Recommending New Items to Ephemeral Groups Using Contextual User Influence

RecSys 2016, Elisa Quintarelli, Emanuele Rabosio and Letizia Tanca

Contents

Motivation

Why Recommendations for Groups?

Some scenarios where group recommendations could be preferable

- Choosing a restaurant
- Choosing which movie to watch in the cinema
- Choosing a TV program

In this article they choose to work with the TV scenario.

Why Ephemeral Groups?

The article discuss two kinds of groups being persistent groups and ephemeral groups.

- Persistent groups have a significant history and can be treated at one user
- Ephemeral groups have little to no history is individual preferences is needed

The argument for selecting the ephemeral approach is that it is more general and is applicable in more scenarios.

Why Concern About Contextual User Influence?

The article argues that some group members are more influential than others depending on the context.
A scenario from the article considers a family with the users parents and kids and the contexts early and late afternoon:

- Early afternoon - The children have most influence
- Late afternoon - The parents have most influence

The Essence of the Article

Let *Users* be a set of users, *Items* be a set of items and *Contexts* be a set of contexts and let u, i and c be subsets thereof. Supposing that the following data are known:

- $score\_ctx(u, i, c)$, assigning score to items based in contexts
- $score\_no\_ctx(u, i, c)$, assigning score to items regardless of context
- A log L recording the history of items chosen by the groups

Then a recommendation of K items, which is available in a given time instance, will be presented for a Group of users based on their given context.

Methodology

Setup Based on the Scenario

Making recommendations based on TV programs limits the available items to given time instances.

Time	Context	Group	Chosen Item	Available Items
t_1	day	{a, b, c}	i_1	$\{i_1, i_2, i_3\}$
t_2	day	{a, c}	i_5	$\{i_4, i_5, i_6\}$
t_3	day	{a, b, c}	i_8	$\{i_7, i_8, i_9\}$
t_4	night	{a, b}	i_{11}	$\{i_{10}, i_{11}, i_{12}\}$
t_5	night	{a, b}	i_{15}	$\{i_{13}, i_{14}, i_{15}\}$

At each time instance t each user have a preferred list of items denoted as:

List of User Preferences

$$TopK(u, c, t)$$

In this article K is in the range from 1 to 3.

Influence

The log is used to calculate a users influence by either using a context aware or unaware approach based on available data.

Influence with Context

$$infl\_ctx(u,c) = \frac{|l_j \in L : c = c_j \wedge u \in G_j \wedge i_j \in TopK(u,c,t_j)|}{|l_j \in L : c = c_j \wedge u \in G_j|}$$

Influence with No Context

$$infl\_no\_ctx(u) = \frac{|l_j \in L : \wedge u \in G_j \wedge i_j \in TopK(u,c,t_j)|}{|l_j \in L : \wedge u \in G_j|}$$

Example on Influence with Context

Influence with Context

$$infl\_ctx(u,c) = \frac{|l_j \in L : c = c_j \wedge u \in G_j \wedge i_j \in TopK(u,c,t_j)|}{|l_j \in L : c = c_j \wedge u \in G_j|}$$

Time	Context	Group	User Preference	Chosen Item	Available Items
t_1	day	{a, b, c}	$a \rightarrow i_1, b \rightarrow i_1, c \rightarrow i_2$	i_1	$\{i_1, i_2, i_3\}$
t_2	day	{a, c}	$a \rightarrow i_5, c \rightarrow i_5$	i_5	$\{i_4, i_5, i_6\}$
t_3	day	{a, b, c}	$a \rightarrow i_9, b \rightarrow i_8, c \rightarrow i_9$	i_8	$\{i_7, i_8, i_9\}$
t_4	night	{a, b}	$a \rightarrow i_{11}, b \rightarrow i_{11}$	i_{11}	$\{i_{10}, i_{11}, i_{12}\}$
t_5	night	{a, b}	$a \rightarrow i_{15}, b \rightarrow i_{13}$	i_{15}	$\{i_{13}, i_{14}, i_{15}\}$

The contextual influence where the context is *day*:

- $infl\_ctx(a, day) = \frac{2}{3}$
- $infl\_ctx(b, day) = \frac{2}{2}$
- $infl\_ctx(c, day) = \frac{1}{3}$

