SYMBOLS: SOFTWARE FOR SOCIAL NETWORK ANALYSIS [*]

UDC (316.3:004.738.12)

Branko Arsić[1], Ljubiša Bojić[2], Ivan Milentijević[3], Petar Spalević[4], Dejan Rančić[3]

[1]Faculty of Science, University of Kragujevac, Republic of Serbia
[2]Faculty of Engineering Sciences, University of Kragujevac, Republic of Serbia
[3]Faculty of Electronic Engineering, University of Niš, Republic of Serbia
[4]Faculty of Technical Sciences, University of Priština, Republic of Serbia

Abstract. *The unique possibilities of online social networks such as real-time data access, knowledge of users' changing preferences and access to their statuses provide the possibility for innovation in the analysis of people's behavior and opinions, when compared to classical offline methods. Literature review shows lack of studies about the use of public Facebook data in Serbia for the improvement of different product sale, political or promotional campaigns, recommender systems, etc. In this paper, we present the way how data from Facebook can be collected in order to gain insight into the individuals' preferences and statuses, as well as their connection to a company's fan pages. In particular, we present data collection framework – Symbols – used for collecting individual specific data. The framework stores data into a local database and involves a module for a graph and content-based analysis of these data. The proposed framework for social network analysis can be used as a decision-making system in users' preferences implementation thus creating a space for business improvements in various areas.*

Key words: *framework, social network analysis, Facebook*

1. INTRODUCTION

Internet social networks, such as Facebook, Twitter, Instagram and LinkedIn have become very popular during the last decade. A great variety of stances are to be found online, independently from the subject of discussion. By using these services, the users share their impressions daily in the form of posts (possibly including text, photos and

videos) about upcoming events and present state of affairs, their problems, plans, novel experiences about the products, political stances, and so on. In this study we focus on Facebook profiles, since Facebook Advertisement service presents potential reach of 3,700,000 people in Serbia for the promotion of products and services. According to the self-reported information from users' Facebook profiles, there are about 48% of women and 52% of men. Information is only available for people aged 18 and older. The largest age group is currently 18-24 with a total of 800 000 users, followed by the users of the age of 25-35 1 200 000. The reported educational level in Serbia has a contribution of 66% for college and 32% for high school. At the same time, percentage for single and married relationship status is 38% and 42%, respectively.

In the recent years the potential of Internet social networks has been increasing, thus a variety of applications for online knowledge extraction is developed. Social media can yield material for sophisticated analyses in the field of marketing and branding [1]. Some examples concern the potential of social media and its effect on customer loyalty [2]. In the paper [3], the authors captured the relationship between personal networks and patterns of Facebook usage. There are also papers related to the role of Facebook social networks on intentional social actions [4], and the effectiveness of user generated content in stimulating sales [5, 6]. Also, social media are increasingly used in political context since the voters share their impressions about political stances daily in the form of statuses [7, 8]. Some of the research questions can be formulated as: 1) Can we collect data from Facebook social network related to local/global brands of some kind? For example: beers, beverages, banks, political parties, cigarettes, football games, concerts and etc. 2) In what way can we store these data and make them usable for later analysis? 3) Is there a possibility to create a network of profiles, then to add weight to corresponding edges thus reflecting the profiles similarity? Can we collect profiles' fanpages and groups they have in common for this purpose? We will provide the answers on these questions throughout the paper.

In this paper we present the Symbols[1], a framework for social network analyses, developed as an idea to potentially reinforce usage of the online social network data for marketing and branding purposes, improving political campaigns, analyzing information about people stances and opinions about social phenomena etc. The framework offers graphical presentation of statistical data for selected brands based on the social network statuses, their implications, as well as target group demographics and territorial structure. Also, we describe how collected data are stored in a database thus we can collect and analyze data about different topics simultaneously. Bearing in mind constant changes of Facebook privacy policies, we also point out how these changes can be handled without changing database scheme. In case some data are not available, which depends on the Facebook API version, data interval can be used to exclude non-relevant data. Although the software data are already used in two case studies [1, 9], the details about framework architecture and its functionalities were not published, since the software was basically developed for commercial usage. Now, the use of framework is free of charge for research tasks, but most settings have to be adjusted with the professionals in this area, sociologists and psychologists, in order to achieve the best framework performances.

---

[1] http://147.91.203.149/index.php/administration/auth/login.html
   login parameters: guest/guest

There are many similar applications on the market, which are focused on five categories of empirical Facebook research [10]: descriptive analysis of users, their motivations for using Facebook, identity presentation, the role of Facebook in social interactions, and privacy and information disclosure. One of them is the Netvizz application [11], a data collection and extraction application that operates over the Facebook social network such as friendship networks, groups and pages. Compared to this application, the Symbols offers a possibility for applying user-defined algorithms for text analysis. The Netvizz support network visualization, while the Symbols uses Gephi[2] – visualization and exploration software for graphs. NameGenWeb[3] application that originated at the Oxford Internet Institute provides the possibility of exporting a user's friendship network. Similarly, the Social Network Importer[4] is a plug-in for the NodeXL network analysis and visualization toolkit which provides a functionality for downloading personal networks and extracting data from Facebook pages. These applications are mainly focused on data collection, while the Symbols additionally provides different tools for data analysis and visualization.

Although the Symbols is mainly developed for the Serbian market, it can be easily extended for other markets. According to the best of our knowledge the Symbols is the only software of this kind in the Serbian market and wider. The main advantage of the framework is a scalable architecture which can start many projects simultaneously, collect and store data, and finally visualize the audience demography and their status messages. The sentiment analysis is implemented according to the user-specified algorithm which is based on specially developed neurolinguistic dictionaries [12, 31].

The rest of the paper is organized as follows: The second section gives an overview of the existing literature of significance for the study area. Information about Facebook API the Symbols is based on and its possibilities are provided in the third Section. The forth section is devoted to overview of Symbols software in detail and its possibilities. Case studies based on the Symbols are provided in the fifth section. The paper concludes with a summary of key points and directions for further work.

2. RELATED WORK

The literature shows the enormous potential of online social networks, since they enable a fast and effective information analysis of the present state of affairs thus reflecting the reality. Some social networks are extremely rich in content, and they typically contain a big amount of data that can be used for analysis [13]. Depending on the purpose of Internet social network, three types of data can be accessible online: 1) static demographic data, 2) semi-static data on the preferences and interests of research participants and 3) dynamic data concerning status messages and links. In general, social media can be explored from two different perspectives, graph- and content-based.

