RSAConference2015
Abu Dhabi | 4–5 November | Emirates Palace

SESSION ID: CIN-W06

Using Security Analytics to Transform Smart Grid Security

Dr. Robert W. Griffin
Chief Security Architect
RSA, The Security Division of EMC

Dr. Silvio La Porta
Lead Research Scientist
EMC Research Europe

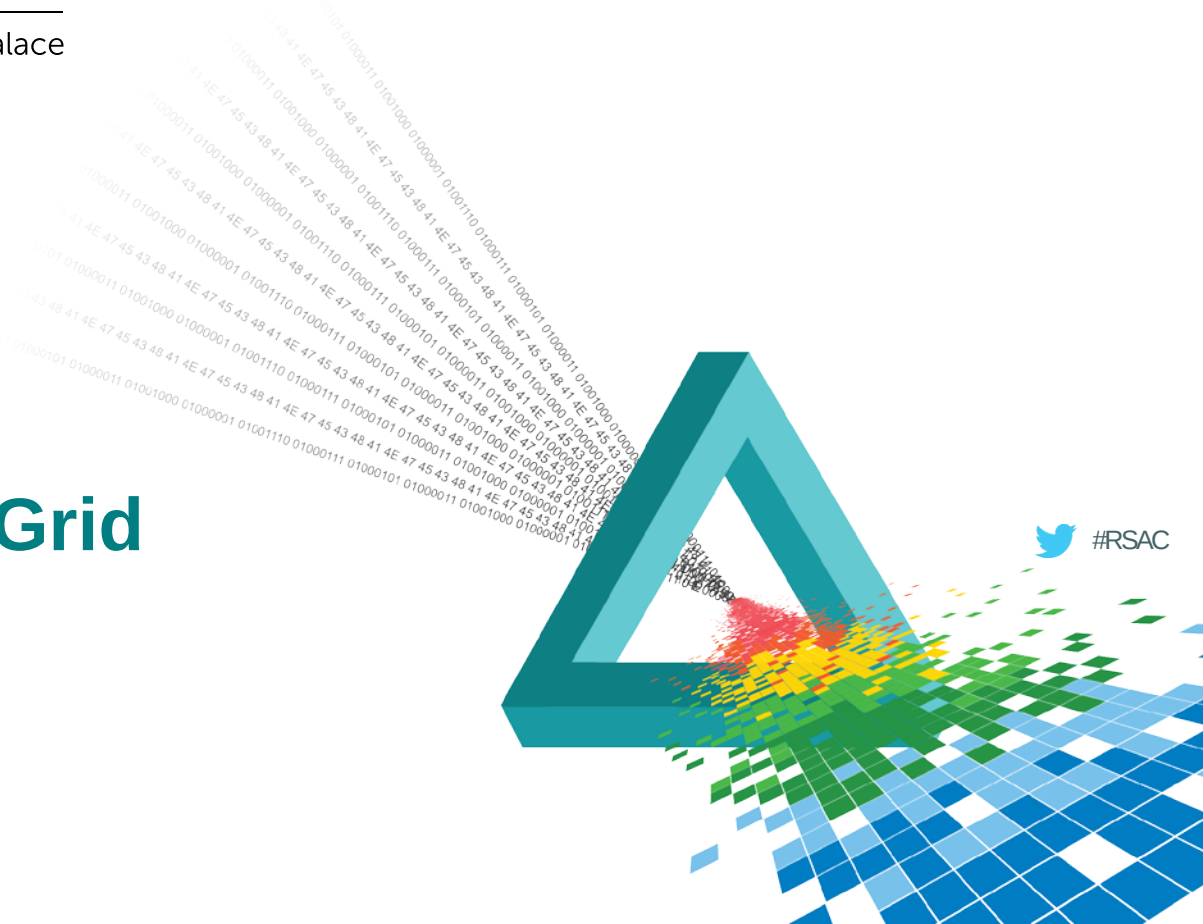
Agenda

- Threats to Smart Grid security
- Cyber risk and the Smart Grid
- Security analytics and the Smart Grid

RSAConference2015

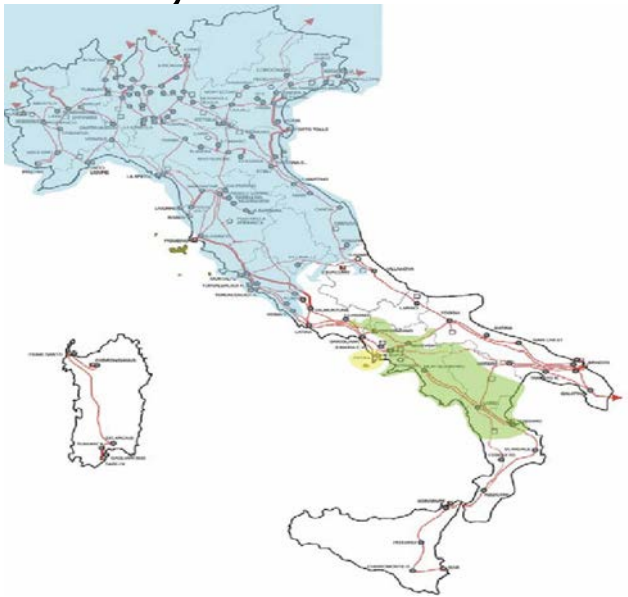
Abu Dhabi | 4–5 November | Emirates Palace

Threats to Smart Grid security



Is Smart Grid Security a real problem?

28th September 2003, Italy (except Sardinia) blackout started on Sunday at 3 am and lasted up to 24 hours.



Cause: transmission line between Switzerland and Italy was cut by tree, and additional minor events