The contextual influence where the context is *night*:

- $infl\_ctx(a, night) = \frac{2}{2}$
- $infl\_ctx(b, night) = \frac{1}{2}$

Since there are no feedback from user c when the context is night, no context influence value can be calculated, hence the need of the non-contextual equation

Example on Influence with No Context

Influence with No Context

$$infl\_no\_ctx(u) = \frac{|l_j \in L : \wedge u \in G_j \wedge i_j \in TopK(u,c,t_j)|}{|l_j \in L : \wedge u \in G_j|}$$

Time	Context	Group	User Preference	Chosen Item	Available Items
t_1	day	{a, b, c}	$a \rightarrow i_1, b \rightarrow i_1, c \rightarrow i_2$	i_1	$\{i_1, i_2, i_3\}$
t_2	day	{a, c}	$a \rightarrow i_5, c \rightarrow i_5$	i_5	$\{i_4, i_5, i_6\}$
t_3	day	{a, b, c}	$a \rightarrow i_9, b \rightarrow i_8, c \rightarrow i_9$	i_8	$\{i_7, i_8, i_9\}$
t_4	night	{a, b}	$a \rightarrow i_{11}, b \rightarrow i_{11}$	i_{11}	$\{i_{10}, i_{11}, i_{12}\}$
t_5	night	{a, b}	$a \rightarrow i_{15}, b \rightarrow i_{13}$	i_{15}	$\{i_{13}, i_{14}, i_{15}\}$

The influence with no regard to context:

- $infl\_no\_ctx(a) = \frac{4}{5}$
- $infl\_no\_ctx(b) = \frac{3}{4}$
- $infl\_no\_ctx(c) = \frac{1}{3}$

Group Preference

Now that we have the item score and the influence value given by:

- $score\_ctx(u,i,c)$ and $infl\_ctx(u,c)$ for context-aware
- $score\_no\_ctx(u,i)$ and $infl\_\_ctx(u)$ for context-unaware

we can make the aggregation in order to get a group score.

Context-aware Scoring for Groups

$$score\_ctx(G,i,c) = \frac{\sum_{u \in G} infl\_ctx(u,c) \times score\_ctx(u,i,c)}{\sum_{u \in G} infl\_ctx(u,c)}$$

Context-unaware Scoring for Groups

$$score\_no\_ctx(G,i) = \frac{\sum_{u \in G} infl\_no\_ctx(u) \times score\_no\_ctx(u,i)}{\sum_{u \in G} infl\_no\_ctx(u)}$$

Example on Context-aware Scoring

Context-aware Scoring for Groups

$$score\_ctx(G,i,c) = \frac{\sum_{u \in G} infl\_ctx(u,c) \times score\_ctx(u,i,c)}{\sum_{u \in G} infl\_ctx(u,c)}$$

Time	Context	Group	Chosen Item	Available Items
t_6	day	{b, c}	??	$\{i_{16}, i_{17}, i_{18}\}$

The users individual preferences for t_6 are:

- $b, i_{16} = 0.8, b, i_{17} = 0.2, b, i_{18} = 0.3$
- $c, i_{16} = 0.2, c, i_{17} = 0.5, c, i_{18} = 0.9$

This individual scores are then aggregated together with the individual users context influence:

- $score\_ctx(\{b,c\}, i_{16}, day) = \frac{1 \times 0.8 + \frac{1}{3} \times 0.2}{1 + \frac{1}{3}} = 0.65$
- $score\_ctx(\{b,c\}, i_{17}, day) = \frac{1 \times 0.2 + \frac{1}{3} \times 0.5}{1 + \frac{1}{3}} = 0.275$
- $score\_ctx(\{b,c\}, i_{18}, day) = \frac{1 \times 0.3 + \frac{1}{3} \times 0.9}{1 + \frac{1}{3}} = 0.450$

Example on Context-unaware Scoring

Context-unaware Scoring for Groups

$$score\_no\_ctx(G,i) = \frac{\Sigma_{u \in G} infl\_no\_ctx(u) \times score\_no\_ctx(u,i)}{\Sigma_{u \in G} infl\_no\_ctx(u)}$$

