RSAConference2020

San Francisco | February 24 – 28 | Moscone Center

HUMAN ELEMENT

SESSION ID: AFD-W11

A Sherlock Holmes Mystery: *AI-Powered Behavioral Forensics*

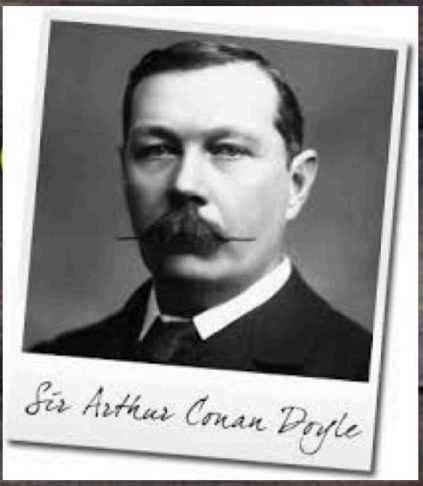
Uri Rivner

Chief Cyber Officer
BioCatch
LinkedIn: Uri Rivner

Erin Englund

Senior Threat Analyst
BioCatch
LinkedIn: Erin Englund

130 years ago…



"MY DEAR fellow," said Sherlock Holmes as we sat on either side of the fire in his lodgings at Baker Street, "life is infinitely stranger than anything which the mind of man could invent. We would not dare to conceive the things which are really mere commonplaces of existence.

A Case of Identity

Meet the Characters

Soundbytes are from a BBC broadcast of A Case of Identity adapted by Peter Mackie

- **Ms. Mary Sutherland (the Victim)**
  - Father died; mother re-married
  - Allowance of 100 pounds / year ($16,000 in today's terms)
- **Mr. James Windibank (the Stepfather)**
  - Draws the funds for Mary, deposits in her mother's account
  - Doesn't allow her to go to the *Gasfitter's ball*
  - *But… she goes anyway*
- **Mr. Hosmar Angel (the Lover)**
  - Met Mary at the ball when stepfather was abroad
  - When the stepfather returned, he was furious at Mary!
  - But love conquers all. Mary and Hosmar fell in love
  - He sent her typewritten love letters
  - They were engaged to marry – *a very serious business in Victorian times* - but he then disappeared mysteriously

RSAConference2020

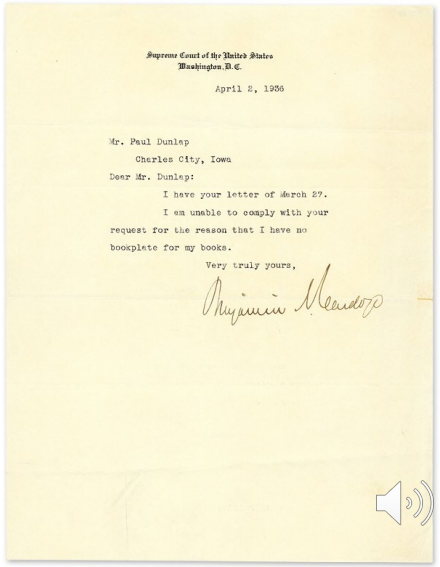
How did Holmes Crack the Case?

Generic Criminal Patterns

- **Spoofing Indicators**
  - Hosmar Angel was using tinted glasses ‘against the glare’
  - His voice had a speech impediment ‘due to a weak throat’
  - Masked his face with a moustache and a pair of bushy whiskers
- **Suspicious Behavior**
  - Did not provide his exact address
  - Preferred taking walks in the evenings
  - Insisted on writing letters with a *typewriter*, including the signature
- **Time Sequence Analysis**
  - Meetings with Hosmar always took place when the stepfather was in France
- **High-Risk events**
  - Engaged to Mary after their first walk (!)
  - Made her swear that whatever happened, she will always be true to him

Individual Criminal Profiling

A letter written by Mr. Windibank to Holmes



Supreme Court of the United States
Washington, D.C.

April 2, 1936

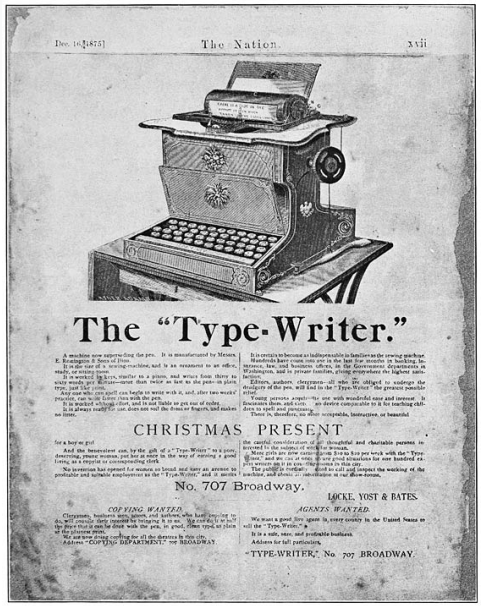
Mr. Paul Dunlap
Charles City, Iowa
Dear Mr. Dunlap:

I have your letter of March 27.

I am unable to comply with your request for the reason that I have no bookplate for my books.

Very truly yours,

Profiling basis: 4 love letters from Hosmar



The Nation. xvii

The "Type-Writer."

CHRISTMAS PRESENT

No. 707 Broadway.

LOCKE, YOST & BATES.

- **So… A Case of Identity is solved**
- **Hosmar Angel is none other than Mr. Windibank, the stepfather**
- **He introduced this traumatic episode so that Mary never falls in love, leave home, and depart him from her generous income**
- *Looking at the data science side: there are n* features in Holme's model
- In Hosmar's profile, 2 are very strong + 14 quite distinct

