Whom to Trust in a Signed Network?

Matúš Medo (with Berno Büchel and Fanyuan Meng)

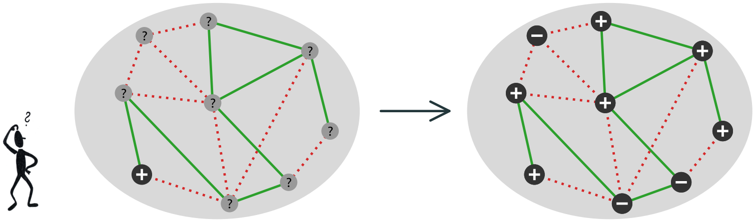
Inselspital, Bern

University of Fribourg, Fribourg

Networks 2021

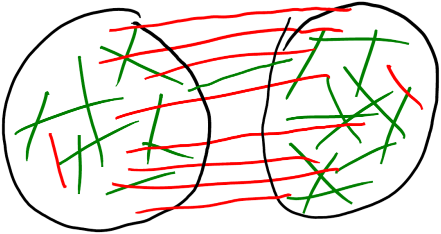
5 July 2021

Problem: An observer who learns from link signs



- Nodes of two types: $\theta_i \in \{-1, +1\}$
- Link signs correlate with node types:
  - $\theta_i = \theta_j$: link is positive with probability $r \geq 0.5$
  - $\theta_i \neq \theta_j$: link is negative with probability $r \geq 0.5$
- Source nodes: The observer knows their types
  - Probability that signal is correct: $q \geq 0.5$

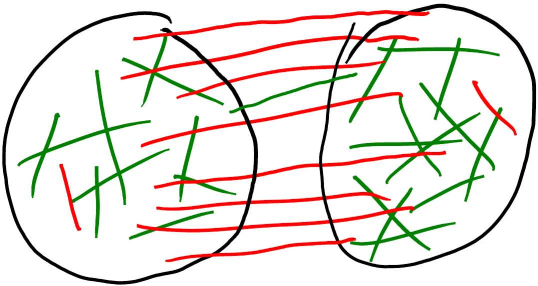
Problem: An observer who learns from link signs



- Nodes of two types: $\theta_i \in \{-1, +1\}$
- Link signs correlate with node types:
  - $\theta_i = \theta_j$: link is positive with probability $r \geq 0.5$
  - $\theta_i \neq \theta_j$: link is negative with probability $r \geq 0.5$
- Source nodes: The observer knows their types
  - Probability that signal is correct: $q \geq 0.5$

Possible applications of the model

1. Two opposing camps: mainstream media and misinformation sources
    - You initially trust in some mainstream media
    - Do you end up trusting other mainstream media and distrusting misinformation sources?



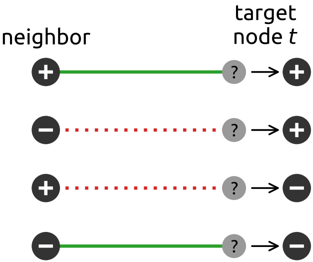
Possible applications of the model

1. Two opposing camps: mainstream media and misinformation sources
    - You initially trust in some mainstream media
    - Do you end up trusting other mainstream media and distrusting misinformation sources?
2. Employee network: Manager attempts to assess hidden qualities of a target employee
3. Inter-firm network: Which other firms to trust
4. Social networks: E.g., find a suitable roommate
5. ...

