

# Using Sentiment Text Analysis of User Reviews in Social Media for E-Tourism Mobile Recommender Systems

Olga Artemenko<sup>1</sup>, Volodymyr Pasichnyk<sup>2</sup>, Nataliia Kunanets<sup>2</sup>, Khrystyna Shunevych<sup>2</sup>

<sup>1</sup>PHEI “Bukovinian University”, Chernivtsi, Ukraine

<sup>2</sup>Lviv Polytechnic National University, Lviv, Ukraine

olga.hapon@gmail.com, vpasichnyk@gmail.com,  
nek.lviv@gmail.com, krishirak@gmail.com

**Abstract** This paper describes main modern tendencies for the design and development of e-tourism recommender systems with sentiment analysis of user generated content in social media. Main goal is to systematize and summarize knowledge about the possibilities of using tourist’s user reviews in social media as a type of e-tourism big data for mobile e-tourism recommender systems. In particular, to analyze the sources and types of tourist feedback data, messages and comments generated by the tourist with his gadget that can be related to e-tourism big data. Developing efficient tools for e-tourism user comments and feedback in social media, combining big data technologies, NLP and smartphone services advantages, can provide e-tourism recommender systems with new better ways of creating more personalized recommendations.

Keywords: e-tourism, mobile recommender systems, trip support, content analysis, sentiment text analysis

## 1 Introduction

There is an increasing interest for the use of content created by consumers of hospitality and tourism services, in particular on social networks and video hosting. Thus, the structure and dynamics of tourists’ preferences can be tracked and analyzed, information about the image and reputation of the tourist product can be received, as well as about the behavior of tourists themselves when traveling. The feedback received from the tourist is not only useful for business, but also can be used by recommender applications as one of the sources for estimation of the alternative item. Two popular e-advice websites Booking and TripAdvisor host users’ opinions since decades. But they are very much moderated. Also not every user leaves feedback on tourism-related review platforms. But every user has a profile in one or more social networks. And there he publishes different aspects of his life, tourist experience included.

Copyright © 2020 for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

## 2 E-Tourism Recommender Systems

Recommender systems are a class of intelligent information retrieval systems designed to filter out, in a abundance of information resources, exactly the instances of data that best meet the interests of a particular user [1]. Diversified e-tourism recommender systems are intensively developing and are very popular among the users. But the problem of getting better, faster, more personalized recommendations is still on the table. One of the resources for improvements is using tourist's user reviews and comments in social media as another kind of recommendation tool.

E-tourism recommender systems can be classified according to different characteristics, such as: architecture, information technology platform, target audience, methods used for generating recommendations, main tasks to be solved, etc [2].