Blue area: supply restored at 12 o‘clock

White area: still unsupplied at 12 o‘clock

Green area: initially successful restoration failed and blackout continued

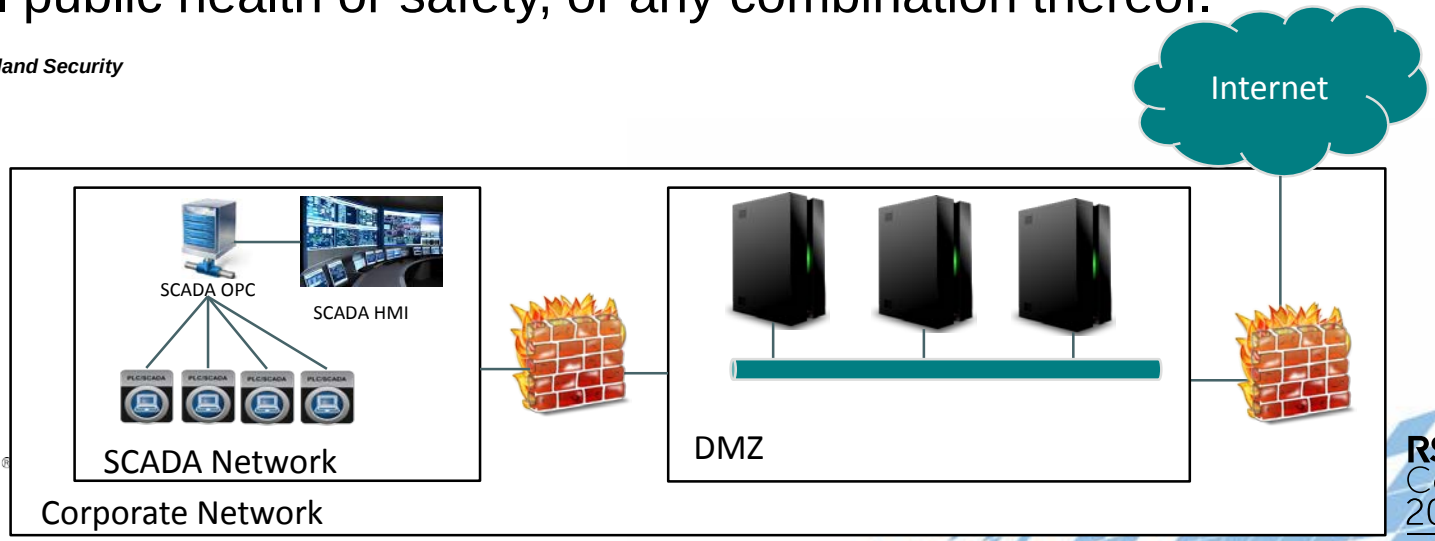
Overall economic damage: **1.180m €**

Household damage: **285m €**

What is a critical infrastructure from attacker point of view? An opportunity!

Critical infrastructure are the assets, systems, and networks, whether physical or virtual, so vital to the United States that their incapacitation or destruction would have a debilitating effect on security, national economic security, national public health or safety, or any combination thereof.

- ***USA Homeland Security***



SandWorm crew

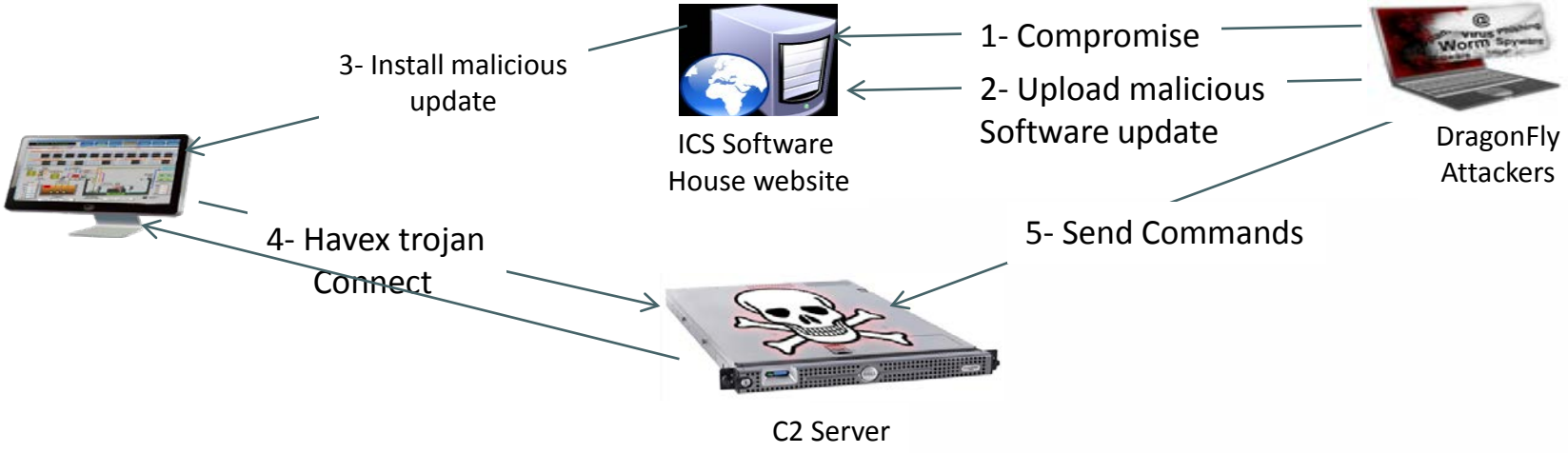
- Has targeted NATO, Western government and military active from 2009 (iSight's report)
- Probably originated in Russia
- A new version of the crimeware Trojan BlackEnergy V2-V3
- They used a zero-day exploit **CVE-2014-4114**
  - Microsoft Windows Vista SP2, Windows Server 2008 SP2 and R2 SP1, Windows 7 SP1, Windows 8, Windows 8.1, Windows Server 2012 Gold and R2, and Windows RT Gold and 8.1 **allow remote attackers to execute arbitrary code via a crafted OLE object in an Office document**, as exploited in the wild with a "Sandworm" attack in June through October 2014, aka "Windows OLE Remote Code Execution Vulnerability."

BlackEnergy

- Well established crimeware tool
- Originally was used as “banking trojan”
- Modular framework, it is possible to use plugins to extend its functions
- Multiple versions exist (BlackEnergyLite, B2 and B3)
- Able to scan IP blocks looking for HMI machine TCP port (**CVE-2014-0751)**
  - Directory traversal vulnerability in CimWebServer.exe (aka the WebView component) in GE Intelligent Platforms Proficy HMI/SCADA - CIMPLICITY before 8.2 SIM 24, and Proficy Process Systems with CIMPLICITY, **allows remote attackers to execute arbitrary code via a crafted message to TCP port 10212**

DragonFly campaign

The recent DragonFly campaign showed how the attackers could use malware to take control of SCADA systems



DragonFly campaign (2/2)

In order to propagate their malware, attackers also use:

- **Watering hole attack** (compromise ICS software house websites with iframes which will redirect victims to Exploit kit websites)
- **Phishing mail campaign** with malicious PDF files targeting specific employees

The malwares used are backdoor or R.A.T.

