Presentation by

Dr. Phil Legg

Senior Lecturer in Computer Science

December 2017

Visual Analytics and Human-Machine Decision Support for Insider Threat Detection

UWE Bristol | University of the West of England

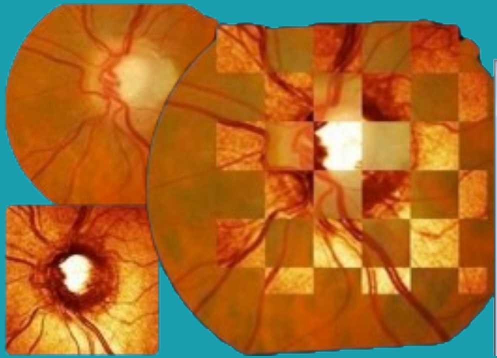
Background

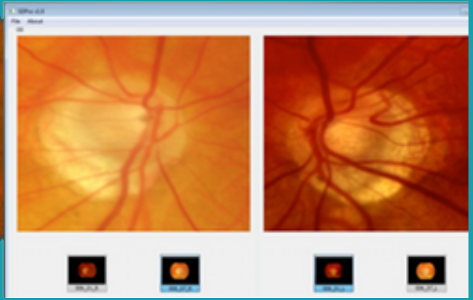
UWE Bristol | University of the West of England

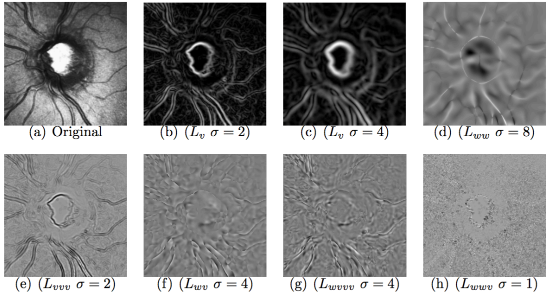
Background

PhD Cardiff University (2006-2010)

- Computer Vision
- Machine Learning
- Search Optimisation
- Grading Classification

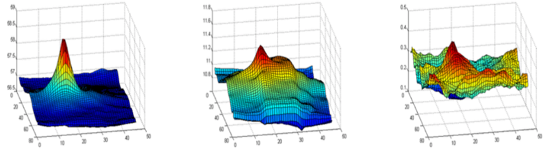






(a) Original (b) $(L_v\ \sigma = 2)$ (c) $(L_v\ \sigma = 4)$ (d) $(L_{ww}\ \sigma = 8)$

(e) $(L_{vvv}\ \sigma = 2)$ (f) $(L_{wv}\ \sigma = 4)$ (g) $(L_{wvvv}\ \sigma = 4)$ (h) $(L_{wwv}\ \sigma = 1)$



(a) FNMI $(L_v\ \sigma\ =\ 2)$ (b) FMI $(L_v\ \sigma\ =\ 2)$ (c) MI (16 bins)

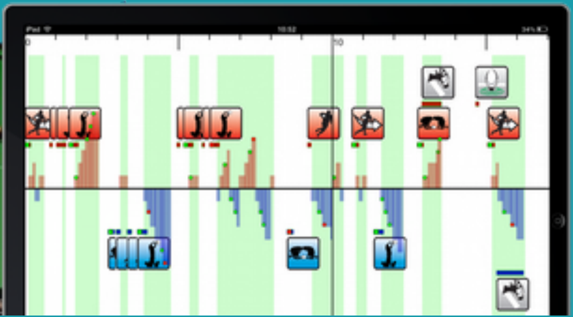
UWE Bristol University of the West of England

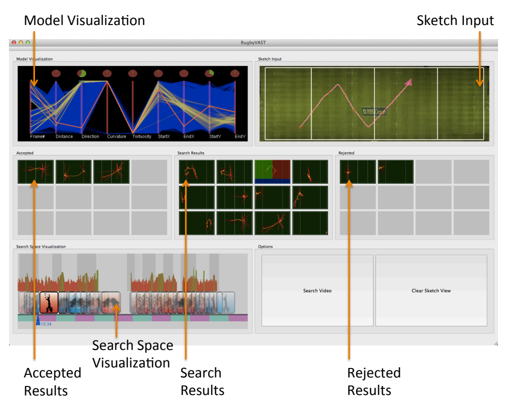
Background

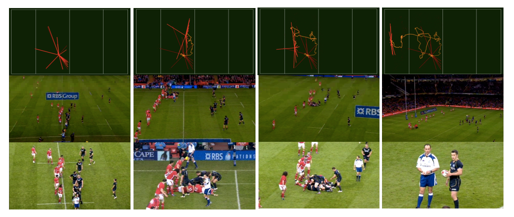
PDRA Swansea University (2010-2013)

- Visualisation
- Data Analytics
- Machine Learning

UWE Bristol University of the West of England



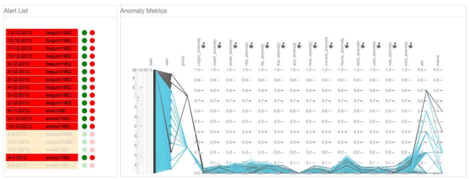


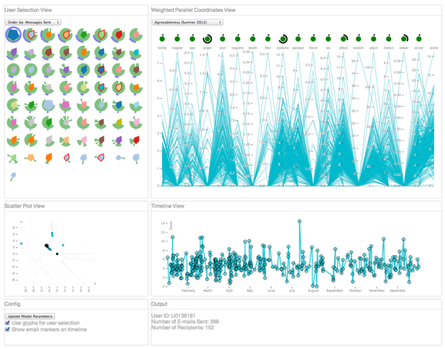


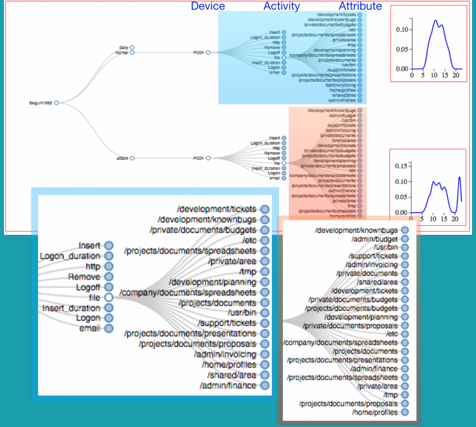
Background

PDRA University of Oxford (2013-2015)

- Cyber security
- Insider Threat Detection
- Machine Learning
- Visual Analytics







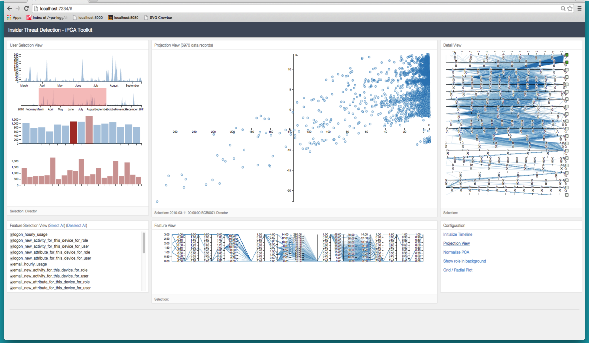
UWE Bristol University of the West of England