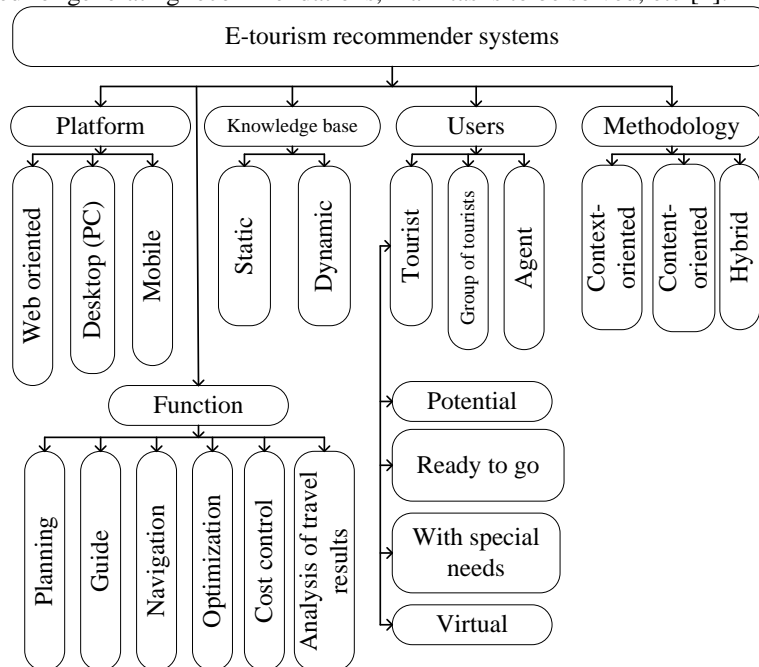


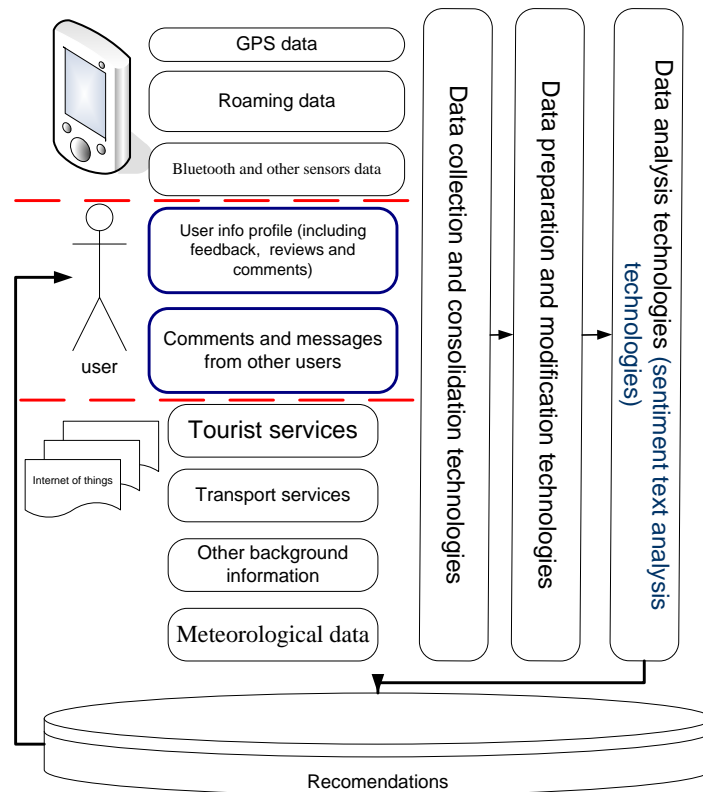
Fig. 1. General classification of e-tourism recommender systems

The basis for successful operation for mobile e-tourism recommender systems is a complete, well-timed and correct data processing. There are certain peculiarities in dealing with input and output data and they should be taken into account when designing tourist recommender systems, in particular mobile ones. There are three main sources of input data for a mobile e-tourism recommender system:

1) The user as an informational source – generates queries, leaves feedback, disseminates messages about himself on social networks. All smart tourism technologies nowadays act in the paradigms "tourist as a sensor" and "every tourist is an expert".

2) The gadget of the user itself - information about the external background of the tourist, contextual data, appearance or disappearance from the operational space of various obstacles, etc.

3) Internet content and internet of things - this is data from referential resources, both tourist and external, work schedules, lists of tourist places and establishments, public transport timetables, etc., including web search data, user net surfing history and online booking data, and more.



**Fig. 2.** Processing user generated big data in mobile e-tourism recommender systems

### 3 Analysis of User Generated Content

The gadget user is a powerful source of information for e-tourism software products. This data belongs to the e-tourism big data category [3]. The development of technology and the phenomenon of social networks have led to the emergence of new concepts, ways, rules and habits of disseminating information in the digital world. Hashtag, emoji, geo-positioning, photo and video content, live streams and pages complemented the classic textual content, which was the main source of tourist feedback [4].

In particular, the text message that a gadget user leaves as a review of a tourism product consumed has changed [5-8].

