

THE INTERACTION BETWEEN PRIVATE UNIVERSITY STUDENTS AND INDUSTRY: COMPETENCE BUILDING FROM SUMMER PRACTICE

Yo-Ping Huang¹ and Chi-Peng Ouyang²

Abstract — This paper bases on a survey from computer science-major students to investigate how to build the competence from summer practice and how to enhance the interaction between private university students and industry. We distinguish the survey for two different groups. One group includes the freshmen and sophomores who have not participated in the summer practice. For students in this group we prefer to know what they expect to learn from the summer practice and how well they think they have prepared for the program. The other group consists of students such as junior, senior and graduate students who have completed the practice. Survey from this group allows us to consider the students' reaction about the program and especially to understand whether they gradually realize the importance of practice to the future career. We also investigate how to associate the relationship in the questionnaire based on the data mining technique. Therefore, in this paper not only conventional statistical methods are applied to analyzing the questionnaire, but also data mining models are exploited to find the association rules in the questionnaire. Finally, students' reaction and the results obtained from data mining models help us to improve the whole program such that the interaction between university and industry can be enhanced.

Index Terms — Association rules, data mining, summer practice, questionnaire, private university

INTRODUCTION

There are more than 135 universities and colleges in Taiwan and most of them are private ones. Thus, private university graduates play a very important role in the Taiwan's industry. Our institution is long-term sponsored by Tatung company. To establish a unique education-industry-integration program in the world, any undergraduate student in our university has to take at least eight-week practical training during the summer vacation for the past thirty years. At the beginning, except the first summer for compulsory military training, students in every department of our institution have to take the summer practice during the second and the third summer vacations. Later, due to the change of educational environment, students hoped that they could spend the third summer vacation to prepare the entrance examinations of graduate schools. We thus shorten the program from two to one eight-week practice. This program is designed for

students to gain more hands-on experience and to put theory into practice. We allow students to choose their own summer jobs. However, the job should be related to the student's major. Besides, one of our faculty will visit the student's working place at least once a month. Of course, most students are reluctant to take the program because the program must be taken during the summer vacation. So, it is a required program. Every student must complete the practice before graduation. We believe if a college student has some working experience before graduation, a company can shorten the training period for an engineer.

The program of summer practice is activated during the summer vacation. Students must decide their working places by the middle of May. The program starts from July 1 and lasts for eight weeks. To understand how the students treat the summer practice and how to improve the program, we design our own questionnaire. Based on the survey, we can realize to what extent the students evaluate the working contents and working environment. We will summarize the results from the survey. To better understand the embedded relationship between students' answers, we exploit the data mining technique [1-6] to find the associations between one or more questions with the others. We will present the association rules from the survey.

QUESTIONNAIRE AND DATA COLLECTION

Since the program of summer practice has been activated for more than thirty years and the information technology becomes more and more important in modern industry, to better evaluate the program from the viewpoints of students we issued the survey last year. To prevent from bothering students too much time, the questionnaire can be easily answered in ten minutes. Two different kinds of questionnaire are designed for our students. Despite the required data to fill, freshmen and sophomores are asked to answer only 15 questions. The questions relate to the working contents, such as type of work, safety problem, and salary. Since students in this group have not participated the program, their answers reflect how important they think those items are in the summer practice.

For those who have completed the practice, i.e., junior, senior, and graduate students, we issue two categories of questionnaire in the survey. One part, including 15 questions, regards the working contents and knowledge accumulation, such as working load and related knowledge learning from

¹ Yo-Ping Huang, Department of Computer Science and Engineering, Tatung University, Taipei, Taiwan 10451 yphuang@cse.ttu.edu.tw

² Chi-Peng Ouyang, Department of Computer Science and Engineering, Tatung University, Taipei, Taiwan 10451 ouyang@mail.ee.sitc.edu.tw

the practice. The other part, also having 15 questions, concerns the working environment, such as working duration, comfort of working place, and safety problems. Each question is divided into five satisfactory levels, ranging from 1 to 5, which correspond from the least to the most satisfaction degrees. Based on the answered questionnaire, we try to collect some useful information to further improve the program.

