

"Something, that is allegedly secure is not necessarily secure, Something, that is allegedly known might turn out to be unknown. Appearance can be deceptive, our senses can deceive us. Even though experience and knowledge can limit errors, reality also limits those."

Inspired by Berthold Brecht

Author: unknown

#### Mittweida

Forensic Science Inv. Lab



Ortrand

Radeburg

Kreischa

Glashütte

Altenberg

National Route

Moritzburg

Radebeul

Wilsdruff Dresden

Freital

Dippoldiswalde

Schv

Königsbri

Radebo

Pir

Ottendorf-Okrilla





Thallwitz

Brandis

Naunhof

**Bad Lausick** 

Wurzen

107

Grimma

Colditz

107

Taucha

Borna

72

Lossatal

14

ditz

Leipzig

Markkleeberg

Zwenkau

oitzsch

Lucka

ue



Mühlberg/Elber

Strehla

Riesa

Ostrau Lommatzsch

Nossen

E40

Großschirn

Freiberg

Brand-Erbisdorf

101

Cavertitz

Oschatz

beln

Dahlen

Mügeln

bigM

Wermsdorf

114

Leisnig

Gröditz

Röderaue

Großenhain

Priestewitz

Meißen

Klingenberg

Frauenstein

171

Rechenberg-Bienenmühle.

#### **16 members 10 PHD students** + Master and Bachelor students

Group Mittweida



DataSys 2017 June 25 - 29, 2017 - Venice, Italy



## P(Security) = 1- P(Crime) Is security the absence of crime?

Dirk Labudde and Michael Spranger Sonntag, 9. Juli 2017



#### Introduction





#### Germanwings Flight 9525, March 24, 2015





2017 terrorist

attack London

#### 2016 attack on a Swiss train





















# What does security mean in this context?

Security means protection of humans or things from humans.

# What does crime mean in this context?

Violating the right of physical and psychological integrity and sacredness of property.

# Can we model security by simulating crime?

P(Security) = 1- P(Crime) Is security the absence of crime?

## **Levels of Security**





#### International security interests

- Terror
- Civil war
- Plane hijacking



#### National security interests

- Extremism
- Amok
- Organized crime



#### Regional security interests

Burglary series



Personal security interests

Burglary

## **Aspects of Security**





Police officers are your friends. They want to help you.

 $\mathbf{\Psi}$ 

## Approaches

## **Top-Down - Approach**



## Approaches for security from the system theory Methods in information gathering

#### Top-down (deductive)

- Creating experimental data
- Analysis of data and formulation of hypotheses (iterative)
- Validate hypotheses by means of experiments (iterative)
- <u>Goal:</u> Identification and characterization of new mechanisms leading to a better understanding of the complex relations in a "populated urban structure"
- Integration in the resilience process

#### **Measurement of overall condition**





A **Simulation** is an approach used for the **analysis** of systems which are too complex to be analyzed theoretically or with formulas. This is mostly the case for **dynamic system behavior**. In a simulation, experiments are performed with a model in order to gain knowledge about the real system.

- Discovering of network structures on different levels
- Gathering of the temporal and spatial dynamic of "cellular/system" components using different (environmental) conditions
- Development of detailed mathematical models
- Visualization of networks and the processes taking place within those networks
  - -> Understanding of the processes as a whole

## **Multi-Agent-Systems**





Cellular automata and multi-agent systems can be used for the simulation of the dynamic of **spatial processes**.

Cellular automata model spatial elements such as streets, parcels of land, and buildings. Those are treated as locally fixed objects, whose state can change at a certain point in time. In a first step, these elements are transferred to a cell of a regular grid and saved as the status of this cell.





A second level includes the individual and collective urban actors, which will be called agents. In comparison to the cells, agents are mobile and can move freely through the cell grid, the cellular room. It is possible to define different types of communication between the agents and between the agents and the cell.



## FACS (Free Agents in a Cellular Space)



**FoSIL** 

Model of an urban system. The system is separated in different levels, which are represented either by the cells of a CA or the agents of the MAS.

## Ideas for Implementation

## P(Security) = 1- P(crime)

## Automata





#### **Cellular Automata**

- Interaction with neighboring cells in a fixed grid
- Neighborly relations have to be determined at the beginning

#### **Graph Automata**

- Neighbor geometry can be flexibly modified
- Neighborly relations can change at any time



## Near-repeat pattern analysis

## Question:

When a criminal event takes place, how does the risk develop in the surrounding area?

 $\rightarrow$  Repeated attacks at the same or nearby places

## $P(A|x,t) \rightarrow P(B|x + \Delta x, t + \Delta t)$

What influence does an assault A at the place x at the time t have at a later point in time  $t + \Delta t$  and a nearby place  $x + \Delta x$ .

