ICIC Express Letters
Volume **10**, Number **6**, June **2016**

 ISSN 1881-803X
pp. **1353–1358**

DATA MINING OF CHINESE STUDENTS' PERFORMANCE IN BEC INTERNATIONAL EXAMINATION

Lian Jiang[1], Jianping Yu[1,*] and Wenxue Hong[2]

[1]College of Foreign Studies
[2]Institute of Electrical Engineering
Yanshan University
No. 438, West of Hebei Avenue, Qinhuangdao 066004, P. R. China
[*]Corresponding author: yjp@ysu.edu.cn

Received November 2015; accepted February 2016

Abstract. *In this paper, the performance of some Chinese students in BEC Vantage is investigated and analyzed based on the approach of formal concept analysis. The data mining of the performance of 88 Chinese participants is conducted. The performance is assessed in four skills: reading, writing, listening and speaking. Both the structural partial-ordered attribute diagram and the focused diagram are generated, and the current situation and the existing problem are revealed and analyzed based on the diagrams. Finally, some suggestions for improving students' skills are given. The result of the study proves the effectiveness and feasibility of the approach of formal concept analysis in linguistic studies.*
Keywords: Data mining, Students' performance, BEC Vantage examination, Formal concept analysis

1. **Introduction.** Business English Certificate (BEC) is recognized by employers, ministries, government bodies and professional organizations throughout the world. This qualification provides a clear proof that the participants have the English skills to succeed in international business. Since the Business English Certificate is important for the students from non-English speaking countries to work in European countries, there are many Chinese students who prepare and take part in the examination. The examination is held twice a year officially. As this examination is influential and popular, it is significant and necessary to know how Chinese students perform in the examination and what problems exist in their performance, in order to find some pedagogical implications for improving the current English teaching and students' English level. However, no studies have been found focusing on these aspects. Many Chinese scholars have studied students' performance in College English Test (CET) [1,2] and Test for English Majors (TEM) [3], while few studies have been focused on Chinese students' performance in international examinations. Therefore, in this paper, we investigate Chinese students' performance in BEC Vantage examination by the approach of formal concept analysis. Formal concept analysis as a methodology of data analysis and knowledge representation, has potential to be applied to a variety of linguistic problems. Priss [4] presented the linguistic applications of formal concept analysis. Hong et al. [5,6] developed the approach of structural partial-ordered attribute diagram which improved the Hasse diagram in formal concept analysis, and the approach simplifies the generation of the diagram and clarifies the relations between the objects and the attributes. Li and Yu [7] used the approach to visualize the negative cultural transfer in Chinese postgraduates' academic writing. Fu et al. [8] studied the grammatical errors in Chinese postgraduates' academic writing. Zhao and Yu [9] investigated the Chinese postgraduate learners' awareness of genre features in academic

writing by formal concept analysis. Yu et al. [10-14] studied the word sense disambiguation and knowledge discovery of semantically complex words mainly focusing on English modal verbs and prepositions. The previous studies have proven the feasibility of the approach of formal concept analysis in linguistic studies and they laid a good foundation for this data mining of Chinese students' performance in BEC Vantage examination. The innovation and contribution of this study are:

1) New in view. Instead of examining students' performances in College English Test (CET) and Test for English Majors (TEM) which are domestic tests, this study examines students' performance in BEC examination which is an international examination. The study may reveal how Chinese students perform in the international examination which, in turn, may provide some more objective pedagogical implications for current English teaching in China.

2) New in methodology. We use the approach of formal concept analysis, which can visualize the data structure. The diagram generated by the approach can not only show the main tendency of the students' grades distribution, but also show the hierarchical relations among the grades in different skills. This gives us a more general view of the students' performance in English usage.

The significance of the study is that it may provide us a feedback of students' English performance from the international angle, which is very important and valuable for improving the current English teaching in China.

The rest of this paper is organized as follows. Section 2 introduces the approach of formal concept analysis. Section 3 explains the preparation of the data used in this study. Section 4 explains the process of data mining and analysis of the distributed data structure by SPOAD. Section 5 comes to the conclusion.

