

R. Sh. Gasanova

EVALUATION OF SCIENTIFIC PAPERS BY MEANS OF NEW ALGORITHM ARTICLERANK

The paper considers new modification of PageRank algorithm intended for evaluation of scientific papers and the model of new type of ArticleRank. In certain cases for automatic evaluation of scientific papers, arranged in one network computer program Delphi with wide possibilities for analysis of qualitative evaluation of scientific papers, is recommended. It is shown that within one network papers, having the same number of references, have different values of ArticleRank.

Key words: PageRank, ArticleRank, impact-factor, Hirsch index.

Introduction

Previously the number of scientific papers was considered to be main index of scientific activity. Nowadays, as a result of considerable increase of the number of published papers there appeared the necessity of their qualitative evaluation. Various systems of evaluation existed for this purpose [1].

Taking into consideration rapid development of computing technologies the elaboration of adequate, comprehensive ratings, applying computational facilities becomes very actual. In recent years various mechanisms for automatic evaluation of scientific activity of scientists have been developed. The example of such mechanism is impact-factor numerical index of scientific journal, annually calculated by the Institute of scientific information since 1960. Nowadays, Thomson corporation has database, where more than 8400 scientific journals from different countries are registered [2]. Besides, H-index for evaluation of researcher output, suggested by American physicist J. E. Hirsch in 2005, can be mentioned [3 – 6]. This index is based on the quantity of publications and references to each of these publications. At present, the list of H-index for scientists, working in various branches of science, is presented in reference database of Scopus or Web of Science [7]. Scientists must apply much efforts and time for impartial evaluation of dissertations. In order to automate this process, some aspects of evaluation as: proximity between the sections in the dissertations and definition of the actuality of the research, are described in the paper [8]. Besides, in the course of search in the Internet the appearance of necessary and unnecessary information indicates that information sites must be filtrated. The most important and useful information sites must be located in the first lines of the list, composed on the results of search. This problem was solved for the first time in 1998 by L. Page and S. Brin. Algorithm PageRank developed by them, was successfully tested in the system Google and led to formation of functional system, where PageRank of this site can be computed only by the input of URL-address [9]. This software, the search system Google is based on, improves the efficiency of information search. In mathematical description of this algorithm, the formula demonstrating the level of importance of PageRank page p_i is used:

$$PR(p_i) = \frac{1-d}{N} + d \sum_{p_j \in B(p_i)} \frac{PR(p_j)}{O(p_j)},$$

where p_1, p_2, \dots, p_N – are considered pages, $B(p_i)$ – is the number of pages referring to page p_i , $O(p_j)$ – is the number of references leaving page p_j , N – is total amount of pages, $d \in [0, 1]$ – is damping factor.

The aim of the given paper is the development of new modification of PageRank algorithm for evaluation of scientific papers. For this purpose PageRank algorithm with application of developments [10, 11], as well as model, representing the importance of the paper u among other papers in the given sphere of science are used

$$AR(u) = \frac{1-d}{N} + d \sum_{v \in B(u)} \frac{AR(v)}{N_v}, \quad (1)$$

where $B(u)$ – is the set of papers, referring to the paper u ; N_v – is total number of references, from the paper v ; d – is damping factor of new type of ArticleRankmodel, with addition to the second term of weight factor:

$$AR(u) = \frac{1-d}{N} + d \cdot w \cdot \sum_{v \in B(u)} \frac{AR(v)}{N_v}, \quad (2)$$

where $w = N_{vu} \cdot IF(v)$ – is weight factor; N_{vu} – is the number of references from paper v to paper u ; $IF(v)$ – impact-factor of journal, where paper v is published.

