

Quantitative Network Security Evaluation and Searching Path Selection Based on Weighted Network Topology

Chunhui Han

College of foreign languages of Qiqihar University, Qiqihar, 161000, China

chunhui_han@yeah.net

Keywords: Quantitative evaluation, Encounter probability, Markov, DTN clustering, Integral coefficient, Corpus.

Abstract. Based on the research object of English translation corpus and its official artificial score, through the corpus annotation and statistical analysis, this paper uses corpus linguistics method to reveal the main factors of rater concerning in the translation test artificial scores. Based on dynamic semi Markov path searching model, this paper proposes a new DTN clustering routing method of vocabulary encounter probability evaluation. This method introduces an integral coefficient in integral, which improves the quantization range and efficiency of translation test. Through the quantitative evaluation analysis of English translation corpus, the algorithm can effectively calculate the class identifier number, number of shape symbols, pictogram category symbols ratio, average sentence length and other parameters, which greatly improves the efficiency of test and evaluation, and it is a new computer quantitative method for English translation.

Introduction

English Chinese translation can not only detect the candidate's comprehensive language ability and cultural accomplishment, but also can test the learning results of language learner. But the current assessment of English Chinese translation test is not optimistic. In the translation test quantitative evaluation method, it has not yet achieved the automatic evaluation of translation, and the translation work requires a lot of manpower and material resources [1-3]. In the translation performance assessment the proportion of human factors is big, which can not accurately reflect the level of the subjects, so the test credibility is lower. This paper adopts dynamic Markov method and designs the DTN clustering routing model of English translation test quantitative evaluation, and the overall design framework is as follows:

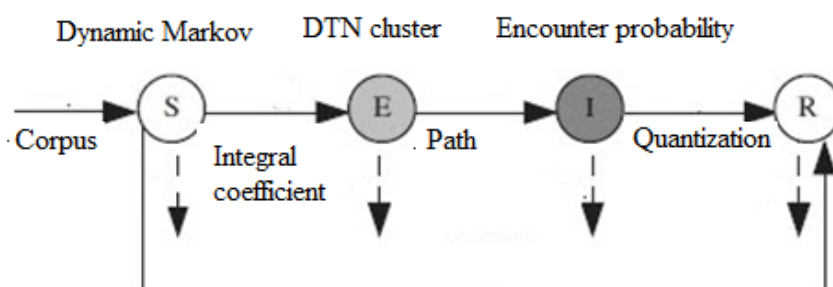


Fig.1: The basic framework design of translation quantification

Figure 1 shows the basic design framework for translation quantitative methods [4]. The basic process is firstly processing corpus data, using dynamic Markov search algorithm introducing into the integral constant, secondly according to the integration constant to design DTN clustering routing, then accounting on the encounter probability of two words, finally the quantitative evaluation result of translation is obtained.

Corpus Quantitative Evaluation Model Optimization and Algorithm Design

In order to realize the quantitative evaluation of corpus, it can do quantization on database according to the occurrences frequency and sort of translation vocabulary [5-7]. For the whole

corpus, it can use probabilistic evaluation function to do quantization evaluation on the translation. The evaluation function can be written as:

$$s(x) = x_p e(x) + x_i \sum_{j=0}^x e(j) + x_d [e(x) + e(x-1)] \quad (1)$$

Among them, $e(x)$ is the probability when the two translation terms appearing simultaneously, T is cycle, x_p is frequency, x_i is integral coefficient, and its expression is as follows

$$x_i = \frac{x_p T}{T_n} \quad (2)$$

Where x_d is differential coefficient, and the expression is as shown in formula (3).

$$x_d = \frac{x_p T_d}{T_n} \quad (3)$$

Based on the clustering model of translation vocabulary nodes path similarity, firstly the whole network weighted is done topology, and the topology is as follows:

