HI or AI?
Which is more relevant?

By Emil Ismalon
CTO & Founder
Collarity Inc.

www.collarity.com

Outline

- *Background*
- *Collarity implicit platform*
  - *Personalization*
  - *Communitization*
- *Explicit graphs vs. implicit graphs*
- *Social network dynamics*
- *Summary and conclusion*

Background

- *Collarity is a relevancy platform incorporating;*
  - *a personalized search engine working above a given index.*
    - *Web search*
    - *Site search for publishers*
  - *content recommendation system leveraging any type of content*
    - *Related stories / videos*
    - *Suggest searches / tags*
    - *Social navigation tools*
  - *Ad targeting both on SERP as well as content pages.*

Background

- *Does search engines algorithms actually learns from user interactions?*

Collarity suggest:

- ☑ *Implicit networks can leverage search, recommendation & Ads targeting.*
- ☑ *Implicit networks can naturally embed Social networks topology, abstractions and semantics.*
- ? *Social networks (e.g. Facebook) dynamics (linking dynamics) might be partially predicted by a combined latent space models and implicit networks models.*

Personalization & Communitization a Gradual perception

- *Language preferences, local/IP based information etc.*
- *Simple Lingual Disambiguation*
  - *The case of Java; is it programming OR island OR coffee?*
- *Form of preferable media*
  - *Textual OR visual?*
  - *Short Summary OR deep analysis?*

Personalization & Communitization a Gradual perception

Subtle ones

- *Understanding the right perception while interacting:*
  - *Pro OR anti*
  - *Believer OR skeptical*
  - *Sentiment*

- *Quantum physics article most suitable to a biologist*

"If you can't explain something to a six-year-old, you really don't understand it yourself".(Richard Feynman)

The ability to personalize

In order to materialize the subtle level of personalization an holistic approach is needed.

Personalization involves "intelligence" characteristic; abstraction, generalization and predictive clustering amongst other.

understanding *the user from many* ***<u>different aspects</u>***

- ***Implicit user profile – User is defined by Graph G_p (v,e) (edge labeled graph)***
- ***Topic based communities*** *-Clustering users around topics*
- ***Like minded communities – $C(\sum G_p) \rightarrow G_c(v,e)$***
  - *Cluster users in a latent information space*

Users implicit communities can serve as a natural enhancement for searching and browsing. (Collaborative filtering).
In this process we can construct ***implicit social networks****. Example: <u>Foo Graph</u>*

Evaluation test July 2008

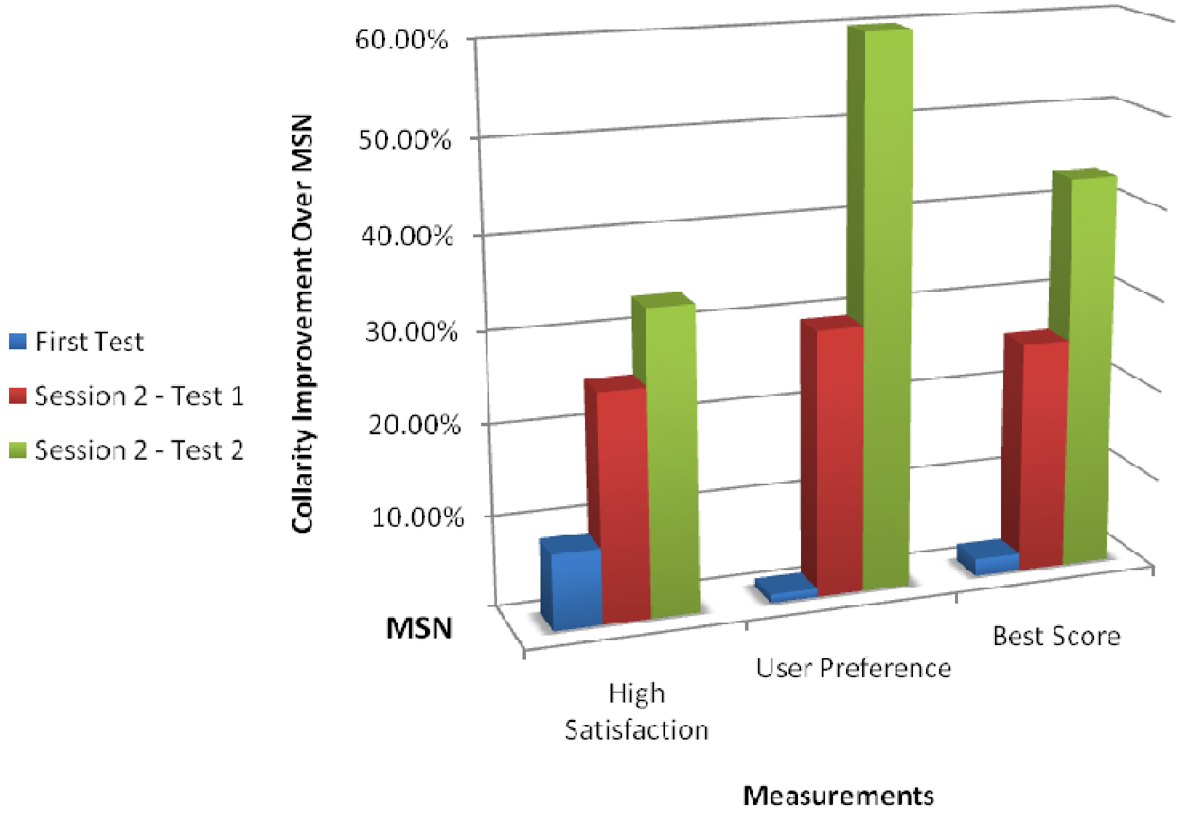
A relevancy comparison test was conducted in July 2008 between Collarity (with MSN index) and MSN, over a two weeks time course as a homework assignment for 36 graduate information students in Bar-Ilan University, Israel

