

**ПРОБЛЕМА АНАЛІЗУ ПОВІДОМЛЕНЬ З МЕТОЮ ВИЯВЛЕННЯ
ЕКСТРЕМІСТСЬКОЇ ІНФОРМАЦІЇ В МЕРЕЖІ ІНТЕРНЕТ**

UDC 004.912

Zhavruk R.**THE PROBLEM OF ANALYSIS OF MESSAGES FOR THE PURPOSE OF
DETECTING EXTREMISTIC INFORMATION ON THE INTERNET**

Метою дослідження є розробка програмного забезпечення, яке за допомогою методів кластеризації текстової інформації буде аналізувати публічні канали новин, видавничі канали та коментарі користувачів для виявлення повідомлень, що несуть терористичні загрози. Питання аналізу текстової інформації на природній мові було розглянуто в роботах [1–7]. Метод класифікації документів за допомогою нейронних мереж, що самонастроюються використовувався у роботі [8]. У роботі [9] метод к-середніх був застосований до кластеризації в задачах розпізнавання повітряних об'єктів в умовах неповних і недостовірних даних. В ході роботи були проаналізовані методи кластеризації текстової інформації, які можуть бути використані для вирішення проблеми аналізу публічних текстових повідомлень з метою виявлення терористичних загроз. В результаті дослідження та порівняння переваг та недоліків методів кластеризації було вирішено використовувати метод карт Кохонена (SOM) для розробки програмного рішення. Подальші дослідження будуть присвячені розробці програмного забезпечення для аналізу публічних текстових повідомлень за допомогою методу SOM.

Література.

1. B. O. Bliznyuk, L. V. Vasiliev, I. D. Strelnikov, D. S. Tkachuk. Modern methods of natural language processing. Bulletin of Kharkiv National University named after V. N. Karazin, 2017.
2. A.Yu. Perevalova. The use of clustering method for information resource classification. Siberian State Aerospace University named after academician M. F. Reshetnev, 2013.
3. Адуенко А. А., Кузьмин А. А., Стрижов В. В. Выбор признаков и оптимизация метрики при кластеризации коллекции документов //Известия Тульского государственного университета. Естественные науки. – 2012. – №. 3.
4. Milos Ilic, Petar Spalevic, Mladen Veinovic. Suffix Tree Clustering – Data mining algorithm. Faculty of Technical Science Kosovska Mitrovica, University of Pristina-temporally seated in Kosovska Mitrovica Faculty of Informatics and Computing, Singidunum University, Belgrade, 2014.
5. Пархоменко П. А., Григорьев А. А., Астраханцев Н. А. Обзор и экспериментальное сравнение методов кластеризации текстов //Труды Института системного программирования РАН. – 2017. – Т. 29. – №. 2.
6. Чугаинов К. В. Методы тематической кластеризации новостных статей //Научно-практические исследования. – 2017. – №. 2. – С. 295-298.
7. Han H., Jung H., Eom H., Yeom H.Y. Scatter-Gather-Merge: An Efficient Star-join Query Processing Algorithm for Data-parallel Frameworks. Cluster Computing, 2011.
8. Андреев А. М., Березкин Д. В., Козлов И. А. Подход к автоматизированному мониторингу тем на основе обнаружения событий в потоке текстовых документов //Информационно-измерительные и управляющие системы. – 2017. – Т. 15. – №. 3. – С. 49-60.
9. Antropov V. V.. Application of k-means and g-means clustering algorithms in objects recognition. St. Petersburg Mining University, 2017