Attack Patterns Recognition Framework

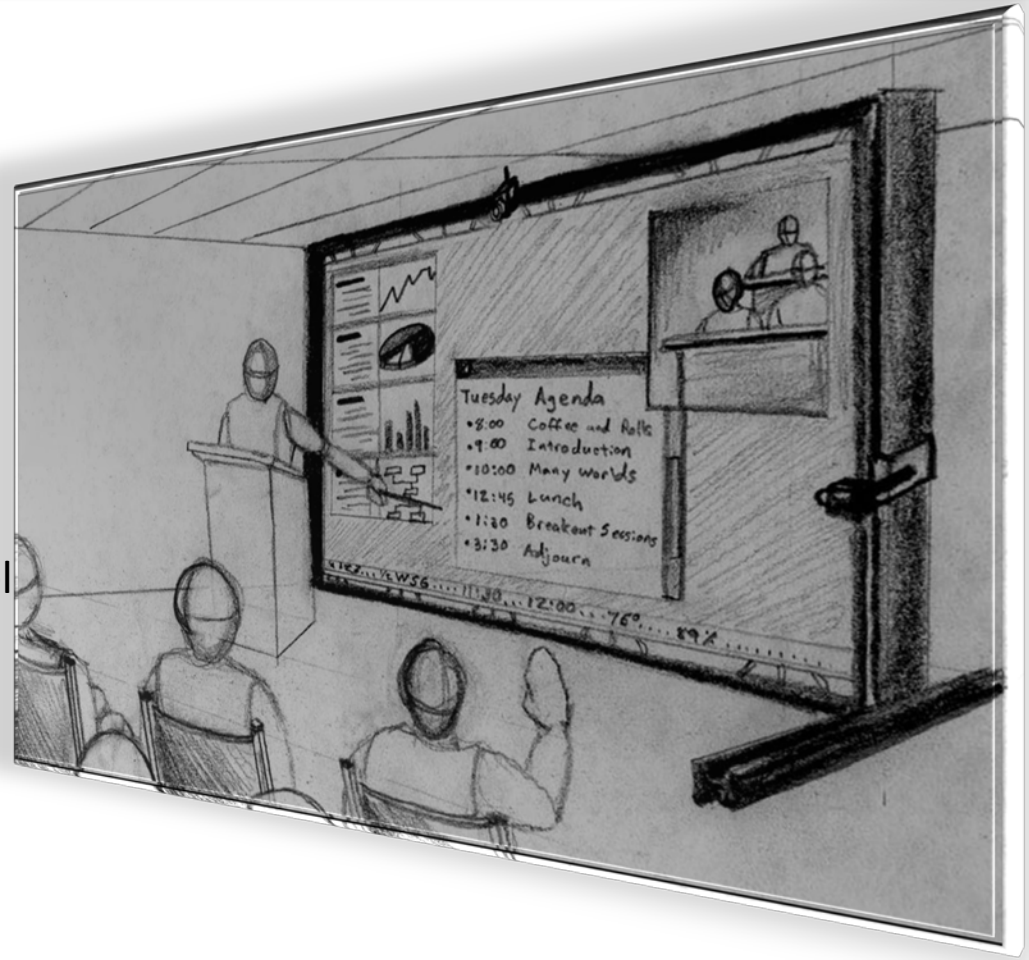
Noor-ul-hassan Shirazi, Alberto Schaeffer-Filho
and David Hutchison

Lancaster University

MSN2012:The Multi Service Networks Workshop
Cosener's House, Abingdon, Oxfordshire, UK
12-13 July 2012

Contents

- Goal
- Motivation
- Attack Pattern Recognition
  - Related Work
  - Proposed Model
  - High Level Design
- Stages of Proposed Model
  - Feature Extraction and Sel
  - Choice of Clustering
  - Aggregation/Fusion
- Future Work
- Conclusion
- References



