Content, Connections, and Context

Daniel Tunkelang
Principal Data Scientist at LinkedIn
Recommendation Products at LinkedIn

**Jobs You May Be Interested In**

**Talent Match**

**CAP**

**Similar Profiles**

**Companies**
Recommendations, similar companies search, peer companies, and company browse maps, company products and services browse maps

**Related search**

**Profile browse maps**

**Network updates**

**Jobs browse maps**

**Connections**

**Groups**
Recommendations, similar groups search

**Similar jobs**

**Referral Engine**

**Events You May Be Interested In**

**Ad matching engine**

\[ pCTR = \frac{\text{number, creative, advertiser, context, inventory, OCTR)} \]

**News**
LinkedIn Today: See all Top Headlines for You

*Image of various LinkedIn product sections*
More than 50%

Recommendations drive:

> 50% of connections

> 50% of job applications

> 50% of group joins
Inputs for Recommender Systems

**Content**

**Social Graph**

**Behavior**

- Queries
- Page Views
- Actions

- Your Groups (48)
  - Visual Analytics
  - The R Project for Statistical Computing
  - Data Mining, Statistics, and Data Visualization
  - Online Experimentation Professionals
Take-Aways

**Content** is king.

**Connections** provide social dimension.

**Context** determines where and when a recommendation is appropriate.
What is the goal of recommendations?

- Growth
- Engagement
- Revenue
What is the goal of recommendations?

- Growth (X)
- Engagement (X)
- Revenue
Users First!

Provide relevant content

and establish social connections

in appropriate context.
Content is King
Goal-Seeking is about Content
Content (Ir)relevance

- No right answer, but many wrong answers.

[Voorhees, 2004]

Example: Related Searches

LinkedIn search results for "web developer" showing suggestions like "web designer", "php", "ruby on rails", "developer", "graphic designer", "web development", "software engineer", and "lamp".
Collaborative Filtering as Content Signal

- Use temporal locality within sessions.
- Find queries with clicks on similar results.
- Look for query overlap.
- Learn more at CIKM! [Reda et al, 2012]
Example: Jobs You Might Be Interested In

Daniel Tunkelang
Principal Data Scientist at LinkedIn
Mountain View, California | Internet

Current: LinkedIn, New York CTO Club
Previous: Google, Endeca, Context Integration
Education: PhD, Computer Science at Carnegie Mellon University

Information Retrieval    Data Mining    Recommender Systems
Machine Learning    Search    Patents    Algorithms
Natural Language Processing    Text Mining    Information Extraction
Text Classification    Text Analytics    Collaborative Filtering    Analytics
Information Visualization    Web Mining    Human-computer Interaction
Predictive Analytics    SEO    Blogging

JOBS YOU MAY BE INTERESTED IN

Business Intelligence Manager
Eliason Corporation - ...

Senior Data Scientist
SCEA - San Francisco, US...

Senior Data Scientist
(Product...
Netflix - Los Gatos, CA

director of Data Science & Analytics
Glassdoor.com - San Fran...

Feedback | See more »
Content Signals Dominate Social Signals

### Job
- Senior search software engineer at LinkedIn
- Location: Mountain View, San Francisco Bay Area
- URL: http://www.linkedin.com
- title: industry
- geo: description
- company: functional area

### User Base
- Filtered
- Candidate
  - General expertise
  - specialties
  - education
  - headline
  - geo
  - experience

### Candidate
- Anthony D'Auria
  - Current Position title
  - summary
  - tenure length
  - industry
  - functional area

### Corpus Stats
- Transition probabilities
- Connectivity
  - yrs of experience to reach title
- education needed for this title

### Matching
- **Binary**
  - Exact matches: geo, industry, ...
- **Soft**
  - transition probabilities, similarity, ...
- **Text**
  - $w_{t,d} = tf_t \cdot \log \frac{|D|}{|\{t \in d\}|}$
  - $\cos \theta = \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\| \|\mathbf{v}_2\|}$
- Similarity (candidate expertise, job description)
  - 0.56
- Similarity (candidate specialties, job description)
  - 0.2
- Transition probability (candidate industry, job industry)
  - 0.43
- Title Similarity
  - 0.8
- Similarity (headline, title)
  - 0.7
Provide relevant content

and establish social connections

in appropriate context.
Connections
Connections are Social Dimensions

Current
Principal Data Scientist at LinkedIn
Member at New York CTO Club

Past
Tech Lead at Google
Chief Scientist at Endeca
Consultant at Context Integration

Education
Carnegie Mellon University
Massachusetts Institute of Technology
St. Hilda's & St. Hugh's
Example: People You May Know

Cameran Evans (2nd)
Manager, Business Intelligence at VMware
San Francisco Bay Area

Miroslav Dudik (2nd)
Researcher at Microsoft
Greater New York City Area

Ledio Ago (2nd)
Engineering Lead, Manager at Splunk
San Francisco Bay Area

Rachael Carbone
Intern: Speech Language Pathologist Assistant at Spaulding Rehabilitation Hospital
Greater Boston Area
Relationships are Social
Shared Connections as a Signal

People You May Know

Linus Upson 2nd
VP, Engineering at Google
San Francisco Bay Area

Connect 8 shared connections

Linus Upson 2nd
See all »
The Power of Social Proof

Tips to increase your likelihood of a response by 50%.

1. Mention people you know in common
   - Adam Nash
   - Craig Nevill-Manning
   - Maria Stone

2. Limit your message to 100 words - make it direct and to the point
Beyond Triadic Closure

- Triads suggest and affect relationships. [Simmel, 1908], [Granovetter, 1973]

- Triangle closing is a Big Data problem. [Shah, 2011]

- Use machine learning to rank candidates.
Provide relevant content and establish social connections in appropriate context.
"to every thing there is a season"
[Ecclesiastes 3:1]
One Platform, Many Users, Many Needs

http://blog.lab42.com/the-linkedin-profile
No Weekend Meetings!

**Weekdays**

**Weekends**

- navigational queries
- exploratory queries
Different Devices for Different Needs
Time and Place for a Career Change

# of colleagues in your network who left in last 6 months
Active vs. Passive Job Seekers

>10x more likely to reply if active

[Rodriguez et al, 2012]
Provide relevant content and establish social connections in appropriate context.
Recap

**Content** is king.

**Connections** provide social dimension.

**Context** determines where and when a recommendation is appropriate.
Thank You!

LinkedIn Members (Millions)

2004 2005 2006 2007 2008 2009 2010 2011

2 4 8 17 32 55 100

175M+

62% non U.S.

2/sec

We’re Hiring!

25th Most visit website worldwide (Comscore 6-12)

>2M Company pages

85%

Fortune 500 Companies use LinkedIn to hire

Learn more at http://data.linkedin.com/