ABSTRACT

Master's Thesis: 76 pages, 15 figures, 26 tables, 1 appendix, 15 sources.

Actuality. Authentication and verification is a unique and, at the same time, a very widespread task, given the possibility of being used in various fields of human activity: to combat plagiarism, to establish the authorship of anonymous texts, to assess and establish a person in criminology and many other tasks, and directions.

Problems of the study of author's style and application of statistical analysis in the study of authorship were considered in the researches of OO Archipova and V.M. Zhuravleva, LA Borisova, Yu.N.Orlov and K.P. Osmina.

The problem is also very complex because of the fundamental problem of forming a set of features that can be used to evaluate the likelihood of a text belonging to a particular author.

The proble is compounded by the fact that until recently, for the developed systems for determining the authorship of texts, a prerequisite for their stable and quality work was the availability of large volumes of copyrighted text in the training sample. Another drawback of the developed models is their qualitative limitation on the number of authors.

The aim of the study. The purpose of the study is to develop an algorithm and implement software for analysis of the author of the text with the help of statistical analysis methods, as well as to establish the influence of the names of characters and words in capital letters as a whole on the frequency characteristics of the text.

Objectives of the study. To achieve this goal, you must complete the following tasks:

- ⁃ - to investigate the works of authors in this field;

-      - analyze the author's existing search systems;

-      - to develop an algorithm for analysis of authorship;

-      - to develop a software application based on the developed algorithm;

-      - analyze the accuracy of the algorithm;

-      - to make conclusions about the expediency of eliminating the algorithm.

The object of study - the artistic texts.

The subject of the study is the analysis of artistic texts in order to determine authorship.

Research methods. We used statistical analysis methods to build the algorithm, object-oriented programming methods to develop a software application.

Scientific novelty. The most significant scientific results of the master's thesis are:

-      - the developed algorithm gives high accuracy in determining the authorship of artistic texts;

-      - the influence of the names of characters and capital letters as a whole on the frequency characteristics of the text was investigated.

Practical meaning. The resulting software allows you to analyze artistic texts and to determine the author of the text with high accuracy.

Relationship with scientific programs, plans and topics. The work was performed at the Department of Automated Information Processing and Management Systems of the National Technical University of Ukraine «Kyiv Polytechnic Institute. Igor Sikorsky "within the topic" Mathematical Models and Technologies in DSS ". State Registration Number 0117U000914.

Approbation. The main provisions of the work were reported and discussed at the 18th International Scientific Conference "Mathematical and Software Software of Intelligent Systems IPAS-2019" and 3 All-Ukrainian Scientific and Practical Conference of Young Scientists and Students "Information Systems and Management Technologies" (ISTU-2019).

Publications. The scientific provisions of the dissertation are published in the theses of the XVIII International Scientific-Practical Conference "Mathematical and Software Software of Intelligent Systems IPAS-2019" and in the Theses of 3 All-Ukrainian Scientific-Practical Conference of Young Scientists and Students "Information Systems and Technologies of Management" (ISTU-2019).

Keywords: Statistical analysis of texts, Definition of authorship.