LOOKINGGLASS

Assessing Targeted Attacks in Incident Response Threat Correlation

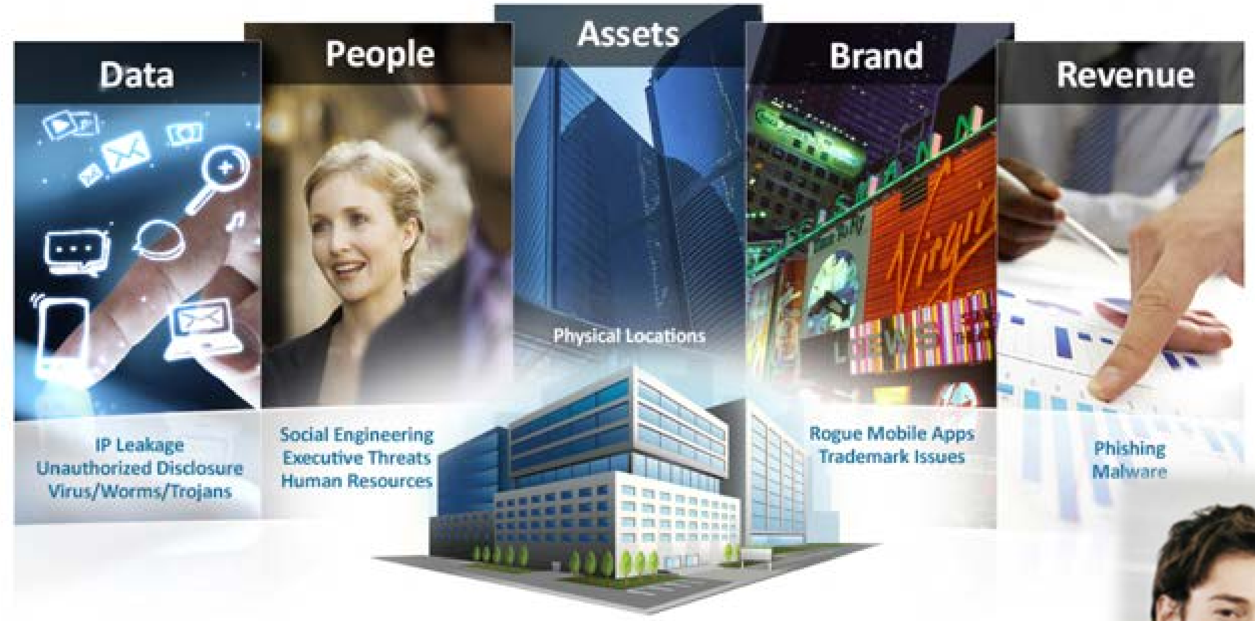
Jan 2017

www.lookingglasscyber.com

PRESENTER:

Allan Thomson, CTO
Dr Jamison Day, Principal Data Scientist

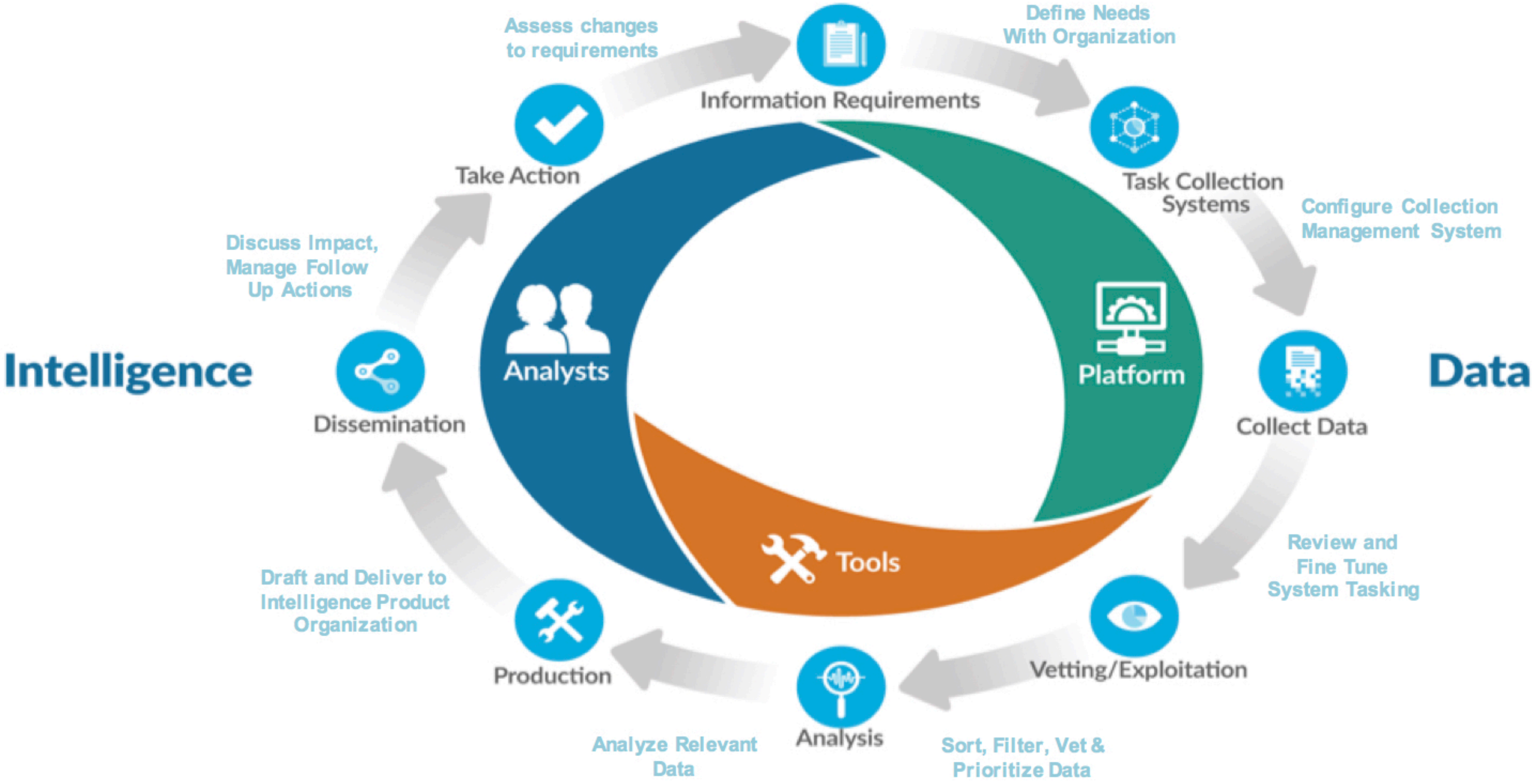
What…threats are targeting?



