

ABSTRACT

Master's dissertation: 92 p., 15 fig., 34 tabl., 1 appendix, 31 sources.

Topicality. Today, the Internet is the main channel for receiving or distributing information. Most people use the internet to stay in social networks. Everyone can feel like a true reporter, and share important news, or spread his thoughts to the masses, and find like-minded people / followers. You can also be a simple reader and draw information from other people.

Among all social networks, Twitter has its niche. At the beginning of 2018, its audience was 336 million people. Users communicate with short messages (up to 280 characters) called tweets. This is very convenient, since you can get a very large amount of information, spending it on not too much time.

That is why it is possible to develop a system that can group interest-based users based on the textual component of their behavior in the social network Twitter.

Relationship with academic programs, plans, themes. Work performed at the Department of ASOIU at the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" within the topic "Methods and technologies of high performance computing and performing of big data". Governments register number 0117U000924.

The aim of the research is improving the quality of sharing social network users into groups by analyzing the textual component of their behavior.

To achieve the goal need to accomplish the following **tasks**:

- analyze the data you can get about the user in the social network twitter;
- choose the criteria by which users are compared;
- analyze methods and means of clustering of large data;
- explore ways to analyze user behavior over a text component;
- to implement software analytical system of division of social network users into interest groups.

Object of research - the process of grouping users into groups of common interests.

Subject of research - methods of clustering users of the social network based on a large number of data on their activities. Scientific novelty of the obtained results.

The scientific novelty of the results - to analyze the existing methods of clustering. The methods of analysis of the text component were also analyzed, since one of the criteria for comparing similarity was the topic of blogging.

Published works.

Bulgar M.M. Clustering users through their interests / M.M. Bulgar // MODS. 2018. pp. 28-29.

Bulgar M.M. Methods of clusterization the users of the social network twitter / M.M. Bulgar // ISMT. 2018. pp. 28-32.

CLUSTERING, SOCIAL NETWORK, BIG DATA, TWITTER, TWEET,
ANALYS