TOPSIS Assisted Selections of the Best Suited Universities for College Applications in Mainland China

Shan Lu and Jie Wang
Department of Computer Science
University of Massachusetts
Lowell, MA 01854, USA
Email: {slu, wang}@cs.uml.edu

***Abstract*—College admissions in mainland China depend mainly on the scores of the standardized annual examination called Gaokao. Students submit a common application to their provincial Gaokao office, on which they are allowed to list a fixed and small number of universities and majors they intend to study. The admission process in a province follows one of the following three admission models: parallel, gradient, and a combination of both. No matter what admission models are used, there is always a possibility that an applicant could end up being rejected by every university listed in the application, even though the applicant could have been accepted by a university not in the list. This process presents a challenge for students to figure out how to select universities to apply so that they can be admitted by a university and major that match their abilities and interests. To many students, and their parents, this is a difficult decision to make and their experience is unpleasant. To help reduce this agony, we present a new approach of applying Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) to generate a personalized selection of the best suited universities and majors that match a student's Gaokao score and meet a list of criteria. We then present case studies to demonstrate the effectiveness of this approach.**

Keywords–TOPSIS; weighted criteria; multi-attribute decision making; recommendation system

I. INTRODUCTION

Gaokao is the standardized annual examination for college entrance in mainland China, which takes place in early June every year. It is mandatory for admission into four-year colleges and universities. Each year during the last 10 years, there were over 9 (sometimes over 10) million high-school graduating students participating in Gaokao according to Chinese Education Online (http://gaokao.eol.cn/). Students must choose one of the two types of exams; namely the Li-Ke exam (meaning the science exam) and the Wen-Ke exam (meaning the liberal-arts exam). The Li-Ke exam must be taken by students for entering the disciplines of science, engineering, agriculture, and medicine; and the Wen-Ke exam must be taken by students for entering the disciplines of arts, humanity, education, and management. Students find out their Gaokao scores in late June, followed by the application and admission process that would last from one to two months.

Students in the same province must complete a common application form either prior to taking Gaokao or after, depending on the province or municipality in which they have official residency. A municipality is a very large city directly under the central government and so is treated as a province.

When the admission process starts, each application is released to a university selected from the universities listed in the application according to certain rules. That is, not all universities listed in an application can see the application simultaneously. If the university that receives the application rejects it, then the application is released to the next university in the list. However, this university may have already filled out its admission quota by then. Recall that students are only allowed to list a fixed and small number of universities and majors on their application forms. Thus, in addition to obtaining good Gaokao scores, students need to figure out which universities and majors they should apply to maximize their chances of acceptance while meeting their education goals. Note that each student can only be accepted by one university, or not accepted at all. This process is fundamentally different from the US where students may apply to universities directly as many as they would like and receive acceptance from multiple universities.

In mainland China, universities are officially categorized into three tiers based on the qualities and the number of programs they offer. National universities are the first tier, provincial universities are the second tier, and regional universities are the third tier. There are about 120 first-tier, 750 second-tier, and 1,550 third-tier universities. Applicants need to specify their preferences according to the official categorization.

Each university sets an admission quota for each province each year, which is broken down into majors. Each province sets its own rules on how universities access applications. These rules may be grouped into three admission models: parallel admission, gradient admission, and hybrid admission.

Mismatched admissions is a common problem. That is, if applicants apply to universities inappropriately, they may end up receiving no offer or an offer that is a poor match of their abilities or interests even though they could have been accepted by a university that presents a better match.

Both Gaokao and college admission are conducted only once a year. Once a student is admitted by a university to a particular major, it is almost impossible to change majors after admission. Thus, it is important to identify best-suited universities and majors to apply to, and students must complete their applications in a short period of time after Gaokao. To help reduce this agony, Lu, Zhang, and Wang [1] presented an automated system using General Morphological Analysis (GMA), based on a proprietary mathematical model for predicting admission scores for each major of each university in the current year, to analyze a large volume of data from previous years of Gaokao and help students make informed decisions based on their Gaokao scores and interests.