Relations for time and space can be derived from statistics.





First "statistic of criminal activity" – people mark crimes on a map. Basic idea – divide in quadrants (grid)

## **BigData – Predictive Policing**







#### Number of criminal offences and their "circumstances"

## Near-repeat pattern analysis

#### **Odds Ratios**

	0 to 15 days	15 to 30 days	30 to 45 days	45 to 60 days	60 to 75 days
Same location	2.6803	0.9033	0.6894	0.4331	0.6208
0 to 121.92 meters	1.2824	1.1158	0.8223	0.9218	1.0476
121.92 to 243.84 meters	1.1851	0.9325	1.0807	0.9207	0.8254
243.84 to 365.76 meters	1.1203	0.9448	1.0811	0.9923	0.851
365.76 to 487.68 meters	1.065	1.0754	0.9932	0.9271	0.8171
487.68 to 609.6 meters	1.149	1.0396	0.9577	1.002	0.9482
609.6 to 731.52 meters	1.0259	1.0772	0.9504	0.9865	1.0612
731.52 to 853.44 meters	1.0309	1.0922	1.046	0.9247	1.0202
853.44 to 975.36 meters	1.0423	1.0727	1.0303	1.0683	0.9943
975.36 to 1,097.28 meters	1.0666	1.0307	1.0631	1.0236	0.9784
1,097.28 to 1,219.2 meters	1.0027	0.9971	0.9834	1.0447	1.04
> 1,219.2 meters	0.9943	0.997	0.9993	1.0004	1.0019

**Near-Repeat Patterns in Philadelphia Shootings**. Jerry H. Ratcliffe and George F. Rengert, Security Journal 2008.



## **Cyclic-load forecasting**

### **Question:**

Are there certain days in the year, certain days of the week or certain times in the day in which more criminal incidents are to be expected?

P(Event|Point in time) = P(Event|Month) + P(Event|Day) + P(Event|Day of the week) + P(Event|Time of day) + ...



Result













Securit of a house in an urban structure with a well defined crime rate

### **Simulation Grid**





### **Simulation Urban Structure**





Primal Representation Modell [Porta et al. 2006]

- Graph G=(V,E)
- Intersections of streets are nodes V
- Enclosed segments of streets are the edges E



**Betweenness Centrality** 



## Modelling patterns of burglary on streetnetwork **Fosil**

model

The risk of a burglary happening at a certain point in time on a specific stretch of road shall be determined -> the risk passes from segments with a high risk to segments with a lower risk

 $R_i(t) = S_i + Q_i(t)$  risk -> static and dynamic components

 $S_i = (D_i + W_i)p_i,$ 

$$W_i = f(C_i).$$

D<sub>i</sub> constant for activities, which

cannot be explicitly captured

Activity of pedestrians (in a wider sense) is a function of betweenness centrality (function-> non-linear behavior)

$$p_i = 1 - e^{-B_i \delta t}.$$

Probability of a criminal offence -> probability within a given period of time, determined by the means of the initial attractiveness including a number of decisions by the offender

Breakdown



Growth of crime

diffussion from link to link

## Modelling patterns of burglary on streetnetwork FosiL

diffusion of the risk



Diffusion in different stretches of roads as an answer to burglaries



Balance at proportional statistical risk to C<sup>B</sup>



Risk is artificially suppressed



Street is taken out
# Modelling patterns of burglary on streetnetwork **FosiL**

Results and future work

- Mathematical network model for the simulation of the spread of crime
- Non-linear effects show the significance -> targeted Policing
- Requirement is the presence of relations between network characteristics and crime rates
- Even though an empiric basis was developed, there have been no analyses with the presented metrics
- Next steps:
  - Creating a model showing the consequences of the predictions made for the organization of police operations
  - statistical analyses
  - Influence of the network configuration on the phenomenon of the near-repeat victimization
  - Implementation in practical police work

# This is what defines us.... Our Identity



**Separation?** 

## Introduction



statista



Source:

Statista 2015



Number of annual cyber attacks in the years 2009 to 2014 (in millions)

Weitere Informationen: Weltweit: 27. März bis 25. Mai 2014; Mehr als 9.700; (Sicherheits )Technische Leiter, Geschäftsführer etc.

# Introduction

**Fo**SIL





2013: using a hacked account of the US news agency AP, hacker spread the news that of an explosion in the White House  $\rightarrow$  within minutes the Dow Jones looses about 1 %

2012: within 45 minutes the publically traded company Knight Capital losses US\$ 440 Million because of a flawed computer software (Malware)  $\rightarrow$  Shares loose 75 % of their value

# Modern communication





Communication and content have changed due to the digitalization.