The main focus of social network analysis is on linkage data [13] i.e. graph-based principle. Linkage data means the graph structure of the social network and the communications between entities. Each user's profile, photos, events, and friends can be treated as the vertices in large graph, where these vertices are connected with different types of edges such as friendship, relationship, shared content, photo tags etc. When the

---

[2] The Open Graph Viz Platform https://gephi.org/
[3] https://apps.facebook.com/namegenweb/
[4] http://socialnetimporter.codeplex.com/

Internet social network is transformed into a mathematical graph, then different graph algorithms can be used for its analysis. We refer the readers to the book [14] for the basic graph concepts. For example, different graph-based clustering approaches such as attributed and median graphs were used recently [15, 16, 17]. Very popular spectral clustering algorithm [18] appears as a much better solution than the previous approaches. In paper [19], the authors described a set of tools, developed to analyze specific properties of Facebook social-network graphs such as degree distribution, centrality measures, scaling laws and distribution of friendships. These data are organized as an undirected graph, comprising millions of collected connections. The authors of the paper [20] studied the structure of the social graph of active Facebook users. Numerous features of the graph including the number of users and friendships, the degree distribution, path lengths, clustering, and mixing patterns were computed in this study in order to characterize the global structure of the graph. In the paper [21] the authors proposed a data collection framework to obtain individual specific data and proposed a methodology to explore users' profiles and identify the profile segments of fans of a well-known professional football club. The study presented in the paper [22] showed how a social network analysis of online data could be leveraged as an efficient approach for measuring predictors of group cohesion with respect to the relationship between structural group properties and people's subjective perceptions of cohesion. Other approaches are also presented, and they are focused on analyzing the structure of the social networks and profiles centrality (e.g. see [23, 24, 25]). An example in which the data collected from the Facebook pages are analyzed is described in the paper [26]. The main objective of the paper [27] is to study the network parameters commonly used to explain social structures. In order to analyze and evaluate the network parameters, data from three real-time Facebook accounts were extracted. All these data are retrieved using the Netvizz Application, but we should take into consideration that API access for a number of elements on public pages was removed on February 2018 and that users can no longer be distinguished in comment files, as well as user-post bipartite graphs can no longer be generated.

Another aspect of analysis is about the content of data. It relates to text, images and other data presented in social networks. In recent years there has been a great interest in the research conducted in the field of Sentiment Analysis (SA) which belongs to one of the sub-fields of Natural Language Processing (NLP). Pang et al. [28] defined the baseline method for sentiment analysis based on the sentiment lexicon. More approaches for sentiment classification of Facebook statuses can be found in papers [29, 30, 31]. An example of how business processes in a Smart City could be improved and automated using knowledge gathered from social media, web services as well as from Internet of Things (IoT) can be found in [32]. Simultaneously, the popularity of Twitter grows with the number of tweets, so an increasing number of authors have decided to test sentiment of this type of short text. An overview of the methods for analyzing sentiment data in Twitter is given in [33]. In the paper [34] the authors examined how the treatment of negation impacts the sentiment of tweets in the Serbian language. A study on social influence produced by the contents published on Facebook was presented in paper [35]. In order to see how social influence can push people to have uniformity of behavior, the collected posts were not analyzed from the NLP perspective, but as a reaction of the friends on the published posts. Additionally, classical Facebook searches using specific words could consist of useful information about world-wide occurrences such as public health information about the disease [36].

3. FACEBOOK APPLICATION PROGRAMMING INTERFACE (API)

Before we present the Symbols architecture, we will describe the software prerequisites. The Facebook Developer Platform is an umbrella term used to describe the set of services, tools, and products provided by the social networking service Facebook for third-party developers to create their own applications and services that access data in Facebook. All these software solutions rely on the Graph API, the core of Facebook Platform, which enables developers to read from and write data into Facebook. In other words, the Graph API represents a simple view of the Facebook social graph where the nodes are e.g., people, photos, events, pages and groups, while the edges between them could be friend relationships, shared content, fanpages likes, photo tags etc.

In order to access these data, Facebook authentication enables developers' applications to interact with the Graph API on behalf of Facebook users, and it provides a single-sign of mechanism across the Web, mobile, and desktop apps. What must be taken into consideration is the fact that the version of Graph API is changing constantly over time. For example, Graph API v2.7 will be deprecated in Oct 05, 2018, when v3.1 will be the only valid version. Basically, the changes we are witnessing are related to the data availability permissions because of constant data sharing privacy violation in recent years. Table 1 shows currently available permissions which users can allow to be read.

Table 1 List of permissions which could be granted or declined by user

Pages	Groups	User info
publish_pages	business_management	user_events
pages_manage_instant_articles	groups_access_member_info	user_likes
manage_pages	publish_to_groups	user_posts
pages_manage_cta	leads_retrieval	user_friends
publish_video		user_link
read_page_mailboxes		user_tagged_places
read_insights	**Features**	user_birthday
	Page Mentions	user_hometown
	Live Video API	user_age_range
	Page Public Content Access	user_gender
	Groups API	user_location
		user_videos/photos

For each of these permissions there are specifications for allowed and disallowed usage. Fig. 1 shows more details for the {user_birthday} permission.
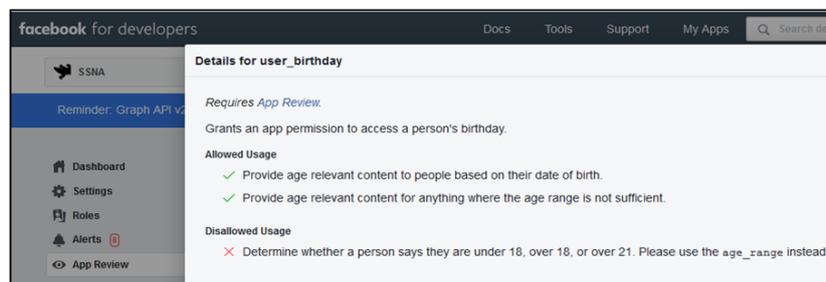
Fig. 1 Details page for selected permission