Schematic description of ArticleRank model. Let us consider the evaluation of the paper u , described in the diagram (Fig. 1):

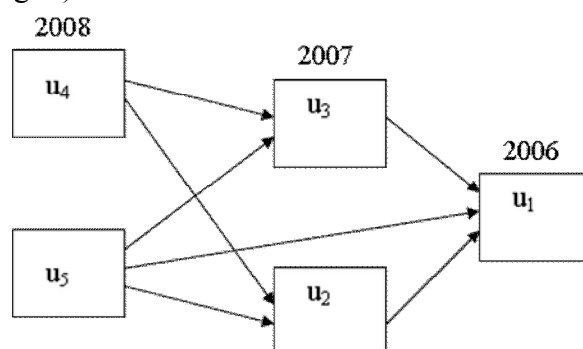


Fig. 1. Schematic description of ArticleRank model

where u_2, u_3, u_4, u_5 – are papers, influencing the paper (i.e., the set of papers, referring to this paper, directly or indirectly).

As it is seen from the diagram at the first stage ArticleRank of the paper, referring to the last year is taken as unit, as in algorithm PageRank, since there are no references, to this paper:

$$AR(u_4) = AR(u_5) = 1.$$

At the second stage by the formula (1) ArticleRank of papers u_2 and u_3 is calculated correspondingly:

$$AR(u_2) = \frac{1-0,85}{5} + 0,85 \cdot \left(\frac{AR(u_4)}{2} + \frac{AR(u_5)}{3} \right) = 0,03 + 0,7055 = 0,7355;$$

$$AR(u_3) = \frac{1-0,85}{5} + 0,85 \cdot \left(\frac{AR(u_4)}{2} + \frac{AR(u_5)}{3} \right) = 0,03 + 0,7055 = 0,7355.$$

Hence, in this example the third stage produces the needed evaluation value of the paper u_1 :

$$AR(u_1) = \frac{1-0,85}{5} + 0,85 \cdot \left(\frac{AR(u_5)}{3} + \frac{AR(u_2)}{1} + \frac{AR(u_3)}{1} \right) = 1,834.$$

Let us describe evaluation, with addition of proposed weight coefficients. For this purpose we include the proposed weight coefficients in the formula (2) and diagram (Fig. 2).

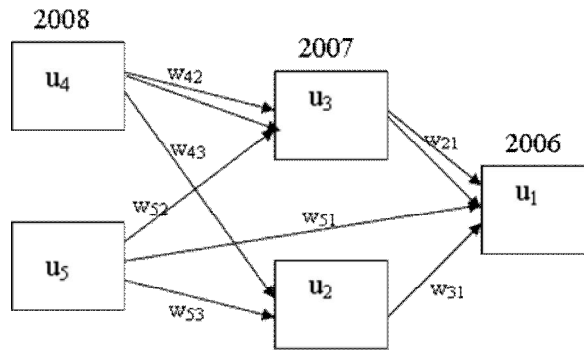


Fig. 2. Addition of weight coefficients in ArticleRank model

We use the value of impact-factor of the journals, where articles, described in the diagram as weight functions, are published. Let us assume, that journals, described in the diagram, have impact factor, described below:

$$IF(u_2) = 0,31, IF(u_3) = 0,71, IF(u_4) = 0,6, IF(u_5) = 0,096$$

Impact-factor of the article u_1 is not needed for the time being, since it does not participate in evaluation of the article itself. We will assign these values to the articles, described in the last scheme (Fig. 2):

$$\begin{aligned} w_{42} &= IF(u_4) \cdot N_{u_4} = 0,6 \cdot 2 = 1,2; & w_{43} &= IF(u_4) \cdot N_{u_4} = 0,6 \cdot 1 = 0,6; \\ w_{51} &= IF(u_5) \cdot N_{u_5} = 0,096 \cdot 1 = 0,096; & w_{52} &= IF(u_5) \cdot N_{u_5} = 0,096 \cdot 1 = 0,096; \\ w_{53} &= IF(u_5) \cdot N_{u_5} = 0,096 \cdot 1 = 0,096; & w_{21} &= IF(u_2) \cdot N_{u_2} = 0,31 \cdot 2 = 0,62; \\ w_{31} &= IF(u_3) \cdot N_{u_3} = 0,71 \cdot 1 = 0,71. \end{aligned}$$