GMA is a method for identifying and investigating the total set of configurations contained in multi-dimensional, non-quantifiable problem complexes. It is "totality research" that attempts to derive all the solutions of any given problem

 ISBN: 978-1-61208-570-8

in an unbiased manner. For a given admission model, we use GMA to identify suitable universities and majors for a student based on the student's Gaokao score and interests [1]. However, GMA does not allow students to specify weights over each interest they are interested to compute the best-suited university and major for the students. We present in this paper a method using Technique for Order Preference by Similarity to an Ideal Solution(TOPSIS) to fill this void. In particular, based on the universities and majors recommended by GMA for a student, we allow the student to specify the weight for each attribute and then use TOPSIS to compute the best-suited university and major. TOPSIS is a method for multi-criteria decision making, which was originally developed by Hwang and Yoon in 1981 [2] with further developments by Yoon in 1987 [3] and Hwang, Lai and Liu in 1993 [4]. TOPSIS is based on the concept that the chosen alternative should have the shortest geometric distance from the positive ideal solution [5] and the longest geometric distance from the negative ideal solution [5]. It compensates aggregation and compares a set of alternatives by identifying weights for each criterion, normalising scores for each criterion, and calculating the geometric distance between each alternative and the ideal alternative. To the best of our knowledge, our study is the first to use TOPSIS to produce Gaokao recommendations.

The remainder of this paper is organized as follows. In Section II, we describe how we identify all the suited universities and majors using GMA. In Section III, we describe TOPSIS and use it to generate the best-suited universities and majors among the suitable solutions found using GMA. In Section IV, we present two case studies and conclude the paper in Section V.

II. Identify All the Suited Universities and Majors using GMA

Lu, Zhang, and Wang [1] used computer assisted GMA to compute all possible combinations of majors and universities that are for students at various degrees according to their Gaokao scores and their interests under the admission model in their province. Let LU denote the set of labels for university slots in an application form, J the number of universities a student is allowed to specify for each tier, and K the number of majors a student is allowed to specify for a university. For example, in application forms for students living in the Fujian province, $J = 4$ and $K = 6$.

For convenience, in what follows we will use Alice to represent a student and XYU to represent a university. Alice enters her Gaokao score and other information to obtain recommendations for X-universities, where $X \in LU$, and we call them X-recommendations. For example, when $J = 4$, we have A-, B-, C-, and D-recommendations, respectively. Universities listed in A-recommendations are competitive for Alice, but Alice still has a chance to be accepted. Universities listed in B-recommendations would present a good match of Alice's ability and interests, which means that Alice would have a good chance to be accepted. Universities listed in C-recommendations are conservative choices for Alice, which means that Alice would have a very good chance to be accepted. Universities listed in D-recommendations are the safest choices for Alice, which means that Alice would have a near 100% chance to be accepted.

The baseline GMA setting consists of 14 parameters, divided evenly into two groups, one group for students and one group for universities. Parameters in the student group are (1) Alice's Gaokao score; (2) the type of the exam that Alice takes; (3) the tier of the universities that Alice wants to attend; (4) locations where Alice wants to go to for college; (5) locations Alice does not want to go to for college; (6) Majors that Alice wants to study; (7) majors that Alice does not want to study.

Parameters in the university group are (1) the lowest admission score of XYU in the past year; (2) the type of majors that XYU offers; (3) XYU's official tier; (4) XYU's location; (5) the majors that XYU offers, including (when possible) the lowest, medium, and highest admission scores for each major in the past year, and the total number of expected enrollment for a major for the current year; (6) the ranking of XYU (the first-tier universities are ranked from 1 to 5 with 1 being the highest one; the second-tier universities are ranked from 6 to 7; and the third-tier universities have one rank of 8); (7) the total enrollment of XYU for the current year (When this number is not known, it uses last year's enrollment number).

III. Identifying the Best-Suited Universities and Majors Using TOPSIS

TOPSIS is a multi-objective decision making method over a hierarchical structure of alternatives with multiple criteria. At the top level is the optimization goal. The next level consists of a list of criteria, which may be decomposed further into several levers of sub-criteria. The bottom level consists of a list of alternatives to be measured against each criterion. The criteria can relate to any aspect of the decision making, tangible or intangible, carefully measured or roughly estimated, well-defined or poorly understood.

Based on the concept that the chosen alternative should have the shortest geometric distance from the positive ideal solution and the longest geometric distance from the negative ideal solution, TOPSIS compensates aggregation through comparisons of alternatives by identifying weights for each criterion, normalizing the scores for each criterion and calculating the geometric distance between each alternative and the ideal alternative, which is the best score in each criterion. TOPSIS allows tradeoffs between criteria, where a poor result in one criterion can be negated by a good result in another criterion. This provides a more realistic form of modeling than non-compensatory methods, which include or exclude alternative solutions based on hard cutoffs [6].

