Methods of public multimedia analysis for a social profile

Alexander M. Bershadsky

doctor of technical sciences, professor, Penza State University, +7 (8412) 36-82-47,
bam@pnzgu.ru

Alexander S. Bozhday

doctor of technical sciences, professor, Penza State University, +7 (8412) 36-82-47,
bozhday@yandex.ru

Alexey Y. Timonin

postgraduate, Penza State University, +79273799248,
c013s017b301f018@mail.ru

Abstract. Analysis of public Internet data is a popular topic for research. Study prospects of using personalized data led the scientists to the problem of constructing a social profile. It is structured set of information, which is able to uniquely characterize a particular person. The social profile building is carried out through the analysis of the filtered Internet open source data. Raw social profile data are subdivided on static and dynamic parts. Dynamic unstructured data includes text and multimedia information and cannot be handled by classical analytic means. The analytical task of social profile data is achieved through mathematical tools of the set theory, Big Data software, NoSQL data stores and analytic tools for social media. Also, mod-ern methods for analysis of the multimedia data are helpful. The techniques review for the analysis of multimedia content (graphics, sound) is offered. Analysis of multimedia resources is complicated by the variety of information processed types. The current work is devoted to combine existing experience by automating non-textual information processing in the task of a social profile building. We consider such areas as Big Data analysis, visual analysis, Optical Character Recognition, speech recognition, neural networks and specialized algorithms for specific objects recognition and linking them with a social profile. Automated processing of multimedia information will improve the completeness and accuracy of the final social profile. In addition, the results of this work can be used to study the social phenomenon of viral media.

Keywords: personal social profile, public data sources, social media, multimedia, unstructured data, data analysis, Big Data, Data Mining.

1 Introduction

Processing of heterogeneous Internet open data covers a very wide range of applicability in various fields of human activity. Application analysis results of such information starts with contextual sampling recommendations and end large-scale scientific society research for the purpose of counteraction to criminality. A social profile building is one of these generally applicable tasks. Social profile [16] - a lot of information that characterizes the social qualities of person and clearly structured for the convenience of automated processing and human perception. This problem is initially reduced to create a mathematical model and optionally data structures for storing personalized information [15]. Network personal identification is made through the determination of its entry points - web resource accounts, which make person stand out from the mass of other users. The collected data are filtered from the outside information and are divided by the degree of structuring to the static (information card) and a dynamic part.

Next step is the information obtained analysis after the person identification on the Internet, collecting and filtering of primary social profile data [14, 15]. This stage purpose is to form a complete picture of the structured social personality. There is no problem with processing of an information card data. However, dynamic data is divided by a processing method (renewable, non-editable, complex graph data) and by its nature (text, graphics, audio, binary data, etc.). Non-relational distributed database HBase [1] used for data storage, not providing editing. These database runs on HDFS (Hadoop Distributed File System) and provides a reliable way to store very large volumes of heterogeneous data. Simple method to store the renewable data is MongoDB [4] usage. It is very popular open source document-based solution. Also, MongoDB does not require the description of the table schema.

Issues of social media analysis profile of a man using a variety of means (Big Data, OCR, image analysis, neural networks, etc.) raise in this paper. In particular, the processing of graphical and audio content, that may be associated with a particular personality, is described. Also, there is attempts to automate the unstructured data analysis. It's necessary to extract the statistical and semantic components from them and to bind them with the social profile data.

2 Background

Studies on analysis of multimedia information are not as common as works related to processing text data. The reasons for this are simple:

- The diversity of multimedia information types;
- Semantic variability of content, depending on the recipient subject;
- A large amount of data, compared with the text;
- The existence of distortion;
- Difficulties of machining processing and structuring the results.

Methodology of speech and visual objects recognition fairly well researched at the current time. But differentiation issues of recognized entities not yet resolved. Also, determining the meaning and symbolism of the certain media image in the particular situation poorly studied. This theme is dealt with in article named "Unveiling the multimedia unconscious: implicit cognitive processes and multimedia content analysis"[7] of Marco Cristani, Alessandro Vinciarelli, Cristina Segalin and Alessandro Perina. A work called "Multimedia mining research – an overview" [17] provides the basic concepts of multimedia mining and its essential characteristics. Multimedia mining architectures for structured and unstructured data, research issues in multimedia mining, data mining models used for multimedia mining and applications are also discussed in their paper.