(Trojan.Karagany , Havex)

- Havex is able to **scan OPC servers** used for controlling SCADA devices in critical infrastructure

RSAConference2015

Abu Dhabi | 4–5 November | Emirates Palace

Cyber risk and the Smart Grid

Disruption and Transformation

Mobile

Cloud

Big Data

Extended Workforce

Networked Value Chains

APTs

Sophisticated Fraud

Infrastructure Transformation

Less control over access device and back-end infrastructure

Business and Legal Transformation

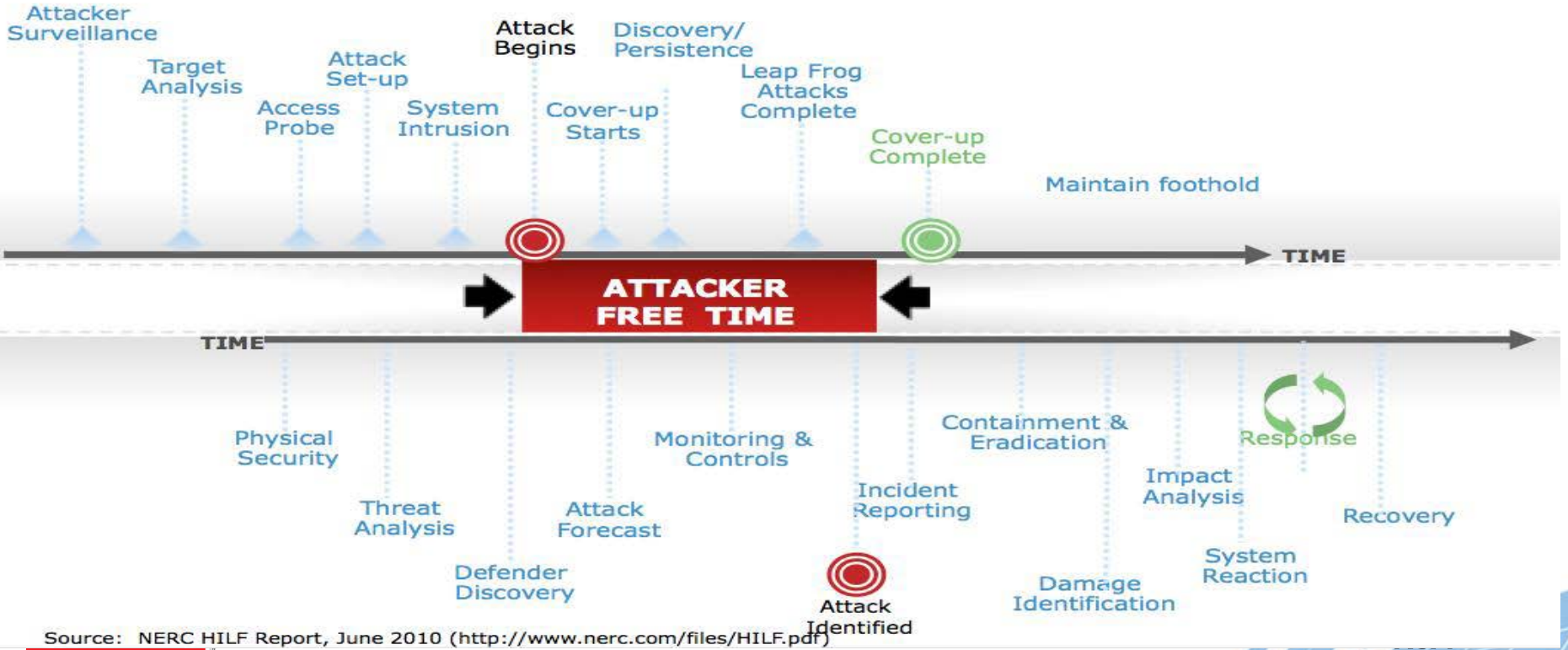
More hyper-extended, more digital

Threat Landscape Transformation

Fundamentally different tactics, more formidable than ever

http://www.emc.com/collateral/industry-overview/h11391-rpt-information-security-shake-up.pdf?pid=sbiclandingpage-sbicspecialreport-122112

Targeted Attacks



Source: NERC HILF Report, June 2010 (http://www.nerc.com/files/HILF.pdf)

Miami Substation Explosion (1993)

1. Power surge destroyed capacitor bank
2. Tripped breaker malfunction caused arc fault
3. Emergency response system did not activate to notify grid dispatcher
4. Coolant in primary transformer overheated and blew seals.
5. Escaping mineral oil vapor ignited at arc fault
6. Vapor explosion ignited oil tank for primary transformer
7. Explosion of substation caused meltdown of all transformers

https://www.youtube.com/watch?v=WkDCS8xeobg

Arkansas Substation Fire (29-Sep-2013)

- Jason Woodring arrested October 2013
  - Alleged to have cut his way into an Entergy substation
  - Used mixture of ethanol and motor oil to burn up a control house
  - «You should have expected U.S.»
- One of several related attacks, including cutting power poles

http://www.cnn.com/2013/10/08/us/arkansas-grid-attacks/

Largest Electric Grid Outages

article	millions of people affected	location	date
July 2012 India blackout	620	India	30 July 2012-31 July 2012
January 2001 India blackout	230	India	2 January 2001
November 2014 Bangladesh blackout	150	Bangladesh	1 November 2014
2015 Pakistan blackout	140	Pakistan	26 January 2015
2005 Java–Bali blackout	100	Indonesia	18 Aug 2005
1999 Southern Brazil blackout	97	Brazil	11 March 1999
2009 Brazil and Paraguay blackout	87	Brazil, Paraguay	10–11 Nov 2009
2015 Turkey blackout	70	Turkey	31 March 2015
Northeast blackout of 2003	55	United States, Canada	14–15 Aug 2003

https://en.wikipedia.org/wiki/List_of_major_power_outages

Using Failure Scenarios to Understand Risk

AMI.1 Authorized Employee Issues Unauthorized Mass Remote Disconnect

Description: An employee within the utility having valid authorization, issues a "remote disconnect" command to a large number of meters. The employee may be bribed, disgruntled, or socially engineered.

Relevant Vulnerabilities:

- Inadequate system and process checks for disconnect commands.

Impact:

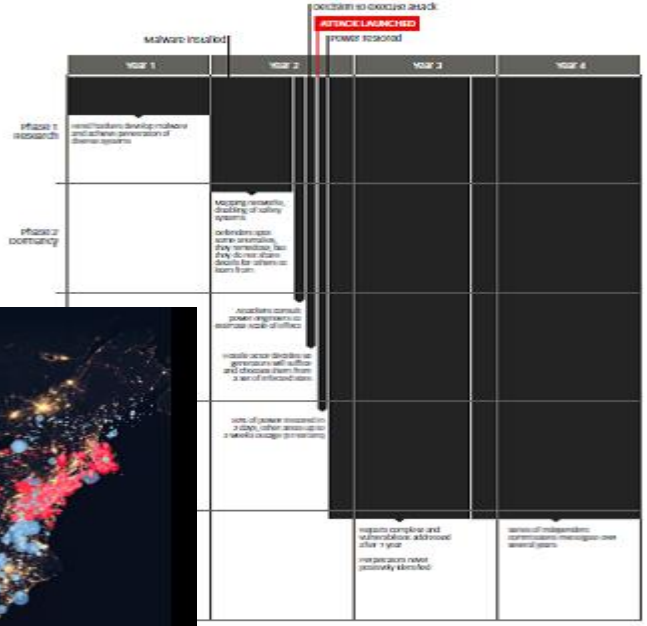
- An instantaneous mass disconnect/reconnect over multiple feeders, if permitted by the system, could cause temporary blackouts due to circuit breaker trips until power on the grid can be rebalanced,
- A small number of disconnects could subvert the Smart Grid deployment and make the utility lose consumer confidence.

