#ml-papers November 2019

A Hidden Markov Model of Customer Relationship Dynamics

Section 1: Intro

1) **what is CRM?**

- Customer Relationship Management: ""improving"" business<->customer relationships to boost business revenue

2) **how do traditional CRM techniques work?**

- Main modeling research areas: how can we affect a customers relationship "state" to boost profits? How do we model LTV and churn likelihood?

3) **how is this approach different?**

- This paper focuses on the concept of business<->customer relationship "states", how they affect customer behavior, and how they evolve with events & external change

  ** note: that's how they frame this paper as being different from existing CRM techniques in the intro, however later in the paper they cite a ton of existing CRM research using state-models/M-chains/HMM techniques to model b<>c relationships & their contribution is a variation on existing techniques. Furthermore they keep using "we suggest" and "we propose" when referencing HMM modeling CRM relationships as a whole & don't explicitly mention their actual contribution till much later on. /rant*

Section 1: Intro

4) **how is this approach better?**
- lets us analyze what individual customer states are @ point in time, and how to most effectively drive them to preferred states

- since our model accounts for:
    - bidirectional relationship events (ie: customer purchases, marketing emails)
    - time
    - user level characteristics (ie: age, geolocation)
    - exogenous factors

.. **we can measure their individual impacts on business<->customer relationship dynamics**

and it feels.. holistic? (to me at least, not an actual argument used by Netzer et al)

5) **how do the authors demonstrate this? what else did they learn about the example problem?**
- they use university-alumni donations with 3 user states (dormant, occasional donor, frequent donor) and events like reunions, volunteer events, etc

<u>example analysis</u>: "*Attending a reunion seems to have a strong impact on moving alumni from the dormant to the occasional donation state and from the occasional to the active state. In contrast to the commonly used highest customer lifetime value approach, using the HMM we find only a small effect of reunion attendance on alumni in the frequent donation state. Volunteering to university roles, on the other hand, seems to have its primary impact on alumni in the dormant and active states, but not on alumni in the occasional state.*"

Before we go on, a quick recap on Markov Models & Hidden Markov Models will be useful imo

❑ Markov Assumption

parameters in stationary model K-ary variables

1st order $$p(\mathbf{X}) = \prod_{i=1}^{n} p(X_n | X_{n-1})$$ $O(K^2)$

mth order $$p(\mathbf{X}) = \prod_{i=1}^{n} p(X_n | X_{n-1}, \ldots, X_{n-m})$$ $O(K^{m+1})$

n-1th order $$p(\mathbf{X}) = \prod_{i=1}^{n} p(X_n | X_{n-1}, \ldots, X_1)$$ $O(K^n)$

≡ no assumptions – complete (but directed) graph

Homogeneous/stationary Markov model (probabilities don't depend on n)

Hidden Markov Models

- **Parameters** – stationary/homogeneous markov model (independent of time t)

Initial probabilities

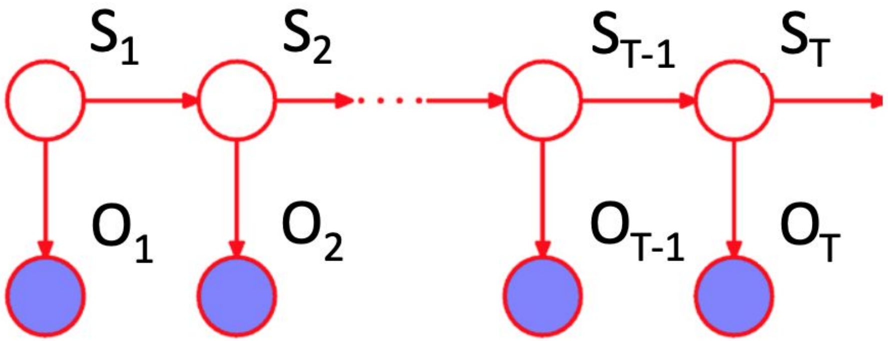
$$p(S_1 = i) = \pi_i$$

Transition probabilities

$$p(S_t = j \mid S_{t-1} = i) = p_{ij}$$

Emission probabilities

$$p(O_t = y \mid S_t = i) = q_i^y$$



$$p(\{S_t\}_{t=1}^T, \{O_t\}_{t=1}^T) = p(S_1) \prod_{t=2}^T p(S_t|S_{t-1}) \prod_{t=1}^T p(O_t|S_t)$$

for each of the T possible states, there is a set of emission probabilities governing the distribution of the observed variable at a particular time given the state of the hidden variable at that time

What would a dynamic HMM look like?

(also a good resource if you want to learn more details)

Section 2: Relationship Marketing and Dynamics in Buying Behavior

2.1 relationship marketing dynamics

- definition of a customer-company "relationship" (according to existing literature):
    - relationship is changed through discrete, bidirectional events
    - .. and environmental changes (less common in literature)

- a relationship can be modeled using discrete relationship states, ie (dormant, occasional donor, frequent donor)
    - if an event surpasses a customers "satisfaction threshold" @ their current state, they may transition into a more positive state

existing research _**mostly**_ fails to account for external impact (aka environment) on customer state, but they include it in their model.

