Journal of Systemics, Cybernetics and Informatics (2024) 22(2), 32-36
https://doi.org/10.54808/JSCI.22.02.32

A Markov Chain approach for the evaluation of global performance of Universities of South Africa

Masikini LUGOMA
Department of Mining, Minerals and Geomatics Engineering
University of South Africa
Pretoria, Gauteng, South Africa

Masengo ILUNGA
Department of Civil and Environmental Engineering and Building Science
University of South Africa
Pretoria, Gauteng, South Africa

ABSTRACT[1]

The Markov chain (MC) technique is used preliminarily in the evaluation of the ranking of selected institutions of higher learning of South Africa. The data related to Scimago is used to illustrate the practical implication of MC. A random selection of universities was conducted for applying the MC technique., i.e., a middle and a bottom performing universities are selected. Different states were defined to carry out the estimation of the steady state of the Markov Chain to predict theoretically the future state of their rankings. These preliminary results showed that the former had 2 states with probability distributions of (0, 1) respectively, whereas the latter had 5 states with probability distribution of (0.892, 0.035, 0.035, 0.035, 0.01). This approach could be replicated to the rest of institutions of higher learning and considering other ranking metrics.

Keywords: Markov Chain, Universities Ranking, probabilistic forecasting, Scimago rankings.

1. INTRODUCTION

Global ranking of institutions of higher education plays a significant role in helping gauge themselves internationally as far as possible. Criteria such as research, teaching and learning are considered in the process. These are recognized as the main criteria, however, there are no methodology or universally agreed indicators for quality assessment as far as university contributions to society is concerned [1]. These authors suggested additional indicators associated with continuing education, technology transfer and innovation. Professional and government bodies have enabled the emergence of higher education rankings as stipulated by UNESCO CEPES [2].

Among the recognized rankings are the Academic Ranking of World Universities (ARWU) or Ranking of Shanghai, QS World University Ranking, Scimago Institutions Rankings SIR and the Web Ranking of Universities-Webometrics [3].
These authors [3] indicated that the purpose of these rankings is to organize the universities according to indicators that should reflect their capacity as an institution, quality of academic activities, production and dissemination of research, innovation, and relations abroad of universities. They are also used to make decisions, from the distribution of research funds to the desired profiles of teachers and researchers. Knowing the characteristics of the rankings offers valuable information for the definition of strategies for the international positioning of universities. The bases of these rankings are further briefly described.

The Academic Ranking of World Universities (ARWU) [4] was first published in June 2003 by the World Class University Centre (CWCU) of Jiao Tong University in Shanghai, China; updated annually. ARWU uses six (6) objective indicators to classify the world's universities based on quality education, faculty quality and research output. As of 2017, universities classified between 501 and 800 are also published as ARWU World Top 500 Candidates. The highest scoring institution is assigned a score of 100 and the rest is calculated as a percentage of the maximum score.

The QS World University Ranking [5] has been published since 2004 with an annual periodicity, and considers academic, employers, students, and international indicators. Among the aspects to be measured are the citations received, the student-teacher ratio, the proportion of international students and foreign professors, the academic reputation, the reputation among employers, and personnel with a doctorate.

[1] The final version of the paper was peer-edited by Dr Israel Dikgwatlhe, who is acknowledged.

This study will use the SIR SCimago Ranking which begins in 2009. It is conducted by the Spanish Scimago research group and is called SCimago Institutions Rankings (SIR). Its periodicity is annual and published until to date.

The SCImago Institutions Rankings (SIR) is a web-based classification of academic and research-related institutions ranked by a composite indicator that combines three different sets of indicators based on research performance, innovation outputs and societal impact measured by their web visibility [6].

It provides a friendly interface that allows the visualization of any customized ranking from the combination of these three sets of indicators. Additionally, it is possible to compare the trends for individual indicators of up to six institutions. For each large sector, it is also possible to obtain distribution charts of the different indicators.

