Credibility

13 December 2024 22:02

User property — confidence that others have on user based on past behavior.

↓

Leads to formation of trust

↙ ↘

CREDIBILITY = node property

TRUST = edge property. → leads to comm.

↓

Users w/ high credibility are less likely to spread disinfo.

↙ Explicit (user rating)

↘ Derived

↙ Local — tweet : retweet

↘ Global — followers + friends

↓

On Twitter:

↙ **Profile cred.**
- user account info
  - age
  - bio length
  - verification
  - geospatial info

↓ **Social Cred**
- no. of followers
- no. of friends

↘ **Behavioral Cred.**
- user's tweet : retweet ratio
- favorites count
- avg. time to retweet.

Myer Briggs :
- Introvert / extrovert
- Sensing / intuition
- Thinking / feeling
- Judging / perceiving

↙ Disinfo dude
→ extrovert
→ feeling

↘ True news :
→ Thinking
→ Judging

↓

Twitter + Hurricane Sandy : 0.3% of users = 90% of fake image retweet.

↳ How they characterized:

Algorithm 1 Compute_Overlap

```
Create_Graph_Retweets()
Create_Graph_Followers()
for each edge in the retweet network do
    num_retweet_edges++
    Insert edge into hashmap, H[1..n]
end for
for each edge in the follower network do
    Insert each edge in hashmap, H[1..n]
    if collision then
        intersections++
    end if
end for
%overlap = (intersections/num_retweet_edges) * 100
```

CredRank — detect coordinated behavior

↳ lower credibility to users who are involved in coordinated behavior

↓

Algo helps find users w/ many social media accounts for misinfo.

Coordinated behavior — herd mentality → low credibility

Algorithm 1 CredRank Algorithm

1: Measure the pairwise similarity between users based on their behavior ($Sim(u_i, u_j)$).
2: Cluster users together if their similarity exceeds the threshold τ.
3: Assign $\omega_{C_i} = \frac{\sqrt{|C_i|}}{\sum_j \sqrt{|C_j|}}$ to each cluster, which is the cluster's weight. Each member in the cluster C_i, has a weight of $\frac{\sqrt{|C_i|}}{|C_i|}$ which is the credibility assigned to the member.

$\sigma(B_i, B_j) = 1$ shows that two users' behaviors are completely similar and $\sigma(B_i, B_j) = 0$ shows that their behaviors are different.