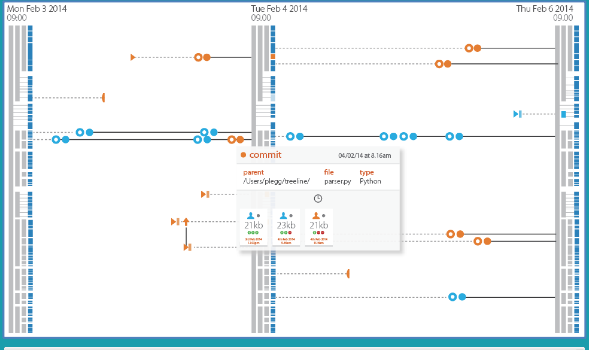
Background

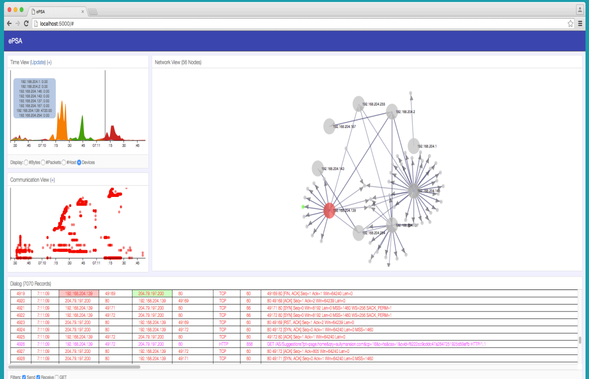
Senior Lecturer at UWE (2015-Present)

- Security Data Analytics and Visualisation
- (Inter) Active Machine Learning
- Human-Machine Collaboration

UWE Bristol | University of the West of England







Today

- Corporate Insider Threat Detection
- Automated Detection of Insider Threat
- Visual Analytics of Insider Threat Detection
- Active Learning for Insider Threat Detection
- Future Directions

Corporate Insider Threat Detection

- What is an insider threat?
  - Someone with **privileged access and knowledge** of an organisation, who uses this in a way that is detrimental to the operation of the organisation
    - E.g., employees, management, stakeholders, contractors.

Corporate Insider Threat Detection

- What is an insider threat?
  - Examples threats may include intellectual property theft, data fraud, system sabotage, and reputational damage.

  - Typically, a threat would be initiated by a **trigger** and a **motive** (e.g., personal financial difficulties result in theft).

2015 Vortmetric Insider Threat Report

- “**93%** of U.S. organisations polled responded as being vulnerable to insider threats”.

- “**59%** of U.S. respondents stated that privileged users pose the biggest threat to their organisation”

Defending Against the Wrong Enemy: 2017 SANS Insider Threat Survey

Which category of insider has the potential to be the most detrimental to your organization? ***Select the best answer.***

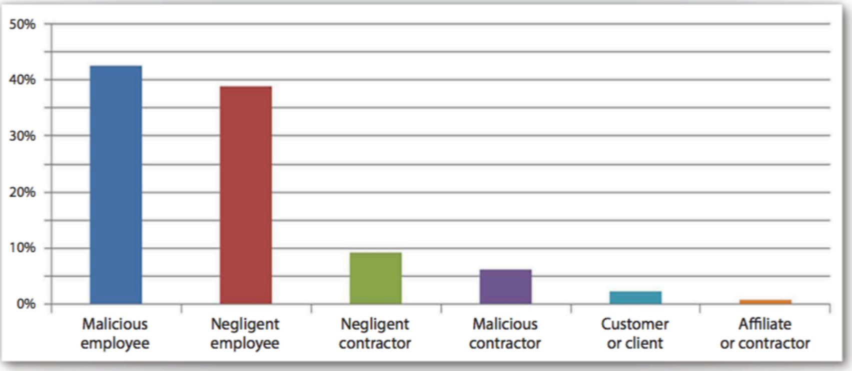
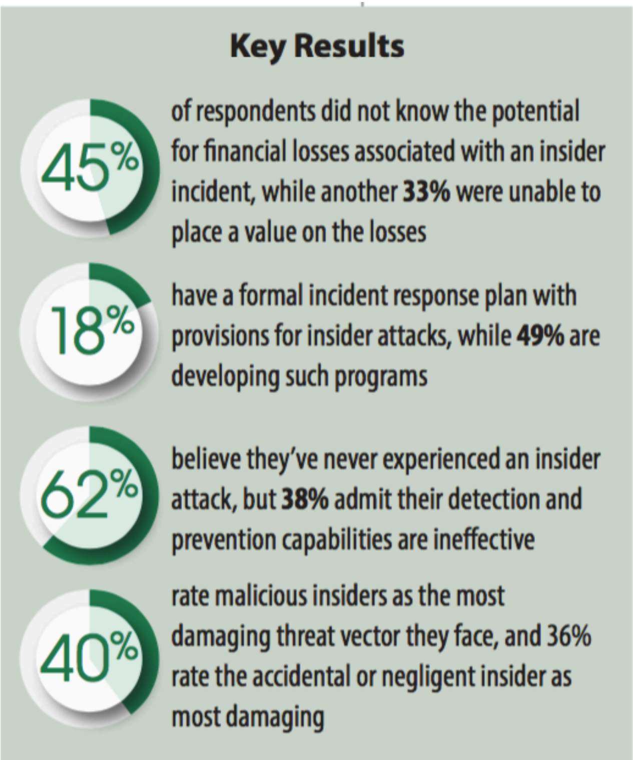


Figure 11. Malicious and Negligent Employees Potentially Damaging



Key Results

45% of respondents did not know the potential for financial losses associated with an insider incident, while another **33%** were unable to place a value on the losses

18% have a formal incident response plan with provisions for insider attacks, while **49%** are developing such programs

62% believe they've never experienced an insider attack, but **38%** admit their detection and prevention capabilities are ineffective

40% rate malicious insiders as the most damaging threat vector they face, and 36% rate the accidental or negligent insider as most damaging

Defending Against the Wrong Enemy: 2017 SANS Insider Threat Survey

Has your organization placed a financial value in U.S. dollars on its potential loss from an insider threat? If so, which of the following ranges best reflects your estimated value of loss?

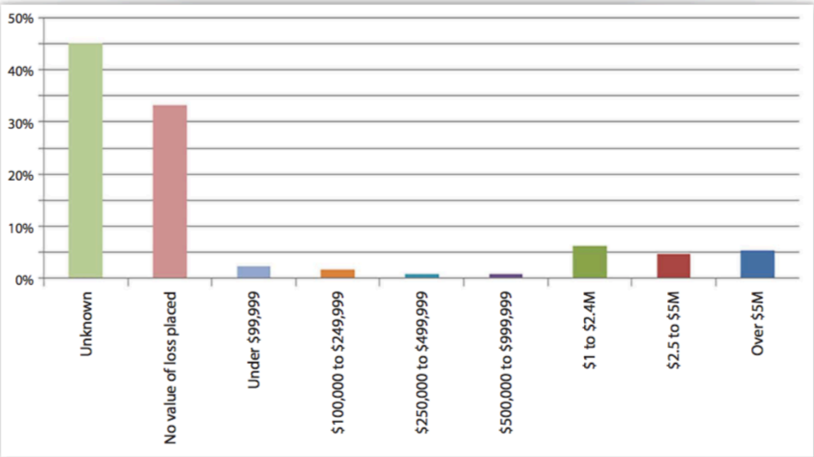


Figure 4. Values of Potential Loss

How effective do you consider your insider threat prevention and detection methods to be?