To use TOPSIS, we need to ensure that the criteria of attributes are either monotonically increasing or monotonically decreasing and use normalisation to compensate incongruous dimensions in multi-criteria problems [7][8].

A. Criteria

After using GMA to obtain all suitable combinations of majors and universities for Alice, we create an evaluation matrix consisting of m alternatives and n criteria, where the alternatives are universities (XYU) and majors suitable for Alice, and the criteria are listed below:

1) Academic environment: This is the academic atmosphere of the city where XYU is located. In this paper we use the number of universities in the city where XYU is located to represent academic environment.

2) Economy: This is the economic growth level of the city where XYU is located. The China Business Network Weekly publishes a ranked list of economic growth for all the cities in China every year. Cities are ranked from the first tier to the eighth tier according to their economic development, with the first being the best (such as Beijing, Shanghai, and Guangzhou).

3) Enrollment: This is the summation of the enrollment figures of all suitable majors of XYU.

4) Interest matching: This is the matching of Alice's favored majors with suitable majors of XYU. In mainland China, areas of studies are officially classified into a hierarchy of three classes. The Class-1 category consists of 11 general areas of studies: (1) Philosophy, (2) Economics, (3) Law, (4) Education, (5) Literature, (6) History, (7) Science, (8) Engineering, (9) Agriculture, (10) Medicine, (11) Management.

Each area in Class 1 (referred to as Class-1 subject) often consists of a number of subjects referred to as Class-2 subjects. For example, Science is a Class-1 subject, which consists of 12 Class-2 subjects: (1) Math, (2) Physics, (3) Chemistry, (4) Astronomy, (5) Geographical Sciences, (6) Atmospheric Sciences, (7) Ocean Sciences, (8) Geophysics, (9) Geology, (10) Biological Sciences, (11) Psychology, (12) Statistics.

Each Class-2 subject further consists of a few subdisciplines referred to as Class-3 subjects. For example, Math is a Class-2 subject, which consists of two Class-3 subjects: (1) Mathematics and Applied Mathematics, (2) Information and Computing Science. Each subject in any class is uniquely identified by a subject code.

We allow students to specify majors at the Class-1 level, Class-2 level (after Class 1 is specified), or the Class-3 level (after Class 2 is specified) [1]. Let (a, b, c) denote a specification of major, where a is a subject in Class 1 (which could be empty), b a subject in Class 2 (which could be empty), and c a subject in Class 3 (which could be empty). Note that if a is empty, then b and c must be empty. Likewise, if b is empty then c must be empty. Given a major specification (a, b, c) entered by s, we define the following terms:

1) We say that a *match* occurs at level 3 for student s with university u if one of the following conditions are satisfied:
   a) The university u offers c.
   b) The university u offers b, and c is empty (in this case, any Class-3 subject offered by u under b is deemed specified by s).
   c) The university u offers a, and b is empty (in this case, any Class-3 subject offered by u under a is deemed specified by s).
   d) The specification (a, b, c) is empty (in this case, any Class-3 subject offered by u is deemed specified by s).
2) We say that a *match* occurs at level 2 for student s with university u, if b is offered by u, but c (not empty) is not offered by u.
3) We say that a *match* occurs at level 1 for student s with university u, if a is offered by u, but b (not empty) is not offered by u.

In addition to matching majors, we would also like to put more weight on university u if it offers more majors under a given Class-2 subject, for it provides more related disciplines of studies for student s. For a particular Class-2 subject b, let n_b denote the number of Class-3 majors u offers under b.

5) Ranking: This is the group ranking of XYU. The first-tier universities are characterized into five groups: Group G_1 consists of the two super universities: Peking University and Tsinghua University. They are the best funded and most reputable universities in China. Both universities are designated by the Chinese government as project-985 universities. There are 39 universities in mainland China with this designation, which are the national key universities; Group G_2 consists of the top ten universities after Peking and Tsinghua. They are also project-985 universities. Group G_3 consists of all the remaining 27 project-985 universities; Group G_4 consists of all the officially designated project-211 universities, excluding project-985 universities. These are universities having top programs in certain areas; Group G_5 consists of the remaining first-tier universities. The second-tier universities can also be further characterized into two groups: Group G_6 consists of provincial key universities; Group G_7 consists of the remaining universities in this tier. Group G_8 consists of all the universities in the third tier.

