

# Detecting Conspiracies

Robert Hilbrich  
Robert@Hilbri.ch

Department of Computer Science  
Humboldt Universität Berlin

"Just because you're paranoid don't mean they're not after you."  
(Woody Allen)

# Introduction

## What you wont see here

...

15:00 Spy vs. Spy: Sharon Daniel's Public Secrets and  
Robert Hilbrich's Terror Detection Algorithms

...

You will see: the Clash of two Worlds!

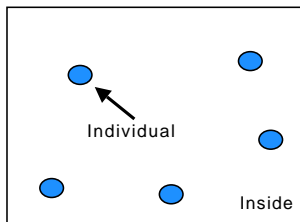
# You will see: the Clash of two Worlds!



## Real World

- Complex
- Love, Hate
- Home of terror

# You will see: the Clash of two Worlds!



Outside

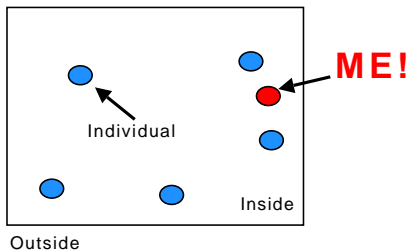
## Real World

- Complex
- Love, Hate
- Home of terror

## Computer Science World

- Simplified
- Digital
- Home of algorithms

# You will see: the Clash of two Worlds!



## Real World

- Complex
- Love, Hate
- Home of terror

## Computer Science World

- Simplified
- Digital
- Home of algorithms

## A first clash ...

**IT'S TIME TO CHECK YOUR NETWORK VALUE**

Your business and personal contacts make up a network, which can exert considerable influence on your professional life and career. If you can recognize the value of your network you can tap into it and use it. So, what's your network value?

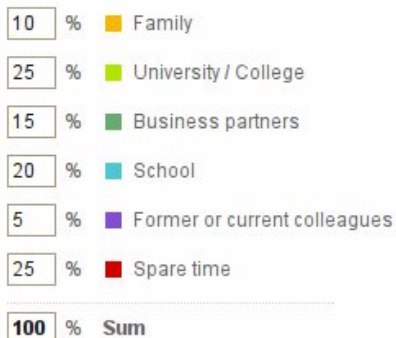
**START**

- > Welcome!
- > Network Value Test
- > Gallery
- > Tell-a-Friend
- > Optimize your network value
- > Network Value Buttons

... taken from <http://mynetworkvalue.com/> (XING AG)

# Questions regarding the Structure

## Where do you know your contacts from?





# Questions regarding the Intensities

## How often do you stay in touch with each of your contacts?

Your network consists of: **15 Persons**

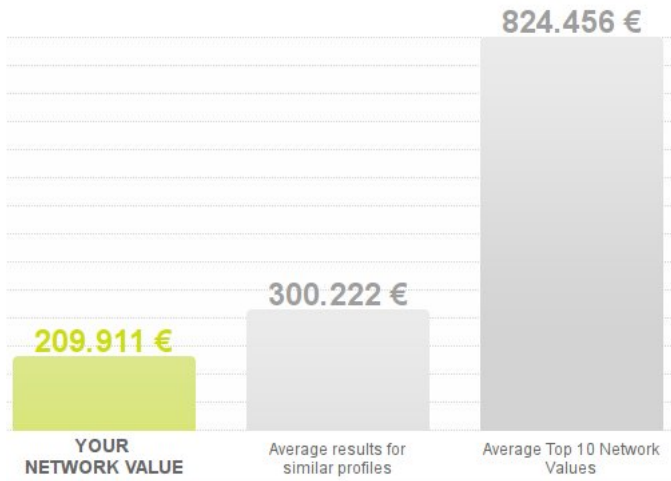


■ **Often** (at least once a week)

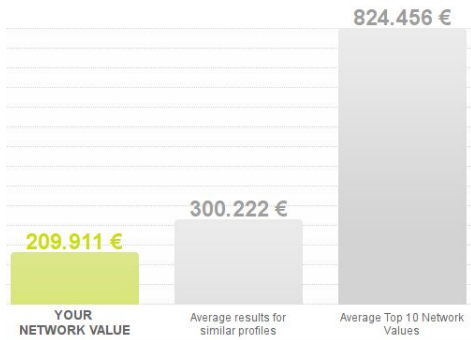
■ **Occasionally** (once a month)

■ **Rarely** (once a year or less)

# Results



# Results



## What is a *Network Value*?

... an estimate of the **potential monetary value of your network**, which you can tap into and capitalize over the next **10 years** ...

