Role Detection in Communication Networks

About specialising protestants and
deceived & deceiving employees

Ramona Roller
10th January 2023

Formal roles guide human behaviour

Enron

Kenneth Lay (**chairman**); Jeffrey Skilling (**CEO**)

Reformation

16th century **scholar**: Johannes Reuchlin

Image sources:

[1] chron.com: Enron's cast of characters

[2] pforzheim.de: Johannes Reuchlin: Der Humanist aus Pforzheim

But what if humans deviate from their formal roles?

Enron: Explain deceit

Successful company

Deceit

Reformation: Explain differentiation

Universal scholar

Theologian

Lawyer
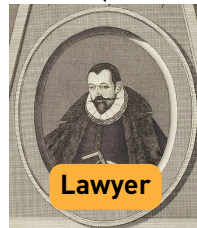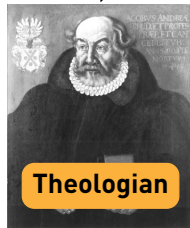
➡ **Which other roles could explain this deviant behaviour?**

Informal roles in communication networks

- ▶ **Informal role**
  - ▶ Emerge through **interactions** and are **context-specific**.

- ▶ **Application**
  - ▶ **Communication network:** nodes = individuals; edges = message exchange
  - ▶ Enron e-mail network;   Reformation letter correspondence network
 - ▶ **Communication role:** Group of nodes with similar topological embeddings

■ Mediation   ■ Reciprocity
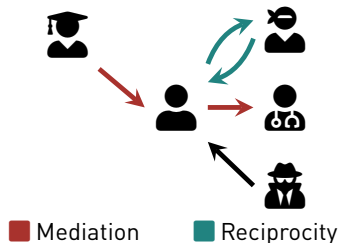
➡ **How can we detect and interpret communication roles?**

➡ **How can we relate communication roles to observed behaviour that deviates from formal roles?**
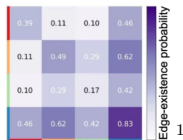
Role detection in networks: common practice (1/2)

Role detection: Procedure for grouping nodes in networks based on topological similarity
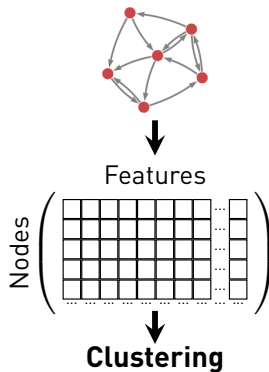
Structural approach
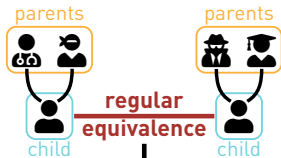
Stochastic Block Model

❯ **Nodes are grouped into blocks (roles) but meaning of blocks is not clear.**

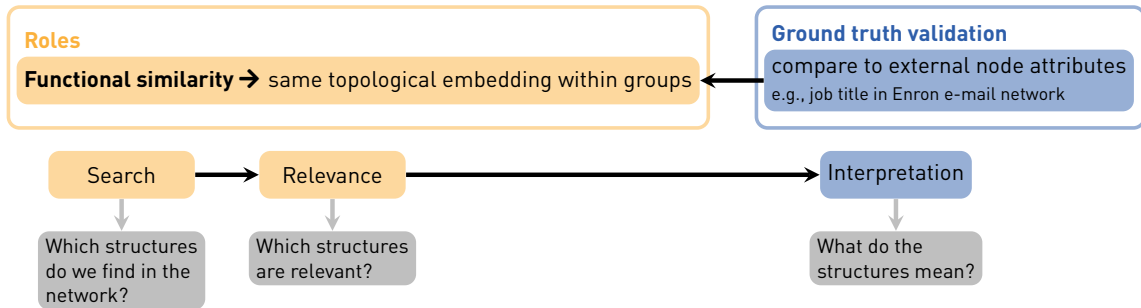❯ **Open question: What do roles mean?**

Feature-based approach

Features

Clustering

❯ **Nodes are grouped into clusters (roles) but meaning of blocks remains technical.**

