Modelling Customer Relationships as Hidden Markov Chains

Edesiri Bridget Nkemnole [1], **Ekene Nwaokoro** [1]

[1] *University of Lagos*
101017 University Road, Akoka, Lagos State, Nigeria

Abstract. Models in behavioural relationship marketing suggest that relations between the customer and the company change over time as a result of the continuous encounter. Some theoretical models have been put forward concerning relationship marketing, both from the standpoints of consumer behaviour and empirical modelling. In addition to these, this study proposes the hidden Markov model (HMM) as a potential tool for assessing customer relationships. Specifically, the HMM is submitted via the framework of a Markov chain model to classify customers relationship dynamics of a telecommunication service company by using an experimental data set. We develop and estimate an HMM to relate the unobservable relationship states to the observed buying behaviour of the customers giving an appropriate classification of the customers into the relationship states. By merely accounting for the functional and unobserved heterogeneity with a two-state hidden Markov model and taking estimation into account via an optimal estimation method, the empirical results not only demonstrate the value of the proposed model in assessing the dynamics of a customer relationship over time but also gives the optimal marketing-mixed strategies in different customer state.

Keywords: customer's relationship management; hidden Markov model; Markov chain; transition probability; estimation.

INTRODUCTION

In interactive marketing, the lifetime value of a customer is very vital and useful, especially given the fact that we are in the age of relationship marketing [27]. An essential attribute of service quality is the direct interaction with customers and having customer-oriented behaviours [29]. Customer relationship management (CRM) is a set of process and enabling systems supporting a business strategy to build long term, profitable relationship with the customers [23]. For [6] CRM is a customer-focused business strategy that creates and adds value to the company and its customers. The author [9] presents the usefulness of CRM for some managerial problems, chiefly the budgeting of marketing expenditures for customer acquisition. CRM is also used to allocate spending across media (mail vs telephone vs television) and programs (gift vs special price); it furthermore informs decisions for retaining existing customers [18]. Author [19] avers that its use helps firms to achieve a strategic competitive advantage. It had developed as an approach based on maintaining positive relationships with customers, increasing customer loyalty and expanding customer lifetime value [2].

Despite the wide adoption of CRM in the business world, the academic community has been lagging in developing models that could help businesses analyze transaction data to assess customer relationships and put forward a support system for marketing decisions.

Lately, marketing modellers started to address this gap by developing models of customer lifetime value [15, 28, 31]. However, far less attention has been given to modelling the dynamics of customer relationships and the effect of relationship marketing actions on customer-brand relationships and the customer's choice behaviour.

In addition to the many available tools to assess customer relationships, one useful tool is the hidden Markov model (HMM). HMMs are widely used in science [22], engineering [29] and many other areas [26] and have been successfully applied in engineering problems such as speech recognition [30]. Advanced models such as higher-order HMM with applications in DNA sequences can be found in [7]. Authors [33]

proposed a methodology based on Partially Observable Markov Decision Processes (POMDPs) with general state and action spaces, under a fractional discounted criterion, a model that combines the Hidden Markov Models with the Markov Decision Process by capturing both dynamics that depend on unobserved states and effects of decisions over time. However, not much work has been done in the field of marketing science.

In an interest in finding patterns in customer relationships and discovering hidden relationship states, this study focuses on the use of the HMM to capture the correct dynamic buying behaviour of customers to services provided by a telecommunication company in Nigeria, proceeding also to find their relationship state over some time. With the relationship states being a finite set which follows the markovian property, transitions between states are determined by the history of interactions between the customers and the brand.

Related Literature

Research in the area of relationship marketing has been emerging in the past decade, both from the consumer behaviour perspective [14] and the empirical modelling perspective [4, 28, 34]. Our focus will be on those relevant to this research.