**A small scale test was also conducted Vs. Yahoo Index showing similar results*

Different metrics where measured explicitly :

1. ***Best result*** *–best URL result chosen from both MSN OR Collarity SERPs.*
2. ***User Satisfaction Score*** *- reflects the relative satisfaction of the user from the SERPs presented to him for a test query.*
3. ***User Preference Score*** *- reflects the comparative preference of the user towards Collarity OR MSN results for a query.*

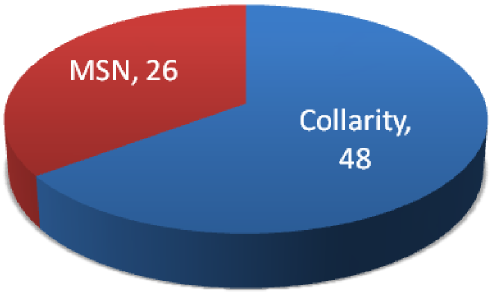
Improvement over time - *The relative improvement of Collarity over MSN in various measures (in percentage) at three different occasions throughout the test (Session 1 Test 1 q 1-20, Session 2 Test 1 q 60-80; Session 2 Test 2 q 80-100).*



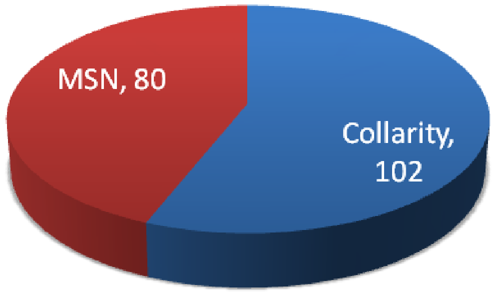
collarity

User Preference Score *- This criteria per query test presents the user's preference of one search engine over the other divided into 'Collarity', 'Somewhat Collarity', 'Neutral', 'Somewhat MSN'. 'MSN'.*

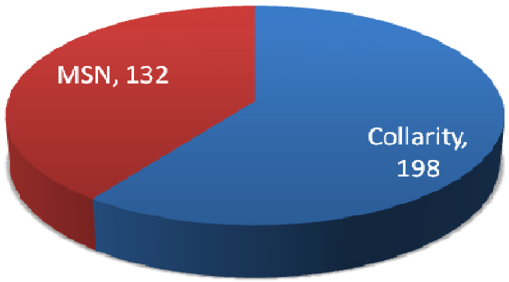
Absolute Preference



Partial Preference



Total Preference Score



The graphs on top show the cases of user preferences for Collarity and MSN for cases with absolute preference on the left and partial preference on the right. In the bottom we have a summary of the above two graphs, giving 2 points for absolute cases and 1 point for partial. The figure shows high percentages of improvement for Collarity over MSN: ***84.62%*** *for absolute preference,* ***27.5%*** *for the partial cases, and overall* ***50%****.*

