ISBN: 978-0-9981331-0-2.
- [13] King, R.A., P. Racherla, and V.D. Bush, What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *Journal of interactive marketing*, 2014. 28(3): p. 167-183, <http://dx.doi.org/10.1016/j.intmar.2014.02.001>.

