Federated Understanding of Security Information Over Networks (FUSION)

Program Manager: Mr. Richard Guidorizzi, DARPA, I2O

DigitalOperatives

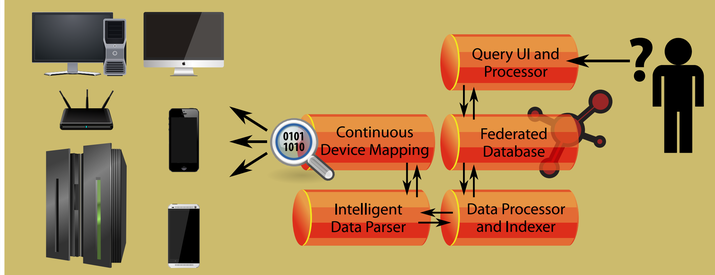
Drexel UNIVERSITY

fusion@digitaloperatives.com

Program Overview

FUSION is an Integrated Cyber Analysis System (ICAS) solution that incorporates agent-less data connectors, machine learning-enabled data parsing, and semantic understanding to put multi-pivotal, relationship-centric queries at the operator's fingertips.

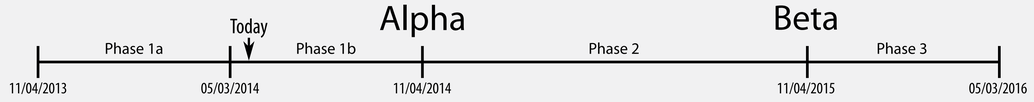
FUSION in a Nutshell



Deployment Advantages

- ✔ FUSION does not require heavy IT investment
- ✔ No software to install on individual workstations
- ✔ No need to alter existing network infrastructure
- ✔ All possible cyber data can be queried at any time

Planned Schedule



Capabilities Today

- Basic Device Discovery Engine
- Basic graphical user interface
- SSH Shell Connector
- Basic data parsing based on machine learning

Planned Capabilities at Alpha

- Windows Shell Connector
- Federated database fulfills simple queries
- Automatic data structure inference for well-structured data
- Cyber data ontology mostly complete with semi-automatic semantic mapping

Planned Capabilities at Beta

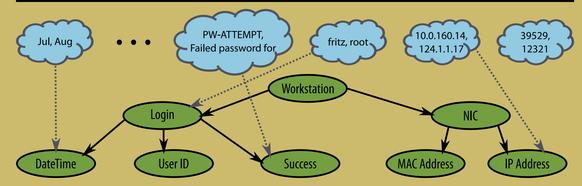
- Completed Device Discovery Engine
- Data Connector generates device-specific scripts
- Automatic data structure inference for semi-structured data
- Federated database complete, supports fully featured queries
- Cyber data ontology completed with automatic semantic mapping

Example Scenario: RSA SecurID Compromise (March 2011)

Discovery and Analysis with FUSION's Unique Capabilities

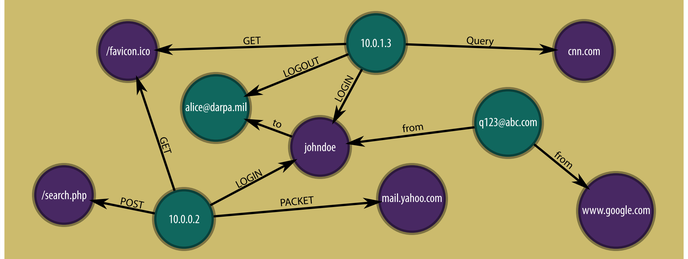
Artificial Intelligence to Understand New Data

Jul	13	09:37:59	basta	sshd[22308]:	PW-ATTEMPT:	fritz							
Jul	13	09:37:59	basta	sshd[22308]:	Failed	password	for	fritz	from	10.0.160.14	port	39529	ssh2
Aug	1	09:38:02	basta	sshd[22310]:	PW-ATTEMPT:	root							
Aug	1	09:38:02	basta	sshd[22310]:	Failed	password	for	root	from	124.1.1.17	port	12321	ssh2



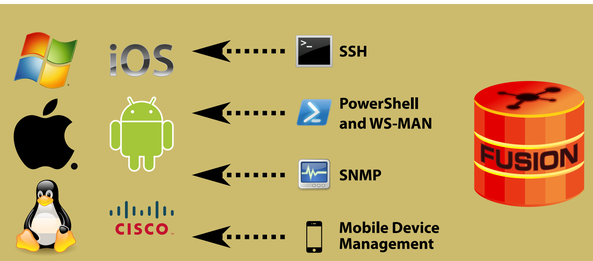
FUSION uses machine learning and a "cyber data" ontology to automatically map heterogeneous data and infer their semantics. This allows FUSION to index any data source, including those not previously seen.

Discover and Explore Complex Relationships in a Graph



FUSION uses a graph database to allow new ways for operators to interact with and query the data. The graph model can dynamically accommodate any relationship within the data and reveal its interconnections.

Automatic Device Connectivity



FUSION reuses existing, trusted remote connectivity and administration tools native to each device. It retrieves local data sources with a combination of high level command line programs and scripts.

Phishing email with malicious Excel attachment

In this example scenario, FUSION automatically **recognizes** and **parses** email server logs and email messages stored on the email server through **machine learning** techniques. The cyber defender correlates that four users received an identical email from an unknown source with an attached file, indicating they were targeted by a spear phishing attack.

The attachment exploits an Adobe Flash vulnerability to run Poison Ivy, a Remote Administration Trojan (RAT)

The beta version of FUSION will interface with various third-party process and memory forensics tools such as Volatility, VMMap, and Process Explorer. This would enable the defender to detect sophisticated threats such as rootkits, injected threads in processes, or hidden network connections.

Poison Ivy connects to a pre-configured, remote server where the attackers wait

In this example scenario, FUSION maps all DNS requests from the endpoint devices as **relationships** in a **visual graph**. The defender learns that two of the endpoint devices requested to resolve *www.usgoodluck.com*, a known C2 point for this attack.

Attackers use Poison Ivy to upload and execute files, gather information, and spread through the network

The beta version of FUSION would correlate information from access logs with internal netflow metadata to display the relationships of users to resources and to potentially related file transfers. Unusual patterns of access and data movement within the network will be apparent in the graph view.

Attackers install backdoors and exfiltrate sensitive data

In this example scenario, FUSION **connects** to the packet capture appliance via its **agent-less** SSH Data Connector. FUSION then searches for pcap (network traffic) files and utilities by which the pcap files can be processed. It finds Xplico, a network forensic analysis tool, uses it to process the pcap files, then retrieves the results.