Random neighbor heuristic (Medo et al, 2021)

Local & easy to apply
("average Joe")

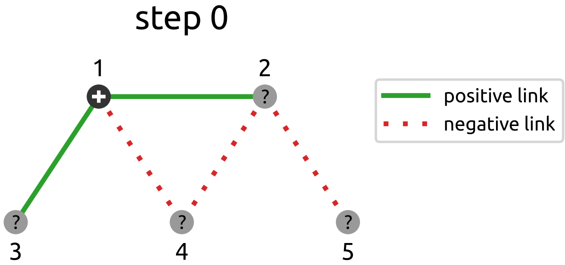
1. Choose target node t at random
2. Opinion on t is made using its random neighbor:



Random neighbor heuristic (Medo et al, 2021)

Local & easy to apply
("average Joe")

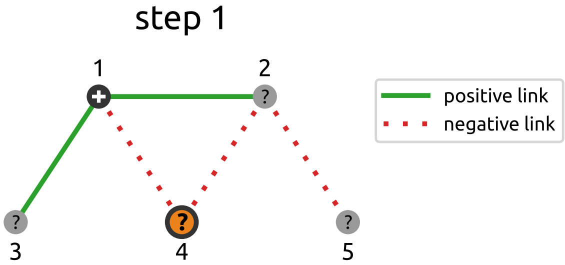
1. Choose target node t at random
2. Opinion on t is made using its random neighbor:



Random neighbor heuristic (Medo et al, 2021)

Local & easy to apply
("average Joe")

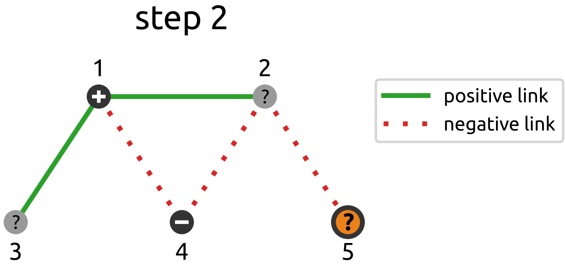
1. Choose target node t at random
2. Opinion on t is made using its random neighbor:



Random neighbor heuristic (Medo et al, 2021)

Local & easy to apply
("average Joe")

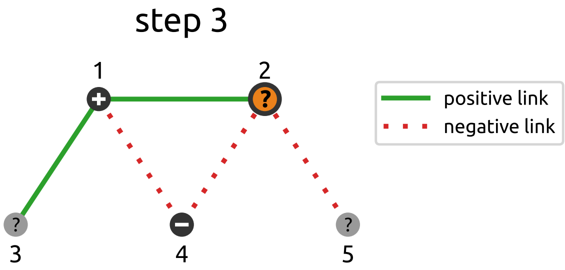
1. Choose target node t at random
2. Opinion on t is made using its random neighbor:



Random neighbor heuristic (Medo et al, 2021)

Local & easy to apply
("average Joe")

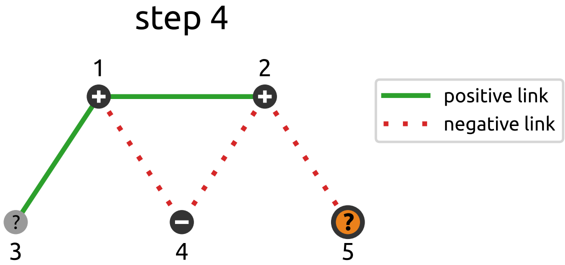
1. Choose target node t at random
2. Opinion on t is made using its random neighbor:



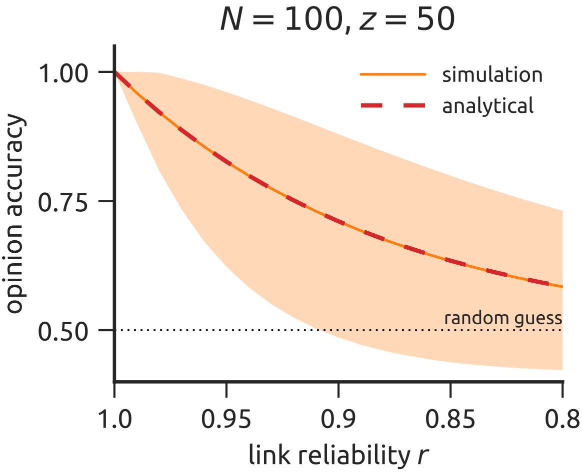
Random neighbor heuristic (Medo et al, 2021)

