Probabilistic Graphical Models for Credibility Analysis in Evolving Online Communities

Subhabrata Mukherjee
Max Planck Institute for Informatics, Germany

smukherjee@mpi-inf.mpg.de
Outline

- Motivation
- Prior Work and its Limitations
- Credibility Analysis
  - Framework for Online Communities
  - Temporal Evolution of Online Communities
  - Credibility Analysis of Product Reviews
- Conclusions
Online Communities as a Knowledge Resource

- Online communities are massive repositories of knowledge accessed by regular users and professionals
  - 59% of adult U.S. population and half of U.S. physicians rely on online resources [IMS Health Report, 2014]
  - 40% of online consumers consult online reviews before buying products [Nielson Corporation, 2016]

- However their usability is restricted due to serious credibility concerns (e.g., spams, misinformation, bias etc.)
Concerns

Misinformation for health can have hazardous consequences

“Rapid spread of misinformation online” --- one of top 10 challenges as per The World Economic Forum
Outline

- Motivation
- Prior Work and its Limitations
- Credibility Analysis
  - Framework for Online Communities
  - Temporal Evolution of Online Communities
  - Credibility Analysis of Product Reviews
- Conclusions
Truth Finding

- Structured data (e.g., SPO triples, tables, networks)
- Objective facts (e.g., Obama_BornIn_Hawaii vs. Obama_BornIn_Kenya)
- No contextual data (text)
- No external KB, metadata

Linguistic Analysis

- Unstructured text
- Subjective information (e.g., opinion spam, bias, viewpoint)
- External KB (e.g., WordNet, KG)
- No network / interactions, metadata
Research Questions

1. How can we *jointly* leverage users, network, and context for credibility analysis in online communities?
2. How can we model users’ *evolution*?
3. How can we deal with *limited* data?
4. How can we generate *interpretable* explanations for credibility verdict?
Contributions

● Credibility Analysis Framework for Online Communities
  ↘ Classification: Health Communities [SIGKDD 2014]
  ↘ Regression: News Communities [CIKM 2015]

● Temporal Evolution of Online Communities
  [ICDM 2015, SIGKDD 2016]

● Credibility Analysis of Product Reviews
  [ECML-PKDD 2016, SDM 2017]
Outline

● Motivation
● Prior Work and its Limitations
● Credibility Analysis
   → Framework for Online Communities
   → Temporal Evolution of Online Communities
   → Credibility Analysis of Product Reviews
● Conclusions
What is Credibility?

“A statement is credible if it is reported by a trustworthy user in an objective language”

“Trustworthy users corroborate each other on credible statements”
Credibility Analysis Framework for Classification

Problem: Given a set of posts from different users, extract credible statements (subject-predicate-object triples like DrugX_HasSideEffect_Y) from trustworthy users.
Credibility Analysis Framework for Classification

Problem: Given a set of posts from different users, extract credible statements (subject-predicate-object triples like DrugX HasSideEffect Y) from trustworthy users.
Network of Interactions: Cliques

Each user, post, and statement is a random variable with edges depicting interactions. Variables have observable features (e.g., authority, emotionality).

A clique is formed between each user writing a post containing a statement.

Statements: An IE tool generates candidate triple patterns like:

Xanax-causes-headache,
Xanax-gave-demonic-feel

Potentially thousands of such triples, with only a handful of credible ones.
Network of Interactions: Cliques

Each user, post, and statement is a random variable with edges depicting interactions

Statements: An IE tool generates candidate triple patterns like:

Xanax-causes-headache,
Xanax-gave-demonic-feel

Potentially thousands of such triples, with only a handful of credible ones

Idea: Trustworthy users corroborate on credible statements in objective language
Conditional Random Field to Exploit Joint Interactions (Users + Network + Context)

Partial Supervision: Expert stated (top 20%) side-effects of drugs as partial training labels. Model predicts labels of unobserved statements.

How to complement expert medical knowledge with large scale non-expert data?
Semi-Supervised Conditional Random Field

1. Estimate user trustworthiness: \[ \text{trust} = \frac{\sum_{\text{Statement}} \text{label(Statement} = \text{True})}{(\text{Statement})} \]

2. Estimate label of unknown statements \( S_u \) by Gibbs Sampling:

\[ Pr(S_u \mid \text{Post, User, S}_{\text{labeled}}; W) \propto \prod_{\Phi \in \text{Clique}} \text{trust(user)} \cdot \Phi(S_u, \text{Post}_\text{features}, \text{User}_\text{features}; W) \]

3. Maximize log-likelihood to estimate feature weights:

\[ W' = \arg \max_{W} \sum_{S_u} Pr(S_u \mid \text{Post, User, S}_{\text{labeled}}; W) \cdot \log Pr(S_u, S_{\text{labeled}} \mid \text{Post, User}; W) \]