4. SYMBOLS - DATA COLLECTOR AND ANALYZER

4.1. Symbols architecture

The Symbols software for social network analyses is developed in PHP 5.3.10 programming language and MYSQL 5.5.37. The software is designed using the Model-View-Controller (MVC) software architecture, separating an application's data model, user interface, and control logic into three distinct components. The main purpose of this web framework is to integrate with Facebook API and to serve as a data collector and analyzer. In this way the software can work with publicly available static and dynamic data left by Facebook users. It is important to mention that users must give permission to Facebook app to access their profiles and collect data from these profiles. These data include gender, age, education level, school, workplace, relationship status, zodiac, birthplace, residence place, and their likes and dislikes of products, services, personalities, etc. The Facebook profiles that accept the software terms of service give us an opportunity to read their dynamic data (statuses, likes and comments) as well as static data (name, birthday, education, etc.) respecting their security and privacy. The Symbols architecture is shown in Fig. 2. In the following text we will explain all software functions and implementation details. Although, the same kinds of applications have similar architecture, the main difference is reflected in storage organization and implemented algorithms (the way data are used for making conclusions).
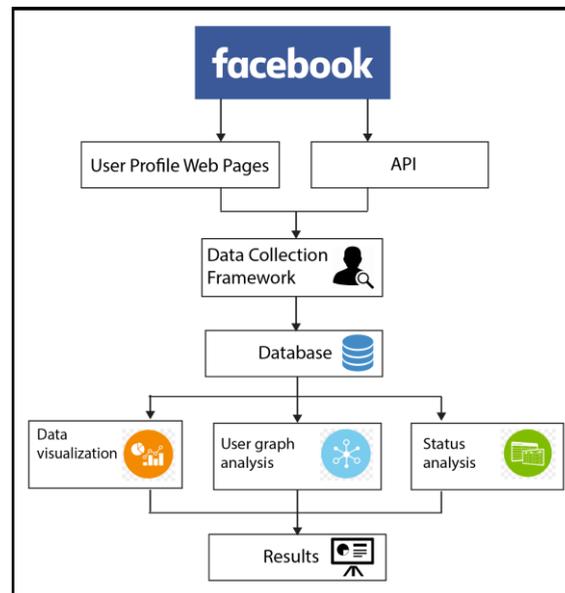
Fig. 2 The Symbols architecture

In this paper we provide all the details about the Symbols based on version 1.0 and its corresponding database scheme. Each new API version requires some database adaptations (starting from the version 2.0), but most of database scheme can be preserved. Throughout

the paper we will point out the necessary database/code changes in order to follow Facebook API changes.

4.2. Facebook application

For the purpose of the Symbols, we developed a Facebook application SSNA (Software for Social Network Analysis) based on Graph API v1.0 (see Fig. 3 for App ID). After the app users are informed that their data will be used only for scientific/analyze purposes and that it will not be given away to any third parties, they have to explicitly agree that the app is permitted to access the part of their data classified into two groups, static and dynamic data. The app had 46 activated, basic Facebook profiles during the data collection period, 12.12.2009 – 30.04.2014. One part of the demanded permissions refers to the friends' data; therefore, the data of 106,434 users can be retrieved. Note that starting from version 2.0 friends' data such as likes, statuses, photos are no longer available. Currently, these data can be collected by engaging all these profiles as basic profile i.e. separately.
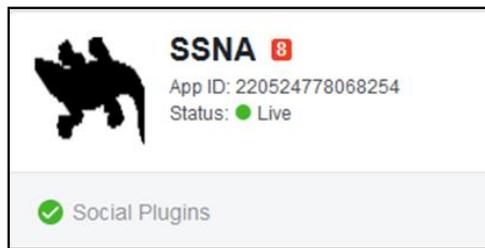
Fig. 3 SSNA Facebook application

4.3. Database scheme

Before we present the functions of the Symbols framework, a database scheme used in the background of the web application (see Fig. 4) is provided. Although the names of some tables and data stored in them can be unknown, we prefer to show a diagram here, since the framework represents a connection between Facebook API and this database. The meaning of the unknown notions will be clarified through the next section. At this point, we provide a brief description of the database structure.

The tables with green boundaries are designed to store all collected data which we want to analyze in certain date-time interval. In these tables, all Facebook data about core profiles, their friends, friends' lists (family, high school …), pages, groups and statuses are stored. Notice that the table "fbprofiles" consists of attributes which the cron job (a time-based job scheduler in Unix-like computer operating systems) uses for automatic run at a specified time and date (in our case on every half an hour).

