Modeling and Evaluating Credibility of Web Applications

Joint WICOW/AIRWeb Workshop on Web Quality (WebQuality 2011)
In conjunction with the 20th International World Wide Web Conference in Hyderabad, India.
March 28, 2011

Sara Guimarães (sara@dcc.ufmg.br)
Adriano C. M. Pereira (adriano@decom.cefetmg.br)
Arlei Silva (arlei@dcc.ufmg.br)
Wagner Meira Jr. (meira@dcc.ufmg.br)
AGENDA

- Some brief words…
- Introduction
- Framework Definition – Credibility Rank;
- Case Study;
- Conclusion and Ongoing work.
Who am I?

- Professor in Computer Engineering;
- Member of Brazilian National Institute of Science and Technology for the Web (INWeb);
- Brazil:
  - The fifth largest country by geographical area;
  - The fifth most populous country in the world.
  - The world's 8th largest economy;
  - Has 26 states and a Federal District.
- Minas Gerais:
  - SouthEast;
  - Close to São Paulo and Rio de Janeiro.
- WWW’2013: will be in Rio de Janeiro, Brazil.
• From the Latin *credibilitate*, credibility means the “quality of what is credible or believable”.

• Thus, we can state that credibility is strongly related to the reliability of an assessment, trust, and also with the knowledge that one has to make value judgments.

• Our life is made of choices...
  • Credibility can be defined as believability.
    • Credible people are believable people;
    • credible information is believable information;
    • In fact, some languages use the same word for these two English terms.
Introduction - Credibility

• Choices:
  • Buy A or B?
  • Option C or D?
  • Can I trust on Person X?

• Credibility is an extremely important concept in everyday life!
Introduction - Credibility

- A scale:
  - The credibility of “something” can be mapped to a scale, like a ranking showing how you can believe (trust) in this “something”.
• **Main motivation:**

Need to acquire information to enforce the credibility on the use of Web applications!
• So... The task of evaluating and quantifying Credibility:
  • Major challenge of this research (number of variables, reliability of the information available, computational challenges);
  • How to do it?
• A new framework for the design and evaluate of credibility models;

**Definition** A credibility model $\mathcal{M}$ is a function that receives a set of services $S = s_1, s_2, \ldots s_n$, where $n$ is the number of services and $s_i$ is a tuple of attributes of the service $i$, and returns a ranking $R$, where services are positioned in terms of credibility based on their attributes. A ranking can be described by a function $R : i \rightarrow \mathbb{N}$, where $R(i)$ is the position of the service $i$ and $0 \leq R(s_j) < n, \forall s_j \in S$. The higher is the credibility of a given service $i$ according to $\mathcal{M}$ the lower is the value of $R(i)$. 
A new framework for the design and evaluate of credibility models;

C++ modules;

The proposal is to provide a tool to model and evaluate different credibility models;
Framework Definition – Credibility Rank

- Data Repository
- Pre-process
- Credibility Criteria
  - Service 1, Supplier x, C11, ..., C1n
  - Service 2, Supplier x, C21, ..., C2n
  - ...
  - Service n, Supplier z, Cn1, ..., Cnn
- Credibility Model
- Applying Model
- Ranking
  - 1. Service x
  - 2. Service 49
  - ...
  - n. Service 37
- Feedbacks
  - Service 1, feedback
  - Service 2, feedback
  - ...
  - Service n, feedback
- Evaluation
- Quality of the Model
Case Study – Dataset Description

• About the dataset:
  • An e-market data from the Largest Latin American ISP and content provider (UOL);
  • Sample of some tens of thousand of transactions;

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#categories (top-level)</td>
<td>32</td>
</tr>
<tr>
<td>#sub-categories</td>
<td>2,189</td>
</tr>
<tr>
<td>Average listings per seller</td>
<td>42.48</td>
</tr>
<tr>
<td>Negotiation options</td>
<td>Fixed Price and Auction</td>
</tr>
</tbody>
</table>
Case Study – Dataset Description
Case Study – Dataset Description

