Modeling Dormitory Occupancy Using Markov Chains

David D. Pokrajac
Delaware State University
1200 N DuPont Hwy
Dover, DE 19901
+1-302-857-7614
dpokrajac@desu.edu

Kimberley Sudler
Delaware State University
1200 N DuPont Hwy
Dover, DE 19901
+1-302-857-7036
krsudler@desu.edu

Diana Yankovich
Delaware State University
1200 N DuPont Hwy
Dover, DE 19901
+1-302-857-6308
dyankovich@desu.edu

Teresa Hardee
Delaware State University
1200 N DuPont Hwy
Dover, DE 19901
+1-302-857-7837
thardee@desu.edu

ABSTRACT

We introduce a Markov chain based model that quantifies university dormitory occupancy as a function of parameters related to university housing policies, students' success and academic progress, and customer satisfaction/dorm availability. The model provides sensitivity of university housing occupancy on change of the parameters. We demonstrated functionality of the model on several case scenarios from a public university.

Keywords

Modeling, dormitory occupancy, university housing, Markov chains, sensitivity, students' success, Banner.

1. INTRODUCTION

In this study, we introduce a housing occupancy model based on Markov chains [e.g., 1]. The model determines relationship between the number of students in dormitories, number of students in incoming class and probabilities quantifying students' retention, advancement between ranks (freshmen, sophomores, etc.), customer satisfaction and availability of housing. The model provides an opportunity for what-if analysis and assessment of change in housing occupancy due to variation of model parameters. The values of model parameters are learned from a transactional database.

We provide a case study based on three years data from Delaware State University, a public comprehensive historically black college/university in Delaware and demonstrate quantitative change of housing occupancy as results of possible changes in housing policy, housing demand and retention. The proposed technique is applicable to universities offering predominantly undergraduate programs and can be easily adapted for universities with substantial graduate programs and participation of international students.

2. METHODOLOGY

2.1 Problem

We consider a university offering undergraduate programs. The students at the university may be of in-state or out-of-state domicile (in-state students are the students whose residence is in the same state as the university). During the course of study, out-of-state students may convert to in-state or vice versa. A new student at the university can be enrolled as a new freshman (NF) or a new transfer (NT). For a student retained at the university, a rank depends on the cumulative number of credits (earned at the university + transferred). The ranks satisfy partial order. Thus, a NF or NT, if retained, may continue as returning freshmen (RF), sophomore (SO), junior (JR) or senior (SR). Retained RF may continue as RF or progress into SO, JR or SR. Retained SO may continue as SO, or progress as JR or SR. Each student in a particular year can be a dorm resident. If retained, a student may change dorm residency status, i.e., a dorm non-resident may become dorm resident or vice versa.

Our goal is to determine the relationship between various parameters characterizing students' population and academic progress and the total number of dorm residents in a particular year.

2.2 Markov Chain Model

We model the considered problem with a time-homogeneous Markov chain [1]. A student at the university can be described by a state $s_{(i,j,k)}$ determined by an ordered triple of indices i, j, and k indicating domicile, rank and dorm residence: $i \in \{InState, OutOfState\}$, $j \in \{NF, NT, RF, SO, JR, SR\}$ and $k \in \{DormResident, NotDormResident\}$. The starting states correspond to $i \in \{InState, OutOfState\}$, $j \in \{NF, NT\}$, $k \in \{DormResident, NotDormResident\}$. The total number of non-absorbing states is 24. In addition, a student can graduate or leave the university, corresponding to an absorbing state, denoted with s_a. The transition between states $s_{(i,j,k)}$ and $s_{(i',j',k')}$ is uniquely determined by transition probability that, under the assumption of time homogeneity is denoted by $p_{(i,j,k),(i',j',k')}$. In addition, the model includes transition probabilities $p_{(i,j,k),a}$ from states $s_{(i,j,k)}$ to the absorbing state.

2.3 Model Implementation

To operationalize the model, we introduce the following assumptions and simplifications:

1) Students can transition only from out-of-state to in-state status;

2) For in-state students who continue to stay in dorms, the transition probability can be expressed as product of probabilities that a student is retained, that a student advanced from rank j to j' and the probability that a student stayed in dorm;

3) For out-of-state students who continue to stay in dorms as out-of-state, the transition probability is expressed as a product of probabilities that a student is retained, that a student does not change out-of-state status, that a student advanced from rank j to j' and the probability that a student stayed in dorm;

4) For out-of-state students who continue to stay in dorms as in-state, the transition probability is expressed as a product of probabilities that a student is retained, that a student changes out-of-state status to instate, that a student advanced from rank j to j' and the probability that a student stayed in dorm;

5) We compute probabilities that a dorm resident with domicile i' and rank j' was a dorm resident in the previous year.