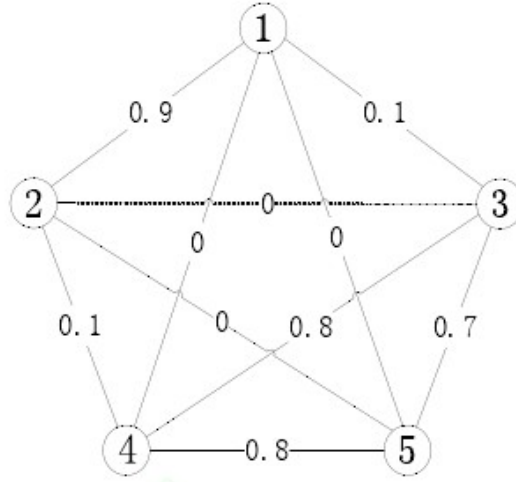


Fig.2: Network weighted topological graph

On the basis of the edge weights setting path similarity, through transform algorithm, it introduces an integral coefficient in the algorithm integral, and the expression is as follows

$$\alpha = \begin{cases} 1 & |e(x)| < \lambda \\ 0 & |e(x)| > \lambda \end{cases} \quad (4)$$

Where λ is the maximum threshold the system allows. In order to optimize the search model of translation lexical frequency, this paper uses dynamic semi Markov algorithm to optimize the path, as shown in Figure 3.

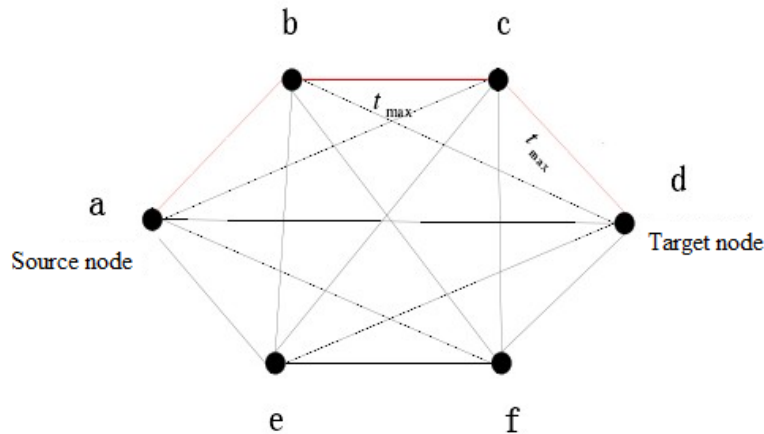


Fig.3: The path search map

As shown in Figure 3, the DTN clustering routing method, based on dynamic semi Markov path search model, not only can predict the meeting possibility of two translation vocabulary, also is able to predict the specific possibility of two words meeting in the same time [8, 9]. So the improved quantitative evaluation model can be written as:

$$s(x) = x_p e(x) + \lambda x_i \sum_{j=0}^x e(j) + x_d [e(x) + e(x-1)] \quad (5)$$

When $\lambda = 0$, it can be written as:

$$\begin{aligned} s(x) &= x_p \left\{ e(x) + \frac{T_D}{T} [e(x) - e(x-1)] \right\} \\ &= x_p \left(1 + \frac{T_D}{T} \right) e(x) - x_p \frac{T_D}{T} e(x-1) \end{aligned} \quad (6)$$

In order to realize the algorithm, this paper uses the VC++ software to design the optimization algorithm of route search, and its main procedure is as follows:

```
void CMovieCaptionMachineDlg::SelectReadTypeA()
{ int flag=0;
  { filepath=m_selectfnameandpath[1].GetAt(fpos);
wfilepath=m_savenameandpath[1].GetAt(wpos);
  AfxMessageBox(wfilepath);
  if(!dic1.Open(filepath,CFile::modeRead))
  { AfxMessageBox("serach"+filepath+"error"); }
  if(!dic2.Open(wfilepath,CFile::modeCreate|CFile::modeWrite))
  { AfxMessageBox("search"+wfilepath+"error"); }
  while(dic1.ReadString(str))
  { if(str.Find("-->")==-1)
    { continue; }
    else
    { while(dic1.ReadString(str)){if(str=="") break;
      else{ dic2.WriteString(str+"\r\n"); } } }
  dic1.Close();}
  dic2.Close();
  AfxMessageBox("Evaluation is completed! ");
  .....
```

Design of English Translation Quantitative Evaluation System based on Corpus

When word frequency statistics, most corpus concordance software can provide two kinds of frequency lists, one is alphabetical order, and the other is according to the frequency size [10, 11]. In order to do quantization evaluation on translation test, this paper selects the second vocabulary alignment method, as shown in Table 1.