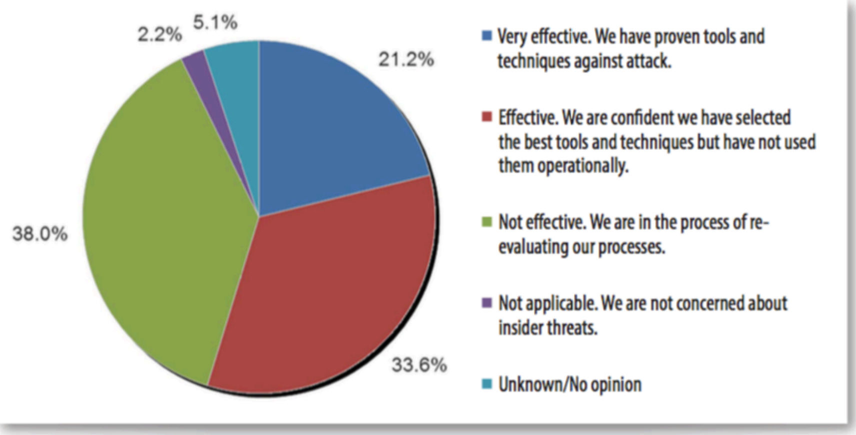
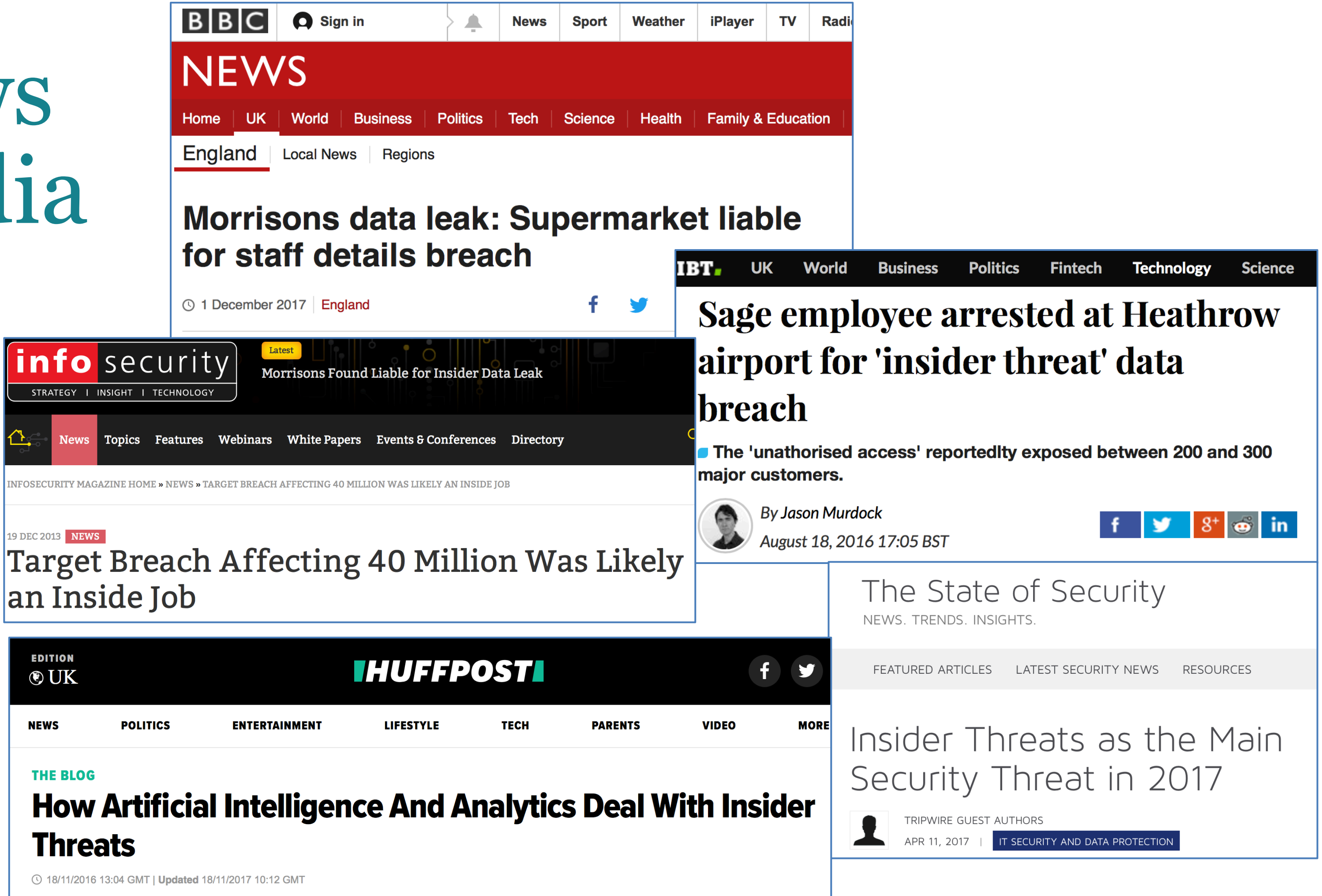


Figure 10. Efficacy of Insider Threat Prevention and Detection[4]

News Media



BBC Sign in News Sport Weather iPlayer TV Radi

NEWS

Home UK World Business Politics Tech Science Health Family & Education

England Local News Regions

Morrisons data leak: Supermarket liable for staff details breach

1 December 2017 | England

info security
STRATEGY | INSIGHT | TECHNOLOGY

Latest
Morrisons Found Liable for Insider Data Leak

News Topics Features Webinars White Papers Events & Conferences Directory

INFOSECURITY MAGAZINE HOME » NEWS » TARGET BREACH AFFECTING 40 MILLION WAS LIKELY AN INSIDE JOB

19 DEC 2013 NEWS

Target Breach Affecting 40 Million Was Likely an Inside Job

IBT. UK World Business Politics Fintech Technology Science

Sage employee arrested at Heathrow airport for 'insider threat' data breach

The 'unathorised access' reportedlty exposed between 200 and 300 major customers.

By Jason Murdock
August 18, 2016 17:05 BST

The State of Security
NEWS. TRENDS. INSIGHTS.

FEATURED ARTICLES LATEST SECURITY NEWS RESOURCES

Insider Threats as the Main Security Threat in 2017

TRIPWIRE GUEST AUTHORS
APR 11, 2017 | IT SECURITY AND DATA PROTECTION

EDITION
UK

HUFFPOST

NEWS POLITICS ENTERTAINMENT LIFESTYLE TECH PARENTS VIDEO MORE

THE BLOG

How Artificial Intelligence And Analytics Deal With Insider Threats

18/11/2016 13:04 GMT | Updated 18/11/2017 10:12 GMT

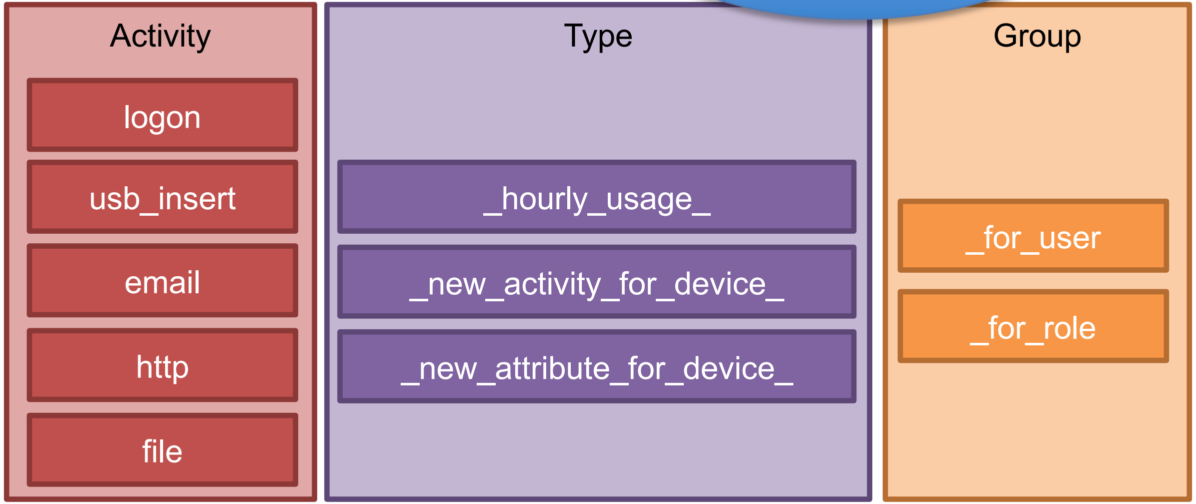
UWE Bristol University of the West of England

How may we attempt to detect insider threat?