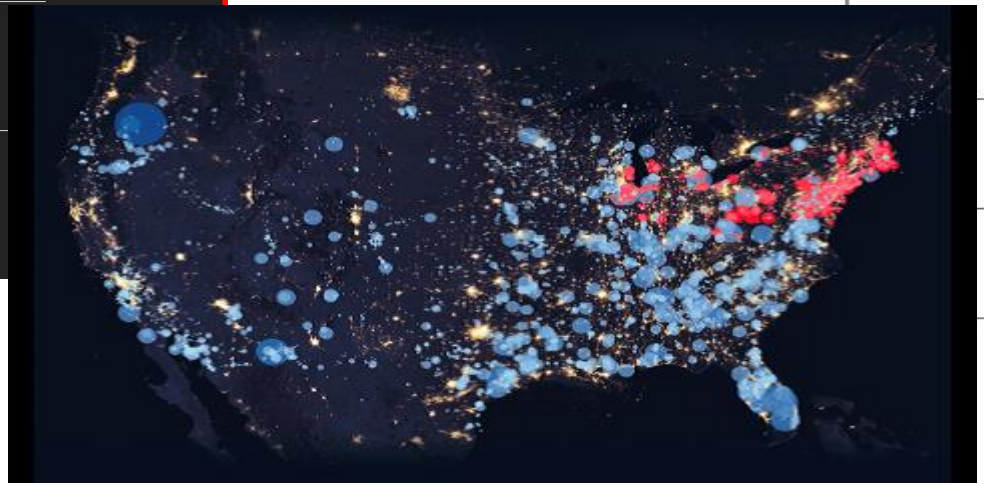
Potential Mitigations:

- *Detect anomalous commands* (anomalous disconnect and reconnect commands) not stemming from the normal Customer Information System (CIS) system,

http://smartgrid.epri.com/NESCOR.aspx

The Lloyds' Erebos Scenario



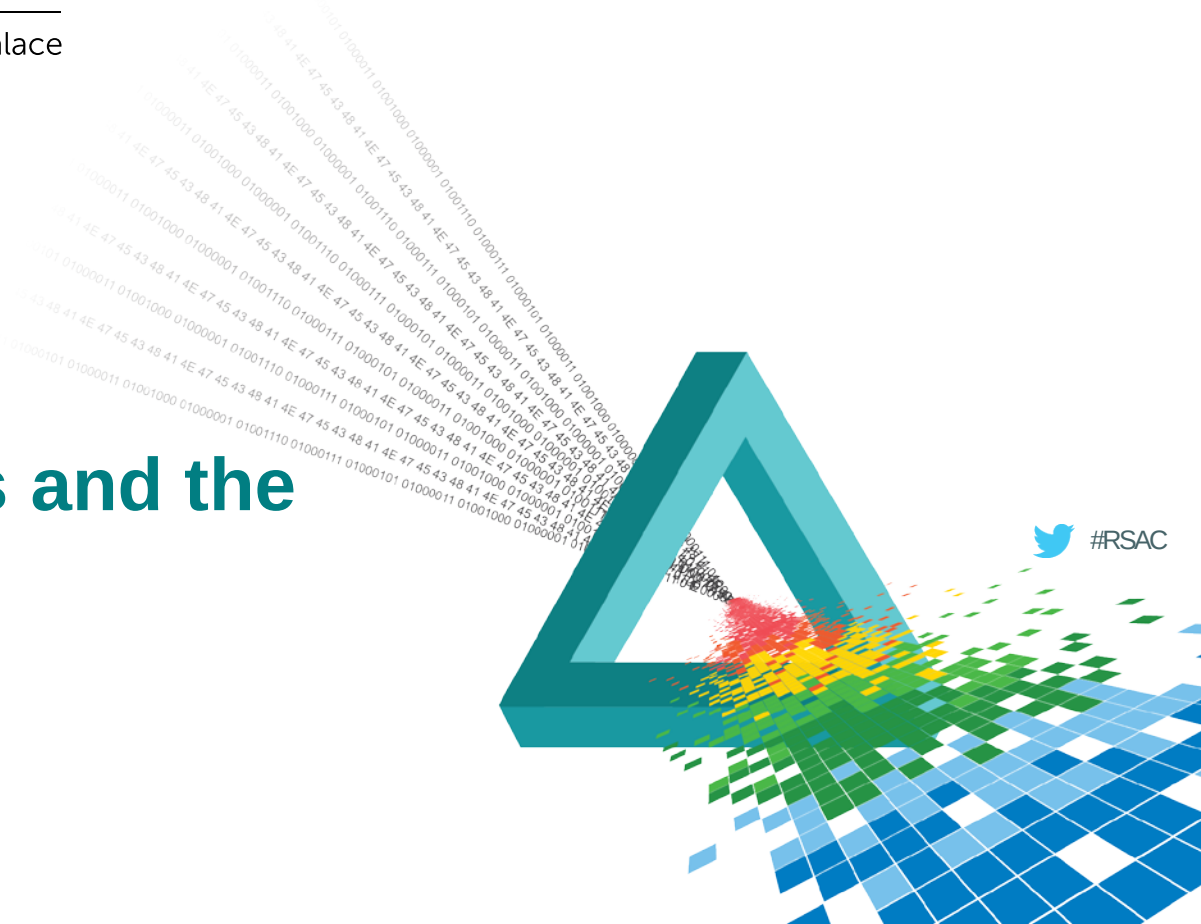


https://www.lloyds.com/lloyds/press-centre/press-releases/2015/07/business-blackout

RSAConference2015

Abu Dhabi | 4–5 November | Emirates Palace

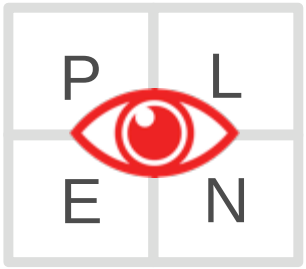
Security analytics and the Smart Grid



Re-thinking Security for Smart Grid

Data Collection and Rationalization

Visibility



Packets, Logs,
Endpoints, NetFlow

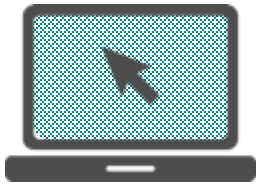
Capture Time
Data Enrichment

Business &
Compliance Context

Generating Information

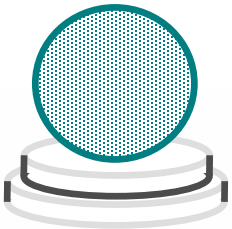
Analysis

Correlate Multiple Data Sources



Endpoint Threat Detection

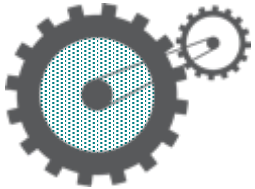
Out-of-the-box Content



Big Data & Data Science

Investigation and Remediation

Action



Prioritized & Unified
Analyst Workflow

Investigate down
to finest details

Integrate SOC
Best Practices

Example: Detect suspicious domain connections

Identifying suspicious domains is difficult – and identifying hosts that have ever communicated with one is even harder.