Table 1: A co-occurrence context database

R1s	...	R12	R11	A
R2s	...	R22	R21	A
...	A
Rts	...	Rt2	Rt1	A

Table 1 shows database data unit based on A co-occurrence context. The studied word R forms a small block of text with surrounding human determined 2s range of words, and the text size is $t(2s+1)$ [12]. Assuming that the research corpus size, i.e. the total number of shape symbols is n .

Concordance - IR-book 1-6.txt.Concordance			
File	Text	Search	Edit
Headwords	Contexts	View	Tools
Help			
Headword	No.	%	Context...
THE	7142	5.482	a top student. This is not necessa...
TO	3810	2.924	course, studying shouldnt occupy...
OF	3508	2.692	occupy all of the free time on
AND	3428	2.631	or a chair at home or in
A	3132	2.404	you should be able to concentrat...
I	2434	1.868	read it more carefully. As you pre...
IN	2094	1.607	the material, you get some idea of
THAT	1764	1.354	your time in class. Listening to wh...
IT	1436	1.102	Take notes to help you remembe...
HE	1386	1.064	in your textbook. If you know wh...
IS	1372	1.053	you know what the teacher will di...
WAS	1330	1.027	material too. This will help you un...
FOR	1102	0.846	you review your notes and textb...
YOU	1068	0.820	test. . Develop a good attitude ...
AS	910	0.698	what you have learned about a s...
BE	858	0.659	study techniques. Share with the...

Fig.4: Corpus frequency table

As shown in Figure 4, in order to realize the quantitative evaluation of corpus, this paper uses the *Make Full Concordance* function to select part of the translation words as a keyword, and the keyword will be put into a text, and uses a regular expression to search the specific requirements of the keyword, finally the result is as shown in Figure 5.

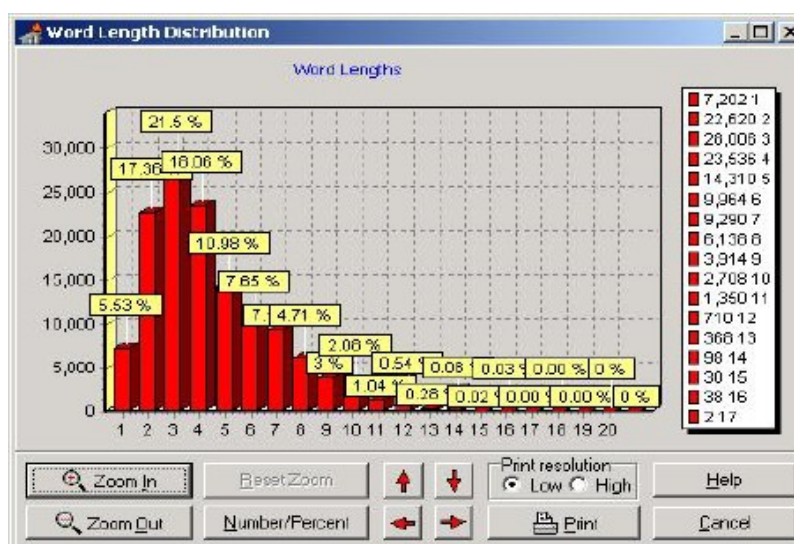


Fig.5: Corpus quantitative computation

Figure 5 shows the test results after the translation is done quantitative evaluation [13, 14]. It includes the class identifier number, the number of shape symbols, pictogram category symbols ratio, average sentence length parameter in the commonly used statistical parameters of corpus. Based on statistical calculation on these parameters, the effect of translation is quantitative represented.

