The University as Educational Lab

Jane M. Fraser, Sadikin Djumin, James J. Mager
University of Southern Colorado/Ohio State University

Abstract

We report on a project that integrated teaching (supervision of a master's student), research (extending the work on Markov chain forecasts of student enrollment), and service (improvement of the university's methods for forecasting enrollment). We give recommendations on how to generate such projects and how to make such projects work well.

1. Introduction

Faculty members have three sets of obligations corresponding to the three areas on which they are evaluated: teaching, research, and service. Demands on time of faculty members usually mean that time spent on one of the three areas is time not spent on the others. We offer here the obvious suggestion that projects that can support more than one area are particularly desirable; indeed, *a project that can support all three areas can help a faculty member tremendously.*

Industrial engineering faculty members bring skills that can help make the operation of any human system more efficient and more effective. Industrial engineering faculty have a special opportunity, perhaps even an obligation, for service to their university. We offer here the obvious suggestion that *industrial engineering skills can be fruitfully applied to the many problems that face our colleges and universities.*

Finally, writing a master's thesis is an intense amount of work for a graduate student, often involving problems of finding a good topic, finding data, performing the work, and writing up the results. We offer here the obvious suggestion that *a real project, with a client who cares about and will help on the project and with pay for the student, can help a student tremendously in completing a worthwhile master's thesis.*

We report here on a project that combined those three suggestions:

- We addressed a problem of significance to Ohio State University and developed a solution that will be of ongoing help.
- A graduate student was employed while working on the project,
- The graduate student wrote up the project as his master's thesis.
- We made a small, but significant, contribution to research in higher education.
- The work will result (we plan) in a publication in a journal in higher education.

In this paper we present information on what was involved in doing the project, the mathematical model that we used, comments on what went well – and what didn't, and our recommendations to others attempting to do such projects.

2. What we did

Forecasting enrollments at Ohio State University (OSU) is a very important function performed by the Office of Enrollment Management. A large portion of the University's budget derives from tuition paid by students and state reimbursement based on enrollment, so forecasting enrollment is equivalent to forecasting a large portion of the University's budget. This forecast must be very accurate; errors of even 1% are unacceptable. The forecasting system that was being used at OSU had been written many years ago and suffered from the usual problems of such software, including the fact that the person who had written had left, and some portions were simply not well understood by current personnel. Also, the factors used to do the forecast had little intrinsic meaning and were hard to interpret or adjust for changing circumstances.

In 1993 two graduate students in industrial engineering[2] compared three methods for forecasting enrollments at OSU: a time-series model, multiple regression, and a Markov chain model. They found the Markov chain model to be superior; in fact, they found the model astoundingly accurate. While no follow up was done on the study immediately, the director of Enrollment Management (Dr. Mager) remembered the result and added it to his list of studies warranting further attention when resources permitted.

In 1996, Dr. Mager and Dr. Fraser continued discussion on the possible model and Dr. Mager reached the conclusion that, with Dr. Fraser's supervision, a student hired by the Office of Enrollment Management should pursue this Markov approach. Several students applied for the announced opportunity and Mr. Djumin, a master's student in industrial engineering with a bachelor's degree in computer and information science, was hired.

The first two tasks proceeded simultaneously: data collection and model building. The 1993 work had laid excellent groundwork for both, but data collection proved to be a very time consuming task. The model building was fairly straightforward, although decisions had to be made about the appropriate states to use.

With the huge number of students at OSU, records and data keeping are obviously highly automated, but the data systems were not created to produce the data needed for this study. We were able to obtain the number of students by rank in each semester for a historical period from 1991 to 1998. Data were available from before 1991, but it was felt that significant changes in university policies meant that earlier numbers were no longer relevant.

Mr. Djumin's questions and examination of the data turned up several definitional issues and led to the need to create a category called "missing students." Because students are counted on the 14th day of the quarter (since changed to the 15th day), students who enroll after that date are not in the official count, but appear in the official count the next quarter as if they had already been enrolled. There were several other similar issues, none of which were difficult, but, when combined, led data collection to take months, rather than days or weeks. The positive outcome is that the system was corrected and adapted in several ways so that future data were more accurate and more easily obtained; data for the model can be now obtained in minutes.