Time	Context	Group	Chosen Item	Available Items
t_7	night	{a, b, c}	??	$\{i_{19}, i_{20}, i_{21}\}$

The users individual preferences for t_7 are:

- $a, i_{19} = 0.9, a, i_{20} = 0.1, b, i_{21} = 0.6$
- $b, i_{19} = 0.2, b, i_{20} = 0.9, b, i_{21} = 0.7$
- $c, i_{19} = 0.3, c, i_{20} = 0.7, c, i_{21} = 0.5$

This individual scores are then aggregated together with the individual users influence:

- $score\_no\_ctx(\{a,b,c\}, i_{19}) = \frac{0.8 \times 0.9 + 0.75 \times 0.2 + \frac{1}{3} \times 0.3}{0.8 + 0.75 + \frac{1}{3}} = 0.515$
- $score\_no\_ctx(\{a,b,c\}, i_{20}) = \frac{0.8 \times 0.1 + 0.75 \times 0.9 + \frac{1}{3} \times 0.7}{0.8 + 0.75 + \frac{1}{3}} = 0.525$
- $score\_no\_ctx(\{a,b,c\}, i_{21}) = \frac{0.8 \times 0.6 + 0.75 \times 0.7 + \frac{1}{3} \times 0.5}{0.8 + 0.75 + \frac{1}{3}} = 0.622$

Evaluation

Dataset

The datasets used are Meetup and Plancast consisting of implicit feedback from individual users and groups watching TV

- 7921 users, 119 channels and 21194 distinct programs
- There are almost 5 million entries
  - 3.5 million individual users
  - 1.5 million groups

The groups are split into training and testing sets 84/16

Compared Methods

In order to determine the recommendation performance the method is compared against excising solutions which is:

- Average scoring
- Least misery
- Maximum satisfaction
- Disagreement
- Expertise

Expertise uses weighted average where the weight is based upon how much a user have watched.

Evaluation and Comparison Metrics

They us recall the see how well the methods match the preferences from the test set. Test are made where the size of *TopK* is from 1-3.

Recall

$$Recall@K = \frac{\text{true positive}}{\text{true positive + false negative}} = \frac{|v \in \text{Test Set} : v_i \in TopK(v_G, v_c, v_t)|}{|v \in \text{Test Set}|}$$

Spearman's footrule is used to measure the heterogeneity of the group based on the *TopK* of each user, where $K = 3$. If users *TopK* are identical the value is 0 but if they are all disjoint the value is 12.

Spearman's footrule

$$F(\sigma) = \sum_i i - \sigma(i)$$

Recall@1

Test set	CtxInfl	Infl	EXP
All	34.12%	32.85%	31.31%
Group size = 2	34.81%	33.66%	32.38%
Group size = 3	31.18%	29.39%	26.87%
Group size = 4	31.71%	29.98%	27.24%
Group size = 5	31.89%	30.58%	26.24%
$h \leq 4$	44.44%	44.03%	42.60%
$4 < h \leq 8$	36.13%	34.88%	33.70%
$h > 8$	26.71%	24.99%	23.03%

Recall@3

Test set	CtxInfl	Infl	EXP
All	62.48%	61.58%	60.67%
Group size = 2	63.36%	62.61%	61.90%
Group size = 3	59.25%	57.72%	56.11%
Group size = 4	57.88%	56.31%	53.84%
Group size = 5	58.91%	56.21%	55.52%
$h \leq 4$	75.25%	75.08%	74.72%
$4 < h \leq 8$	65.41%	64.84%	64.16%
$h > 8$	52.86%	51.21%	49.77%

These recalls are based on a subset of the test set where every user have in a group have provided feedback for the current context in a given time instance.

Test set	Recall@3	
	CtxInfl	Infl
All	64.60%	63.40%
Group size = 2	65.35%	64.38%
Group size = 3	61.90%	59.77%
Group size = 4	59.44%	57.18%
Group size = 5	61.95%	57.92%
$h \leq 4$	79.06%	78.83%
$4 < h \leq 8$	67.44%	66.70%
$h > 8$	54.48%	52.30%

Conclusion

Conclusion

Based on the result one can conclude that:

- Since the three methods performing best takes influence into account, this have a positive effect.
- Context influence performs better than plain influence.