Digital identity and real identity

# Modern communication and crime



Modern ways of communication is not only used in order to perform criminal acts, yet especially to plan criminal activities.



Consideration of digital communication for the modeling and simulation of criminal offences.



Prediction of criminal offences (tendencies), which are planned in virtuality and executed in reality ...

P(security) = 1- P(crime)





# **Crime scene hypotheses cycle**

# **Computational Forensics**



#### Human Individual Digital Footprint (HIDF) Network



- Communication
- Information transfer
- Feelings
- Emotions
- Sentiments
- Statements

. . . .



- Analyses
- Validation
- HIDF in groups, sub-groups
- Evaluation

# Individual Digital Footprint





Human Individual Digital Footprint (HIDF) Network

# **Digital Forensics**





An ontology is a specification of a conceptualization.



Definition of syntax of terms and symbols in a network of associations

# Social Networks



#### Human Individual Digital Footprint (HIDF) Network



CrossMark

# Computer-based personality judgments are more accurate than those made by humans

Wu Youyou<sup>a,1,2</sup>, Michal Kosinski<sup>b,1</sup>, and David Stillwell<sup>a</sup>

<sup>a</sup>Department of Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom; and <sup>b</sup>Department of Computer Science, Stanford University, Stanford, CA 94305



# Social networks -Sentiment analyses of groups in social networks and



#### **Sentiment Detection:**

- sub-area of text mining
- automatic analysis of texts with the aim of identifying an expressed attitude as positive or negative

#### trend analysis

Statistical analysis starts from a basic set of concepts (or n-grams), which is associated positive or negative sentiments.

- Occurrences of positive or negative sentiments
- Comparison
- Designation of trends, opinions Statements

## Inside the Web --- feelings, beliefs and personal opinions!

# Social networks



#### **Generation of person landscapes from Facebook**

PEGIDA – Patriotische Europäer gegen die Islamisierung des Abendlandes (Patriotic Europeans Against the Islamisation of the West)





"Slogan"

- Asylindustrie alysum industry
- Lügenpresse lies press
- Volksverräter public traitor
- Der Untergang des Abendlandes
- "the decline of the West"
- Islamisierung ... Islamic





PEGIDA – Patriotische Europäer gegen die Islamisierung des Abendlandes (Patriotic Europeans Against the Islamisation of the West)

A political movement

- far-right, conservative, nationalist
- anti-Islam
- formed in October 2014
- present in Saxony region of Germany, many offshoots exist(ed) in other German cities and countries

Causes for the group's sudden uprising and gain in popularity:

- Nationalist anti-EU movements had grown to significant sizes in entire Europe in 2014
- refugee crisis
- Acts of Islamist terrorism and violence in Europe
- $\rightarrow$ Increase of Islamophobia

# Social networks



#### **Generation of landscapes from Facebook**





#### **Generation of person landscapes from Facebook**

#### Information exchange via FB

f

Graph API application programming interface





#### **Generation of landscapes from Facebook**

#### Comments



Marika Heimer Hallo Merkel..tschüss bitte....Sie sind der Gefolgsmann der Industrie und ein Verräter See Translation Like · Reply · 🕑 3 · 3 hrs · Edited





Kerstin Muschke Diese Frau hat jeglichen Bezug zur Realität verloren, ich frage mich, WER oder WAS sie noch stoppen kann ... See Translation



#### Definition of a general Data Model – based on the Graph API

#### Information from FB via API -> Data model



#### Search relations and correlations

#### and sentence





IMMM2015

# Das Phänomen Massendaten – Soziale Netzwerke



#### **Generierung von Wissenslandkarten aus Facebook**



South - Social Here	NOIN PHIBIJ2EI								- ·
File Help									
Navigation TermTree									
Navigation									
FC N	Jew > ync Project Delete Project	Projec	t Dok Analysis						
Visualization Pr	roperties								
Time Range									
Activate	16.2017 ~ To: 16.06.2	017 ~							
Filter									
Activate  Relevant co Relevant pi Relevant us Show topic	omments only osts only sers only ts								
Search Results		- 0							
Search		6	Outline						
			Post			Comments			
		Search	Metadata			All			
Scoper	Ontions		Date:	Author:		filter comments	filter authors		
Users	Ignore Case		Merrana			Comment		Author	
Posts	Whole-Word		ine sough						
Comments	Repex								
Belevant only	Phonetic					Selected			
Chercian cong						Date:	 Author:		
									^



etch			
Facebo	ok Page ID/URL		
Citizer	IsForTrump		
Limits		-	
Time	e Range	O Amount	
From:	12.06.17 12:14	Posts:	5
To:	14.06.17 12:14	Comments	5
	No time limits for comments, links, and likes	Photos:	5
		Videos:	5
LDA			
🗌 Extra	act topics from posts and comments		
(Can to	ike several hours with large ammounts of data.		
munut	cuse, extraction after download is recommended.)		



