Web Page Classification Information System as an information software system reducing access to Internet resources using keywords with wages

Piotr Wojcik
PJA,
Warszawa, `piotr.wojcik@pja.edu.pl`

July 8, 2018

1 Abstract

Keywords: *Internet, classificator, classification, keywords with wage, DNS*

2 Introduction

3 DNS

4 Keywords with wages

5 Keywords wages update and datamining

Alg 1. Podstawowy algorytm klasyfikacji stron internetowych

```
1
2  //input:
3  //     phrase − analysed phrase,
4  //     analysedPage − web page with URL and other characteristics,
5  //     compromisedWebPages − colection of compromised web pages,
6  //     redFlagPhraseDictionary − colection of phrases
7  //                              that compromise web page
8  //output:
9  //     score − points of page classification
10 //     compromisedWebPages − modyfied input object
11
12 mostCompromisedPages = compromisedWebPages.top();
13 foreach (phrase in webPageTextContent)
14 {
15   if((phrase in redFlagPhraseDictionary)
```

```
16        || (phrase like redFlagPhraseDictionary))
17   {
18     redFlagPhraseDictionary.computeWage(phrase, mostCompromisedPages);
19     score += redFlagPhraseDictionary.wage(phrase);
20     compromisedWebPages += analysedPage;
21   }
22 }
```

1.

2.

3.

6   Web pages category

-
-
-
-
-
-
-
-
-
-
-

7   Web Page Classification Information System modules

8   Web Page Classification Information System appliance

9   Conclusion and future plans