The Markov chain model has four requirements:

1. The system has discrete states or categories.
2. Changes in state occur only at discrete intervals.
3. The probability of moving from one state to another at any period of time depends only on the current state. Mathematically, the Markovian property for a Markov Chain is

$$P[Xi = j \mid Xn_1 = i_1, Xn_2 = i_2, \ldots, Xn_k = i_k] = P[Xi = j \mid Xn_1 = i_1],$$
$$\textit{for} \ldots i > n_1 > n_2 > \ldots > n_k \textit{ and every sequence } j, i_1, i_2, \ldots, i_k.$$

4. The transition probabilities remain constant over the period being considered.

After examining several possibilities, we decided to classify students as rank 1, 2, 3, and 4, with each subdivided into part-time and full-time. Rank is defined by the number of credit hours a student has completed; part and full-time by the number of credit hours being undertaken in the current quarter. An additional rank was defined as CRST or Rank V, including undergraduate transient students, non-degree students, conditional students and other undergraduate students who are not in the four ranks above; this category was also divided into part and full-time. These 10 states plus 6 other states: New First Quarter Freshmen (NFQF), Regional Transfers, Extra-University Transfers, Returning Students, Missing Students, and Left the University (by graduating or dropping out) made up the 16 states in the Markov chain model. Note that all states are transient, except the last, which is absorbing. We chose not to try to model the return of students who stopped out to the university. While 256 transitions are possible, only 129 can be nonzero since a student cannot move back in rank.

OSU is on the quarter system so the obvious choice of time for each step was one quarter.

We chose not to examine the Markovian property, believing that the real test of our model would be its forecasting ability. Obviously, however, a student who has been in rank 1 for several quarters has a higher probability of moving to rank 2 than a student who is in rank 1 for the first time does.

Finally, the most important decision we made was to drop assumption four, that is, we chose to create a nonstationary Markov model in which the transition probabilities change over time. While the Markov chain model has been used in higher education to forecast and to track student flows,[3] all previous applications have used a stationary model; indeed, most of the published articles in this area do not seem to recognize that assumption four is not strictly required. In the published higher education literature, a nonstationary Markov chain has never been used to forecast student enrollment.

In our nonstationary model, the matrix changes in two ways. First, the transition matrix is different for fall-to-winter, winter-to-spring, spring-to-summer, and summer-to-spring transitions. (Actually, since many students do not attend OSU in the summer, we predicted fall enrollment from a spring-to-fall transition matrix for students enrolled in the spring and a summer-to-fall matrix for students who first enrolled in the summer.) Also, each matrix, for example, the fall-to-winter transition matrix, changes from year-to-year.

Our problem became one of forecasting the transition matrix. We used our historical information to create historical matrices (for each quarterly transition matrix) and examined several methods for forecasting these rates, from simple linear regression through double exponential smoothing. We also considered how to incorporate human judgment into the forecasts. Simple exponential smoothing performed best most often, but we found that different methods worked better for

different transition probabilities. The selection of different methods for different probabilities in the same row of the transition matrix meant that we had to select a way to ensure that each row summed to one; we used the leaving rate (the transition to the absorbing state) to balance the row. This paragraph describes only briefly a huge amount of work done by Mr. Djumin.[1]

As the final and most important step, we performed a test of the model. In all steps involving creation of the model, the most recent year's data were withheld. We then tested the model by predicting the enrollment for those four quarters. The results were gratifying, with the nonstationary Markov model projection being more accurate than a stationary projection (that is, a simple average of historical data) and more accurate than the University's own system had been. Errors were well under the maximum 1% error. However, the rank-by-rank predictions were not as accurate and work will continue to see if they can be improved. (State reimbursement varies by rank of the student so such forecasts are important).

Research always opens the possibilities of more research. Continued refinement will be needed to improve the forecast of enrollment by rank. Also, a model should be created to forecast the return of students who have stopped out. We believe that attention should be given to jointly forecasting the elements of the matrix; an understanding of why one transition rate shows a particular trend may lead to better predictions about another transition rate. While Dr. Fraser has left OSU, it is possible that Dr. Mager will work with another faculty member in the industrial engineering department at OSU to continue along these lines.

3. What went well – and what didn't

Data collection proved to be a much more serious undertaking than predicted, but, as all of the readers of this paper know, that statement can be made of almost every industrial engineering project. Indeed, the graduate student learned that fact very well and it can be argued this learning was a very valuable educational outcome. Also, the difficulty of data collection led to improvements in the data system so that the model data can now be obtained in minutes. Indeed, this improvement to the data system may be one of the most long lasting positive effects of the project.

Many of the team involved in the project, including staff people in Dr. Mager's office, knew each other. Dr. Mager, director of OSU's Office of Enrollment Management, holds a Ph.D. in industrial engineering from the department in which Dr. Fraser was a faculty member. While Dr. Mager completed his degree before Dr. Fraser joined the faculty, they knew each other through various faculty governance committees. In particular, Dr. Fraser was a member of the University Senate Committee on Enrollment and Student Progress, which works closely with Dr. Mager. The OSU staff person in charge of the data systems that produce the 14-day report (the data systems that were crucial for the input to this model) had taken a class from Dr. Fraser while a student in the OSU college of business several years before this study. Mr. Djumin had taken classes from Dr. Fraser. The result was that cooperation was forthcoming in many ways.