RSAConference2020

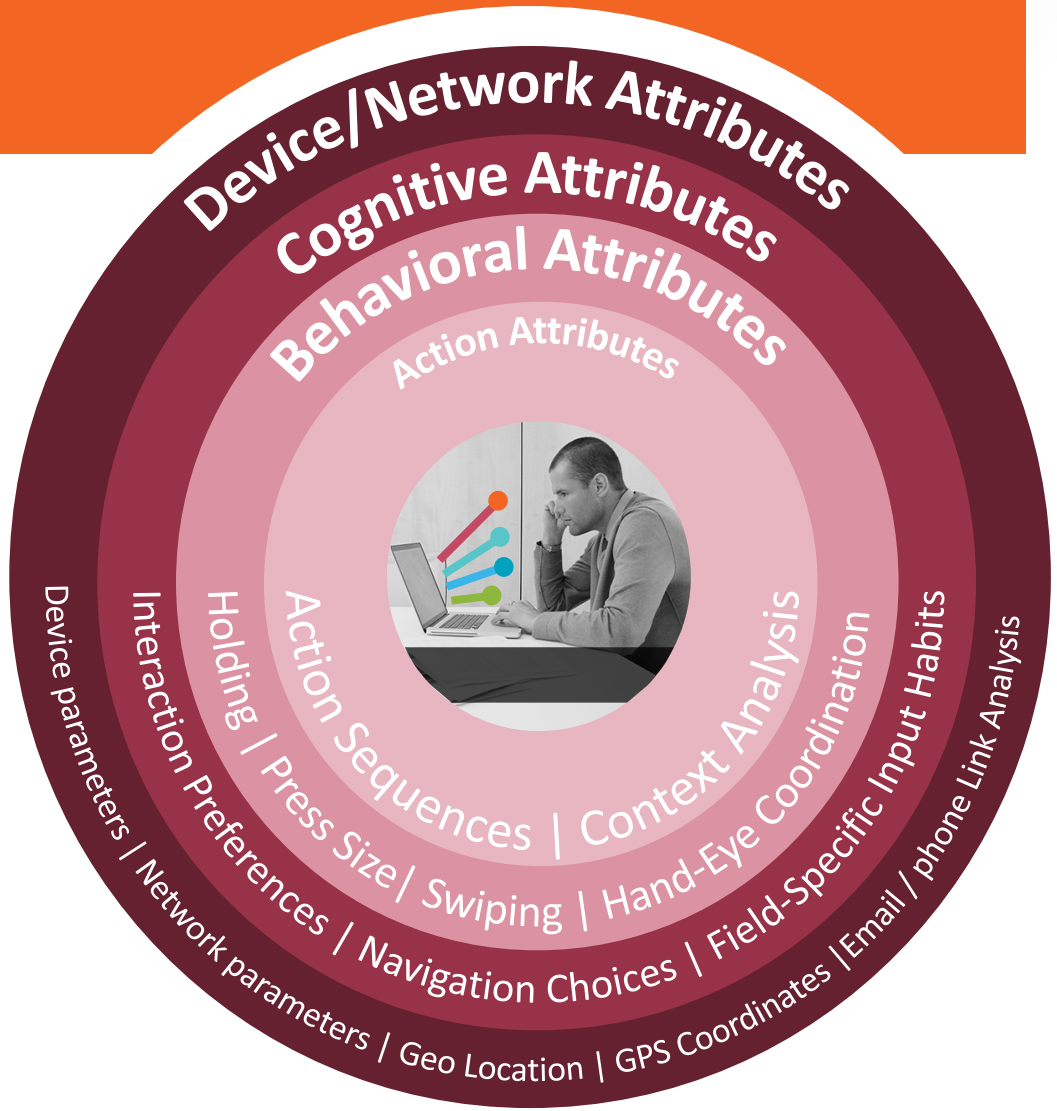
2020: using AI for criminal profiling

How to use AI in digital criminal profiling

- **Data**
  - Device fingerprints
  - Network, Geo location, GPS coordinates
  - Interaction (mouse, keyboard ,accelerometer, touch)
  - Context data – what is the criminal doing? Attack sequence
- **AI/ML**
  - **Benefits over blacklists, link analysis and rules**:
    - Adapts as criminals adjust resources and behaviors
    - Allows looking at a very large number of independent variables, even if each individually is not too 'incriminating'
    - Even if they knew what is being tracked, how quickly and easily can criminals change *all* of their behaviors?
- **Sharing**
  - Consortium data sharing allows tracking criminals across the industry

Digital Data Sources: Deep Dive

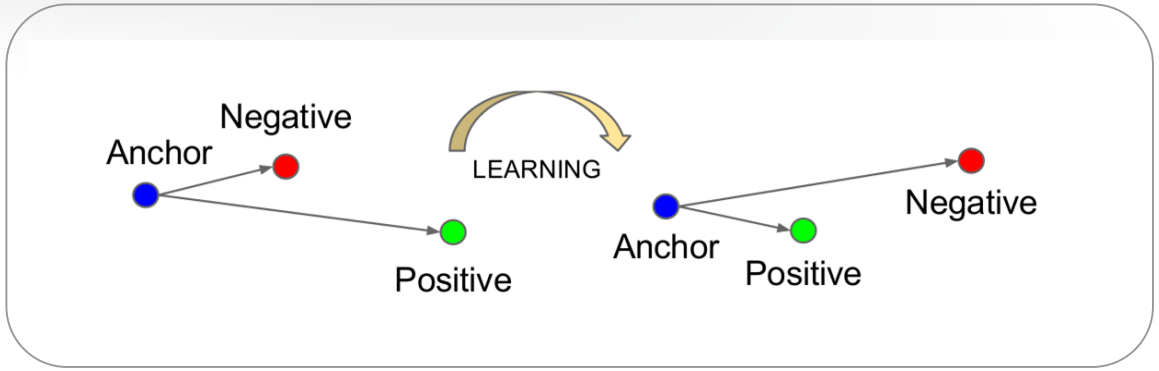
Building a **Criminal Digital Profile**

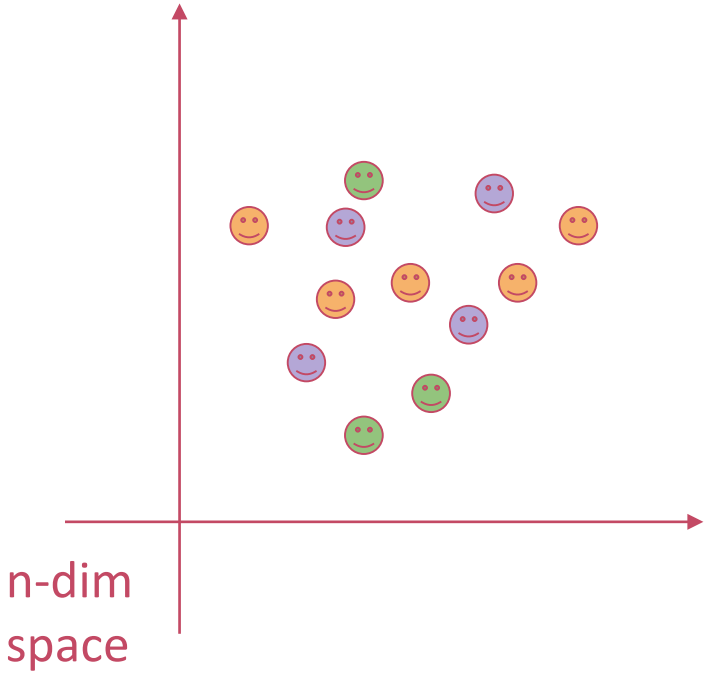


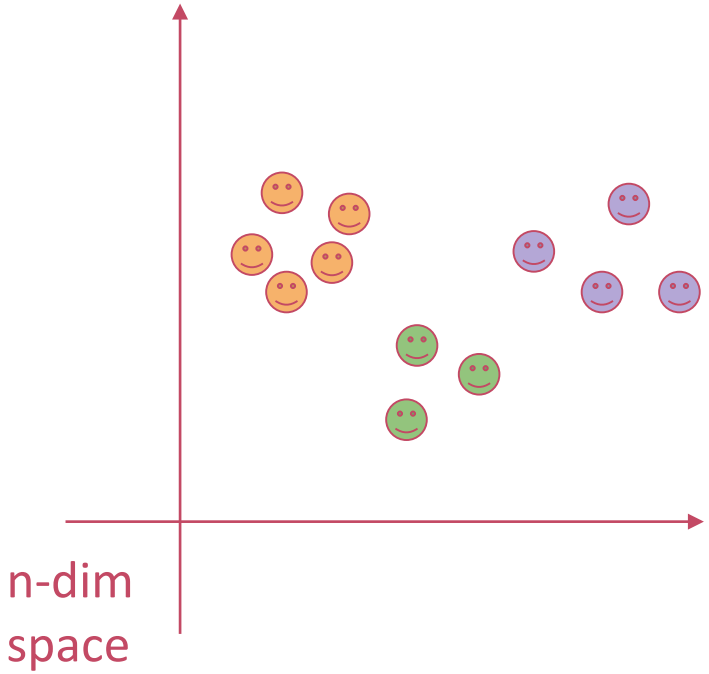
Training the Model

When the process starts, the samples (in our case, criminal sessions) are not separated from each other and crimes done by the same criminal are not clustered.