Goal

❑ Attack Pattern Recognition

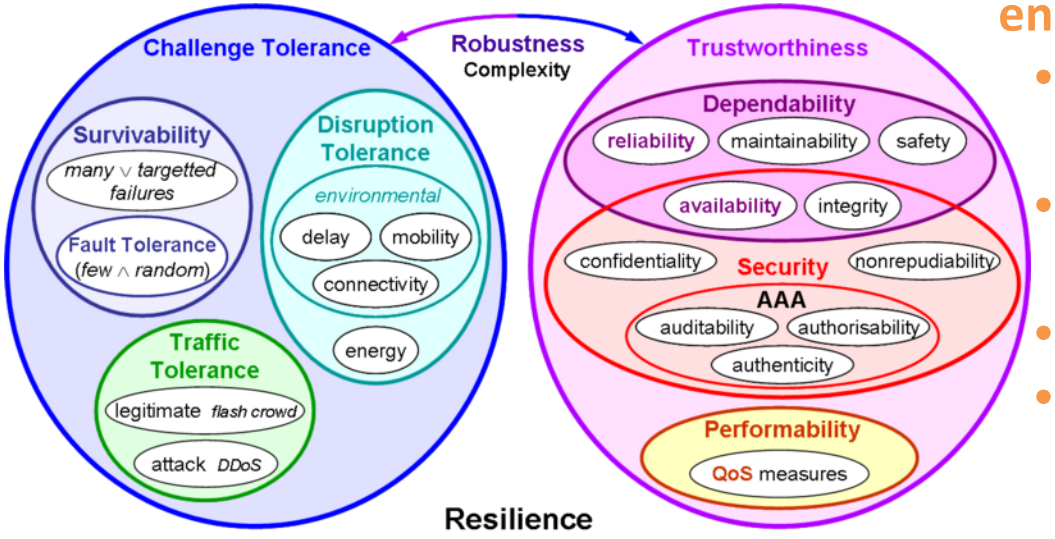
- The goal of this position paper is to propose a framework for attack pattern recognition by collecting and correlating cyber situational information vertically across protocol-levels, and horizontally along the end-to-end network path.

- To analyse cyber challenges from different viewpoints and to develop effective countermeasures.

Motivation:Network Resilience?

"The ability of the network to provide and maintain an acceptable level of service in the face of various faults and challenges."

Ref: ResumeNet: Resilience and Survivability for Future Networking: Framework, Mechanisms, and Experimental Evaluation (FP7)



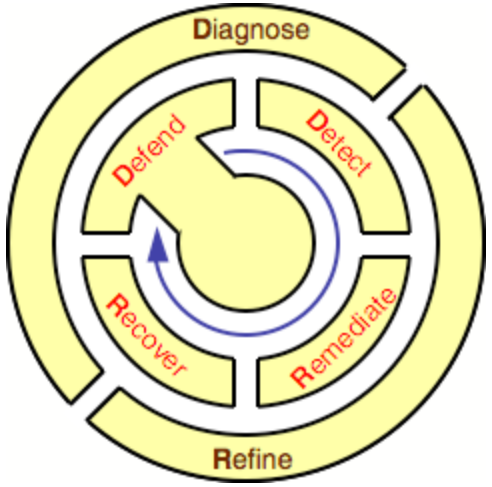
- **Network resilience is difficult to ensure and it is a wide topic**
  - **Tackles important Future Internet issues.**
  - **Configuration of systems is complex.**
  - **Spans across several levels.**
  - **Subject to a wide range of challenges.**

Motivation

Network security and resilience framework: D2R2 + DR

- **Real-time control-loop (D2R2)**
- Defend against challenges to normal operation
- Detect when an adverse event has occurred.
- Remediate the effects of the adverse event
- Recover to original and normal operations

- **Offline control-loop (DR)**
- Diagnose what caused the challenge
- Refine operation to prevent it from happening again



