Adaptive Collaborating Filtering

Onkar Dabbeer
Qualcomm Technologies Inc.
San Diego, CA

Abstract—In this paper, we study collaborative filters that adapt future recommendations based on feedback (like/dislike) received for past recommendations. We consider a mathematical model where users and items are clustered, and ratings are noisy. We apply sequential decision techniques to study structure of collaborative filters that maximize the time average of the expected ratings. While our previous work considered the case with unknown clusters and a moving horizon reward, in this paper we study the case of known clusters and the long run average reward (that is, the ergodic cost). We show that our model fits the Markov decision process framework. We show that if a solution to the dynamic programming equation exists, then the optimal cost is achieved for the problem. Moreover, we identify a simple collaborative filter that is optimal in the low noise regime. Moreover, empirically we find that this collaborative filter is near-optimal over the entire range of noise values.

I. INTRODUCTION

Recommendation systems are commonly used in e-commerce to recommend relevant content to users. Most practical recommendation systems exploit a variety of data sources to achieve their end goal and a collaborative filter, which exploits past user-item rating data, is usually an important ingredient. Collaborative filters have been studied extensively in the past decade both empirically and theoretically (see references [1], [2], [6]-[8], [10]-[15] and references therein). Specifically, in the last five years, for a number of mathematical models of the user-item rating matrix (such as the low-rank matrix model, or the neighborhood model) have been used to study fundamental limits of collaborative filtering and get insight into structure of near-optimal collaborative filters. One aspect that has received relatively less attention is the “exploration-exploitation” tradeoff in collaborative filters: in practice, collaborative filters are used to make recommendations, user ratings are obtained for these recommendations, which are then used for future recommendations. Thus recommendations should not only be relevant at the current time, but they should also aim to explore/sample the user-item matrix so that quality of future recommendations improves. We call such collaborative filters as adaptive collaborative filters (ACFs).

From the empirical standpoint, one of the main difficulties in studying ACFs is that an actual online implementation that directly makes recommendations to users and collects there feedback is necessary. An alternative is to consider mathematical models similar to those already used in the analysis of collaborative filters. While such mathematical models may not capture all the niceties of real world data, they often provide structural insight about good ACFs. In practice, several such simpler ACFs are usually combined together to form a more powerful recommender ([11]).

In this paper, we continue our study in [8], where we explicitly modeled user feedback and considered a sequential decision approach to the problem. Specifically, for a mathematical model of clustered users and items but with unknown clusters and noisy ratings, [8] studied optimization of a moving horizon reward, which is an approximation to the long term average reward problem [9]. In this paper, we consider the case of known clusters which allows for a more fuller treatment. Clustering of items and users using metadata about items/users is a well studied field, and our setup is applicable when the clustering is based on such metadata. For the case of known clusters and noisy feedback ratings, we consider optimization of the long-term average recommendation (that is, the ergodic reward). We show that the problem can be cast as an average reward Markov decision process and if a solution for the dynamic programming equation (DPE) exists, then the average reward must equal the best possible. In the low noise case, we give a solution to the DPE, which yields a simple collaborative filter: recommend items of types whose number of 1 ratings are greater than the number of 0 ratings (and recommend any item if no such item exists). While we are unable to solve the DPE for a non-zero noise level, we provide empirical evidence that this simple collaborative filter is close to optimal over the entire range of noise levels.

The details of our mathematical model, which is similar to that in [1], [2], [8], are given in Section II, the main results are stated and discussed in Section III, the conclusions are given in Section IV, and the proofs are given in the Appendix.

II. DATA MODEL AND PROBLEM FORMULATION

We index users with $m \in \{1, 2, \ldots\}$ and items by $n \in \{1, 2, \ldots\}$. At discrete times $t = 1, 2, \ldots$ an user $M_t$ seeks a recommendation. If the collaborative filter recommends item $N_t$, then immediately a feedback $Y_{M_t, N_t} \in \{0, 1\}$ is obtained depending upon whether the user likes the recommendation or not. The collaborative filter is strictly causal and we assume it is updated each time a recommendation is made, though in general, the updates may happen once every few $t$.