• Characterization of several attributes that could be Credibility criteria:
  • Price;
  • Views;
  • Percentage of Positive Qualifications;
  • Global Score;
  • Average Negotiated value;
  • Etc.
Case Study – Methodology

• Each attribute can be used to define a function \( \rightarrow \) Credibility Model;

• Simple Strategy:
  • Combine the attributes;
  • Choosing best \( k \) models; and
  • Continue process until:
    • \( N \) iterations;
    • Minimum gain at each step.

• Exponential possibilities, but fast.
Case Study – Methodology

- Evaluation:
  - Compare to baselines:
    - Global Score;
    - % of Positive Feedback;
    - Combination of them.
  - Compare with SVM-Rank.
Case Study – Methodology

- Quality of model:
  - Probability of receiving negative feedback (focus on some ranking segments);
  - Graph inclination and Area Under the Curve (AUC);
  - Credibility Indicator (CI) = 1/AUC.

- Ranking:
  - Top and bottom: most important.

- Apply CredibilityRank to evaluate the dataset.
Case Study – Experiments / Results

- Credibility Models: **Top** of the Rank

![Graph showing the comparison of different credibility models](image)
Case Study – Experiments / Results

- Credibility Models: **Bottom** of the Rank
<table>
<thead>
<tr>
<th>Credibility Models</th>
<th>Credibility Indicator Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type</td>
</tr>
<tr>
<td>Top 3 TOP</td>
<td>Top_1</td>
</tr>
<tr>
<td></td>
<td>Top_2</td>
</tr>
<tr>
<td></td>
<td>Top_3</td>
</tr>
<tr>
<td>Top 3 BOTTOM</td>
<td>Bottom_1</td>
</tr>
<tr>
<td></td>
<td>Bottom_2</td>
</tr>
<tr>
<td></td>
<td>Bottom_3</td>
</tr>
<tr>
<td>Baselines</td>
<td>BaseLine_1</td>
</tr>
<tr>
<td></td>
<td>BaseLine_2</td>
</tr>
<tr>
<td></td>
<td>BaseLine_3</td>
</tr>
<tr>
<td>SVM-Rank</td>
<td>SVM_1</td>
</tr>
<tr>
<td></td>
<td>SVM_2</td>
</tr>
<tr>
<td></td>
<td>SVM_3</td>
</tr>
</tbody>
</table>
Conclusion

- Model and evaluate some credibility models (functions) for e-Business (e-market dataset);
- Apply a framework (CredibilityRank) to this actual dataset;
- Compare results with baselines and with a SVM-Rank algorithm;
- Top of the rank (most “credible” services / users) and bottom of the rank;
- Consider Probability of Negative Feedback as a quality indicator.
Conclusion

- The results:
  - Top of the rank:
    - 116.8% better than baseline;
    - 36.4% over the SVM-Rank;
  - Bottom of the rank:
    - 24.6% over the baseline;
    - 37.8% better than SVM-Rank;
- Promising results, but much more to improve;
- A good model: not necessarily need many combined attributes;
Conclusion

- Ongoing work:
  - Improve the evaluation / analysis of credibility models (metrics);
  - New credibility models based on machine learning and genetic algorithms;
  - Fraud detection project (e-market / e-payment systems);

- Acknowledgements:
  - INWeb and Brazilian Gov. Agencies;
  - UOL Inc.
Thank you!
Questions? Any suggestions?

Joint WICOW/AIRWeb Workshop on Web Quality (WebQuality 2011)
In conjunction with the 20th International World Wide Web Conference in Hyderabad, India.
March 28, 2011

Adriano C. M. Pereira
Federal Center of Technological Education of Minas Gerais.
e-mail: adriano@decom.cefetmg.br

www.inweb.org.br