Visitor Navigation Pattern Prediction Using Transition Matrix Compression

Ei Theint Theint Thu, Aye Aye Nyein, Hlaing Htake Khaung Tin

Faculty of Information Science,University of Computer Studies, Hinthada, Myanmar

***Abstract*: - As an increasing number of cities consists of an increasing number of visiting places, it is more difficult for the visitors to consider. Meanwhile, the system tries to introduce recommendation features to their visitors. The main aim of this paper is to only implement visitor navigation pattern prediction system in Myanmar. The paper uses traveling paths that assist visitors to navigate the visiting places based on the past visitor's behavior. To cluster the paths with similar transition behavior and compress the transition matrix to an best size for efficient probability calculation in paths, transition probability matrix compression has been used. In this paper, Visitor Navigation Pattern Prediction Using Transition Matrix Compression is developed. It uses data mining techniques for recommending a visitor which (next)paths is closely the most popular paths in Myanmar. By looking at the traveling paths in the organization, the system can know the popular paths(places).**

I. INTRODUCTION

Myanmar is one of the most interesting countries to explore in Asia. The Travels and Tours is a Myanmar-based travel agency. In other words, traveling to Myanmar is very similar to traveling back and forth to other geographical areas, not just one trip. Travel is the movement of people between distant geographical locations. It can be done by foot, bicycle, train, boat, bus, airplane and ship. Travelers may use such as bus, trains, airplanes and ships. Among them, they may travel by bus [4].

The large number of cities on many visiting places in Myanmar has raised. In Myanmar, visitors often have the first navigational question such as where can I go? The organization can advise you the popular paths in there. The system refers to tracking the visitor's past behavior. By gathering increasing amounts of visitor's information in organization, it can predict the next visiting places [4].

By viewing the visitor's navigation in a Travels and Tours company, Directed Graph can construct for path forecast created on past visitors visit behavior recorded in the company's documentation. The cities to be visited by a visitor in the forthcoming are strong-minded by visiting history in the company. Directed Graph consists of nodes representing cities, directed lines representing paths and weights on the directed lines representing the numbers of traversals on the paths. By counting the weights on the directed lines as past visitor's implicit feedback of their preferences in the paths, Directed Graph can be used to calculate a transition probability matrix. To cluster the cities with similar transition performances together to get a compact transition matrix, transition probability matrix compression is used. The transition matrix compression can make the paths prediction more efficient [1][2].

II. RELATED WORK

The system will evaluate to generate in visitor's visits from organization's visitor data and assist visitor's navigation. The paths pattern forecast is dynamic to response the varying interests of the travelers.

A transition matrix compression has been used to calculate based on transition's behaviors states in the matrix. Transition matrices calculated from system can be compressed for smaller state spaces. The Directed Graph Theory is used to measure the transition similarities between cities in this work and compress the probability transition matrix to an optimal size for well-organized path prediction.

Forward Path method is maximal to recover the accuracy of path prediction results by elimination the effect of backward visits by visitors. The path prediction calculation has been improved by taking into account more steps in the future to deliver more insight into the future. Travel recommender system includes traveling presentation and traveling navigation support. The system can be viewed as a kind of traveling presentation to help visitor's visiting pattern [1].

III. THEORITICAL BACKGROUND

Data mining techniques that aim at learning and extracting hidden information in data stored on the organization. Another important purpose of data mining is to deliver a mechanism to make the data access more powerfully and effectively. The organization owner can take advantage of data mining technique to gain insights about the visitor behavior when visiting the places and use the acquired knowledge to improve the design of the organization [3].

IV. DIRECTED GRAPH THEORY

4.1 Construct Directed Graph

The directed line information contained in the organization to construct a directed line structure, called a directed graph. The directed graph consists of directed line from city to city. In this method, only the paths of the requested city and the corresponding referrer are used for directed graph construction [1]. It has a simplified set

WLs={(r,p,w)}

where,

w= the number of traversals form r to p

r,p= paths of the referrer and the requested city respectively.

4.2 Transition Probability Matrix

One-step transition probability from city i to city j can be calculated by using directed graph and by considering the similarity between a directed graph [1].

$$P_{i,j} = \frac{W_{i,j}}{\sum_k W_{i,k}}$$

where,

$P_{i,j}$ =one-step transition probability from city i to city j.

$W_{i,j}$ =weight on the directed line from i to j.

$W_{i,k}$ =weight on the directed line from i to k.

4.3 Transition Matrix Compression

Transition Matrix Compression is used to compress Transition Probability Matrix. The time occupied by compression can be reimbursed by all subsequent possibility computations for path prediction [1].