A common belief that makes collaborative filtering feasible is that several users have similar tastes and several items have similar characteristics. One way to mathematically model this is to assume that the users and items are clustered. We assume $K$ ($L$) distinct user (item resp.) types and denote the type of user $m$ (item $n$ resp.) by $U_m$ ($V_n$ resp.). The $U_m$ or $V_n$ may be obtained from metadata associated with users and items,
but they may also be obtained from past ratings as in [1], [8]. In this paper we assume that $U_m, V_n$ are known, which is an approximation for the case when the clustering is performed at a slower time scale than the collaborative filter using metadata.

The rating $Y_{m,n}$ depends only on the types of the user $m$ and the item $n$. Suppose that $\{A(k,l)\}_{k,l=1}^{K,L}$ are i.i.d. Bernoulli(1/2) and define the matrix $X$ such that its $(m,n)$th entry $X_{m,n} = A_{U_m,V_n}$. The matrix $Y$ is obtained by passing the entries of $X$ through a memoryless binary symmetric channel with error probability $\delta$, that is, given $X$, the entries of $Y$ are independent with

$$P(Y_{m,n} = X_{m,n}|X) = 1 - \delta,$$

$$P(Y_{m,n} \neq X_{m,n}|X) = \delta.$$

We note that the entries of $Y$ are not known and the only way to reveal them is by making recommendations. The noisy in the ratings captures both the noisy nature of user preferences as well as the modeling errors.

**Remark:** The assumption that $\{A(k,l)\}_{k,l=1}^{K,L}$ are i.i.d. Bernoulli(1/2) ensures that data from a single user (that is, a single row) cannot be used to predict the rating of that user for unseen items. Thus, under this assumption, it is, a single row) cannot be used to predict the rating of that user for unseen items. Thus, under this assumption, it is necessary to pool information from other users to make recommendations that are better than random guessing, that is, the “collaborative” aspect is a must.

### III. Main Results

Our goal is to find an ACF that maximizes the ergodic reward

$$\hat{R} = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[Y_{M_t, N_t}]$$

over the class of strictly causal ACFs. The standard framework for such optimization problems is the average reward Markov decision process framework. Our first result shows that our case does fit into this framework. To simplify notation, define the following past data,

$$A_{t-1} := \{m_{s,n_s} = y_s, 1 \leq s \leq (t-1)\},$$

$$B_{t-1} := \{M^{t-1} = m^{t-1}, N^{t-1} = n^{t-1}\},$$

where $m^n$ denotes the sequence $m_1, \ldots, m_t$. Let $S_{t-1} = \{A_{t-1}, B_{t-1}\}$ denote all the past data; we are interested in strictly causal ACFs that make a recommendation at time $t$ based on information $S_{t-1}$. We also wish to keep track of the number of positive and negative ratings received by items of type $l$ from users of item $k$. For this purpose, let $J^l(t)$ be the matrix whose entry $J^l_{k,l}$ equals the number of items of type $l$ that have been rated $i$ by users of type $k$. Let $D(t) = J^1(t) - J^0(t)$, which is the difference between the number of positive and negative ratings.

Our first result gives an expression for the expected rating when item $n$ is recommended to user $m$ conditioned on all past data.

**Lemma 1:** Under the assumptions of the model,

$$E[Y_{m,n}|S_{t-1}, M_t = m] = F(D_{U_m,V_n}(t-1); \delta)$$

where,

$$F(x; \delta) = \frac{\delta x + (1 - \delta)x}{\delta x + (1 - \delta)x}.$$

**Proof:** The proof is given in Appendix I.

In words, the expected reward given all the past depends only on $D(t - 1)$. It is easy to check that $\{D(t)\}_{t \geq 0}$ is a Markov chain and maximizing the long run average cost is related to solving the appropriate DPE. To write the DPE, we note that the recommendations for users of type $k$ do not have any impact on users of other types and the collaborative filter decouples into one for each type of user. Similarly, the reward only depends on the type of the item recommended. So without loss of generality, to simplify notation, we focus on user 1 with type 1 and the relevant state variable is denoted $d$ whose component $l$ is $d_l = D_{1,l}$. For this case, the DPE is