Since it is the first time to evaluate the program, we do not have the history data for comparison. Although the data are collected from the computer science-major students only and the sample space, especially from the graduate students, may not have strong representative, we believe the questionnaire can faithfully reflect the students opinions about what they have faced in the past or what they will face in the future. Our sample space consists of 86, 75, 22, 19, 19, 18, and 9 answered questionnaire from freshmen, sophomores, juniors practicing in Tatung company, juniors practicing in other companies, seniors practicing in Tatung company, seniors practicing in other companies, and graduate students. Note that it is unnecessary for graduate students to take the summer practice. This is why we have only a few samples from graduate students.

ANALYZING THE SURVEY BY DATA MINING TECHNIQUE

Data mining is a new area for database research in discovering the interesting rules from large collections of data. The process of data mining can be divided into two steps: (1) find all frequent item sets; and (2) generate strong association rules from the frequent item sets [4]. Two approaches, i.e., minimum support and confidence, are commonly predefined to find the frequent item sets. The support value from item-A to item-B is defined as follows:

$$\text{support}(A \rightarrow B) = \frac{\#_tuples_containing_both_A_and_B}{\#_tuples_containing_A}$$

Similarly, we can define the confidence value from grouping item-A and item-B as follows:

$$\text{confidenc}(A \Rightarrow B) = \frac{\#_tuples_containing_both_A_and_B}{\#_tuples_containing_A}$$

Based on the defined support and confidence values, we can find whether two or more items appear frequent enough in the transaction databases. Depending on the available size of the transaction records, normally both minimum support and minimum confidence can be predefined by designers. The concept of mining the association rules in databases is applied to investigating how to associate the answered questions in the questionnaire. Therefore, in this paper not only conventional statistic methods are applied to analyzing the questionnaire, but also data mining models are exploited to find the association rules in the questionnaire. We will make some suggestions on how to build the competence from practice. How to enhance the interaction between university and industry will be discussed in detail in our paper.

When students answered the questionnaire it was assumed that there was no mutual interference between their answers. Also assumed is that the order of questions is not arranged on purpose. The results from the answers are not only used to summarize how satisfactory in average students with a question in conventional research. Here, we use the data mining method to analyze the relationship between the questions. Based on the analysis, we intend to understand how students answer the questions and the embedded satisfactory degree of one question with the others from the survey.

Based on each student's answers to the questionnaire, we have to transform the answers into a transaction table as commonly used in data mining. Each question has five choices, ranking from 1 to 5 that corresponds to the satisfaction level from low to high. Thus, for example, a student gave the first problem a score 3, the second a score 4 and the third a score 2. This answer sheet will be transformed into a transaction as {A3, B4, C2}. Since either the working content or working environment has fifteen questions in the questionnaire, a student's answers to each part of questionnaire is transformed into a transaction record with fifteen items. Table 1 gives a simple example of how the answers are transformed into a transaction record.

RESULTS AND DISCUSSION

Fig. 1 plots the statistical results from what freshmen and sophomores expect in the summer practice. As mentioned before, we intend to know how the students in this group think about the summer practice. The average scores for freshmen and sophomores are 4.19 and 4.1, respectively. Such high scores imply what their intuitive reaction to the importance of working contents in the future. Based on the survey, we found that some students are reluctant to take the summer practice. As a result, question number 9 has the lowest score as shown in Fig. 1. However, if they must take it, they feel that both the working contents and working environment are very important. Fig. 2 depicts the statistical results about working contents from students who have completed the summer practice in Tatung company. The average scores are 3.29, 2.82, and 3.37 for juniors, seniors, and graduate students, respectively. Based on the results, questions 9 and 10 received lower scores. This implies that students thought what they were assigned to work did not satisfy what they expected. This deserves for further investigation why students did not satisfy their summer job. One reason may happen that students thought what they have learnt in the past two years is not adequate for their summer job. Statistical results about working environment from students who have completed the summer practice in Tatung company are plotted in Fig. 3. The average scores are 3.52, 3.30, and 3.13 for juniors, seniors, and graduate students, respectively. From the result, it is quite clear that question number 8, which is the salary item, had the lowest score. It means that the most unsatisfied thing is their salary. We

