



# Personalization in Local Search

Personalization of Content Ranking in the Context of  
Local Search

Philip O'Brien, Xiao Luo, Tony Abou-Assaleh, Weizheng Gao, Shujie Li

Research Department, GenieKnows.com

September 17, 2009



# About GenieKnows.com

- Based in Halifax, Nova Scotia, Canada
- Established in 1999
- ~35 People
- Online Advertising Network
  - 100 to 150 million searches per day
- Search Engines (local, health, games)
- Content Portals



What e.g. "pizza"

genieknows.com

Where e.g. "grifton st, halifax"

halifax, nova scotia, canada

GO

Welcome guest Log in or Sign up  
Halifax, NS Weather: Light Drizzle, 18 °C

one search box

Home > Canada > Nova Scotia > Halifax > Genieknows.com

Categories

- > Advertising (1)
- > Computer Software (1)

Results for: genieknows.com in Halifax

Sort Results: Rank Street Name

1 **Genieknows.com**  
 1567 Argyle St, Halifax, NS  
 +1 902-431-4847

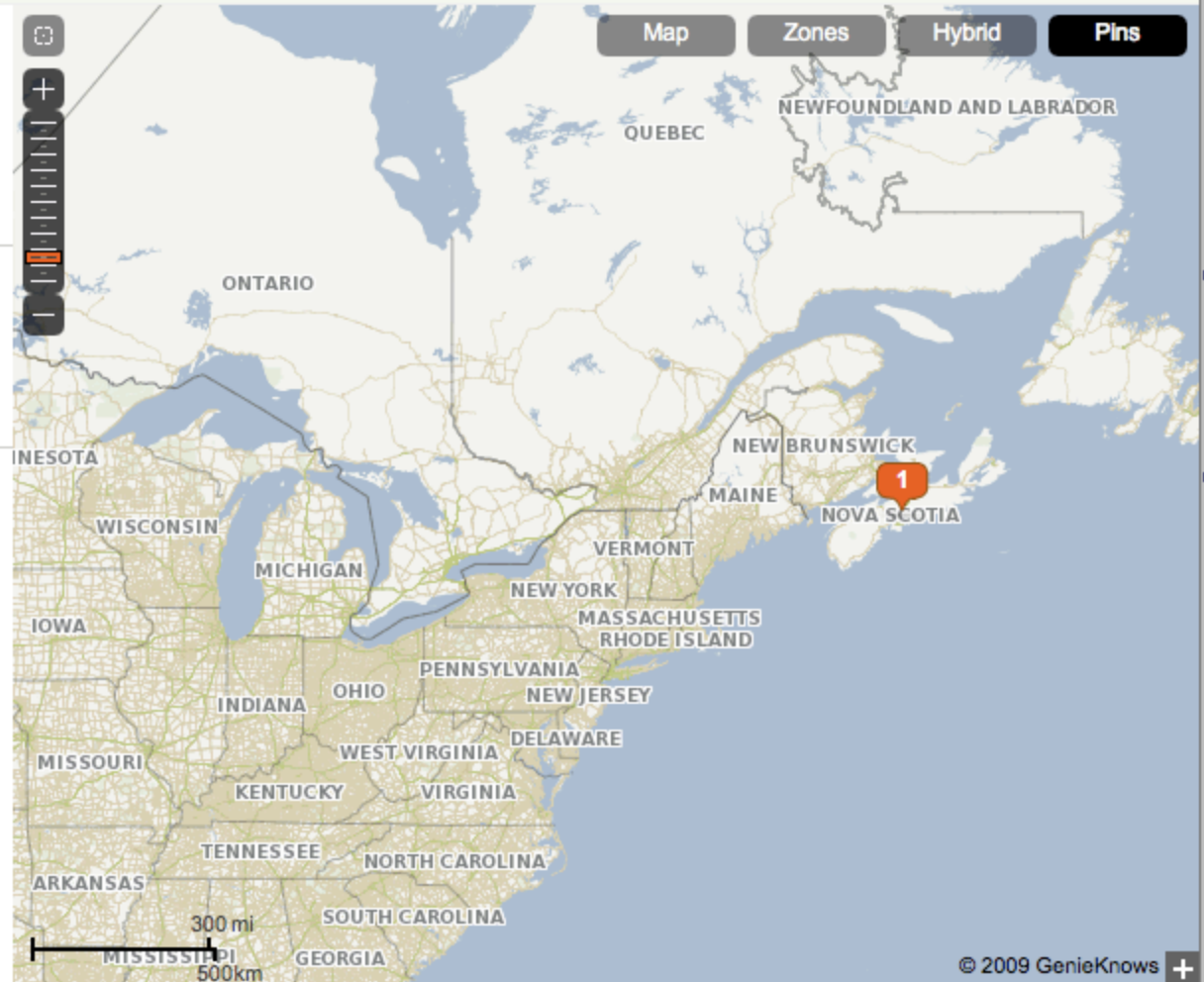


map this search nearby

Advertise with us

Add your business

previous 1 next



© 2009 GenieKnows



# About Tony Abou-Assaleh

- Director of Research at GenieKnows
  - Since 2006
  - Build search engines
  - Other internal R&D initiatives
- Lecturer at Brock University, St. Catharines, Canada
  - 2005 - 2006
- GNU grep official maintainer

# Agenda

- Introduction
- Related Work
- Our Approach
- Experiments
- Conclusion & Future Work

# Agenda

- **Introduction**
- Related Work
- Our Approach
- Experiments
- Conclusion & Future Work

# Introduction

- Local Search
  - What? Why?
- Personalization
  - What? How? Why?
- Assumptions
- Objectives

# What is Local Search?

- Local Search vs. Business Directory
- Contains:
  - Internet Yellow Pages (IYP) Business Directory
  - Enhanced business listings
  - Map
  - Ratings and Reviews
  - Articles and editorials
  - Pictures and rich media
  - Social Networking





What e.g. "pizza"

restaurants

Where e.g. "grifton st, halifax"