2. **The Approach of Formal Concept Analysis.** In [5,6], Hong et al. presented the approach of generation of structural partial-ordered attribute diagram. The approach is based on the theory of formal concept analysis. The theory of formal concept analysis was proposed by Wille [15] as a theory of applied mathematics. The idea is based on the philosophical understanding of a concept: a concept is composed of extent and intent. The extent is understood as a set of all the objects of a concept, and the intent is a set of common attributes of all the objects. All the concepts and their relations in generalization and specification may constitute a lattice which may be visualized by a Hasse diagram. The diagram shows the complex network relations between the objects and the attributes. The structural partial-ordered attribute diagram approach proposed by Hong et al. [5,6] improved the Hasse diagram and it can better visualize the hierarchical and conceptual relations between objects and attributes. The advantageous function of the approach makes it suitable for many applications, such as the analysis of data structure, data mining, pattern recognition and knowledge discovery. Therefore, in this paper, we use the approach to investigate the situation of Chinese students' performance in BEC examination in order to find the existing problem and the solutions.

3. **Data Preparation.** The data used in this study are collected from the Statement of Results of 2014 BEC Vantage examination. Altogether 88 Statements of Results are collected for 88 participants from China. They are all college or university students. The Statements of Results reported the performance of the students in reading, writing, listening and speaking. Based on the Statement of Results for each student, we categorize 4 degrees for each skill as the following:

RE – exceptional in reading
RG – good in reading
RB – borderline in reading
RW – weak in reading

LE – exceptional in listening
LG – good in listening
LB – borderline in listening
LW – weak in listening

WE – exceptional in writing
WG – good in writing
WB – borderline in writing
WW – weak in writing
SE – exceptional in speaking
SG – good in speaking
SB – borderline in speaking
SW – weak in speaking

All together 16 degrees are categorized for the 4 skills. Taking the 16 degrees as the attributes and the 88 students as the objects, a formal context can be formed as shown in Table 1.

Table 1. Formal context of students' performance in BEC examination

oj \ ai	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	a15	a16
o1			1			1						1			1	
o2			1			1						1			1	
o3			1			1						1		1		
o4				1				1				1				1
o5				1			1					1			1	
o6		1				1						1		1		
o7			1			1						1			1	
o8			1				1					1			1	
o9				1		1						1		1		
o10			1			1						1				1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
o79			1				1					1				1
o80		1				1					1			1		
o81			1			1						1	1			
o82			1			1					1			1		
o83		1				1						1			1	
o84			1			1						1		1		
o85				1			1					1			1	
o86				1			1					1			1	
o87			1			1					1			1		
o88				1		1						1		1		

a1-RE; a2-RG; a3-RB; a4-RW; a5-WE; a6-WG; a7-WB; a8-WW; a9-LE; a10-LG; a11-LB; a12-LW; a13-SE; a14-SG; a15-SB and a16-SW.

