When and How to Diversify: A Multicategory Utility Model for Personalized Content Recommendation

Changrong XIAO

School of Economics and Management
Tsinghua University

July 9, 2020

Contents

Online News Recommendation

- Online news.
  - Great accessibility: challenge to discover interesting contents.
- Readers' needs for content discovery platforms and personalized recommendation tools.
- Content consumption is different from product consumption.
  - Not limited demands.
  - Readings rarely "done"
- Readers' variety-seeking behavior.

Diversity Recommendation

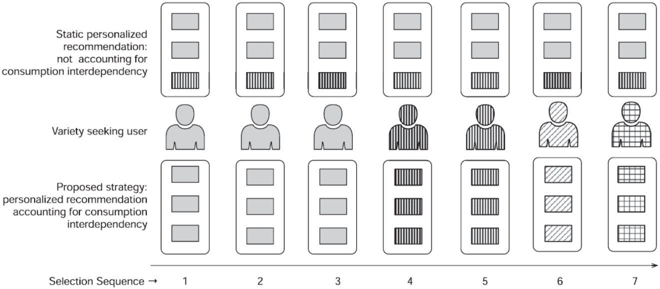
- Traditional recommendation approaches
  - Without considering the interdependence among consumptions over time. Time-invariant
  - Lead to narrow set of recommendations, and potentially losing relevance in practice.
- Extant diversification approaches
  - Provide a diversified list, instead of a single product.
  - Advantages: Reduce the risk to predict consumer's uncertain desires; recommendations are more useful for consumers.
  - Challenges: Do not learn how diversity can lead to greater relevance.

Tradeoffs between Accuracy and Diversity!

Diversity Recommendation (Cont'd)

- Improved diversification approach in this paper
  - Must we sacrifice accuracy to offer readers the diversity they want? Not necessarily
  - Learns how quickly the reader satiates with a category of content and wishes to substitute it. Time-variant!

Figure 1. The Intuition of the Proposed Recommendation Strategy as Illustrated over One Session



Note. In the proposed approach (per the lower panel), the recommendation is most accurate for the consumer at each stage of the session while exhibiting diversity over the entire session.

Context

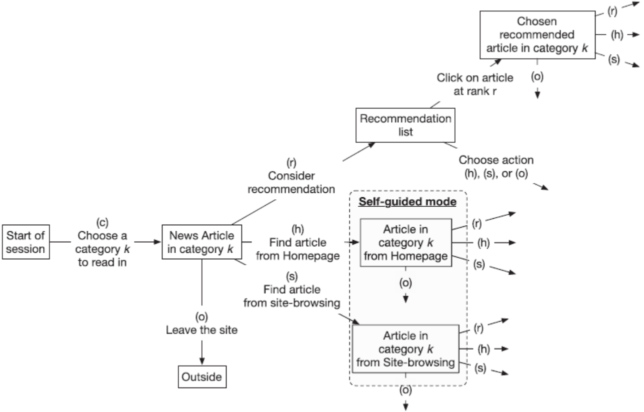
- Chosen content-discovery platform: Outbrain.
- Three types of recommendations:
  - Links with thumbnails
  - Text links
  - Text links to external sites
- Focus on contents consumed during each online session, instead of all consumption occasions.
  - The variety-seeking behavior may get masked in the aggregated consumption over a long period of time that includes multiple sessions.

Model

- Aim: Provide a sequential recommendation strategy in a single session with both diversity and accuracy.
- Assumption: When choosing a category at each step of the session, each reader maximizes the expected utility.
- Recommend the most relevant items at the right time $\Rightarrow$ Learning a consumer's utility function from historical choices

Consumer's Decision Process

- A consumer's decision process within a session.



Utility from Consuming Content

- Constant elasticity of substitution (CES), a type of multicategory utility function

$$U(\mathbf{x}_u^t) = (\sum_{k=1}^{K} a_{ku} {x_{ku}^t}^{\frac{s_u - 1}{s_u}})^{\frac{s_u}{s_u - 1}}$$

  - $\mathbf{a}_u = \{a_{ku}\}$ is a $K \times 1$ vector of utility coefficients from different categories of content for reader u. $log(\mathbf{a}_u) \sim Normal(\mu^a, \mathbf{\Sigma}^a)$
  - s_u is the elasticity of substitution among categories. $log(s_u) \sim Normal(\theta^a, \sigma^a)$
  - s_u can model the perfect complements and substitutes
  - **Decreasing marginal utility property**. $\frac{s_u - 1}{s_u} < 1$