believe that most students expect more from the company because they think they make the same contribution to the company as other employees do. Therefore, they feel that they are cheap laborers. Actually, it is an interesting outcome. Nowadays, some students do not consider what they can do for the company; instead, they just care how much they can be paid. From the educational viewpoint, we have to teach students from considering the aspect of what they can learn and envisioning the future, not just caring the salary problem, from the summer practice. However, many students realize that the practice has much help to find their first job.

For comparison, Fig. 4 is the statistical results about working contents from students who have completed the summer practice in other companies. The average scores are 3.31, 3.55, and 3.23 for juniors, seniors, and graduate students, respectively. Similar to those who worked for the Tatung company, students gave both questions 9 and 10 lower scores than the rest. This also indicates that students thought that they have not well prepared for the summer job. Fig. 5 gives the statistical results about working environment from students who have completed the summer practice in other companies. The average scores are 3.36, 3.57, and 3.33 for juniors, seniors, and graduate students, respectively. Similarly, students did not satisfy their salaries. One more interesting result from the survey shows that students did not satisfy the extra-activity held by their working companies.

Based on the statistical data plotted in Fig. 1 to Fig. 5, we obtained the following interesting results:

- (1) Juniors who have completed summer practice in Tatung company expressed higher satisfaction in both working contents and environment than seniors. On the contrary, those who worked in other companies, seniors were more satisfied than juniors. This indicates that juniors appreciated the working contents and environment provided by Tatung company. However, seniors preferred to work for other companies.
- (2) For both working contents and environment, seniors obtained higher satisfaction from other companies.
- (3) Juniors thought that Tatung company provided nearly the same working contents as other companies. But, Tatung company had a much better working environment.
- (4) If the economical growth rate is good enough, our students can easily find companies for summer practice. The sponsored company can alleviate the headache of accommodating our institution students for practice. If the economical recession occurs, it is more difficult for students to find a company for summer practice, the sponsored company will face the pressure to provide more vacancies for summer practice.

Since we have many samples from freshmen and sophomores questionnaire, we set the minimum support and confidence to 20 and 80%, respectively. The large itemsets derived from the data mining technique based on the freshmen and sophomores questionnaire are plotted in Fig. 6. Fig. 7 depicts the large itemsets from questionnaire of

working contents answered by students having completed the summer practice. Note that the sample space for students having completed the practice is not large enough, the minimum support and confidence are set to 5 and 60%, respectively. Large itemsets of working environment are demonstrated in Fig. 8. Association rules about working contents and environment from juniors practicing in Tatung company are shown in Fig. 9 and Fig. 10, respectively.

To illustrate how the association rules benefit the analysis of freshman questionnaire, we set minimum support and confidence to 35 and 80%, respectively. Based on such thresholds, the large 1-itemsets are plotted in Fig. 11. In Fig. 11, we found that the sum of support values for A4 (the score of the first question being 4) and A5 (the score of the first question being 5) is 77. This means that the atmosphere in working place is very important. Similarly, the sum of J4 and J5 is 80. This indicates that the colleague's interaction is also very important. Safety is most concerned problem; therefore, F5 has 68 records in our data. If we continue the same process to find the large itemsets, we have the results shown in Fig. 12. From the large itemsets, we can derive some interesting combinations such as {A5, F5, K5, L5}, {A5, J5, K5, L5}, {C5, F5, K5, L5}, {D5, F5, K5, L5}, {F5, J5, K5, L5}. Based on the combinations we found that many students treat the atmosphere of working place, safety problem, their specialties, and their interests to the work as very important factors.