An analyst's preferences which describe the way data will be analyzed are saved in the tables marked with the red borders. First, we can load specially-defined dictionaries which will be used for sentiment analysis of the posts. The content of dictionaries, words they consist of, meaning and categories weights will be explained later. The table "algorithms" consists of information for the set statuses which will be involved in the analysis – those statuses in which brands of interest are mentioned. This goal is achieved by
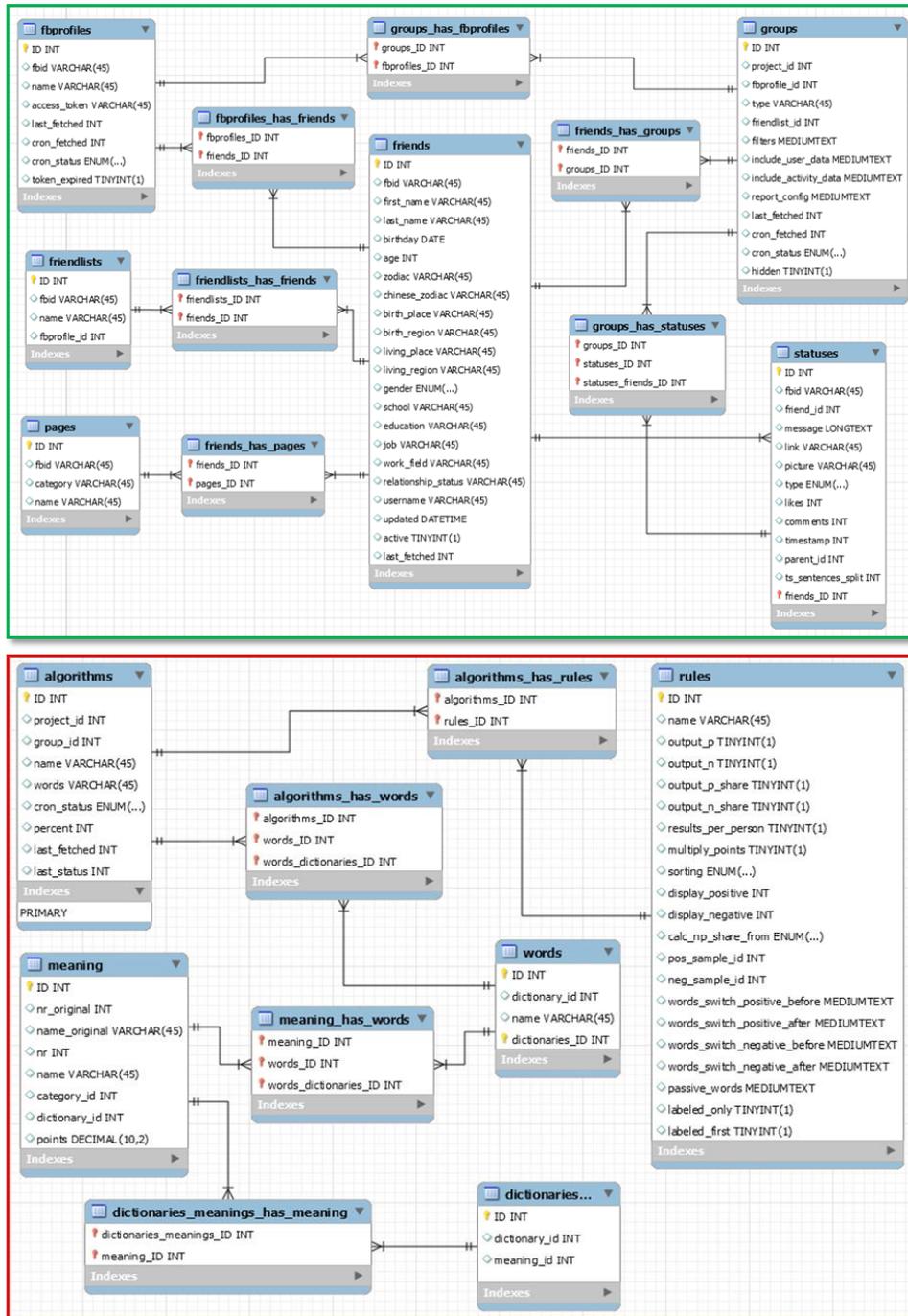
Fig. 4 Symbols database scheme

using specific keywords related to the specific brands, stored in the table rules. The table "rules" is designed for statuses sentiment analyses according to the specific words found in sentences (as parts of a status) and their position regarding the words from dictionaries. While table "fbprofiles" is used for novel statuses collecting, the cron job uses table "algorithms" for initializing new status analysis, starting with the statuses collected after the last one recorded. There are also additional tables for the results in database. Depending on the way in which data are analyzed, these tables can be modified.

4.4. Data collection framework - web application

The Symbols consists of three main functionalities: 1) Input, 2) Processing, and 3) Output.

Input functionality

Fig. 5 shows the webpage for parameters settings according to which the framework searches and collects data. Since Facebook API forbids access to the fanpages sensitive data, profiles and then personal data reading as well as corresponding fanpages likes are selected first. For example, if we want to check which type of profiles like specific
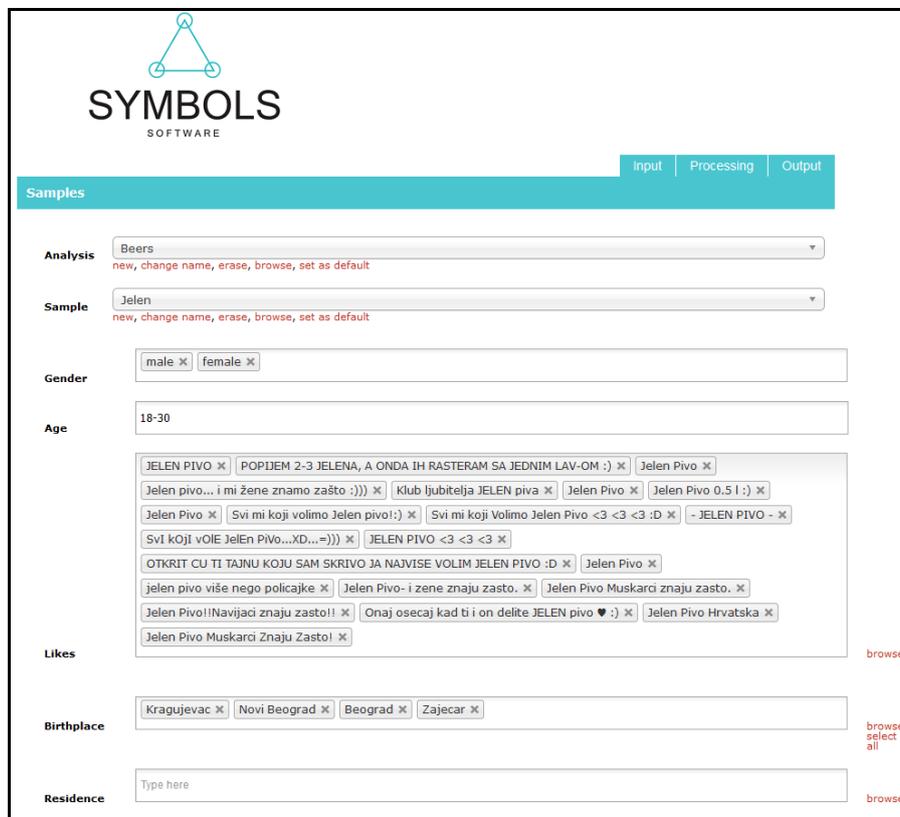
Fig. 5 One part of the possible options for users' selection

beverage, we select all grabbed users' fanpages which consist of brand label in their name. Notice that we can select many fanpages which correspond to the same brand (in Fig. 5 it is the same beer brand – "Jelen"). As default, all basic Facebook profiles are selected, as well as all their friends (in our case a sample consists of 106 434 profiles).

After the profiles data are grabbed, the cron job starts to collect their statuses every half an hour. In this way, for one analysis (for example, "Beverages"), we can create many samples corresponding to brands of interest such as Jelen, Lav, Tuborg, Heineken, etc. Besides person's age, gender, and likes, there are filters for education, job, relationship status and others.