1567 Argyle Street, Halifax, NS

GO

one search box

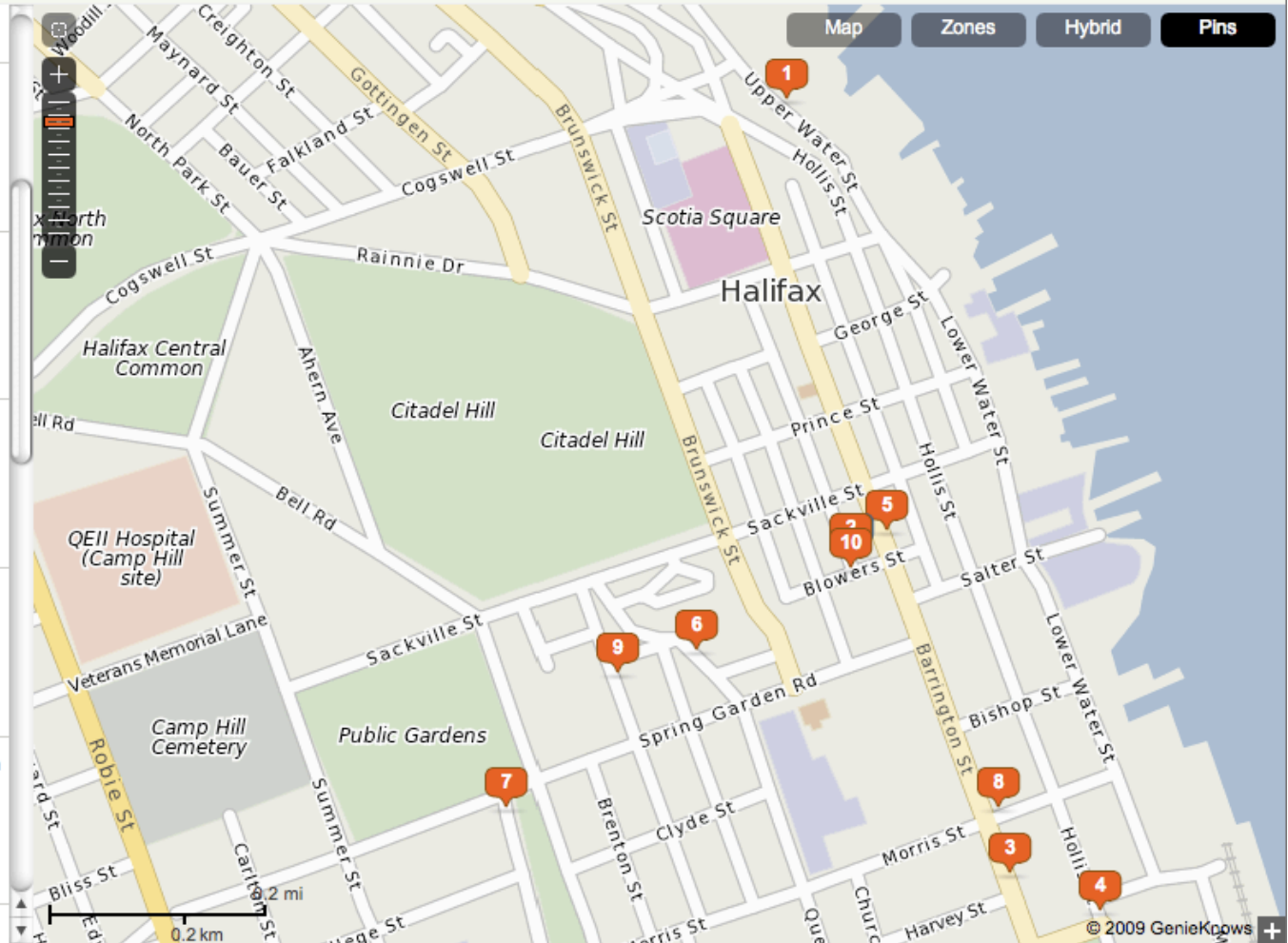
Welcome guest Log in or Sign up

Halifax, NS Weather: Light Drizzle, 18 °C

Home > Canada > Nova Scotia > Halifax > 1567 Argyle St > Restaurants

Sort Results: Rank Street Name Distance

- 1** [44 North Restaurant](#) 0.68 km  
 1919 Upper Water St, Halifax, NS  
 +1 902-428-7852  
 ★★★★★ map this search nearby
- 2** [Opa Restaurant](#) 0.00 km  
 1565 Argyle St, Halifax, NS  
 +1 902-492-7999  
 ★★★★★☆ map this search nearby
- 3** [Talay Thai Restaurant](#) 0.53 km  
 1261 Barrington St, Halifax, NS  
 +1 902-404-3700  
 ★★★★★☆ map this search nearby
- 4** [Trinity Restaurant & Catering](#) 0.65 km  
 1333 South Pk, Halifax, NS  
 +1 902-423-8428  
 ★★★★★☆ map this search nearby
- 5** [Star Anise Vietnamese Noodles Restaurant](#) 0.06 km  
 1571 Barrington St, Halifax, NS  
 +1 902-422-2602  
 ★★★★★☆ map this search nearby
- 6** [Chabaa Thai Restaurant](#) 0.27 km



# Why Local Search?

- Good for end users
- Good for businesses
- Good for our company
- Interesting research problems
- No market leader
- Could be the next big thing

# What is Personalization?

- No personalization:
  - Everybody gets the same results
- Personalization:
  - User may see different results
- Personalization vs. customization

# What to Personalize?

- Ranking
- Snippets
- Presentation
- Collection
- Recommendations

# How to Personalize?

- Search history
- Click history
- User profiles - interests
- Collaborative filtering