[1] Image: Faskowitz et al., 2018: Weighted Stochastic Block Models ..., Nature.
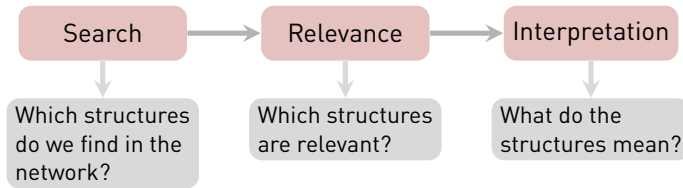
Role detection in networks: common practice (2/2)

Roles

Functional similarity → same topological embedding within groups

Ground truth validation
compare to external node attributes
e.g., job title in Enron e-mail network

Search — Which structures do we find in the network?

Relevance — Which structures are relevant?

Interpretation — What do the structures mean?

▸ **Problem:** Ground truth may not be relevant for detected roles[1].

❯ **Role detection is easy (1), interpretation is not (2).**

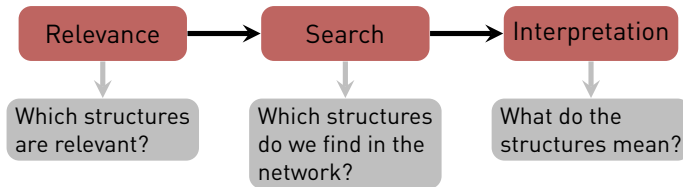[1] Peel et al., 2017: The ground truth about metadata…, Science Advances.

Role detection in networks: New practice

| Search | → | Relevance | → | Interpretation |

Common practice

| Which structures do we find in the network? | Which structures are relevant? | What do the structures mean? |

| Relevance | → | Search | → | Interpretation |

New practice

| Which structures are relevant? | Which structures do we find in the network? | What do the structures mean? |

➔ **Select roles that are relevant in communication networks.**

Role detection in networks: applications

Enron e-mail network

- ▶ Enron: fraudulent commodities company
- ▶ Nodes: 144 employees
- ▶ Edges: 21,701 sent e-mails
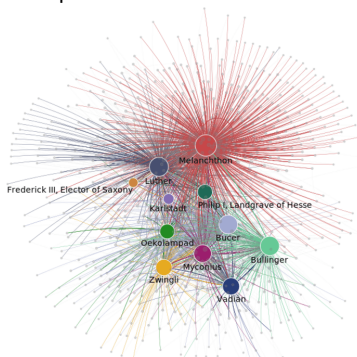- ▶ Period: 1998–2002
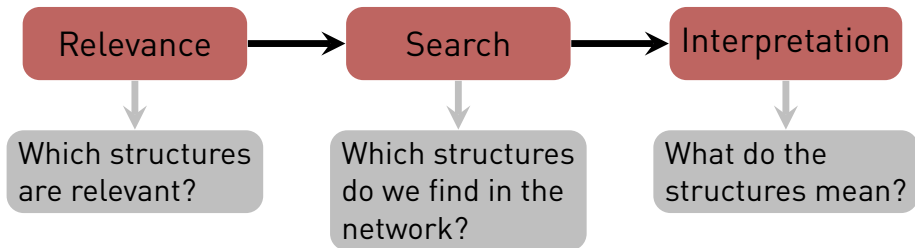- **RQ** Does deceit manifest in communication roles?
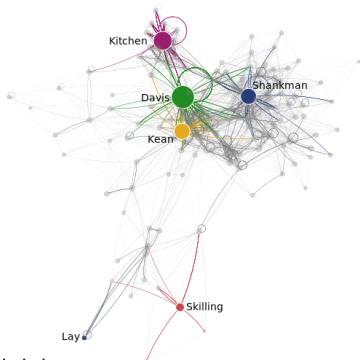
Reformation letter correspondence network