$Sim_{i,j} = Sim_{i,j}$ (out − directed line) × $Sim_{i,j}$ (in − directed line)

$Sim_{i,j}$(out – directed line) $= \sum_y \propto_{i,j}(y)$

$\beta_{i,j}(x) = \frac{|m_i \times P_{x,j} - m_j \times P_{x,i}|}{m_i + m_j} \propto_{i,j}(y) = |P_{i,y} - P_{j,y}| m_i = \sum_l P_{l,i}, m_j = \sum_l P_{l,j}$

where,

m_i and m_j = the sums of possibilities on the in−directed line of city i and j respectively.

$Sim_{i,j}$ (out – directed line) = the sum of the out−directed line probability difference between i and j.

$Sim_{i,j}$ (in – directed line) = the sum of the in−directed line probability difference between i and j.

$Sim_{i,j}$= product of their in−directed line and out−directed line similarities.

$$P_{k,ivj} = P_{k,i} + P_{k,j}$$

$$P_{ivj,k} = \frac{m_i \times P_{i,k} + m_j \times P_{j,k}}{m_i + m_j}$$

where,

P_{ivj}= transition probabilities between the new state i v j.

P_k = transition probabilities of the remaining state k in the transition matrix.

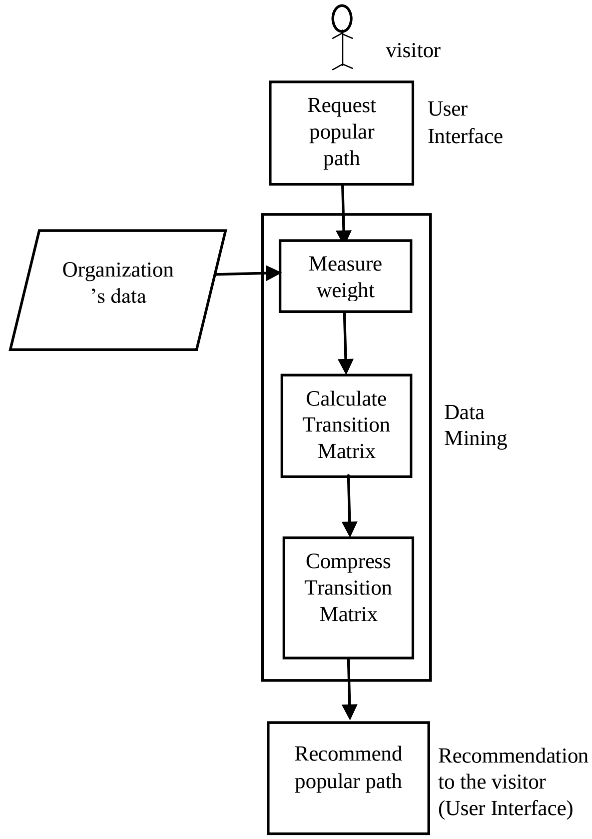

Figure 1. The flowchart of system

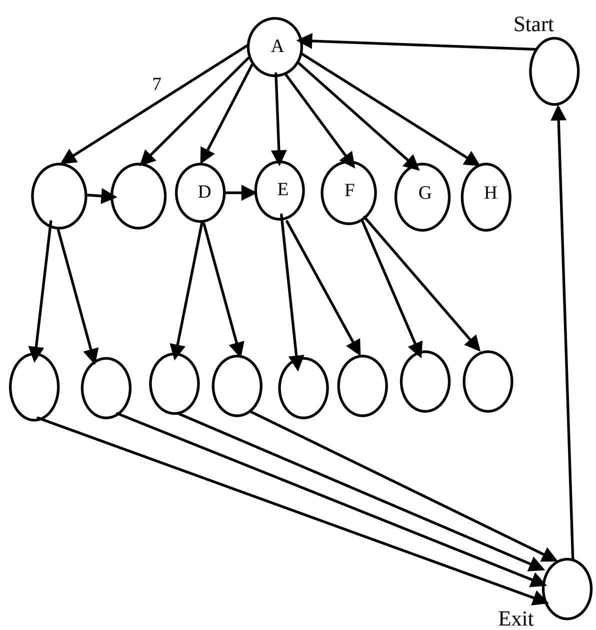

Figure 2. A directed graph constructed from organization's data

Figure 2 shows a directed graph that is constructed using organization's data. This organization's data is a collection of cities. In this approach, nodes suppose cities. The edges represent directed line with from city to city and weights represent the number of traversals on the directed line.

No	Car Name	Travel Time	From City	To City	The number of Travelers(Visitors)
1	001	5 AM	A	B	50
2	002	1 AM	A	D	50
3	003	2 PM	A	E	50
4	004	10AM	B	C	50
5	005	3 AM	D	E	50
6	006	8 PM	B	I	50
7	007	5 PM	B	J	50
8	008	10 PM	D	K	50
9	009	2 AM	D	L	50
10	010	4 AM	A	C	50
11	001	5 AM	A	-	-

Figure 3. Weight measure using organization's data

Figure 3 shows Weight measure that is obtained by using organization's data. This organization's data is used to construct directed graph using weight matrix.