- What data can we gather about users?
  - Log-on, E-mail, USB, File access, Web access?
  - Job role (any other HR related data)?

- What kind of 'features' can we calculate based on users?

This describes 30 numerical 'features' for each user per day to characterize the user behaviour



Activity

logon

usb_insert

email

http

file

Type

_hourly_usage_

_new_activity_for_device_

_new_attribute_for_device_

Group

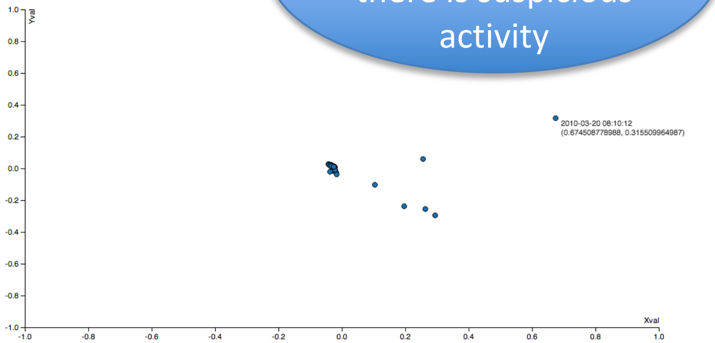
_for_user

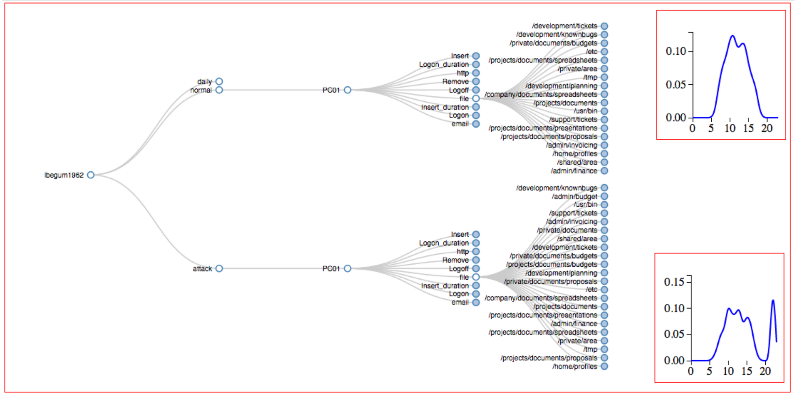
_for_role

How may we attempt to detect insider threat?

- How can we ‘detect’ from our features?
  - ○ Calculate distances from the norm?
  - ○ Use dimensionality reduction (e.g., PCA, t-SNE) to assess distances?
  - ○ Thresholding to flag suspicious activities for user

An outline suggests that there is suspicious activity





UWE Bristol University of the West of England

How may we attempt to detect insider threat?

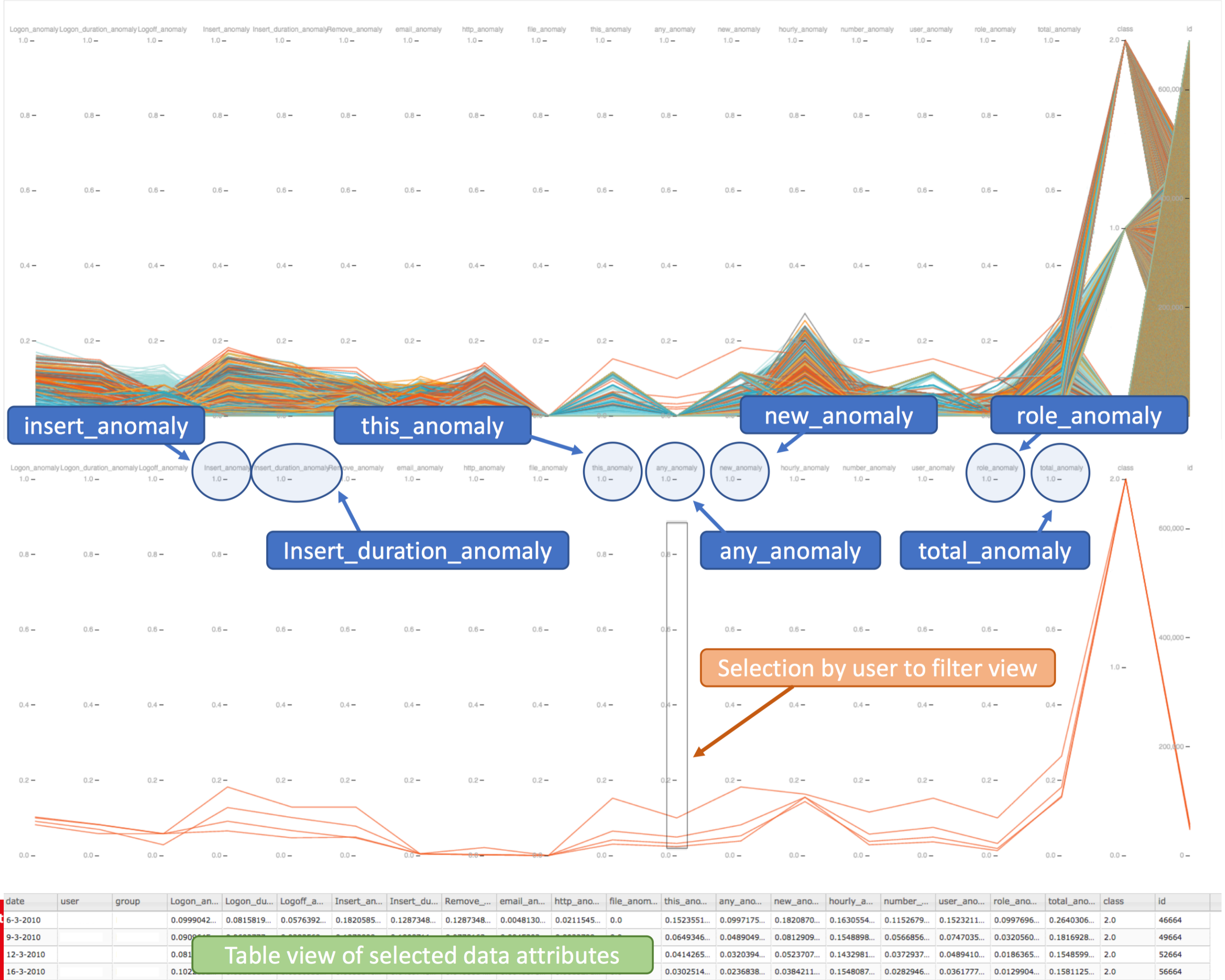
- How do we know that our features are well-selected?
- How can we inspect the performance of this detector?
- How do we know that the ‘score’ of the detector is valid?
- How can we report false positives to inform our system?