Who…is impacted by targeted threats?

Why automation is critical to success…

Security data is not intelligence. Intelligence is data that has been refined, analyzed or processed such that it is ***relevant***, ***actionable*** and ***valuable***.

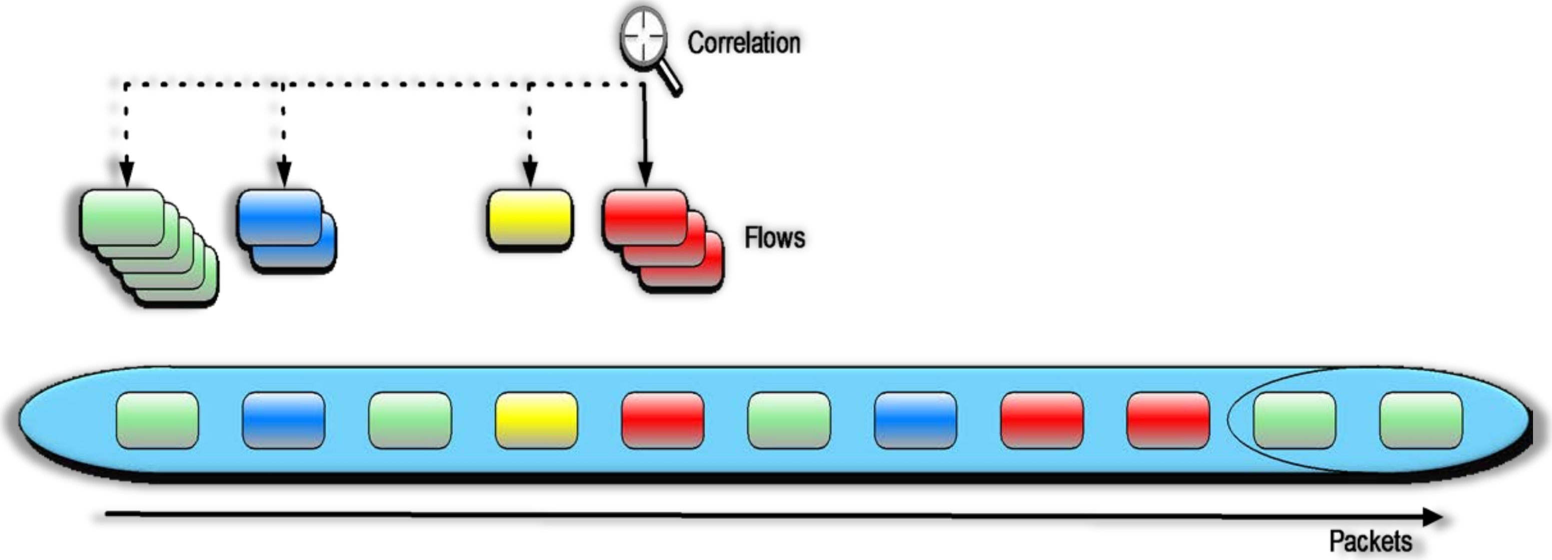

Assess changes to requirements
Information Requirements
Define Needs With Organization
Take Action
Task Collection Systems
Configure Collection Management System
Discuss Impact, Manage Follow Up Actions
Intelligence
Dissemination
Analysts
Platform
Collect Data
Data
Tools
Review and Fine Tune System Tasking
Draft and Deliver to Intelligence Product Organization
Production
Vetting/Exploitation
Analysis
Analyze Relevant Data
Sort, Filter, Vet & Prioritize Data

Choosing Threat Intelligence Feeds

- Ensure **rich context:** Vulnerabilities, TTPs, Indicators, Actors
- Ensure **broad coverage**: Surface web, Dark web, Social media, Human & Automated
- Ensure **Timely**: Real-time is important; Hourly and frequent updates

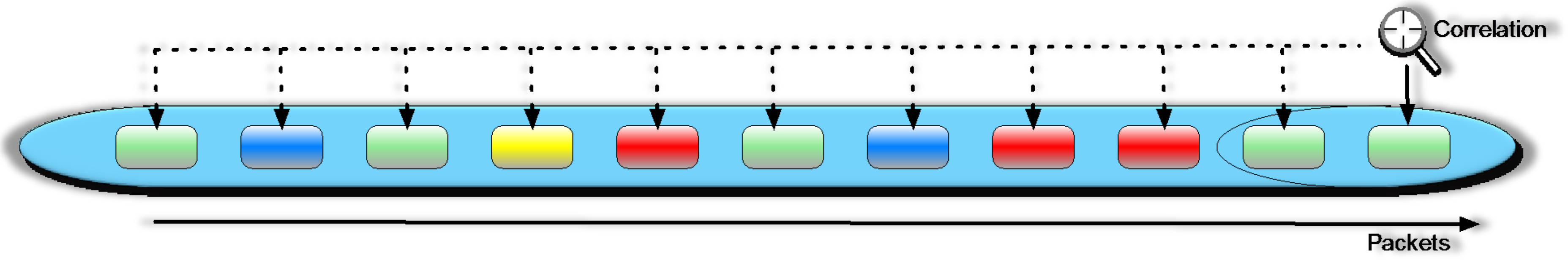
Choosing Threat Correlation Telemetry - Flows

- Provides network session context
- Typically done as a non-inline correlation process to enable identification of behaviors and patterns over time
- Often uses automated techniques defined later in the presentation