By applying the same minimum support and confidence to analyzing the sophomore questionnaire we obtain the results as shown in Fig. 13. Some interesting associations are {A4, B4, G4}, {A4, B4, O4}, {A4, C4, J4}, {A4, C4, O4}, {A4, G4, O4}, and {A4, J4, M4}. These combinations reflect that most sophomores treat atmosphere of working place, working duration, type of work, and distance to the working place as important factors. The results from working contents of seniors practicing in other companies indicate that the working contents should have closer relationships with their specialties, private time, and type of work.

CONCLUSION

Based on a survey from computer science-major students, we investigated how to build the competence from summer practice and how to enhance the interaction between private university students and industry. We distinguish the survey for two different groups. One group includes the freshmen and sophomores who have not participated in the summer practice. For students who have not involved in summer practice, we gained some useful messages from the survey. For example, some students are reluctant to take the summer practice. This is why the item in the questionnaire received the lowest score as shown in Fig. 1. However, if they must take it, they realize that both the working contents and working environment are very important. The other group consists of students such as junior, senior and graduate

students who have completed the practice. Survey from this group allows us to consider the students' reaction about the program and especially to understand whether they gradually realize the importance of practice to the future career. We also investigate how to associate the relationship in the questionnaire based on the data mining technique. Therefore, in this paper not only conventional statistic methods are applied to analyzing the questionnaire, but also data mining models are exploited to find the association rules in the questionnaire. Also, students' suggestions help teachers modify their teaching materials in the classrooms. This in turn can improve the whole program such that the interaction between university and industry can be enhanced.

ACKNOWLEDGMENT

This work was supported in part by the National Science Council, Taiwan, R.O.C., under Grant NSC89-2213-E-036-050 and in part by Tatung University, Taipei, Taiwan under Grant B89-24.

REFERENCES

- [1] Agrawal, R. and Srikant, R., "Fast algorithms for mining association rules in large databases," *Proc. 20th Int. Conf. Very Large Data Bases*, Sept. 1994, pp.478-499.
- [2] Chen, M.S, Han, J., and Yu, P.S, "Data mining: an overview from a database perspective," *IEEE Trans. on Knowledge and Data Engineering*, Vol. 8, No. 6, Dec. 1996, pp.866-883.
- [3] Han, J. and Fu, Y., "Mining multiple-level association rules in large databases," *IEEE Trans. Knowledge and Data Engineering*, Vol. 11, No. 5, Sept. 1999, pp.788-805.
- [4] Han, J. and Kamber, M., *Data mining: concepts and techniques*, Morgan Kaufmann Publishers, San Francisco, CA, 2001.
- [5] Srikant, R. and Agrawal, R., "Mining generalized association rules," *Proc. of the 21st Int. Conf. Very Large Data Bases*, Zurich, Switzerland, 1995.
- [6] Wur, S.Y. and Leu, Y.H., "An effective Boolean algorithm for mining rules in large database," *Proc. 6th Int. Conf. on Database Systems for Advanced Applications* 1999, pp.179-186.

TABLE 1

A STUDENT'S ANSWERS WAS TRANSFORMED INTO A TRANSACTION RECORD

Problem number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Satisfaction level	3	4	2	5	4	4	4	3	5	4	2	2	4	3	3
Transaction item	A3	B4	C2	D5	E4	F4	G4	H3	I5	J4	K2	L2	M4	N3	O3