After computation of weighted functions, we may pass to calculation of ArticleRank of the paper. Values of the paper, published in 2008 equal unit:

$$AR(u_4) = AR(u_5) = 1.$$

At the second stage the following calculations are carried out:

$$\begin{aligned} AR(u_2) &= \frac{1-0,85}{5} + 0,85 \cdot \left(\frac{AR(u_4)}{N_{u_4}} \cdot w_{42} + \frac{AR(u_5)}{N_{u_5}} \cdot w_{52} \right) = 0,03 + 0,85 \cdot \left(\frac{1}{3} \cdot 1,2 + \frac{1}{3} \cdot 0,096 \right) = 0,3972; \\ AR(u_3) &= \frac{1-0,85}{5} + 0,85 \cdot \left(\frac{AR(u_4)}{N_{u_4}} \cdot w_{43} + \frac{AR(u_5)}{N_{u_5}} \cdot w_{53} \right) = 0,03 + 0,85 \cdot \left(\frac{1}{3} \cdot 0,6 + \right. \\ &\quad \left. + \frac{1}{3} \cdot 0,096 \right) = 0,2272; \\ AR(u_1) &= \frac{1-0,85}{5} + 0,85 \cdot \left(\frac{AR(u_2)}{N_{u_2}} \cdot w_{21} + \frac{AR(u_5)}{N_{u_5}} \cdot w_{51} + \frac{AR(u_3)}{N_{u_3}} \cdot w_{31} \right) = 0,03 + 0,85 \times \\ &\quad \times \left(\frac{0,3972}{2} \cdot 0,62 + \frac{1}{3} \cdot 0,096 + \frac{0,2272}{1} \cdot 0,71 \right) = 0,299. \end{aligned}$$

Software of ArticleRank model in Delphi. Computer program, providing the choice of

interconnected papers in the base and automatic computation of ArticleRank is written on the base of constructed model in Delphi. Main data for the software are set in the following manner: in the window "Adding" (Fig. 3) the titles of the papers and their references are to be introduced in increasing order of years.