- **Recommendations**
  - Should include both northbound and east-west traffic flows to detect external and cross-domain traffic behaviors
  - **If possible include payload extraction and correlation across packets**
  - IPFIX (Netflow v10) supports much context beyond traditional 5-tuple
  - Gather unsampled flow rather than sampled flow especially if you are doing behavioral analysis

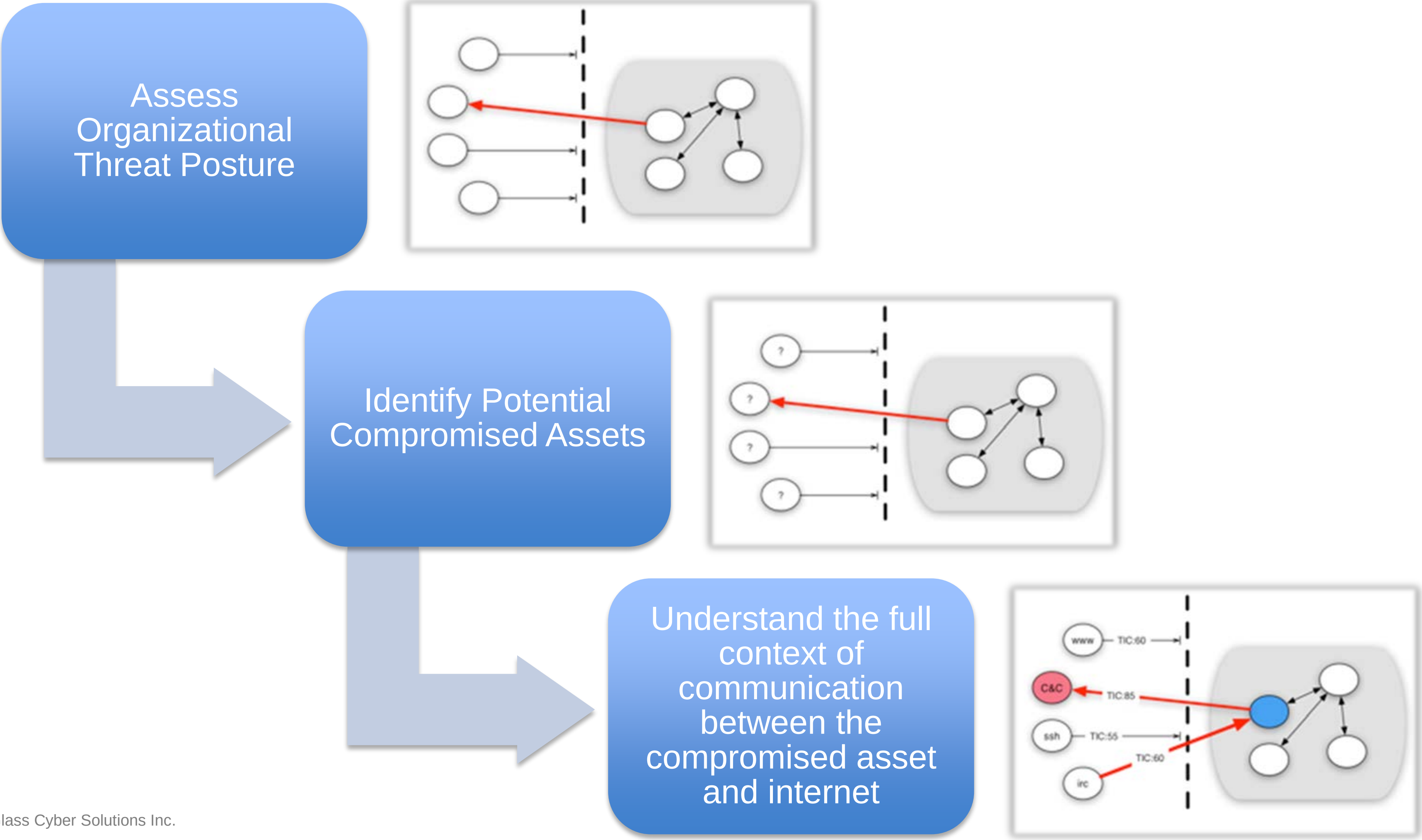
Choosing Threat Correlation Telemetry - Packets



- Provides ability to identify content in every packet that matches specific patterns
- Typically network inspection devices are programmed with rules to identify regex, signatures and payload that may be malicious

- **Recommendations**
  - Must focus on inline data rate inspection
  - Ability to correlate at line rate

Workflow Supporting Correlation Steps: 1 of 2



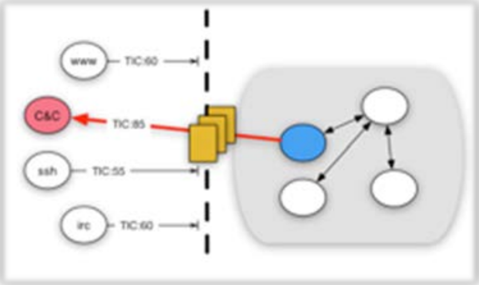
Assess Organizational Threat Posture

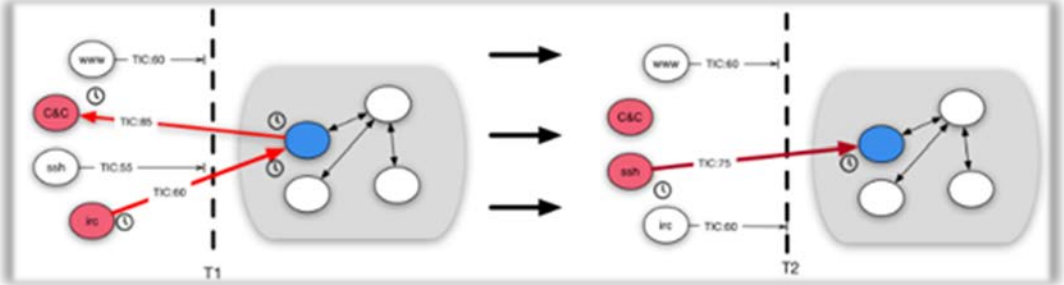
Identify Potential Compromised Assets

Understand the full context of communication between the compromised asset and internet