B. Weights

Alice enters a weight value w_j for each criterion C_j, where $w_j \in \{0, 1, \ldots, 10\}$ with the following meanings: 0 is to ignore this criterion; 2 is to consider this criterion with no importance, 4 is moderately important; 6 is strongly important, 8 is very strongly important, 10 is extremely important, and 1, 3, 5, 7, 9 are between the above scales.

C. TOPSIS Steps

Step 1. Create an evaluation matrix U with universities as alternatives and the criteria set up above.

$$U = \begin{array}{c} \\ A_1 \\ A_2 \\ \vdots \\ A_m \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & C_2 & \cdots & C_n \end{array} \\ \begin{pmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,n} \\ u_{2,1} & u_{2,2} & \cdots & u_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{m,1} & u_{m,2} & \cdots & u_{m,n} \end{pmatrix} \end{array}$$

Step 2. Normalize matrix U to form the matrix $R = (r_{ij})_{m \times n}$, where for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$,

$$r_{ij} = u_{ij} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} u_{ij}^2 \right)^{-1/2}. \quad (1)$$

Step 3. Normalize weights entered by Alice such that the new weights, still denoted by w_j for criterion C_j with $j = 1, \ldots, n$, satisfies $\sum_{j=1}^{n} w_j = 1$. This is often referred to as linear normalization.

Step 4. Multiply the weights to each of the column entries in the matrix R to obtain a new matrix $T = (t_{ij})$, where $t_{ij} = w_i r_{ij}$, for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. That is,

$$T = \begin{pmatrix} w_1 r_{1,1} & w_2 r_{1,2} & \cdots & w_n r_{1,n} \\ w_1 r_{2,1} & w_2 r_{2,2} & \cdots & w_n r_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_1 r_{m,1} & w_2 r_{m,2} & \cdots & w_n r_{m,n} \end{pmatrix}$$

 ISBN: 978-1-61208-570-8

Step 5. Determine the best alternative vector, denoted by $\boldsymbol{A}_b$; and the worst alternative vector, denoted by $\boldsymbol{A}_w$. Write

$$\boldsymbol{A}_b = (t_{b1}, t_{b2}, \ldots, t_{bn})^T, \quad (2)$$

$$\boldsymbol{A}_w = (t_{w1}, t_{w2}, \ldots, t_{wn})^T. \quad (3)$$

Let $J_- = \{j \mid C_j \text{ is the smaller the better}, j = 1, \ldots, n\}$ and $J_+ = \{j \mid C_j \text{ is the larger the better}, j = 1, \ldots, n\}$. Then for $j = 1, 2, \ldots, n$,

$$t_{bj} = \begin{cases} \max\{t_{ij} \mid i = 1, 2, \ldots, m\}, & \text{if } C_j \in J_+, \\ \min\{t_{ij} \mid i = 1, 2, \ldots, m\}, & \text{if } C_j \in J_-; \end{cases} \quad (4)$$

$$t_{wj} = \begin{cases} \min\{t_{ij} \mid i = 1, 2, \ldots, m\}, & \text{if } C_j \in J_+, \\ \max\{t_{ij} \mid i = 1, 2, \ldots, m\}, & \text{if } C_j \in J_-. \end{cases} \quad (5)$$

Step 6. For each alternative A_i with $1 \le i \le m$, calculate the Euclidean distance between $(t_{i,1}, t_{i,2}, \cdots, t_{i,n})^T$ and $\boldsymbol{A}_b$, denoted by d_{ib}, and between $(t_{i,1}, t_{i,2}, \cdots, t_{i,n})^T$ and $\boldsymbol{A}_w$, denoted by d_{iw}, as follows: $d_{ib} = \sqrt{\sum_{j=1}^{n}(t_{ij} - t_{bj})^2}$, and $d_{iw} = \sqrt{\sum_{j=1}^{n}(t_{ij} - t_{wj})^2}$, where $i = 1, 2, \cdots, m$.