# Why Personalization?

- One size does not fit all
- Ambiguity of short queries
- Improve per-user precision
- Improve user experience
- Targeted advertising \$\$\$

# Assumptions

- Interests are location dependent
- Long-term interests
- Implicit relevance feedback
- Relevance is location dependent
- Relevance is category dependent
- User cooperation
- Single-user personalization

# Objectives

- General framework for personalization of spatial-keyword queries
- User profile representation
- Personalized ranking
- Improve over baseline system



# Agenda

- Introduction
- **Related Work**
- Our Approach
- Experiments
- Conclusion & Future Work

# Related Work

- User Profile Modeling
- Personalized Ranking

# User Profile Modeling

- Topic based (Liu et al, 2002)
  - Vector of interests
  - Explicit: how to collect data?
  - Implicit: relevance feedback
- Click based (Li et al, 2008)
  - Implicit feedback from click through data
  - Require a lot of data
- Ontological profiles (Sieg et al, 2007)
- Hierarchical representations (Huete et al, 2008)

# Personalized Ranking

- Web, desktop, and enterprise search
- Local search?
- Strategies:
  - Implicit
  - Clicks as relevance feedback
  - Query topic identification
  - Collaborative filtering
  - Learning algorithms

# Agenda

- Introduction
- Related Work
- **Our Approach**
- Experiments
- Conclusion & Future Work

# Our Approach

- Problem formulation
- Ranking Function Decomposition
- Business Features
- User Profile
- User Interest Function
- Business-specific Preference Function