Authors of "Triangulating Social Multimedia Content for Event Localization using Flickr and Twitter" [11] carry out the connection between real events and media content on example of messages about natural disasters. Yilin Yan, Qiusha Zhu, Mei-Ling Shyu, Shu-Ching Chen in their article named "A Classifier Ensemble Framework for Multimedia Big Data Classification" [19] develop Spark ensemble system for multimedia big data processing. A paper called "Distributed Multimedia Content Analysis with MapReduce" [15] of Arto Heikkinen, Jouni Sarvanko, Mika Rautiainen and Mika Ylianttila introduces a scalable solution for distributing content-based video analysis tasks using the emerging MapReduce programming model. They present a novel implementation utilizing the popular Apache Hadoop MapReduce framework for both analysis job scheduling and video data distribution.

Jinglan Zhang, Kai Huang, Mark Cottman-Fields and others present an overview of techniques for collecting, storing and analyzing large volumes of acoustic data efficiently, accurately, and cost-effectively in their work called "Managing and Analyzing Big Audio Data for Environmental Monitoring" [20].

A paper named "Fusing audio, visual and textual clues for sentiment analysis from multimodal content" [12] introduce a novel methodology for multimodal sentiment analysis, which consists in harvesting sentiments from Web videos by demonstrating a model that uses audio, visual and textual modalities as sources of information. Authors used both feature- and decision-level fusion methods to merge affective information extracted from multiple modalities.

Research work called "Mining Melodic Patterns in Large Audio Collections of Indian Art Music" [3] by Sankalp Gulati, Joan Serrà, Vignesh Ishwar and Xavier Serra is devoted to the selection of music from a wide variety of audio templates, that further be used in challenging computational tasks such as automatic raga recognition, composition identification and music recommendation. Article named "YouTube as a source of chronic obstructive pulmonary disease patient education" [13] authors explored the potential of thematic YouTube videos to increase health literacy among patients at ex-ample diseased COPD.

Experience of works discussed above can be useful for handling multimedia social profile data.

3 Preparation to the analysis of social profile media

It is need to analyze the existing text data before proceeding the multimedia processing. Filling information card produced during collection and filtering stage of source data. Dynamic text data processed by subsystem of social entities and relationships analysis. It uses specialized tools of NLP (natural language processing), a tonality definition of the text and the predictive text mining methods. IBM ContentAnalytics software [5] was used for this purpose. However, unstructured text data may contain hidden information, which is determined only indirectly. Visual analysis means are most suitable to address these issues. A good choose is use of the IBM i2 software tools [2].

Next step is to create a mathematical model for the construction of iBase social profile database. Model based on the results of the textual data analysis using BigInsights and Content Analytics. Social graph is built on this basis. It specifies the possible relationships between the considered person and social profile entities (mentioned persons - information about persons, associated to the considered person in any context; organizations - information about various facilities related to personalities from the social profile; events - information about events, bringing together a group of people by some common features; contact details of the person; activities, specialization, attainments and hobbies of person). A such graph schematic example is shown in Figure 1.

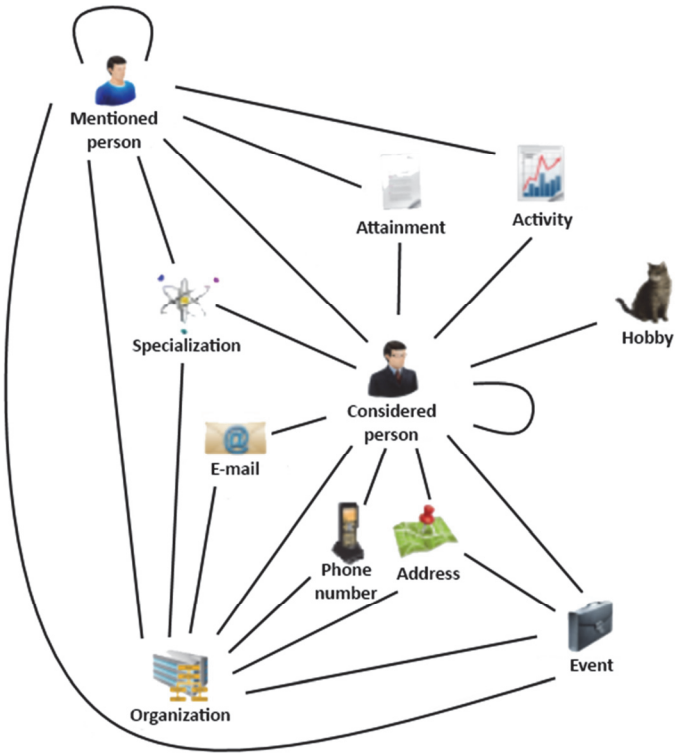

Fig. 1. Sample graph types of social profile entities.