Section 2: Relationship Marketing and Dynamics in Buying Behavior

2.2 dynamics in buying behavior and hidden markov models (part 1)

- difficulty in capturing individual & market dynamics (ie env impact on relationship state?) largely stems from lack of data volume

- if dynamics are smooth we can use AR-like models, but authors say since we have discrete events we need discrete states.. not sure why they can't have the relationship dynamics be smooth and have both discrete and continuous events? you can adjust a continuous parameter by a discrete amount right?

> past purchases.
> However, the continuous state space is inadequate to capture dynamics that are postulated to develop in a discrete manner such as an instantaneous regime shift in the market conditions or consumer preferences (e.g., due to an inclusion or a drop of a brand from the consumer's consideration set). One could model such dynamics, by allowing consumers (or markets) to transition over time between a set of discrete states. Probably the simplest demonstration of

(a better reason to use discrete state space is data volume imo but they don't mention this again?)

Section 2: Relationship Marketing and Dynamics in Buying Behavior

2.2 dynamics in buying behavior and hidden markov models (continued)

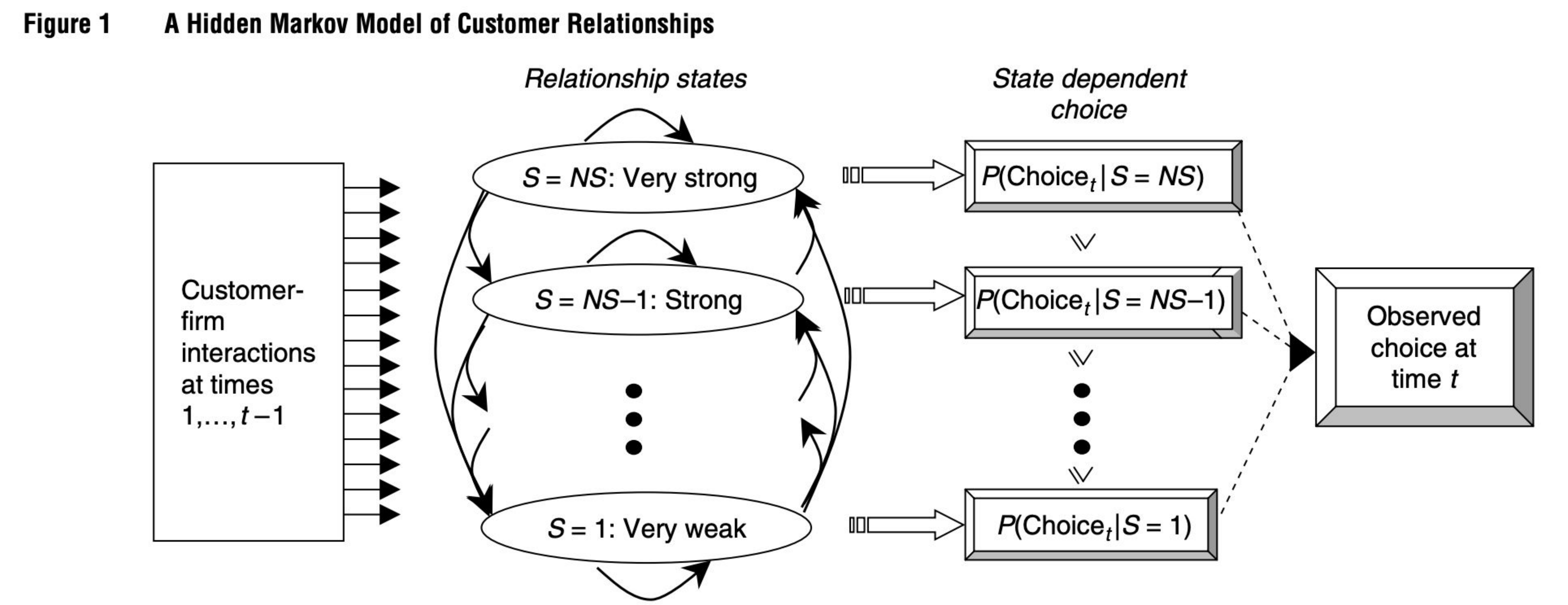
- Because of the hard Markov assumption MC models are too restrictive with regards to exogenous variables (each state can only be a function of the previous state)

- HMMs address this issue by creating unobservable/latent states & transitions to explain observed relationship patterns

- blah blah blah modeling dynamic change in latent segment membership is an ongoing challenge, but here's what they did different from existing work:
    - relax typical HMM assumption of stationary transitions between latent states, instead transitions are a function of time varying covariates (aka non-homogeneous MC).. "this is key if you want to understand the drivers of relationship dynamics" -> **if the impact of events on relationship dynamics is static in time the order in which they occur doesn't affect outcome and we can't model the impact of consecutive events**

    - they use individual level data on the customer to isolate user preference differences from relationship dynamic (this has been done in other papers)

Section 3: The Markov Model



A User's internal state (to purchase, to donate etc.) is based solely on their prev. customer-firm interactions.

