

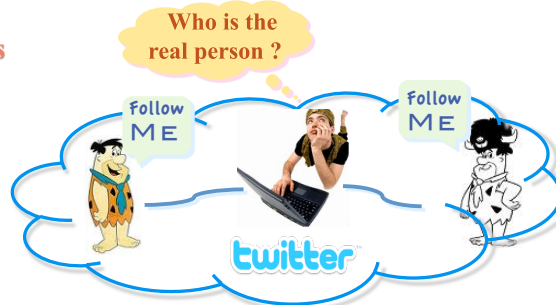
A User Authentication Service for Twitter

Bing Wei, Alessandra Sala, Ben Y. Zhao@UCSB

Is the online user the real person?

Problems with fake identities

- Send spam
- Broadcast false information
- Legal consequence
- Gain revenue by cheating
- Make advertisement



Easy to create identities

- In large numbers
- With low cost

Difficult to validate users

- Similar profiles/behaviors
- Lack authorities

Solution

Leverage Twitter network characteristics to verify a user's identity

Intuition

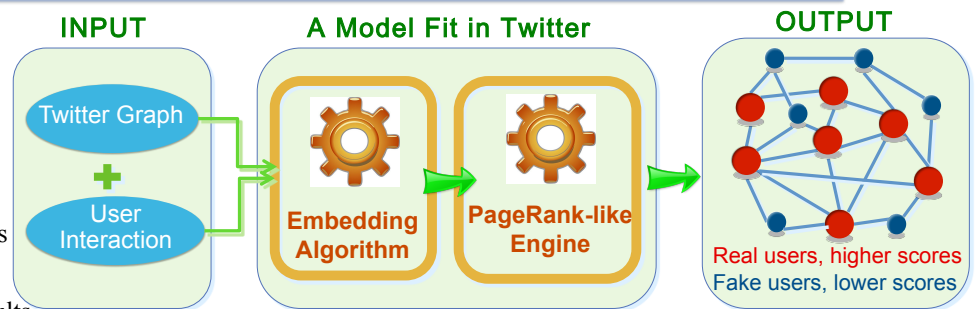
- Trust information propagate along connected links

Model

- Map semantic information into friendship graph
- Compute trust scores to identify users

Usage

- Integrate trust scores with search results

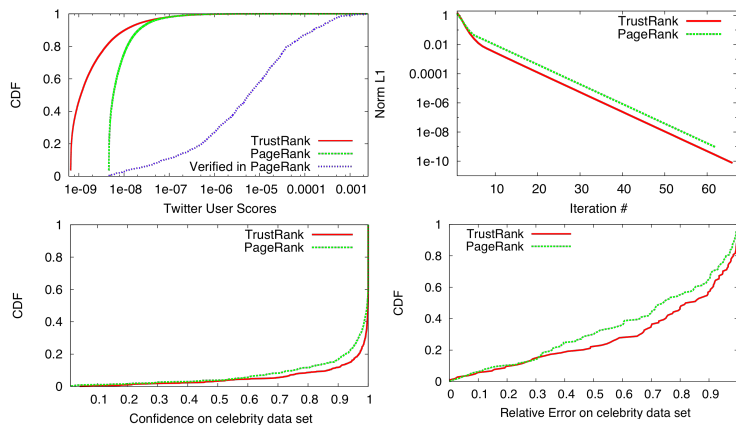


Preliminary Results

Computation

- PageRank on friendship graph of 41M users
- TrustRank on the same graph
- Trust originate from Twitter verified accounts

Data Set	Total #	PageRank Correct%	TrustRank Correct%
Twitter Verified Accounts	1348	90.50%	99.11%
Celebrity Data Set	914	85.89%	88.62%



Evaluation

- Two data sets of ground truth
 - Twitter officially verified accounts
 - Celebrity real accounts from truthtweet.com & twitter-athletes.com
- **Metrics**
 - **Correct:** Given a person's name, score of the real account is higher than those of fake accounts
 - **Confidence:** Score difference between the real account and the highest fake account

$$confidence = \frac{Score(real) - Score(highest_fake)}{Score(real)}$$

- **Incorrect:** Given a person's name, score of the real account is NOT the highest

- **Relative Error:** Score difference in incorrect cases

$$relative_error = \frac{Score(highest_fake) - Score(real)}{Score(highest_fake)}$$

Ongoing Work

- Analyze user interaction
- Generate meaningful weighted graph of interaction
- Adjust PageRank-like engine to prevent manipulation