Example: Triplet Loss Optimization







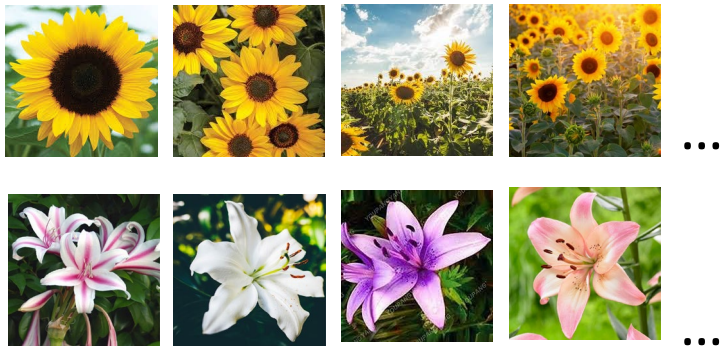
Now the crimes are neatly separated. Any new crime should fit an existing cluster (so we know it's done by that criminal), or start a new cluster (unknown criminal, and future crimes by same person will automatically match the new cluster)

ML vs AI

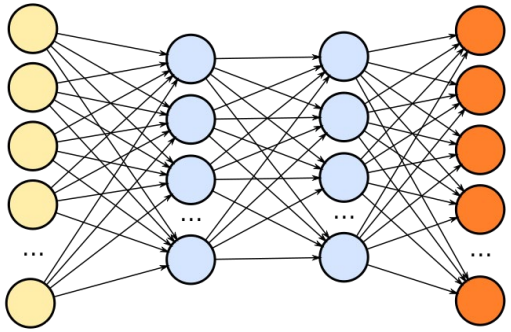
Define features

Machine Learning

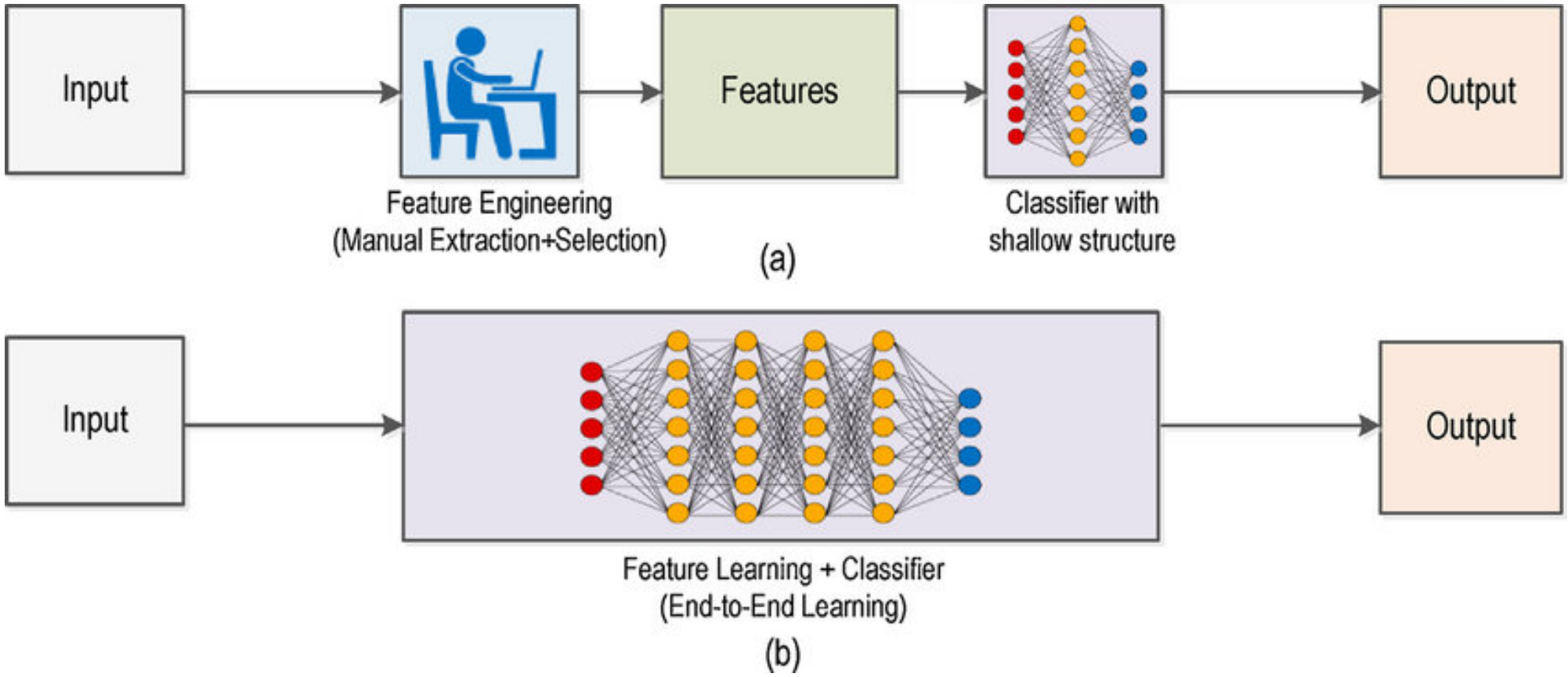
- Color
- Height
- Number of petals



AI: work on raw data

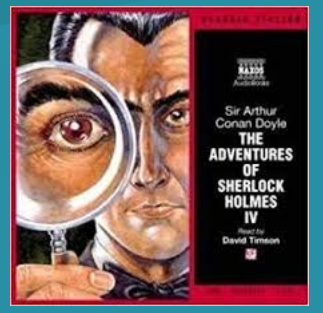


ML vs AI



ML: have data science build features and create criminal profiles (supervised ML)
AI: run deep learning – requires a large setting of known fraud cases, and not 'explainable'

RSAConference2020

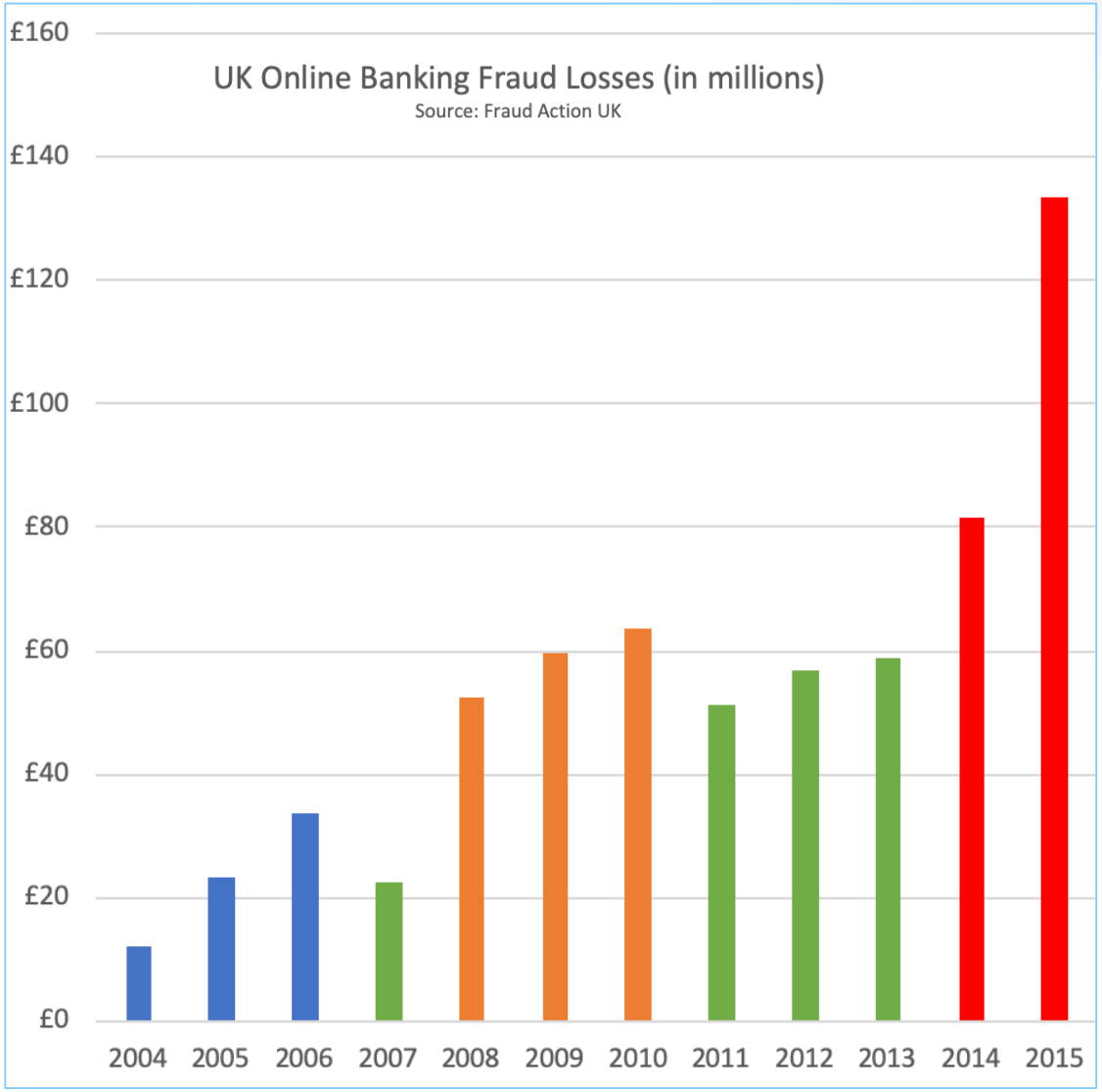


Digital Forensics: Real World Case Studies

The Trouble with Police Lineups

"Number 4, step forward and yell WHOOMAA, WHOOMAA!!!"