    - We need the **human analyst**,
      and we need **visual analytics** to assist them

Insider Threat Visualisation

UWE Bristol | University of the West of England

How can we assess features to identify anomalies?



insert_anomaly

this_anomaly

new_anomaly

role_anomaly

Insert_duration_anomaly

any_anomaly

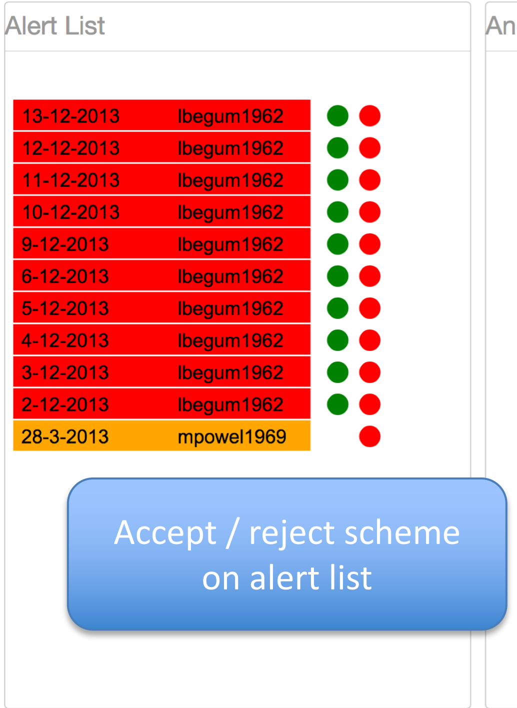
total_anomaly

Selection by user to filter view

Table view of selected data attributes

date	user	group	Logon_an...	Logon_du...	Logoff_a...	Insert_an...	Insert_du...	Remove_...	email_an...	http_ano...	file_anom...	this_ano...	any_ano...	new_ano...	hourly_a...	number_...	user_ano...	role_ano...	total_ano...	class	id
6-3-2010	[redacted]	[redacted]	0.0999042...	0.0815819...	0.0576392...	0.1820585...	0.1287348...	0.1287348...	0.0048130...	0.0211545...	0.0	0.1523551...	0.0997175...	0.1820870...	0.1630554...	0.1152679...	0.1523211...	0.0997696...	0.2640306...	2.0	46664
9-3-2010	[redacted]	[redacted]	0.090	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0649346...	0.0489049...	0.0812909...	0.1548898...	0.0566856...	0.0747035...	0.0320560...	0.1816928...	2.0	49664
12-3-2010	[redacted]	[redacted]	0.081	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0414265...	0.0320394...	0.0523707...	0.1432981...	0.0372937...	0.0489410...	0.0186365...	0.1548599...	2.0	52664
16-3-2010	[redacted]	[redacted]	0.102	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0302514...	0.0236838...	0.0384211...	0.1548087...	0.0282946...	0.0361777...	0.0129904...	0.1581125...	2.0	56664

How can we report false positives and reconfigure our model?



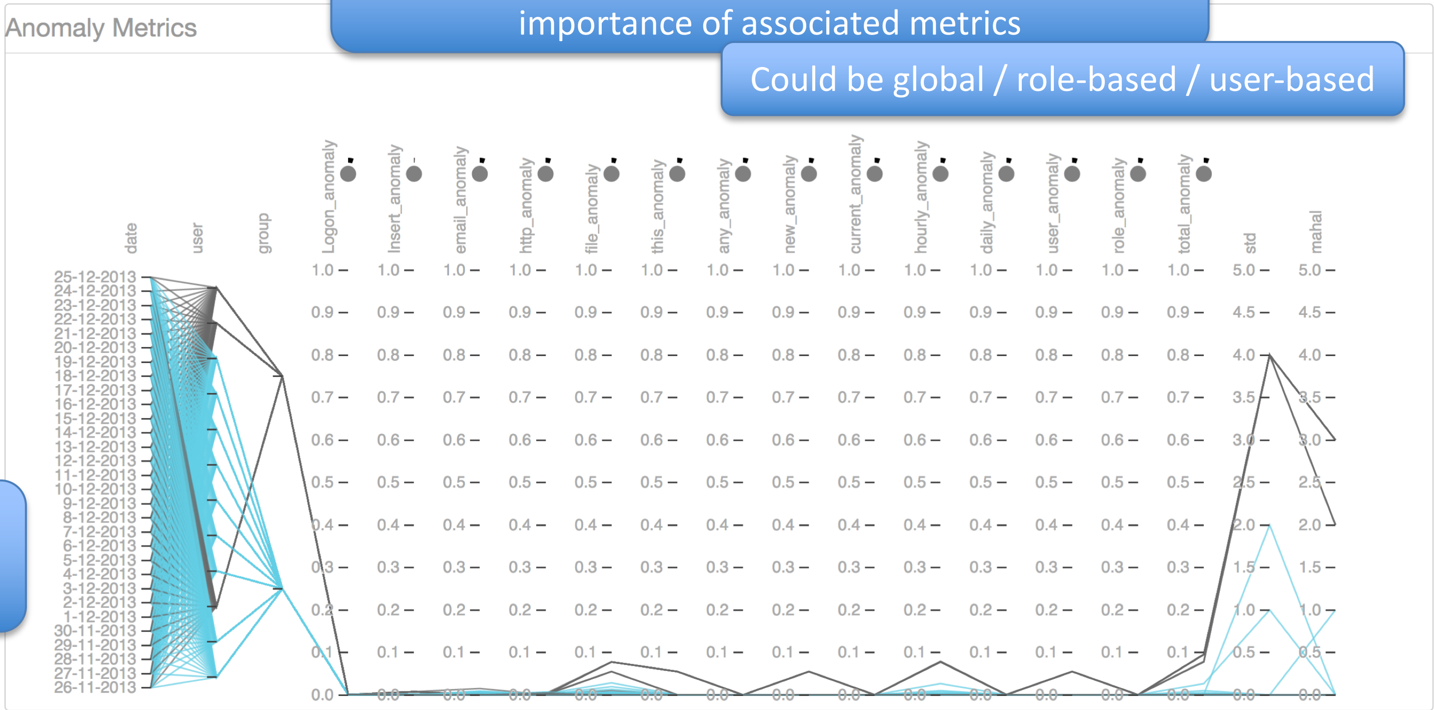
Alert List

13-12-2013	lbegum1962	🟢	🔴
12-12-2013	lbegum1962	🟢	🔴
11-12-2013	lbegum1962	🟢	🔴
10-12-2013	lbegum1962	🟢	🔴
9-12-2013	lbegum1962	🟢	🔴
6-12-2013	lbegum1962	🟢	🔴
5-12-2013	lbegum1962	🟢	🔴
4-12-2013	lbegum1962	🟢	🔴
3-12-2013	lbegum1962	🟢	🔴
2-12-2013	lbegum1962	🟢	🔴
28-3-2013	mpowel1969		🔴