Authors [17] applied the probabilistic method of Markov chain to the systematic study of brand switching and brand loyalty. After that this application developed slowly. Researchers tried to use it in other dimensions of CRM, such as analyzing and predicting unique customer behaviour [12, 29]. Authors [37] applied a Markov chain model in quality function deployment to analyze customer requirements. Their proposed approach provides a decision-maker to investigate and then satisfy both present and future customers' needs. There probabilistic nature of Markov chain has a crucial impression on decision making [25]. Authors [20] proposed a data mining model that considers e-customer activities through a discrete-time semi-Markov process. Researchers [21] use this modelling technique in a macro marketing decision level, applying the Markov decision process to forecast markets share and determine customer's decision pattern in the diverse circumstances that are common.

Authors [28] constructed and estimated a nonhomogenous HMM to model the transitions among hidden relationship states. They used HMM to overcome the problem of the unobserved state, as well as to describe a set of hidden conditions and changes between them; also, they translated these hidden states into observed customers' behaviours. HMM often encountered researchers with estimating the models' parameters. There are several estimation methods and algorithms, such as expectation-maximization (EM). Authors [36] used this algorithm to estimate HMM's parameters and utilized the Viterbi algorithm to find some path which was describing customers migration pattern in an online retailer. Also, [32] according to Bayesian rules obtained the conditional probability and calculated the Equation that was referred as likelihood function and then designed a classifier based on HMM for discovering which customer is loyal and which is not dedicated. Authors [10, 24] used HMM as a portfolio optimization technique. Their basic idea is to describe the essential movement of the stock price using HMM and to calculate the optimal portfolio using HMM's recursive algorithms. Authors [37] constructed a specific HMM for web browsing, which could prospect whether the users have the intention to purchase in real-time. It can be useful to be ready for customers' needs in the buying process and predicting repetitive customer attitude [5, 16, 29].

Model Development

The hidden Markov model. Although initially introduced and studied as far back as 1957 and early 1970, the recent popularity of statistical methods of HMM is not in question. An HMM is a bivariate discrete-time process $\{X_k, Y_k\}_{k \geq 0}$ where $\{X_k\}_{k \geq 0}$ is a homogeneous Markov chain which is not directly observed but can only be observed through $\{Y_k\}_{k \geq 0}$ that produce the sequence of observation. $\{Y_k\}_{k \geq 0}$ is a sequence of independent random variables such that the conditional distribution of Y_k solely depends on X_k [1]. The underlying Markov chain $\{X_k\}_{k \geq 0}$ is called the state.

An HMM is characterized by the following: N, the number of states in the model. We denote the state set as $S = \{S_1, S_2, \cdots, S_N\}$ and the state at time t as x_t, $x_t \in S$.

The state transition probability distribution $A = \{a_{ij}\}$ where
$$a_{ij} = p(x_t = S_j | x_{t-1} = S_i),$$
$$a_{ij} \geq 0, 1 \leq i, j \leq N$$

The observation probability distribution B. $Y_t \sim g(Y_t|X_t, B)$ $Y_t \sim f(Y_t \mid X_t, B)$ is observation density at time t, $B = \{b_j(k)\}$ gives the conditional probability distribution of each observation symbol within a given hidden state with the definition of

$b_i(k) = p\{Y_t \mid X_t = S_t\}$ denotes the kth observation symbol per state.

The initial state distribution $\pi = \{\pi_i\}$, $\pi_i = p(x_i = s_i), 1 \le i \le N$. For convenience, we used the compact notation $\lambda = (A, B, \pi)$ to indicate the complete parameter set of the model. Given the form of the HMMs, the goal is to find the best model for a given time series through optimally adjusting model parameters $\lambda = (A, B, \pi)$ [1].

Hidden Markov Model of Customer Relationships

The HMM presented here is an individual-level model of buying behaviour. We consider a set of customers, each of whom is involved in repeated interactions of the service provided by a telecommunication company. Accordingly, we define a set of hidden relationship states, which differ for the strength of the relationship between the customer and the brand and the conditional likelihood of choice given the relationship state. A first-order random walk model characterizes the stochastic relationship-states model. The transitions between the Markovian states are probabilistically determined and are affected by relationship encounters [28].