Molinari and Molinari [7] developed a methodology that complements the h-index, for comparing the scientific production of institutions of higher learning, laboratory facilities or journals.

It is difficult for a university to know exactly in advance what will be its ranking. Ranking could be associated with uncertainty, hence may be associated with a degree of probability or randomness. Most importantly, the current performance of a university is likely to influence the next performance, in terms of ranking. Generally, a random process, called stochastic process is well described by Markov chain in [8].

2. MARKOV CHAIN CONCEPTS

The MC deals with stochastic processes known as random events such that the probability of the future state of a system is not a function of the previous states but of the present state [9]. The considerations in this section have similarities with previous studies, e.g. [10], [11]. Given a state space S = 1, 2, ……m, the MC is defined by a series of random variables Y_k € S where k = 1,2, …m

$$\Pr [Xq+1 = xq+1 / X1 = x1, \ldots\ldots, Xq = xq] = \text{Prob} [Xq+1 = x_{q+1} / X1 = x_1, \ldots, Xq = xq] = \Pr [Xk+1 / Xk] \quad (1)$$

This expression is valid for homogeneity characteristics of MC. It implies that given $X_1, X_2, X_3, \ldots\ldots, Xq, Xq+1$, the conditional distribution of Xk+1 is only a function of the value of Yk, not a function of the previous values Y_1, ……Yk-1. This form of conditional probability is derived from the memorylessness principle. The probability matrix of the MC corresponding to the finite space state is of dimension nxn and is called the transition matrix. The transition matrix summarises all probabilities p_{ij} obtained from Equation (2), into the following matrix.

$$A = [P_{kj}] \quad (2)$$

Where P_{kj} is the transition probability to move from Si to state Sk, k, j = 1, 2, …, m.

The transition matrix is instrumental in obtaining probability distribution at different times. An absorbing Markov chain has one or more absorbing states, which do not allow a subject to leave, meaning that once a subject enters an absorbing state, it remains trapped. An absorbing state is characterised by a maximum probability of occurrence. A state different from absorbing states is said to be transient or non-absorbing state. From a non-absorbing state, it is possible to reach absorbing states in one step or several steps.

The following concepts of Markov Chain are discussed and their implications in the analysis of the rankings of the universities: Ergodic, Irreducible and Periodic Markov Chains [12].

Ergodic Markov Chain (EMC)

An *Ergodic Markov Chain* is a special type of Markov Chain that possesses certain properties that make it particularly useful for analysis. In an Ergodic Markov Chain, the system reaches a state of equilibrium over time, where the probabilities of transitioning between states stabilize and remain constant. This equilibrium state is often referred to as the stationary distribution or the steady state.

In an Ergodic Markov Chain, every state is reachable from any other state, meaning there are no transient states. This ensures that the chain will eventually reach the equilibrium state regardless of its initial state. Additionally, the chain must be irreducible, which means there are no subsets of states that are isolated from the rest of the chain.

The key characteristic of an Ergodic Markov Chain is that it satisfies the ergodicity property. This property states that as the number of transitions increases, the probabilities of being in each state converge to fixed values. These fixed values represent the long-term behavior of the Markov Chain and are independent of the initial state.

By modelling the ranking transitions as an Ergodic Markov Chain, we can analyze the long-term behavior and equilibrium of the university's ranking. The ergodicity property ensures that over time, the university's ranking will reach a steady-state distribution where the probabilities of being in each ranking interval stabilize.

With an Ergodic Markov Chain, we can estimate the long-term probabilities of the university being in each ranking interval. This information can provide valuable insights into the stability or volatility of the university's ranking position. It allows us to understand the likelihood of the university moving between different ranking intervals and the overall dynamics of its ranking performance over the period under study.

Irreducible Markov Chain

An *irreducible Markov Chain* is one in which every state is reachable from any other state. In the context of university rankings, this means that there are no subsets of ranking intervals that are isolated from the rest. In other words, it is possible for a university to transition between any two ranking intervals over the analysed period. This property is important because it ensures that the Markov Chain has a single, unified behavior and that there are no disconnected portions of the ranking system.