1. The user response has become shorter. First, there are limits to the number of characters for a single message in different systems, such as Twitter; screen size of the smartphone – there is an unspoken rule "what wasn't fit on one screen will not be read" [9].

2. The space for posting reviews has also changed. Traditionally, users have left posts on specialized sites, travel forums, travel agency blogs, and more [10]. To do this, the user has either logged in or left an anonymous comment. But since the last decade, a tourist with a gadget leaves a comment anywhere in the social media space [11].

3. The structure of the comment has also changed: text is now being supplemented or even partially replaced by graphic, audio and video content. Emoji, stickers, animated elements convey the emotional tinge of user feedback. Video stories and live streams may contain text captions and subtitles to increase the content of the response [12-14].

4. The user may not plan and prepare the response text in advance, his story may be devoted to a completely different topic, and his own impression of the consumed tourist product will "slip through" among other things. Such reviews are the hardest to follow, but they also create a reputation for the tourist product [15].

5. Using a hashtag for text and geo-positioning for images and videos allows you to uniquely identify the tourism product [16]. Making it easier to find and analyze data.

6. Option of personalized feedback from the author of the review. From the official owner's profile of the tourist product can be added in response to the tourist post a gratitude for the positive feedback or an apology in case of complaints [17-18]. In this way, thanks to the social media space, the product seller is able to reach the customer's information territory and attract his (and his social environment) attention. It is also possible to supplement user profile of the recommendation application with new review facts.

7. The language used by the tourist: in the forums and official pages of the tourist objective (classic space for creating reviews), as a rule, one language is used, or in the case of regional information resources, two: English and the language of the region [19]. The social media space enables the user to express his or her thoughts in the language best suited to them [20]. That is, it is likely that the tourism product provided by one country will have reviews in five, ten or more different languages. Which complicates the analysis of the text.

Therefore, travel product reviews need to be collected not only on specialized resources, but also increasingly in the social media space. Analyzing the sentimental content of such reviews is complicated, first, by multilingualism and, second, by the presence of such graphic elements as emoji and stickers. Consumer feedback now needs to be maintained on users territory – on social media.

Accordingly, the analysis of the sentimental filling of tourist feedback on tourist products is not only a source of data for mobile e-tourism recommender systems, but

also transforms from the classic text-mining task to the task of analyzing big data, not only text [21].

Finding and retrieving useful information from user reviews of a tourism product in the social media space poses a number of challenges to developers of recommender applications. In particular:

1. How to properly treat sentimental tint of an emoji in reviews? Is negative content related to the mood of the user, the weather, the day in general or the quality of the tourist product consumed this day? Is it possible to use for comparison as a description of previous bad experiences with another product? Should emoji be considered equivalent to keywords in reviews?
2. How situational and implicit (no hashtag and location) reviews can be well tracked and consolidated?
3. How to effectively extract text content from photos, videos and audio messages?
4. Should the publication a tourist photo or video related to the tourist product but without supplementing the text message be considered as a feedback and how to classify it: as positive, negative or neutral?

These and other problems need to be solved to create efficient sentiment analysis technologies for mobile e-tourism recommender systems.

#### 4 Using Sentiment Text Analysis for E-Tourism Technologies

Natural Language Processing (NLP) is a field of Computer Science that studies the use of automatic ways to process natural language. Sentiment text analysis is a fast growing element of NLP [22]. Automatic processing of e-tourism text data due to the large amount of content generated by users every minute is becoming more complicated.