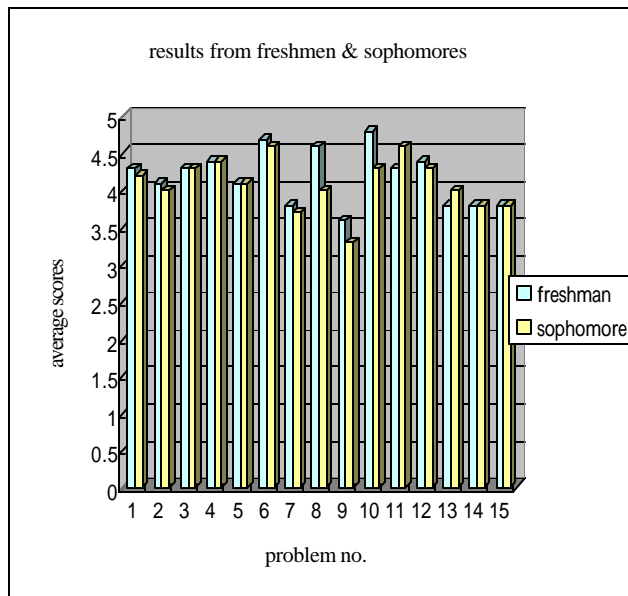


FIGURE 1
STATISTICAL RESULTS FROM WHAT FRESHMAN AND SOPHOMORE EXPECTATION IN THE SUMMER PRACTICE

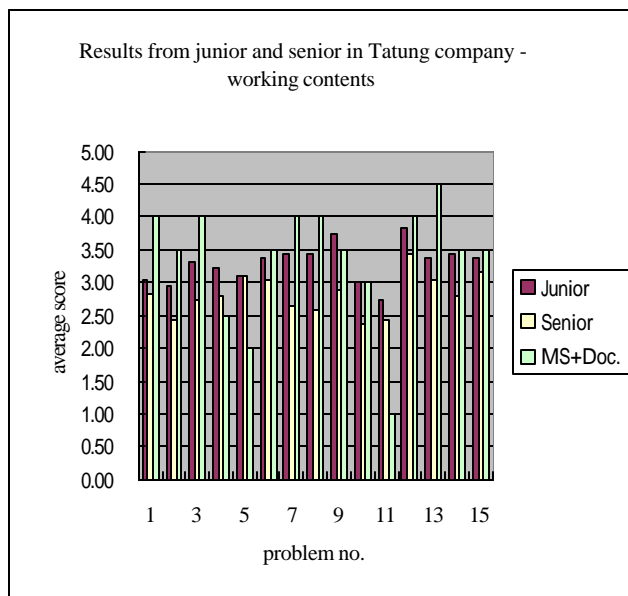


FIGURE 2
STATISTICAL RESULTS ABOUT WORKING CONTENTS FROM STUDENTS HAVING COMPLETED THE SUMMER PRACTICE IN TATUNG COMPANY

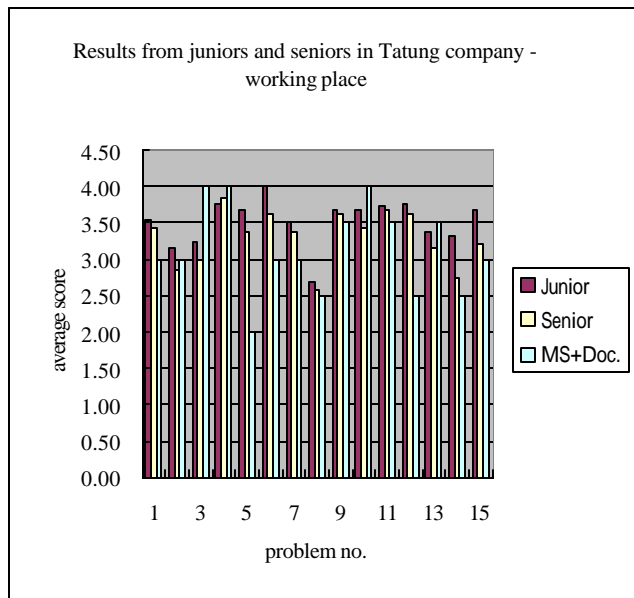


FIGURE 3

STATISTICAL RESULTS ABOUT WORKING ENVIRONMENT FROM STUDENTS HAVING COMPLETED THE SUMMER PRACTICE IN TATUNG COMPANY