An Adaptive Race



Dyre: the scariest Trojan in 2015

Top 5 UK retail bank

00:00
Session start

Phase 1: Preparing for the Act

automatically being changed to allow criminal to receive OTP

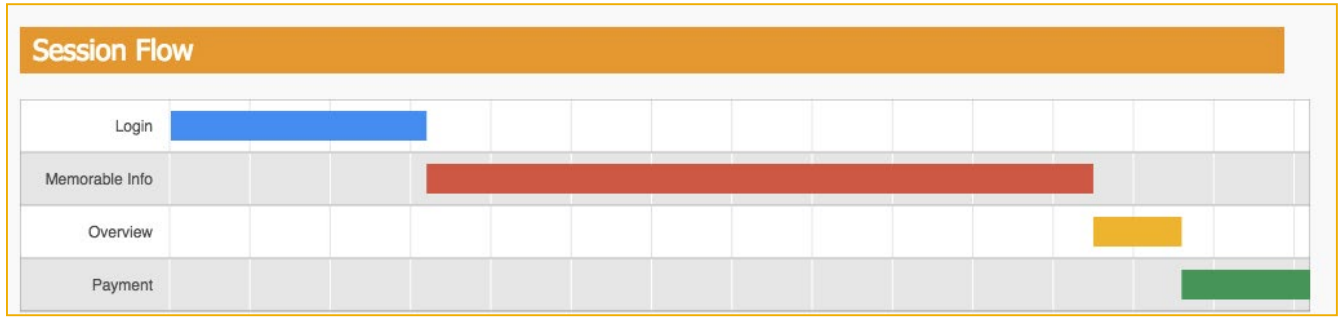
user is redirected
o they can continue as

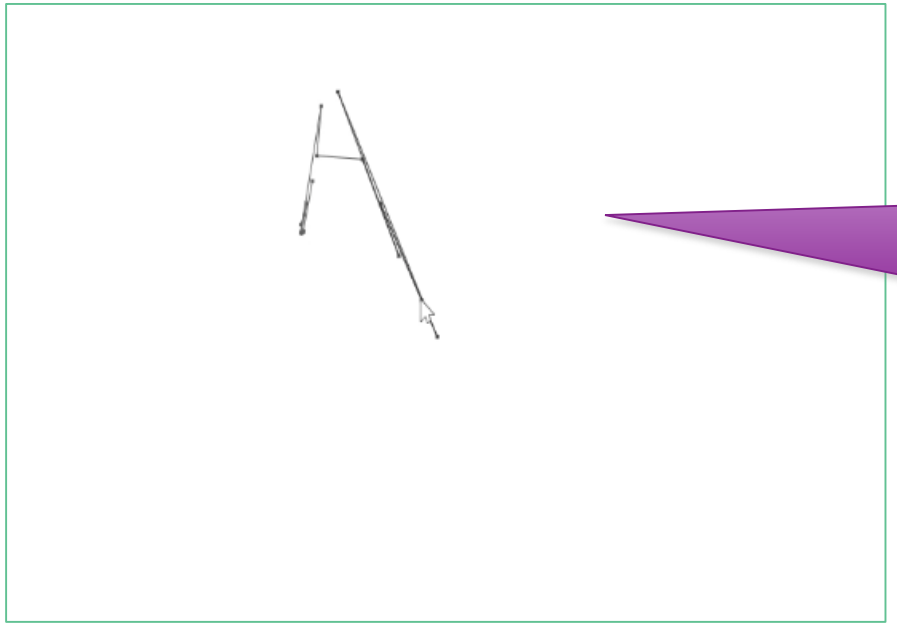
- User completely unaware of change of phone number

Into the mind of a Dyre Operator

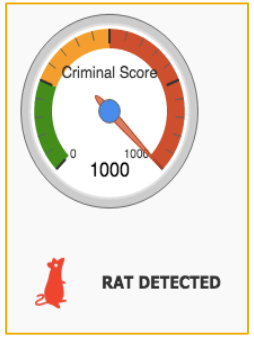
Top 5 UK retail bank

Phase 2: RAT attack two weeks later



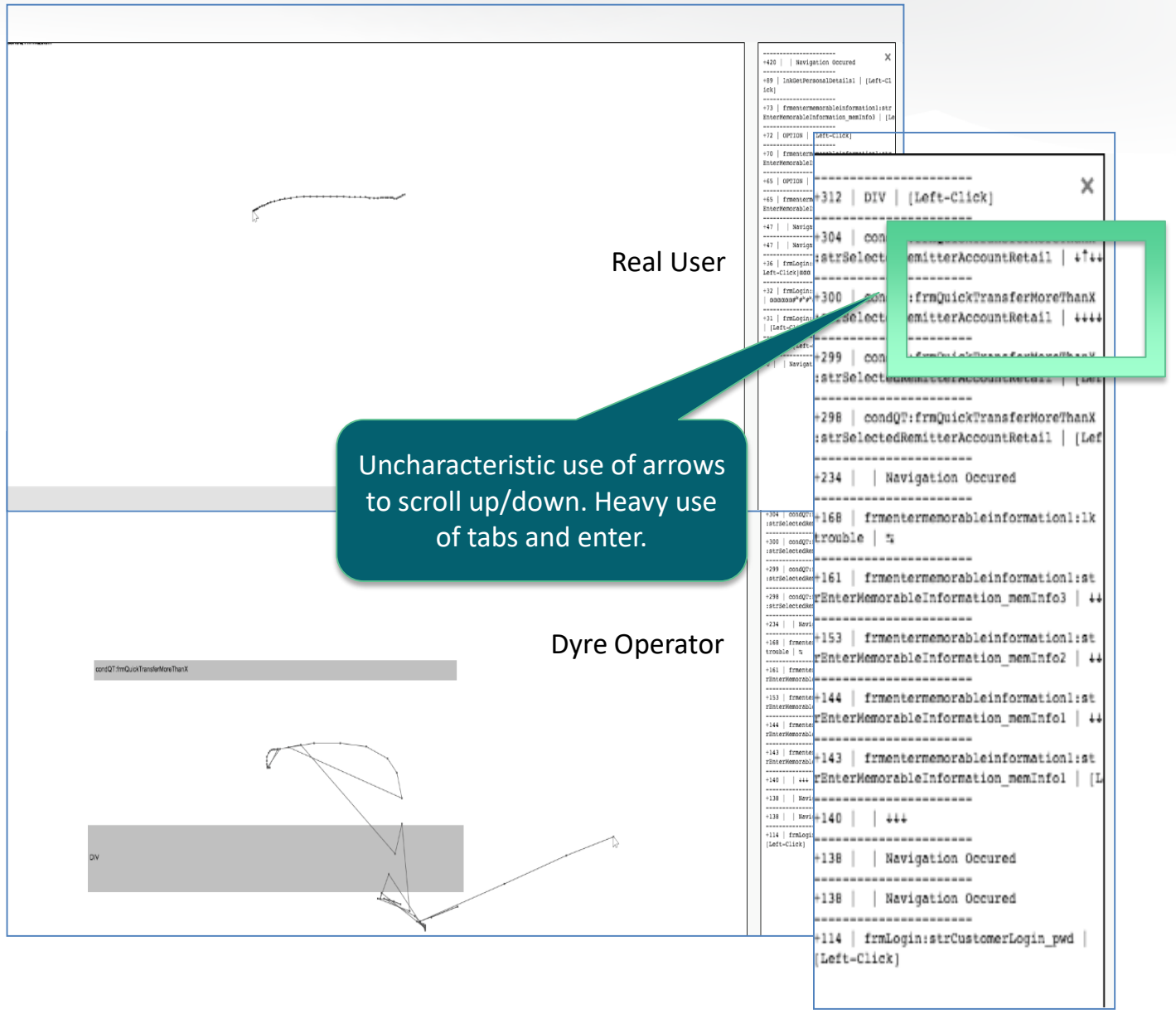


Hand-eye coordination shows strong signs of VNC back-connect capability. This is a nasty RAT case