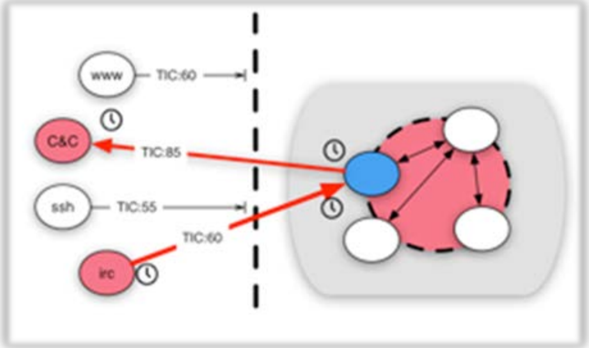
Workflow Supporting Correlation Steps 2 of 2

Identify any data exfiltration or impact on compromised asset

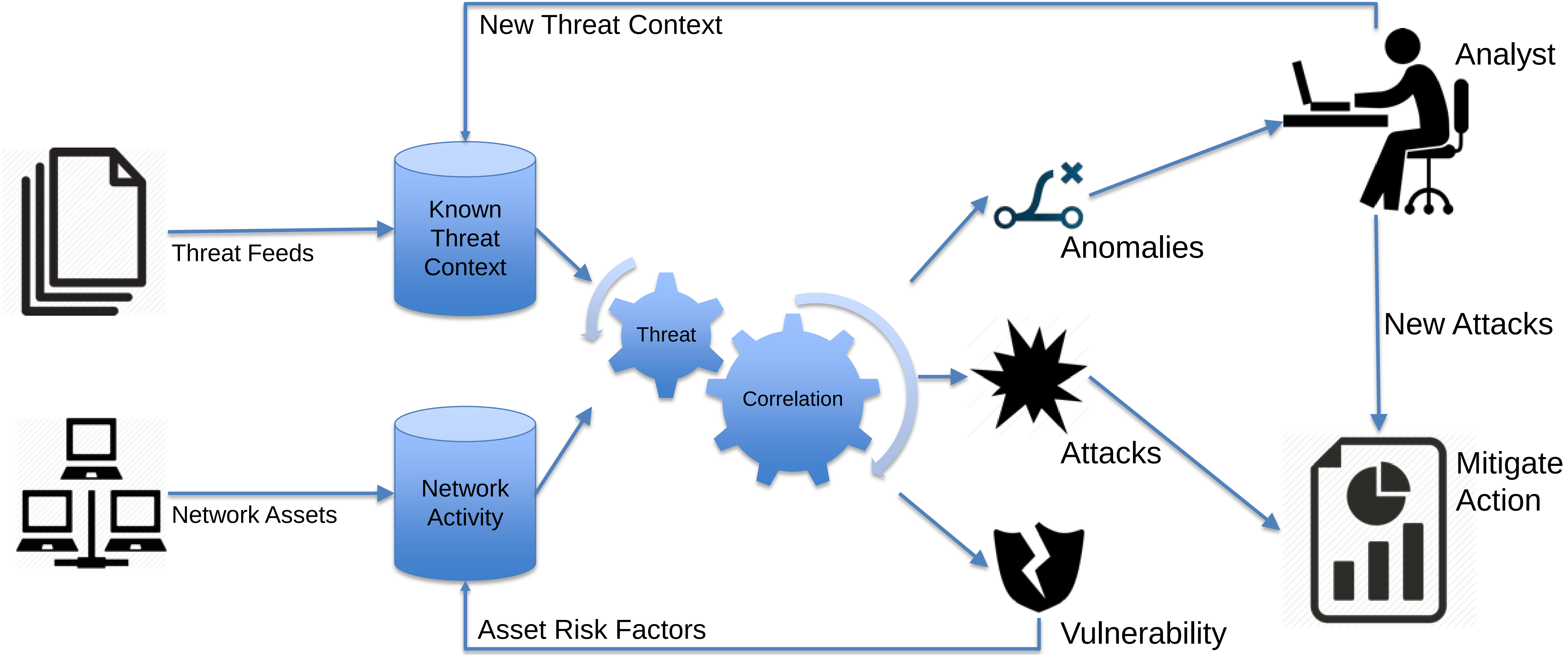




Identify the spread of any threat within the perimeter



Threat Correlation in Your Cyber Security Ecosystem



LOOKINGGLASS

Threat Correlation Approaches

Threat Correlation Approaches

Threat Correlation
Identifies new cyber threat insights by associating events from multiple data sources