Processing functionality

In the processing tab specific settings for short text analysis can be filled or changed (Table of symbols). In our case these short texts are publicly available statuses which are shared by Facebook profiles, the ones that were selected in the previous step. In this part we can add newly developed neurolinguistics dictionary which can be used for status analysis. Fig. 6 shows the webpage for new dictionary upload which is most often saved in a file with ".csv" extension.
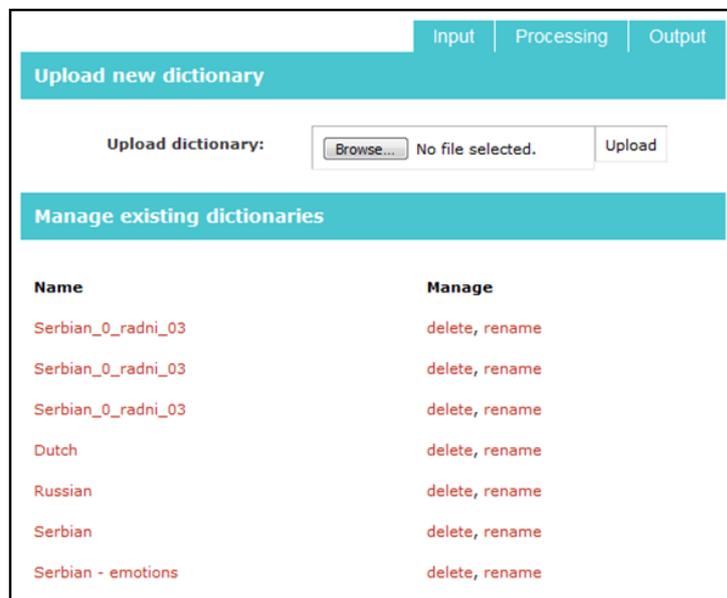
Fig. 6 Page for dictionaries upload

Each dictionary consists of word categories which have specific weights for text analysis (see Fig. 7). These categories are used to express certain levels of positive and negative emotions for status analysis. We omit this part in this paper, since more details about psychological aspects of dictionaries, its interpretation, and algorithms for status analysis could be found in paper [37].

Fig. 7 Loaded words categories with corresponding weights

The term „Algorithm" is used for user-selected words related to specific brands. In this way we filter users' statuses in order to decide whether the status about a certain brand is positive or negative. For the selected analysis, and each sample within its scope, we select the words which should be as a substring of users' statuses. Algorithmic part is based on psychological algorithms and social theories developed by Bojic et al. [38], while an overview of all negation rules in Serbian language is given in the paper [12]. The implemented algorithms are based on "Table of symbols" i.e. selected neurolinguistic dictionary which is used for final explanation of profile stances about certain topic. The outputs of the algorithms are presented in the following subsection.

Output functionality

In output menu there are three different options:

I. **Quantitative analysis** – Here, we can see some results for user profiles status analysis (see Fig. 8). For the selected date interval, and set of trademark brands we can see their mentions and tags, positive or negative connotations, etc. This part and mathematical formulas are described in more detail in the paper [37].

The key concepts and indicators of the Symbols are:

- *P Share* is an analytical tool, which measures levels of positive attitudes, emotions and moods of the targeted population.
- *N Share plus/minus* is an analytical tool, which measures levels of negative attitudes, emotions and moods of the targeted population.
- *PN Ratio* is the ratio of positive and negative attitudes/emotions in communication.
- *NP Ratio* is the ratio of negative and positive attitudes/emotions in communication.

The increase of the NP ratio, as the best early warning indicator, indicates that the company/government/other entity is starting to lose support and to obtain a negative feedback. At the same time this may indicate that the given entity may face more serious challenges, which can also be predictable by using SYMBOLS's resources.
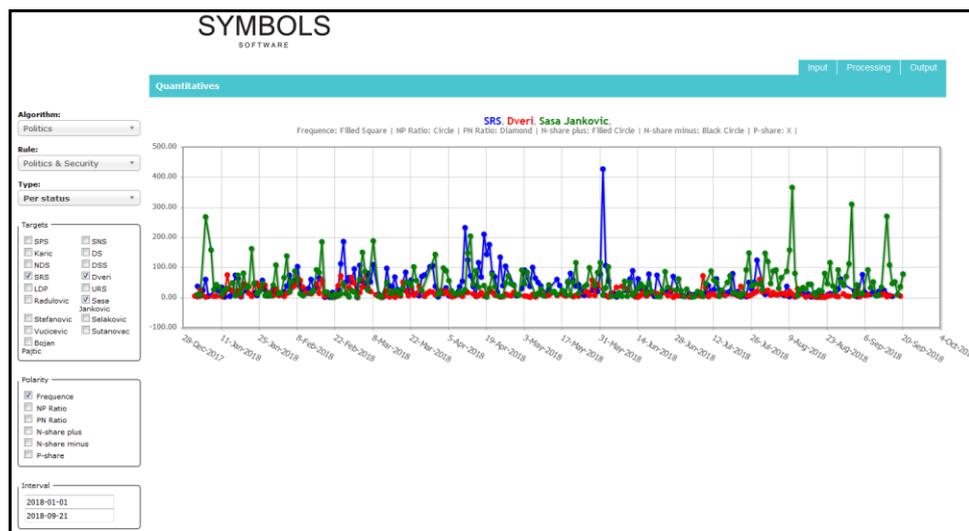

Fig. 8 Quantitative analysis of user statuses

II. **Qualitative analysis** – In this part we can obtain a list of collected statuses and their connections to specific brand, either for politics, beverages, food or banks, sport events and so on. The user can select specific brands, time interval, the number of statuses to show, sentiment (positive or negative) etc. (see Fig. 9).