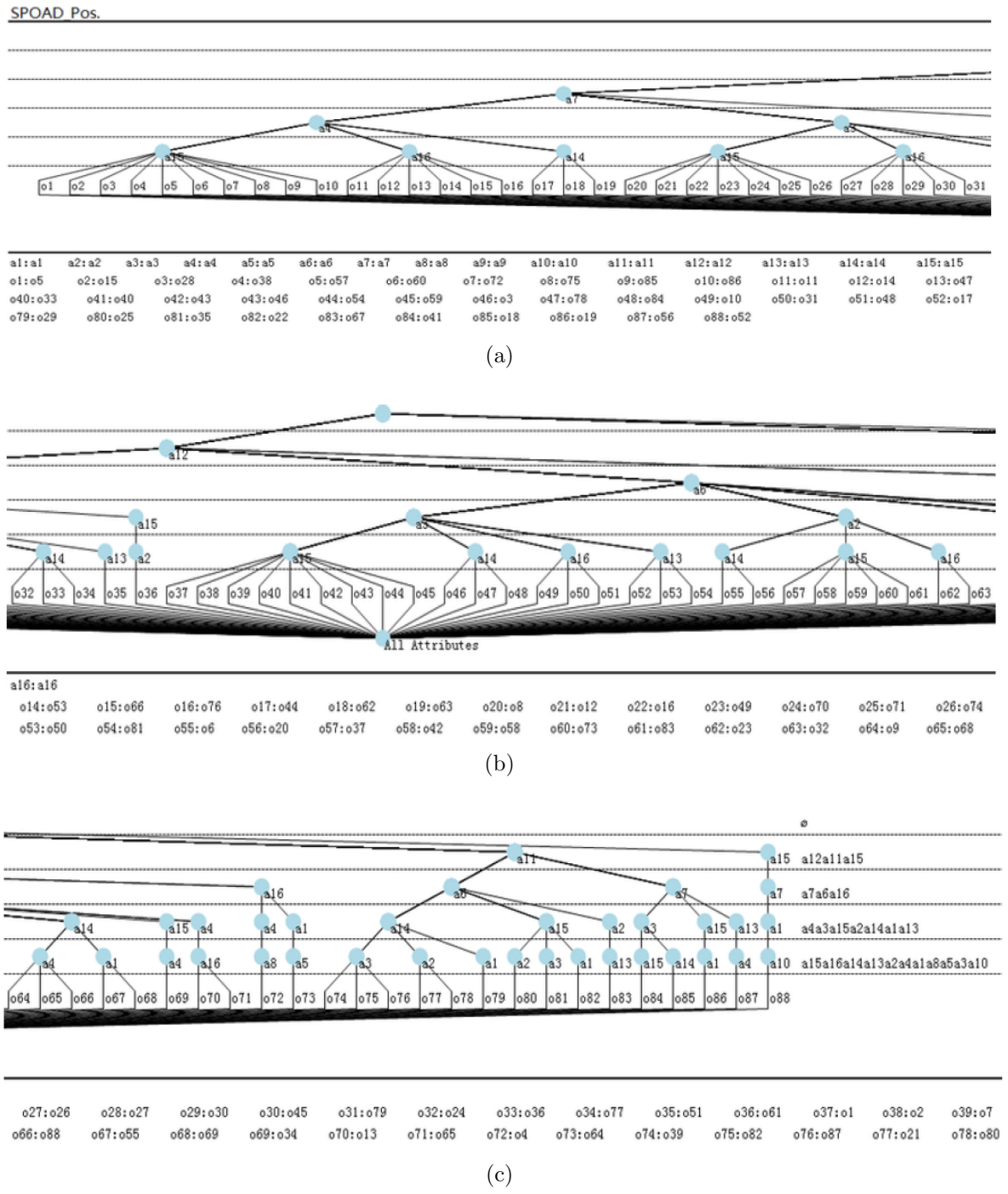
By now the data for data mining are ready and the formal context for investigating the general performance of the students in the BEC Vantage examination is generated.

4. **Data Mining of Chinese Students in BEC Vantage Examination.** Based on the formal context, the structural partial-ordered attribute diagram (SPOAD) of the students' performance in BEC Vantage examination is generated by the SPOAD software [6] as shown in Figure 1, and a focused diagram is also generated as shown in Figure 2.

As can be seen from Figures 1 and 2, the first layer of the SPOAD diagram is composed of 3 nodes: a12 (LW – weak in listening); a11 (LB – borderline in listening) and a15 (borderline in speaking). Since node a15 covers only 1 object, we can say almost all the objects fall in the a12 (83%) and a11 (16%) classes. This implies that almost all the students are not so good in listening skill.

We can also see from Figure 2 that, on the second layer, a7 (borderline in writing) and a6 (good in writing) take the prominence. This indicates that most students fall into the grades of good and borderline in writing, which implies that their writing skill is generally acceptable.