- ▶ Reformation: overthrow of Catholic Church
- ▶ Nodes: 3,000 scholars
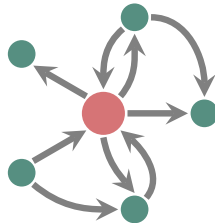- ▶ Edges: 30,000 sent letters
- ▶ Period: 1510–1565
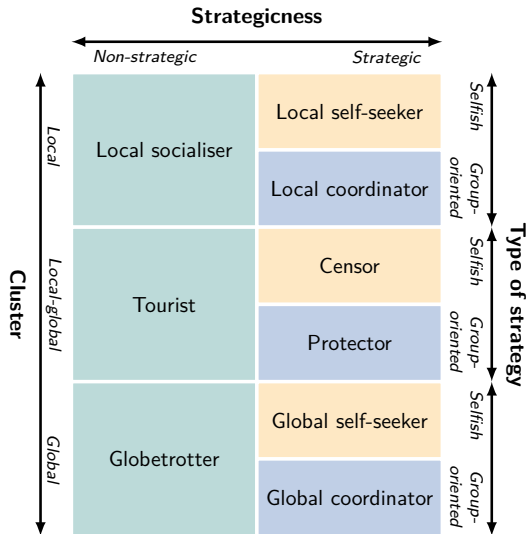- **RQ** Does differentiation manifest in the specialisation of communication roles?

Our pipeline

Relevance	→	Search	→	Interpretation

Which structures are relevant?

Which structures do we find in the network?

What do the structures mean?

Relevance: Which roles are relevant?[1]

Strategicness

	Non-strategic	Strategic	
Local	Local socialiser	Local self-seeker	Selfish
		Local coordinator	Group-oriented
Local-global	Tourist	Censor	Selfish
		Protector	Group-oriented
Global	Globetrotter	Global self-seeker	Selfish
		Global coordinator	Group-oriented

**Cluster** / **Type of strategy**

Global coordinator
Mediates between local and distant parts of the network. Promotes information exchange between other nodes.

➲ **Literature review of communication networks reveals to-be expected roles.**

➲ **Roles are relevant for communication networks and generalisable across them.**

---

[1] Roller and Schweitzer, 2021: A new theory of social roles in networks, `/.../socarxiv/xkafb/`

Enron

# Search: Which roles do we find in the network? (1/2)
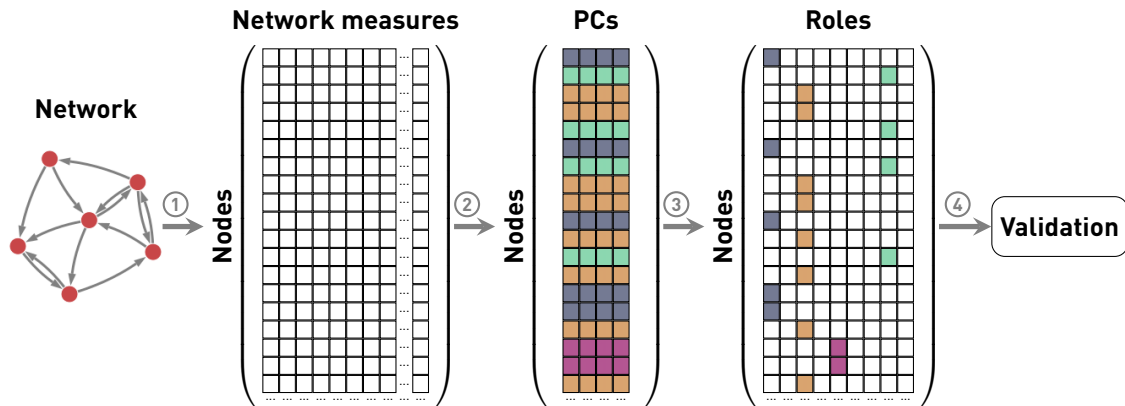
① Topological network measures for all nodes ② Dimensionality reduction (PCA) ③ Clustering

➔ **Roles: clusters of nodes based on non-redundant topological network measures**