# Social networks – Detection of Information





- User activities
- Topics
- Subgroups
- User-specific content
- Sentiments of the network

Participants FB\_ID  $\rightarrow$  user profiles





All information are labeled by time and user!

Analyze der opinion in der Group

- a. positive
- b. negative
- c. neutral

#### Sentiment Analyze



Time and geographical Monitoring



# **Opinions and moods (Sentences)**



# Digitized traces - geographical Monitoring

**Digital traces** 

## **Digitized traces - geographical Monitoring**





#### Foto BigData





#### Video BigData





# What makes pictures and videos from the perspective of Forensics so interesting?

- Database forensics
- Mobile device forensics
- Forensic video
- Forensic audio
- Manipulation of pictures
- artificial generation of Pictures
- class of devices (digital camera, scanner, etc.)

# Content



## Pictures – Fotos - Content



#### **Basic Image Information**

Target file: IMG\_20151018\_142938734.jpg

Camera:	Motorola XT1072						
Lens:	3.5 mm (Max aperture f/2) (shot wide open)						
Exposure:	Auto exposure, Program AE, 1/38 sec, f/2, ISO 80						
Flash:	Off, Did not fire						
Date:	October 18, 2015 2:29:39PM (timezone not specified) (14 days, 14 hours, 16 minutes, 6 seconds ago, assuming image timezone of US Paville)						
Location:	Latitude/longitude: 53° 35' 53.7" North, 9° 45' 7.7" East (53.598240, 9.752140)						
	Location guessed from coordinates: Feldweg 85 24, 22559 Hamburg, Germany						
	Map via embedded coordinates at: Google, Yahoo, WikiMapia, OpenStreetMap, Bing (also see the Google Maps pane below) Altitude: 90 meters (295 feet)						
File:	<b>3,264 × 1,836</b> JPEG (6.0 megapixels) 3,543,061 bytes (3.4 megabytes)						
Color Encoding:	<b>WARNING:</b> Color space tagged as sRGB, without an embedded color profile. Windows and Mac browsers and apps treat the colors randomly.						
	Images for the web are most widely viewable when in the sRGB color space and with an embedded color profile. See my Introduction to Digital-Image Color Spaces for more information.						

P P III Extracted 512 × 288 53-kilobyte "Composite: ThumbnailImage" JPG Displayed here at 88% width (<sup>1</sup>/<sub>53</sub> the area of the original)



Click image to isolate; click this text to show histogram

#### camera- und shooting-related data:

- Aperture
- Time
- Place
- Camera type etc.

#### EXIF Metadata (TIFF or JPEG)



#### Geo-location data

#### **Digitized traces - geographical Monitoring**



#### View 2 more comments

Herta Hofeditz Das ist sonnenklarl Sonst würden die illegalen Glücksrifter nicht nur Deutschland, Schweden und Österreich amsteuern. Dann wären die auch nach Tschechien und Polen und Rumänien gegangen. Alles sichere Länder. Das liegt alles klar auf der Hand! See Translation

Like · Reply · 🙆 8 · 1 hr

Dieter Krause Das Sozialsystem wird kollabieren. Muss es, da reicht es nur kurz drüber nachzudenken. Genau deshalb handelt die Merkel verantwortungslos. Es wird dann alle treffen. See Translation

Like · Reply · 🙆 3 · 48 mins





🚯 61% 📱 🔹 Chat (2)





# comparison to famous persons





#### facial proportions

$$\mathbb{R}^{m \times n} = \begin{pmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m,1} & r_{m,2} & \cdots & r_{m,n} \end{pmatrix}$$

$$r_{i,j} = \frac{a_{i,j}}{a_{i+1,j}}$$





$$d(\mathbf{j}, \mathbf{k}) = d(\mathbf{k}, \mathbf{j}) = \sqrt{(r_{1,j} - r_{1,k})^2 + (r_{2,j} - r_{2,k})^2 + \dots + (r_{m,j} - r_{m,k})^2}$$
$$= \sqrt{\sum_{i=1}^m (r_{i,j} - r_{i,k})^2}.$$

# Euclidean distance to the pairwise distances of faces j and k





All information are labeled by time, locations and user!