The proposed HMM consists of three main components:

(1) the initial state distribution (π),

(2) a sequence of random walk transitions (Q) that is express in a probabilistic manner, the likelihood that the series of customer-brand interactions in the previous period were strong enough to transition the customer to an adjacent state,

(3) a vector of state-dependent choice probabilities given the relationship states (m). The possibility of observing a sequence of choices is defined [28] by:

$$P_i(Y_{i1} = y_{i1}, \cdots, Y_{it} = y_{it}) =$$
$$\sum_{s_1=1}^{N}\sum_{s_2=1}^{N}\cdots\sum_{s_T}^{N}\left[P(S_{i1} = s_1)\prod_{t=2}^{T} P(S_{it} = s_t \mid S_{it-1} = s_{t-1})\prod_{t-1}^{T} P(Y_{it} = y_{it} \mid S_{it} = s_t)\right] \quad (1)$$

where,

S_{it} is customer i's state at time t in a Markov process with N states, and Y_{it} is customer i's choice at time t.

Following Equation (1) the three components of the HMM are defined as follows:

1) The initial state distribution – the probability that customer i in state s at the time one can be defined as:

$P(S_{i1} = s) = \pi_{is}$.

2) The transitions – the probability that a customer transitions from the state s_{t-1} at the time $t-1$ to state s_t at time t can be defined as: $P(S_{it} = s_t \mid S_{it-1} = s_{t-1}) = q_{is_{t-1}s_t}$. The transition matrix is defined as,

State at t-1	State at t					
	1	2	3	...	N-1	N
1	q_{11}	q_{12}	0	...	0	0
2	q_{21}	q_{22}	q_{23}	...	0	0
⋮	⋮	⋮	⋮	⋱	⋮	⋮
N	0	0	0	...	q_{NN-1}	q_{NN}

where $q_{itss'} = P(S_{it} = s' \mid S_{it-1} = s)$ is the conditional probability that individual i moves to state s' at time t given in state s at the time $t-1$ and where $0 \le q_{itss'} \le 1$ $\forall s, s'$ and $\sum_{s'} q_{itss'} = 1$.

Following the random walk transition, all terms that are two states or more away from the diagonal are set to zero. Each one of the matrix elements represents a probability of change to an adjacent state or a likelihood of staying in the current state.

3) The state-dependent choice shows the probability that the customer will choose the product at time t conditioned on her state is defined [28] as:

$P(Y_{it} = 1 \mid S_{it} = s_t) = m_{is_t}$

Most times, the model parameters are obtained by using the EM algorithm, which results in a

nonlinear optimization problem [7, 22, 28]. Still, the EM algorithm does not work for the HMM here, so an optimal estimation method is used for obtaining the transition probabilities among the hidden states which guarantee convergence and global optimum.

Estimation procedure

Here, the work presents the procedure used to estimate the model described above. In choosing the estimation procedure, we focus on properly accounting for observed and unobserved heterogeneity. An optimal estimation procedure, according to [8], is used to estimate the model. The parameter estimation method is an HMM with two hidden states and two observable states.

The number of m hidden states and the stationary distribution of the hidden states is given by:

$$\alpha_t = (\alpha_1, \alpha_2, \cdots, \alpha_m)$$

The number of n observable state and the stationary distribution of the observable states is

(pi1, pi2, . . . , pin).

Given an observed sequence of the observable states, the occurrences of each state in the series are calculated and hence the experimental distribution q. If the hidden stats are ignored, the visual states follow the one-step transition probability matrix:

$$\overline{P}_2 = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_2 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_m & \alpha_m & \cdots & \alpha_m \end{pmatrix} \begin{pmatrix} p_{11} & p_{21} & \cdots & p_{m1} \\ p_{12} & p_{22} & \cdots & p_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ p_{1n} & p_{2n} & \cdots & p_{mn} \end{pmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} p$$

where

$$p = \left(\sum_{k=1}^{m} \alpha_k p_{k1}, \sum_{k=1}^{m} \alpha_k p_{k2}, \cdots, \sum_{k=1}^{m} \alpha_k p_{kn} \right)^T$$

$$\overline{P}_2 p = p \text{ and } \sum_{k=1}^{m} p_k = 1$$

The vector p is the stationary probability distribution of $\overline{P}_2$. Therefore, the transition probabilities of the hidden states $\alpha_t = (\alpha_1, \alpha_2, \cdots, \alpha_m)$

are obtained by solving $\min_{\alpha} \| p - q \|_2^2$

subject to $\sum_{k=1}^{m} \alpha_k = 1$ and $\alpha_k \geq 0$

This is a standard constrained least square problem when $\|.\|$ is chosen to be the square of the L2-norm. When $\|.\|$ is determined to be L1-norm or the La? z-norm, the resulting optimization problem are transformed into a linear programming problem.