Step 7. For each alternative A_i, calculate the similarity to A_w as follows:

$$s_{iw} = d_{iw}/(d_{ib} + d_{iw}), \quad (6)$$

where $0 \le s_{iw} \le 1,\ i = 1, 2, \cdots, m$. The alternative A_i with the largest value s_{iw} is the best alternative. In other words, that university is the optimal one for Alice.

IV. Case Studies

We present two case studies, where Tom and Bob are high school seniors in the JiangSu Province. In JiangSu, each student must select two subjects from Politics, History, Physics, Chemistry, Geology, Biology, and Technology; and take them in addition to the must-take Gaokao subjects of Chinese, Mathematics, and English. The two selected subjects will be graded with a letter grade for A+, A, B+, B, C, or D. The three Gaokao subjects are graded numerically. Students who take the Li-Ke exam must choose Physics and students who take the Wen-Ke exam must choose History in their two selected subjects, respectively. Note that the selected subjects are criteria for GMA, not for TOPSIS. For a student to be admitted to a university, a good Gaokao score and good grades of selected subjects are a must. We will analyze the first-tier optimal recommendation for Tom and the second-tier optimal recommendation for Bob using TOPSIS, and justify that they make sense. We use data in the year of 2015 for demonstration.

A. Case A

Tom obtains a Gaokao score of 407 in the Li-Ke exam and A+ for both of his selected subjects. He is not interested in Literature, Economics, Law, Medicine, Management or Engineering, and he likes to study in Beijing or the Hubei province. The weights he gives to academic environment, economy, enrollment, interest matching, and ranking are 5, 4, 1, 1, and 7.

For the 2015 Gaokao in the Jiangsu province, the highest mark in the Li-Ke exam was 425 out of 480 (the student with this mark went to Tsinghua University) and the highest mark in the Wen-Ke exam was 418 out of 480 (the student with this mark went to Peking University). The mark of 407 in the Li-Ke exam was ranked from the 80th to the 94th among about 180,000 students who took the Li-Ke exams. Students in this range were admitted to top universities including University of Chinese Academy of Sciences and Renmin University of China in Beijing; and Wuhan University and Huazhong University of Science and Technology in Hubei. Table I shows the values for TOPSIS on the following universities recommended by GMA based on Tom's Gaokao score and interests, where AE stands for Academic Environment, IM for Interest Matching:

- China Agricultural University (CAU),
- Huazhong University of Science and Technology (HUST),
- University of Chinese Academy of Sciences (UCAS),
- Beijing Institute of Technology (BIT),
- Beijing University of Posts and Telecommunications (BUPT),
- Beihang University (BHU),
- North China Electric Power University (NCEPU),
- Wuhan University (WU),
- Central University of Finance and Economics (CUFE),
- Beijing Normal University (BNU),
- Renmin University of China (RUC),
- Zhongnan University of Economics and Law (ZUEL).

TABLE I. UNIVERSITY VALUES FOR TOPSIS

University	AE	Economy	Enrollment	IM	Ranking
CAU	10.0	10.0	1.0	10.0	7.50
HUST	10.0	7.50	2.0	10.0	7.50
UCAS	10.0	10.0	1.0	10.0	5.00
BIT	10.0	10.0	2.0	10.0	7.50
BUPT	10.0	10.0	1.0	10.0	6.25
BHU	10.0	10.0	2.0	10.0	7.50
NCEPU	10.0	10.0	1.0	10.0	6.25
WU	10.0	7.50	4.0	10.0	8.75
CUFE	10.0	10.0	1.0	10.0	6.25
BNU	10.0	10.0	2.0	10.0	7.50
RUC	10.0	10.0	1.0	10.0	8.75
ZUEL	10.0	7.50	1.0	10.0	6.25

Tables II and III represent, respectively, the linear normalization and criterion weight normalization. Table IV represents the process of multiplying the weights to each of the column entries in the normalized matrix of university values for TOPSIS. Tables V and VI show the best alternative vector and the worst alternative vector, respectively. Table VII displays the Euclidean distance between alternatives and the best alternative vector, and the worst alternative vector, respectively. Table VIII shows the similarities to the worst possible alternatives.