- ▶ Result of clustering nodes along topological network measures

- ▶ **9 Columns:** Extracted roles

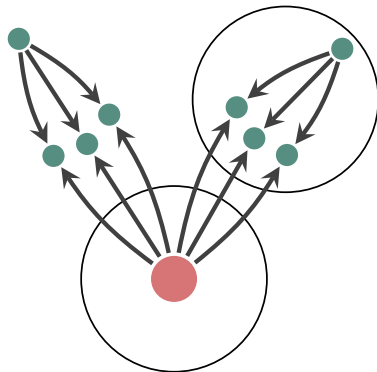- ▶ **7 Rows:** topological measures in lower dimensional space (principal components (PCs))

① ② ③　④　⑤　⑥　⑦　⑧ ⑨
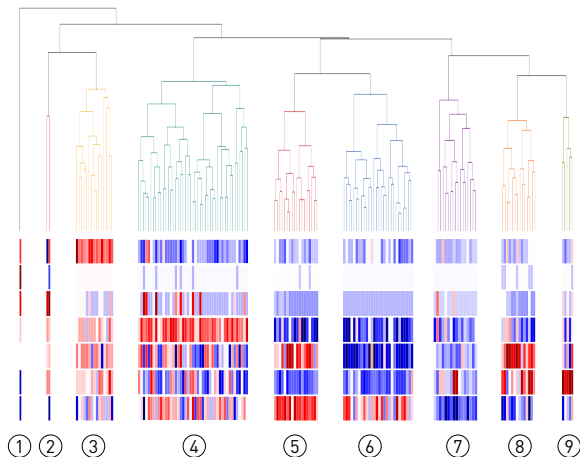
➲ **Nodes of the same role have similar values along the topological measures.**

Interpret topological measures in the lower dimensional space

| Topological network measure	Factor loading
global binary out coreness	0.22
between binary out coreness	0.15
between mean in dice similarity	0.13
global mean out dice similarity[*]	0.83
between mean out dice similarity[*]	0.94
global out neighbours 1st order	0.17

[*] Used for interpreting PC as communication constituent

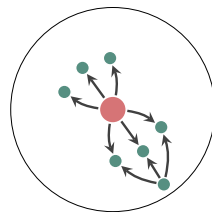➡ **Global action seeking:** Sending e-mails to those who also receive e-mails from others in the global network.

Interpret roles

▶ Per cluster identify pattern across topological measures in lower dimensional space (PC)

▶ Combine pattern with interpretation of PCs to yield a communicatio pattern → **role**
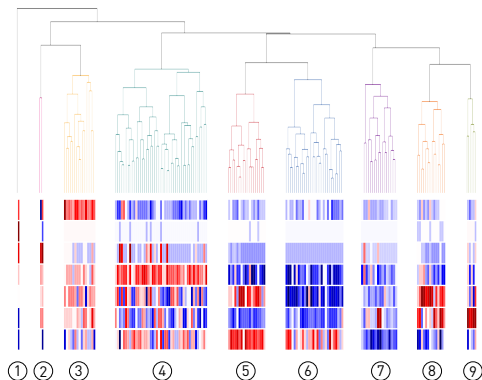
Role 5

▶ **PC 1:** $-/--$
▶ **PC 2:** $-$
▶ **PC 3:** $--$
▶ **PC 4:** $--/---$
▶ **PC 5:** $+/++$
▶ **PC 6:** $--/---$
▶ **PC 7:** $+++$

❱ **Role 5:** Contacts active communicators locally without receiving replies. → **Local socialiser**

Roles of Enron employees

The bosses: Skilling & Lay
▶ Fraudulent behaviour, prison sentence
▶ Skilling: **local coordinator**;   Lay: **censor**
❓ **Too local?**   Roles ➡ lost awareness of leadership responsibilities?

Similar yet different employees: Kitchen & Shankman
▶ Kitchen: bonus + acquitted of charges; **global coordinator**
▶ Shankman: fired + prison sentence; **globetrotter**
❓ **Too unfocused?**   Roles ➡ presence/lack of clear tasks?