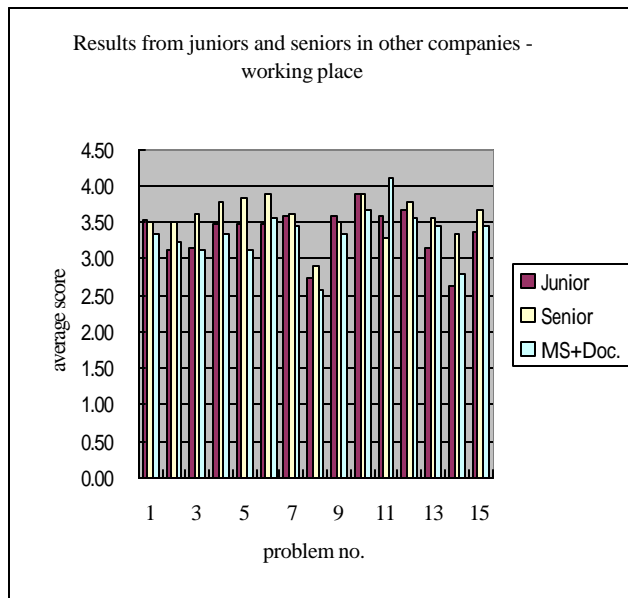


FIGURE 5

STATISTICAL RESULTS ABOUT WORKING ENVIRONMENT FROM STUDENTS HAVING COMPLETED THE SUMMER PRACTICE IN OTHER COMPANIES

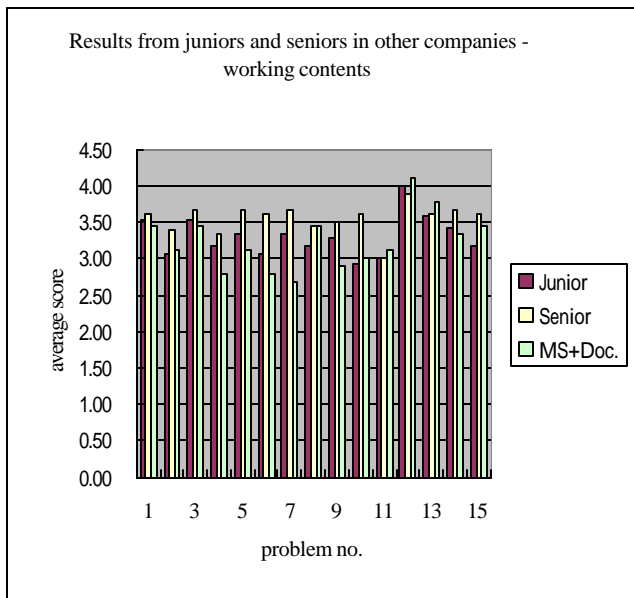


FIGURE 4

STATISTICAL RESULTS ABOUT WORKING CONTENTS FROM STUDENTS HAVING COMPLETED THE SUMMER PRACTICE IN OTHER COMPANIES

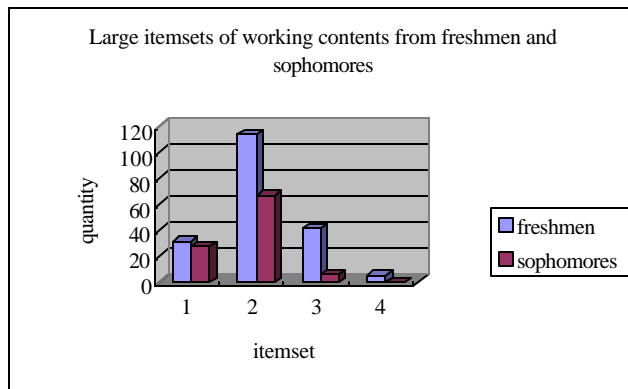


FIGURE 6

LARGE ITEMSETS FROM FRESHMEN AND SOPHOMORES (MINIMUM SUPPORT=20, CONFIDENCE=80%)