The main work on the identification of relationships and dependencies is performed in IBM i2 TextChart program. Data analysis is performed using the following algorithm:

1. Raw social profile data shall be input in the TextChart project via import CSV-file with source information.
2. The text is allocated first read important information, then search is carried for repetition and synonyms throughout the text by using the Find tool.
3. The results are added to the project as social profile entities (considered person, activities, etc.).
4. Words are highlighted in the processed text, expressing the relationships between the generated objects. Then corresponding objects are selected in the dependency graph window.
5. The desired connection type is selected by clicking the Insert Link button, then the expressions are added as the connections.
6. Similarly, new objects and attributes are added to existing ones by means of navigation between entity search results.
7. Entries calculated for each option when finding conflicting information. Then it concludes about their truth: false information is removed from the social profile or additional specifying search is performed.

The results of visual analysis are the social graph and iBase social profile database. It is possible to find implicit relationships built on the graph in addition to the hidden information directly identified during the analysis. Analyze the social profile text in such way will form the basis for relations with the social objects. These objects obtained in the analysis of the multimedia information.

4 Methods overview for analysis of the social profile media

In contrast to textual data, multimedia content is difficult to analyze using traditional means. Therefore, it is necessary to resort the implementation of big data solutions, machine learning and neural networks. Examples of such systems can be considered Google Analytics, MS Azure, Multimedia Mining Marvel, Quaero. The most common multimedia information are the images and sound. There are a number of multimedia data analysis strategies [8]. Holistic strategy is prefered to handle unstructured information of social profile.

Multimedia content can appear in two different versions as part of social profile building:

1. Multimedia data, viewed and created by a considered person. That information can speak about the activities and preferences of the person. Also it applies to the author's content. Analysis issue consists in comparing the multimedia objects with the existing samples in the Internet and a social profile (e.g. definition of personal musical preferences across multiple recordings).
2. Content that contains the information about the considered person itself. The analysis purpose is to identify the essential information from the direct media object

(e.g. emotion recognition in the photo, the semantic expressions selection of audio).

Content-based analysis method appropriate to use for multimedia information processing of the first category [18]. Its essence is to split the data into constituent parts and their subsequent direct comparison. The following algorithm would run for the mentioned example of the musical preferences definition.

1. The existing audio records ID3 tags of social profile are checked for the presence of Artists and Genres of Music positions.
   (a) If such tags are found, they are recorded in the preference table.
   (b) Otherwise, the comparative analysis is carried out. Audio record compare with the samples from the Internet (by means of AudioTag, Shazam or Google Sound Search utilities. Required tags are added to the preference table after finding a match.
2. Tag counting produced after completion processing of all available audio records.
3. The conclusion about the predominance of a particular genre or artist in the sample is output.

User content may contain different identification labels in the service data (information about the author, the developer device, spatial data) and within the multimedia entity. Examples of these labels can be: the author's signature or watermark, typical author's style of the object, author's mention in comments, etc. Some of this information is implicit, so the possibility of automatic processing for user-generated content is very limited. The use of visual analysis is desirably. Processing algorithm will be different from one presented in paragraph 3 of this article that, you need to consider abstractions in addition to text information extracted from the multimedia objects. These abstractions are mined manually and have a fuzzy interpretation.

Another object of the user-generated content analysis is to define options for using multimedia entities. In this case, the analyst should link a particular media object to the text definition of the event meaning in which the object is mentioned. Big Data means searches for all possible situations in Internet of the multimedia entity use. Further, statistical analysis and tone selection are performed for the accompanying text. Also, it is recommended to apply visual analysis means as in the previous case.

Let's take a second data category. Content-interpretative method is used for its analysis [18]. It consists that parts of multimedia data are assigned to the concepts in the formal language, and then the links line up between them. Approaches to audio and graphics information analysis are different.