$$\rho + h(d) = \max_{e_i} \left\{F(d_l) + F(d_i)h(d + e_i) + (1 - F(d_i))h(d - e_i) \right\},$$

where $e_{l,p} = 1$ for $p = l$ and 0 otherwise, and we have to solve for the pair $(\rho, h(\cdot))$. From [5], if a solution to (1) exists, then under suitable conditions, the argument that maximizes the RHS yields the optimal recommendation and the constant $\rho$ is the ergodic reward for the optimal collaborative filter while $h(\cdot)$ is a potential function. In our model, even if we $X$ is known, due to the noise we cannot get an ergodic reward greater than $(1 - \delta)$, that is, $\rho \leq 1 - \delta$. In fact, we have the following result.

**Proposition 1:** If the DPE (1) has a solution, then $\rho = 1 - \delta$.

**Proof:** The proof is given in Appendix II.

The above proposition implies that if a solution to (1) exists, then it must be the case that optimal policy attains the highest possible average reward of $(1 - \delta)$. Unfortunately, to find the optimal policy, we need to find $h(\cdot)$, and we do have been able to do so analytically. However, we get some insight by looking at the noiseless case. As $\delta \to 0$, $F(x; \delta) \to \text{sign}(x)$, where at $x = 0$ the sign function takes the value 1/2. For this limiting value of $F(x; \delta)$, the optimality equation becomes

$$1 + h(d) = \max_{e_i} \left\{\text{sign}(d_l) + \text{sign}(d_i)h(d + e_i) + (1 - \text{sign}(d_i))h(d - e_i) \right\}.$$

It is easy to check that the following is a solution to this equation:

$$h_s(d) = -\sum_{l=1}^{L} d_l, \quad \text{if } d_l < 0 \text{ for all } l$$

$$= 0.5, \quad \text{if } d = 0$$

$$= 1, \quad \text{otherwise}.$$

The collaborative filter corresponding to this solution is also easy to describe: if $d < 0$, then any item can be recommended, else any item of a type $l$ with $d_l \geq 0$ can be recommended. Surprisingly, we find empirically that this simple collaborative filter works well even for $\delta$ close to 1/2. This is illustrated
Fig. 1. Simulations show that the ergodic reward of the δ → 0 optimal ACF is close to the optimal for all noise levels δ.

in Figure 1, where we plot the average cost obtained for 1000 recommendations for L = 50 as δ varies from 0 to 0.5. It is clear that the proposed collaborative filter is near optimal even for δ close to 0.5 (even though it was derived by considering the δ → 0 case). We also find empirically that just recommending an item of type with highest δ_1 also works equally well.

IV. CONCLUSIONS

Mathematical models in a data-driven field such as collaborative filtering serve to provide insights into the structure of good collaborative filters. In [8] and this paper, we studied structure of optimal ACFs for the case of clustered users and items. In [8] the structural insight was in the form of a metric for clustering users and items based on the ratings feedback, which is followed by recommendation of items most popular amongst users of the same type. The setup in this paper aims to address the case where the clustering is primarily obtained from metadata and is updated on a much slower time scale than the ratings. The structural insight in this case of known clusters is that only the difference between the number of 1 ratings and 0 ratings matters. Moreover, the magnitude of the difference appears to be unimportant: just recommending items from types with positive difference leads to performance close to optimal over the whole range of noise levels. Several variations on our basic model as well on the ACF derived are possible and may be worth exploring further.

