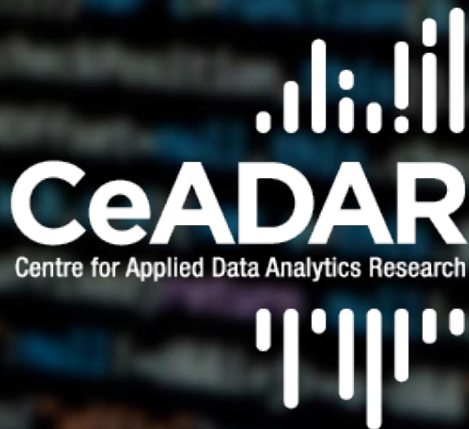


Predict

DATA ANALYTICS FOR DIGITAL FORENSICS AND CYBERSECURITY

DR MARK SCANLON



UCD Forensics and
Security Research Group

Agenda

2

- ▶ Challenges Facing Digital Forensics
- ▶ Challenges Facing Cybersecurity
- ▶ Where Data Analytics Can Help
 - ▶ [Sample Current Research Projects]



Data Analytics for Digital Forensics



Digital Forensic Challenges

- ▶ The consistency and correlation problem
 - ▶ Results from the fact that existing tools are designed to find fragments of evidence, but not to otherwise assist in investigations.
- ▶ The unified time lining problem
 - ▶ Multiple sources present different time zone references, timestamp interpretations, clock skew/drift issues, and the syntax aspects involved in generating a unified timeline.
- ▶ The diversity problem
 - ▶ Results from ever-increasing volumes of data
 - ▶ Lack of standard techniques to examine and analyse the increasing numbers and types of sources, which bring a plurality of operating systems, file formats, etc.

Digital Forensics: The Volume Challenge

- ▶ The number of cases whereby digital evidence is deemed pertinent is ever increasing.
- ▶ An increase in the number of devices that are seized for analysis per case.
- ▶ The volume of potentially evidence-rich data stored on each item seized is also increasing.



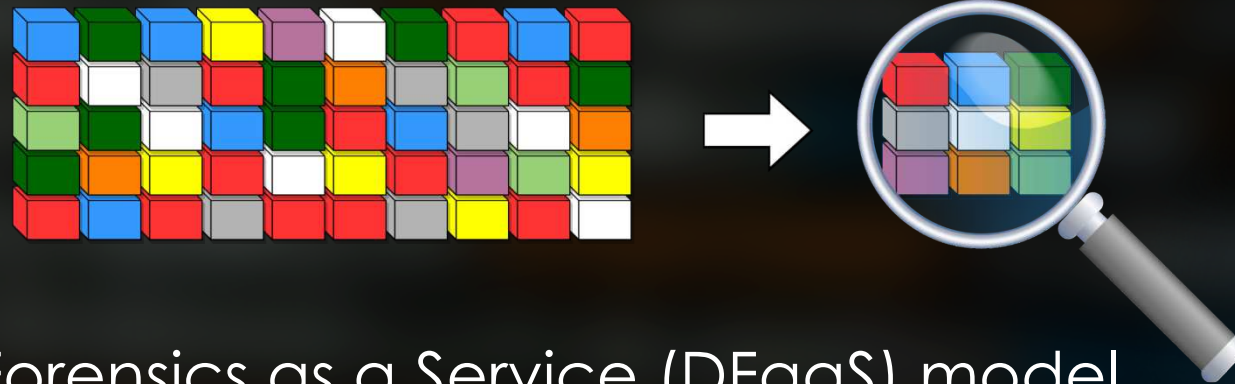
Digital Forensic Backlog

- ▶ Backlogs have become commonplace in recent years
 - ▶ Commonly exceeds 18 months
 - ▶ Often exceeds 2 years
- ▶ According to a report by An Garda Síochána, delays of up to four years
 - ▶ "Seriously impacted on the timeliness of criminal investigations" in recent years.
 - ▶ In some cases, these delays have resulted in prosecutions being dismissed in courts.