The sponsor of the project, Dr. Mager, was enthusiastic and controlled the resources necessary to get the project done. Because of the 1993 study, Dr. Mager was, in fact, the person who most insistently initiated the project. The germination period was long, but the earlier study planted a seed that eventually bore fruit.

The selection of the student to be employed by the Office of Enrollment Management was done by the usual staff procedures in that office, not by Dr. Fraser. The opportunity (including the opportunity to get a master's thesis from the project) was announced via a flyer in the mailboxes of all graduate students in the department. After interviews with the people who would be involved in the project, the candidates were evaluated by the OSU staff. Dr. Fraser provided input at that point, but did not select the student. As a result of using the traditional staff selection process in that office, the staff in Dr. Mager's group quickly accepted the student.

4. Recommendations

It is easy to give general recommendations: look for opportunities for you and your graduate students to apply industrial engineering skills in service to your university. You can thus make contributions in all the areas on which faculty are evaluated: teaching, research, and service. We have tried below to give more specific recommendations about how to generate such opportunities and about how to make the opportunities work after you have generated them.

Perhaps the crucial point in the success of this project was the Dr. Mager has a Ph.D. in industrial engineering. He understood the methodology and appreciated the contributions that industrial engineering could make. Our first recommendation is to work with such people and even to try to make sure that your university hires such people. In the usual help you provide to your graduates in job hunting, be sure they consider positions open at your university.

Dr. Fraser's involvement in University committees created the opportunity for the department to be involved in this study. Therefore, we recommend that as you perform your service obligations on university committees you select involvement where your expertise might be of use and then that you volunteer your services.

The 1993 study was a less thorough model, but was a clear demonstration that the Markov chain approach was worth attention. This study was done as part of a graduate level course project. We recommend that class projects be used in this preliminary sales approach. Even without the direct cooperation of a university office, some data can be obtained and used to demonstrate feasibility. The report should be transmitted to the university office. The students learn from such a project and the paper may be useful in persuading the university office that a more serious study will be worthwhile. Such a preliminary sales approach does require patience, since the gestation period can be years.

Mr. Djumin was selected by and paid by the OSU Office of Enrollment Management as a part-time employee while he continued course work and work on his master's thesis. He thus was accepted by them and was able to work with those staff people very well. He also received financial support for his graduate work. We recommend that the sponsor select the student (with professor's input) and pay and treat the student as a regular employee.

Another possible outcome of interaction of university staff with faculty and graduate students is the recruitment of new graduate students. As a result of interaction during this project, it is possible that one of the staff persons involved will continue graduate work in industrial engineering. We recommend that faculty members recruit graduate students from University staff.

The work should result in a published paper. While the nonstationary model is not a contribution to industrial engineering, its use in higher education is, we believe, a significant improvement to enrollment forecasting practice. Thus we recommend that industrial engineering faculty recognize the contributions they can make by exporting well known industrial engineering results to fields where these results are not well known. We are writing a paper that reports on the use of a nonstationary Markov model to forecast enrollment and will submit it to a research journal in higher education.

5. Conclusions

Applying industrial engineering to University projects offers an opportunity to encourage the use of fact-based decision making by your administration. Even a project involving collection and analysis of data, without extensive model building, can help focus discussion.

We began this paper by noting the three sets of obligations of faculty members. A result of a successful project such as the one reported here is that the lines among those areas become blurred. Indeed, it was difficult to know where to report this project in an annual report divided into teaching, research, and service. Such a reintegration of the three areas of obligations may help faculty members make better use of their time.

Bibliography

1. Djumin, Sadikin. *A Non-Stationary Markov Chain Model for Enrollment Projection for Undergraduate Students at the Ohio State University*, M.S. thesis, 1998.
2. Mitchell, Christine A., and Karen Hartz Semer, *Forecasting Enrollments at the Ohio State University*, December 1993.
3. Thonstad , Tore. *Education and Manpower, Theoretical Models and Empirical Applications*, University of Toronto Press, 1968.

JANE M. FRASER

Jane M. Fraser is Chair of the Department of Engineering at the University of Southern Colorado. She received the B.A. in mathematics from Swarthmore College, and M.S. and Ph.D. from the University of California at Berkeley. She has been on the faculty at Purdue University and Ohio State University.

SADIKIN DJUMIN

Sadikin Djumin was born in Indonesia. He received the B.S. in Computer and Information Science and the M.S. in Industrial and Systems Engineering from Ohio State University. He was employed as a Graduate Research Assistant in the Office of Enrollment Management 1997-1998.

JAMES J. MAGER

James J. Mager is Assistant Vice President for Enrollment Management at Ohio State University. He received the Ph.D. in Industrial Engineering from Ohio State University.