Tackling Cold-Start Users in Recommender Systems with Indoor Positioning Systems

Emanuel Lacic, Dominik Kowald, Matthias Traub,
Granit Luzhnica, Joerg Simon and Elisabeth Lex
{elacic, dkowald, mtraub, gluzhnica, jsimon}@know-center.at, elisabeth.lex@tugraz.at

The cold-start problem

Users who have **not a single** or only **very few** item interaction data (e.g., ratings) available.

Typical approaches:

- **Unpersonalized** recommendation (e.g., MostPopular)
- **Personalized** recommendation when a **minimum number** of user-item interactions **is available**
  - For example, Matrix Factorization **fails** in "extreme" cold-start settings when there are no item interactions
- **Interaction surveys** (e.g., MovieLens):
  - Users need to fulfill a predefined number of interactions before getting recommendations
  - It is **often annoying** and **hard to immediately come up with a representative list** of item ratings

Reference & Framework

[1] E. Lacic, D. Kowald, M. Traub, G. Luzhnica, J. Simon and E. Lex. Tackling cold-start users in recommender systems with indoor positioning systems. In *Proc.* RecSys '15.

Test setup

Foursquare dataset:

$\#Items$	1,143,092	$\#Users$	2,153,471
$\#Ratings$	2,809,581	$\#Check-Ins$	1,021,970

Cold-start setup:
The interactions of all users that interacted with 10 items (= 2,783 out of 2,153,471 users) were put into the test set to be predicted.

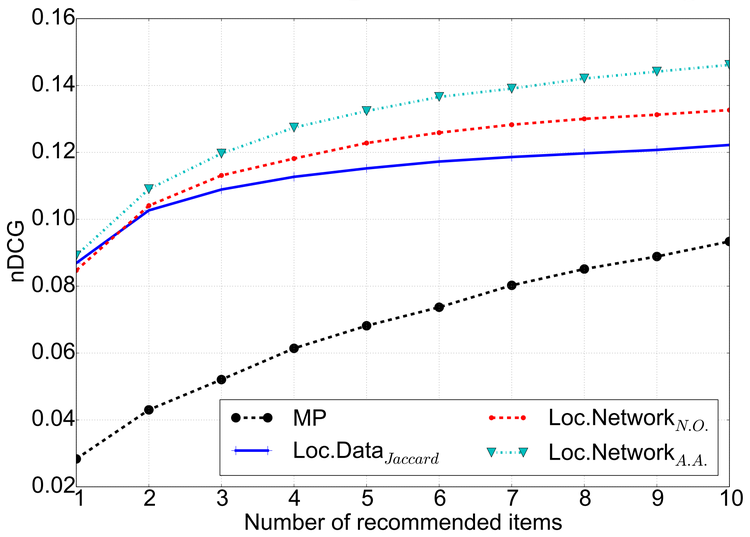
Approach

IPS-based Recommender

There exist a number of easily attainable technologies, or indoor positioning systems (IPS), to track indoor locations. Among them, BLE (Bluetooth Low Energy) beacons have gained importance and popularity, especially after Apple introduced the iBeacon protocol.

Preliminary Evaluation

Evaluation with respect to nDCG using the FourSquare dataset:



Approach	Algorithm	Similarity Name	Similarity Metric
MP	MostPopular	-	-
Loc.Data	UB CF	Jaccard's Coefficient	$sim(u,v) = \frac{\lvert\Delta(u) \cap \Delta(v)\rvert}{\lvert\Delta(u) \cup \Delta(v)\rvert}$
Loc.Network*	UB CF	Neighbourhood Overlap	$sim(u,v) = \frac{\lvert\Gamma(u) \cap \Gamma(v)\rvert}{\lvert\Gamma(u)\rvert + \lvert\Gamma(v)\rvert}$
Loc.Network*	UB CF	Adamic Adar	$sim(u,v) = \sum_{z \in \Gamma(u) \cap \Gamma(v)} \frac{1}{log(\lvert\Gamma(z)\rvert)}$

* Ties between two users are existent if they visited the same location within the same day and hour.