Loners: Kaminski
▶ voice of reason + not accused; **protector**
❓ Do roles mirror personal values?

➡ **No clear role pattern for deceitful behaviour: individual interpretations are necessary**

➡ **Missing information on behaviour of (crucial) employees**
e.g., Andrew Fastow (CFO) is not part of the data

[1] Skilling, Lay: chron.com
[2] Kitchen: bbc.com: Enron's internet monster
[3] Shankman: ladailypost.com: Ex-Enron Executive...
[4] Kaminski: https://business.rice.edu/: Personal page

Reformation

Reformation: Role detection pipeline

▶ Are communication roles subject to social differentiation?
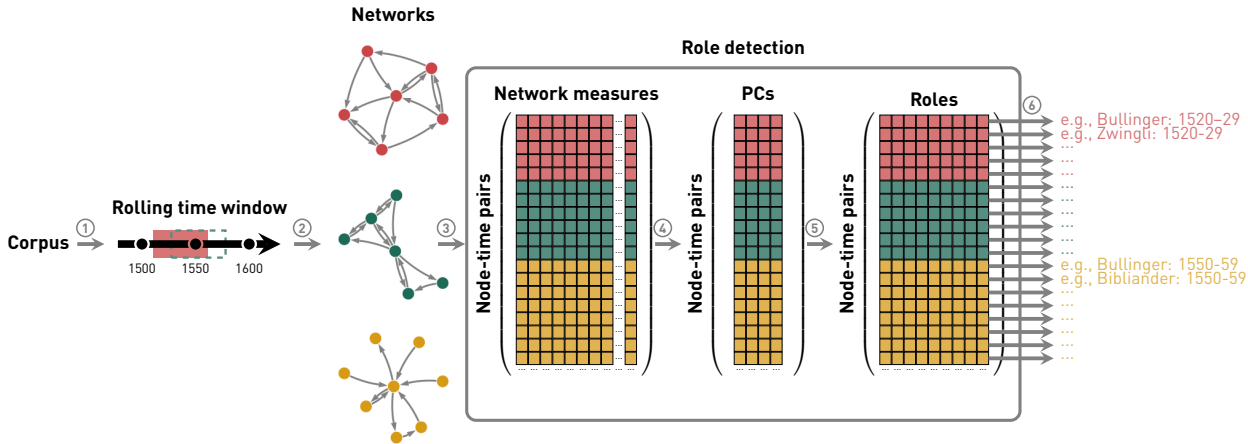
❸ **Result of role detection:** Probability distribution across roles per reformer and subperiod.

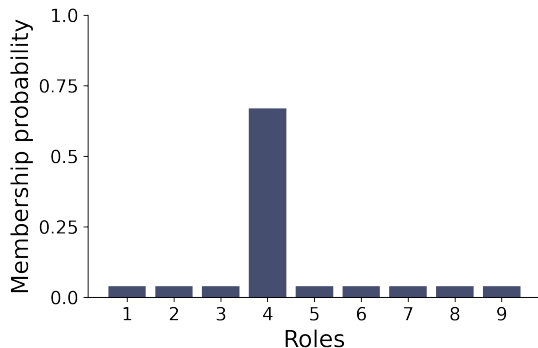❸ Probabilities indicate to what extent reformer fulfills roles.