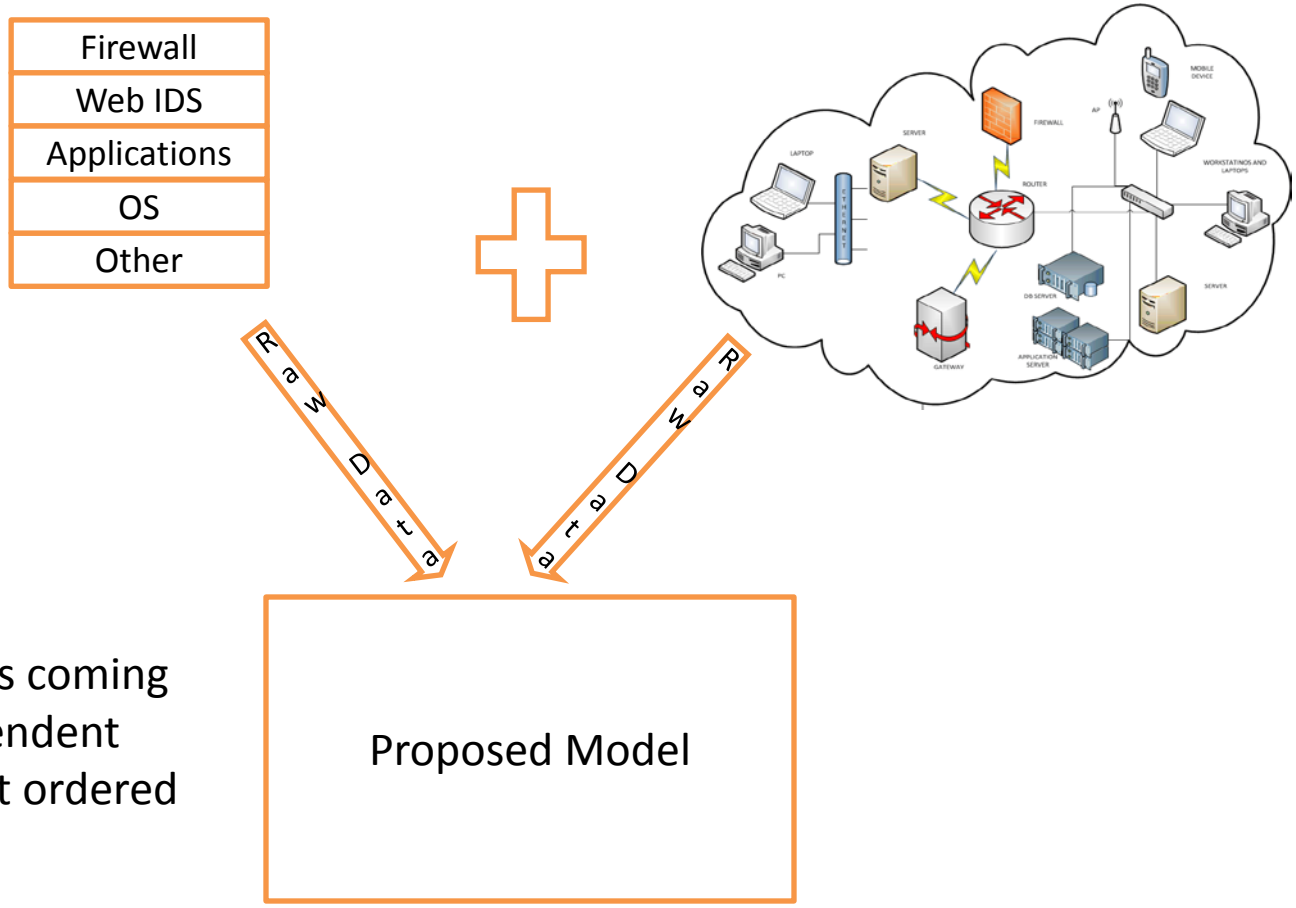
- **Conceptual framework.**
- **Network- and service-level mechanisms.**
- **Systematic approach to resilience.**
- **Blueprint for designing resilient system.**