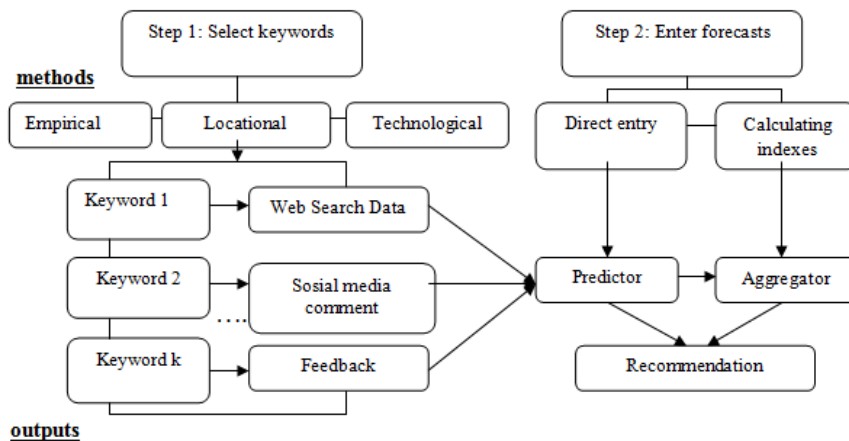


Fig. 3. Using web search data and social media feedback texts to predict tourists' preferences

The keys to solve it lie in combining big data technologies, NLP and smartphone services advantages [23].

Different domains and types of texts have different information extraction requirements and thus require different NLP tasks and tools [24]. Developing efficient tools for sentiment analysis of specific type of text – e-tourism user comments and feedback in social media can provide e-tourism recommender systems with new better ways of creating more personalized recommendations.

There are researches and discussions on the mechanisms behind reviewing tourists behavior and it's connection with the data of the reputation sites, hotels, attractions and destinations have online, and how this affects tourist behavior and purchasing decisions. Social media feedback data bring new context and new challenges to this topic. But also they bring new perspectives and resources.

## 5 Sentiment Text Analysis of E-Tourism User Reviews

In the first stage, a list of key components of the response was compiled: keywords, hashtags, emoji, and the order of punctuation was drawn. The keywords were divided into classes, as well as Ukrainian, English and Romanian, since these three languages are used by tourists to provide feedback on Bukovina tourist services, as shown in Tables 1, 2 and 3.

**Table 1.** Fragment of Keywords list (Ukrainian)

<i>Keywords</i>									
	Дуже позитивні	Позитивні	Нейтральні	Негативні	Дуже негативні				
x1	дуже сильно сподобалося	y1 прикольн о	z1 аби як	a1 не сподобалось	b1 нічого не сподобалося				
x2	надзвичайно сподобалося	y2 круто	z2 50/50	a2 звичайні відчуття	b2 зовсім не сподобалося				
x3	дуже романтичні відчуття	y3 сподобалось	z3 непогані відчуття	a3 не романтичні відчуття	b3 погані відчуття				
x4	надзвичайно гарно	y4 нормальні відчуття	z4 може бути	a4 не гарно	b4 зовсім не гарно				
x5	дуже гарно	y5 все сподобалось	z5 не зовсім сплановано	a5 не сплановано	b5 нічого не гарно				
x6	дуже-дуже гарно	y6 романтичні відчуття	z6 досить цікаво	a6 ніяково	b6 зовсім не зручно				

**Table 2.** Fragment of Keywords list (Romanian)

<i>Comentariile utilizatorilor</i>				
<b>Foarte pozitive</b>	<b>Pozitive</b>	<b>Neutru</b>	<b>Negative</b>	<b>Foarte negative</b>
mi-a plăcut foarte mult sentimente foarte romantice foarte frumos foarte, foarte frumos bine planificat	mi-a plăcut  senzații normale	50/50  nu-i rău	nu mi-a plăcut senzații obișnuite	nu mi-a plăcut nimic nu mi-a plăcut deloc
extrem de convenabil foarte in- terasant incredibil de interesant	senzații bune  destul de frumos  destul de bine	poate nu chiar planifi- cat fără precedent	nu e bine nu este planificat incomod	Senzație de rău nu e frumos  nimic nu este bun
		nu chiar con- fortabil destul de gândit	nu mă interesează nici o impresie așteptările nu s-au îndeplinit	nu este con- venabil nimic nu este gândit nemulțumiți