≠

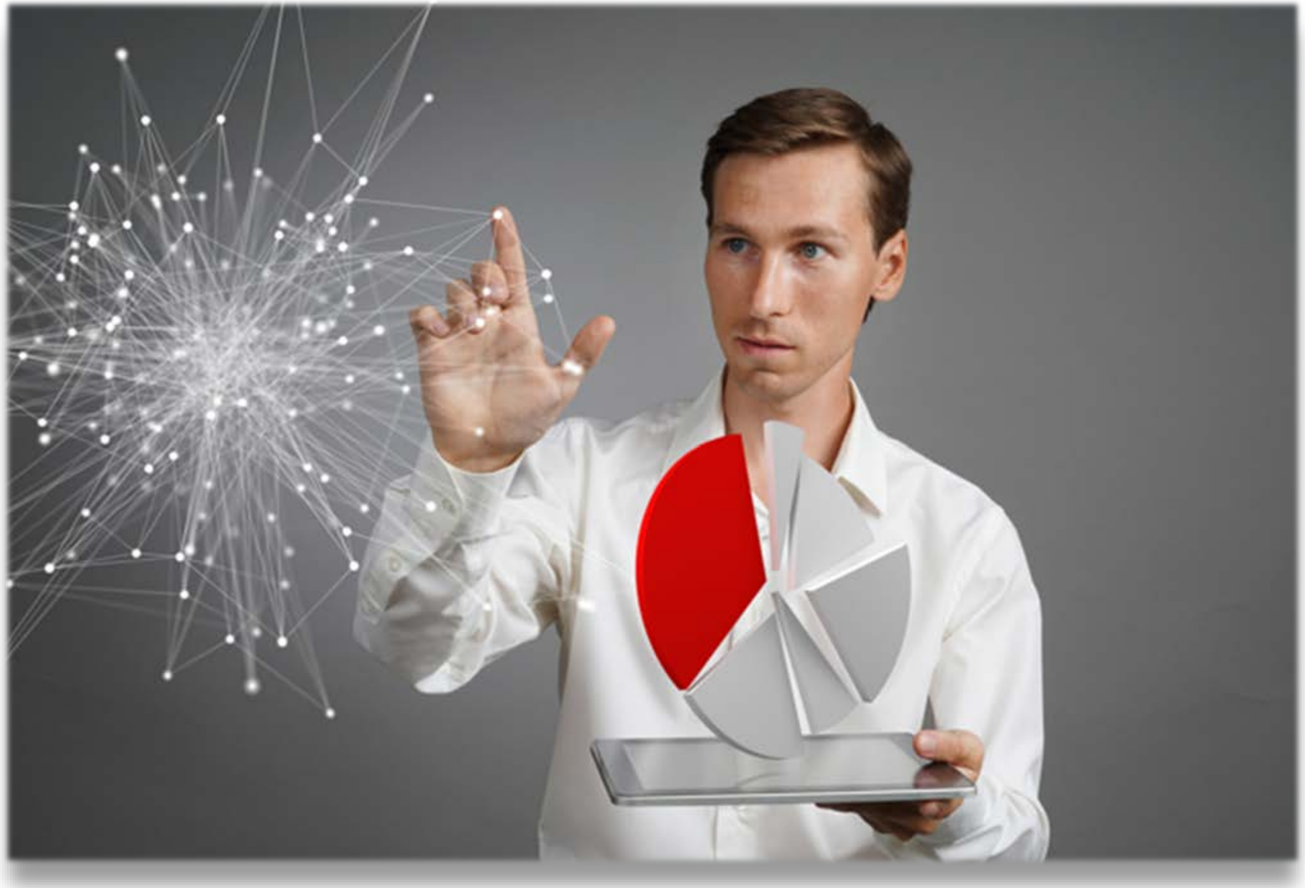
Statistical Correlation
Measures the similarity in fluctuations between two variables.

Approaches
Manual Threat Correlation
Field Comparison
Rules-Based Matching
Fuzzy Matching
Machine Learning

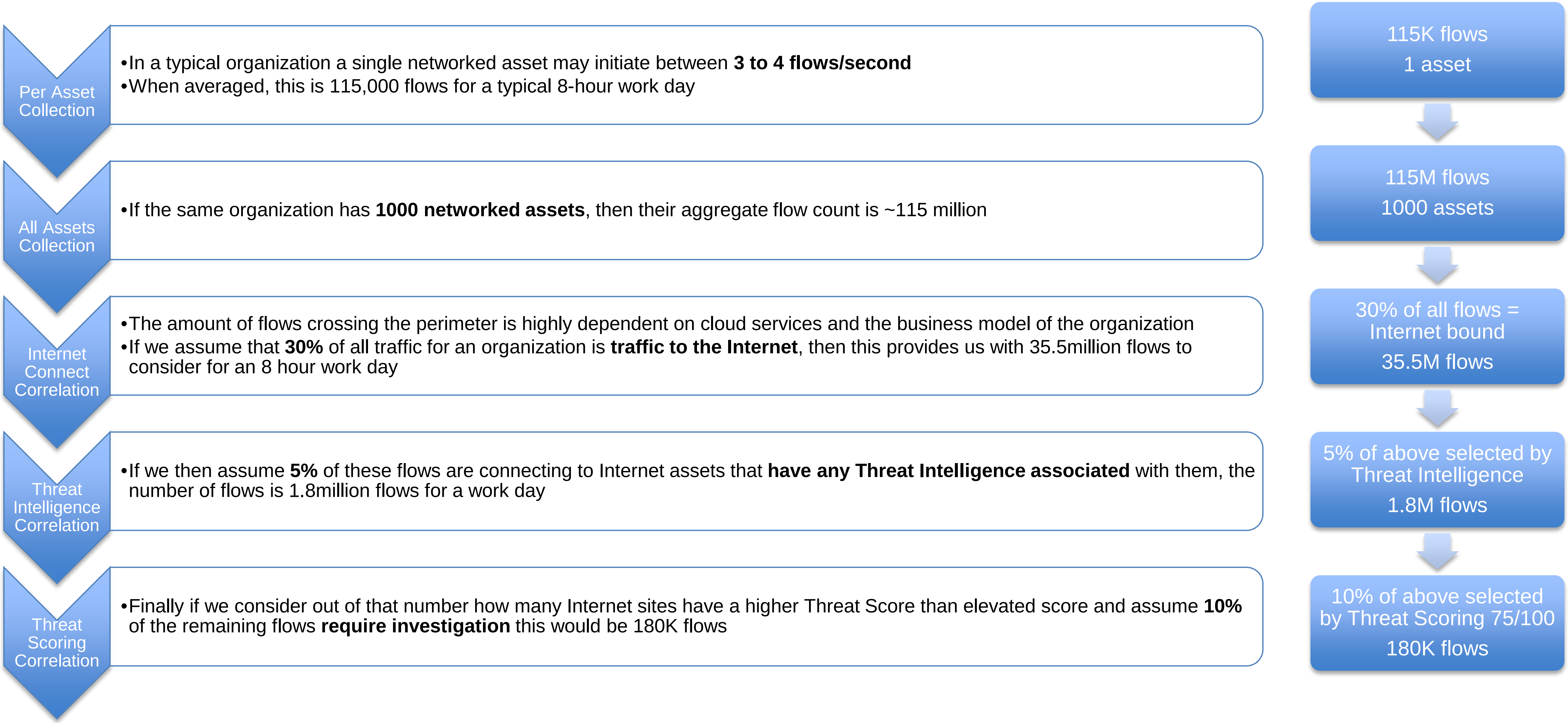
Manual Threat Correlation

- Human comparison of data from multiple sources to identify threat-related events