Role distributions of Heinrich Bullinger

1520–29: Uniform distribution

1550–59: Nonuniform distribution
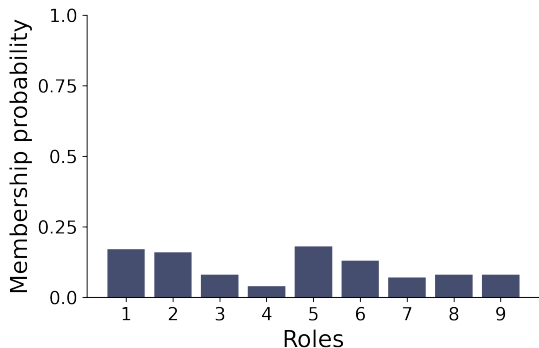
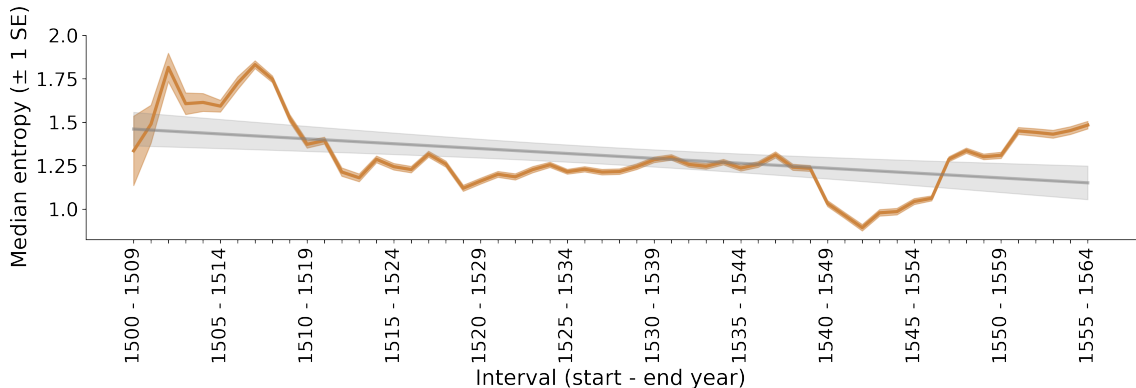- ▶ Bullinger fulfills roles to *same* extent.
- ▶ Large entropy

- ▶ Bullinger fulfills roles to *different* extents.
- ▶ Small entropy

❷ **Entropy distinguishes between shapes of probability distributions.**

Trend of entropy time series

Entropy time series

▶ Per subperiod: average entropies of probability distributions of reformers

▶ Result: one entropy value per subperiod

❖ **Negative trend:** reformers fulfill fewer roles over time ➡ become more specialised in their communication roles

Conclusion

Our approach to role interpretation
- ▶ Roles are relevant for communication networks
- ▶ Roles are generalisable across communication networks
- ▶ **Enron:** local + non-strategic roles for fraudulent actors
- ▶ **Reformation:** individuals specialise in their roles ➡ differentiation
- ▶ **Next step:** Theory building and falsification
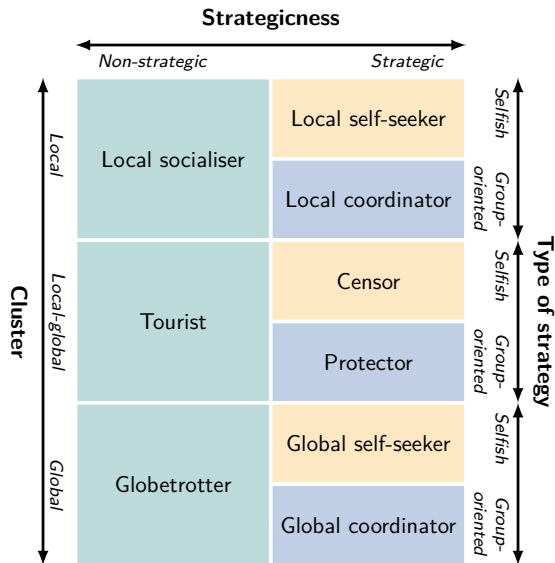
MSc course: Data Science for Social Challenges

Simple inference	Causal inference	Complex inference
▸ OLS regression ▸ Gender pay gap ▸ Logistic regression ▸ Energy crisis ▸ Panel regression ▸ Loneliness	▸ Causal concepts ▸ Assessing the gain ▸ Diff-in-diffs ▸ Poverty ▸ Matching ▸ Political participation	▸ Network null models ▸ Human migration ▸ Measures and motifs ▸ Legalization of cannabis ▸ Network regressions ▸ Information overflow

For further information on the course (**syllabus**, **course catalog**)
visit https://www.sg.ethz.ch/teaching/
data-science-for-social-challenges/