# Problem Formulation

- Query: keywords + spatial (geographic) context
- Ranking function:

Relevant Results  $\times$  User Profiles  $\times$  Location  $\rightarrow$  Real Number

- Online personalized ranking:
  - Optimization of an objective function over rank scores

# Ranking Function Decomposition

- Final rank = weighted combination of:
  - Baseline rank
  - User rank
  - Business rank



# Ranking Function Decomposition

- Final rank = weighted combination of:
  - **Baseline rank**
  - User rank
  - Business rank

# Baseline Rank

- Okapi BM25F on textual fields
- Distance from query centre
- Other non-textual features

# Business Features

- List of categories
  - 18 top level, 275 second level
- Terms
  - Vector-space model
- Location
  - Geocoded address
- Meta data
  - Year established, number of employees, languages, etc.

# User Profile

- Local Profile
  - For each geographic region (city)
  - For each category
  - Needs at least 1 query
- Global Profile
  - Aggregation of local profiles
  - Used for new city and category combination

# Local Profile

- Category interest score
  - Fraction of queries in this category
  - Fraction of clicks in this category
- Number of queries
- Terms vector-space model
- Clicks (business, timestamp)

# Global Profile

- Estimated global category interest score
  - Aggregated over all cities
  - Weighted combination of interest scores
  - Weights derived from query volume
  - Estimated using a Dirichlet Distribution

# Ranking Function Decomposition

- Final rank = weighted combination of:
  - Baseline rank
  - **User rank**
  - Business rank

# User Interest Function

- Rank (business, user, query) =  
Category interest score  $\times$  Term similarity  $\times$  Click count
- Averaged over all categories of the business
- Term similarity: cosine similarity
- Click count: capture navigational queries



# Ranking Function Decomposition

- Final rank = weighted combination of:
  - Baseline rank
  - User rank
  - **Business rank**

# Business-specific Preference Function

- Rank (business, user, city, category) =  
Sum of query dependent click scores  
+  
Sum of query independent click scores
- Click scores are time discounted
  - 1 year windows
  - 1 week intervals
- Parameter to control relative importance of query-dependency

# Agenda

- Introduction
- Related Work
- Our Approach
- **Experiments**
- Conclusion & Future Work

# Experiments

- Data
- Procedure
- Results
- Discussion

# Data

- 22 Million businesses
- 30 participants
- Only 12 with sufficient queries
- 2388 queries
- 1653 unique queries