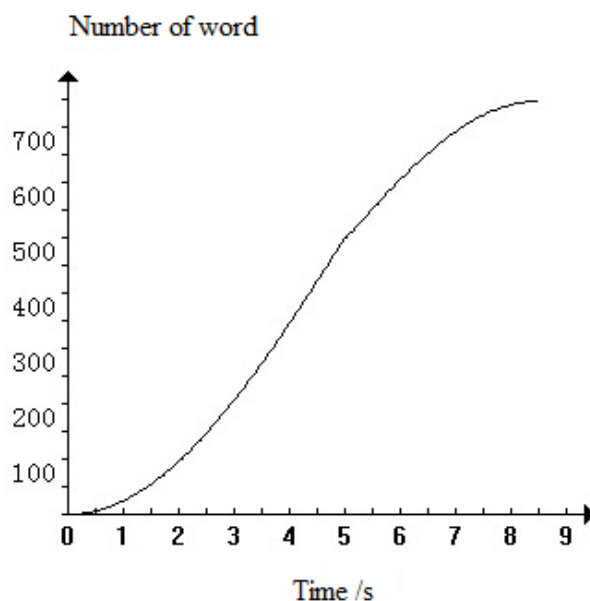


Fig.6: Translation evaluation quantitative results

Figure 6 shows using the quantitative evaluation system to evaluate translation for a period of time, the overall quantitative evaluation result curve is obtained [15]. From the chart it can be seen, at the time of 9s, quantization vocabulary has reached 800, quantitative evaluation is higher efficiency, which meets the need of translation evaluation system development and is a high efficient of computer software quantization method.

Summary

According to the basic principle of dynamic Markov optimization search path, this paper presents a new translation test corpus quantitative evaluation method, and the algorithm uses the DTN clustering routing of lexical encounter probability, introducing an integral coefficient in integral, which greatly improves the quantization range and efficiency. This paper uses VC++ software to design the quantization algorithm, and through quantitative measurement tool, the translation test corpus is evaluated, obtained the number of symbols, and the number of shape symbols, symbols of form type and token ratio, average sentence length of translation evaluation parameters are obtained, which provides a new computer method for the research on measurement translation evaluation method.

References

- [1] J.L. Jiang, Q.F. Wen. Comparative study of N tuelles and translation unit automatic scoring function to the translation. *Modern foreign languages*, 2012, 4(2): 112-115.
- [2] J.Q. Wang, Q.F. Wen. Research on model and the construction of Chinese student mass translation testing machine aided scoring. *Modern foreign languages*, 2013, 2(4): 24-27.
- [3] M.C. Liang, Q.F. Wen. Review and its implications to foreign language automatic scoring. *Foreign composition teaching system*, 2013, 2(5): 67-70.
- [4] J.H. Fan. The application of task-based teaching mode in the university teaching of English listening and speaking. *Academy*, 2014, 2(17): 87-91.
- [5] J.H. Liu. The thinking of university English translation teaching. *Heilongjiang Education (theory and practice)*, 2014, 2(4): 91-93.
- [6] Y.Q. Zhang. Explicit instruction and implicit comparison in teaching university English writing teaching. *Campus English*, 2013, 4(9): 41-45.
- [7] X. Wang. New period of university English teaching reform and development. *Youth literature*, 2013, 4(20): 76-79.

- [8] Q. Zhao. Review of corpus translation studies in foreign countries. Chongqing Jiaotong University (SOCIAL SCIENCE EDITION), 2012, 3(3): 45-48.
- [9] X.M. Luo, Q. Huang, J. Zhang. The culture of university translation teaching reform and translation ability. Foreign language teaching, 2012, 4(1): 13-16.
- [10] M.C. Liang, Q.F. Wen. Review and its implications of foreign composition automatic scoring. Foreign language teaching system, 2012, 4(5): 61-64.
- [11] S.L. Ge, X.X. Chen. Study of foreign automated essay scoring technique. Language teaching, 20012, 4(5): 91-93.
- [12] J. Wang, H.P. Cai, Y. Luo. Study on experience English writing mode based on network platform. Foreign language teachin, 2012, 3(4): 76-78.
- [13] W.J. Zhou. Study on English pan ecological based on the "cloud" resources. Foreign language teaching, 2012, 4(4): 43-46.
- [14] L.F. Yang. The application of mobile learning in university English vocabulary learning. Language teaching, 201, 4(4): 54-58.
- [15] T.L. Tang, Y. Wu. Review of online English writing language automatic evaluation system. Foreign language teaching and research, 2012, 4(4): 32-35.