

Projection Model of Postgraduate Student Flow

Rahela Rahim¹, Haslinda Ibrahim², Maznah Mat Kasim³ and Farah Adibah Adnan⁴

^{1,2,3}School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah Malaysia

⁴Mathematical Engineering Institute, Universiti Malaysia Perlis, 02000 Kuala Perlis, Perlis, Malaysia

Received: 1 Oct. 2012, Revised: 22 Dec. 2012, Accepted: 25 Dec. 2012

Published online: 1 Jun. 2013

Abstract: An enrolment projection model based on the Markov chain is developed for postgraduate students at the College of Arts and Sciences in Universiti Utara Malaysia. Four years worth of data of student enrolments at the college were studied. The Markov chain model produced results close to the actual data for the first three years but deviated towards the fourth year. The model is helpful for the college's future planning in matters regarding postgraduate student enrolments.

Keywords: Student flow analysis, Enrolment projection model, Markov chain, Simulation.

1 Introduction

The average time taken to complete a course and the numerical success rate in higher education vary for each university and this has been a matter of discussion among education policy-makers. Indirectly, this will affect the internal planning of the universities in terms of predicting the number of student enrolments for each semester, estimating the number of lecturers needed to supervise research students, implementing classroom-planning and foreseeing other student needs such as spaces for car park, residential placements, cafeterias, etc. With the high rate of dropouts, deferred programs, unfinished studies, and new student admissions in universities, without proper planning, the managements will face problems in making any strategic planning and coming up with a decision. Due to this matter, decisions in educational policies and planning require the utilization of the student enrolment projection model. This model will translate into the demand for courses, personnel planning, facilities and other services needed. In addition, by understanding the flow of students' progressions during their studies, we can track any existing problems and take remedial actions to improve the matter. Accurate enrolment forecasts are crucial for universities to remain competitive, while inaccurate ones can lead to poor allocation of funds and resources [1]. Hence, this study aims to analyse future student flow based on the current one and try to answer questions such as the average time students spend to complete their studies as well as the probability of them

dropping out or graduating from their programs. This study is also able to be used to predict the number of intakes in the future. The results are hoped to provide guidance to the university administrative as well as the academicians in planning for the best programs.

Over the years, each department at the university college has admitted postgraduate students without any policy for the number to be admitted to various programs. Conventionally, the students are admitted based on their qualifications and the availability of lecturers to supervise their dissertations. Even though the enrolments each year show a generally stable condition, the management still cannot foresee the overall flow of students when this information is required for future planning. This is especially essential in recent time as some changes have taken place where the number of international student enrolments has tremendously increased. This phenomenon appears to affect the recruiting of students in ways that may change historical student flows.

Because of these and other factors, it was decided that a study of postgraduate student's flows throughout the university programs was necessary so that some basis would appear to understand how admissions can relate to available program resources and what kind of adjustments might be useful if admissions have increased.

There are various forecasting models used to estimate future enrolments such as cohort, regression, ratio, Markov and simulation. Among these techniques, the Markov chain seems to be the most suitable model for

* Corresponding author e-mail: rahela@uum.edu.my

this study. This is due to the specialty of the Markov chain method that not only can estimate promotion and repetition rates, but it can also estimate the number of dropouts, graduates and death rates in the matrix [2], [3], [4]. The Markov chain method can also measure detailed information on the students' progress such as the average time students spend in an education system whereby other techniques like regression and ratio are unable to measure this [5],[1],[6],[7]. Moreover, the Markov chain model is effective for predicting enrolment for a long range planning horizon which is over 5 years [8]. The capability of certain approaches to make a long-range planning is important for any institutions to give them directions and establish control over its own future.

The important element of the Markov chain is known as state. State can be divided into two; transient state and absorbing state. Transient state refers to the state where subject moves from state i during one period to state j in the next period, while the term absorbing state refers to the state where subject enters one state which is impossible to leave [9]. The transient states used in this study are the year of course enrolment [10], [11], [8], [12], [13] and the age of course enrolment [14], [12], [13], while the absorbing states for this study are dropouts and graduates from courses [14], [11], [8], [3], [12], [13]. A one-year period is used as the period of transition states [3], [12], [13], [8]. In this study, the specified states and year of duration are used to develop the Markov chain model for postgraduate students flow.

2 The Model

The Markov chain is a stochastic process where the probability of occurring events is not predetermined but it depends on the previous history events [9]. Markov chains have the computational properties that are able to calculate the probability of progressing from one state to another. This property will enable us to calculate the estimation of the probability of the students who have been previously admitted to enroll, withdraw or advance to the next semester.