Utility from Consuming Content (Cont'd)

- Marginal utility function:

$$\Delta U\left(k;\mathbf{x}_u^t\right) = \left(\sum_{k'=1}^{K} a_{k'u} {x_{k'u}^t}^{\frac{s_u-1}{s_u}}\right)^{\frac{1}{s_u-1}} a_{ku} {x_{ku}^t}^{-\frac{1}{s_u}}$$

- With switching cost:

$$\Delta U\left(k;\mathbf{x}_u^t\right) = \left(\sum_{k'=1}^{K} a_{k'u} {x_{k'u}^t}^{\frac{s_u-1}{s_u}}\right)^{\frac{1}{s_u-1}} a_{ku} {x_{ku}^t}^{-\frac{1}{s_u}} - C_u^{switch}\mathbf{D}(k',k)$$

  - Dissimilarity matrix $\mathbf{D}(k',k) = 1 - \mathbf{T}(k',k)$
  - $C_u^{switch} \sim LogNormal(\theta_{switch}, \sigma_{switch})$, parameters shared by all readers.

Reader's Belief About Content Quality

- Model two levels of quality: good and bad.
- Consider *recommendation* channel:
  - Assume the quality of an article $\sim Bernoulli(p_{ru})$
  - Assume $p^{\tau}_{ru} \sim Beta(\alpha^{\tau}_{ru}, \beta^{\tau}_{ru})$
  - Reader updates the belief over time in a Bayesian manner: $\alpha^{\tau}_{ru} = \alpha^{0}_{ru} + n^{\tau}_{rug}, \quad \beta^{\tau}_{ru} = \beta^{0}_{ru} + n^{\tau}_{ru} - n^{\tau}_{rug}$.
- Quality belief from *site-browsing* channel and *homepage* channel: $p^{\tau}_{su} \sim Beta(\alpha^{\tau}_{su}, \beta^{\tau}_{su})$, $\quad p^{\tau}_{hu} \sim Beta(\alpha^{\tau}_{hu}, \beta^{\tau}_{hu})$
  - These two self-guided mode beliefs are not updated.
- Expectation of marginal utility over the article quality:

$$\Delta V^{r}(k; \mathbf{x}^{t}_{u}) = \{ (\sum_{k'=1}^{K} a_{k'u} {x^{t}_{k'u}}^{\frac{s_u - 1}{s_u}})^{\frac{1}{s_u - 1}} a_{ku} {x^{t}_{ku}}^{-\frac{1}{s_u}} \} \bar{p}^{\tau}_{ru} - C^{switch}_{u} \mathbf{D}(k', k)$$

Position Bias

- The position of an item in a list can affect the probability of the item being selected (*Day 1969, Blunch 1984*).
- Cascade model
  - The reader inspects links on a list starting from the top, clicks on the first satisfactory link, and does not inspect any link positioned lower.
- Model each click decision using a binary logit model.

$$P\left(\Delta V_i^r\left(k;\mathbf{x}_u^t\right)+\epsilon<0\right)=\frac{1}{1+e^{\Delta V_i^r\left(k;\mathbf{x}_u^t\right)}}$$

  - ϵ is the Type I extreme value error capturing factors affecting the expected utility of the reader, observed by the reader but unobserved by the econometrician.

Channel Choices

- The expected value from an article on the *recommended list*: $\left\langle \Delta V^r \left(\mathbf{x}_u^t\right)\right\rangle = \sum_k P_u^r(k) \left\langle \Delta U \left(k; \mathbf{x}_u^t\right)\right\rangle_{p_n^\tau}$
- The expected value of choosing to inspect the *recommendation list*:

$$\langle\langle \Delta V^r(\mathbf{x}_u^t)\rangle\rangle = \left(1 - \prod_i^R \frac{1}{1 + e^{\langle \Delta V^r(x_u^t)\rangle}}\right) \cdot \langle \Delta V^r(\mathbf{x}_u^t)\rangle - C_u^r$$