**Table 3.** Fragment of Keywords list (English)

<i>User feedback</i>				
<b>Very positive</b>	<b>Positive</b>	<b>Neutral</b>	<b>Negative</b>	<b>Very negative</b>
I liked it very much I really liked it very ro- mantic feelings Very beau- tiful very, very beautiful well- planned extremely convenient	cool  hard-boiled  liked  sensations are normal romantic feeling nicely quite beauti- ful	to how  50/50  good feeling  maybe  reasonably interesting not very con- venient elaborate	not like  normal feelings  not romantic feelings  not like  confusedly  uncomfortable  discomfort	I didn't like any- thing I didn't like it at all bad feelings  not pretty at all  not at all conven- ient not at all interest- ing no way

Since Protégé cannot write hashtags via "#", we wrote them using the letter "h". The hashtags were divided into "Very Positive", "Positive" and "Neutral" as well as being Ukrainian, English and Romanian as shown in the tables. 4, 5 and. 6.

**Table 4.** Fragment of hashtags list (Ukrainian)

<i>Хештеги користувачів</i>			
<b>Дуже позитивні</b>	<b>Позитивні</b>	<b>Нейтральні</b>	<b>Ім'я змінної</b>
h_дуже	h_щастя	h_мандрівка	h.x1
h_дужесмачно	h_щастяє	h_мандруй	h.x2
h_дужевесело	h_цікавімісця	h_мандриукраїною	h.x3

h_дуже дешево	h_щастя в дрібницях	h_мандрівки	h.x4
h_дуже круто	h_щастя поруч	h_мандрівник	h.x5
h_дуже красиво	h_щастя_е	h_мандрівниця	h.x6
h_дуже гарно	h_щасливі	h_мандрівники	h.x7
h_дуже дуже	h_щасливий день	h_мандруй дешево	h.x8
h_дуже гарне місто	h_щастя в простих речах	h_мандруй сміливо	h.x9
h_веселуха	h_щасття	h_мандруємо україною	h.x10
h_супер	h_весело	h_мандрівники україною	h.x11
h_божественно	h_цікаво	h_мандрую україною	h.x13
h_чудовий день	h_цікава країна	h_мандри подорожі	h.x14
h_чудово	h_цікаві місця України	h_мандрувати легко	h.x15
h_чудовий настрій		h_мандруй з нами	h.x16
h_чудовий ранок		h_мандруй активно	h.x17
h_чудовий день		h_подорож	h.x18
h_чудовий вечір		h_подорожі	h.x19
h_класно		h_подорожі україною	h.x20
		h_подорожуй україною	h.x21
		h_подорожі_україною	h.x22
		h_подорожуємо	h.x23
		h_подорожуй знами	h.x24
		h_подорожувати	h.x25
		h_подорожуємо морозом	h.x26
		h_подорожуй_україною	h.x27
		h_подорожуючи україною	h.x28

Table 5. Fragment of hashtags list (Romanian)

<i>Hashtag-urile utilizatorului</i>		
<b>Foarte pozitive</b>	<b>Pozitive</b>	<b>Neutru</b>
h_foartegustos	h_fericire	h_calator
h_foartevesel	h_fericirea	h_calatoreste
h_foarteieftin	h_fericit	h_calatorestecudrag
h_foarte_frumos	h_fericirea exista	h_calatorii
h_foarte_tare	h_interesant	h_calatori
h_foarte_amuzant	h_interesantelocuri	h_calatorie
h_foarte foarte	h_fericirea in lucruri simple	h_calatorintaramea
h_vesel	h_fericirea in lucruri mici	h_calatorinromania
h_divin	h_fericirea in lucruri marunte	h_calatoriicugust
h_perfect		h_calator_in_romania
h_perfectazi		h_calator_in_tara_mea
h_perfectadimineata		h_calator_prin_lume
h_perfecta_zi		h_calator_prin_romania
h_perfectaseara		h_calatori_in_viata
h_perfecta_dimineat		h_calatori_prin_lume
a		
h_pecta_seara		h_calatorii_cu_zambet
		h_itur