Periodic Markov Chain

A *periodic Markov Chain* is one in which the chain returns to certain states in fixed intervals of time, known as the period. In the context of university rankings, this would mean that the university's ranking periodically cycles through a specific set of ranking intervals. This can happen if there are certain patterns or factors that influence the university's ranking performance and cause it to consistently move between specific ranking intervals. A periodic Markov Chain can have different periods, such as a period of 1 (no periodicity) or a period greater than 1 (indicating a repeating pattern).

Understanding whether the Markov Chain representing the university rankings is irreducible or periodic can provide insights into the dynamics and behavior of the rankings. An irreducible Markov Chain ensures that all ranking intervals are interconnected, allowing for potential transitions between any two intervals. A periodic Markov Chain suggests the presence of recurring patterns in the university's ranking movements.

3. DATA AND METHODS

Data availability

The data used were extracted from the "Scimago Institutions Rankings" website [6]. The data varied from 2009 and 2023 and were selected for University of Johannesburg [UJ] and University of Limpopo [UL], which are middle and lower universities in terms of ranking, as said earlier. From the data used, it was noticed that none of the universities was within the first 500 universities, however, they ranged between 501 and 7300.

Methods

The different ranking intervals, which range in intervals of 1000, can be considered as the states of the Markov Chain. In this case, the rankings range from 1-1000, 1001-2000, 2001-3000, and so on. Hence 8 states were considered that a university can move between different states over the 15-year period.

The different states are described in Table 1:

TABLE I: Defined States of the Markov Chain for the two selected Universities

	University of Johannesburg		University of Limpopo	
Year	Ranking	State	Ranking	State
2009	3175	S4	3720	S4
2010	3220	S4	3976	S4
2011	3470	S4	4141	S5
2012	3436	S4	4285	S5
2013	3287	S4	4556	S5
2014	3089	S4	4637	S5
2015	2807	S3	4788	S5
2016	2277	S3	4746	S5
2017	2444	S3	4973	S5
2018	2675	S3	5302	S6
2019	2550	S3	4407	S5
2020	2475	S3	5474	S6
2021	2383	S3	6192	S7
2022	2513	S3	6685	S7
2023	2655	S3	7550	S8

From Table 1, the different frequencies of each possible state (S3 to S8) can be estimated by counting the number of occurrences.

The probability of moving from one state to another state is approximated by the frequency. The probabilities thus defined allows to establish the probability transition matrix.

The steady state is determined by solving the following system of equations:

$$\begin{cases} Ay = y \\ \sum y_i = 1 \end{cases} \tag{3}$$

The first equation emanates from linear Algebra, where y is the eigen vector, considering an eigen value of 1, and the last equation satisfies the probability formalism, where "y" corresponds to the probability distribution of the steady state and "A" the transition probability matrix.

4. RESULTS AND DISCUSSION

Based on the data, 8 states were established (Table 1), with an interval of 1000, and frequency of each state were calculated between 2009 and 2023. The probability of remaining in the state was evaluated where possible and transition probabilities as well. The MC structures for UJ and UL were then established as shown in Figure 1 below.

From Table 1, UJ moves from S4 into S4, 5 times out of the 15 years. Thus, the frequency is 5/15. Thereafter, UJ moves to S3, and the frequency is 10/15 to comply with the principle of MCs, as the sum of frequencies on S4 must equal 1. Still from the table, it can be counted that S3 returns to S3 8 times, thus 8/15 as a frequency.