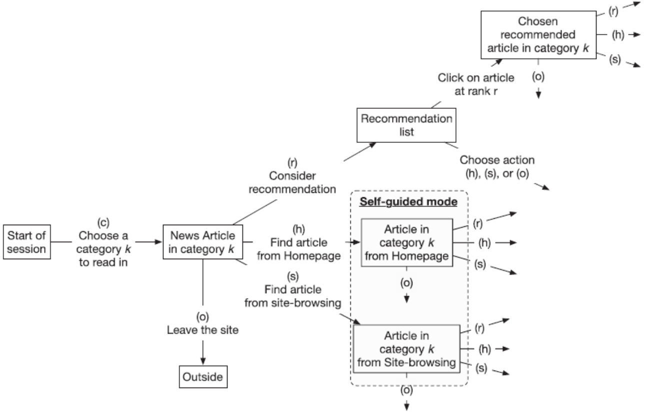
  - *Homepage*: $\langle\langle \Delta V^h(\mathbf{x}_u^t)\rangle\rangle = \sum_k P^h(k) \left\langle \Delta U \left(x_u^t\right)\right\rangle_{p_{hu}^r} - C_u^h$
  - *Site-browsing*: $\langle\langle \Delta V^s(\mathbf{x}_u^t)\rangle\rangle = \sum_k P^s(k) \left\langle \Delta U \left(x_u^t\right)\right\rangle_{p_{su}^r} - C_u^s$
  - $C_u \sim LogNormal(\theta, \sigma)$ is reader-specific searching cost.

Model Summary

- Multicategory utility with CES (MCUwCES) model:
  1. For the first article, choose a category using multinomial logit model with expected value of each category $\Delta U(k;0) = a_{ku}^{\frac{s_u}{s_u-1}}$
  2. Repeat till the end of the session:
     a. Choose one from the four possible channels using multinomial logit model with: *recommendations* $\langle\langle \Delta V^r(\mathbf{x}_u^t)\rangle\rangle$, *homepage* $\langle\langle \Delta V^h(\mathbf{x}_u^t)\rangle\rangle$, *site-browsing* $\langle\langle \Delta V^s(\mathbf{x}_u^t)\rangle\rangle$, and leave 0.
     b. If *recommendations*, the probability of clicking the n^{th} article: $\frac{1}{1+e^{-\Delta V_n^r(k;x_u^t)}} \prod_i^{n-1} \frac{1}{1+e^{\Delta V_i^r(k;x_u^t)}}$. If the reader choose none, turn to the remaining three channels.
     c. If *homepage* or *site-browsing*, determine the category using multinomial logit model with $\Delta U\left(k;\mathbf{x}_u^t\right)$.
     d. If *leave*, end the session.
- Multinomial logit model: $\Pr\left(Y_i = c\right) = \frac{e^{\boldsymbol{\beta}_c \cdot \mathbf{X}_i}}{\sum_{k=1}^{K} e^{\boldsymbol{\beta}_k \cdot \mathbf{X}_i}}$

Model Summary (Cont'd)

- Readers' decision process recall



Data and Estimation

- Data source: a clickstream data set on an large international news website over a period of three months.
- Train/Test partition: $90\%/10\%$
- Estimation method: MCMC (Markov chain Monte Carlo)

Alternative Models

- Alternative models:
  - Utility functions.
  - Recommendation consideration structures.
  - Different mechanisms to account for the effect of day-to-day variation in content supply and demand.
- Model performance
  - The first two sets of models have lower likelihood and higher BIC. $\Rightarrow$ MCUwCES model is a better fit
  - The third set of models have higher likelihood but higher BIC. $\Rightarrow$ MCUwCES model is a better fit after accounting for model complexity

Recommending a Category

- Compare to 9 alternatives:

Table 4. Average Accuracies of Category-Level Predictions

Recommendation approach	Average accuracy, %
Random category	6.25
Most frequently read category	20.85
Recommend current category	22.04
DailyLearner	25.19
Coverage-based	26.94
User–user	31.74
CCF[7]	31.92
Fossil	30.17
Katz-CWT	34.75
MCUwCES (proposed approach)	40.90

Note. The standard error for each average is less than 0.01%.

- The accuracy of MCUwCES is higher and stable both at the beginning of the sessions and when the satiation occurs.

Recommending an Article

- Category-level + item-level algorithms
- MCUwCES + optimized item-level improves 8%-11% accuracy.
- While list-diversification *(Ziegler et al. 2005)* + optimized item-level reduces 4%-18% accuracy.