The basic development of the Markov chain model is in matrix form. The general form of the Markov chain model is given by $n(t) = Z(t)c + D(t)$ where $n(t)$ is the vector of the number of students in each state at time t . $Z(t)$ is the matrix of the number of students moving from state i at time t to state j at time $t + 1$ and c is the unit vector of 1 and $D(t)$ is the number of students from transient state i at time t to absorbing state j at time $t + 1$. The equation indicates that the number of students at the commencing year of study consists of those who will survive to the next year and those who will leave the program in that year.

By adopting the formulation in Shah and Burke (1999), the transition matrix (Q) and the absorbing matrix (R) are used to calculate the estimated average time a postgraduate

student spend in system (T) and estimate the probability of graduation and termination (B) as shown below:

The matrix of transition (Q), given by $Q(t) = \tilde{n}^{-1}(t) Z(t)$ where $Q(t)$ is the probability of students making a transition from state i at time t to state j at time $t + 1$, and $\tilde{n}(t)$ is a diagonal matrix whose diagonal elements are the elements of $n(t)$.

The matrix of absorption (R), given by $R(t) = \tilde{n}^{-1}(t) D(t)$ where $R(t)$ is the probability of students who are in state i at time t moving to absorbing state j at time $t + 1$ and $D(t)$:the number of students from transient state i at time t to absorbing state j at time $t + 1$.

The fundamental matrix (N) of an absorbing Markov chain acts as a tool to the assessment of the student completion attributes

$$N = (I - Q)^{-1}$$

where N is the mean time (in years) a student commences in state i remains in state j before departing the program and I is the identity matrix. The average time a student spends in the system is calculated by,

$$T = N1$$

where T is the mean time a student remains in the program and the commencement of study is state i . Finally, the probability of a student moving into absorbing state (B) is given by $B = NR$.

3 The Data

The years from year 2005 until 2009 were chosen for data. The data containing students' information such as age, field of study, students' status whether they have graduated, dropped out, or deferred from their programs were examined. Based on the list, a transition matrix was calculated for each successive year by tabulating the number of students with respect to the state they were in the previous year and the state they were in the subsequent year. A transition matrix was calculated for each year by taking the average number in each state. The enrolment projection was made by the college system and separated by the students' status (part time and full time) and program (master and PhD).

4 Computational Result

The results of the proposed methods are presented in this section. In this study, we discuss the results obtained from the College of Art and Business (CAS).

Table 1 presents the transition process of full time Master students at CAS from one state to another in one academic year (1 year) over the 4-year duration of their studies that is from 2005 to 2009. In Table 1, the last column shows the total number of students in each state who began their studies in July 2005 (A051) and

November 2005 (A052) and continued (to the next year) or ended (dropped out or graduated) their studies in July 2006 (A061). For example, CAS had a total of 116 students, of which 54 students were between the age range of 26-30. Each row in the matrix indicates where students started and ended in a particular state. For example in row three, it can be interpreted that 54 students commenced the course at the age range of 26-30 in A051, 35 students continued to the next year (A061) in the same age range, while 13 other students were between the age range of 31-35. Both groups of students in these states were in their second year of enrolment. From 54 students who commenced at the age range of 26-30, 5 students dropped out from their studies while 1 student graduated at A061. For the values in the table, follow the same definition.

Table 1 Full time Master Students (CAS) Enrolment Transitions by Age Classification from July 2005 (A051) to July 2006 (A061) Session

To\From	21-25A	21-25B	26-30A	26-30B	31-35A	31-35B	36-40A	36-40B	41-45A	41-45B	46-50A	46-50B	Over 50A	Over 50B	Drop out	Grad	Total
From A051																	
21-25A																	
21-25B																	
26-30A				35	13										5	1	54
26-30B																	
31-35A					22	1									5	4	33
31-35B																	
36-40A							8	2							1	1	14
36-40B																	
41-45A									5	2					2	0	9
41-45B																	
46-50A											3				0	0	3
46-50B																	
Over 50A															3	0	3
Over 50B																	
Dropped															0	0	0
Grad															0	0	0
Total															5	0	116

Transition Probabilities of CAS' Full time Master Students

Table 2 presents the transition probabilities in matrix form which have been obtained from Table 1. The result in each entry in Table 2 is obtained by dividing each value in each row to the corresponding row total in Table 1. For example, value 0.65 in the third row and fourth column in Table 2 is obtained by dividing 35 with total value 54 from Table 1. Other values in the tables below follow the same definition. Lines are drawn here to differentiate the submatrices *Q*, *R*, *O*, *I* within the overall matrix.