Local & easy to apply
("average Joe")

1. Choose target node t at random
2. Opinion on t is made using its random neighbor:



Random neighbor heuristic: The outcome



Shaded area: 10th–90th percentile range

Random neighbor heuristic: The outcome

- Even few misleading links ($r < 1$) cause low expected accuracy and high accuracy variability
- As N grows, expected accuracy approaches 0.5 as

$$E(A) - 0.5 \sim N^{1-r}$$

Random neighbor heuristic: The outcome

- Even few misleading links ($r < 1$) cause low expected accuracy and high accuracy variability
- As N grows, expected accuracy approaches 0.5 as

$$E(A) - 0.5 \sim N^{1-r}$$

To make sense
of a complex world
is difficult

See *M. Medo, M. S. Mariani, L. Lü, Communications Physics 4, 1, 2021* for more

Two new solutions (Meng et al, 2021)

1. Bayesian solution:
    - The probability of a vector of node types, $\boldsymbol{\theta}$, when source node signals are $\boldsymbol{\sigma}$ and observed link signs, $\boldsymbol{R}$: $P[\boldsymbol{\theta}|\boldsymbol{\sigma}, \boldsymbol{R}]$
    - The probability that given target node t is of positive type is

$$P(\theta_t = +1) = \sum_{\boldsymbol{\theta}:\theta_t=+1} P[\boldsymbol{\theta}|\boldsymbol{\sigma}, \boldsymbol{R}]$$

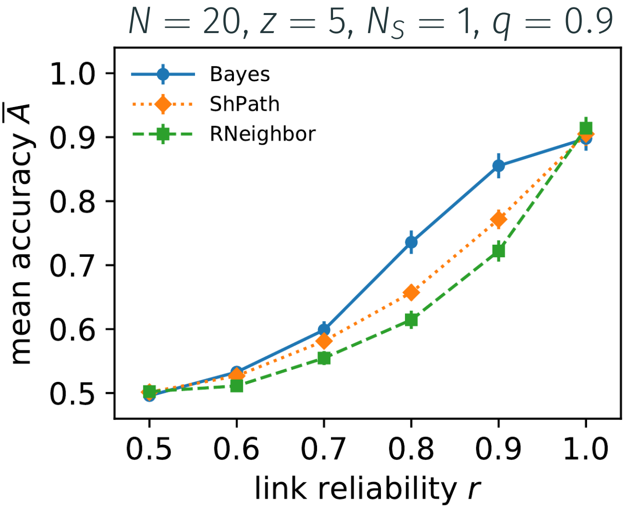
Two new solutions (Meng et al, 2021)

1. Bayesian solution:
   - The probability of a vector of node types, $\boldsymbol{\theta}$, when source node signals are $\boldsymbol{\sigma}$ and observed link signs, $\boldsymbol{R}$: $P[\boldsymbol{\theta}|\boldsymbol{\sigma}, R]$
   - The probability that given target node t is of positive type is

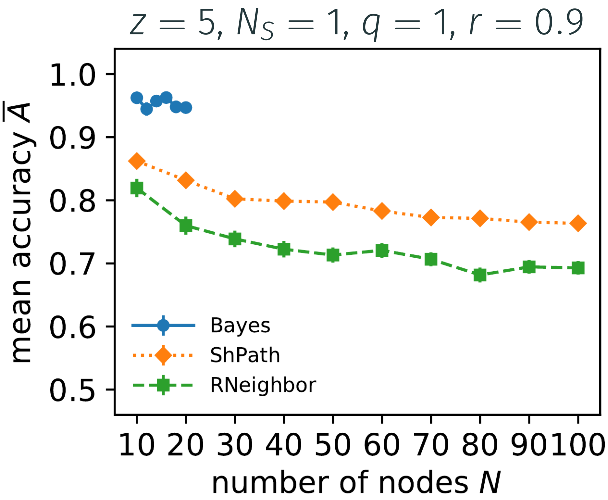
$$P(\theta_t = +1) = \sum_{\boldsymbol{\theta}:\theta_t=+1} P[\boldsymbol{\theta}|\boldsymbol{\sigma}, R]$$