For the model having two hidden states (m = 2),

$$\overline{P}_2 = \begin{pmatrix} \alpha & 1-\alpha \\ \alpha & 1-\alpha \\ \vdots & \vdots \\ \alpha & 1-\alpha \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \end{pmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} p$$

where

$$p = (\alpha \, p_{11} + (1-\alpha) p_{21}, \alpha \, p_{12} +$$
$$(1-\alpha) p_{22}, \cdots, \alpha \, p_{1n} + (1-\alpha) p_{2n}$$

$$\overline{P}_2 p = p \text{ and } \sum_{k=1}^{n} p_i = 1$$

and therefore p is the steady-state probability distribution.

Given the observed distribution q of the observable states, then α is estimated by the following minimization problem [8]:

$$\min_{\alpha} \| p - q \|_2^2$$

subject to $0 \leq \alpha \leq 1$ $0 \leq \alpha \leq 1$ or equivalently

$$\min_{0 \leq \alpha \leq 1} \sum_{i=1}^{n} \{\alpha \, p_{1k} + (1-\alpha) p_{2k} - qk\}^2 \quad \min_{0 \leq \alpha \leq 1}$$

$$\tau = \frac{\sum_{j=1}^{n} (q_j - p_{2j})(p_{1j} - p_{2j})}{\sum_{j=1}^{n} (p_{1j} - p_{2j})^2}$$

Let

Then the optimal value of α is given as follows:

$$\alpha = \begin{cases} 0 & if\ \tau \leq 0 \\ \tau & 0 < \tau < 1 \\ 1 & if\ \tau \geq 1 \end{cases}$$

noting that

$$\tau = \frac{<(q-p_2),(p_1-p_2)>}{<(p_1-p_2),(p_1-p_2)>} = \frac{\left\| q-p^2 \right\|_2 \cos(\theta)}{\left\| p_1 - p^2 \right\|_2}$$

Here, $<...,\cdots>$ is the standard inner product on the vector space R_n

$p_1 = (p_{11}, p_{11}, \cdots, p_{1n})^T$ and $p_2 = (p_{21}, p_{22}, \cdots, p_{2n})^T$

θ is the angle between the vectors $(q - p_2)$ and $(p_1 - p_2)$.

Empirical Results and Discussion

Data Employed. The data records two services offered by a telecommunication company in Lagos State, Nigeria and represented by 'V' (voice) and 'D' (data). From the customer database, information is obtained on the expenditure distribution of 93 randomly chosen customers on these services for four months. The customers are classified into two groups based on responses observed. The data is split into two, which is the 'train data' to build an HMM and the 'test data' to validate the constructed HMM. Given the expenditure distribution of a customer observed from the customer's sequence of purchases, and the average expenditure distributions, the estimation procedure above is used to compute the transition probability α in the hidden states. The parameter estimation procedure is done using the R statistical software.

Through observing the responses of customers, customers are classified into two groups, among which 60 customers being loyal (Group 1) and the other 33 customers not-loyal (Group 2). For the train data, two-third of the customers are randomly selected from Group 1 and Group 2. This implies that for the train data, 40 customers are randomly selected (these customers are labelled in the first 40 customers in Table 1) from Group 1, and 22 customers are selected from Group 2. The remaining 20 customers from Group 1 and 11

customers from Group 2 are used for validating the constructed HMM. An HMM of two observable states (V and D) and two hidden forms (Group 1 and Group 2) is built. From the information of the customers in Group 1 and Group 2, the probability distributions of customers' purchases and the emission probability distributions for both groups are computed in Tables 1 and 2.