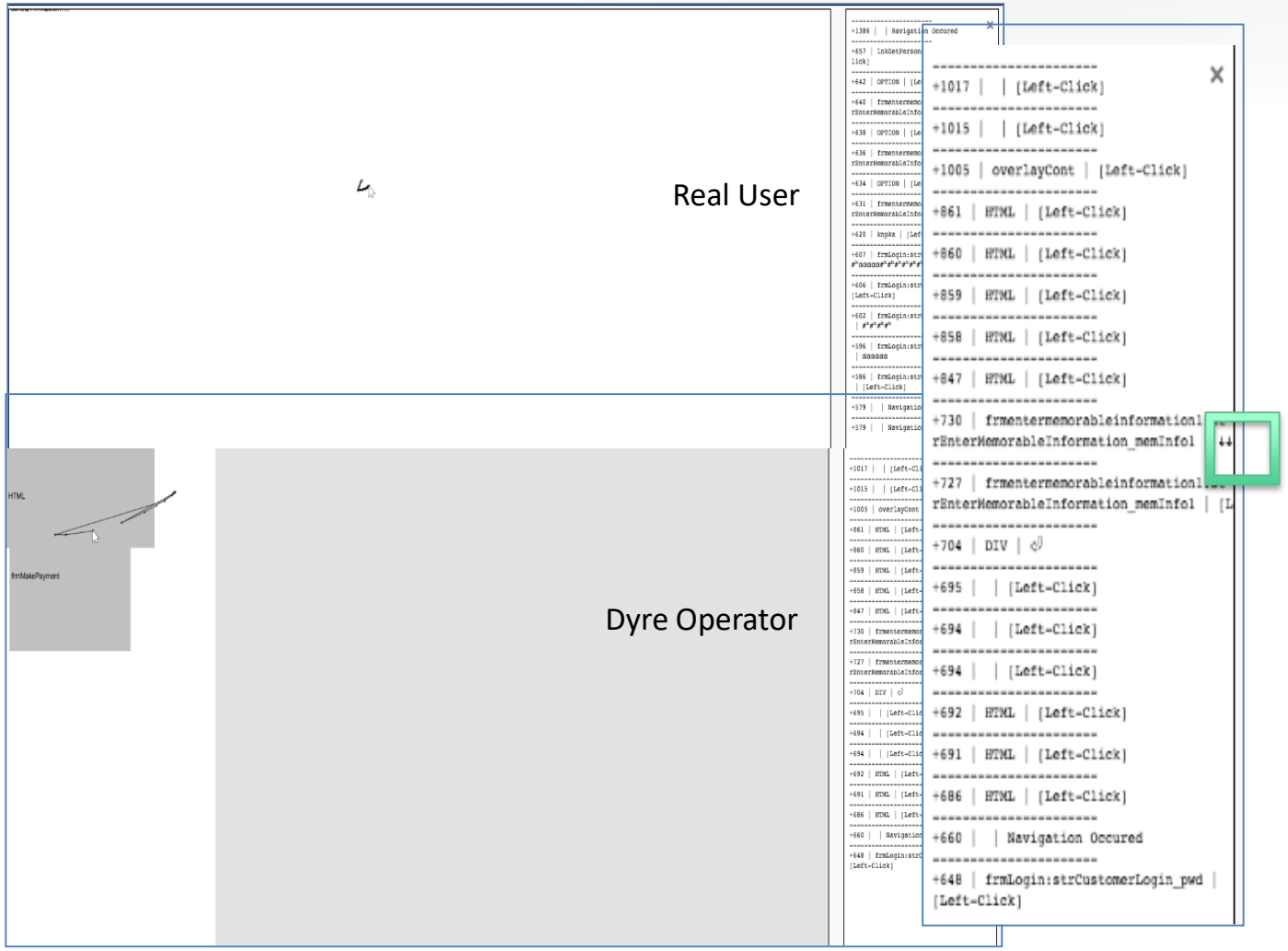
The Dyre Operator wrote 4 “letters”…

Case #1



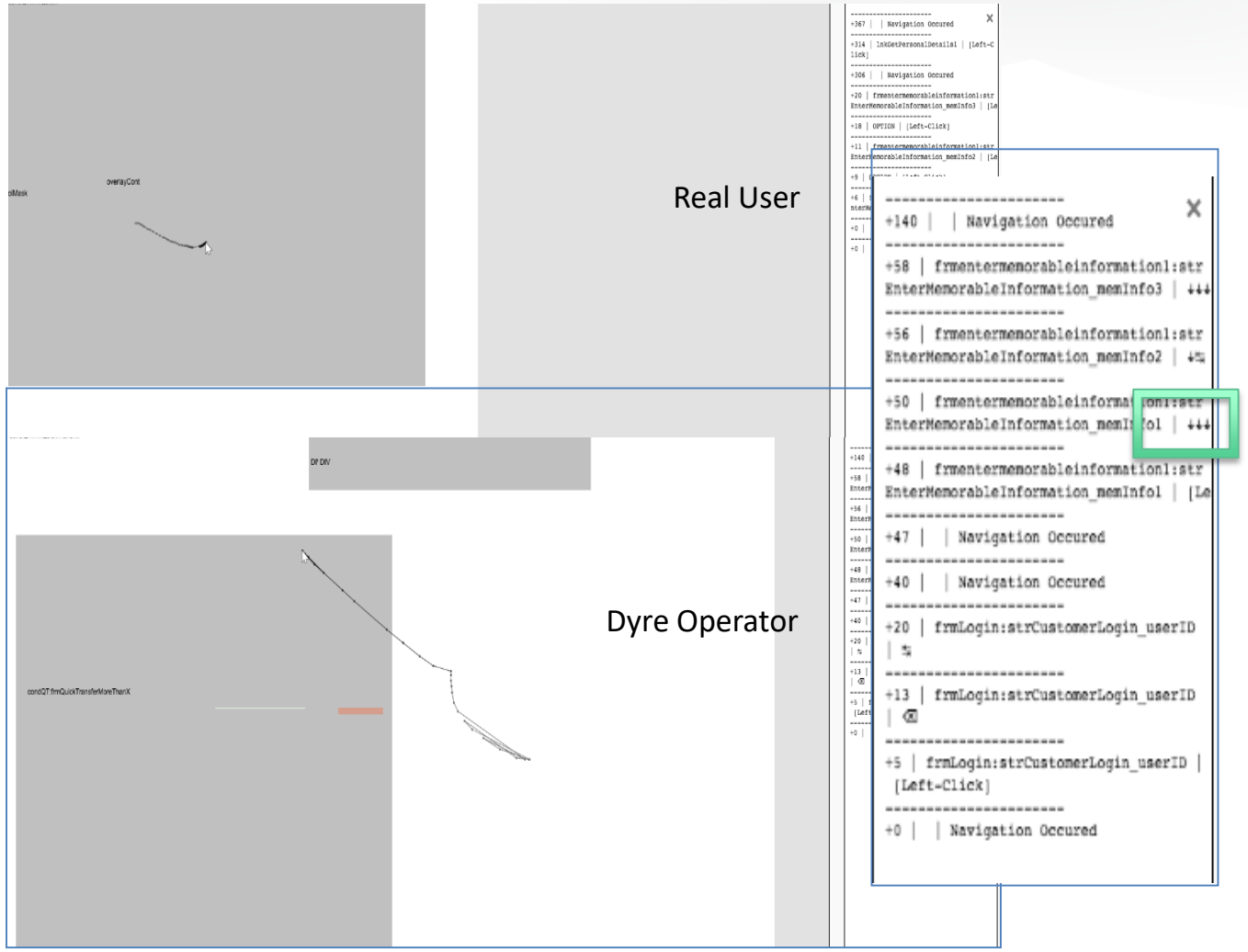
The Dyre Operator in Action…

Case #2



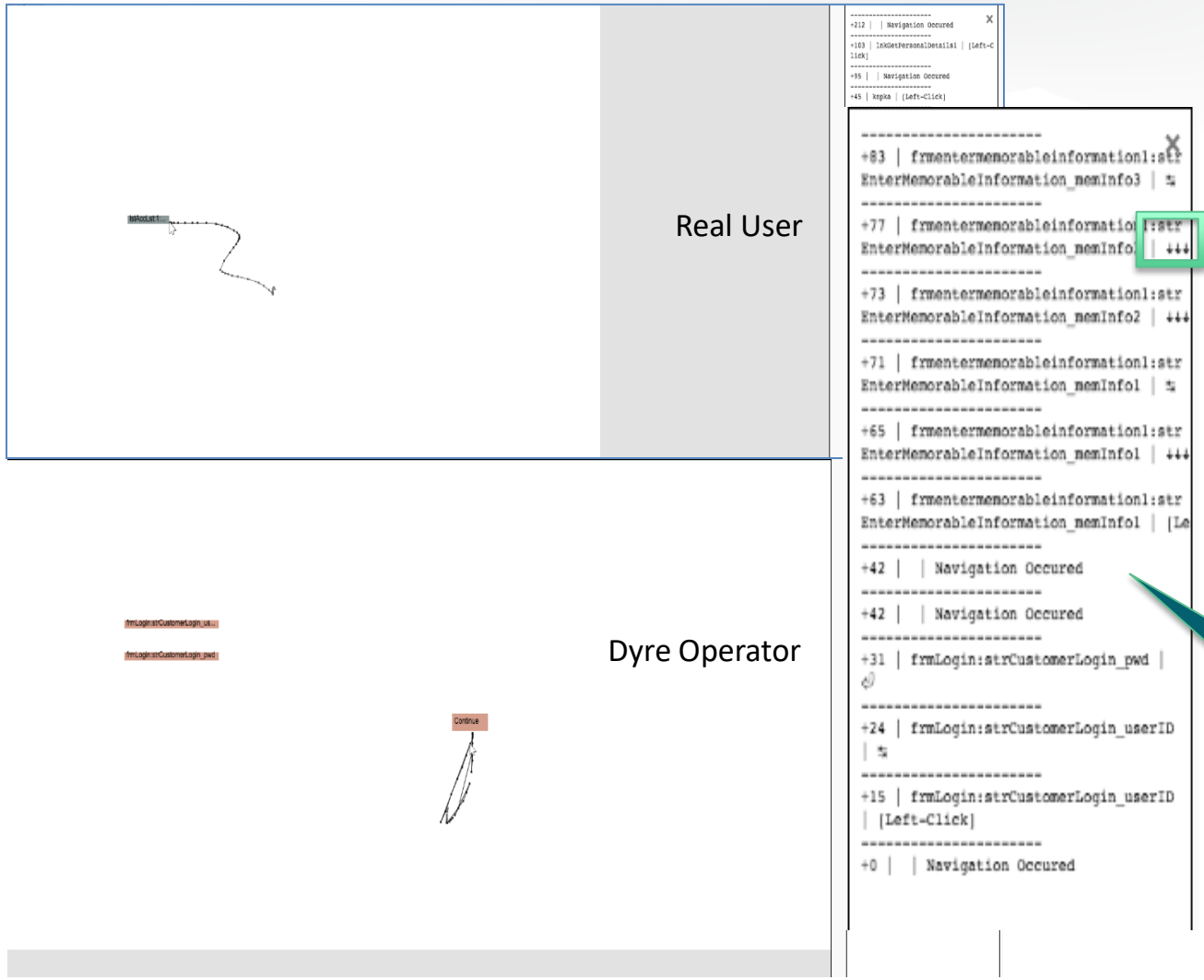
The Dyre Operator in Action…

Case #3



The Dyre Operator in Action…

Case #4



Bottom line:
Multiple rare characteristics seen in all 4 crimes…
Conclusion:
This is the same ***individual*** *criminal*

Serial Killer

Link Analysis (Device/IP): 1400 accounts accessed June 2017 to Jan 2018 – probably by same gang

Event Time	Event Type	CUSTOMER_ID	IP	Geo
2017-11-14 13:19:01.525 UTC	account_creation	1709733973	86.187.165.28	GB
2017-11-14 12:55:59.960 UTC	login	1269454403	86.187.165.28	GB
2017-11-14 12:41:22.121 UTC	login	1700554834	86.187.165.28	GB
2017-11-14 12:11:17.436 UTC	login	1707142348	86.187.165.28	GB
2017-11-13 17:21:21.717 UTC	login	1780457969	86.187.161.56	GB
2017-11-13 16:29:42.84 UTC	login	1252987799	86.187.161.56	GB
2017-11-13 16:04:16.905 UTC	login	1206510933	86.187.161.56	GB
2017-11-13 14:04:50.314 UTC	login	1812643420	86.187.161.56	GB
2017-11-13 13:58:52.897 UTC	login	1228778211	86.187.161.56	GB
2017-11-13 12:36:58.155 UTC	login	1751519369	86.187.161.56	GB
2017-11-13 11:04:05.574 UTC	login	1269231715	86.187.161.56	GB
2017-11-13 11:01:56.188 UTC	login	1347667687	86.187.161.56	GB