The third and the fourth layers mainly distribute the grades in reading (a1-a4) and speaking (a13-a16). As we can see from the statistic results that 40 objects (45% of the students) are in borderline in reading and 27 (31%) are in weak in reading. In speaking, 39 (44%) objects are in borderline and 20 (23%) objects are in weak, which implies that



(a)

(b)

(c)

FIGURE 1. Structural partial-ordered attribute diagram of students performance in BEC examination

the students are relatively poor in reading and speaking. From the top-down observation of Figure 1, we can also see that there are 41% of the objects (36 of them) who are either in borderline or in weak in all the 4 skills, which means that their general English level is low. Some of the objects are good or exceptional in 1 or 2 skills, but there are few of them. Based on the above data mining we can summarize the students' general situation of the 4 skills in English according to their performance in BEC Vantage examination as the following. Here, the mark ">" represents "better than".

Writing > speaking > reading > listening

As can be seen from the order, the listening skill ranks the last. This may be related to the current teaching situation of the non-English major students. In many universities

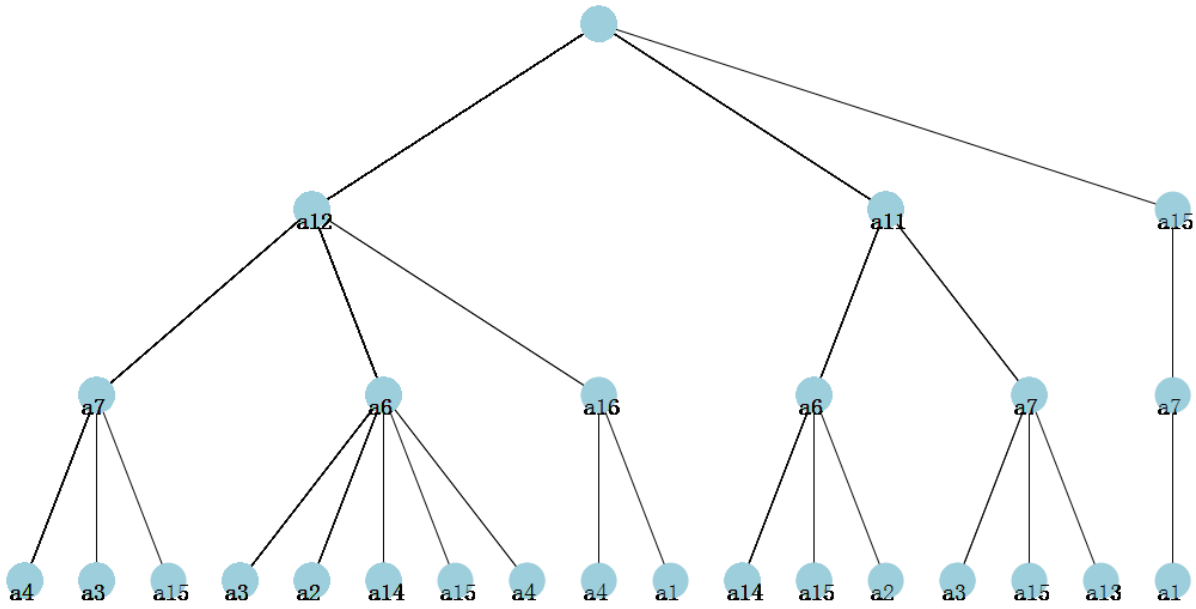


FIGURE 2. Focused diagram of the first 3 layers in Figure 1

or colleges, non-English major students do not have enough training in listening and speaking. Based on this order, we can decide the emphasis in English teaching. We should make especially great effort to improve students' listening skill, such as to create more environments for them to practice listening. We should also work hard in improving students' reading and speaking skills in order to help them to perform better in the international examinations.

5. **Conclusions.** The approach of formal concept analysis is used in the data mining of some Chinese students' performance in BEC Vantage examination. The approach is proven effective in investigating the general situation of the students' English skills. It is found that, generally, the students' listening skill is poor, their writing skill is acceptable, and their speaking and reading skills are relatively poor. This finding is significant. It can help us to decide the emphasis in future English teaching. We should put great emphasis on improving students' listening ability, meanwhile, make effort to improve their abilities of speaking and reading. In further study, we will investigate the reasons for the imbalanced levels of the students' English skills and find out the solutions for them.