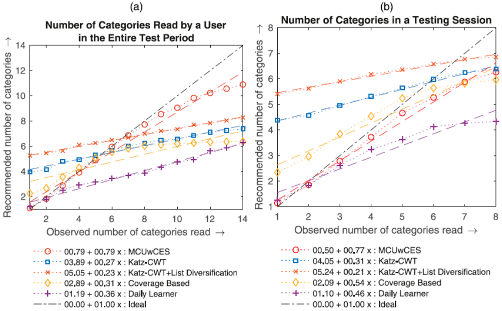
Table 5. Average Accuracies of Article-Level Predictions

	No additional diversification		List diversification		Multicategory diversification	
Article recommender system	Top 1, %	Top 6, %	Top 1	Top 6, %	Top 1, %	Top 6, %
Popularity	2.42	5.81	Same as no diversification	5.10	2.80	6.51
DailyLearner	2.14	5.87		4.83	2.24	6.10
User–user	2.51	6.00		5.25	2.84	6.52
Coverage-based	2.23	6.06		5.84	2.26	6.15
Spreading algorithms	2.76	6.14		5.42	3.12	6.73
CCF	2.54	6.06		5.35	2.80	6.55
Fossil	2.49	6.10		5.82	2.60	6.32
Katz-CWT	2.85	6.31		5.46	3.16	6.80

Note. The standard error for each average is less than 0.01%.

Analysis of Diversity

- Diversity measurement:
  - Plotting the average number of categories recommended against the number of categories consumed.
- MCUwCES outperforms all other approaches
  - Common problems: over-diversify in small number of categories, and under-diversify in more categories.



Note. A straight line is fitted to each plot, and the linear equation is shown to help assess the closeness of the recommendation diversity to the diversity of an ideal recommendation.

Potential Impact on Utility

- The evaluations above use readers' choices given the current recommender system, but the readers' choices and utilities may be altered if the proposed system is implemented.
- The utility of an article selected in self-guided modes could be higher than the utility of an article selected from recommendations
  - Readers click on recommendation lists only in 4.3% of all page visits.

Potential Impact on Utility (cont'd)

- Using policy simulation method.
- MCUwCES increases the net utility from the recommended links, even higher than that from self-guided channels.

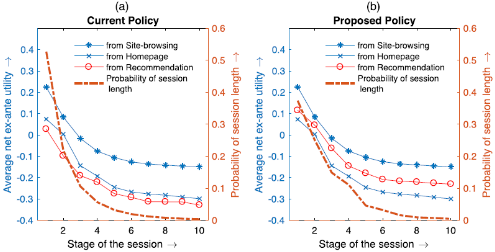
Table 6. Average Estimated *Ex Post* Utilities from the Selected Articles Under (a) the Current Recommendation Policy and (b) the Proposed MCUwCES Policy

	(a) Current recommendation policy			(b) Proposed MCUwCES policy		
	Site browsing	Home page	Recommended	Site browsing	Home page	Recommended
Gross utility	6.433	4.963	4.243	6.527	5.023	7.326
Cost	1	0.654	0.572	1	0.654	0.572
Net utility	5.433	4.309	3.671	5.527	4.369	6.754

Note. The standard error of each average is less than 0.004.

Potential Impact on Engagement

- Under the proposed strategy, the average net *ex ante* utility from a recommended article is significantly higher than what it is under the currently implemented strategy.
- MCUwCES leads to more page views from the recommended items. $\Rightarrow$ Higher engagement with the site
- Each selected article would provide higher utility to the reader on average.



Conclusion

- Dual objective: **Accuracy and diversity**.
- MCUwCES model: Learning how a reader selects content from different categories over time in a session, and using this knowledge to make sequential recommendations in a path-dependent manner
  - Accuracy: Category choosing and channel switching.
  - Diversity: Decreasing marginal utility property of multicategory utility function.
- Result: Improve accuracy by adaptively diversifying recommendations during the session
  - 8%-11% more accurate than only the collaborative filters.
  - 25% more accurate than collabarative filters + list-diversification strategy.

Limitations and Future Studies

- Limitations
  - Many sessions start from the homepage, while not all categories appear on it.
  - The diversity of consumption in the self-guided mode may not be true preference for diversity of the reader.
  - May be affected by the editorial decision to display certain categories on the home page.
- Future Studies
  - Recommend individual items without relying on other item-level recommender systems.
  - Consider the situation when only one item is consumed in a session.
  - Consider time lapses between consumption sessions.

Q&A