Accept / reject scheme on alert list

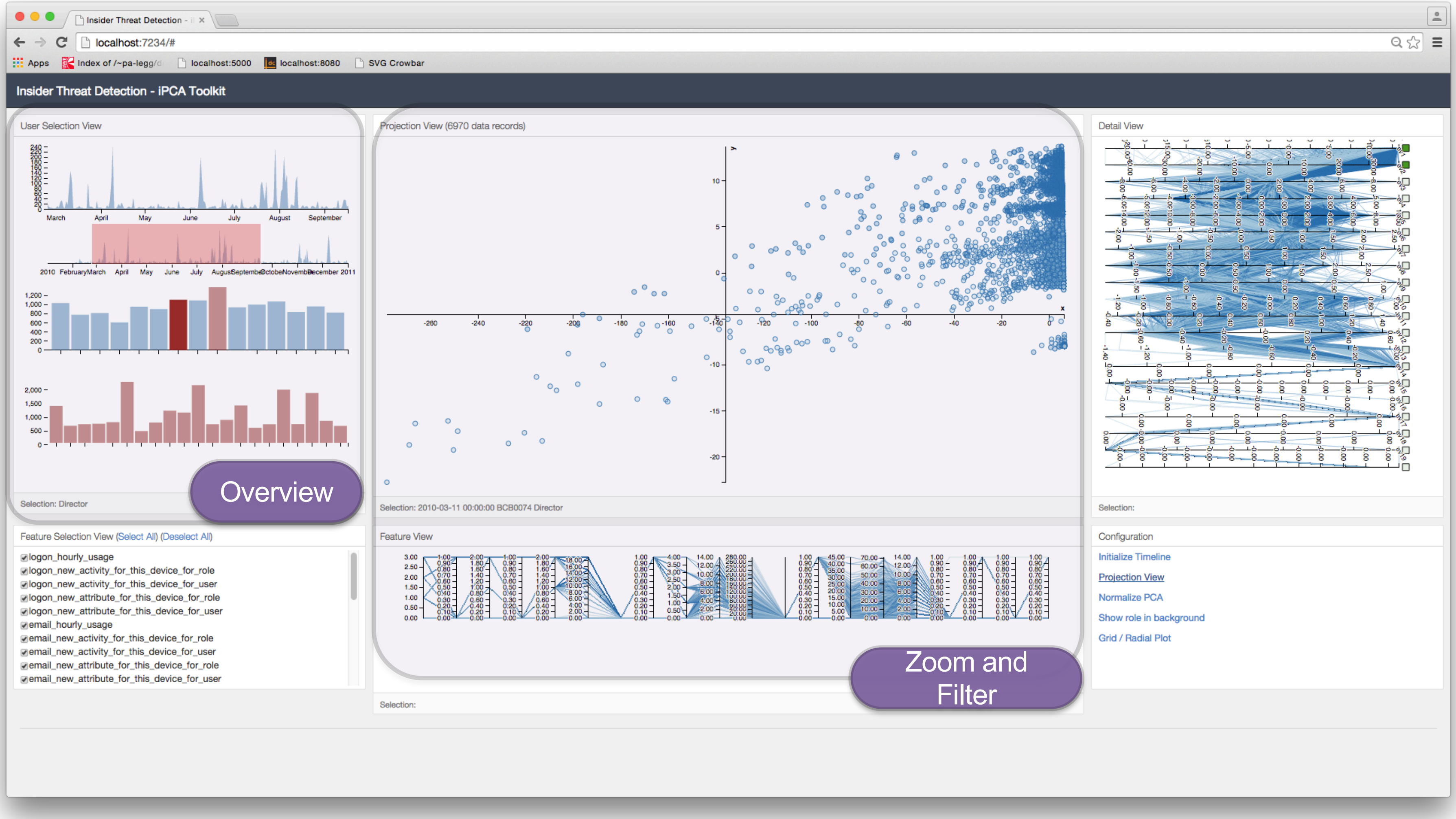
Weighted contribution of each anomaly type – acceptance/rejection will increase/decrease importance of associated metrics

Could be global / role-based / user-based



Anomaly Metrics

How can we assess detection model in context of activity?



Insider Threat Detection - iPCA Toolkit

User Selection View

Selection: Director

Projection View (6970 data records)

Selection: 2010-03-11 00:00:00 BCB0074 Director

Detail View

Selection:

Feature Selection View (Select All) (Deselect All)

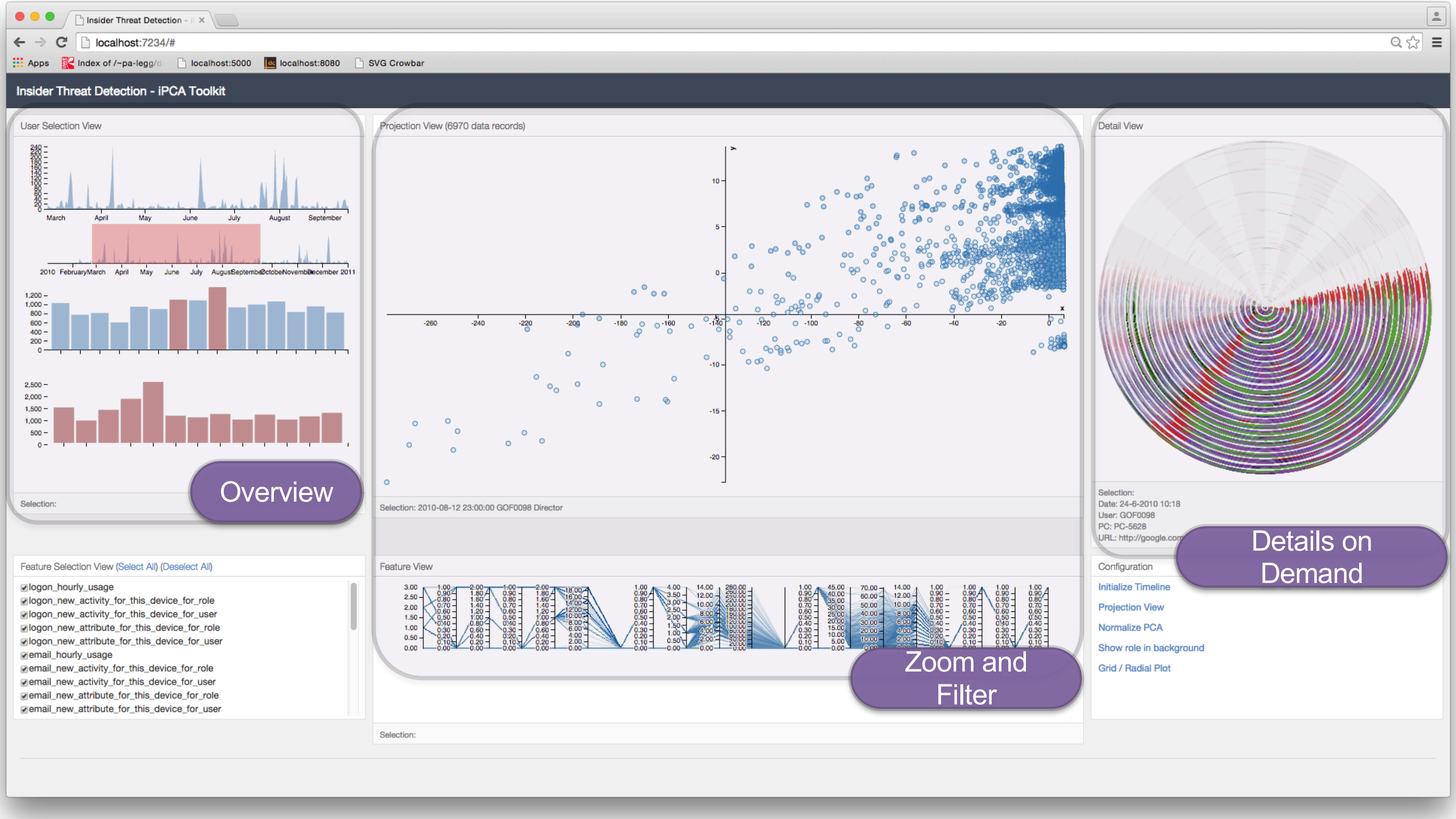
☑logon_hourly_usage
☑logon_new_activity_for_this_device_for_role
☑logon_new_activity_for_this_device_for_user
☑logon_new_attribute_for_this_device_for_role
☑logon_new_attribute_for_this_device_for_user
☑email_hourly_usage
☑email_new_activity_for_this_device_for_role
☑email_new_activity_for_this_device_for_user
☑email_new_attribute_for_this_device_for_role
☑email_new_attribute_for_this_device_for_user

Feature View

Selection:

Configuration

Initialize Timeline
Projection View
Normalize PCA
Show role in background
Grid / Radial Plot



Selection:

Selection: 2010-08-12 23:00:00 GOF0098 Director

Selection:
Date: 24-6-2010 10:18
User: GOF0098
PC: PC-5628
URL: http://google.com