Table 2 Transition Matrix for Full time Master Students (CAS) by Age Classification from July 2005 (A051) to July 2006 (A061) Session

July 2006 (A061) Session																	
To\From	21-25A	21-25B	26-30A	26-30B	31-35A	31-35B	36-40A	36-40B	41-45A	41-45B	46-50A	46-50B	Over 50A	Over 50B	Drop out	Grad	Total
From A051																	
21-25A																	
21-25B																	
26-30A			0.65	0.24											0.09	0.02	1.00
26-30B																	
31-35A					0.64	0.01									0.13	0.22	1.00
31-35B																	
36-40A							0.57	0.11							0.22	0.07	1.00
36-40B																	
41-45A									0.56	0.22					0.22	0	1.00
41-45B																	
46-50A											1.00				0	0	1.00
46-50B																	
Over 50A														1.00	0	0	1.00
Over 50B																	
Dropped															1.00	0	1.00

Analysis of the Transition Probabilities of CAS' Full time Master Students

From the matrix of transition probabilities constructed in Table 2, the transition probabilities are assumed to be constant over time since it gives a close result within the 4-year duration of studies. The steady state probabilities (Matrix *Q*) in Table 2 is the heart of this study since it will be used to predict future enrolments.

Enrolment Projection of Continuing Students

Table 3 presents the enrolment projection of continuing students to the next year of commencement who started their studies in July 2005. The steady state probabilities obtained in Table 2 is used to make the enrolment projection for the coming sessions (A061 until A081).

To project the enrolment for A061 (Table 3), the transition probabilities in Table 2 (steady state) should be multiplied by the individual total states from the A051 intake (Table 1). For example, the probability of 0.65 indicates that 65% of students between the age ranges of 26-30 will progress to the next year of commencement in the following year, so therefore the number of students involved is forecasted by $0.65 \times 54 = 35$. The projection for the next session (A071) will use forecasted data of A061 as the inputs, and the projection for the subsequent year (A081) follows the same manner. It should be noted here that the figures in fraction must be rounded to whole numbers.

Table 3 Projected Enrolment of Continuing CAS' Full time Master Students for July 2007 Intake (A081) by Age Classification

To Age	21-25.A	26-30.A	31-35.A	36-40.A	41-45.A	46-50.A	Over 50.A	Total
From A061								
21-25.A								
26-30.A		27.8	20.5					48.3
31-35.A			21.1	5.6				26.7
36-40.A				4.1	1.0			5.1
41-45.A					3.0	1.2		4.2
46-50.A						8		8
Over 50.A							1	1
Over 50.B								
Total								83.9

Comparison of Enrolment Projection

The prediction for the future number of postgraduate students in the session includes students who are commencing the course for the first time (new students) and also students who are continuing their studies to the next semester (continuing students). Table 4 and Figure 1 show the enrolment projection from the A061 until A081 sessions. To measure the accuracy of the model, a comparison with the historical data using mean absolute percent error (MAPE) is developed for results projection of the A061, A071 and A081 sessions. The results in Table 4 below show that the percentage errors are small (less than 10%) at the A061 and A071 sessions, but exceed 10% at the A081 session. This may be due to the fact that many students have graduated in this session. Hence, the predicted value is far from the actual enrolment of about 39 students. Therefore, we suggest that the results of dropouts and graduates obtained at this session are more meaningful for analysis.

Table 4 Comparison of Results between the Actual Data and the Markov Chain model of CAS' Full time Master Students

	A061			A071			A081			A081		
	New	Cont	Total	New	Cont	Total	New	Cont	Total	New	Cont	Total
Actual	130	0	130	196	181	377	230	63	293	356	27	383
Predicted (Markov)				186	95	281	230	73	303	305	86	391
Deviation						96			16			39
MAPE (%)						2.62			5.26			10.89

5 Conclusion

In this study, the Markov chain model is proposed for postgraduate students flow at the College of Arts and Sciences in Universiti Utara Malaysia. Historical data of the number of postgraduates in the college can be used to develop transitions to and from each semester session for four consecutive years. The results yield better estimates

of the actual data for the first three years but are less accurate for the fourth year where most postgraduate students are expected to have graduated. Therefore, the Markov chain model is proposed for the postgraduate projection model for the first three years of enrolment and for the fourth year, other analyses such as the probability of dropouts and graduates is expected to be more significant.