Table 1 – Probability distribution of observations (Two-third of the data used to build the HMM)

	Customer	V	D		Customer	V	D
1	3	0.96	0.04	21	52	1.00	0.00
2	8	0.99	0.01	22	55	1.00	0.00
3	12	0.81	0.19	23	56	0.85	0.15
4	19	0.92	0.08	24	58	0.96	0.04
5	20	1.00	0.00	25	59	1.00	0.00
6	23	0.94	0.06	26	61	0.93	0.07
7	24	1.00	0.00	27	65	0.95	0.05
8	25	0.88	0.12	28	66	0.88	0.12
9	26	1.00	0.00	29	67	0.94	0.06
10	29	0.98	0.02	30	68	0.98	0.02
11	32	0.98	0.02	31	69	1.00	0.00
12	33	0.99	0.01	32	71	1.00	0.00
13	34	0.91	0,09	33	79	1.00	0.00
14	36	0.98	0.02	34	86	0.86	0.14
15	37	0.91	0.09	35	87	0.98	0.02
16	38	0.96	0.04	36	90	1.00	0.00
17	39	0.95	0.05	37	92	0.99	0.01
18	42	0.90	0.10	38	93	1.00	0.00
19	49	0.80	0.20	39	95	0.84	0.16
20	51	0.97	0.03	40	96	1.00	0.00
41	1	0.17	0.83	52	60	0.63	0.37
42	2	0.28	0.72	53	73	0.51	0.49
43	6	0.43	0.57	54	74	0.04	0.96
44	7	0.72	0.28	55	81	0.74	0.26
45	11	0.76	0.24	56	82	0.08	0.92
46	17	0.78	0.22	57	83	0.08	0.92
47	30	0.76	0.24	58	88	0.39	0.61
48	46	0.65	0.35	59	94	0.57	0.43
49	48	0.44	0.56	60	97	0.29	0.71
50	50	0.59	0.41	61	99	0.53	0.47
51	54	0.7	0.3	62	100	0.56	0.44

Table 2 – Emission probabilities

Group	V	D
1	0.95	0.06
2	0.49	0.51

To find the corresponding optimal state given the observation sequence, the transition probability

α in the hidden states is computed with the function 'optimize' after defining a process α which is

$$\min_{0\le\alpha\le1} \sum_{i=1}^{n}\left\{\alpha\, p_{1k} + (1-\alpha)p_{2k} - q_k\right\}^2$$

where p is the stationary distribution,

$p = $ 0.94975α + 0.48636(1-α) 0.05025α + 0.51363(1-α)

and q is the probability distribution of the observation sequence of each customer.

Table 3 – Train data - Transition probability α in the hidden states

	Customer	V	D	α		Customer	V	D	α
1	3	0.96	0.04	1.00	21	52	1.00	0.00	1.00
2	8	0.99	0.01	1.00	22	55	1.00	0.00	1.00
3	12	0.81	0.19	0.70	23	56	0.85	0.15	0.78
4	19	0.92	0.08	0.94	24	58	0.96	0.04	1.00
5	20	1.00	0.00	1.00	25	59	1.00	0.00	1.00
6	23	0.94	0.06	0.99	26	61	0.93	0.07	0.96
7	24	1.00	0.00	1.00	27	65	0.95	0.05	1.00
8	25	0.88	0.12	0.85	28	66	0.88	0.12	0.85
9	26	1.00	0.00	1.00	29	67	0.94	0.06	0.98
10	29	0.98	0.02	1.00	30	68	0.98	0.02	1.00
11	32	0.98	0.02	1.00	31	69	1.00	0.00	1.00
12	33	0.99	0.01	1.00	32	71	1.00	0.00	1.00
13	34	0.91	0,09	0.91	33	79	1.00	0.00	1.00
14	36	0.98	0.02	1.00	34	86	0.86	0.14	0.81
15	37	0.91	0.09	0.91	35	87	0.98	0.02	1.00
16	38	0.96	0.04	1.00	36	90	1.00	0.00	1.00
17	39	0.95	0.05	1.00	37	92	0.99	0.01	1.00
18	42	0.90	0.10	0.89	38	93	1.00	0.00	1.00
19	49	0.80	0.20	0.68	39	95	0.84	0.16	0.76
20	51	0.97	0.03	1.00	40	96	1.00	0.00	1.00
41	1	0.17	0.83	0.00	52	60	0.63	0.37	0.31
42	2	0.28	0.72	0.00	53	73	0.51	0.49	0.05
43	6	0.43	0.57	0.00	54	74	0.04	0.96	0.00
44	7	0.72	0.28	0.50	55	81	0.74	0.26	0.55
45	11	0.76	0.24	0.59	56	82	0.08	0.92	0.00
46	17	0.78	0.22	0.63	57	83	0.08	0.92	0.00
47	30	0.76	0.24	0.59	58	88	0.39	0.61	0.00
48	46	0.65	0.35	0.35	59	94	0.57	0.43	0.18
49	48	0.44	0.56	0.00	60	97	0.29	0.71	0.00
50	50	0.59	0.41	0.22	61	99	0.53	0.47	0.09
51	54	0.7	0.3	0.46	62	100	0.56	0.44	0.16