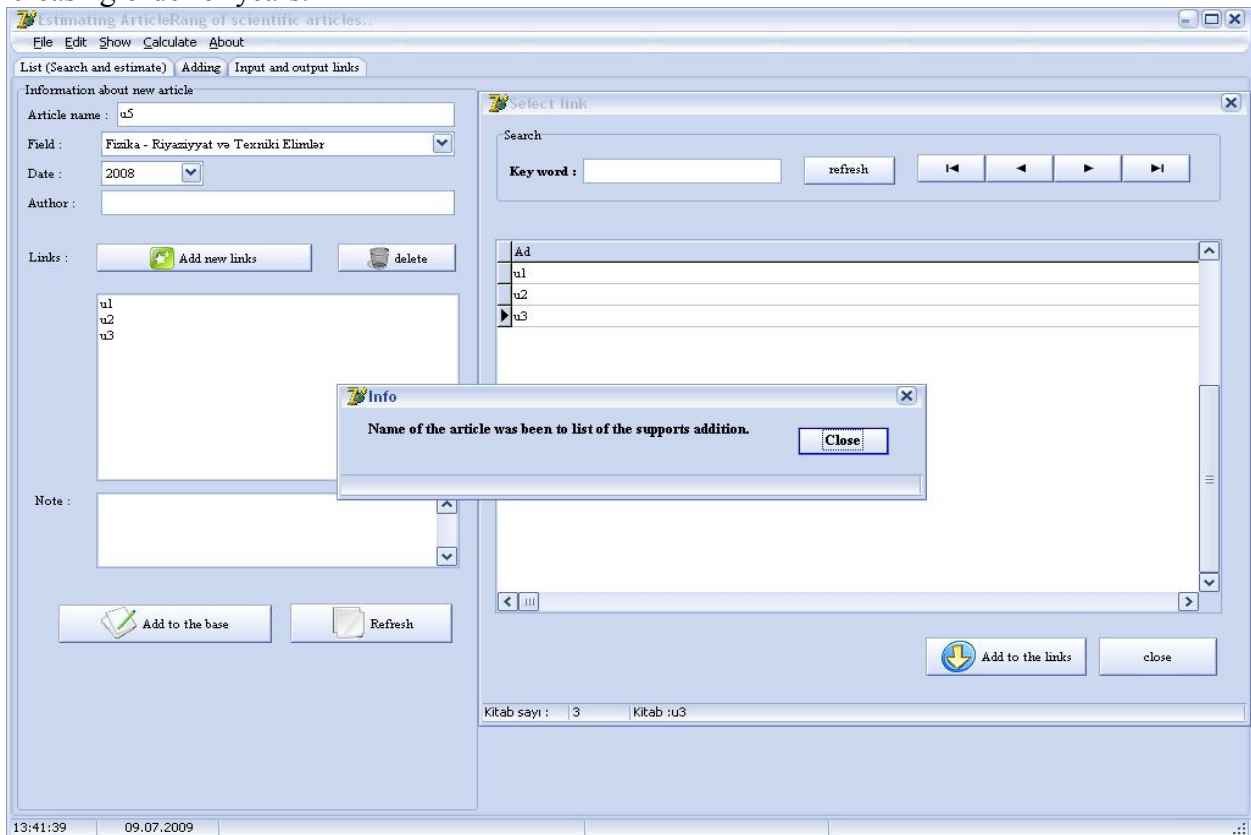


Fig. 3. Addition of papers and references in the base

The text of this fragment in Delphi is presented in the following manner:
procedure TEsas.TntBitBtn7Click(Sender: TObject);

```
Var
  i:integer;
begin
  If (TntEdit3.Text<>"")and(TntEdit4.Text<>"")and(TntComboBox3.Text<>"") then
  Begin
    TntBitBtn7.Enabled:=False;
    Esas.AdoQueryh.SQL.Text:='INSERT INTO siyahi(ad,tarix,deqiqlik,muellif+
      ',elave) VALUES(:p0,:p1,:p2,:p3,:p4)';
    Esas.AdoQueryh.Parameters.ParamByName('p0').Value:=TntEdit3.Text;
    Esas.AdoQueryh.Parameters.ParamByName('p1').Value:=TntEdit4.Text;
    Esas.AdoQueryh.Parameters.ParamByName('p2').Value:=TntComboBox3.ItemIndex;
    Esas.AdoQueryh.Parameters.ParamByName('p3').Value:=TntEdit7.Text;
    Esas.AdoQueryh.Parameters.ParamByName('p4').Value:=TntMemo1.Text;
    Esas.AdoQueryh.ExecSQL;
    If TntListBox1.Items.Count>0 then
      For i:=0 to TntListBox1.Items.Count-1 do
      Begin
        AdoQueryh.SQL.Text:='INSERT INTO istinadlar(kid,iid) VALUES(:p0,:p1)';
        AdoQueryh.Parameters.ParamByName('p0').Value:=NewRecordNo;
        AdoQueryh.Parameters.ParamByName('p1').Value:=TntListBox2.Items.Strings[i];
        AdoQueryh.ExecSQL;
```

```

End;
ElaveEdildi;
SiyahiniYaz;
TntBitBtn7.Enabled:=True;
ShowInfo('Info',Sozler.Siyahi.Lines.ValueFromIndex[0]);
End else
ShowInfo('Info',Sozler.Siyahi.Lines.ValueFromIndex[1]);
end;

```

In the window “Input and output links” (Fig. 4) input and output references of the selected paper are presented

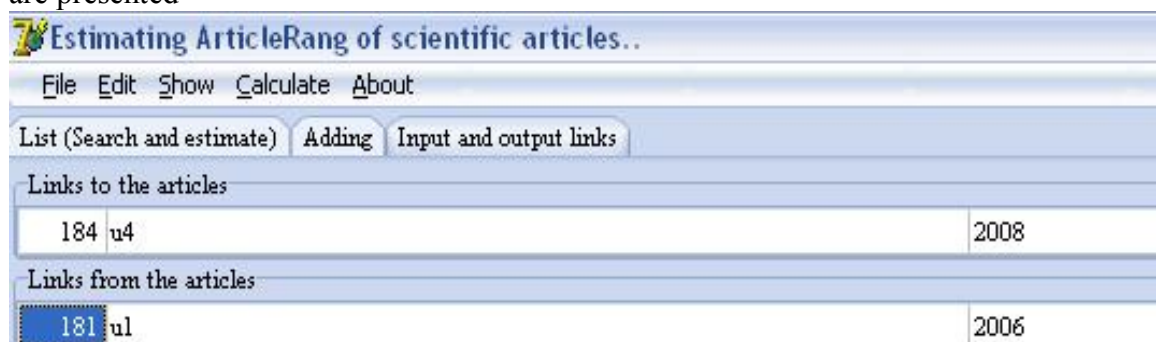


Fig. 4. List of input and output references of the paper

Text of the program:

Procedure ButunCedvel(Id:integer);

Begin

```

Esas.ADOQuery4.SQL.Text:='SELECT Id,Ad,Tarix,Sbk,Elave FROM siyahi WHERE '+'
'(id in( SELECT iid FROM istinadlar WHERE '+'kid =:p0))';

```

```

Esas.ADOQuery4.Parameters.ParamByName('p0').Value:=id;

```

```

Esas.ADOQuery4.Open;

```

```

Esas.ADOQuery3.SQL.Text:='SELECT Id,Ad,Tarix,Sbk,Elave FROM siyahi WHERE '+'
'(id in( SELECT kid FROM istinadlar WHERE '+'
'iid =:p0))';

```

```

Esas.ADOQuery3.Parameters.ParamByName('p0').Value:=ID;

```

```

Esas.ADOQuery3.Open;

```

```

Esas.TntDBGrid3.Columns[0].Width:=40;

```

```

Esas.TntDBGrid3.Columns[1].Width:=400;

```

```

Esas.TntDBGrid3.Columns[2].Width:=90;

```

```

Esas.TntDBGrid3.Columns[3].Width:=50;

```

```

Esas.TntDBGrid3.Columns[4].Width:=100;

```

```

Esas.TntDBGrid4.Columns[0].Width:=40;

```

```

Esas.TntDBGrid4.Columns[1].Width:=400;

```

```

Esas.TntDBGrid4.Columns[2].Width:=90;

```

```

Esas.TntDBGrid4.Columns[3].Width:=50;

```

```

Esas.TntDBGrid4.Columns[4].Width:=100;

```

End;

In the window List (Search and estimate) (Fig. 5) ArticleRank of the paper can be calculated and title of the paper and other data can be edited.