References

- [1] M. T. Healey and D. J. Brown, *Forecasting university enrollments by ratio smoothing*, *Higher Education*. **7** (1978) 417-429.
- [2] J. N. Johnstone, *Mathematical models developed for use in educational planning: A review*, *Review of educational research*. **44** (1974) 177-201.
- [3] V. M. H. Borden and J. F. Dalphin, *Simulating the effect of student profile changes on retention and graduation rates: A Markov chain analysis*, paper presented at the 38th annual forum of the association for institutional research, Minneapolis. (1998).
- [4] R. L. Armacost and A. L. Wilson, *Using Markov chain models to project university and program level enrolment [power point slides]*, annual conference. (2004).
- [5] F. E. Kinard and A. S. Krech, *Projected degree-credit enrollments through 1985 in South Carolina college and universities*. (1977).
- [6] R. S. Grip and J. W. Young, *Predicting school enrollments using the modified regression technique*, paper presented at the annual meeting of American Educational Research Association, Montreal, Canada. (1999).
- [7] S. Guo, *Three enrollment forecasting models: Issues in enrollment projection for community colleges*, presented at the 40th RP Conference, California. (2002).
- [8] M. G. Nicholls, *Assessing the progress and the underlying nature of the flows of doctoral and master degree candidates using absorbing Markov chain*, *Higher Education*. **53** (2007) 769-790.
- [9] S. Ahmad, *Markov chains theory (Unpublished master's thesis, Indiana State University, Terre Haute, Indiana*. (1985).
- [10] M. D. Orwig, P. K. Jones and O. T. Lenning, *Enrollment projection models for institutional planning*, *Higher Education*. **1** (1972) 435-448.
- [11] M. G. Nicholls, *A Markovian evaluation of a tertiary education faculty*, *Higher Education*. **12** (1983) 721-730.
- [12] C. Shah and G. Burke, *An undergraduate student flow model: Australian higher education*, *Higher Education*. (1999) 359-375.
- [13] I. Adyda and H. Hashibah, *Markov chain in modeling Universiti Utara Malaysia undergraduate student flow*. (2007).
- [14] W. L. Li, *A demographic model of student progression*, *International Review of Education*. **17** (1971) 408-424.



R. Abdul Rahim was born in Penang in 22 April 1969. Received her Bachelor of Mathematics degree from University of Malaya, Malaysia in 1993, Masters in Information Management from University of Sheffield,

United Kingdom in 1995 and Doctor of Philosophy from University Utara Malaysia, Sintok, Kedah in 2004. She has been with University Utara Malaysia for sixteen years. Her primary research areas of interest include mathematical modeling, queueing model and decision science. She has published numerous articles in research publications including *The Journal of Information and Communication Technology*, *Journal of Quality Measurement and Analysis* and *Journal of Modern Science*. Before joining School of Quantitative Sciences, she was a lecturer at School of IT. Her current research interest is in mathematical modeling. Dr Rahim is a member of Lab Publication Committee at University Utara Malaysia. The task includes promote the publication activities among scholars in her departments.

H. Ibrahim

obtained her bachelor degree in Mathematics in 1992(Michigan State University, USA), Master in Mathematics in 1996(Indiana State University) and PhD in Mathematics in 2003 (Southern Illinois

University, USA). She is well known for her profound contributions in combinatorics domain especially in Combinatorial Design Theory. Her current research interests include graph theory, combinatorial design theory, enumerative theory and optimization theory. Her work experiences include deputy dean of postgraduate, head of Mathematics Department, and Coordinator for Postgraduate program. She has been awarded several awards such as University Best Research Award (Universiti Utara Malaysia), Dissertation Research Award (Southern Illinois University, USA), Outstanding PhD Teaching Assistant (Southern Illinois University, USA) and Bronze Medal Award, 9th Malaysia Technology Expo: Invention and Innovation Award (Malaysia). She is also a member for American Mathematical Society (AMS) and Malaysia Mathematical Science Society (PERSAMA). She is also an editorial board for Journal of Mathematics Research. She has published over 50 articles in journals and conference proceedings.



M. Mat Kasim was born in Perlis on 18 May 1962. Received her Bachelor of Mathematics degree from Wichita State University of Malaya, Kansas, USA in 1984, Masters in Mathematics from the same university in 1986 and Doctor

of Philosophy from Universiti Kebangsaan Malaysia, Bangi, Selangor in 2008. She has been with Universiti Utara Malaysia for fifteen years. Her primary research areas of interest include multi-criteria decision making, and performance evaluation. She has published numerous articles in research publications including *The Journal of Information and Communication Technology*, *WSEAS Transactions on Business and Economic*, *Sains Malaysiana*, *Journal Of Business Case Studies and Analysis* and *International Journal of Digital Content Technology and its Applications(JDCTA)*. Before joining School of Quantitative Sciences, she was a lecturer at School of Information Technology. Her current research interest is performance evaluation. Dr Maznah Mat Kasim is a member of International Journal of Management Studies editorial board since 2009.

F.A Adnan

received her Bachelor of Decision Science with honours in 2008 and Master of Science (Decision Science) from Universiti Utara Malaysia. She began her professional career as a lecturer of Engineering Mathematics at the Institute

of Engineering Mathematics, Universiti Malaysia Perlis (UniMAP) in 2010. Her principal areas of research include operational research and optimization, Markov theory and design of experiment.