2017-11-13 11:00:58.557 UTC	login	1769702297	86.187.161.56	GB
2017-11-13 09:36:09.927 UTC	login	1181569171	86.187.161.56	GB
2017-11-13 09:35:32.808 UTC	login	1179587048	86.187.161.56	GB
2017-11-13 09:31:44.448 UTC	login	1785552971	86.187.161.56	GB
2017-11-13 09:17:58.984 UTC	login	1212648632	86.187.161.56	GB
2017-11-10 12:42:49.228 UTC	login	1159007220	86.187.175.83	GB
2017-11-10 12:07:26.97 UTC	login	1236327817	86.187.175.83	GB
2017-11-10 11:51:52.269 UTC	login	1230473965	86.187.175.83	GB
2017-11-10 09:17:57.990 UTC	login	1716227150	86.187.162.75	GB
2017-11-10 09:09:05.317 UTC	login	1344947061	86.187.162.75	GB
2017-11-10 09:04:29.27 UTC	login	1275063724	86.187.162.75	GB
2017-11-10 09:02:28.869 UTC	login	1667391770	86.187.162.75	GB
2017-11-10 09:01:37.505 UTC	login	1291591795	86.187.162.75	GB
2017-11-10 09:00:46.356 UTC	login	1162362599	86.187.162.75	GB
2017-11-10 08:58:36.699 UTC	login	1866977182	86.187.162.75	GB
2017-11-10 08:57:11.678 UTC	login	1629842597	86.187.162.75	GB
2017-11-10 08:56:27.15 UTC	login	1744623964	86.187.162.75	GB

Criminal Patterns

Device

- New Device for User: 100%
- Device associated with Multiple Users: 100%

Behavior

85% of sessions had a unique combination:

- Use of Arrows to navigate up/down
- Use of both keypad and numpad to type numbers
- Use of double-click