Related Work

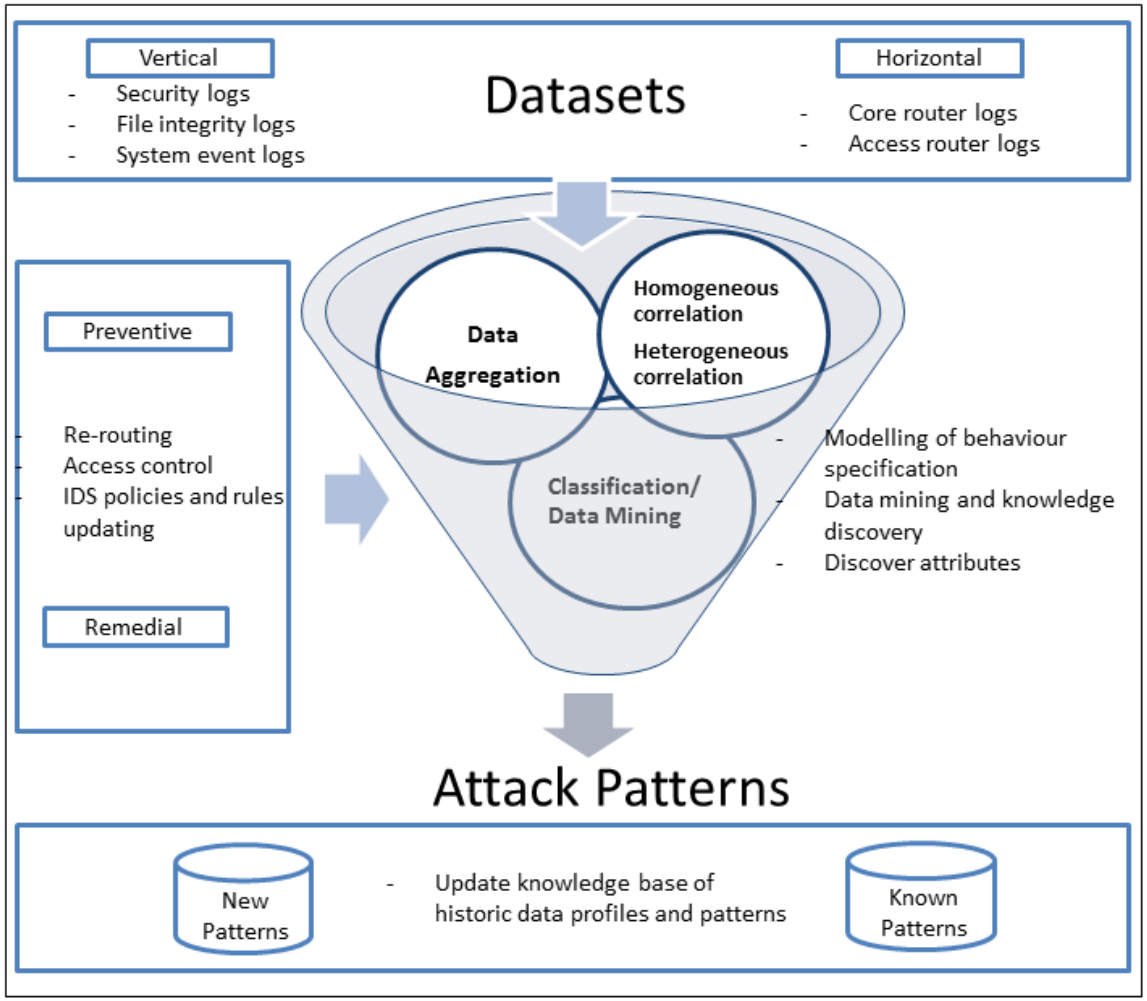
- Attack detection and classification has been investigated by using individual datasets *(Web IDS logs, Net Flow etc)*
- Honeynet traffic analysis: our work is different because we will be using spatial distribution and model the behaviour of attacks found in different correlated events from multiple datasets. *(Honeynet Traffic Analysis)*
- Botnet Tracking: We aim to develop more general model that can be applied to the detection and classification of a range of cyber-attacks as opposed to specialized technique targeted at single type of attack.*(BotMiner)*
- Event Correlation: Currently used for network management and we aim to extend this to other domain such CSA across multiple levels.*(GrIDS, Snort)*
- Darknet: Primarily used to analyse specific phenomenon that are essentially related to worm propagation.*(Team Cymru Darknet, Internet Motion Sensor)*