This value of α was used to classify a customer: if α is close to 1, the customer is likely to be a loyal customer, and if α is close to 0 then the customer is expected to be a not-loyal customer. Based on the values of α for all the 62 customers listed in table III, the values of α of the first 40 customers falls under Group 1 as it lies in the interval [0.68, 1.00]. In contrast, the importance of α between [0.00, 0.63], of the remaining customers, falls under Group 2. Thus, the two groups of customers are separated by setting the cutoff value β to be 0.67.

Furthermore, the decision rule is that a customer is classified to Group 1 if $\alpha \ge \beta$, otherwise the customer is classified to Group 2.

Table 4 shows the HMM and the decision rule applied to the "test data" among which contains 20 customers belonging to Group 1 and 11 belonging to Group 2.

Table 4 – The remaining one-third of the data for the validation of the HMM

	Customer	V	D	α		Customer	V	D	α
1	9	1.00	0.00	1.00	11	44	0.97	0.03	1.00
2	10	1.00	0.00	1.00	12	45	0.94	0.06	0.98
3	13	0.85	0.15	0.78	13	53	0.92	0.08	0.94
4	16	0.90	0.10	0.89	14	57	0.96	0.04	1.00
5	22	1.00	0.00	1.00	15	63	0.92	0.08	0.94
6	27	0.97	0.03	1.00	16	64	0.97	0.03	1.00
7	28	0.95	0.05	1.00	17	72	0.83	0.17	0.74
8	31	1.00	0.00	1.00	18	77	0.86	0.14	0.81
9	35	0.80	0.20	0.68	19	84	0.85	0.15	0.78
10	41	0.98	0.02	1.00	20	98	0.98	0.02	1.00
21	4	0.64	0.36	0.33	27	47	0.51	0.49	0.05
22	5	0.35	0.65	0.00	28	62	0.47	0.53	0.00
23	18	0.78	0.22	0.63	29	70	0.58	0.42	0.20
24	21	00.57	0.43	0.18	30	76	0.58	0.42	0.20
25	40	0.63	0.37	0.31	31	78	0.77	0.23	0.61
26	43	0.77	0.23	0.61					

Table 5 – The transition probability

States	1	2
1	0.94	0.06
2	0.24	0.76

Table 6 – Stationary distribution

States	1	2
1	0.80	0.20
2	0.80	0.20

Tables 5, 6 are the transition probabilities and the long-term behaviour of the hidden Markov chain. The pattern of the one-step transitions of the customers based on their purchases with cutoff 0.67 separating the two related groups shows a higher percentage of loyal customers and a frequent movement from loyal customers to loyal customers.

Empirical Findings

Based on the analysis, the emission probability of the observations showed a 95% probability of a loyal customer making a purchase in service 'V' and a 49% probability of an unloyal customer purchasing service 'V' (see Table II). In measuring the loyalty of the customers to this service, the transition probability showed a 94% probability of a loyal customer remaining loyal and a 6% probability of a loyal customer becoming unloyal.