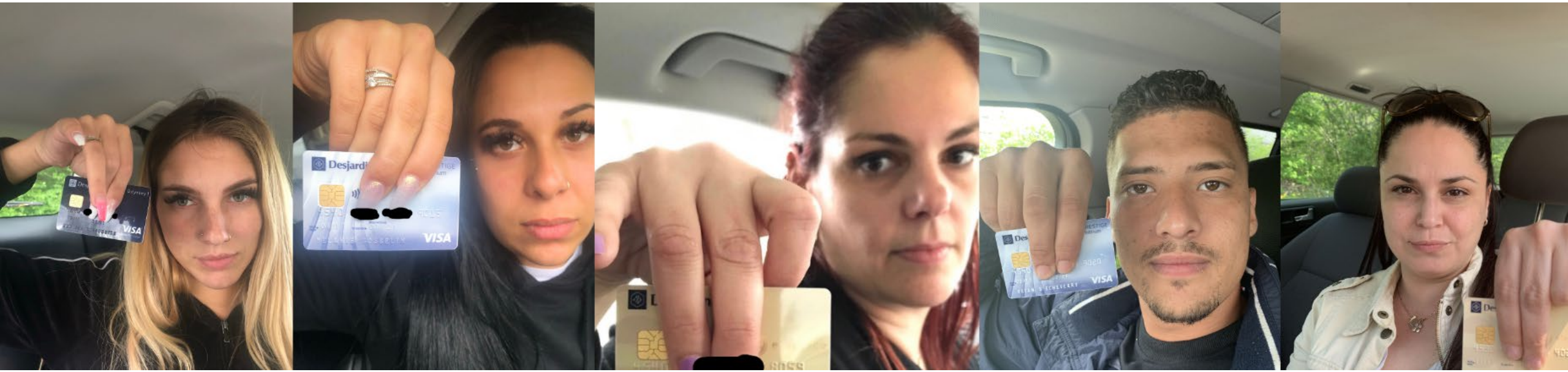
- Looks like vast majority of access is done **by a single person**.

⇒ **Tracking criminal even if they move to totally new devices/IPs**
⇒ **Identifying immediate change in genuine user's behavior**

The Bitcoin Crew

Source: Simplex.com

- 5-10 Bitcoin orders per week during May-August ‘19
- All made from newly created bank accounts, mostly a Top 5 Canadian issuer
- Young buyers, passing verification easily



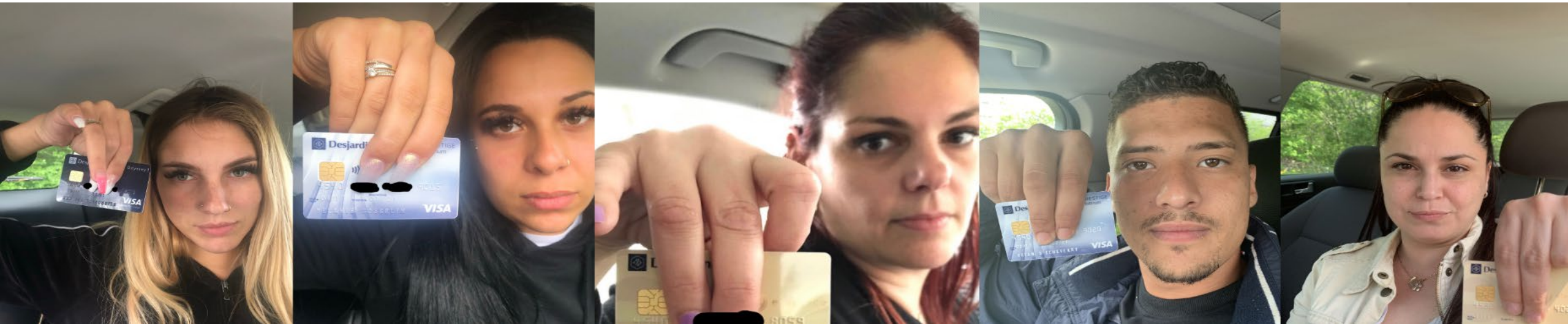
The Bitcoin Crew

- Deeper analysis shows that *what seemed like independent shoppers had too much in common to be a coincidence*
- Link analysis on meta data shows cluster of **same GPS location for multiple images**

Photo IDs
were all taken
while placed on
same table…

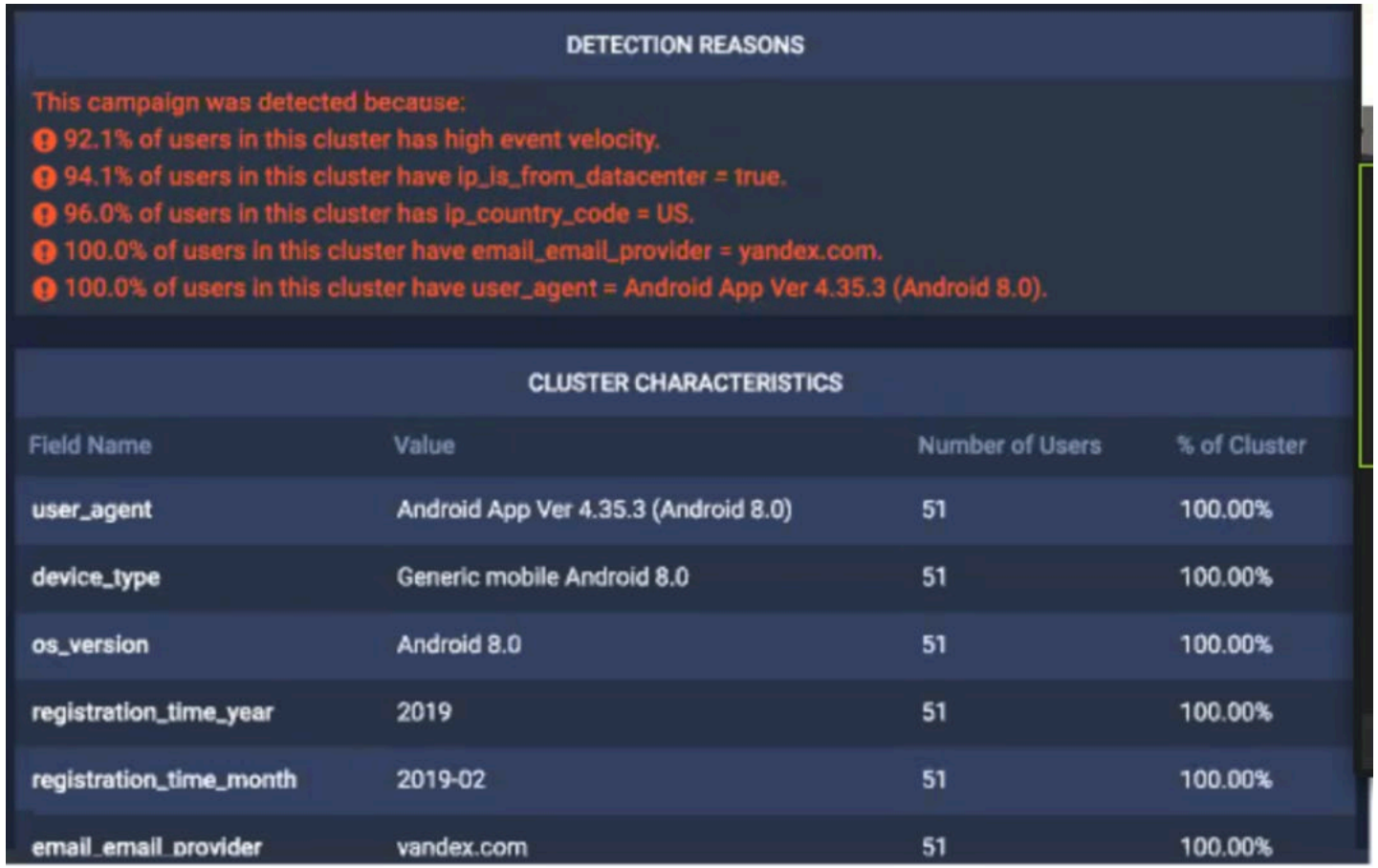
The Bitcoin Crew

- Also, all the selfie verification photos are taken in the *same car*
- Forensics exposes chain of accounts opened online for money laundering and fraud
- Reviewing a single account does not raise suspicion; however, deep link analysis reveals that something isn't quite right



For more info: gilits@simplex.com

Analyzing Resources: Email, Device



DETECTION REASONS

This campaign was detected because:

- 92.1% of users in this cluster has high event velocity.
- 94.1% of users in this cluster have ip_is_from_datacenter = true.
- 96.0% of users in this cluster has ip_country_code = US.
- 100.0% of users in this cluster have email_email_provider = yandex.com.
- 100.0% of users in this cluster have user_agent = Android App Ver 4.35.3 (Android 8.0).