Datasets

- Detection technologies have matured over time.
- Computer Networks have become more accessible and great deal of monitoring tools providing wealth of information.
- Non Determinism-Events coming from all different independent sources and they are not ordered and analysed together.
- Available in the forms of logs

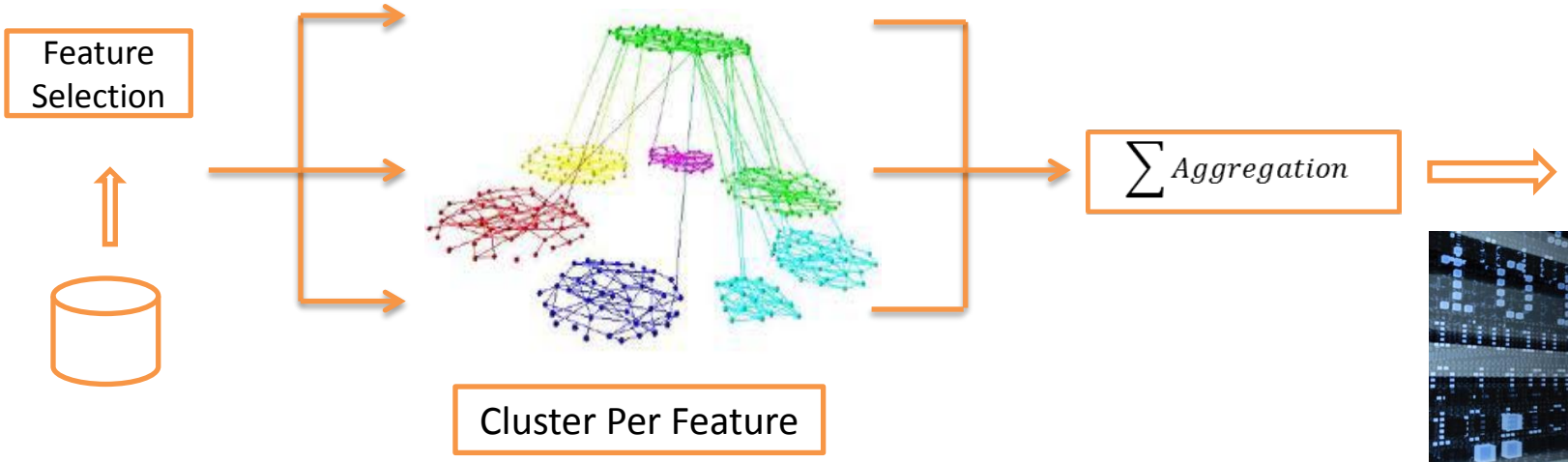


High Level Design



- High level design.
- Aim to extract specific features from datasets .
- Clustering and Classification.
- Aggregation of these clusters
- Store patterns into database.
- Not tailored to one specific dataset.
- Depending what dataset we feed, we aim to get complete insight into attack phenomenon such as attack attribution.

Application of our Model



Cyber Situational Awareness

- Collect real-world attack traces from a number of distributed sensors
  - Network of honeypots = “Honeynet”
- Analysis
  - Collect “attack events” from each sensor
    - Extract relevant information← (with expert-defined features- CAPEC)
      - Using appropriate clustering
    - Synthesizing those pieces of information, to create “concepts” describing the attack phenomena
      - Using Aggregations

Feature Selection and Extraction

- In many data mining procedures, one of the very first steps consists in selecting some key characteristics from data sets.
- Extract and combine features from security data sets such as : *Origins of attack, timing, behaviour etc.*
- Feature selection is the process of identifying, within the raw data set, the most effective subset of characteristics to use in clustering.
- Pattern representation refers to the number of categories, or variables available for each feature to be used by clustering algorithm.
- we characterize each object of the data set according to this set of extracted features