APPENDIX I

Proof of Lemma 1

Since Y_m,n is a binary random variable

\[ E[Y_m,n|S_{t-1}, M_t = m] \]
\[ = P(Y_m,n = 1|S_{t-1}, M_t = m) \]
\[ = \frac{P(Y_m,n = 1, A_{t-1}|B_{t-1}, M_t = m)}{P(A_{t-1}|B_{t-1}, M_t = m)}. \] (2)

First we calculate the numerator. For simplicity of notation, let n_t = n, m_t = m, and y_t = 1. If we condition on X, then we know that the Y_m,n depends only on X_m,n and is independent of all other random variables. Therefore, conditioning on X and using the fact that the noise is i.i.d., the numerator equals,

\[
\text{Num} = E \left[ \prod_{k,l \in I_{k,l}} (1 - \delta) \sum_{m,n} A_{k,l,m} = y_k \delta \sum_{m,n} A_{k,l,m} \neq y_k \right].
\]

But \{A_{k,l}\} are i.i.d. Bernoulli(1/2), and hence

\[
\text{Num} = E \left[ \prod_{k,l \in I_{k,l}} (1 - \delta)^{J_{k,l}^0(t)} \delta^{J_{k,l}^1(t)} \right].
\]

Pulling out the factor \delta^{J_{k,l}^0} + J_{k,l}^1 and noting that \sum_{k,l}(J_{k,l}^0 + J_{k,l}^1) = t, we get

\[
\text{Num} = \delta^t \left[ \prod_{k,l} \left( e^{J_{k,l}^0(t)} + e^{J_{k,l}^1(t)} \right) \right].
\]

The denominator has a similar expression

\[
\text{Den} = \frac{\delta^{t-1}}{2L} \left[ \prod_{k,l} \left( e^{J_{k,l}^0(t-1)} + e^{J_{k,l}^1(t-1)} \right) \right].
\]

Since m_t = 1, n_t = 1 and y_t = 1, J_{U_m,V_n}^0(t) = J_{U_m,V_n}^1(t - 1) + 1 and all other entries of J(t) equal those of J(t-1) for i = 0, 1. Therefore,

\[
E[Y_m,n|S_{t-1}, M_t = m] = \delta \times \frac{e^{J_{U_m,V_n}^0(t-1)} + e^{J_{U_m,V_n}^1(t-1)+1}}{e^{J_{U_m,V_n}^0(t-1)} + e^{J_{U_m,V_n}^1(t-1)}}
\]
\[
= F(D_{U_m,V_n}(t-1); \delta),
\]
APPENDIX II
PROOF OF PROPOSITION I

From (1),
\[ \rho + h(d) \geq F(d) + F(d)h(d + e_t) + (1 - F(d))h(d - e_t). \] (3)
Noting that \( \lim_{d_t \to \infty} F(d_t) = 1 - \delta \) and taking limit \( d_t \to 0 \) on both sides, we get,
\[ \lim_{d_t \to 0} \inf h(d) \geq 1 - \delta - \rho + \limsup_{d_t \to 0} h(d). \] (4)
Since \( (1 - \delta - \rho) \geq 0 \), it follows that \( \lim h(d) \) exists. If the limit is finite, then from (4), we get that \( \rho = 1 - \delta \). Next we show that \( \lim h(d) \) is indeed finite. Rewriting (3),
\[ h(d + e_t) - h(d) \leq \frac{(1 - F(d))}{F(d)} \left( h(d) - h(d - e_t) \right) + \left( \frac{\rho}{F(d)} - 1 \right). \]
For sufficiently large \( d_t \), \( \frac{(1 - F(d))}{F(d)} \) is bounded by a constant less than 1 (since its limit \( \delta/(1 - \delta) < 1 \)). Also the second term in the bound is finite. Therefore, it follows that the sequence \( \{h(d + e_t) - h(d)\}_{d_t=0}^{\infty} \) decays exponentially fast for \( d_t \) sufficiently large, and its partial sum sequence - namely \( \{h(d)\}_{d_t=0}^{\infty} \) - is a bounded sequence. Hence \( \lim h(d) \) is finite and the proof that \( \rho = 1 - \delta \) is complete.

REFERENCES