Explicit graphs vs. Implicit graphs☑

Bringing the pieces together

- *A social network is basically a graph G_s (v,e) where each vertex is a member connected to his/her friends based on explicit friendship connection.*
- *An information based network is a graph $G_i(U, v, e)$ where each member is connected to each other depending on their shared content interest (implicit profiles).*
- *Embedding one graph to another is straightforward. Consider the foo camp graph (implicit graph): embedding into it a social graph (the real social connections of the scientists) is just adding new edges → recommendation, search and browsing using the new connection (collaborative filtering) just get better.*

Social Dynamics

- We can impose dynamics on the social graph (implicit or/and explicit) as a basis for recommendation.
- One option: imagine a social widget that enables users to see their most closest like-minded neighbors. Obviously, some users will connect to their content-friends → we bring graph G_s closer to graph G_i, and thus enable better information flow.

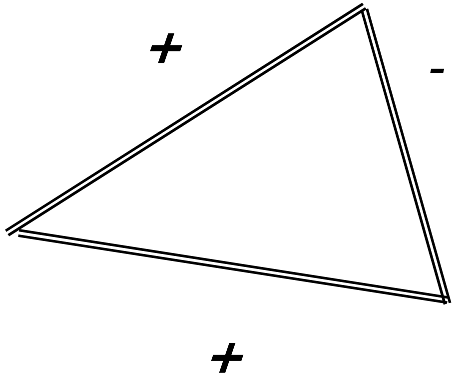
- We experimented with a model based on This *similarity criterion* for social influence is an example of social comparison theory in which individuals are most influenced by others who are similar:

$$p_{ind} \propto + f(dist)$$

$$p_{clust} \propto - f(dist^{-1});$$

Show video

Another option: define an objective function, which is the driving force of the social network, intended to bring the system closer to a maximal flow scenario.
For instance: ***Heider Balance theory*** *(1958) – in terms of like minded users.*



my friend's friend is my friend
my friend's enemy is my enemy
my enemy's friend is my enemy
my enemy's enemy is my friend

Define a balance Index :

$$\beta = \frac{T_{balanced}}{T_{tot}}$$

where T is the number of triads

1) *Marcov process - change each triad randomly and accept change if balance Index increases → path dependent, non-ergodic, local minima.* [1]

2) *Define a continuous approach based on differential equations*[1]
3) *Statistical physics approach – Ising Hamiltonian model*

1) Some recent attempts to simulate the Heider balance problem*, Krzysztof Ku lakowski, 2008*
http://arxiv.org/abs/physics/0612197v2

2) Spontaneous coalition forming. Why some are stable? *Serge Galam, 2002*
http://arxiv.org/PS_cache/cond-mat/pdf/0212/0212444v1.pdf

Summary and Conclusion

- Holistic approach for personalization can give rise to better understanding of users' interest and intents
- Implicit communities can serve as a base for collaborative filtering models aiming for better search results
- Implicit information networks can be easily embedded into social networks (explicit), increasing content relevancy.
- AI↔HI achieves better results

Further readings

Pros & cones personalized search

http://searchengineland.com/070309-081324.php

Dynamic Social Network Analysis using Latent Space Models

http://books.nips.cc/papers/files/nips18/NIPS2005_0724.pdf

Constraint-based Personalization Model: Multi-Channel Messaging

http://www.research.att.com/~rjana/TothNagboth.pdf

A Hybrid Web Personalization Model. Based on Site Connectivity.

Miki Nakagawa, Bamshad Mobasher ... on non-sequential models, such as association rules and ...

http://maya.cs.depaul.edu/~mobasher/papers/NM03b.pdf

Social Balance Theory

The Psychology of Interpersonal Relations, F. Hieder, 1958

Revisiting Heider's Balance Theory for many agents

http://www.au.af.mil/au/awc/awcgate/lanl/social_balance_0405041.pdf

Social Dynamics:

Some recent attempts to simulate the Heider balance problem, Krzysztof Ku lakowski, 2008

http://arxiv.org/abs/physics/0612197v2

Spontaneous coalition forming. Why some are stable? Serge Galam, 2002

http://arxiv.org/PS_cache/cond-mat/pdf/0212/0212444v1.pdf