**Table 6.** Fragment of hashtags list (English)

<i>User hashtags</i>		
<b>Very positive</b>	<b>Positive</b>	<b>Neutral</b>
h_very	h_happy	h_ journey
h_very_delicious	h_happytime	h_travel
h_very_fun	h_happiness	h_traveling
h_very_cheap	h_happiness_in_the_little_things	h_travelling
h_very_good	h_happiness_nearby	h_travels
h_very_good	h_happiness_exists	h_traveller
h_super	h_happy_moments	h_traveler
h_wonderful	h_happy_day	h_travel_drops
h_wonderful_location	h_happy_night	h_travelboddy
h_wonderful_vacations	h_happy_morning	h_travel_drops_
h_wonderful_day	h_fun	h_travel_capture
h_wonderful_night	h_interesting	h_travel_europe
h_wonderful_morning	h_interesting_places	h_tarvel_captures
h_wonderful_mood		h_travel_
h_wonderfulvacations		h_travel_tourist
h_wonderfulday		h_travel_life
h_wonderfulnight		h_lifesjourney
h_wonderfulmorning		h_thejourney
h_wonderfulmood		h_journeys
h_very_nice		h_travel_wonderful
h_very_beautiful		h_travel_world
h_very_delicious_food		h_travel_magic
h_cool		h_travel_love
		h_travel_time

	h_travel_is_life h_travellife h_travelgoals
--	---

Since we can't add emoji to Protégé, we wrote them through the letter "e". Emoji were divided into "Very Positive," "Positive," "Neutral," "Negative," and "Very Negative." Table 7.

**Table 7.** Fragment of Emoji list

<i>Emoji</i>	<i>Transcription</i>	<i>Variable name</i>
	е_дуже сильно сподобалось	e.x1
	е_надзвичайно гарно	e.x2
	е_дуже задоволені	e.x3
	е_дуже романтичні відчуття	e.x4
	е_найкращі емоції	e.x5
	е_дуже весело	e.x6
	е_на високому рівні	e.x7
	е_розкішно	e.x8

Punctuation marks are used to denote such a dismemberment of a written language that cannot be transmitted either by morphological means or by the order of the words in the sentence.

An exclamation point (!) Is a punctuation mark that is placed at the end of a sentence to express outrage, a call for strong feelings, anxiety, and more. It can also be doubled, tripled or used many times to express greater expression and emotionality in grammatical abuse.

Question mark (?) Is a punctuation mark, usually placed at the end of a sentence to express a question or doubt.

In user reviews, punctuation marks such as question mark (?) And exclamation mark (!) Are very common, they can be for positive feedback as well as negative feedback, it all depends on the words found before punctuation marks, positive keywords or negative .

Users use punctuation to express or displease tourist services. If after a positive keyword there are three exclamation marks then the keyword refers to very positive feedback, but if after a positive keyword there are three question marks then the keyword refers to very negative feedback. A single exclamation mark after neutral keywords means that the keyword refers to positive responses, but if one question mark after a negative keyword means the keyword refers to neutral responses. For example, a user posted the following comment: Like it! this keyword is not a positive but a very

positive one, because there are three exclamation points after it, or the user left a "Dear !!!" this keyword refers not to negative but very negative feedback, or the user left a response: "Why is it so expensive?", the keyword here is "expensive", since after the keyword one question mark, the response refers to neutral feedback.

According to the keyword tables, hashtags and emoji built a hierarchy of ontology classes and subclasses with Protege software. The classes in Protege are displayed as a class hierarchy (Class Hierarchy). Initially, they created base classes according to the hierarchy. Instances were created for each class as shown in Figure 4.