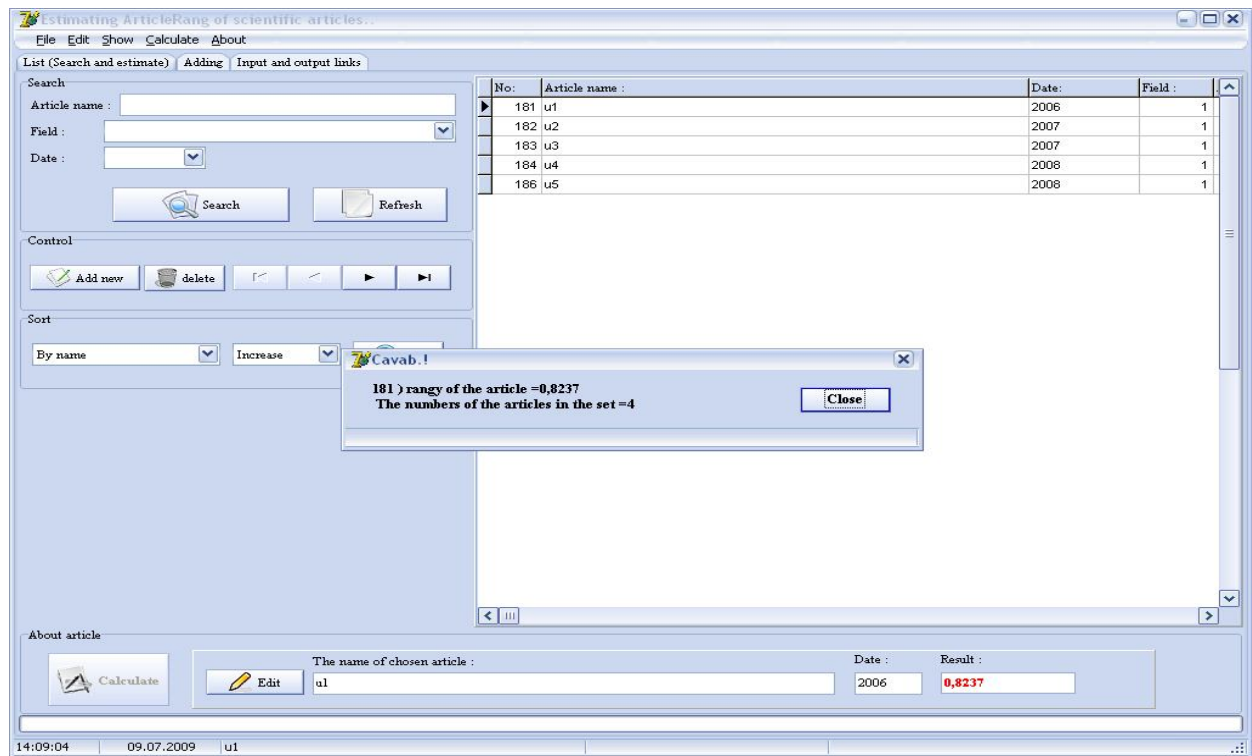


Fig. 5. ArticleRank calculation of the allocated paper

The text of the program of the given window will be the following:

```
Function RandomName:String;
Var
  s:string[11];
Begin
  Randomize;
  s:=chr(random(26)+65)+chr(random(26)+65)+chr(random(26)+65)+chr(random(26)+65)+chr(random(26)+65)+c
hr(random(26)+65);
  RandomName:=s;
End;
Procedure SebekeleriYenile;
Begin
  Esas.TntPageControl1.Pages[0].Show;
  Esas.ADOQueryH.SQL.Text:='Select id,tarix from siyahi Order by tarix asc';
  Esas.ADOQueryH.Open;
  Esas.ADOQueryH.First;
  SebekeniYenileN:=Esas.ADOQueryH.RecordCount;
  SebekeniYenileI:=0;
  Esas.TntProgressBar1.Max:=SebekeniYenileN ;
  Esas.TntProgressBar1.Min:=SebekeniYenileI;
  Esas.SebekeniYenileT.Enabled:=True;
End;
//*****
Function SebekdekiJurnallar(id:integer):integer;
Var sbk:string[10];
Begin
  Esas.ADOQueryH.SQL.Text:='SELECT sbk FROM siyahi WHERE id=:p0';
  Esas.ADOQueryH.Parameters.ParamByName('p0').Value:=id;
  Esas.ADOQueryH.Open;
```

```

    sbk:=Esas.ADOQueryH.Fields.Fields[0].Value;
    Esas.ADOQueryH.SQL.Text:='SELECT id FROM siyahi WHERE sbk=:p0';
    Esas.ADOQueryH.Parameters.ParamByName('p0').Value:=Sbk;
    Esas.ADOQueryH.Open;
    SebekdekiJurnallar:=Esas.ADOQueryH.RecordCount;
End;
// *****//
//           Estimating           //
// *****//
Procedure MassiviYenile;
Var
  i:integer;
Begin
  For i:=0 to MaxCount do
    count[i]:=0;
  End;
Function AR(n,k,id:integer):real;
Var
  c,rar:real;
  j:byte;
Begin
  Esas.ADOQueryH.SQL.Clear;
  Esas.ADOQueryH.SQL.Text:='SELECT id,Ad FROM siyahi WHERE '+'
    'id in( SELECT kid FROM istinadlar WHERE '+'
    'iid=:p0)';
  Esas.ADOQueryH.Parameters.ParamByName('p0').Value:=id;
  Esas.ADOQueryH.Open;
  Esas.ADOQueryH.First;
  count[k]:=Esas.ADOQueryH.RecordCount;
  If count[k]>0 then
    Begin
      For j:=0 to count[k]-1 do
        Begin
          data[k,j].id:=Esas.ADOQueryH.Fields.Fields[0].AsInteger;
          data[k,j].icount:=IstinadSayi(Esas.ADOQueryH.Fields.Fields[0].AsInteger);
          Esas.AdoQueryH.Next;
        End;
        c:=0;
      For j:=0 to count[k]-1 do
        c:=c+ar(n,k+1,data[k,j].id)/data[k,j].icount;
        rar:=(1-d)/n+d*c;
      End else
      Begin
        rar:=1;
      End;
      Ar:=rar
    End;
  Procedure Hesabla;
  Var
    rAR:real;
    sAR:string;