2. Shortest-path heuristic:
   - For all source nodes, $s \in S$, find the shortest path $s \to t$
   - Compute the probability $P(\theta_t = +1|s \to t)$ for each path
   - Aggregate information from all paths as if they were not overlapping

Simulation results



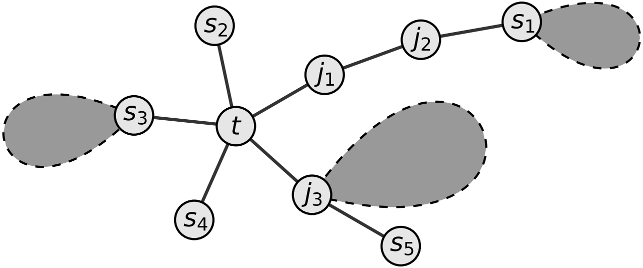
Simulation results



Theoretical results

Theorem (Equivalence)

If path from s to t is unique for all $s \in S$ and the paths from distinct source nodes do not overlap, the Bayesian rule and the shortest path rule are equivalent.



Theoretical results

Theorem (Ordering)

For a given network, set of source nodes S and target node t, the expected accuracies of the three rules are be ordered as

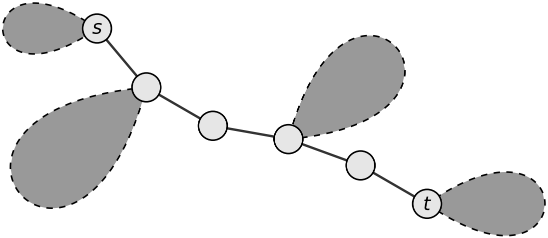
$$E[A^{Bayes}] \geq E[A^{ShPath}] \geq E[A^{RNeighbor}].$$

Theoretical results

Theorem (Unique Path)

If there is only one source and the path from s to t is unique, then all three rules yield the same expected accuracy

$$E[A^{Bayes}] = E[A^{ShPath}] = E[A^{RNeighbor}].$$



Theoretical results

Theorem (Shortest Path Accuracy)

When the source and target node are chosen at random in an Erdös-Rényi network, the expected accuracy of the shortest path decays with the number of nodes, N, as $E[A^{ShPath}] - 0.5 \sim N^{-\gamma_{ShPath}}$ *where*

$$\gamma_{ShPath} = -\ln(2r-1)/\ln z$$

and z is the mean degree.

Recall: For the random neighbor rule, the scaling exponent is

$$\gamma_{RNeighbor} = 1 - r.$$

Summary

- Opinion/trust formation on a signed network
- Different from other opinion formation models (voter model, DeGroot,...)
  - One agent (observer), N subjects

Summary

- Opinion/trust formation on a signed network
- Different from other opinion formation models (voter model, DeGroot,...)
  - One agent (observer), N subjects
- For simple formation mechanisms, the results are sensitive to noise in the system
- More sophisticated mechanisms yield better results at higher computational costs
- See more here:

1. M. Medo, M. S. Mariani, L. Lü, The fragility of opinion formation in a complex world, Communications Physics 4, 1 (2021)
2. F. Meng, M. Medo, B. Buechel, Whom to Trust in a Signed Network? Optimal Solution and two Heuristic Rules, preprint (2021)

Further questions

1. Consider more than two node types
2. Correlate link presence with node types
   (here only link signs depend on types)

---

3. Which rules yield accurate opinions without being excessively demanding?
4. Which spurious links distort the results most?
5. Combination with social opinion-formation models
6. How to empirically study belief formation on signed network data?
7. ...

Thank you for your attention!

`http://www.ddp.fmph.uniba.sk/~medo/physics/`
matus.medo@unifr.ch