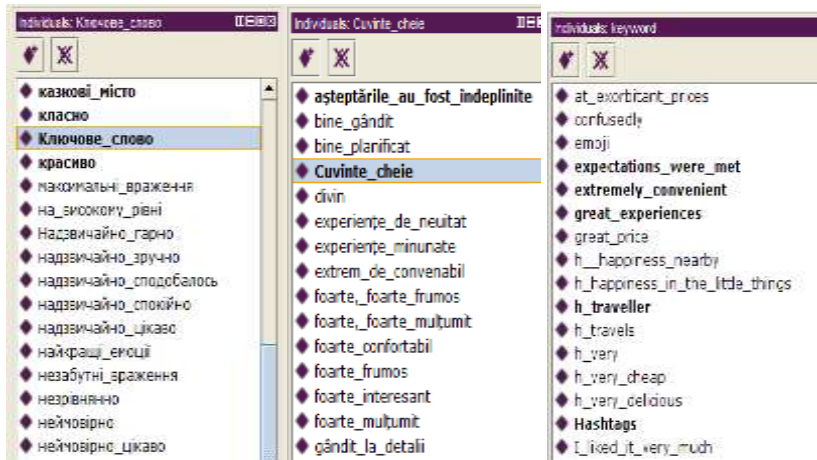


Fig. 4. Instances (Ukrainian, Romanian, English)

The ontology properties were created, corresponding to the areas of definition and areas of value of the hierarchical ontology. Figure 5.

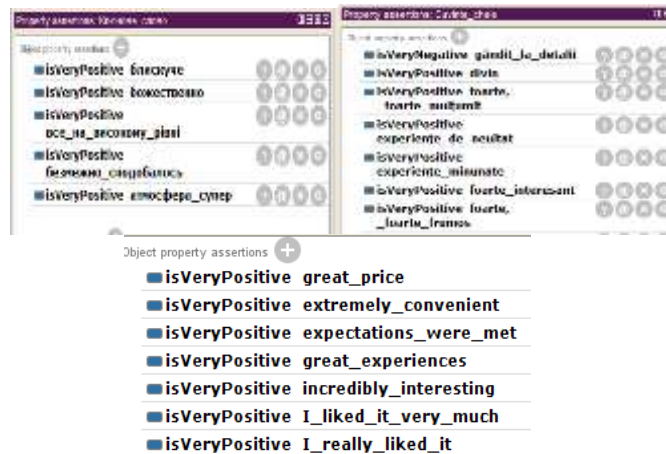


Fig. 5. Specific relations between classes (Ukrainian, Romanian, English)

## 6 Conclusions

This study is an attempt to systematize and summarize knowledge about the possibilities of using tourist's user reviews in social media as a type of e-tourism big data for mobile e-tourism recommender systems. In particular, to analyze the sources and types of tourist feedback data, messages and comments generated by the tourist with his gadget, that can be related to e-tourism big data.

