Detection And Analysis Of Author`S Geographical Trackin Social Media Content

Iryna Khmil
Dept. of Social Communication
and Information Activity,
Lviv Polytechnic National University
Lviv, Ukraine
Khmil.iryna@gmail.com

Maria Komova
Dept. of Social Communication
and Information Activity,
Lviv Polytechnic National University
Lviv, Ukraine
maria.komova@gmail.com

Abstract— In this paper the algorithm for detection and analysis of author`s geographical track in social media is considered.

Keywords—algorithm; geographical track; web-content; socail media.

I. INTRODUCTION

Analysis social media textual content in present give an opportunity to form consolidated resource of relevant information about geospatial object without direct contact with a large number of experience subjects – authors in social media. Active and continuous analysis of social media text content can provide with a retrospective view of previous experience as one author, and the overall picture of experience which concerning selected geospatial object - that is the general information image.

To get as close to reality image of the information necessary to carry out: 1) sorting on the basis of author`s relevant content, that will enable us to obtain a level of precision of obtained consolidated experience; 2) make a comparative analysis of characteristics changes of each post of each author with the change of the time.

II. ALGORITHM DESCRIPTION

Algorithm for detection and analysis of the author`s geographical track is an extension of the approach to the analysis of the author's activity in social media [1].In this article under consideration geographical track is only using tools based on automatic / automated instruments of social services and which can be used for navigation systems [2] that with great certainty can help get information about where author was at the time of publication posting. But it is not informative in the case about the geographical object as text post can not apply it at all. Therefore, we should consider how geographic information that can be obtained about the author of the content virtual communities, namely copyright texts posts. These characteristics are Residence (does author live near geospatial object or not) and Consumption (Involvement – author were there or not).

Algorithm for detection and analysis of author`s geographic trackin social media text contentconsists in the following steps (fig. 1):

A. Identification of Residence and Consumptionmarkers in the texts of the posts.

These markers are part of the geographical track, and are the first step in obtaining the level of accuracy, relevance and adequacy in the formation characteristics, and further assessments and relative evaluations of descriptors that are extracted from the text of one or more posts of the author.

The presence of at least one marker of consumptionor residence within the text of one author`s post means that we give to this post (and therefore all components that relate to this post – markers, descriptors, authors – [3]) value 1, in the absence - the value 0 to R, C indicators.

B. Inspection of posts to the possibility of analysis.

If post does not contain any marker of Residence or Consumption, i.e. indicators R:=0, C:=0, then we do not take it into account since it is not possible to identify other future characteristics, and the relevance of the post to the original information query.

C. Identification of the author.

Identification of the author is conduct according to the algorithms used in [4]. In our case we add author identification author activity and the availability of informationin database.

D. Identification of time differences.

Since the identification of the author at the preliminary stage it is impossible if there were two authors with the same nickname, the discrepancy between the informationabout one author demonstrates the unreliability of his posts, and therefore and therefore unreliability of defined characteristics of his consumption or residence.Such unreliability can exclude amount of time for which the author`s characteristics can change in connection with the change of status of the geospatial object, and therefore the initialauthor`s experience has to be irrelevant. This time interval chosen 0,5 year as in the most dynamic industry concerning geospatial objects - tourism - even minor changes occur seasonally - for example, winter and summer periods.

In this case change of the author characteristics with indicating of temporal changes`and additional data to the time-dependent database should be entered.

Otherwise - time period less than 0.5 year - all previous components of this author`sposts and data on it should be removed from the database as unreliable and participation in shaping the future of complex