It also showed a 76% probability of an unloyal customer remaining unloyal and 24% of an unloyal customer becoming loyal (see Table 5). This led to a higher possibility of an unloyal customer becoming reliable than a loyal customer becoming unloyal. Loyal and unloyal customers were separated using cutoff value 0.67 which implied that a customer is loyal if $\alpha \geq 0.67$ and is unloyal if $\alpha \leq 0.67$. A likely future state for a loyal or unloyal customer is predicted with an 80% possibility of a loyal customer remaining loyal and a 20% possibility of an unloyal customer remaining unloyal (see Table 6).

CONCLUSION

This study is based on modelling customer relationships as hidden Markov chains. The importance of good customer relationships to any company cannot be over-emphasized. Every business requires customers, but it is not enough to have customers, companies must be able to manage their customers and build good relationships with them. In this study, a Hidden Markov model (HMM) is utilized to study customer's relationship dynamics of a telecommunication service company using data on purchases made by these customers. The HMM developed in this research helped to relate the relationship-states Markov chain process to the observed buying behaviour by identifying customers purchases for two services provided and then obtained the hidden relationship states of customers through the observed

sequence of investments made over some time. The customers were further classified as loyal or not based on these relationship states. By accounting for the functional and unobserved relationship with a two-state hidden Markov model and taking estimation into account via an optimal estimation method for the transition probabilities, the empirical results show that the model not only takes care of the dynamic effect of interactions between the customer and the brand on the customer's relationship state but also helps to establish a proficient structure for customer relationship management.

REFERENCES

1. Ahani, E., & Abass, O. (2010). A Sequential Monte Carlo Approach of Online Stock Market Prediction using Hidden Markov Model. *Journal of Modern Mathematics and Statistics, 4*(2), 73–77.

2. Ahnn, Y. J., Kim, K. S., & Han, S. K. (2003). On the Design Concepts for CRM Systems. *Industrial Management and Data Systems, 103*(5), 32-43.

3. Bhar R., & Hamori S. (2004). *Hidden Markov Models: Application to Financial Economics*. Dordrecht: Kluwer Academic Publishers.

4. Bolton, R. N., & Lemon, K. (1999). A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent of Satisfaction. *Journal of Marketing Research, 36*(2), 171–186.

5. Burez, J., & Van Den Poel., D. (2007). CRM at a pay-TV company, using Analytical models to reduce customer attrition by targeted marketing for subscription services. Expert Systems with Application, 32 (2), 277-288.

6. Chelmeta, R. (2006). Methodology for Customer Relationship Management. *The Journal of Systems and Softwares, 79*, 1015–1024.

7. Ching, W. K., Fung, E. S., & Ng, M. K. (2003). Higher-Order Hidden Markov Models with Applications to DNA Sequences. *Lecture Notes in Computer Science*, 535–539.

8. Ching, W. K., Wong, K. (2004). Hidden Markov models and their applications to customer relationship management. *IMA Journal of Management Mathematics , 15*, 13–24.

9. Courtheaux, R. J. (1986). *Database Marketing: Developing a Profitable Mailing Plan*.

10. Elliot, R., & Hinz, J. (2002). Portfolio Optimization, Hidden Markov models and Technical Analysis of P&F charts. *International Journal of Theoretical and Applied Finance, 5*(4), 385-399.

11. Fan Wu, I-Han Chiu, & Jiunn-Rong Lin. (2005). Prediction of the intention of purchase of the user surfing on the Web using hidden Markov model. *Proceedings of ICSSSM '05. 2005 International Conference on Services Systems and Services Management, 2005*.

12. Feinberg, F. M., Kahn, E. B., & McAlister, L. (1994). Implication and Relative Fit of Several First-Order Markov models of Customer Variety Seeking. *European Journal of Operational Research, 76*(2), 309-320.

13. Ford, D. (1980). The Development of Buyer-Seller Relationships in Industrial Markets. *European Journal of Marketing, 14*(5/6), 339–353.

14. Fournier, S. (1998). Consumers and Their Brands: Developing Relationship Theory in Consumer Research. *Journal of Consumer Research, 3*, 343-373.