5 Audio information processing

A speech and intonation recognition are the main focus of audio analysis in the construction task of social profile. The results of the audio analysis are: the sound characteristics of the considered person voice, attached to the social profile; recognized text linked to the original record. The characteristics of the other people voices also may be taken from the analyzed audio file for comparison with other social profiles. It should be noted that the sound voice characteristics can`t be regarded as an infor-

mation card elements. They can vary considerably depending on the age and condition of the per-son, environment and recording quality.

Currently, there are a sufficient number of freely distributed speech recognition systems with open source: CMU Sphinx, Julius, HTK, Praat, SHoUt, VoxForge and others. Many of them are based on the use of hidden Markov models and neural networks. Sound voice characteristics include spectral-temporal, cepstral, amplitude-frequency and signs of nonlinear dynamics.

It should be mentioned that there are seven types of intonation construction in Russian language (3 interrogative, 2 exclamatory, 2 narrative). Intonation recognition of human speech can take place in three steps [6, 10]:

- Record of human speech and its division into finished intonation constructions;
- Selection of voice tone in the each of record parts;
- Development of the intonation constructions classifier.

Accuracy of this algorithm depends on the quality and duration records, characteristics of speech, etc.

6 Graphic data analytical approaches

Analysis of the graphical information includes: image, text recognition and compare the results with the creation date of the file under consideration. The approaches of the graphic information recognition are divided into three categories: iteration methods, artificial neural networks, and object search paths to the further study of their properties. OCR technology (Optical Character Recognition) allows you to find the printing and handwriting (rarely) text. The original image quality has a strong influence on the final result of the recognition algorithms. As in the case of the audio analysis, the image-recognized text should be attached to the original object. Further it will be treated by means of text analytics. Findings data are reduced in the resulting table after completion of the image processing. That table contains the following key parameters: detected faces, their emotions, the list of labels, environmental data (the recognized objects), service image information (size, creation date, name, etc).

Pattern recognition in the task of a social profile building is divided into: search people in the images, the determination of their sentiments and the environmental elements allocation. Face Detection Services provide by projects such as ASID, FaceID, FindFace, Vissage Gallery and others. Emotion recognition is a more complicated procedure, major problems which are to determine the face position and color, quality of illumination, foreign objects in the image foreground. However, despite this, there are ready-made solutions: FaceReader, FaceSecurity, etc. Development of systems for determining environmental elements is not sufficiently well-researched now. There-fore, it is advisable to use specially trained neural network to solve this problem.

7 Conclusion

The degree of analytical subsystem maturity for social profile building affects the information content and correctness of the final results. This is especially true for multimedia processing, because these data are diverse and difficult to automate processing. A review of existing approaches to the multimedia content analysis is provided in this paper. Also, possibility of their application in the task of a social profile building is considered. It was revealed that is possible to use existing algorithms for processing audio data, while the analysis of graphic information requires improving recognition technology.

References

1. Apache HBase ™ Reference Guide (2016), http://hbase.apache.org/book.html
2. Hidden communications identification on the basis of the textual analysis with i2. Center of competence for IBM Big Data technology, Moscow (2014)
3. Gulati, S., Serrà, J., Ishwar, V., Serra, X.: Mining Melodic Patterns in Large Audio Collections of Indian Art Music. In: Signal-Image Technology and Internet-Based Systems (SITIS), 2014 Tenth International Conference on, pp. 264-271. Publisher: IEEE (2015). DOI= https://doi.org/10.1109/SITIS.2014.73
4. MongoDB for GIANT Ideas | MongoDB (2017), https://www.mongodb.com
5. The analysis of the structured and unstructured data with the Content Analytics. Center of competence for IBM Big Data technology, Moscow (2014)
6. Boykov, I. V., Ivanov, A. I., Kalashnikov, D. M.: An algorithm for constructing a statistical description of the discrete-continuum duration meaningful speech speaker sound stream. In: Proceedings of higher educational institutions. Volga region. Technical science, №4, pp. 64-78. Penza: PSU Publisher, Penza (2015)
7. Cristani, M., Vinciarelli, A., Segalin, C., Perina, A.: Unveiling the multi-media unconscious: implicit cognitive processes and multimedia content analysis. In: Proceedings of the 21st ACM international conference on Multimedia, pp. 213-222. ACM New York, NY, USA (2013) DOI= https://doi.org/10.1145/2502081.2502280
8. Davydov, A.A.: Systemic Sociology: an analysis of multimedia information on the Internet. In: Official site of SI RAS (2009), http://www.isras.ru/publ.html?id=1257
9. Heikkinen, A., Sarvanko, J., Rautiainen, M., Ylianttila, M.: Distributed Multimedia Content Analysis with MapReduce. In: 24th International Symposium on Personal, Indoor and Mobile Radio Communications: Services, Applications and Business Track, pp. 3502-3506. Publisher: IEEE (2013). https://www.researchgate.net/profile/Mika_Ylianttila/ publication/257641284_Distributed_Multimedia_Content_Analysis_with_MapReduce/links/57399f8008ae9ace840d90d7.pdf
10. Levin, A.I., Minin, P.E., Egorov, A.D.: Recognition of intonation in human continuous speech. In: XIX International telecommunication conference of young scientists and students "Youth and science" Theses of reports. edited by O.N. Golotyuk, pp. 109–110. Moscow: National research nuclear university "MIFI" (2015)
11. Panteras, G., Wise, S., Lu, X., Croitoru, A., Crooks, A., Stefanidis, A.: Triangulating Social Multimedia Content for Event Localization using Flickr and Twitter. In: Transactions