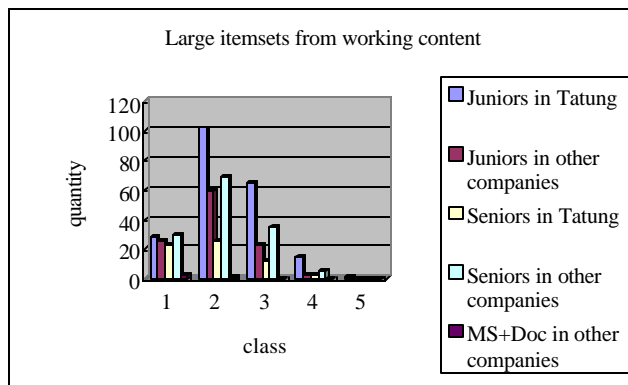


FIGURE 7

LARGE ITEMSETS OF WORKING CONTENTS FROM STUDENTS HAVING COMPLETED SUMMER PRACTICE (MIN. SUPPORT=5, CONFIDENCE=60%)

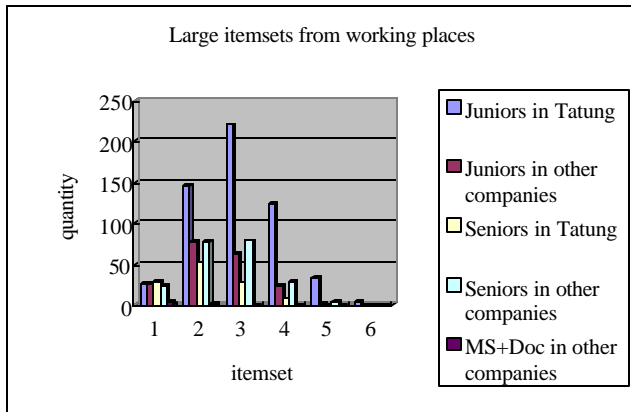


FIGURE 8

LARGE ITEMSETS OF WORKING CONTENTS FROM STUDENTS HAVING COMPLETED SUMMER PRACTICE (MIN. SUPPORT=5, CONFIDENCE=60%)