CLUSTER CHARACTERISTICS

Field Name	Value	Number of Users	% of Cluster
user_agent	Android App Ver 4.35.3 (Android 8.0)	51	100.00%
device_type	Generic mobile Android 8.0	51	100.00%
os_version	Android 8.0	51	100.00%
registration_time_year	2019	51	100.00%
registration_time_month	2019-02	51	100.00%
email_email_provider	vandex.com	51	100.00%

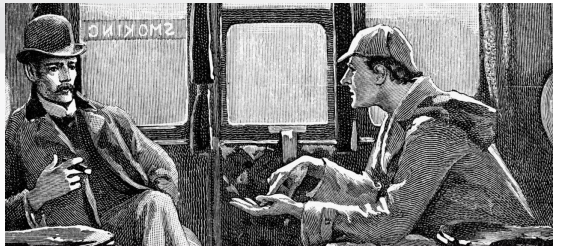
Analyzing Resources: Phone account

Synthetic Identities

Phone Intelligence

- Line Tenure: June 14, 2010
- Line Type: MVNO (Tracfone)
- Service Provider Tenure: May 16, 2014
- SIM Tenure:
- Device Tenure:
- **Call Activity: No activity**
- **Login Activity: No activity**

Synthetic ID Cyber Gang Behavior



00:56	Text Box › phone-…	› Typing	111
00:56	2 seconds of inactivity		
00:59	Text Box › phone-…	› Typing	111
01:02	Text Box › phone-…	› Typing	11
01:04	Text Box › phone-…	› Typing	11

01:10	Text Box › ssn-72	› Typing	111
01:12	Text Box › ssn-72	› Typing	11
01:14	Text Box › ssn-72	› Typing	11
01:16	Text Box › ssn-72	› Typing	11

09:03	Element › income…	› Left click
09:04	Element › dropdo…	› Left click

TOTAL ANNUAL INCOME ⓘ

$

NON-TAXABLE ANNUAL INCOME (optional) ⓘ

$

INCOME SOURCE

- Choose One
- Employed
- Retired
- Self-Employed
- Unemployed
- Military
- Business Owner
- Other

Phone, SSN typed
in a specific rhythm and grouping

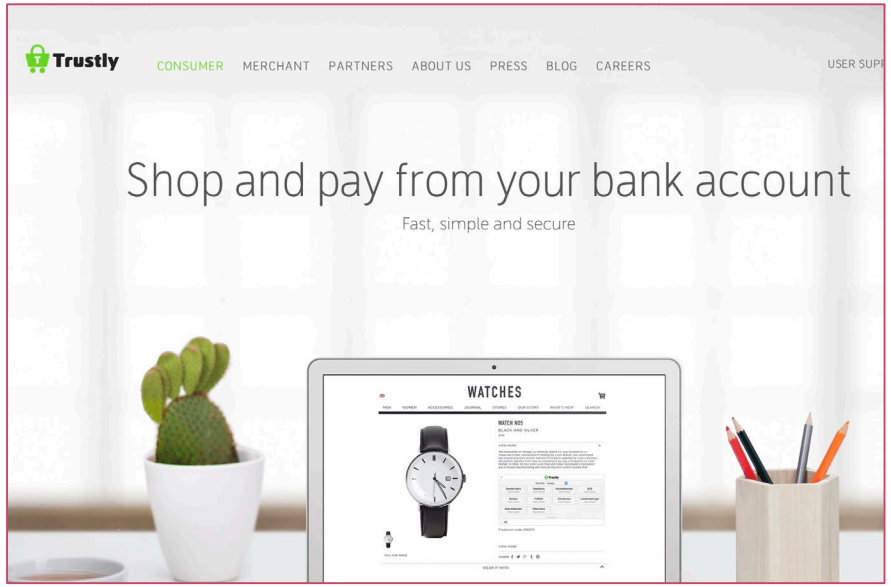
High familiarity with Site:
Super fast interaction with Annual Income, Income Source

Profiling Bots

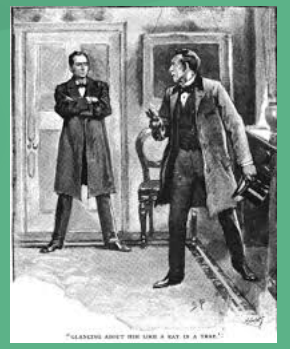
Bot moves through session, makes payment on behalf of user

Bot profiling elements:

- IP and geo-location
- Server ownership
- Device fingerprint
- Device profile (OS, browser, version, user agent analysis)
- Spoofing analysis
- Blacklists and Link analysis
- Consortium checks
- Event sequence
- Behavioral fingerprints

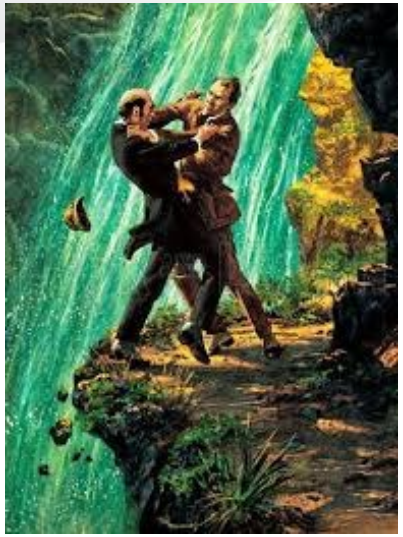


RSAConference2020



Closing Notes

Whodunit?



- In fraud detection, the goal is to prevent crimes in the most cost-effective way possible, not go after the criminal
- In criminology, profiling behavior is useful in actually catching the bad guys
- Holmes' technique relied heavily on human intuition and manual searching
- AI-powered criminal profiling can…
  - Prevent future attacks
  - Track threat actors across enterprises using consortium databases
  - Build strong cases against individual hackers within an actor group
  - Assist law enforcement agencies, financial crime investigators, cyber intelligence researchers
- Can the data be shared across enterprises?
- Can criminals expect privacy?

Summary

- 1891: A Case of Identity – generic, individual criminal profiling
- 2020: AI can look at hundreds of device, location, behavior parameters and create unique digital criminal profiles
- Diversity of data: device/network, image/documentation, behavior/cognitive
  - Criminals can change resources
  - It's more difficult to change behaviors
  - It's also more difficult to understand and neutralize *all* that is being monitored

Apply What You Have Learned Today

- Next week you should:
  - Map the data elements and link processes in your online fraud organization
- In the first three months following this presentation you should:
  - Research your historic data for any linkage and profiling elements that could point to specific criminals
  - Benchmark your linking and criminal profiling capability by talking to colleagues in the industry / online fraud analysts
  - Assess whether your current layers of visibility and analysis tools are adequate for your risk management needs
  - Identify any gaps in your criminal profiling strategy (data layers and tools)
  - Check whether you can link to consortium-based criminal profiling databases
- Within six months you should:
  - Build a plan for closing those gaps, if there's a business case to do it

RSAConference2020

Q&A

LinkedIn: Uri Rivner
LinkedIn: Erin Englund