Acknowledgment. This work is supported by the National Social Science Foundation under Grant No. 12BYY121, the Social Sciences Foundation of Hebei Province under Grant No. HB15YY002 and the Humanities and Social Sciences Foundation of the Ministry of Education of China under Grant No. 12YJA740096. It is also partially supported by National Natural Science Foundation of China under Grant No. 61074130. The authors gratefully acknowledge the supports.

REFERENCES

[1] L. Quan and Y. Zheng, A study of oral communication strategy used by Chinese non-English majors, *Journal of Guangdong University of Foreign Studies*, vol.37, no.5, pp.67-71, 2014.

[2] J. Cai, Reorientation and study of CET in China, *Computer Assisted Foreign Language Education*, vol.32, no.5, pp.3-10, 2011.

[3] Q. Xu, The washback effect of the TEM-8, *Foreign Language World*, vol.32, no.3, pp.21-31, 2012.

[4] U. Priss, Linguistic application of formal concept analysis, in *Formal Concept Analysis: Foundations and Applications*, B. Ganter, G. Stumme and R. Wille (eds.), Germany, Springer-Verlag Berlin Heidelberg, 2005.

[5] W. Hong, S. Li, J. Yu and J. Song, A new approach of structural partial-ordered attribute diagram, *ICIC Express Letters, Part B: Applications*, vol.3, no.4, pp.823-830, 2012.
[6] W. Hong, S. Li, J. Yu, J. Song, X. Wang and J. Liu, A new approach of generation of structured partial ordered attribute diagram based on covering, *ICIC Express Letters, Part B: Applications*, vol.6, no.4, pp.1049-1054, 2015.
[7] H. Li and J. Yu, Visualization analysis of negative cultural transfer in academic writing based on structural partial-ordered attribute diagram approach, *ICIC Express Letters*, vol.7, no.12, pp.3229-3235, 2013.
[8] J. Fu, J. Yu and H. Liu, Data mining of grammatical errors in Chinese postgraduates' academic writing, *ICIC Express Letters*, vol.7, no.12, pp.3221-3227, 2013.
[9] S. Zhao and J. Yu, Investigation of postgraduate learners' awareness of genre features in academic writing by formal concept analysis, *ICIC Express Letters, Part B: Applications*, vol.6, no.5, pp.1295-1301, 2015.
[10] J. Yu, C. Li, W. Hong, S. Li and D. Mei, A new approach of rule extraction for word sense disambiguation by features of attributes, *Applied Soft Computing*, vol.27, no.1, pp.411-419, 2015.
[11] J. Yu, N. Chen, W. Hong, S. Li and T. Zhang, Interactive relations between semantic and syntactic features in word sense disambiguation of semantically complex words, *International Journal of Innovative Computing, Information and Control*, vol.9, no.9, pp.3627-3638, 2013.
[12] J. Yu, W. Hong, S. Li, T. Zhang and J. Song, A new approach of word sense disambiguation and knowledge discovery of English modal verbs by formal concept analysis, *International Journal of Innovative Computing, Information and Control*, vol.9, no.3, pp.1189-1200, 2013.
[13] J. Yu, N. Chen, R. Sun, W. Hong and S. Li, Word sense disambiguation and knowledge discovery of English modal verb *can*, *ICIC Express Letters*, vol.7, no.2, pp.577-582, 2013.
[14] J. Yu, N. Chen, W. Hong and Y. Shi, Solution of semantic mergers of English modal verbs, *Information*, vol.16, no.7, pp.4991-5006, 2013.
[15] R. Wille, Restructuring lattice theory: An approach based on hierarchies of concepts, in *Ordered Sets*, I. Rival (ed.), Dordrecht, D. Reidel Publishing Company, 1982.