```

```

n:integer;
begin
If Esas.ADOQuery1.Fields.Fields[0].AsString<>" then
Begin
Esas.HesablaB.Enabled:=False;
n:=SebekdekiJurnallar(Esas.AdoQuery1.Fields.Fields[0].Value);
Esas.TntPageControl1.Pages[0].Show;
MassiviYenile;
rAR:=AR(n,0,Esas.AdoQuery1.Fields.Fields[0].AsInteger);
sAR:=Copy(FloatToStr(rAR),1,6);
Esas.ARText.Text:=sAR;
ShowInfo('Cavab.!',Esas.AdoQuery1.Fields.Fields[0].AsString+Sozler.Siyahi.Lines.Strings[5]+sAR+#13#10+Sozler.Siyahi.Lines.Strings[8]+IntToStr(n));
Esas.HesablaB.Enabled:=True;
End;
End;

```

Conclusions

On the base of the suggested models and their software realization, the following results have been obtained:

1. Algorithm ArticleRank for evaluation of scientific papers analogous to PageRank algorithm, intended for evaluation of well-pages has been developed.
2. Weight functions, determining references from other papers on analyzed paper are introduced in the model and used, their influence on ArticleRank computation is revealed.
3. In certain cases, for automatic evaluation of scientific papers, located in one network, computer program Delphi is recommended, the program possesses wide possibilities for carrying out the analysis of qualitative evaluation of scientific papers.
4. Within one network, papers having equal number of references, have different values of ArticleRank, since each reference has its own weight coefficient.

REFERENCES

1. Бернал Дж. Наука в истории общества. / Дж. Бернал. – М., 1956. – 311 с.
2. Thomson Reuters [Электронный ресурс] // Режим доступа: <http://www.thomsonreuters.com>.
3. Hirsch J. E. An index to quantify an individual's scientific research output / J. E. Hirsch // Proceedings of the National Academies of Science. – 2005. – Vol. 102. – No. 46. – P. 16569 – 16572.
4. Hirsch, J. E. (2007). Does the h-index have predictive power? [Электронный ресурс] // Режим доступа: http://arxiv.org/PS_cache/arxiv/pdf/0708/0708.0646v2.pdf.
5. Bar-Ilan J. Which h-index? / J. Bar-Ilan // A comparison of WoS, Scopus and Google Scholar Scientometrics. – 2008. – Vol. 74. – No. 2. – P. 257 – 271.
6. Cronin B. Using the h-index to rank influential information scientists / B. Cronin, L. I. Meho // Journal of the American Society for Information Science and Technology. – 2006. – Vol. 57. – No. 9. – P. 1275 – 1278.
7. Jacso Peter. The pros and cons of computing the h-index using Google Scholar”/ Peter Jacso // Online. Information Review. – 2008. – Vol. 32. – No. 3. – P. 437 – 452.
8. Гасанова Р. III. Автоматизація оцінювання дисертаційних робіт / Р. III. Гасанова // Вісник ВПІ. – 2009. – № 1. – С. 89 – 91.
9. Этичное продвижение сайтов. [Электронный ресурс] // Режим доступа: <http://linkclub.ru>.
10. Brin S. The anatomy of a large-scale hypertextual web search engine / S. Brin, L. Page // Computer Networks ISDN Systems. – 1998. – vol. №№ 1 – 7. – P. 107 – 117.
11. Ma N. Bringing PageRank to the citation analysis / N. Ma, J. Guan, Y. Zhao // Information Processing and Management. – 2008. – vol. 44. – № 2 – P. 800 – 810.

Gasanova Rakhilia Shaban kyzy – Researcher with the Institute of Information Technologies of National Academy of Science, Azerbaijan, tel.: (994 12) 4397213, e-mail: rahasanova@gmail.com.
National Academy of Science.