	A	B	C	D	E	I	J	K	L	.	Z	Exit	Start
A		7	2	5	3								
B			1			3	5						
C													
D					3			7	1				
E													
I												1	
J												1	
K												1	
L												1	
:													
Z													
Exit													1
Start	1												

Figure 4. Construct directed graph from Weight measure

Figure 4 shows Directed Graph that is constructed from Weight measure. Weight measure is used to count the number of cities in organization.

	A	B	C	D	E		I	J	K	L	.	Z	Exit	Start
A		0.41	0.11	0.29	0.18									
B			0.11				0.33	0.56						
C														
D					0.27				0.63	0.09				
E														
I													1.0	
J													1.0	
K													1.0	
L													1.0	
:														
Z														
Exit														1.0
Start	1.0													

Figure 5. Calculate Transition Probability Matrix

Figure 5 shows Transition Probability Matrix by calculating from the Directed Graph. In a transition probability matrix, row icovers one-step transition possibilities from i to all conditions. Row i sums up to 1.0. Column icovers one-step transition possibilities from all conditions to i.

	A	B	C	D	E	I	J	K	L	.	Z	Exit	Start
A	0.0												
B	1.03	0.0											
C	0.36	0.15	0.0										
D	0.73	0.0	0.12	0.0									
E	0.57	0.3	0.0	0.2	0.0								
I	0.9	0.7	0.12	0.6	0.38	0.0							
J	1.4	0.9	0.15	0.76	0.6	0.0	0.0						
K	0.8	1.03	0.18	0.76	0.1	0.8	0.4	0.0					
L	0.3	0.3	0.12	0.7	0.06	0.07	0.07	0.0	0.0				
:										0.0			
Z											0.0		
Exit												0.0	
Start													0.0

Figure 6. Calculate Similarity Matrix

Figure 6 shows that the Similarity Matrix is calculated from Transition Probability Matrix. Cities sharing more in-directed line, out-directed line and having equal weights will meet the similarity threshold. When the similarity is below a given threshold, the effect of compression on the transition behavior of the states will be controlled.

The compressed matrix for Transition Matrix is shown in Figure 7. The compressed matrix is thicker than the original transition matrix. This is the compression process for Transition Matrix Compression.

Compressed state D into state B(similarity 0.0)(states: B D)

Compressed state E into state C(similarity 0.0)(states: E C)

Compressed state J into state I(similarity 0.0)(states: IJ)

Compressed state L into state K(similarity 0.0)(states: K L)

Finished compression.

Have compressed 11 states to 7.

Figure 7 shows Transition Matrix Compression. Transition Matrix Compression can be used to compress transition probability matrix. States with similar transition behaviors are aggregated together to form new states. If the similarity is close to zero, the error resulted from compression is close to zero. The threshold value is indicated between 0.08 and 0.15 yielded good compression with minimal error for this compression process.

	A	(D,B)	(E,C)	(J,I)	(L,K)	…….	Z	Exit	Start
A		0.7	0.29	0.0	0.0				
(D,B)			0.2	0.5	0.7				
(E,C)				0.0	0.0				
(J,I)					0.0			1.0	
(L,K)								1.0	
: :									
Z									
Exit									1.0
Start	1.0								

Figure 7. Compress Transition Matrix

This is an example of recommendation of popular path for organization. The organization will begin to travel from Magway Division.

City A	Magway Division
City B and D	City A
City E and C	City A
City I and J	City B
City K and L	City D

City A = Magway

City B =Hinthada

City C = Yangon

City D = Lashio

City E = NayPyiTaw

City I = PaThein

City J = Mandalay

City K = MawLaMyaing

City L = PhaAnn

V. CONCLUSION

Clustering method in Directed Graph Theory has been proven very suitable for traveler's navigation in organization. This method helps in better understanding the visitor's navigation behavior and improving visitor's request serving (by decreasing the lengths in organization's navigation pathways). This means that the system will predictto go the most visitor's traveling places. Using this method, theinformation can be used to help visitors to discoverwanted information in organization effectively and efficiently. For these reason, visitors can be recommended by organization owners to visit popular places.

REFERENCES

[1]. Jianhan Zhu, Jun Hong, and John G. Hughes. "Using Markov Chains for Link Prediction in Adaptive Web Sites", pp.60-73.
[2]. Yu Ya Win, "Prefetching Based On Web User Clustering", PhD thesis, University of Computer Studies, Yangon.
[3]. https://www.Guru99.com
[4]. https://www.wikipedia.org.com