4. Apply E-Step and M-Step till convergence
Healthforum Dataset

- Healthboards.com community (www.healthboards.com) with 850,000 registered users and 4.5 million posts

- Expert labels about drugs from MayoClinic (www.mayoclinic.org)
  - 6 widely used drugs for experimentation
What constitutes credible language?

compunction
anxiety
embarrassment
misery
distress

certainty
sympathy
self-esteem
eagerness
coolness
What constitutes credible language?

determiner (this, that, ..)
question (what, why, ..)
conditional (if)
adverb (maybe, probably, ..)
modality (might, could, ..)

determiner (this, that, ..)
negation (not, never, ..)
second person (you, ..)
conjunction (therefore, consequently, ..)
Credibility Analysis Framework for Regression

In many online communities users rate items on their quality
Credibility Analysis in News Communities

Sources
- trunews.com

Topics
- Climate Change

Sources / Users
- Scientificamerican.com
- snopes.com
- user-donald

Articles
- “Global warming is a hoax”

Reviews & Ratings
- scientific analysis, 1.5/5, conspiratory theory

However, user feedback is often subjective; influenced by their bias and viewpoints
Credibility Analysis Framework for Regression

We use CRF to capture these mutual interactions in news communities (e.g., newstrust.net, digg, reddit) to jointly rank all of the underlying factors.

Idea: Trustworthy sources publish objective articles corroborated by expert users with credible reviews/ratings.
Related to Ensemble Learning, Learning to Rank
How to incorporate *continuous ratings* instead of discrete labels in CRF?

Probability Mass Function for discrete labels:

\[ p(y|X) = \frac{\exp(\Psi)}{\sum_y \exp(\Psi)} \]

Probability Density Function for continuous ratings:

\[ p(y|X) = \frac{\exp(\Psi)}{\int_{-\infty}^{\infty} \exp(\Psi) \, dy} \]
Energy Function to Combine All

\[
\psi(y, s, d, \langle u \rangle, \langle r \rangle) = -\sum_u \alpha_u I_u(d) (y - \text{SVR}_u)^2
\]

-partitions the user space

user expertise

error of predictor SVR

\[
-\sum_s \beta_s I_s(d) (y - \text{SVR}_s)^2 - \gamma_1 (y - \text{SVR}_L)^2 - \gamma_2 (y - \text{SVR}_T)^2
\]

source trustworthiness language objectivity topical perspective
How to incorporate continuous ratings instead of discrete labels in CRF?

- We show that a certain energy function for clique potential --- geared for reducing mean-squared-error --- results in multivariate gaussian p.d.f.!!!

\[
P(y|X) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2} (y - \mu)^T \Sigma^{-1} (y - \mu)\right)
\]

- Constrained Gradient Ascent for inference

Subhabrata Mukherjee and Gerhard Weikum: CIKM 2015
**Predicting Article Credibility Ratings in Newstrust.net**

<table>
<thead>
<tr>
<th>Model</th>
<th>Only Title MSE</th>
<th>Title &amp; Text MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language Model: SVR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language (Bias and Subjectivity)</td>
<td>3.89</td>
<td>0.72</td>
</tr>
<tr>
<td>Explicit Topics</td>
<td>1.74</td>
<td>1.74</td>
</tr>
<tr>
<td>Explicit + Latent Topics</td>
<td>1.68</td>
<td>1.01</td>
</tr>
<tr>
<td>All Topics (Explicit + Latent) + Language</td>
<td>1.57</td>
<td>0.61</td>
</tr>
<tr>
<td><strong>News Source Features and Language Model: SVR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>News Source</td>
<td>1.69</td>
<td>1.69</td>
</tr>
<tr>
<td>News Source + All Topics + Language</td>
<td>0.91</td>
<td>0.46</td>
</tr>
<tr>
<td><strong>Aggregated Model: SVR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Users + All Topics + Language + News Source</td>
<td>0.43</td>
<td>0.41</td>
</tr>
<tr>
<td><strong>Our Model: CCRF+SVR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User + All Topics + Language + News Source</td>
<td>0.36</td>
<td>0.33</td>
</tr>
</tbody>
</table>

*Progressive decrease in mean squared error with more network interactions, and context*
Take-away

● Semi-supervised and Continuous CRF to jointly identify trustworthy users, credible statements, and reliable postings in online communities

● A framework to incorporate richer aspects like user expertise, topics / facets, temporal evolution etc.
Motivation

Prior Work and its Limitations

Credibility Analysis

Framework for Online Communities

Temporal Evolution of Online Communities

Credibility Analysis of Product Reviews

Conclusions
Temporal Evolution

- Online communities are dynamic, as users join and leave; acquire new vocabulary; evolve and mature over time