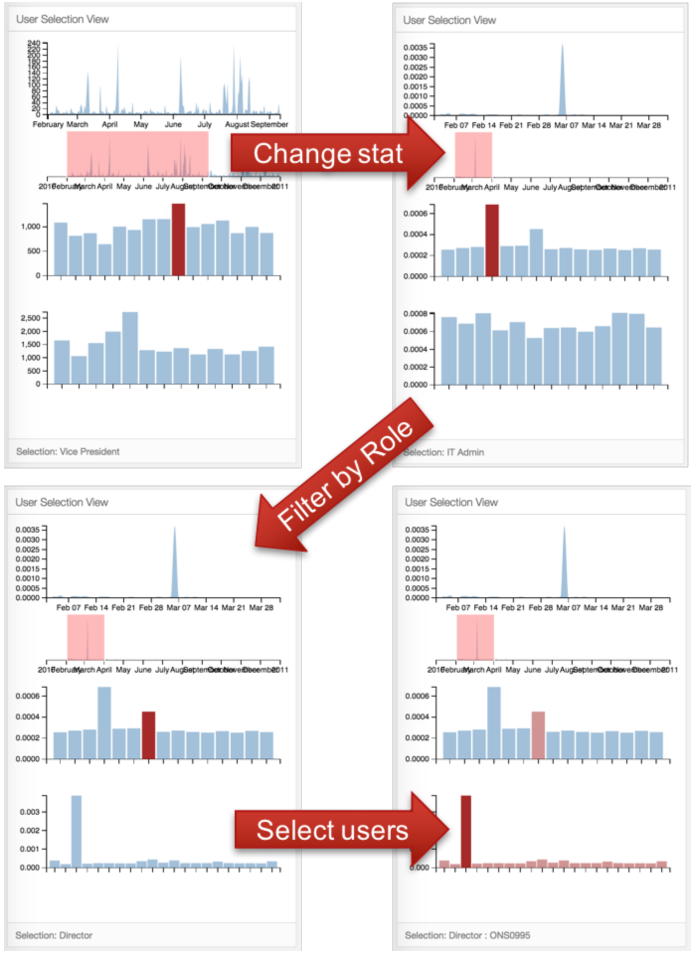
Feature Selection View (Select All) (Deselect All)

☑ logon_hourly_usage
☑ logon_new_activity_for_this_device_for_role
☑ logon_new_activity_for_this_device_for_user
☑ logon_new_attribute_for_this_device_for_role
☑ logon_new_attribute_for_this_device_for_user
☑ email_hourly_usage
☑ email_new_activity_for_this_device_for_role
☑ email_new_activity_for_this_device_for_user
☑ email_new_attribute_for_this_device_for_role
☑ email_new_attribute_for_this_device_for_user

Selection:

Overview

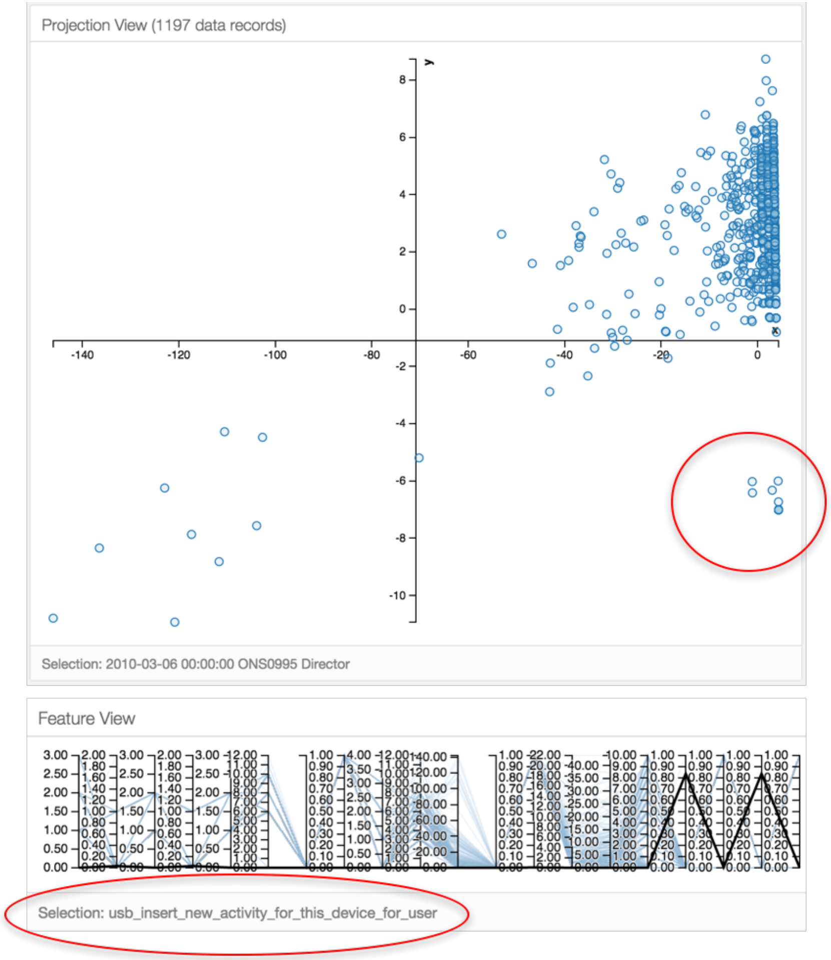
- Charts provide an interactive overview of selected summary statistics (e.g., amount of activity, deviation of activity).
- Support filtering (date range, selection).
  - ○ Zoomed view of activity by date.
  - ○ Contextual view of activity by date.
  - ○ Activity bar chart by job role.
  - ○ Activity bar chart by individual.


User Selection View
February March April May June July August September
2010 February March April May June July August September October November December 2011
Selection: Vice President
Change stat
User Selection View
Feb 07 Feb 14 Feb 21 Feb 28 Mar 07 Mar 14 Mar 21 Mar 28
Selection: IT Admin
Filter by Role
User Selection View
Selection: Director
Select users
User Selection View
Selection: Director : ONS0995

UWE Bristol
University of the West of England

Filter and Zoom

- Interactive PCA [Jeong *et al.*]
  - Scatter plot view of user daily activity based on PCA.
  - Parallel co-ordinates shows linked view between plot and profile features.
  - Can identify groups of outliers, and what features contribute towards the groupings.



Projection View (1197 data records)

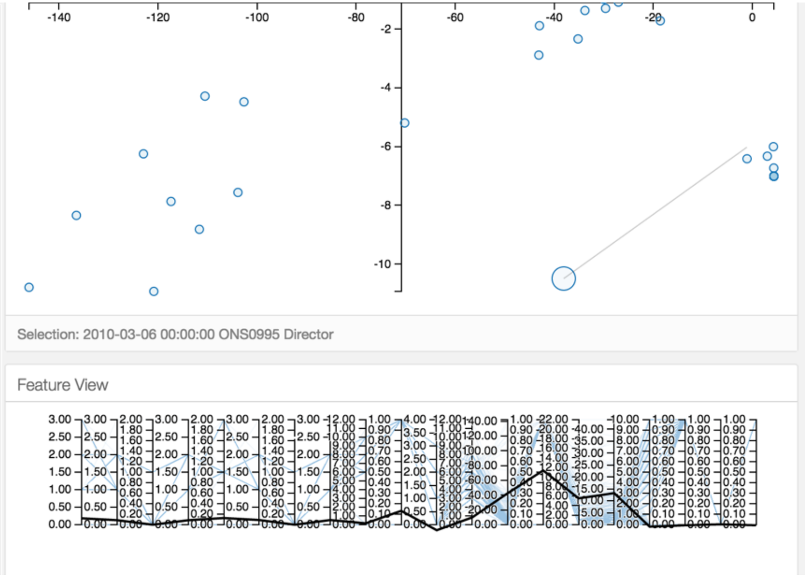
Selection: 2010-03-06 00:00:00 ONS0995 Director

Feature View

Selection: usb_insert_new_activity_for_this_device_for_user

Filter and Zoom

- Dragging points on scatter plot performs inverse PCA.
    - Analyst can examine relationship between the projection space and the original feature space.
    - Can be used to identify the contribution or *'usefulness'* of each feature for refinement of detection model (e.g., apply weighting function to PCA).