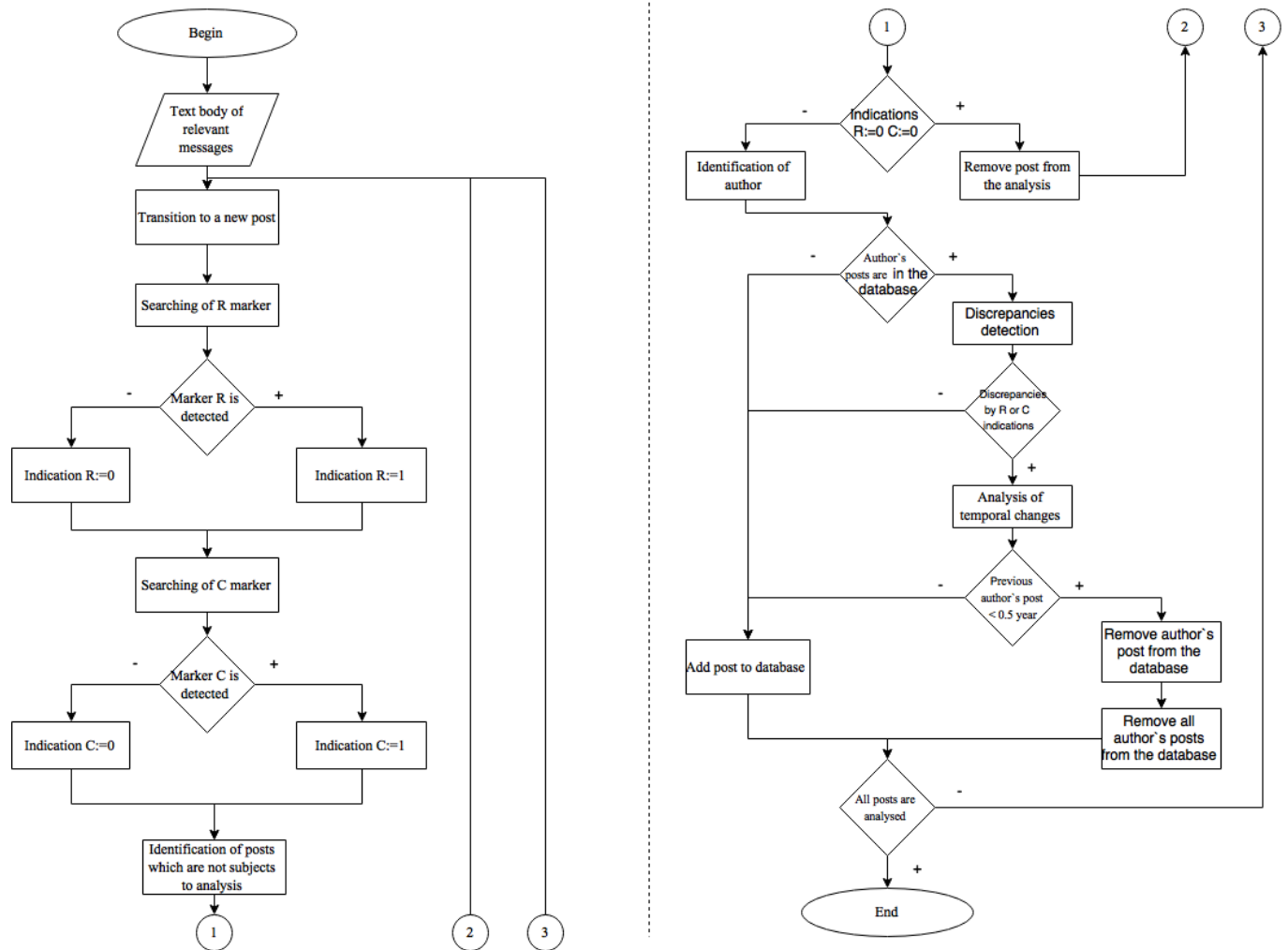geospatial information image of the object will not be taken.

Fig. 1. *Flowchart of detection and analysis of geografical track in social media content*

REFERENCES

[1] A. Stefanidis, A. Crooks, Ja. Radzikowski Harvesting ambient geospatial information from social media feeds, GeoJournal, Volume 78, Issue 2, April 2013, pp 319-338

[2] I. Khmil Content of social media as a sourse of geographic infommation, Conference Proceedings 'MKCI' Kharkiv, Ukraine, vol. 9, April 2015, pp. 54-55.

[3] I. Khmil Analysis of the interaction of geospatial objects` image information components in virtual communities. Management of Development of Complex Systems, Issue 21, 2015, pp. 87-91.

[4] A .Peleshchyshyn, Yu. Serov, S. Fedushko Development of registration and validation algorithm of personal information of Web community members, Computer Science and Information Technologies, Vol. 686, 2010, pp. 238-244.