III. **Demographic analysis** – shows demography and personal information of the profiles which are connected to the specific brands (see Fig. 10 and Fig. 11). For the demographic parameters a user can select the following options: hometown, residence, country, education, relationship, job status and many more. Notice that the number of options depends on the Facebook API version and supported permissions. Some graphics are very complicated because of the large number of varieties, such as hometown.
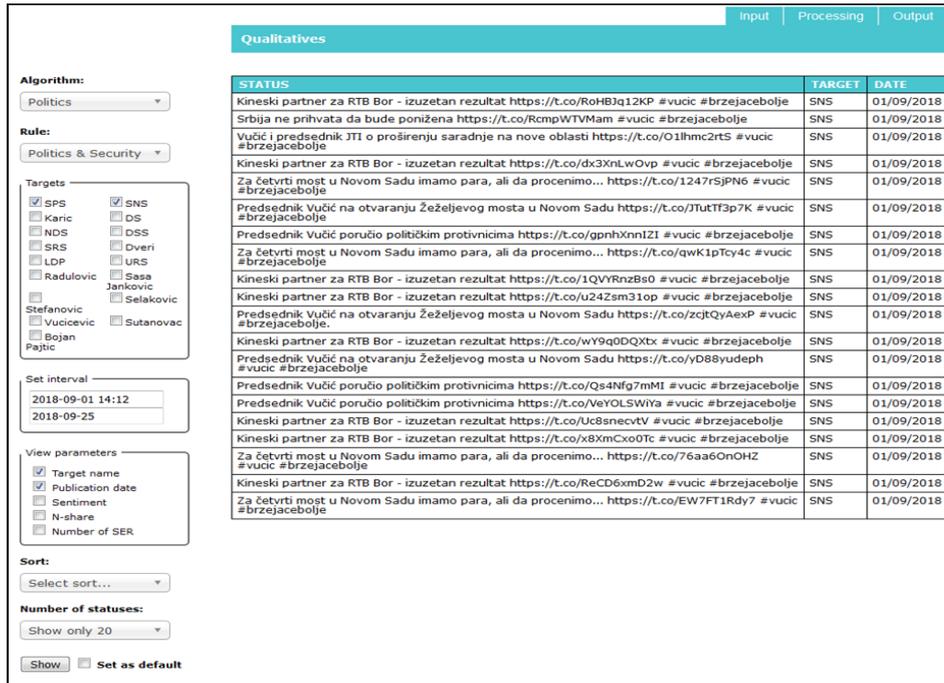
Fig. 9 Qualitative analysis of user statuses

**Fig. 10** Personal information for beer brands – education
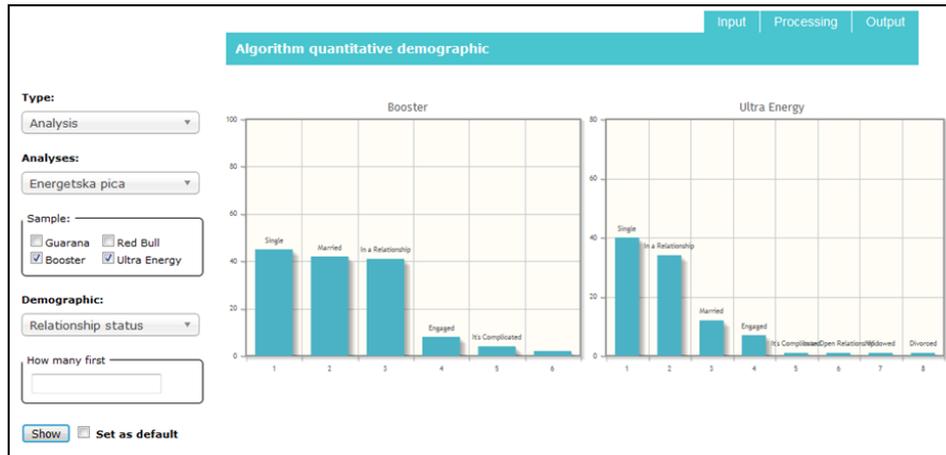
Fig. 11 Personal information for energy drink brands – relational ship status

In the meantime, the Symbols is adjusted to work with Twitter API as well. The presented database scheme can be easily adopted for these data, but our recommendation is to use separate databases for practical reasons. Compared to Facebook and LinkedIn, Twitter is imagined as a short-text social network. For example, the users do not share their personal information, even their basic profile data stay blank very often. Also, the network is directed graph, because the users do not necessarily follow one another. The details about Twitter API and data collection are omitted because this is beyond the scope of this paper.

5. CASE STUDIES

The strength of the framework is verified with three publicly available case studies. We do not describe the used methodologies in detail separately, since a reader is referred to read the papers for more information about obtained results.

The results presented in the paper [1] give an affirmative answer to the research question whether it is possible to use Facebook demographic data for product popularity improvement and point out the novel demands appearing in logistics. This study was based on a demographic representation and status analysis shared by Facebook profiles. In our previous paper [9], we applied clustering algorithm to social network profiles with the aim to obtain separate groups of people with different opinions about political views and parties. The proposed algorithm was tested on real Facebook data, and we showed that similar conclusions could be obtained in a much faster way when compared to the research conducted by marketing agencies engaged for the same purpose and tasks.

Here, we present the results of novel case study related to the political scene in 2018. One recent event that has generated a significant amount of buzz within domestic Facebook community was the 2017 Serbian presidential elections. Using these posts many questions can be answered. Reflecting the election itself, there were significant differences across the political party spectrum, and this is especially evident in last twenty years. Here, we analyze whether the public opinion from Facebook has a clear inference related to one political party