Number of domain name owners associated with an IP address

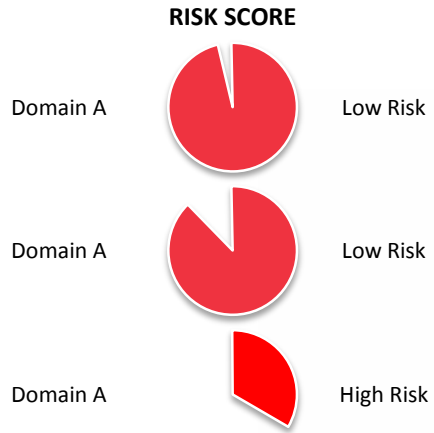
A high number of domain owners associated with a system is suspicious

Number of IP addresses

Malicious domains use many IP addresses to evade static IP watchlists.

Traffic content types

Suspicious domains often host many services on the same server.



Number of users hitting a domain relative to complexity

A complex domain that few people access is more likely to be malicious.

GETS vs PUT/POSTs

Domains where the ratio of POSTs to GETs is high are more likely malicious.

Example: Discover beaconing hosts

Traffic from hosts doing automated connections to (beaconing) command and control hosts can look like normal traffic. Data science helps identify outliers.

Use of referrer strings

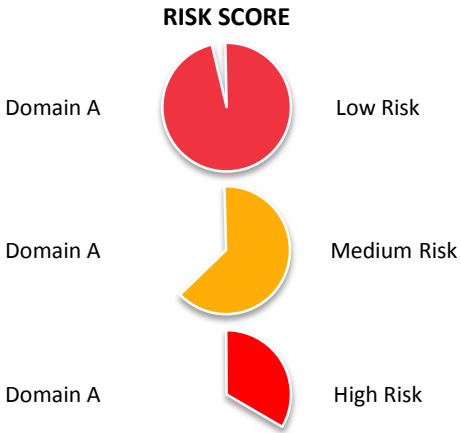
Most web sessions come from clicking on another link, resulting in a “referrer string”. Malicious sessions seldom do.

Bytes uploaded vs. downloaded

Malicious sessions often upload far more than just a URL request.

Use of cookies

Malicious sessions seldom use cookies.



URL lengths

Malicious attacks often embed themselves deep in web servers, resulting in unusually long URL lengths.

Other

RSA uses several other identifiers to determine the risk score.

Communication Valley Reply (Italy) Leveraging Security Analytics

- Requirements:
  - Efficient, cost-effective management and reporting of security
  - Reduce cost of services delivery
  - Improved MSSP service as competitive advantage
- Solution:
  - Automatically tracked and reported on client risk and compliance
  - Enhanced incident triage
  - Improved event analysis

http://www.emc.com/collateral/customer-profiles/h11982-reply-cp.pdf

Evolving Role of Data Analytics in Security

Data goes in, hard to extract value

Fast queries over large data

Single source metrics, single correlation, rule based, high false positive

Leverage full contextual info, multi-source, automatic, for low false positives

SPARKS Project Consortium

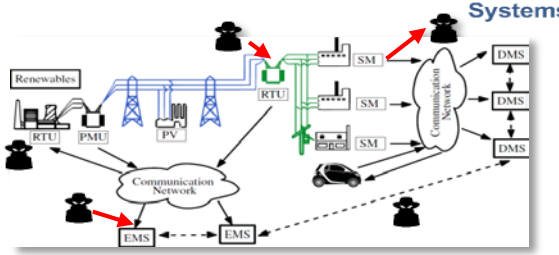
SPARKS Mini-projects

Intrusion Detection for SCADA Systems

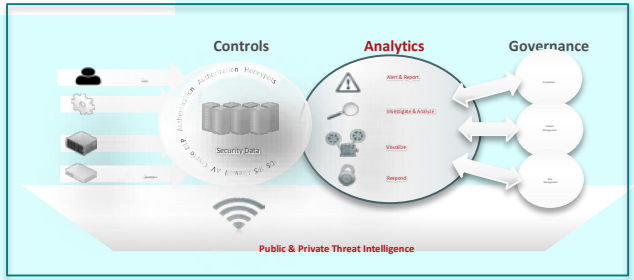
Novel Authentication and Key Management Technology

Security Technologies

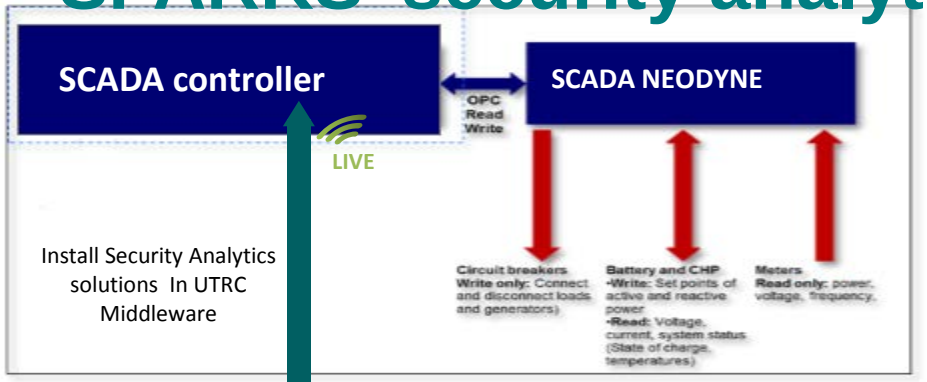
Resilient Control Systems



Security Information Analytics



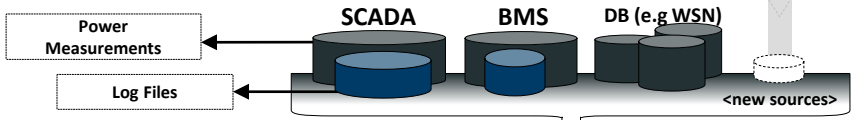
SPARKS' security analytics test environment