- **Advantages**
  - Pattern Recognition
  - Language Abilities
  - Creative Thinking
  - Flexible Inference
  - Intuition/Guessing
- **Drawbacks**
  - Slow step-by-step instruction execution
  - Imprecise, Unpredictable, Reproducibility Issues
  - Bias/Prejudice



Real World Example: Data Processing Reduction



Per Asset Collection

- In a typical organization a single networked asset may initiate between **3 to 4 flows/second**
- When averaged, this is 115,000 flows for a typical 8-hour work day

All Assets Collection

- If the same organization has **1000 networked assets**, then their aggregate flow count is ~115 million

Internet Connect Correlation

- The amount of flows crossing the perimeter is highly dependent on cloud services and the business model of the organization
- If we assume that **30%** of all traffic for an organization is **traffic to the Internet**, then this provides us with 35.5million flows to consider for an 8 hour work day

Threat Intelligence Correlation

- If we then assume **5%** of these flows are connecting to Internet assets that **have any Threat Intelligence associated** with them, the number of flows is 1.8million flows for a work day

Threat Scoring Correlation

- Finally if we consider out of that number how many Internet sites have a higher Threat Score than elevated score and assume **10%** of the remaining flows **require investigation** this would be 180K flows

115K flows
1 asset

115M flows
1000 assets

30% of all flows = Internet bound
35.5M flows

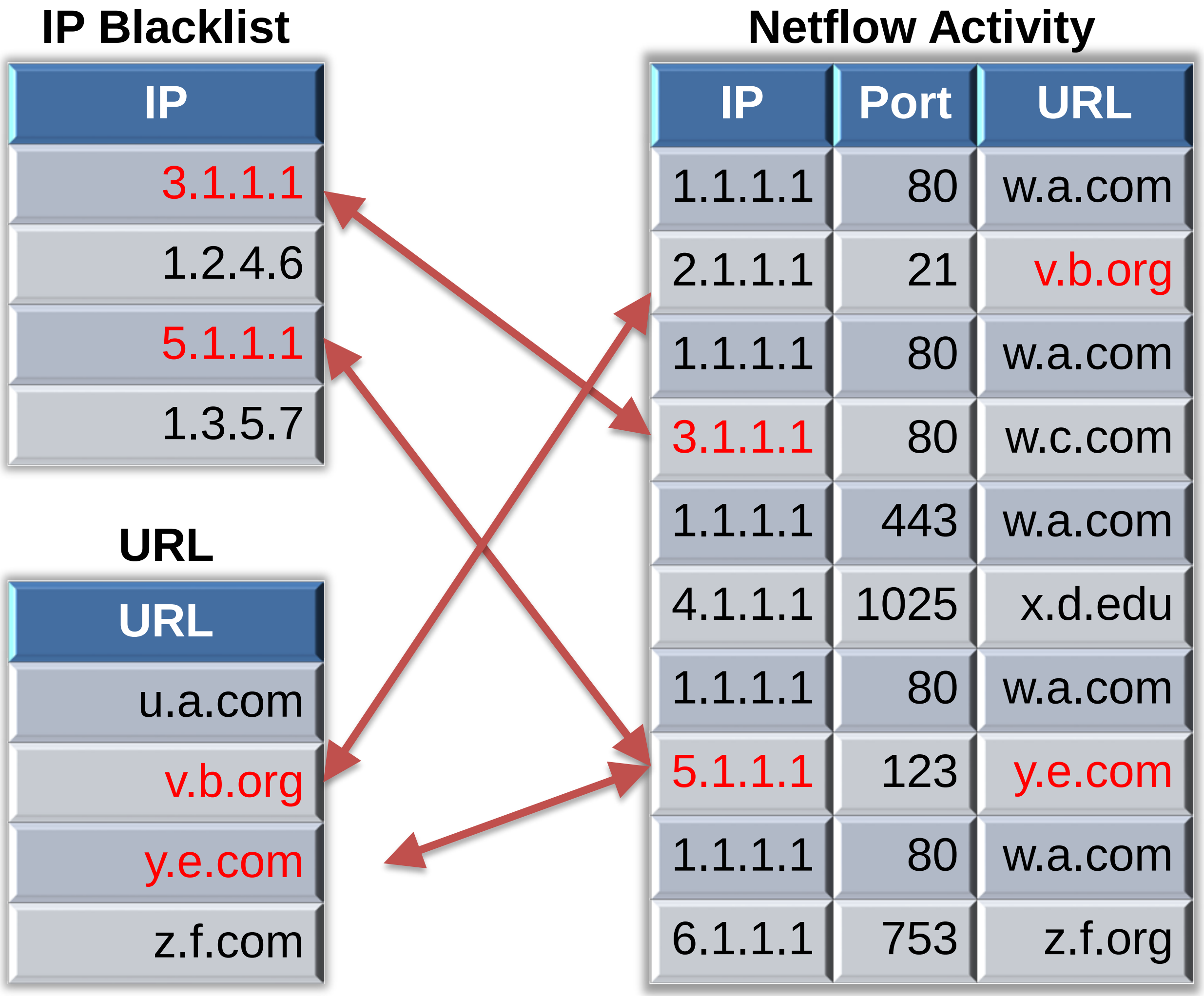
5% of above selected by Threat Intelligence
1.8M flows

10% of above selected by Threat Scoring 75/100
180K flows

Field Comparison

Identical features seen in fields of different datasets

- **Advantages**
  - Simple to Implement & Update
  - Very Fast
  - Very Scalable
- **Drawbacks**
  - Naïve Approach
  - Misses Sophisticated Attacks



IP Blacklist

IP
3.1.1.1
1.2.4.6
5.1.1.1
1.3.5.7

URL

URL
u.a.com
v.b.org
y.e.com
z.f.com

Netflow Activity

IP	Port	URL
1.1.1.1	80	w.a.com
2.1.1.1	21	v.b.org
1.1.1.1	80	w.a.com
3.1.1.1	80	w.c.com
1.1.1.1	443	w.a.com
4.1.1.1	1025	x.d.edu
1.1.1.1	80	w.a.com
5.1.1.1	123	y.e.com
1.1.1.1	80	w.a.com
6.1.1.1	753	z.f.org

