УДК 004.912

**РАЗРАБОТКА ПРОГРАММНЫХ СРЕДСТВ ОПРЕДЕЛЕНИЯ ПСИХОЛОГИЧЕСКОГО ПОРТРЕТА ЛИЧНОСТИ НА ОСНОВЕ АНАЛИЗА ЭЛЕКТРОННОГО ТЕКСТА**

Н.И. УТКИНА, А.М. РУСАКОВ

*Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский технологический университет»*

*119454, Москва, Россия*

В психологии для составления психологического портрета личности по почерку, как правило, применяют графологический способ, который на данный момент является наиболее эффективным [1]. Каждый человек обладает своим уникальным почерком, по которому можно много рассказать о его обладателе и возможно построить довольно подробный психологический портрет. Но текст не всегда может быть написан от руки. Все чаще приходится работать электронным текстом [1, 2].

В настоящее время люди большую часть времени проводят в Интернете, в так называемом «Виртуальном мире». Сейчас в Интернете можно найти практически любую информацию. Даже общение между людьми теперь проходит в различных Интернет-сервисах, таких как социальные сети, форумы, мессенджеры и др. Человек окружает себя электронными устройствами — «гаджетами» и пользуется ими не только для развлечения, но и для делового общения. Поэтому возникает потребность понять каков ваш собеседник и что он на самом деле из себя представляет [5]. Дело в том, что пользователи Интернета могут выдавать себя за тех, кем они на самом деле не являются. Ваш виртуальный собеседник может выдать себя за другого человека и сообщить заведомо ложную информацию [2, 3], так как в Интернете нет эффективных способов проверки ее на достоверность. При этом, так или иначе сам по себе электронный текст, который пишет ваш виртуальный собеседник, возможно проанализировать и на основе этого анализа получить психологический портрет конкретной личности. И с некоторой долей вероятности идентифицировать кто на самом деле перед вами. Характерные особенности личности, которые ее идентифицируют, невозможно скрыть, так же как невозможно скрыть истинный почерк человека.

Таким образом, анализ электронного текста и построение по нему психологического портрета личности предлагается выполнить в четыре этапа, которые представлены на рисунке 1.



Рисунок 1 – Основные этапы анализа электронного текста.

Сначала необходимо получить исходный электронный текст. Потом выполнить его фильтрацию. В общем случае фильтрация, сводится к удалению лишних пробелов и знаков препинания. Система посимвольно перебирает текст и удаляет лишние символы, которые могут помешать дальнейшему его анализу. Далее выполняется частотный анализ текста, который состоит в определении встречаемости буквы на основе статистики общей частотности букв в русском языке и речи [6, 7]. И наконец система отображает полученный результат в виде диаграммы.

Анализ текста происходит в виде расчета характеристик на основе метода Журавлева [6]. По звукобуквам рассчитывается средняя арифметическая оценка слова по шкале характеристики. Журавлев выделил всего 25 шкал характеристик: «хороший – плохой», «красивый – отталкивающий», «безопасный – страшный», «радостный – печальный», «светлый – темный», «легкий – тяжелый», и др. характеристики,
описанные в работе [6].

Далее анализ текста выполняется по формуле расчета характеристики слова по какой-либо шкале значимости [6, 7] имеет следующий вид:

 $F= \frac{\sum\_{}^{}f\_{i}k\_{i}}{\sum\_{}^{}k\_{i}}$ (1)

где F — фонетическая значимость слова; $f\_{i}$, — фонетическая значимость очередного (i-того) звука слова; $k\_{i}$ — коэффициент для очередного (i-того) звука, который рассчитывается по формуле:

$k\_{i}= \frac{P\_{max}}{P\_{i}}$ (2)

где $P\_{max}$ — максимальная частотность звука в данном слове, $P\_{i}$ – частотность любого звука. Учитывая положение каждого звука в слове, коэффициент ($k\_{i}$) первого звука в слове увеличивается в 4 раза. Для ударного звука в слове коэффициент ($k\_{уд}$) увеличивается в два раза см. работу [6].

Далее проведя нормирование и калькуляцию значимостей характеристик по формулам (1) и (2) можно, например, получить следующую диаграмму (рисунок 2).

Рисунок 2 – Пример определения психологического портрета личности на основе анализа её электронного текста.

Из диаграммы на рисунке 2, можно заключить, что человек, который писал анализируемый электронный текст имеет добрый настрой, но достаточно груб, и возможно имеет некоторый злой умысел. Таким образом, применение предложенного подхода к анализу электронного текста может повысить общий уровень информационной безопасности за счёт определения истинного психологического портрета виртуального собеседника.

**Список использованных источников**

1. IBM, The science behind the service [Электронный ресурс] // Last Updated: 12.10.2017. URL: <https://console.bluemix.net/docs/services/personality-insights/science.html#science> (Дата обращения: 1.04.2018)

2. Lightcaster, Генератор текста на основе триграмм (python) [Электронный ресурс] // 24.03.2010. URL: <https://habrahabr.ru/post/88514/> (Дата обращения: 3.04.2018)

3. An Empirical Study of Smoothing Techniques for Language Modeling. Stanley F. Chen & Joshua Goodman, Harvard University // 1998 (Дата обращения: 01.05.2018)

4. Speech and Language Processing. Daniel Jurafsky & James H. Martin, Stanford University // 2014 (Дата обращения 10.05.2018)

5. BrainApps, Что может сказать о состоянии человека печатный текст [Электронный ресурс] // 1.04.2016. URL: <https://brainapps.ru/blog/2016/04/pechatniy-tekst/> (Дата обращения: 01.04.2018)

6. А.П.Журавлев, ЗВУК И СМЫСЛ [Электронный ресурс] // Издание 2-е, исправленное и дополненное, МОСКВА «ПРОСВЕЩЕНИЕ» 1991, URL:http://svitk.ru/004\_book\_book/14b/3142\_juravlev-zvuk\_i\_smisl.php (Дата обращения: 20.03.2018)

7. Левицкий В.В, ФОНЕТИЧЕСКА МОТИВИРОВАННОСТ СЛОВА // 1994 г. URL: [http://www.issuesinlinguistics.ru/pubs/фонетическая-мотивированность-слова](http://www.issuesinlinguistics.ru/pubs/%D1%84%D0%BE%D0%BD%D0%B5%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B0%D1%8F-%D0%BC%D0%BE%D1%82%D0%B8%D0%B2%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%BD%D0%BE%D1%81%D1%82%D1%8C-%D1%81%D0%BB%D0%BE%D0%B2%D0%B0) (Дата обращения: 21.03.2018)