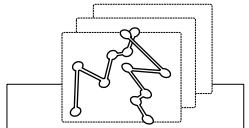
Install Security Analytics solutions In UTRC Middleware

Enable New (?) attacks

Additional Data



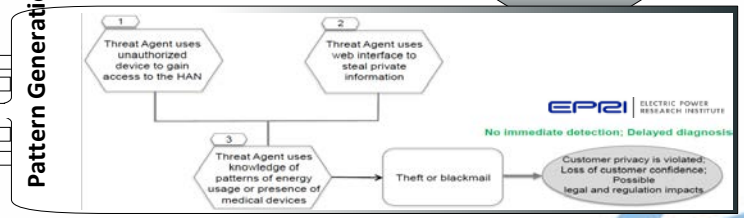
NESCOR attack trees Formulation



Demo site attack down-selection

Final demonstration

Pattern Generation



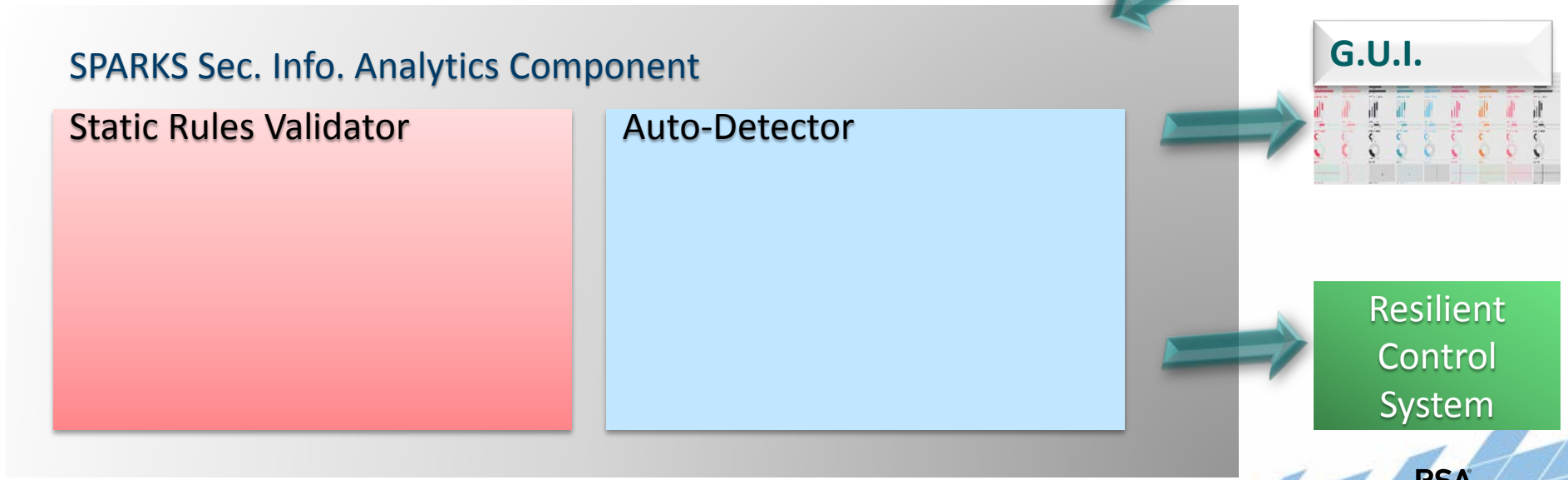
Use security analytics as inputs for designing resilient control algorithms

Example of AMI.29: Unauthorized Device Acquires HAN Access and Steals Private Information

SPARKS Sec. Info. Analytics Component

The module is composed of two main components

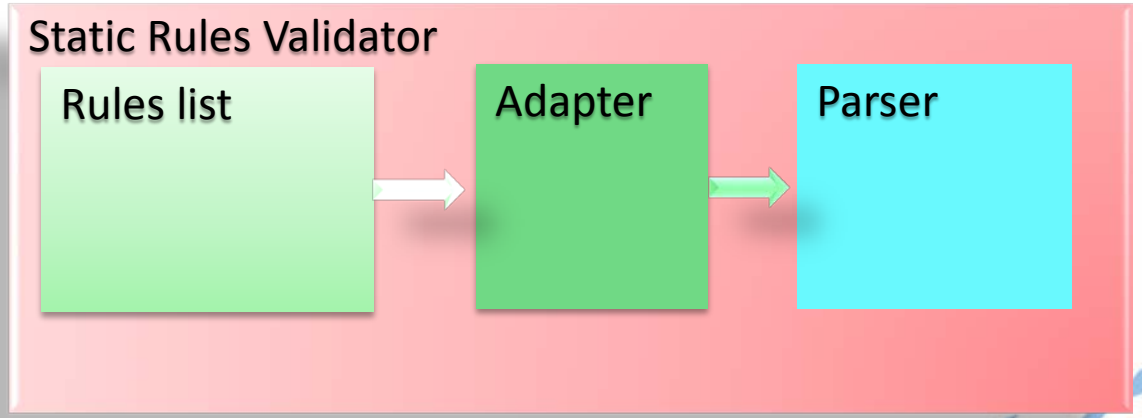
- Static Rules Validator
- Auto-Detector



Static Rules Validator

The component will search for systems' asserts violations

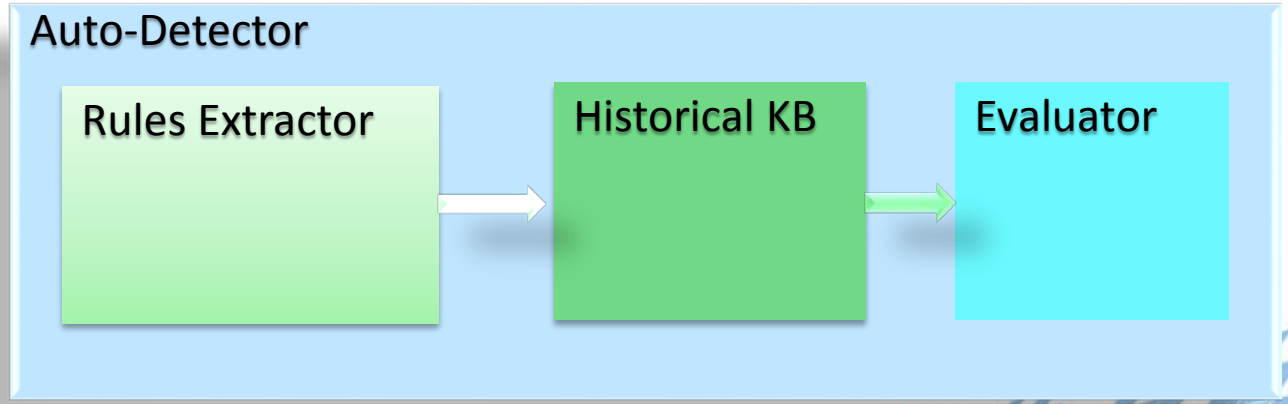
- *Rules List* contains the assertions to verify
- *Adapter* translate the rules in common language
- *Parser* get the rules and search for negative or positive outliers