Rules-Based Matching

Specific features seen in combination across datasets

- **Advantages**
  - Identifies complex interactions
  - Scalable
- **Drawbacks**
  - Requires managing a large number of pre-defined rules
  - New threats require new rules

Threat Intelligence Feed Records & Signatures

IP	Port	Protocol	Regex
1.1.1.1	53	UDP	^\w+@[a-zA-Z_]+?\.[a-zA-Z]{2,3}$
2.1.1.1	80	TCP	((\(\d{3}\) ?)|(\d{3}-))?\d{3}-\d{4}

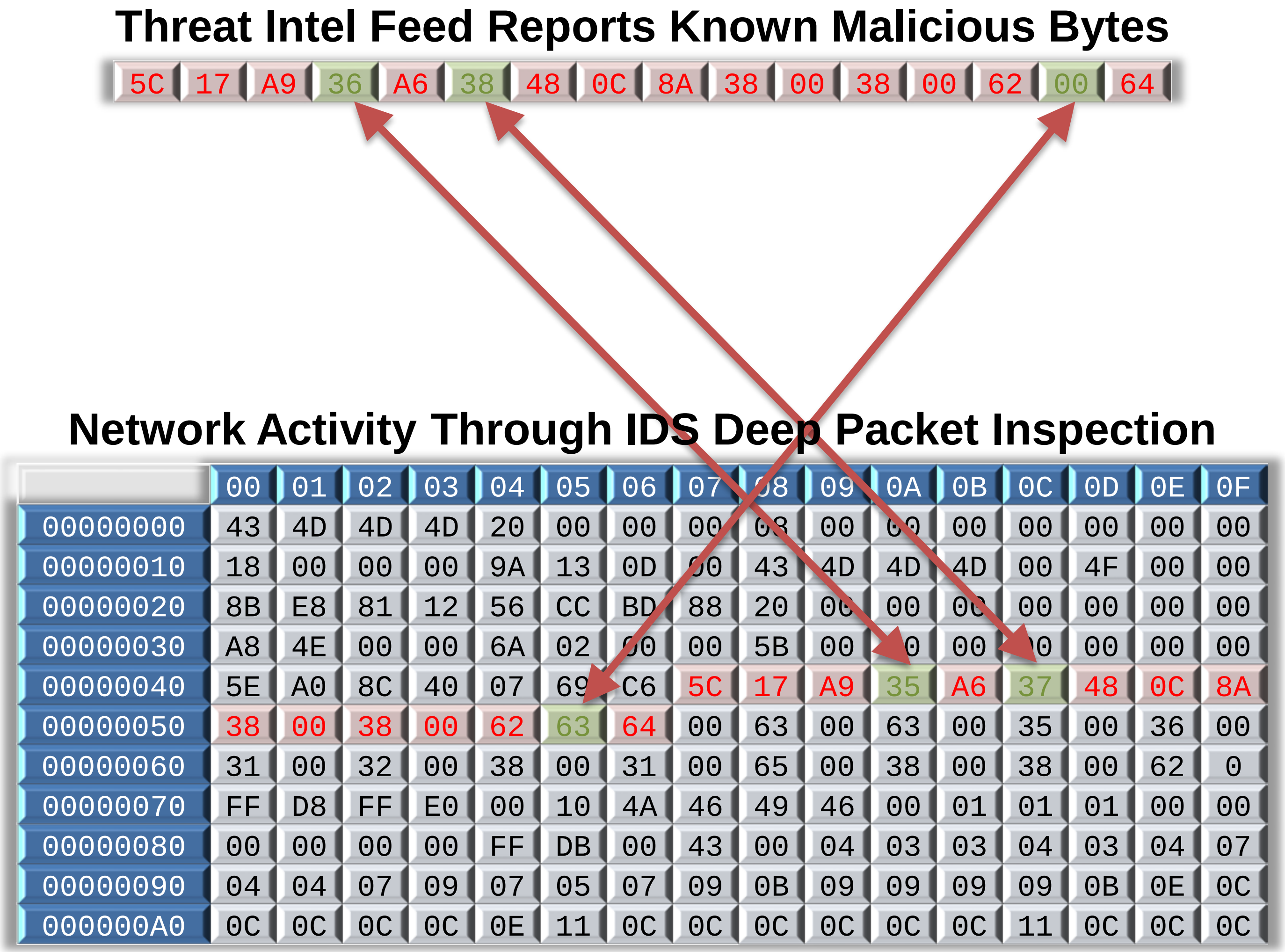
Netflow Activity

IP	Port	Protocol	Regex
1.1.1.1	53	UDP	bad@malware.net
2.1.1.1	80	TCP	(800) 800-1337
2.1.1.1	53	TCP	really.bad@malware.net

Fuzzy Matching

Approximate features seen in combination across datasets

- **Advantages**
  - Helps identify new tactics in complex interactions
  - Captures issues with minor changes
- **Drawbacks**
  - Fuzzier → more false positives
  - Requires feedback for refinement
  - Computationally expensive



Threat Intel Feed Reports Known Malicious Bytes

5C	17	A9	36	A6	38	48	0C	8A	38	00	38	00	62	00	64

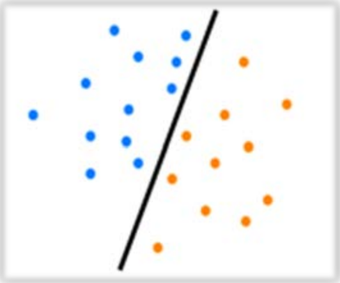
Network Activity Through IDS Deep Packet Inspection