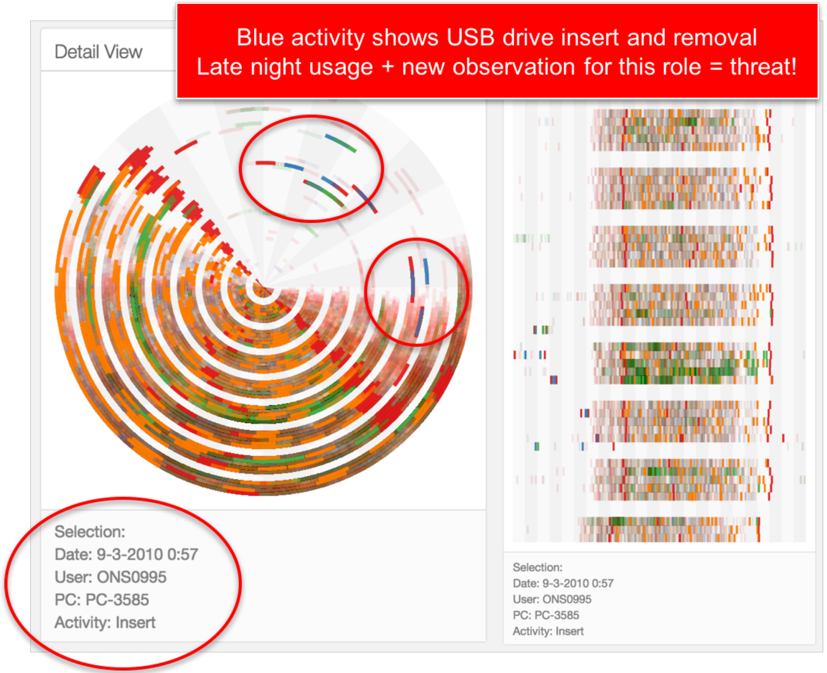


Selection: 2010-03-06 00:00:00 ONS0995 Director

Feature View

Detail View

- Activity plot that maps user and role activity to time (supports either polar or Cartesian grid layout).
- Comparison of user activity on a daily basis, and against others in the same job role.
- Could potentially be used in conjunction with other data if available (e.g., HR records, performance reviews).



Blue activity shows USB drive insert and removal
Late night usage + new observation for this role = threat!

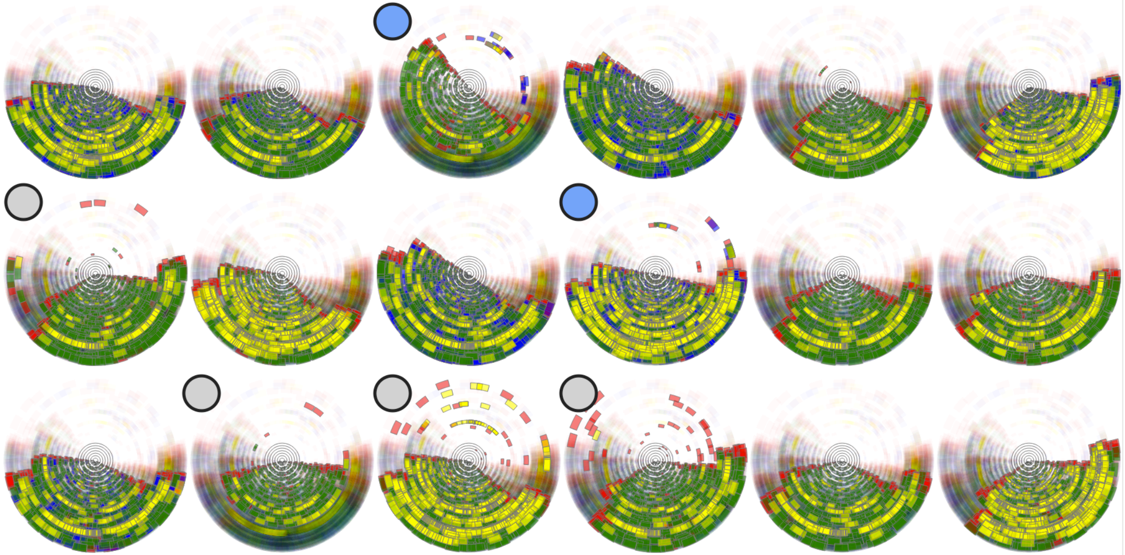
Selection:
Date: 9-3-2010 0:57
User: ONS0995
PC: PC-3585
Activity: Insert

Selection:
Date: 9-3-2010 0:57
User: ONS0995
PC: PC-3585
Activity: Insert

Role Overview

- We can use a ‘glyph’ visualisation approach for observing role overviews.

- 6 of the 18 roles have ‘activity of interest’ (flagged by circle)
  - 2 of these also contain USB activity during night (shown in blue)



Challenges and Limitations

Gathering Data

- How can we observe ‘all’ activity to inform insider threat detection?
- *Is it even ethical to observe ‘all’ activity to inform insider threat detection?*

- How realistic are synthetic datasets, and how can organisations work with academia to share data in this area? *(Recognising that no organisation wants to admit the issue of insider threat – let alone share data)*

- How can we be sure that we have an accurate model of normality? *(Some companies suggest they do not have a ‘normality’)*

Anomaly != Malicious

- How do we identify malicious activity – rather than only anomaly?
  - Requires human knowledge to distinguish – *yet we may struggle to train a classifier to recognise 'all' forms of malicious activity…*

  - How can the human analyst be more engaged to understand how the machine processes for detection / prediction are performed?

  - How do we separate out responsibility and decisions across multiple users? *(Who guards the guards?)*

Prevention better than cure

- Is it possible to prevent insider threat rather than detect it after the attack?
  - Requires understanding likely behavioural pre-cursors.

- Can we assess behavioural pre-cursors to a potential attack?
- What data is required for this (e.g., e-mail sentiment)?
- How do we (can we even) address this appropriately and *ethically*?

- If users understand how their data is being used (e.g., GDPR), can they not just 'game' the detection/prevention systems?

Thank you

Phil.Legg@uwe.ac.uk
@dr_plegg
2Q17, Frenchay, UWE

http://go.uwe.ac.uk/phil
http://www.plegg.me.uk

Related References:

- Legg, P. (2017) Human-machine decision support systems for insider threat detection . In: Palomares, Iván, Kalutarage, H. and Huang, Y., eds. (2017) Data Analytics and Decision Support for Cybersecurity: Trends, Methodologies and Applications. Springer. ISBN 9783319594385 [In Press] Available from: http://eprints.uwe.ac.uk/31385
- Legg, P. A. (2015) Visualizing the insider threat: Challenges and tools for identifying malicious user activity . In: IEEE Symposium on Visualization for Cyber Security, Chicago, Illinois, USA, 26 October 2015. IEEE Symposium on Visualization for Cyber Security (VizSec) 2015: IEEE Available from: http://eprints.uwe.ac.uk/27441
- Legg, P. A., Buckley, O., Goldsmith, M. and Creese, S. (2015) Caught in the Act of an Insider Attack: Detection and Assessment of Insider Threat. In: *IEEE International Symposium on Technologies for Homeland Security*, Waltham, USA, 14th-16th April 2015. Available from: http://eprints.uwe.ac.uk/26244
- Legg, P., Buckley, O., Goldsmith, M. and Creese, S. (2015) Automated insider threat detection system using user and role-based profile assessment.*IEEE Systems Journal*, 11 (2). pp. 503-512. ISSN 1932-8184 Available from: http://eprints.uwe.ac.uk/25809

UWE Bristol University of the West of England