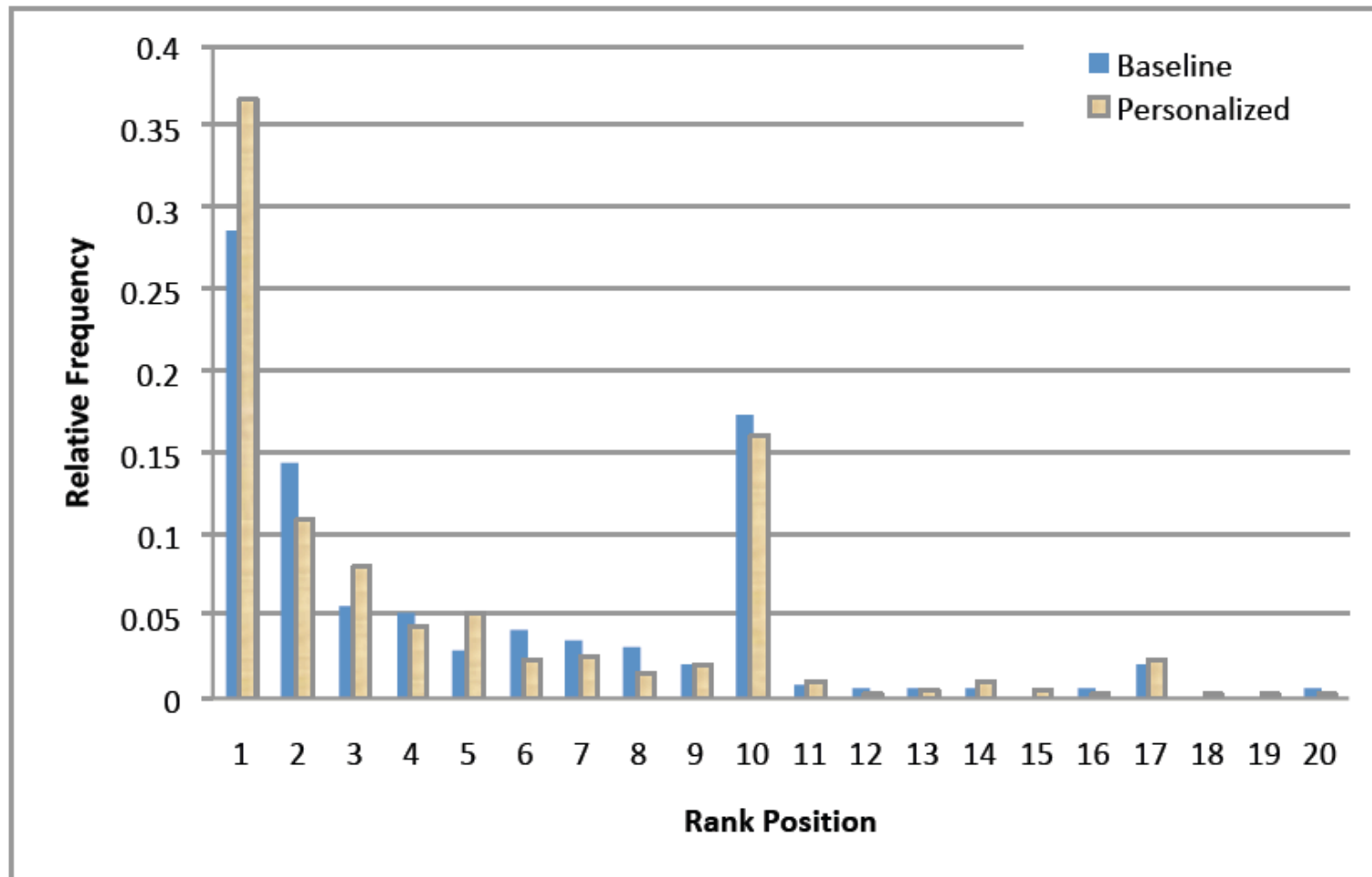
# Procedure

- Types of tasks:
  - Navigational, browsing, information seeking
- 5-point explicit relevance feedback
- Ranking algorithm
  - Baseline vs. personalized
  - Alternates every 2 minutes
  - Identical interface
  - No bootstrapping phase

# Results

- Measures:
  - Mean Average Precision - MAP
  - Mean Reciprocal Rank - MRR
  - Normalized Discounted Cumulative Gain - nDCG

# Results





# Results

<i>Metric</i>	<i>Baseline</i>	<i>Personalized</i>	<i>Improvement</i>
MAP	0.612117453	0.774551253	26.5%
MRR	0.811232333	0.930912438	14.8%

- Welch two-sample t-test:
  - Significant improvement
  - MAP:
    - 95% confidence,  $p=0.04113$
  - MRR:
    - 95% confidence,  $p=0.02192$

# Results

- nDCG@10
- 16 randomly selected queries
- Not significant

<i>Query</i>	<i>Baseline</i>	<i>Personalized</i>
Golf	0.974	0.953
Photo	0.958	0.992
Real estate	0.946	0.921
Apartment	0.903	0.852
Coffee	0.950	0.960
Photography	0.591	0.864
Restaurants	0.859	0.922
Books	0.978	0.986
Home inspection	0.980	0.979
Camping	0.873	0.724
Hotels	0.946	0.976
Sushi	1.000	0.944
Spa	0.908	0.957
Parks	0.940	0.987
Martial arts	0.945	0.965
Clothes	0.840	0.952
<b>Average:</b>	0.912	0.934

# Agenda

- Introduction
- Related Work
- Our Approach
- Experiments
- **Conclusion & Future Work**

# Contributions

- Personalization framework for spatial-keyword queries
- Model for user profiles
- Local and global profiles
- Address data sparseness problem
- Personalized ranking function
  - Interests, clicks, terms
- Empirical evaluation
  - Significant improvement over the baseline system

# Future Work

- Modeling of short-term interests
- Modeling of recurring interests
- “Learning to Rank” algorithms
- Multi-user personalization
  - Recommender system
- Incorporate on [www.genieknows.com](http://www.genieknows.com)



# Thanks you!

<http://www.genieknows.com>

<http://tony.abou-assaleh.net>

[taa@genieknows.com](mailto:taa@genieknows.com)

@tony\_aa

# Questions

- Can I access your data?
- Did you do parameter tuning?
- Did users try to test/cheat the system?
- What is the computational complexity?
- Any confounding variables?