- Trustworthiness and expertise of users evolve over time

How to capture evolving user expertise?
Illustrative Example for Review Communities

- Consider following camera reviews by the same user John:

  “My first DSLR. Excellent camera, takes great pictures with high definition, without a doubt it makes honor to its name.”  
  [Aug, 1997]

  “The EF 75-300 mm lens is only good to be used outside. The 2.2X HD lens can only be used for specific items; filters are useless if ISO, AP,... . The short 18-55mm lens is cheap and should have a hood to keep light off lens.”  
  [Oct, 2012]
The EF 75-300 mm lens is only good to be used outside. The 2.2X HD lens can only be used for specific items; filters are useless if ISO, AP,... . The short 18-55mm lens is cheap and should have a hood to keep light off lens.

Consider following camera reviews by John:

“My first DSLR. Excellent camera, takes great pictures with high definition, without a doubt it makes honor to its name.”

How can we quantify this change in users’ maturity / experience?

How can we model this evolution / progression in users’ maturity?

Mukherjee et al.: ICDM 2015, SIGKDD 2016
Prior Work: Discrete Experience Evolution

1. Users at similar levels of experience have similar facet preferences, and rating style (McAuley and Leskovec: WWW 2013)

2. Additionally, our work exploits similar writing style (Mukherjee, Lamba and Weikum: ICDM 2015)

Assumption: At each timepoint a user remains at the same level of experience, or moves to the next level
Language Model (KL) Divergence Increases with Experience

Experienced users have a distinctive writing style different than that of amateurs
Prior Work: Discrete Experience Evolution

1. Users at similar levels of experience have similar facet preferences, and rating style (McAuley and Leskovec, WWW 2013)

2. Additionally, our work exploits similar writing style (Mukherjee, Lamba and Weikum, ICDM 2015)

Assumption: At each timepoint a user remains at the same level of experience, or moves to the next level
Continuous Experience Evolution
(Mukherjee, Günnemann and Weikum, SIGKDD 2016)
Continuous Experience Evolution: Assumptions

- **Continuous-time process**, always positive
- **Markovian assumption**: Experience at time $t$ depends on that at $t-1$
- **Drift**: Overall trend to increase over time
- **Volatility**: Progression may not be smooth with occasional volatility
  E.g.: series of expert reviews followed by a sloppy one
Geometric Brownian Motion

We show these properties to be satisfied by the continuous-time stochastic process: Geometric Brownian Motion
Language Model (LM) Evolution

- Users' LM also evolve with experience evolution
- Smoothly evolve over time preserving Markov property of experience evolution
- Variance of LM should change with experience change
- Brownian Motion to model this desiderata:

\[ \beta_{t,z,w} \sim \text{Normal} \left( \beta_{t-1,z,w}, \sigma \cdot |e_t - e_{t-1}| \right) \]
Inference

- Topic Model (Blei et al., JMLR ’03)
- Users (Author-topic model, Rosen-Zvi et al., UAI ’04)
- Continuous Time (Dynamic topic model, Wang et al., UAI ’08)
- Continuous Experience (this work)
Sampling based Inference

Gibbs Sampling for Facets

Language Model
Words (Observed) at (Observed) Timepoints

Kalman Filter for LM evolution

Metropolis Hastings for Exp. evolution

Facets (Latent)

Experience (Latent)

E.g.: The smell of grains a malts on the nose with the slight hop aroma... The taste of the beer is crisp
Sampling based Inference

- Kalman Filter for LM evolution
- Metropolis Hastings for Exp. evolution
- Gibbs Sampling for Facets

Diagram showing a graphical model with nodes and edges representing variables and their dependencies.
## Dataset Statistics

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#Users</th>
<th>#Items</th>
<th>#Ratings</th>
<th>#Time (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beer (BeerAdvocate)</td>
<td>33,387</td>
<td>66,051</td>
<td>1,586,259</td>
<td>16</td>
</tr>
<tr>
<td>Beer (RateBeer)</td>
<td>40,213</td>
<td>110,419</td>
<td>2,924,127</td>
<td>13</td>
</tr>
<tr>
<td>Movies (Amazon)</td>
<td>759,899</td>
<td>267,320</td>
<td>7,911,684</td>
<td>16</td>
</tr>
<tr>
<td>Food (Yelp)</td>
<td>45,981</td>
<td>11,537</td>
<td>229,907</td>
<td>11</td>
</tr>
<tr>
<td>Media (NewsTrust)</td>
<td>6,180</td>
<td>62,108</td>
<td>89,167</td>
<td>9</td>
</tr>
<tr>
<td>TOTAL</td>
<td>885,660</td>
<td>517,435</td>
<td>12,741,144</td>
<td>-</td>
</tr>
</tbody>
</table>
Can we recommend items better, if we consider users’ experience to consume them?
Log-likelihood, Smoothness, and Convergence
### Interpretability: Top Words* by Experienced Users