# How do they know that?

**Social Network Analysis (SNA)** ... is a **key technique** in:

*modern sociology, anthropology, sociolinguistics, geography, social psychology, communication studies, information science, organizational studies, economics, biology ...*

# How do they know that?

**Social Network Analysis (SNA)** ... is a **key technique** in:

*modern sociology, anthropology, sociolinguistics, geography, social psychology, communication studies, information science, organizational studies, economics, biology ...*

## Common Goal

Analyze the structure of a social network to **infer knowledge** about an individual or a group.

# How do they know that?

**Social Network Analysis (SNA)** ... is a **key technique** in:

*modern sociology, anthropology, sociolinguistics, geography, social psychology, communication studies, information science, organizational studies, economics, biology ...*

## Common Goal

Analyze the structure of a social network to **infer knowledge** about an individual or a group.

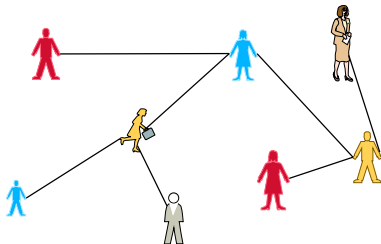
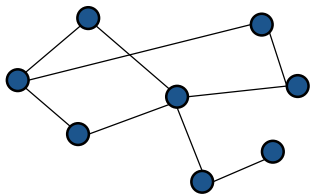
## Goal for Today

Can we use **Social Network Analysis** to detect **conspiracies**?

# Modelling Social Networks

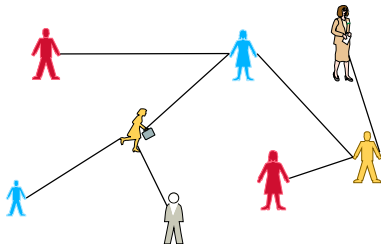
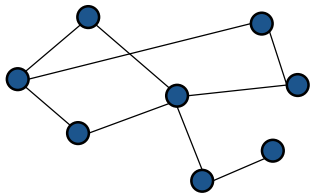
How to represent Social Networks for proper analysis?

# What are Graphs?





# What are Graphs?

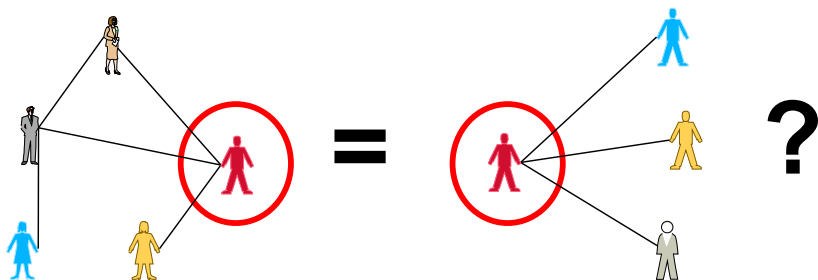


Graphs are ...

**Graphical structures** to model **pairwise relations** between objects.

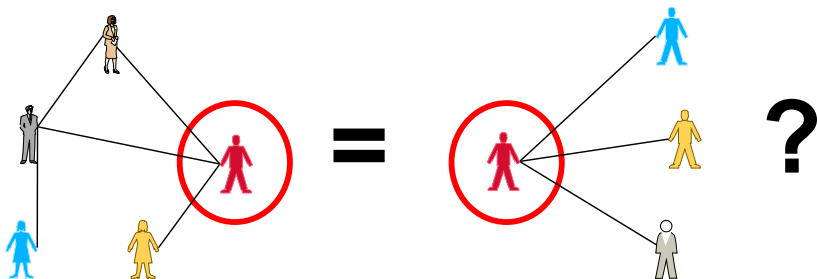
- **nodes** to represent objects ( $\rightarrow$  *individuals*)
- **edges** to express relations ( $\rightarrow$  *communication paths*)