in GIS, Volume 19, Issue 5, pp. 694–715 (2014). DOI= http://onlinelibrary.wiley.com/doi/10.1111/tgis.12122
12. Poria, S., Cambria, E., Howard, N., Huang, G.-B., Hussain, A.: Fusing audio, visual and textual clues for sentiment analysis from multimodal con-tent. In: Neurocomputing, Volume 174, Part A, pp. 50–59 (2016). DOI= http://dx.doi.org/10.1016/j.neucom.2015.01.095
13. Stellefson, M., Chaney, B., Ochipa, K., Chaney, D., Haider, Z., Hanik, B., Chavarria, E., Bernhardt, J.M.: YouTube as a source of chronic obstructive pulmonary disease patient education. In: Chronic Respiratory Disease Journal, Volume 11, issue 2, pp. 61-71. (2014) DOI= https://doi.org/10.1177/1479972314525058
14. Timonin, A.Y., Bozhday, A.S., Bershadsky A. M.: Research of filtration methods for reference social profile data. In: EGOSE '16 Proceedings of the International Conference on Electronic Governance and Open Society: Challenges in Eurasia, pp. 189-193. ACM New York, NY, USA (2016) DOI=https://doi.org/10.1145/3014087.3014090
15. Timonin, A.Y., Bozhday, A.S., Bershadsky A. M.: The Process of Per-sonal Identification and Data Gathering Based on Big Data Technologies for Social Profiles. In: Digital Transformation and Global Society. DTGS 2016. Communications in Computer and Information Science, vol. 674, pp. 576-584. Springer, Cham (2016) DOI=https://link.springer.com/chapter/10.1007/978-3-319-49700-6_57
16. Timonin, A.Y., Bozhday, A.S.: The use of Big Data technologies to build a human social profile on the basis of public data sources. In: Bulletin of Penza State University, № 2 (10), pp. 140–144 (2015) http://elibrary.ru/item.asp?id=24097671
17. Vijayarani, S., Sakila, A.: Multimedia mining research – an overview. In: International Journal of Computer Graphics & Animation (IJCGA), Vol.5, No.1, pp. 69-77. (2015) DOI= https://doi.org/10.5121/ijcga.2015.5105
18. Yakovlev, V.E.: Macromedia: multimedia information analysis. M-Lang. In: Journal "Young Scientist", №4. vol.1, pp. 105-108 (2011)
19. Yan, Y., Zhu, Q., Shyu, M.-L., Chen, S.-C.: A Classifier Ensemble Framework for Multimedia Big Data Classification. In: Information Reuse and Integration (IRI), 2016 IEEE 17th International Conference on, pp. 615-622. Publisher: IEEE (2016). DOI= https://doi.org/10.1109/IRI.2016.88
20. Zhang, J., Huang, K., Cottman-Fields, M., Truskinger, A., Roe, P., Duan, S., Dong, X., Towsey, M., Wimmer, J.: Managing and Analysing Big Audio Data for Environmental Monitoring. In: Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference on, pp. 997-1004. Publisher: IEEE (2014). DOI= https://doi.org/10.1109/CSE.2013.146