One Solution: Deduplication

7



- ▶ Digital Forensics as a Service (DFaaS) model
- ▶ Centralisation of digital forensic processing
- ▶ Elimination of duplicated effort from the typical forensic process:
 - ▶ Eliminate duplicated acquisition
 - ▶ Eliminate duplicated storage
 - ▶ Eliminate duplicated analysis and processing



Intelligent Automated Evidence Processing

- ▶ Research towards automated evidence processing
- ▶ Leverages centralised record of evidence analysis
 - ▶ Learns what makes evidence pertinent/non-pertinent
- ▶ Photographic and Video Human/Object Identification
 - ▶ Biometric estimation; ageing, height, weight, etc.
 - ▶ Location determination



Data Analytics for Cybersecurity



Information Overload

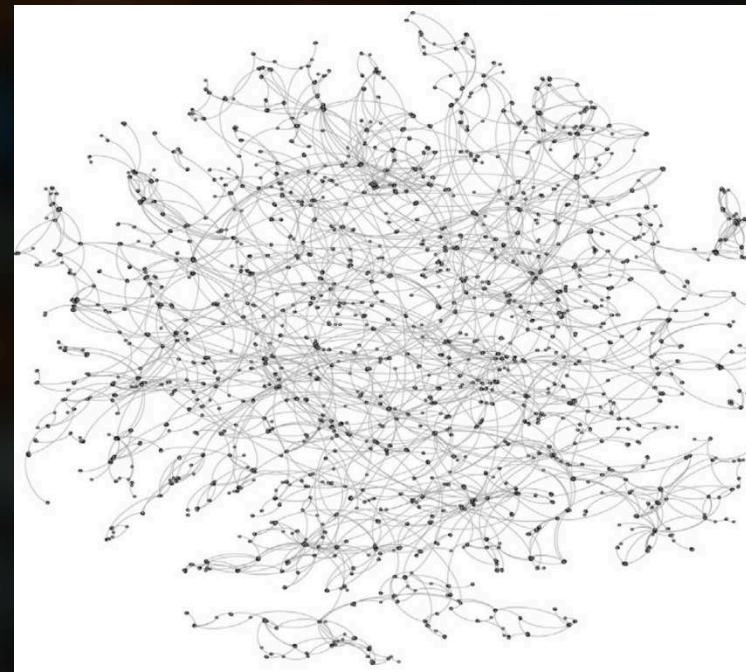
- ▶ Information overload facing cybersecurity professionals
 - ▶ False positive alert rate is too high
- ▶ Attack Sophistication
 - ▶ Difficult to identify anomalies
- ▶ Data Analytics can enable behavioural anomaly detection
 - ▶ Similar premise to what antivirus systems followed to combat polymorphic and metamorphic malware



Network Behavioural Analysis

- ▶ Build a baseline of each node's activity on the network
- ▶ Categorise nodes based on their normal behaviour
- ▶ Alert when a deviation from this norm is identified

- ▶ Intrusion Detection
- ▶ Botnet Investigation



User Behavioural Analysis

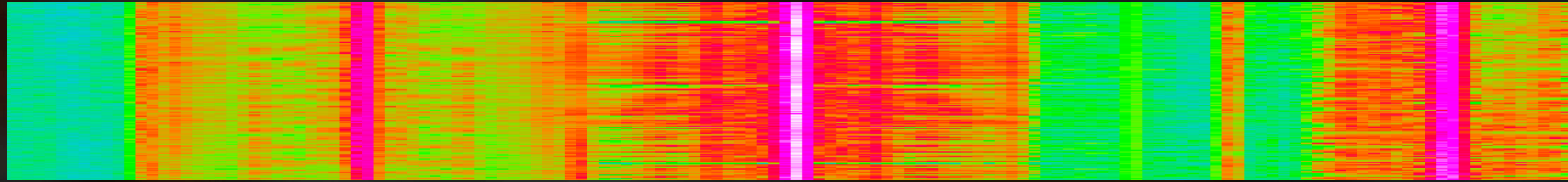
12

- ▶ Effectively the same idea! Includes:
 - ▶ Network Traffic
 - ▶ Device Utilisation
 - ▶ Correlation between devices
- ▶ Can identify specific users in multiuser environments