## References

1. Ricci F., Rokach L., Shapira B.: *Recommender Systems Handbook: Second Edition*. Springer Science+Business Media New York, p. 1003. (2015)
2. Artemenko O., Kunanets O., Pasichnyk V.: E-tourism recommender systems: a survey and development perspectives. *Econtechmod an international quarterly journal*, Vol. 6. No. 2, 91-95. (2017)
3. Li, J., Xu, L., Tang, L., Wang, S., & Li, L.: Big data in tourism research: A literature review. *Tourism Management*, 68, 301–323. (2018). doi:10.1016/j.tourman.2018.03.009
4. Su, K.-W., Liu, C.-L., & Wang, Y.-W.: A principle of designing infographic for visualization representation of tourism social big data. *Journal of Ambient Intelligence and Humanized Computing*. doi:10.1007/s12652-018-1104-9 (2018).
5. Guan, D., & Du, J.: Cross-Media Big Data Tourism Perception Research Based on Multi-Agent. *Lecture Notes in Electrical Engineering*, 353–360. (2015). doi:10.1007/978-3-662-48386-2\_37
6. Delic, A., Neidhardt, J., Nguyen, T. N., & Ricci, F.: An observational user study for group recommender systems in the tourism domain. *Information Technology & Tourism*, 19(1-4), 87–116. (2018). doi:10.1007/s40558-018-0106-y
7. Huertas, A.: How live videos and stories in social media influence tourist opinions and behaviour. *Information Technology & Tourism*, 19(1-4), 1–28. (2018). doi:10.1007/s40558-018-0112-0
8. Sertkan, M., Neidhardt, J., & Werthner, H.: What is the “Personality” of a tourism destination? *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0135-6
9. Iinuma, S., Nanba, H., & Takezawa, T.: Investigating the effectiveness of computer-produced summaries obtained from multiple travel blog entries. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0132-9
10. Lalicic, L., Huertas, A., Moreno, A., & Jabreel, M.: Which emotional brand values do my followers want to hear about? An investigation of popular European tourist destinations. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0134-7
11. Höpken, W., Eberle, T., Fuchs, M., & Lexhagen, M.: Google Trends data for analysing tourists' online search behaviour and improving demand forecasting: the case of Åre, Sweden. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0129-4
12. Scaglione, M., Johnson, C., & Favre, P.: As time goes by: last minute momentum booking and the planned vacation process. *Information Technology & Tourism*, 21(1), 9–22. (2018). doi:10.1007/s40558-018-0133-8
13. Leung, D. Denis Trček: Trust and reputation management systems: an e-business perspective. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0118-7

14. Rossetti, M., Stella, F., & Zanker, M.: Analyzing user reviews in tourism with topic models. *Information Technology & Tourism*, 16(1), 5–21. (2015). doi:10.1007/s40558-015-0035-y
15. Krawczyk, M., & Xiang, Z.: Perceptual mapping of hotel brands using online reviews: a text analytics approach. *Information Technology & Tourism*, 16(1), 23–43. (2015). doi:10.1007/s40558-015-0033-0
16. García-Pablos, A., Cuadros, M., & Linaza, M. T.: Automatic analysis of textual hotel reviews. *Information Technology & Tourism*, 16(1), 45–69. (2015). doi:10.1007/s40558-015-0047-7
17. Francalanci, C., & Hussain, A.: Discovering social influencers with network visualization: evidence from the tourism domain. *Information Technology & Tourism*, 16(1), 103–125. (2015). doi:10.1007/s40558-015-0030-3
18. Tao, M., Nawaz, M. Z., Nawaz, S., Butt, A. H., & Ahmad, H.: Users' acceptance of innovative mobile hotel booking trends: UK vs. PRC. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0123-x
19. Fazzolari, M., & Petrocchi, M.: A study on online travel reviews through intelligent data analysis. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0121-z
20. Zhang, W., & Fesenmaier, D. R.: Assessing emotions in online stories: comparing self-report and text-based approaches. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0122-y
21. Ghahramani, L., Khalilzadeh, J., & KC, B.: Tour guides' communication ecosystems: an inferential social network analysis approach. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0114-y
22. Van der Zee, E., & Bertocchi, D.: Finding patterns in urban tourist behaviour: a social network analysis approach based on TripAdvisor reviews. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0128-5
23. Al-Ghossein, M., Abdessalem, T., & Barré, A.: Open data in the hotel industry: leveraging forthcoming events for hotel recommendation. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0119-6
24. Qi, S., Wong, C. U. I., Chen, N., Rong, J., & Du, J.: Profiling Macau cultural tourists by using user-generated content from online social media. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0120-0
25. Chan, I. C. C., & Law, R.: Tanja Schneider, Karin Eli, Catherine Dolan, and Stanley Ulijaszek (editors): Digital food activism. *Information Technology & Tourism*. (2018). doi:10.1007/s40558-018-0117-8