# What are Graphs Metrics?



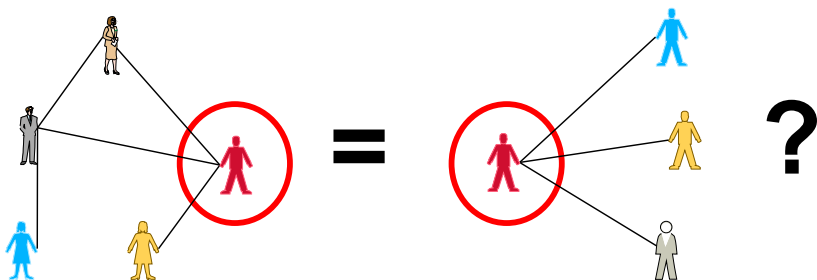
Want to **compare** Nodes with other Nodes!

# What are Graphs Metrics?



Want to **compare** Nodes with other Nodes!  
Need **measures** for comparisons → **Graph Metrics!**

# What are Graphs Metrics?



Want to **compare** Nodes with other Nodes!  
Need **measures** for comparisons → **Graph Metrics!**

## Example

Task: Search for **popular** individuals in a graph

Popularity of a Person (node) corresponds to **# edges**

# Conspiracies

Can we locate conspiring groups in a Social Network?

## Approach taken

Research from *21st Century Technologies*<sup>1</sup>



### Idea

- Given a **graph** containing communication paths
- Find path structures that are **uncommon** for human interaction

---

<sup>1</sup><http://www.21technologies.com/>, DARPA funded

## Approach taken

Research from *21st Century Technologies*<sup>1</sup>



### Idea

- Given a **graph** containing communication paths
- Find path structures that are **uncommon** for human interaction

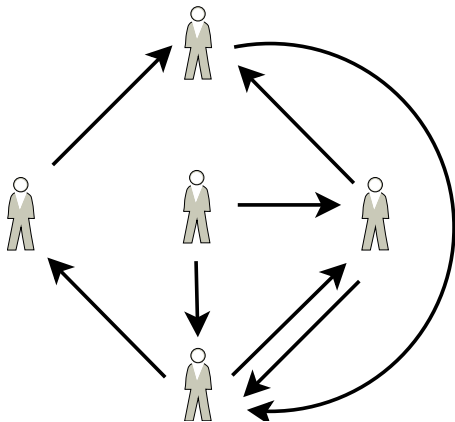
### Assumption

**Regular** human interaction differs from **Conspiracy** interaction

---

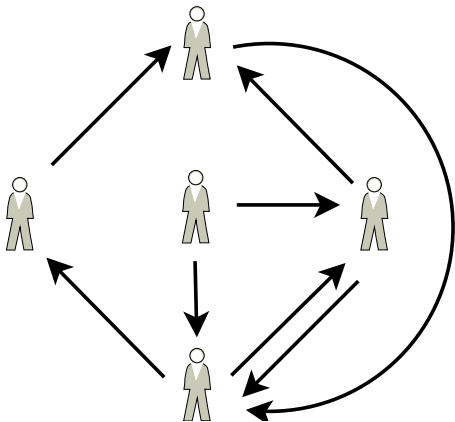
<sup>1</sup><http://www.21technologies.com/>, DARPA funded

# Regular communication structures





## Regular communication structures



### Characteristics

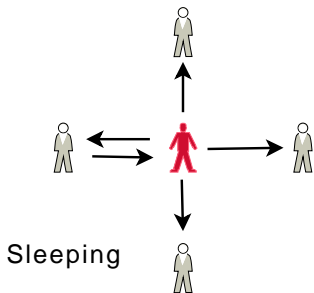
- **non-mediated** structure
- **shortest path** preferred
- **redundant** paths