	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F
00000000	43	4D	4D	4D	20	00	00	00	08	00	00	00	00	00	00	00
00000010	18	00	00	00	9A	13	0D	00	43	4D	4D	4D	00	4F	00	00
00000020	8B	E8	81	12	56	CC	BD	88	20	00	00	00	00	00	00	00
00000030	A8	4E	00	00	6A	02	00	00	5B	00	00	00	00	00	00	00
00000040	5E	A0	8C	40	07	69	C6	5C	17	A9	35	A6	37	48	0C	8A
00000050	38	00	38	00	62	63	64	00	63	00	63	00	35	00	36	00
00000060	31	00	32	00	38	00	31	00	65	00	38	00	38	00	62	0
00000070	FF	D8	FF	E0	00	10	4A	46	49	46	00	01	01	01	00	00
00000080	00	00	00	00	FF	DB	00	43	00	04	03	03	04	03	04	07
00000090	04	04	07	09	07	05	07	09	0B	09	09	09	09	0B	0E	0C
000000A0	0C	0C	0C	0C	0E	11	0C	0C	0C	0C	0C	0C	11	0C	0C	0C

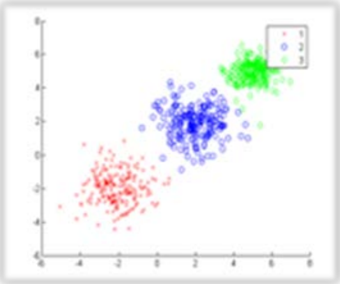
Machine Learning

Program computers to learn which dataset features are relevant

- **Advantages**
  - Identifies correlations humans haven't yet made
  - Can learn new tactics
- **Drawbacks**
  - Slow(ish)
  - Some ML approaches are not very scalable
  - Does not help build intuition
  - Tough to tune false positives/negatives



Classification



Clustering

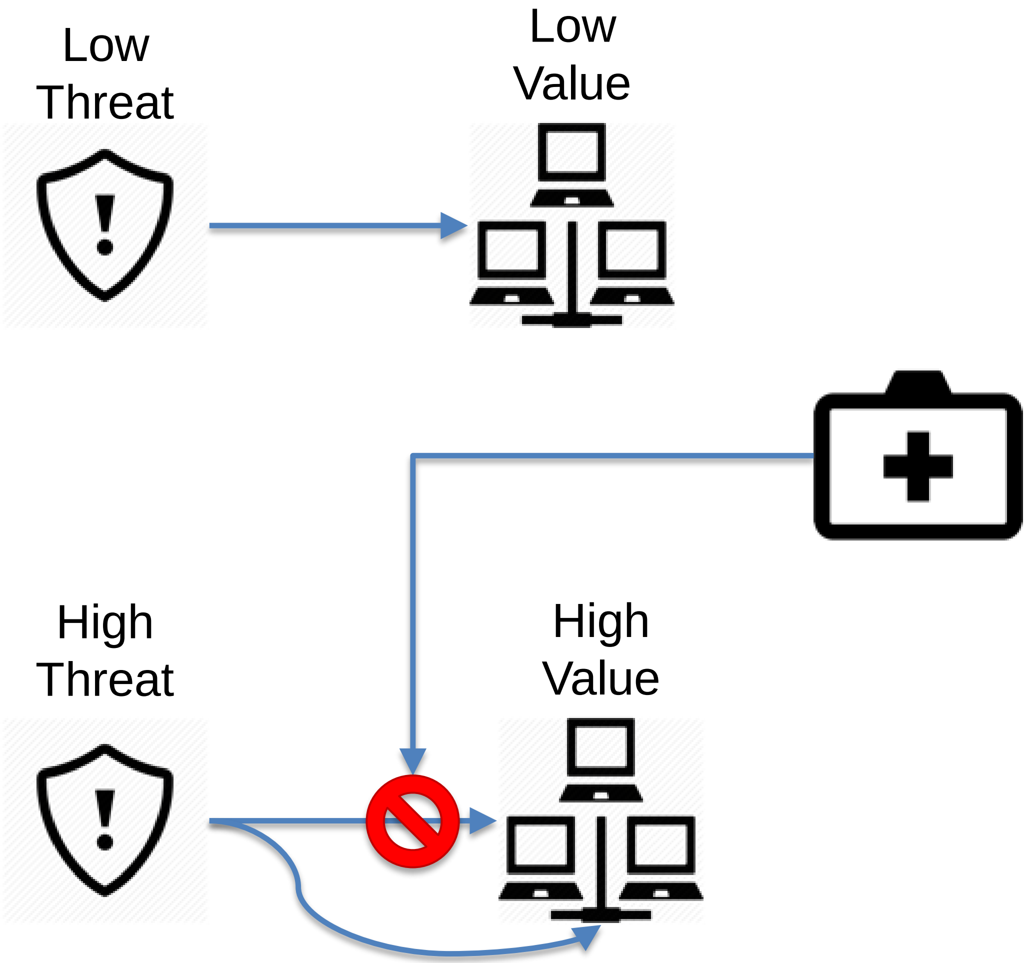
Neural Networks

How Can Hackers Evade Threat Correlation Detection?

Threat Correlation Approach	Common Evasion Tactics	Level of Effort
Manual Threat Correlation	• Increase amount of traffic to overwhelm humans	Low
Field Comparison	• Rotate use of unique identifiers (such as IP addresses & domains)	Low
Rules-Based Matching	• Rotate use of unique identifiers • Slight modifications to tools	Moderate
Fuzzy Matching	• Rotate use of unique identifiers • Significant modifications to tools	High
Machine Learning	• Rotate use of unique identifiers • Significant modification to tools • Continuously change tactics	Very High

Assessing Targeted Attacks

- Automating correlation of threat & network information can help your organization:
  - Identify active attacks
  - Assess attack severity
  - Prioritize response and mitigation activity
  - Identify important new threats & anomalies



Recommendations

Determine which threat intelligence feeds are best for your organization

Integrate threat intelligence into your automated threat management

Capture & **analyze** your network activity

Automate correlation of network activity with threat intelligence

Maximize impact with feedback loops within your threat management activities to continuously improve your organization's abilities

LOOKINGGLASS

Thank you

Web: www.lookingglasscyber.com

Twitter: @LG_Cyber