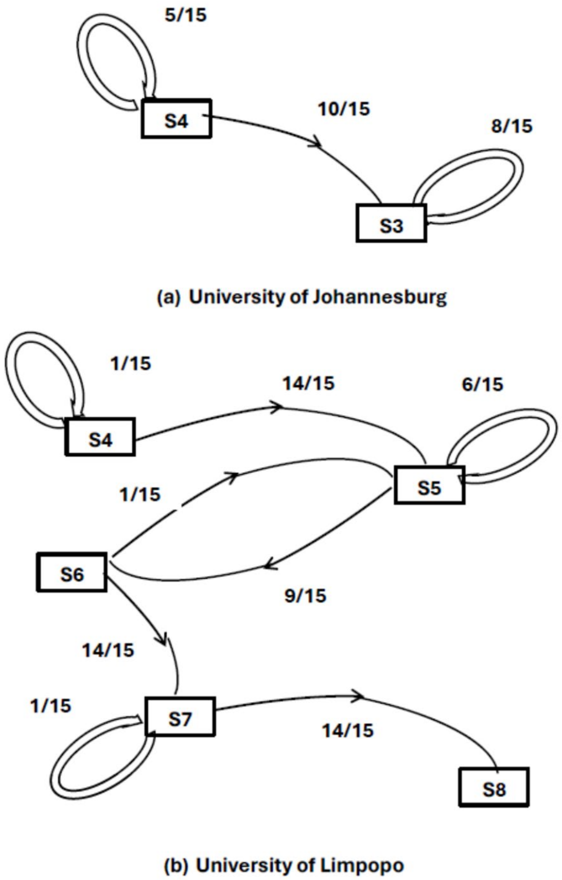

Fig. 1. (a), (b): Markov Chain Structures for UJ and UL.

The transition probability matrices for UL and UJ were:
For UL,

$$\begin{bmatrix} 0.06 & 0.933 & 0 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0.067 & 0 & 0.933 & 0 \\ 0 & 0 & 0 & 0.067 & 0.933 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore, their equations are given by the following system:

$$\begin{cases} 0.067y_1 + 0.933y_2 = y_1 \\ 0.400y_2 + 0.933y_4 = y_2 \\ 0.067y_2 + 0.933y_4 = y_3 \\ 0.067y_4 + 0.933y_5 = y_4 \\ y_1 + y_2 + y_3 + y_4 + y_5 = 1 \end{cases}$$

For UJ, the probability transition matrix is as follows:

$$\begin{bmatrix} 1/3 & 2/3 \\ 0 & 8/15 \end{bmatrix}$$

and the resulting system of equations:

$$\begin{cases} 0.533y_1 = y_1 \\ y_1 + y_2 = 1 \end{cases}$$

The solution to this problem where the steady state is reached is shown below.

For UJ, state 2 = 100% probability, could mean that this university could likely remain in this position, as steady state, based on the data provided. The data revealed that UJ has remained in this state for quite some time. And for UL, the future probability distribution for the different states was (0.892, 0.035, 0.035, 0.035, 0.01). More weight is on S_4, hence there could be high probability the institution being dominated by this state in the future, whereas S_8 is likely negligible.

5. CONCLUSIONS

This study has demonstrated the application of the Markov chain technique for evaluating the global performance and ranking of selected universities in South Africa, namely, the University of Johannesbourg (UJ) and the University of Limpopo (UL). The two universities were ranked by Scimago as middle and bottom performing universities in the global rankings.

Eight different states were defined to carry out the estimation of the steady state of the Markov Chain to predict theoretically the future state of their rankings. These results showed that UJ had 2 states with probability distributions of (0, 1) respectively, whereas UL had 5 states with probability distribution of (0.892, 0.035, 0.035, 0.035, 0.01).

On one hand, for UJ, state S4 was found to have 100% probability. This could mean that the university could likely remain in this position, as steady state, based on the data provided.

On the other hand, for UL, the future probability distribution for the different states was (0.892, 0.035, 0.035, 0.035, 0.01). More weight is on S4, hence there could be high probability the institution being dominated by this state in the future, whereas S8 is likely negligible.

By analysing the transition probabilities and steady states of university rankings, valuable insights have been gained into the stability and volatility of these rankings.