2.4 Model Sensitivity

After the parameter values are estimated, the sensitivity s_l of the number of students in dorms on a particular parameter π_l can be determined as: $s(\pi_l) = \frac{\Delta N^y}{\Delta \pi_l}$, where ΔN^y is change of number of students in dorms, due to change $\Delta\pi_l = \pi_l^{new} - \pi_l$ of a parameter. Subsequently, the influence of change of particular model parameters on the model output—the number of students in dorms can be linearized such that: $\Delta N^y = \sum_l s(\pi_l)\Delta\pi_l$.

3. RESULTS

3.1 Data Set

We estimated the model discussed in Section 2 on data from Delaware State University (DSU), a historically black college/university (HBCU) located in Dover, DE, USA. DSU utilizes Banner® Version 8 (Ellucian, Fairwax, VA, USA) as a higher education enterprise resource planning (ERP) system. The dataset contained the total of 13,709 records from years 2013/14—2015/16. Each record had the values of attributes: StudentID, Year, Rank, DormResidence, Domicile. StudentID is a unique identifier of a student and together with Year comprise the primary key of the extracted table.

3.2 What-if Analyses

In this section we analyze realistic cases for changes of some of the model parameters and their influence on the change of number of students in dormitories.

Case 1. Due to policy change, *all* new freshmen and new transfers are expected to stay at university housing *regardless* whether they are in-state or out-of-state. We can easily obtain the increase of the number of students in dormitories of ΔN^y=467.

Case 2. Due to implementation of initiatives to address needs of incoming and returning freshmen, the retentions of in-dorm new and returning freshmen increase to 80%. This leads to the increase of ΔN^y=175 students in dorms.

Case 3. Owing to improvement of dorm facilities, the demand for dorm housing for upper rank students increases. This can, thus, be considered as a result of increased customer satisfaction. As a consequence, this leads to the increase of ΔN^y=83.

4. DISCUSSION

The proposed model makes it possible to account for retention that is frequently a key performance indicator related to university strategic plans and one of common quantitative measures of students' success. Further, the model involves parameters related to academic progress of students. Also, we can indirectly model housing satisfaction and availability. The model makes it possible to consider in-state and out-of-state students separately, as the two groups of students that may have different demography, socio-economical conditions and academic success. Also, it is possible to evaluate the relationship between the size of the incoming class (new freshmen and transfers) and the housing occupancy.

The model considers only two categories of students: in-state and out-of-state students. For universities with substantial numbers of international students, they can be added as an additional category and treated similarly as out-of-state students. The model assumes that in-state students cannot become out-of-state. However, the assumption can be relaxed by introducing a non-zero probability that in-state students of rank j become out-of-state.The assumptions 2—4 (probability independencies) may be contingent on university policies (distribution of students within dorms and on-campus housing allocation across student classes/ranks). Hence, they should be validated prior to the application of the proposed models at another institution of higher education. The current model assumes that the students who leave the university without graduating do not come on a later date. In reality, some students may leave the university temporarily and return ("stop-outs"). Note that we utilized point estimates, hence the accuracy of parameter estimates (e.g., standard deviation) has not been addressed. Future work will include the development of interval estimates for model parameters as well as an application of validation techniques (e.g., leave-one-out cross-validation) to more strictly justify predictive ability of the model.

5. CONCLUSION

We proposed a Markov chain-based model of university housing occupancy and demonstrated it in a case study of a public university. We have shown that the proposed model can be useful in quantifying what-if scenarios related to changes in housing policy, retention and customer satisfaction. The model is developed for a university offering primarily undergraduate programs. It can be extended to graduate program offering institutions, with a challenge that graduate (especially PhD) programs are typically less structured (as evidenced in lack of ranks corresponding to sophomores, juniors, seniors in undergraduate programs). We demonstrated the use of a model with parameters estimated from data readily available on an industry-standard ERP system (Banner). As such, the model can be easily deployed at an institution of higher education that utilizes this or similar technology.

6. ACKNOWLEDGMENTS

This work has been supported through a grant from the Bill and Melinda Gates foundation.

7. REFERENCE

[1] Grinstead, C.M. 1997. *Introduction to Probability*, 2nd edn. American Mathematical Society, Providence, RI.