(politician) more than the other. If the presidential elections were organized at the end of 2018, we would anticipate potential winners according to the posts related to the candidates and their positive or negative content. These results should be taken with caution since the number of voters and active Facebook profiles is different. Request has been made through Facebook API in regard to all posts containing certain keywords from profiles that have given adequate permissions to the Symbols. These keywords are related to the current political leaders from Serbia, but we avoid using their names. Instead of names we use labels Candidate 1, Candidate 2, Candidate 3 etc. Posts based on each keyword made separate group for analysis. In the analysis segment, sentences written by Facebook users were analyzed in terms of their relation to different political leaders. All analyzed posts were posted in Serbia, which was one of the parameters given to the Facebook API when request was made. The analyses were conducted from January to December of 2018. Outputs included Frequency, P share, N share, PN ratio and NP Ratio for the sentences (posts) mentioning different political leaders. For simplicity purposes only Frequency and N-share plus statistics for the four leaders of the opposition will be presented here (see Fig. 12). According to the results of the analysis it was found that Facebook users from Serbia had the most positive relation to the Candidates 1 and 3. These graphics provide an insight for the frequencies and negative context of the posts per each day, thus comparing the influence their statements caused on public opinion. Also, the obtained outputs could have been helpful to a political party intending to enter the Serbian political sphere.
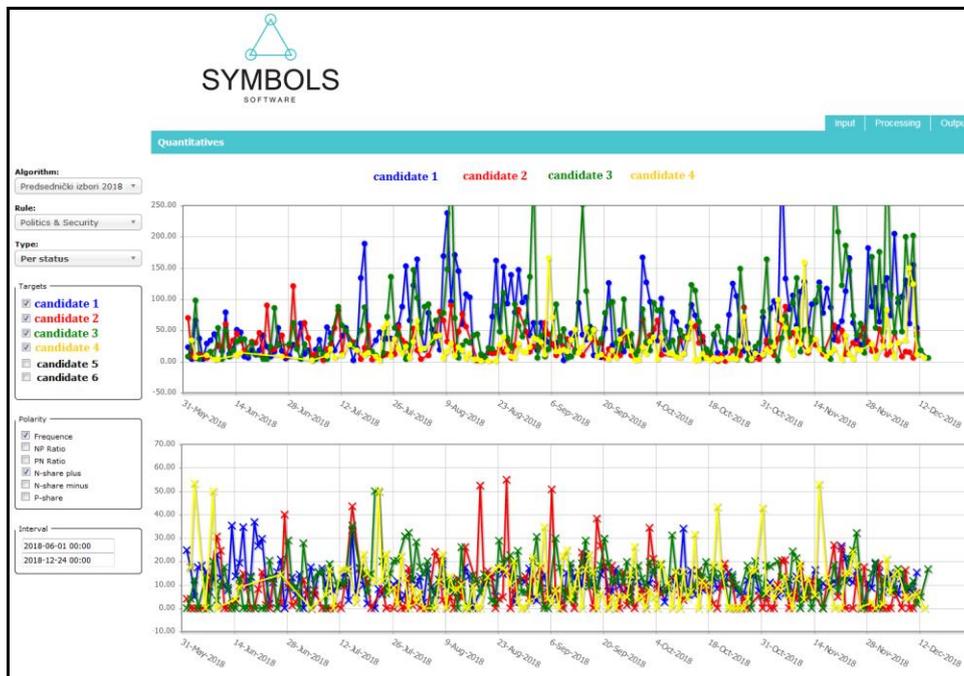

Fig. 12 Quantitative analysis of the posts for the possible presidential elections

6. CONCLUSION

Internet social networks have become unavoidable parts of people's lives. People share contents about almost every aspect of their life, from opinions on global problems, comments on events, to criticism of political parties and their leaders. During these daily online activities people have an insight into the news in local or global community, and very often they are encouraged to exchange opinions thus creating decentralized mini-campaigns about educational, cultural, political and sports novelties. This motivated us to develop a universal framework to collect and analyze these data with the focus on Serbian market.

In this paper, we presented a data collection framework Symbols for retrieving profile data from Facebook fanpages, groups and users. A component for social network analysis which includes data visualization tool, status filtering and status analysis functionalities is an integral part of the framework. The framework can analyze data from two main perspectives - as a social graph or through the users' preferences by using their statuses. The results obtained can be used for decision-making systems in marketing area, logistics and supply chains, political campaigns, recommender systems and so on. The main strength of the framework is flexibility for the Facebook API changes, but also a scalability which is reflected in an easy implementation of new algorithms. The strength of the framework is verified with public case studies.

REFERENCES

[1] B. Arsić, P. Spalević, Lj. Bojić, A. Crnišanin, "Social networks in logistics system decision-making," in *Proceedings of 2nd Logistics International Conference (LOGIC 2015)*, pp. 166–171, Serbia, May 21–23, 2015 (ISSN: 978–86–7395–339–7).

[2] C.H. Baird, G. Parasnis, "From social media to social customer relationship management," Strategy and Leadership, vol. 39, no. 5, pp. 30–37, 2011. Available: https://doi.org/10.1108/10878571111161507

[3] N. Park, S. Lee, J.H. Kim, "Individuals' personal network characteristics and patterns of Facebook use: a social network approach," Computers in Human Behavior, vol. 28, no. 5, pp. 1700–1707, 2012. Available: https://doi.org/10.1016/j.chb.2012.04.009

[4] C.M. Cheung, M.K. Lee, "A theoretical model of intentional social action in online social networks," Decision Support Systems, vol. 49, no. 1, pp. 24–30, 2010. Available: https://doi.org/10.1016/j.dss.2009.12.006

[5] C. Forman, A. Ghose, B. Wiesenfeld, "Examining the relationship between reviews and sales: the role of reviewer identity discloser in electronic markets," Information Systems Research, vol. 19, no. 3, pp. 291–313, 2008. Available: https://doi.org/10.1287/isre.1080.0193

[6] T. Hennig-Thurau, K. P. Gwinner, G. Walsh, D. D. Gremier, "Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the Internet," Journal of Interactive Marketing, vol. 18, no.1, pp. 38–52, 2004. Available: https://doi.org/10.1002/dir.10073

[7] D. Zeng, H. Chen, R. Lusch, S. H. Li, "Social media analytics and intelligence," Intelligent Systems, IEEE, vol. 25, no. 6, pp. 13–16, 2010. Available: https://doi.org/10.1109/MIS.2010.151

[8] S. Wattal, D. Schuff, M. Mandviwalla, C. B. Williams, "Web 2.0 and politics: the 2008 US presidential election and an e-politics research agenda," Mis Quarterly, vol. 34, no.4, pp. 669–688, 2010. Available: https://doi.org/10.2307/25750700

[9] B. Arsić, M. Bašić, P. Spalević, M. Ilić, M. Veinović, "Facebook profiles clustering," in *Proceedings of 6th International Conference on Information Society and Technology (ICIST 2016)*, Serbia, 28 February – 2 March 2016, pp. 154–158 (ISBN: 978–86–85525–16–2).

[10] R. E. Wilson, S. D. Gosling, L. T. Graham, "A review of Facebook research in the social sciences," Perspectives on psychological science, vol. 7, no. 3, pp. 203–220, 2012. Available: https://doi.org/10.1177/1745691612442904

[11] B. Rieder, "Studying Facebook via data extraction: the Netvizz application," In *Proceedings of the 5th annual ACM web science conference*, Paris, France, May 02–04, pp. 346–355. ACM, 2013.