15. Gupta, S., Lehmann, D. R., & Stuart, J. A. (2004). Valuing Customers. *Journal of Marketing Research, 41*(4), 417-418.

16. Hadden, J., Tiwari, A. Roy, R., & Ruta, D. (2007). Computer Assited Customer Churn Management: State-of-the-art and Future Trends. *Computers and Operations Research, 34*(10), 2902-2917.

17. Harary, F., & Lipstein, B. (1962): The dynamics of Brand Loyalty: A Markovian approach. *Operation Research, 10*(1), 19-40.

18. Hughes, A. M. (1997). Customer Retention: Integrating Lifetime Value into Marketing Strategies. *Journal of Database Marketing, 5*(2), 171–178.

19. Jackson, D. R. (1996). Achieving Strategic Competitive Advantage Through the Application of the Long-Term Customer Value Concept. *Journal of Database Marketing, 4*(2), 174–186.

20. Jenamani, M., Mohapatra, P., & Chose, S. (2003): A Stochastic Model of e-customer behavior. *Electronic Commerce Research and Applications, 2*(1), 82-94.

21. Jonker, J. J., Piersma, N., & Van den Poel, D. (2004). Joint Optimization of Customer Segmentation and Marketing policy to Maximize Long-term Profitability. *Expert Systems with Applications, 27*, 159-168.

22. Koski, T. (2001). *Hidden Markov Models for Bioinformatics*. Dordrecht: Kluwer.

23. Ling, R., & Yen, D. C. (2001). Customer Relationship Management: An Analysis Framework and Implementations Strategies. *Journal of Computer Information Systems, 41*, 82-97.

24. Liu, H. (2010). Dynamic Portfolio Choice Under Ambiguity and regime Switching mean Returns. *Journal of Economic Dynamics and Control, 35*(4), 623-640.

25. Ma, M., Li, Z., & Chen, J. (2008). Phase-type Distribution of Customer Relationship with Markovian Response and Marketing Expenditure Decision on the Customer Lifetime Value. *European Journal of Operational Research, 187*(1), 313-326.

26. MacDonald, I. L., & Zucchini, W. (1997). *Hidden Markov and Other Models for Discrete-valued Time Series*. London: Chapman and Hall.

27. Morgan, R., Hunt, S. (1994). The Commitment-trust theory of relationship marketing. *Journal of Marketing, 58*(3), 20–38.

28. Netzer, O., Lattin, J., & Sirinvasan, V. (2005). A Hidden Markov Model of Customer Relationship Dynamics. *Marketing Science, 27*(2), 185-204.

29. Nour-Mohammad, Y., Mahdi, S., & Jamshid, M. (2011). Improving Service Quality by using Organizational Citizenship Behaviour: Iranian Evidence. *Iranian Journal of Management Studies, 4*(2), 79–97.

30. Rabiner, L. R. (1989). *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*. Retrieved from https://web.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/tutorial%20on%20hmm%20and%20applications.pdf

31. Reinartz, W., & Kumar, V. (2003). The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration. *Journal of Marketing, 67*(1), 77-99.

32. Shen, H., & Zhao, J. (2006). Classification Customer Loyalty Based on hidden Markov model. *International Journal of Internet and Enterprise Management, 4*(1), 57-64.

33. Tanaka, T. (2017). A partially observable discrete-time Markov decision process with a fractional discounted reward. *Journal of Information and Optimization Sciences, 38*(1), 21-37.

34. Thomas, J. (2001). The Importance of Linking Customer Acquisition to Customer Retention. *Journal of Marketing Research, 39*, 262-268.

35. Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transaction on Information Theory, 13*(2), 260–269.

36. Wang, S., Gao, W., Huang, T., Ma, J., Li, J., & Xie, H. (2000). Adaptive Online Retail Web Site Based on Hidden Markov Model. *Lecture Notes in Computer Science*, 177–188.

37. Wu, H. H., & Shieh, J. I. (2006). Using a Markov Chain Model in Quality Function Development to Analyse Customer Requirements. *International Journal of Manufacturing Technology, 30*(1), 141-146.