Auto-Detector

The component will use machine learning technique to evaluate the entire system state

- *Rules Extractor* get data from last readings
- *Historical KB* compare the new feature with system history
- *Evaluator* use tolerance to reduce FP and noise



Algorithm Summary

- Four *static* components:
  - Rule Validator
  - Variable Validator
  - Kullback Distance
  - Dead Sensor Clustering
- Future *dynamic* component:
  - Machine learning outlier detector

Variable outlier

- Outliers against a predefined bound
  - E.g. Voltages should not fluctuate very much
- Examine voltages and frequency only

Rule Outlier

- Calculate physical relationships between variables
- 18 separate equations

$$\cos^{-1}\frac{V_A^{\,2}+V_B^{\,2}-V_{AB}^{\,2}}{2V_A\,V_B}+\cos^{-1}\frac{V_B^{\,2}+V_C^{\,2}-V_{BC}^{\,2}}{2V_B\,V_C}+\cos^{-1}\frac{V_C^{\,2}+V_A^{\,2}-V_{CA}^{\,2}}{2V_C\,V_A}=360°$$

- Measurement is asynchronous
  - Use difference between RHS and LHS (*error*)
- Determine probability of error from historical data
- Flag when below some threshold

Kullback Leibler

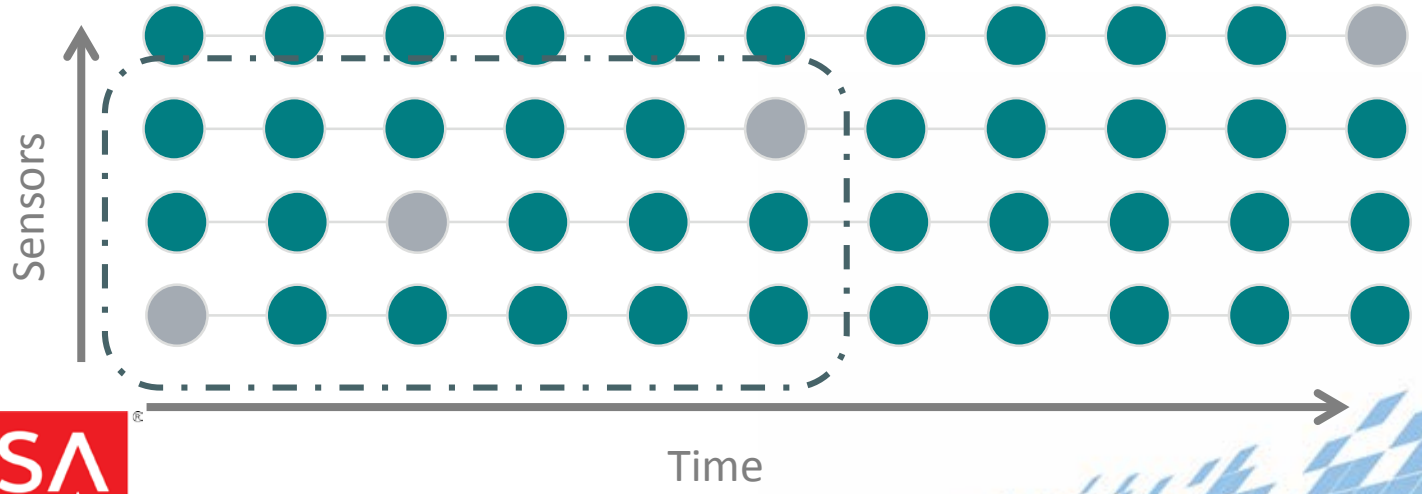
- Symmetrized KL distance on rule errors
  - Symmetrisation due to Kullback & Leibler

$$D_{KL} = d_{KL}(j,i) - d_{KL}(i,j)$$

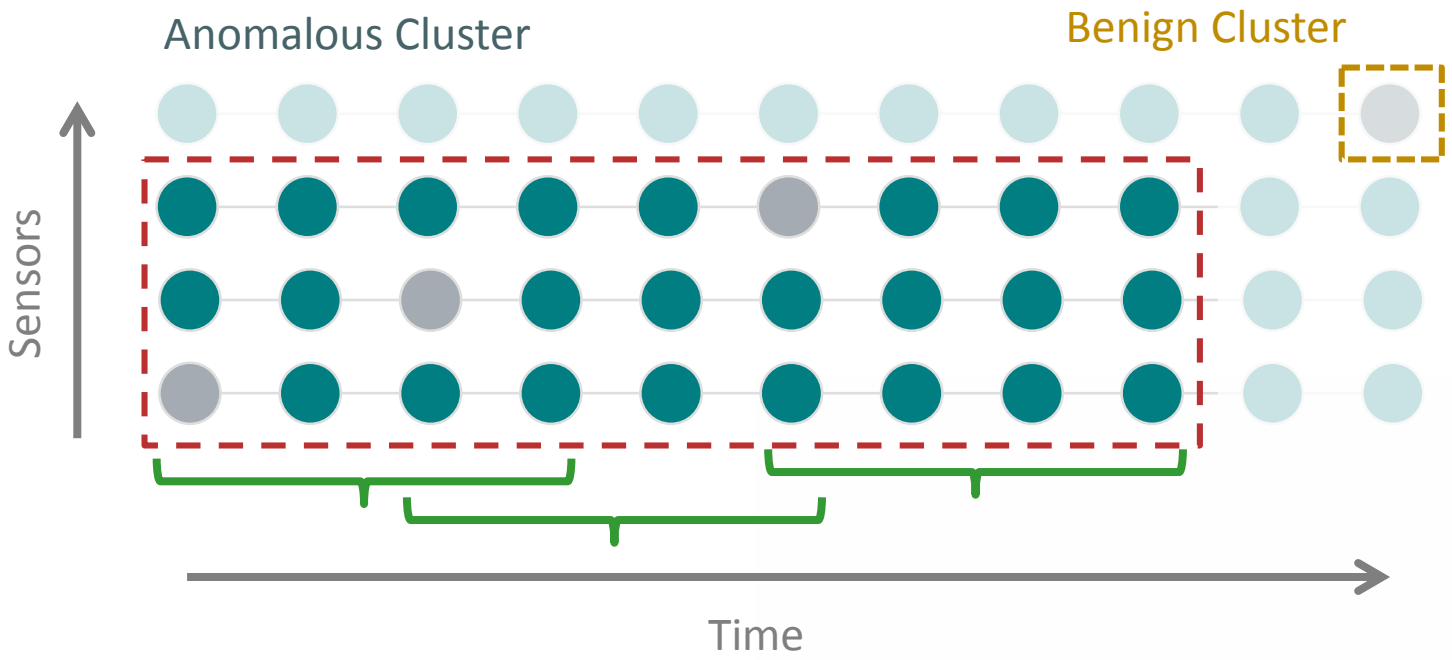
- Historical data (baseline) vs Current measurement
- Anomaly when value above some threshold