# Sentiment Analyze

+ Posible geo-information Digital traces Digital traces Digital traces Digital traces Digital traces Digital traces Digitized traces Digitized traces Digitized traces Digitized traces

**Digitized traces** 



# Prediction

**BigData – Predictive Policing** 

# Questions





# **BigData – Predictive Policing**









Number of crimes and their "conditions"


# Towards Predictive Policing: Knowledge-based Monitoring of Social Networks

Michael Spranger, Florian Heinke, Steffen Grunert and Dirk Labudde

University of Applied Sciences Mittweida Mittweida, Germany Email: {*name.surname*}@hs-mittweida.de

Abstract—Increasing the resilience of the society against disorders, such as disasters, attacks or threatening groups, is one of the biggest challenges. Recent events highlight the importance of a resilient society and steps which are required to be taken in resilience engineering. A priori the optimal way to handle such adverse events is to prevent them, or at least provide appropriate courses of preparation. The essential requirement for every kind of preparation is information about relevant upcoming events. Such information can be gained for example from social networks and can form the basis for a long-term and short-term strategic planning by security forces. For that purpose, an application framework for knowledge-based monitoring of social networks is proposed, which is able to predict short-term activities as well as the long-term development of potentially dangerous groups, and discuss its basic concepts.

Index Terms-forensic; text processing; resilience engineering





- Extraction of profiles for monitoring
- Extraction of post or comment content relating to the threat ontology and a sentiment analysis
- $\rightarrow$  enables short-term reaction
- Simulation of temporal development of groups and hot-spots
- →enables long-term resource and strategic planning
- Increasing resilience



# **Opinions and moods (Sentences)**

## rights rioters in Leipzig-Connewitz









- comprised by a set of independent statistical testing units
- the role of each unit is to test the degree of change between sentiment statistics obtained for two points in time

Sentiment analyses of user comments made on officials PEGIDA's facebook page  $\rightarrow$  per day 'temper' tracking



The networking energy is the sum of all unit energy values → Representation of 'network awareness'

## Social networks – Predictive Policing





- an energy value is assigned to each unit
- for each observation (sentiment and comment frequencies), it is decided whether a given unit conducts a statistical test on the new data depending on its energy value
- in the test, the new data is compared to the data the unit processed during the last test!
  → the network memorizes the data and underlying dynamics!



Sentiment analyses of user comments made on officials PEGIDA's facebook page  $\rightarrow$  per day 'temper' tracking



Are energy peaks correlated to major criminal incidents with anti-Islam and anti-refugee background?

The networking energy is the sum of all unit energy values → Representation of 'network awareness'

## Social networks – Predictive Policing



# Sentiment analyses of user comments made on officials PEGIDA's facebook page $\rightarrow$ per day 'temper' tracking



The networking energy is the sum of all unit energy values → Representation of 'network awareness'



Fos

# Answer ..... next time

# **One Publications**



#### **CONFERENCE-PAPER**

"Semantic Tools for Forensics: Approaches in Forensic Text Analysis"

Autoren: Michael Spranger, Dirk Labudde

veröffentlicht: IMMM 2013, The Third International Conference on Advances in Information Mining and Management; ISBN: 978-1-61208-311-7

## **CONFERENCE-PAPER**

"Establishing a Question Answering System for Forensic Texts"

Autoren:

Michael Spranger, Dirk Labudde

## veröffentlicht:

3rd International Conference on Integrated Information, IC-ININFO ISSN: 1877-0428

### PAPER

"Semantic Tools for Forensics: A Highly Adaptable Framework"

Autoren: Michael Spranger, Stefan Schildbach, Florian Heinke, Steffen Grunert, Dirk Labudde veröffentlicht: ThinkMind ISBN: 978-1-61208-227-1

## **CONFERENCE-PAPER**

"Semantic Tools for Forensics: Towards Finding Evidence in Short Messages"

Autoren:

Spranger Michael, Dirk Labudde

veröffentlicht:

The Fourth International Conference on Advances in Information Mining and Management :: IMMM2014

## PAPER

"Towards Establishing an Expert System for Forensic Text Analysis"

Autoren:

Spranger Michael, Dirk Labudde

### veröffentlicht:

International Journal On Advances in Intelligent Systems, v 7 n 1&2 2014

CONFERENCE-PAPER "Towards Predictive Policing: Knowledge-based Monitoring of Social Networks" Autoren: Michael Spranger, Florian Heinke, Steffen Grunert, Dirk Labudde veröffentlicht: IMMM2015

# FEEL FREE TO ASK QUESTIONS



VISIT US AT: www.bioforscher.de/FoSIL