The findings of this research highlight the potential of the Markov chain approach in predicting future rankings and providing evidence-based decision-making tools for stakeholders in the higher education sector. The analysis of distinct probability distributions and steady states for middle performing and bottom performing universities has shed light on the dynamics of rankings and the potential for fluctuations in performance.

The practical implications of this study are significant. The knowledge gained from this research can inform strategic planning, resource allocation, and improvement initiatives for universities in South Africa. By understanding the stability and volatility of rankings, institutions can identify areas of improvement, benchmark against their peers, and develop strategies to enhance their overall performance.

Furthermore, the application of the Markov chain technique in evaluating university rankings can contribute to the broader field of higher education management. It provides a quantitative framework for assessing performance trends, making informed decisions, and fostering excellence in academic institutions.

While this study focused on a specific sample of institutions, the methodology and findings can be replicated for a broader range of universities in South Africa and extended to additional ranking metrics. Further research can build upon this study to deepen our understanding of the factors influencing rankings and explore more advanced modelling techniques.

6. ACKNOWLEDGMENT

Professor Akeem Ayinde Raheem (Ladoke Akintola University of Technology) and Dr. Michel Kalenga Wakalenga (University of Johannesburg) are acknowledged as peer reviewers of this paper.

7. REFERENCES

[1] Marhl M, Pausits A., “Third mission indicators for new ranking methodologies. Evaluation in Higher Education”. 2011 Jun 1;5(1):43-64.

[2] UNESCO CEPES, European Centre for Higher Education. Higher education ranking systems and methodologies: how they work, what they do. Available online: http://www.cepes.ro/hed/policy/ranking.htm [13 January 2006].

[3] M. Torres-Samuel, C.L. Vásquez, A. Viloria, N. Varela, L. Hernández-Fernandez, and R. Portillo-Medina, “Analysis of patterns in the university world rankings webometrics, Shanghai, QS and SIR-SCimago: Case Latin America”. In Data Mining and Big Data: Third International Conference, DMBD 2018, Shanghai, China, June 17–22, 2018, Proceedings 3 (pp. 188-199). Springer International Publishing.

[4] Shanghai Jiao Tong University. The Academic Ranking of World Universities (ARWU), Shanghai Jiao Tong University Ranking. Available : http://www.shanghairanking.com/ARWU2017.html

[5] QS Quacquarelli Symonds. The QS World University Rankings. Available : http://www.qs.com/world-university-rankings-2018/

[6] SCIMAGO INSTITUTIONS RANKINGS, Available: https://www.scimagoir.com/rankings.php?sector=Higher+educ.&country=ZAF&ranking=Research&year=2013

[7] J.F. Molinari, and A. Molinari, “A new methodology for ranking scientific institutions”. Scientometrics, Vol. 75, No. 1 (2008) 163–174. DOI: 10.1007/s11192-007-1853-2

[8] G.S. Walde, “Triple absorbing Markov chain model to study the flow of higher education students.” Journal of Physics: Conference Series 1176 042066, pp.1-8, 2019

[9] Nyandwaki, M.J. and Kennedy, G., 2016. “Statistical Modelling of the Kenyan Secondary School Students Enrolment: An Application of the Markov Chain Model”. IOSR Journal of Mathematics, Vol. 12, Issue 2 Ver. II, pp. 11-18.

[10] S.W. Gikie, S. Karoki, and S. Kinyanjui, “Evaluation of the impact of the free secondary education on students’ progression using Markov chain model: a case study of Kiambaa sub-county, Kiambu county”. Bulletin of Mathematics and Statistics research, Vol.8, Issue 2, pp. 17-31, 2020.

[11] M.G. Voskoglou, “Smart Learning Systems: A Markov Chain Approach”. Egyptian Computer Science Journal Vol. 45 (2), pp. 24-29.

[12] C.M. Grinstead, and J.L. Snell, “Introduction to probability”. American Mathematical Soc. 1997