## *Conspiring* communication structures<sup>2</sup>

---

<sup>2</sup>based on Leninist-cell model

# Conspiring communication structures<sup>2</sup>

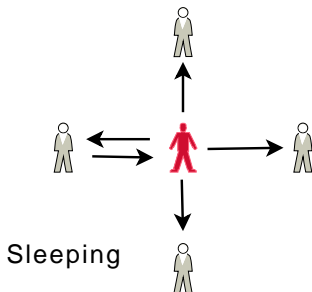


## Characteristics

- Structural Change: **Sleeping**  $\Rightarrow$  **Active**

<sup>2</sup>based on Leninist-cell model

## Conspiring communication structures<sup>2</sup>

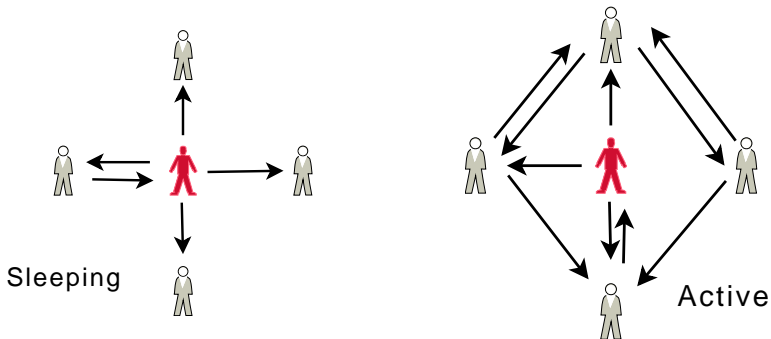


### Characteristics

- Structural Change: **Sleeping**  $\Rightarrow$  **Active**
- Sleeping: Security demands **mediator** and **no redundancies**

<sup>2</sup>based on Leninist-cell model

## Conspiring communication structures<sup>2</sup>



### Characteristics

- Structural Change: **Sleeping**  $\Rightarrow$  **Active**
- Sleeping: Security demands **mediator** and **no redundancies**
- Active: efficiency demands **short paths** and **redundancies**

<sup>2</sup>based on Leninist-cell model

# Experiments & Results

## Approach

- Given a **graph** with communication paths
- Define structural differences with graph **metrics**<sup>a</sup>
- Train **algorithms** to search for these **irregular** communication patterns

---

<sup>a</sup>density, characteristic path length, average cycle length

# Experiments & Results

## Approach

- Given a **graph** with communication paths
- Define structural differences with graph **metrics**<sup>a</sup>
- Train **algorithms** to search for these **irregular** communication patterns

---

<sup>a</sup>density, characteristic path length, average cycle length

## Results

- Decision **terrorist** or **non-terrorist**

# Experiments & Results

## Approach

- Given a **graph** with communication paths
- Define structural differences with graph **metrics**<sup>a</sup>
- Train **algorithms** to search for these **irregular** communication patterns

---

<sup>a</sup>density, characteristic path length, average cycle length

## Results

- Decision **terrorist** or **non-terrorist**
- Accuracy rate of **93%**



# Effects and Evaluation

Should we use SNA? What are the drawbacks?

## What's the point?



- fast, simple, cheap
- only need the **graph** data (FaceBook, Law, ...)
- High **accuracy rate**

# What's the point?



- fast, simple, cheap
- only need the **graph** data (FaceBook, Law, ...)
- High **accuracy rate**



- *Leninist cell* - model
- 7% false classifications (7% of 80'000'000 = 5'600'000)
- Using a graph puts everybody under **general suspicion**
- Found a pattern = a conspiracy? (**false positives**)
- Capturing of **complex phenomena** in simple metrics?
- Effects from observation **noise**?

## However ...

### Government agencies

Results are from 2004, so expect this approach to be used **today**.

## However ...

### Government agencies

Results are from 2004, so expect this approach to be used **today**.

### Business companies

- *AT&T* is funding research in this area to detect **fraud**
- *Amazon* uses a similar approach for **reputation management**  
*"8 from 8 customers found this review helpful"*

## However ...

### Government agencies

Results are from 2004, so expect this approach to be used **today**.

### Business companies

- *AT&T* is funding research in this area to detect **fraud**
- *Amazon* uses a similar approach for **reputation management**  
*"8 from 8 customers found this review helpful"*

### Open questions

- How to exploit SNA for your own good?
- Can we define a **standard** behaviour?
- Effects on cultural **diversity**?

# Questions or Comments

# Bibliography



Thayne Coffman, Seth Greenblatt, and Sherry Marcus.  
Sensitivity of social network analysis metrics to observation noise.

*In Proc. IEEE Aerospace Conf., 2004.*



Thayne Coffman and Sherry Marcus.  
Dynamic classification of groups using social network analysis and hmms.

*In Proc. IEEE Aerospace Conf., 2004.*



Corinna Cortes, Daryl Pregibon, and Chris Volinsky.  
Communities of interest.

*Lecture Notes in Computer Science, 2189:105–??, 2001.*