<table>
<thead>
<tr>
<th></th>
<th>Most Experience</th>
<th>Least Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BeerAdvocate</strong></td>
<td>chestnut_hued near_visous cherry_wood sweet_burning faint_vanilla woody_herbal citrus_hops mouthfeel</td>
<td>originally flavor color poured pleasant bad bitter sweet</td>
</tr>
<tr>
<td><strong>Amazon</strong></td>
<td>aficionados minimalist underwritten theatrically unbridled seamless retrospect overdramatic</td>
<td>viewer entertainment battle actress tells emotional supporting</td>
</tr>
<tr>
<td><strong>Yelp</strong></td>
<td>smoked marinated savory signature contemporary selections delicate texture</td>
<td>mexican chicken salad love better eat atmosphere sandwich</td>
</tr>
<tr>
<td><strong>NewsTrust</strong></td>
<td>health actions cuts medicare oil climate spending unemployment</td>
<td>bad god religion iraq responsibility questions clear powerful</td>
</tr>
</tbody>
</table>

*Learned by our generative model without supervision*
### Interpretability: Top Words* by Experienced Users

<table>
<thead>
<tr>
<th>Platform</th>
<th>Most Experience</th>
<th>Least Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>BeerAdvocate</td>
<td>chestnut_hued near_viscous dark sweet_burning faint_vanilla woody_herbal citrus_hops mouthfeel</td>
<td>pleasant</td>
</tr>
<tr>
<td>Amazon</td>
<td>aficionados minimalist underwritten theatrically unbridled seamless retrospect</td>
<td>overdramatic</td>
</tr>
<tr>
<td>Yelp</td>
<td>smoked marinated savory signature contemporary selections delicate texture</td>
<td></td>
</tr>
<tr>
<td>NewsTrust</td>
<td>health actions cuts medicare oil change spending unemployment</td>
<td></td>
</tr>
</tbody>
</table>

*Learned by our generative model without supervision
Take-away

- Insights from Geometric Brownian Motion trajectory of users:
  - Experienced users *mature faster* than amateurs
  - Progression depends more on *time* spent in community than on activity
- Users' experience evolve *continuously*, along with language usage
- Recommendation models can be improved by considering users’ maturity
- Learns from *only* the information of *users reviewing products at explicit timepoints* --- no meta-data, community-specific / platform dependent features --- easy to *generalize* across different communities
Outline

- Motivation
- Prior Work and its Limitations
- Credibility Analysis
  - Framework for Online Communities
  - Temporal Evolution of Online Communities
  - Credibility Analysis of Product Reviews
- Conclusions
Can we use this framework to find helpful product reviews?

Distributional Hypotheses

- Users with similar facet preferences and expertise are likely to be equally helpful.

- Reviews (e.g., camera) with similar facet-sentiment distribution (e.g., bashing “zoom” and “resolution”) are likely to be equally helpful.

---

Subhabrata Mukherjee, Kashyap Popat, Gerhard Weikum: SDM 2017
Consistency Analysis of Product Reviews

We analyze consistency of embeddings from previous models to detect fake / anomalous reviews with discrepancies like:

1. Rating and review description (promotion/demotion)
   Excellent product... technical support is almost non-existent ... this is unacceptable. [4]

2. Rating and Facet description (irrelevant)
   DO NOT BUY THIS. I can’t file because Turbo Tax doesn’t have software updates from the IRS “because of Hurricane Katrina”. [1]

3. Temporal bursts (group spamming)
   Dan’s apartment was beautiful, a great location. (3/14/2012)[5]
   I highly recommend working with Dan and... (3/14/2012) [5]
   Dan is super friendly, confident... (3/14/2012) [4]
Future Work

★ Going beyond topics and bag-of-words features / lexicons
Learning linguistic cues from embeddings

★ Incorporating richer facets like multi-modal interactions, stance, influence evolution etc.

★ Applications to tasks like Anomaly Detection, Community Question-Answering, Knowledge-base Curation etc.
Conclusions

1. How can we \textit{jointly} leverage users, network, and context for credibility analysis in online communities?
2. How can we model users’ \textit{evolution}?
3. How can we deal with \textit{limited} data?
4. How can we generate \textit{interpretable} explanations for credibility verdict?

Interactional Framework for Credibility Analysis
Collaborators and Co-authors
Acknowledgments (3/3)

Databases and Information Systems Department at Max Planck Institute
1. How can we jointly leverage users, network, and context for credibility analysis in online communities?
2. How can we model users’ evolution?
3. How can we deal with limited data?
4. How can we generate interpretable explanations for credibility verdict?

**THANKS!**

**Interactional Framework for Credibility Analysis**