From Table VIII, we can see that RUC (Renmin University of China) in Beijing has the largest value. From Table II, we can see that RUC has the top score for each criterion except Enrollment, and from Table III, we can see that Tom cares more about Ranking, Academic Environment, and Economy, and he does not care much about Enrollment or Interest Matching. Thus, all things considered, we conclude that RUC is the best-suited university for Tom to apply. Based on the admission data in 2015, we confirm that RUC did admit students with Gaokao scores similar to Tom's.

TABLE II. NORMALIZED MATRIX OF UNIVERSITY VALUES FOR TOPSIS

University	AE	Economy	Enrollment	IM	Ranking
CAU	0.0389	0.0311	8.0E-4	0.0078	0.0409
HUST	0.0389	0.0234	0.0016	0.0078	0.0409
UCAS	0.0389	0.0311	8.0E-4	0.0078	0.0273
BIT	0.0389	0.0311	0.0016	0.0078	0.0409
BUPT	0.0389	0.0311	8.0E-4	0.0078	0.0341
BHU	0.0389	0.0311	0.0016	0.0078	0.0409
NCEPU	0.0389	0.0311	8.0E-4	0.0078	0.0341
WU	0.0389	0.0234	0.0031	0.0078	0.0477
CUFE	0.0389	0.0311	8.0E-4	0.0078	0.0341
BNU	0.0389	0.0311	0.0016	0.0078	0.0409
RUC	0.0389	0.0311	8.0E-4	0.0078	0.0477
ZUEL	0.0389	0.0234	8.0E-4	0.0078	0.0341

TABLE III. NORMALIZED CRITERION WEIGHTS

Criterion number	Criterion name	Normalized criterion weight
1	Academic Environment	0.25
2	Economy	0.20
3	Enrollment	0.05
4	Interest Matching	0.05
5	Ranking	0.35

TABLE IV. NORMALIZED WEIGHTED DATA

University	AE	Economy	Enrollment	IM	Ranking
CAU	0.0389	0.0311	8.0E-4	0.0078	0.0409
HUST	0.0389	0.0234	0.0016	0.0078	0.0409
UCAS	0.0389	0.0311	8.0E-4	0.0078	0.0273
BIT	0.0389	0.0311	0.0016	0.0078	0.0409
BUPT	0.0389	0.0311	8.0E-4	0.0078	0.0341
BHU	0.0389	0.0311	0.0016	0.0078	0.0409
NCEPU	0.0389	0.0311	8.0E-4	0.0078	0.0341
WU	0.0389	0.0234	0.0031	0.0078	0.0477
CUFE	0.0389	0.0311	8.0E-4	0.0078	0.0341
BNU	0.0389	0.0311	0.0016	0.0078	0.0409
RUC	0.0389	0.0311	8.0E-4	0.0078	0.0477
ZUEL	0.0389	0.0234	8.0E-4	0.0078	0.0341

TABLE V. THE BEST POSSIBLE

AE	Economy	Enrollment	IM	Ranking
0.0389	0.0311	0.0031	0.0078	0.0477

TABLE VI. THE WORST POSSIBLE

AE	Economy	Enrollment	IM	Ranking
0.0389	0.0234	8.0E-4	0.0078	0.0273

TABLE VII. DISTANCE TO THE BEST POSSIBLE (DTBP) AND THE DISTANCE TO THE WORST POSSIBLE (DTWP)

University	DTBP	DTWP
CAU	0.0072	0.0157
HUST	0.0105	0.0136
UCAS	0.0206	0.0078
BIT	0.0070	0.0157
BUPT	0.0138	0.0103
BHU	0.0070	0.0157
NCEPU	0.0138	0.0103
WU	0.0078	0.0206
CUFE	0.0138	0.0103
BNU	0.0070	0.0157
RUC	0.0023	0.0219
ZUEL	0.0159	0.0068

B. Case B

Bob's Gaokao score is 375 in the Wen-Ke exam and his two selected subjects are both B. Recall that the full mark of the Wen-Ke exam in 2015 was 480 and the highest mark was 418. He is interested in Literature but not in Agronomy, Medicine, Management, Law, or Economics. He likes to study in the Jiangsu province, Liaoning province, or Shanghai. He enters the following weights of 4, 2, 4, 5, 3 on Academic Environment, Economy, Enrollment, Interest Matching, and Ranking.