Dead Sensor Clustering

- Cluster sensors that stop recording in time
- User configurable time window
- Anomalous when cluster size > threshold



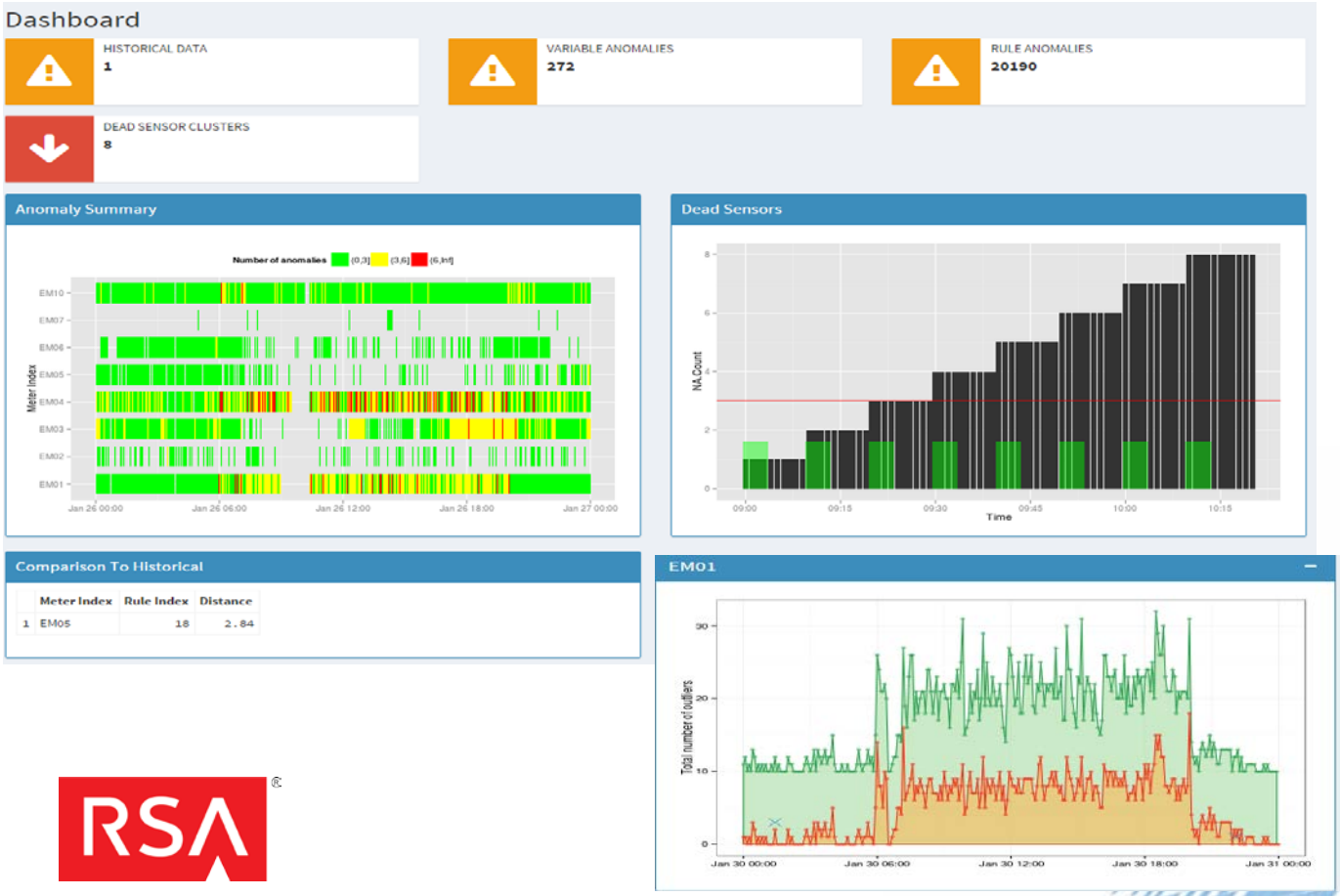
Dead Sensor Clustering

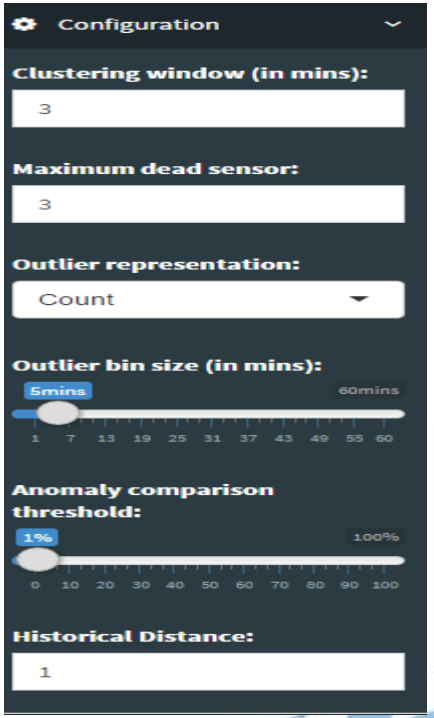


ML Outlier detection

- Prototyping phase
- Abundance of *normal* data. Little to no *outlier* data
- Train a one-class SVM using only normal data
- *Provisional:* Group similar sensors and train a model for each sensor using only
- Early studies show good performance but modelling needs more work

Some Screenshots of SPARKS' S.I.A.





Work in progress

Data analytics Algorithm basic features :

- Patterns Detection and Patterns Violation (example battery is charged everyday between 7am-12am and discharged between 6pm-10pm)
- Inter Meter checks

Already clear demonstration of the value of security analytics for Smart Grid!

Applying this Session

- Evaluate your current approach to responding to cyber threats in the light of the kinds of attacks that are being waged on Smart Grid
- Assess whether there is an area in which security analytics could provide significant improvement in your respond to cyber threats
- Prototype of pilot security analytics in that area

Thank you!

silvio.laporta@emc.com

robert.griffin@rsa.com
blogs.rsa.com/author/griffin
project-sparks.eu/blog/
@RobtWesGriffin
www.linkedin.com/pub/robert-griffin/0/4a1/608