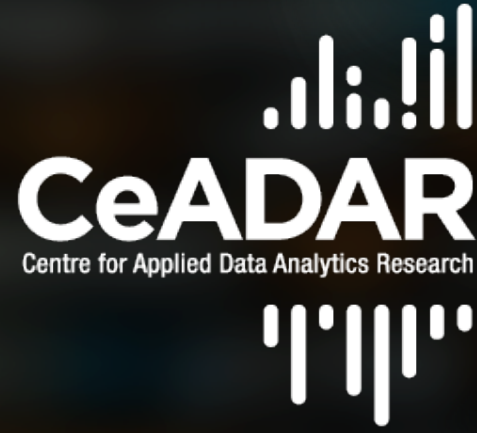
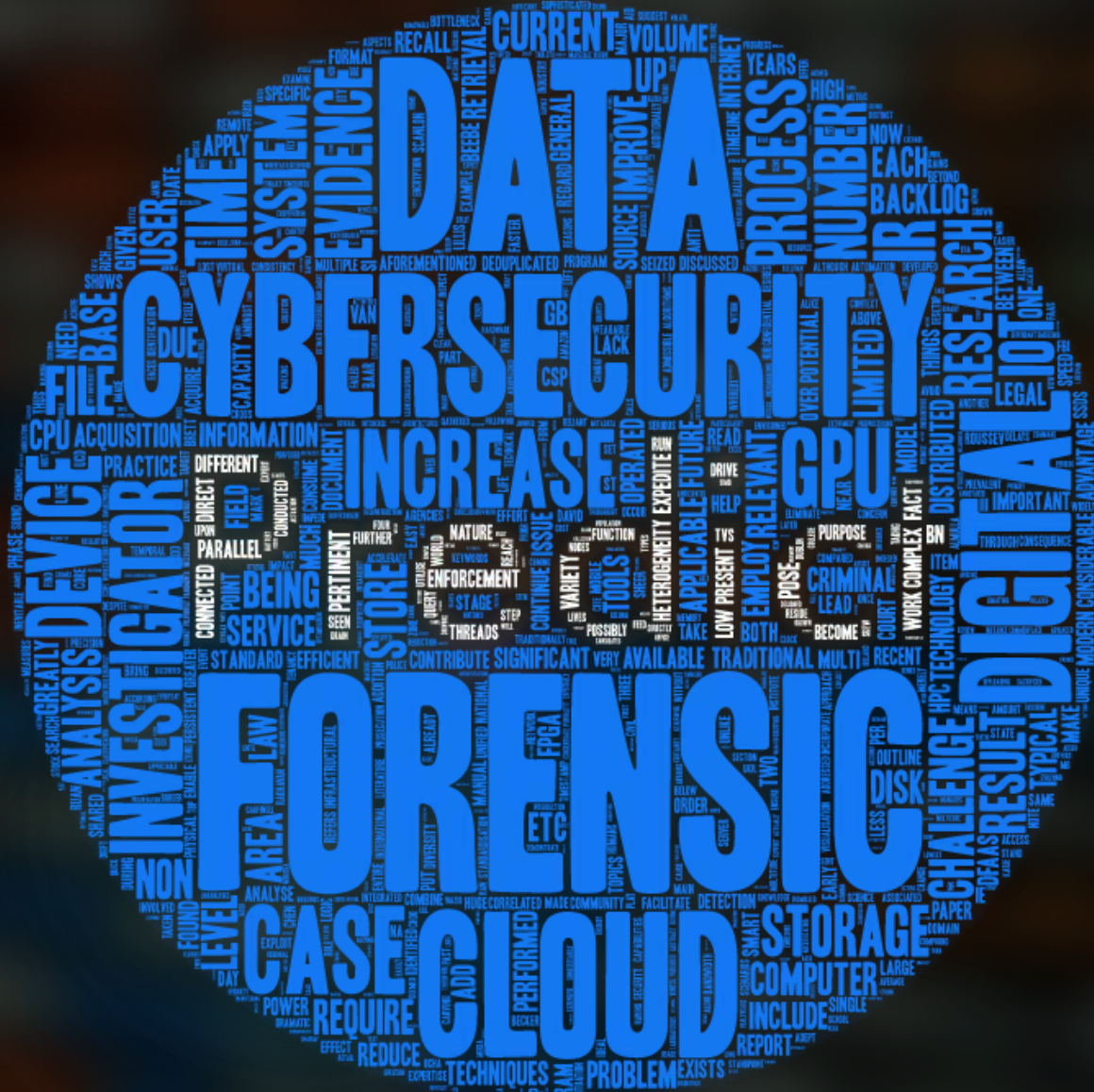
Data Analytics to Break Encryption

13



- ▶ Software defined radio to capture leaking CPU electromagnetic radiation
- ▶ Becomes a Big Data/Data Analytics problem





MARK.SCANLON@UCD.IE



WWW.FORENSICSANDSECURITY.COM



@MRKSCN / @FORSECRESEARCH



UCD Forensics and Security Research Group