F = {Fk}, k = 1, . . . ,n (e.g., by creating feature vectors for each object);

Choice of Clustering Approach

- Clustering real data sets can be a difficult task, and different clustering methods will probably yield different results.
- Our current analysis indicates that our best bid is for graph based clustering approach and this is motivated this choice due to following reasons:
  - Simplicity to formulate the problem, i.e., by representing the graph by its adjacency matrix (or proximity matrix).
  - Graph-based approach does not require a number of clusters as input.
  - Can be coded in a few lines of any high-level programming language, and it could be easily Implemented in a parallel network, if scalability becomes an issue.
  - Different graphs (obtained for different attack features) can be easily combined using different types of aggregation functions (e.g., averaging functions, fuzzy integrals, etc).

Cluster Ck, is created regarding every feature Fk, based on similarities.

Aggregation/Fusion

- Last component of our model takes advantage of different aggregation techniques., such that multiple attack features can be effectively combined without prior knowledge.

- The aim of this aggregation is to determine, for each pair of events, how likely it is that those events are linked to the same root phenomenon, given the set of relations obtained by assessing different attack features.

- we combine different clusters of attack features using an aggregation function that models the expected behaviour of the phenomena under study.

- Choice of our Aggregation function require further research in order to find detailed attack phenomenon.

Aggregated Function = Aggregated Clusters w.r.t each Fk= ∑

Attack Profiles

- All sources will be clustered into “attack (profiles)” based on certain network characteristics:
  - targeted port sequence
  - No of packets
  - Attack duration
  - Packet payload

Viewpoints

- We need to identify salient features for the creation of meaningful viewpoints
  - Expert defined characteristics for each dimension
- Geo-location
  - Botnet located in specific regions
- IP Blocks
  - Cluster of compromised machines
- Time series
  - Synchronized activities targeting different sensors

Future Work

- Integration of relevant attack features.
- Generation of higher-level concepts describing real world phenomenon.
- Knowledge engineering.
- Due to uncertainty and little prior knowledge of attack events, most suitability of clustering and classification in order to find security problem require further research.
- Implementation of proposed model.

References

❑ResumeNet: Resilience and Survivability for Future Networking: Framework, Mechanisms, and Experimental Evaluation (FP7).

http://www.resumenet.eu/

❑MITRE manages federally funded research and development centres (FFRDCs), partnering with government sponsors to support their crucial operational mission. CAPEC- CybOX is managed by MITRE.

http://www.mitre.org/ ; http://capec.mitre.org/

❑Barnum, S. "Common Attack Pattern Enumeration and Classification (CAPEC) Schema Description", Cigital Inc.

http://capec.mitre.org/documents/documentation/CAPEC_Schema_Description_v1.3.pdf

❑Barnum, S. and Sethi, A. "Introduction to attack patterns" Technical report, U.S. Dept. of Homeland Security.

http://capec.mitre.org/about/documents.html.

❑The Team Cymru. Home page of "The Team Cymru darknet" project.

http://www.team-cymru.org/Services/darknets.html

❑G.Gu, R. Perdisci, J. Zhang and W. Lee. "BotMinier: Clustering Analysis of Network Traffic for Protocol – and Structure Independent Botnet Detection", In proceedings of the 17th USENIX Security symposium, 2008.

❑IETF Policy Framework Working Group

http://WWW.ietf.org/html.charters/policy-charter.html

❑DMTF Information Service Level Agreement (SLA) Working Group

http://www.dmtf.org/info/sla.html

❑Cabinet Office

http://cabinetoffice.gov.uk/resource-library/best-management-practice-portfolio.html

❑Information Technology Infrastructure Library (ITIL):

http://www.itil-officialsite.com/

LANCASTER UNIVERSITY

Attack Patterns Recognition Framework

Noor Shirazi

n.shirazi@lancaster.ac.uk

Thank You

InfoLab21