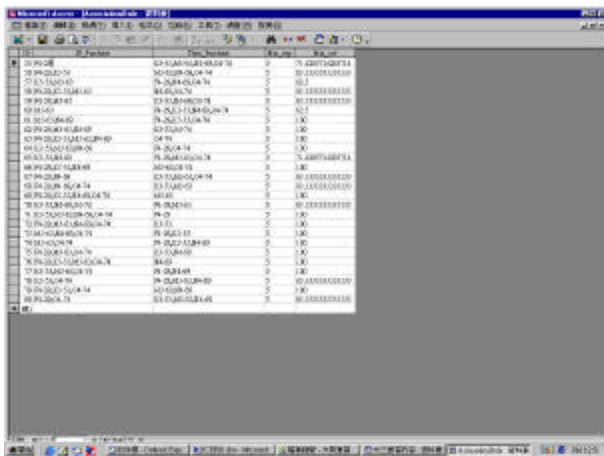


FIGURE 9

ASSOCIATION RULES ABOUT WORKING CONTENTS FROM JUNIORS HAVING COMPLETED THE SUMMER PRACTICE IN TATUNG COMPANY

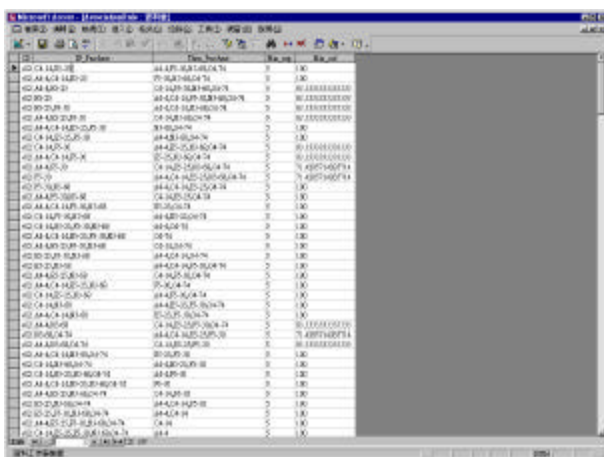


FIGURE 10

ASSOCIATION RULES ABOUT WORKING ENVIRONMENT FROM JUNIORS HAVING COMPLETED THE SUMMER PRACTICE IN TATUNG COMPANY

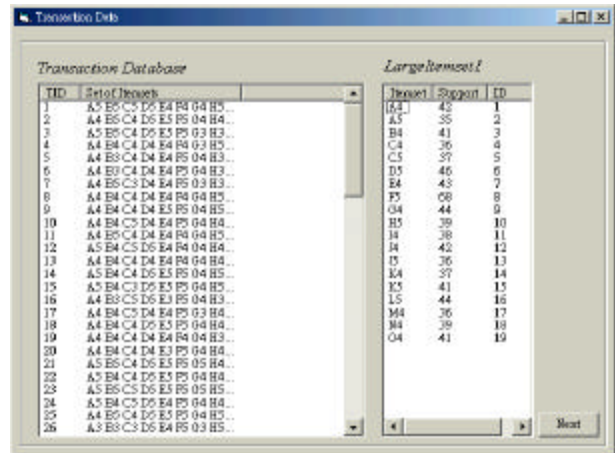


FIGURE 11

LARGE 1-ITEMSETS FROM FRESHMEN QUESTIONNAIRE WHEN MINIMUM SUPPORT=35 AND CONFIDENCE=80%

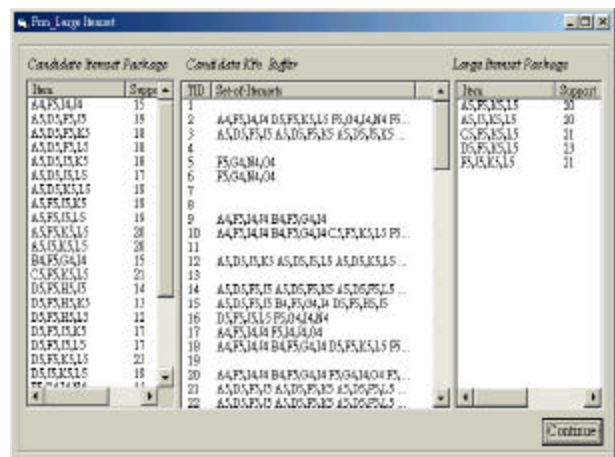


FIGURE 12

LARGE ITEMSETS FROM FRESHMEN QUESTIONNAIRE WHEN MINIMUM SUPPORT=20

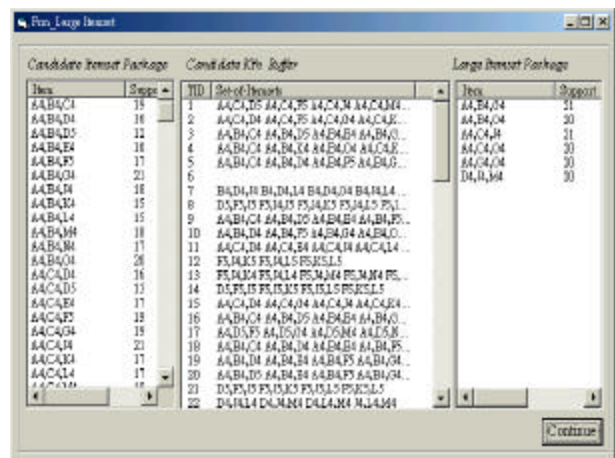


FIGURE 13

LARGE ITEMSETS FROM SOPHOMORE QUESTIONNAIRE WHEN MINIMUM SUPPORT=20