Section 3: Transition Matrix

$$
\mathbf{Q}_{i,\,t-1\to t} = \begin{array}{c|cccccc|}
 & \multicolumn{6}{c}{\text{State at } t} \\
\text{State at } t-1 & 1 & 2 & 3 & \cdots & NS-1 & NS \\
\hline
1 & q_{it11} & q_{it12} & q_{it13} & \cdots & q_{it1NS-1} & q_{it1NS} \\
2 & q_{it21} & q_{it22} & q_{it23} & \cdots & q_{it1NS-1} & q_{it2NS} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
NS & q_{itNS1} & q_{itNS2} & q_{itNS3} & \cdots & q_{itNSNS-1} & q_{itNSNS} \\
\hline
\end{array}, \quad (1)
$$

The transition matrix contains a probability for “sudden death.” From each state the customer/alumni could move to an adjacent state or drop immediately to dormancy.

Section 3: Estimation Procedure

Initial States

The initial state distribution is commonly defined as the stationary distribution of the transition matrix

HMM Estimation

We estimate our HMM using a standard hierarchical Bayes estimation procedure using two sets of parameters:

1. random-effect parameters
2. parameters that are common across individuals

Section 4: Gift Giving Empirical Evidence

The variables of this data set can be divided into three categories:

1. Alumni-university interactions (reunion attendance, volunteering)
2. Influence Attempts
3. Choice Behavior (incidence of donation)

Tested 4 models:

1. **Full HMM**
2. **HMM with no heterogeneity:** This model is similar to Model 1, but with common parameter across individuals
3. **Nondynamic model:** This is the latent class model (Kamakura and Russell 1989)
4. **Recency-frequency model:** This model is frequently used in relationship management applications (e.g., Bult and Wansbeek 1995)

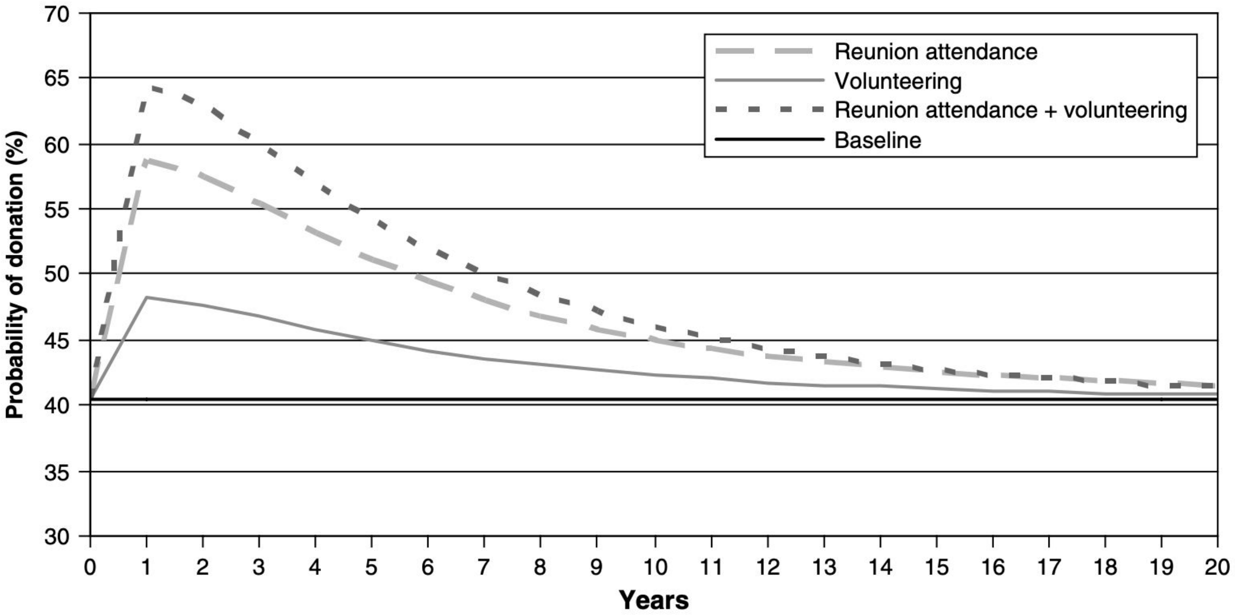
Section 4: Gift Giving Empirical Evidence

Reunion Attendance has a strong impact on moving alumni from the occasional to the active state;

1. Dormant state moves to the occasional state (10%→ 68%)
2. Occasional State to Active State (28 → 53%)
3. Likelihood of dropping off (14 → 5%)

Volunteering to university roles, on the other hand, has its primary impact on alumni in the dormant and active states, whereas the effect on alumni in the occasional state is minimal.

Figure 2 The Long-Term Effects of the Time-Varying Covariates



Section 5: General Discussion

The paper describes using a non-homogenous hidden Markov model (HMM) to model the changes in hidden state of a customer, when only the customer's actions over time are known. They discuss the example of a alumni donation dataset, where at each time point we can observe whether or not the alumni donated as well as their actions such as attended a reunion, volunteered, etc

Using these methods, it is possible to group customers into different states, and gain insight into understanding which marketing activities are most effective to transition customers into an active state.