TABLE VIII. SIMILARITY TO THE WORST POSSIBLE (STWP)

University	STWP
CAU	0.6854
HUST	0.5661
UCAS	0.2746
BIT	0.6922
BUPT	0.4280
BHU	0.6922
NCEPU	0.4280
WU	0.7254
CUFE	0.4280
BNU	0.6922
RUC	0.9035
ZUEL	0.3004

Bob's Gaokao score in the Wen-Ke exam in 2015 was ranked from the 754th to the 834th. Students with Gaokao scores in this range were admitted by Shanghai Normal University, Nanjing University of Posts and Telecommunications in Jiangsu, and Dalian University of Foreign Languages in Liaoning. Table IX shows the university values for TOPSIS on the following universities recommended by GMA:

- Shanghai Second Polytechnic University (SSPU),
- Nanjing Technical University (NTU),
- Shanghai Customs College (SCC),
- Nanjing University of Posts and Telecommunications (NUPT),
- Shanghai University of Political Science and Law (SUPSL),
- Shanghai Normal University (SNU),
- Yangzhou University (YU),
- Shanghai Finance University (SFU),
- Nantong University (NU),
- Dalian University of Foreign Languages (DUFL),
- Shanghai Lixin University of Accounting and Finance (SLUAF),
- Shanghai Ocean University (SOU).

TABLE IX. UNIVERSITY VALUES FOR TOPSIS

University	AE	Economic level	Enrollment	IM	Ranking
SSPU	10.0	10.0	1.00	9.79	3.75
NTU	10.0	8.75	1.00	9.79	5.00
SCC	10.0	10.0	1.00	9.79	3.75
NUPT	10.0	8.75	6.00	10.0	5.00
SUPSL	10.0	10.0	2.00	10.0	2.50
SNU	10.0	10.0	1.00	9.79	3.75
YU	1.04	5.00	10.0	10.0	5.00
SFU	10.0	10.0	1.00	9.79	3.75
NU	1.11	5.00	2.00	9.79	3.75
DUFL	5.85	8.75	2.00	10.0	3.75
SLUAF	10.0	10.0	1.00	9.79	2.50
SOU	10.0	10.0	1.00	9.79	2.50

Table X shows linear normalization. Table XI displays criterion weight normalization. Table XII represents the process of multiplying the weights to each of the column entries in the normalized matrix of university values for TOPSIS. Tables XIII and XIV show the best alternative vector and the worst alternative vector, respectively. Table XV shows, respectively, the Euclidean distance between alternatives and the best alternative vector, and the worst alternative vector. Table XVI shows the similarity to the worst possible alternatives.

TABLE X. NORMALIZED MATRIX

University	AE	Economy	Enrollment	IM	Ranking
SSPU	0.1711	0.1711	0.0171	0.1675	0.0642
NTU	0.1711	0.1497	0.0171	0.1675	0.0855
SCC	0.1711	0.1711	0.0171	0.1675	0.0642
NUPT	0.1711	0.1497	0.1027	0.1711	0.0855
SUPSL	0.1711	0.1711	0.0342	0.1711	0.0428
SNU	0.1711	0.1711	0.0171	0.1675	0.0642
YU	0.0177	0.0855	0.1711	0.1711	0.0855
SFU	0.1711	0.1711	0.0171	0.1675	0.0642
NU	0.0190	0.0855	0.0342	0.1675	0.0642
DUFL	0.1001	0.1497	0.0342	0.1711	0.0642
SLUAF	0.1711	0.1711	0.0171	0.1675	0.0428
SOU	0.1711	0.1711	0.0171	0.1675	0.0428

TABLE XI. NORMALIZED CRITERION WEIGHTS

Criterion number	Criterion name	Normalized criterion weight
1	Academic Atmosphere	0.20
2	Economic Level	0.10
3	Enrollment	0.20
4	Interest Matching	0.25
5	Ranking	0.15

TABLE XII. NORMALIZED WEIGHTED DATA

University	AE	Economy	Enrollment	IM	Ranking
SSPU	0.0342	0.0171	0.0034	0.0419	0.0128
NTU	0.0342	0.0150	0.0034	0.0419	0.0128
SCC	0.0342	0.0171	0.0034	0.0419	0.0096
NUPT	0.0342	0.0150	0.0205	0.0428	0.0128
SUPSL	0.0342	0.0171	0.0068	0.0428	0.0064
SNU	0.0342	0.0171	0.0034	0.0419	0.0096
YU	0.0035	0.0086	0.0342	0.0428	0.0128
SFU	0.0342	0.0171	0.0034	0.0419	0.0096
NU	0.0038	0.0086	0.0068	0.0419	0.0096
DUFL	0.0200	0.0150	0.0068	0.0428	0.0096
SLUAF	0.0342	0.0171	0.0034	0.0419	0.0064
SOU	0.0342	0.0171	0.0034	0.0419	0.0064