[12] M. Kovačević, *Sintaksička negacija u srpskome jeziku*. Izdavačka jedinica Univerziteta u Nišu, 2002.

[13] CC. Aggarwal, "An Introduction to Social Network Data Analytics," In: Aggarwal CC (eds.) Social Network Data Analytics. New York: Springer US, pp. 1–15, 2011. Available: https://doi.org/10.1007/978–1–4419–8462–3_1

[14] D. Cvetković, P. Rowlinson, S. Simić, *An Introduction to the Theory of Graph Spectra*. London Mathematical Society Student Texts. Cambridge University Press, 2010.

[15] S. Günter, H. Bunke, "Self-organizing map for clustering in the graph domain," Pattern Recognition Letters, vol. 23, no. 4, pp. 405–417, 2002. Available: https://doi.org/10.1016/S0167–8655(01)00173–8

[16] F. Serratosa, R. Alquézar, A. Sanfeliu, "Synthesis of function described graphs and clustering of attributed graphs," International Journal of Pattern Recognition and Artificial Intelligence, vol. 16, no. 6, pp. 621–655, 2002. Available: https://doi.org/10.1142/S0218001402001915

[17] X. Jiang, A. Münger, H. Bunke, "An median graphs: properties, algorithms, and applications," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 23, pp. 1144–1151, 2001. Available: doi.ieeecomputersociety.org/10.1109/34.954604

[18] U. von Luxburg, "A tutorial on spectral clustering", Statistics and Computing, vol. 17(4), pp. 395–416, 2007. Available: https://doi.org/10.1007/s11222-007–9033–z

[19] S. A. Catanese, P. De Meo, E. Ferrara, G. Fiumara, A. Provetti, "Crawling facebook for social network analysis purposes," In *Proceedings of the International Conference on Web Intelligence, Mining and Semantics*, Sogndal, Norway, May 25–27, pp. 52, ACM, 2011.

[20] J. Ugander, B. Karrer, L. Backstrom, C. Marlow, "The anatomy of the Facebook social graph", arXiv preprint arXiv:1111.4503, 2011.

[21] J. W. van Dam, M. Van De Velden, "Online profiling and clustering of Facebook users," Decision Support Systems, vol. 70, pp. 60–72, 2015. Available: https://doi.org/10.1016/j.dss.2014.12.001

[22] M. Tulin, T. V. Pollet, N. Lehmann-Willenbrock, "Perceived group cohesion versus actual social structure: A study using social network analysis of egocentric Facebook networks," Social Science Research, vol. 74, pp. 161–175, 2018.

[23] S. Catanese, P. De Meo, E. Ferrara, G. Fiumara, A. Provetti, "Extraction and analysis of Facebook friendship relations," In Computational Social Networks, A. Abraham (Eds.) Springer London, 2012, pp. 291–324. Available: https://doi.org/10.1007/978–1–4471–4054–2_12

[24] M. G. Everett, S. P. Borgatti, "The centrality of groups and classes," The Journal of mathematical sociology, vol. 23, pp. 181–201, 1999. Available: https://doi.org/10.1080/0022250X.1999.9990219

[25] J. Scott, *Social network analysis*. Sage, London, 2012.

[26] C. Thirumalai, K. S. Sree, H. Gannu, "Analysis of cost estimation function for Facebook web click data," In *Proceedings of the International conference of Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, pp. 172–175, 2017.

[27] P. Dey, K. Dutta, "Social Network Analysis of Different Parameters Derived From Real-Time Facebook Profiles," In *Graph Theoretic Approaches for Analyzing Large-Scale Social Networks*, pp. 95–108, IGI Global, 2018.

[28] B. Pang, L. Lee, H. Rd, S. Jose, "Thumbs up? Sentiment Classification using Machine Learning Techniques," In *Proceedings of the ACL-02 conference on Empirical methods in Natural Language Processing (EMNLP '02)*, Philadelphia, PA, USA, July 6–7, vol. 10, pp. 79–86, 2002.

[29] J. Akaichi, Z. Dhouioui, M. J. L. H. Pérez, "Text mining Facebook status updates for sentiment classification". In *Proceedings of the 17th International Conference on System Theory, Control and Computing (ICSTCC)*, Sinaia, Romania, October 11–13, pp. 640–645, IEEE, 2013.

[30] J. K. Ahkter, S. Soria, Sentiment analysis: Facebook status messages. Master's thesis, Stanford, CA, 2010.

[31] C. Troussas, M. Virvou, K. J. Espinosa, K. Llaguno, J. Caro, "Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning". In *Proceedings of the Fourth International Conference on Information, Intelligence, Systems and Applications (IISA)*, pp. 1–6. IEEE, 2013.

[32] E. Ugljanin, D. Stojanović, E. Kajan, Z. Maamar, "Re-engineering of smart city's business processes based on social networks and Internet of Things," Facta Universitatis, Series: Automatic Control and Robotics, vol. 16, no. 3, pp. 275–286, 2018. Available: https://doi.org/10.22190/FUACR1703275U

[33] A. Giachanou, F. Crestani, "Like it or not: A survey of Twitter sentiment analysis methods," ACM Computing Surveys (CSUR), vol. 49(2), no. 28, pp. 1–41, 2016. Available: https://doi.org/10.1145/2938640

[34] A. Ljajić, U. Marovac, "Improving Sentiment Analysis for Twitter Data by Handling Negation Rules in the Serbian Language," Computer Science and Information Systems. Available: https://doi.org/10.2298/CSIS180122013L, to be published.

[35] E. J. Villota, S. G. Yoo, "An Experiment of Influences of Facebook Posts in Other Users," In *Proceedings of the International Conference on eDemocracy & eGovernment (ICEDEG)*, Ambato, pp. 83-88, 2018.

[36] M. Sharma, K. Yadav, N. Yadav, K. C. Ferdinand, "Zika virus pandemic–analysis of Facebook as a social media health information platform," American journal of infection control, vol. 45(3), pp. 301-302, 2017.

[37] S. Branković, "Napredna istraživanja komunikacije na društvenim mrežama: jedan analitički model," Časopis za komunikaciju i medije (communication and media journal), vol. 32, pp. 63–82, 2014.

[38] L. Bojic, J.-L. Marie, S. Brankovic, "Reception and Expression Capabilities of Media Addicts in Serbia," Kultura polisa, vol. 10, no. 22, pp. 353–368, 2013.