TABLE XIII. THE BEST POSSIBLE

AE	Economy	Enrollment	IM	Ranking
0.0342	0.0171	0.0342	0.0428	0.0128

TABLE XIV. THE WORST POSSIBLE

AE	Economy	Enrollment	IN	Ranking
0.0035	0.0086	0.0034	0.0419	0.0064

From Table XVIII, we can see that NUPT (Nanjing University of Posts and Telecommunications) has the largest value. From Table XI, we can see that NUPT has the top score for all the criteria except Economy. From Table XII, we can see that Bob cares more about other criteria than Economy. Thus, we can conclude that NUPT is the best-suited university for Bob to apply. Based on the admission data in 2015, we confirm that NUPT did admit students with Gaokao scores similar to Bob's.

TABLE XV. DISTANCE TO THE BEST POSSIBLE (DTBP) AND DISTANCE TO THE WORST POSSIBLE (DTWP)

University	DTBP	DTWP
SSPU	0.0310	0.0320
NTU	0.0309	0.0320
SCC	0.0310	0.0320
NUPT	0.0139	0.0363
SUPSL	0.0281	0.0320
SNU	0.0310	0.0320
YU	0.0318	0.0315
SFU	0.0310	0.0320
NU	0.0419	0.0047
DUFL	0.0311	0.0183
SLUAF	0.0315	0.0318
SOU	0.0315	0.0318

TABLE XVI. SIMILARITY TO THE WORST POSSIBLE (STWP)

University	STWP
SSPU	0.5081
NTU	0.5088
SCC	0.5081
NUPT	0.7237
SUPSL	0.5326
SNU	0.5081
YU	0.4971
SFU	0.5081
NU	0.1007
DUFL	0.3708
SLUAF	0.5029
SOU	0.5029

V. CONCLUSION AND FUTURE WORK

Gaokao is a unique and major annual event in mainland China, which affects the lives of about 10 million graduating high-school students each year, and attracts tremendous attentions by their parents, relatives, and teachers. We presented an automated tool using computer assisted TOPSIS over big data to identify the best-suited university that matches the student's Gaokao score and interests. Our case studies showed that the recommendation provided by TOPSIS makes sense. Using computer assisted TOPSIS, students can easily figure out the best universities to apply under different sets of weights for the criteria. In a future project, we plan to fully develop this tool by allowing students to specify their own criteria.

REFERENCES

[1] S. Lu, H. Zhang, and J. Wang, "EEZY: A Gaokao Recommendation System Using General Morphological Analysis over Big Data." Journal of Acta Morphologica Generalis, vol. 12, 2016, pp. 5(1): 1–12, ISSN: 2001-2241.

[2] C. L. Hwang and K. Yoon, Multiple Attribute Decision Making: Methods and Applications. New York: Springer-Verlag, 1981.

[3] K. Yoon, "A reconciliation among discrete compromise situations." Journal of Operational Research Society, vol. 10, 1987, pp. 277–286.

[4] C. L. Hwang, Y. Lai, and T. Liu, "A new approach for multiple objective decision making." Computers and Operational Research, vol. 11, 1993, pp. 20:889–899.

[5] A. Assari, T. Mahesh, and E. Assari, " Role of public participation in sustainability of historical city: usage of TOPSIS method." Indian Journal of Science and Technology, vol. 6, 2012, pp. 5(3), 2289–2294.

[6] R. Greene, R. Devillers, J. Luther, and B. Eddy, "GIS-based multi-criteria analysis." Geography Compass, vol. 21, 2006, p. 5/6: 412432.

[7] E. Zavadskas, A. Zakarevicius, and J. Antucheviciene, " Evaluation of Ranking Accuracy in Multi-Criteria Decisions." Journal of Informatica, vol. 18, 2006, p. 17 (4): 601618.

[8] K. Yoon and C. Hwang, Multiple Attribute Decision Making: An Introduction. California: SAGE publications., 1995.