An Evolutionary Approach of Attack Graph to Attack Tree Conversion

Md. Shariful Haque ,Miclain Keffeler, Travis Atkison

THE UNIVERSITY OF
ALABAMA®

Presentation Outline

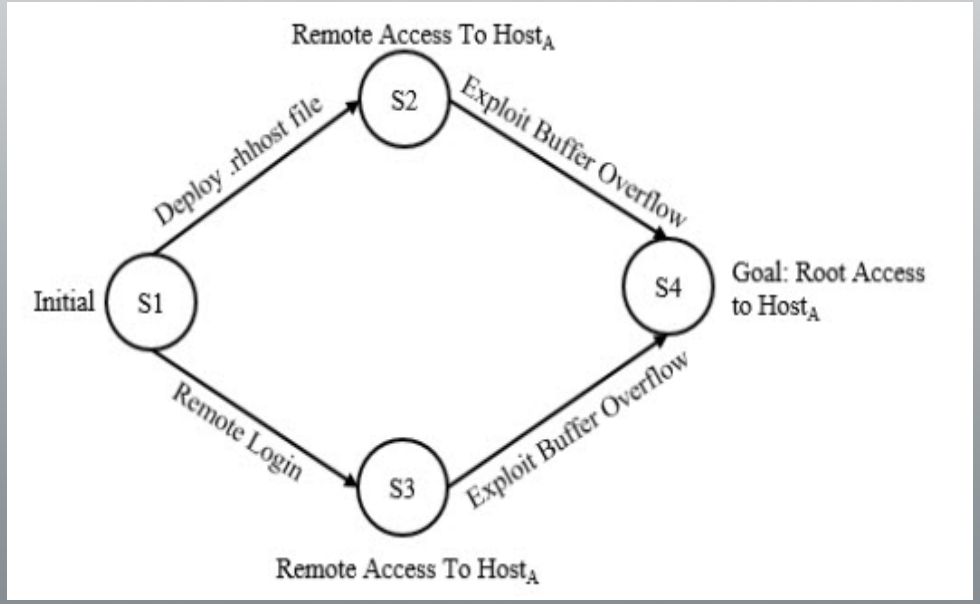
- **Attack Representation Models**
- Research Goal
- Analysis of selected papers
- Open Questions
- Conclusion

Attack Representation Model

- Attack Representation Model
  - Attack scenario recognition system
  - Represents alerts and their relationship or network/host configuration
- Purpose
  - Determine the path of attack
  - Do not describe attack steps
  - Generate attack reports
- Types
  - Attack Graph
  - Attack Tree

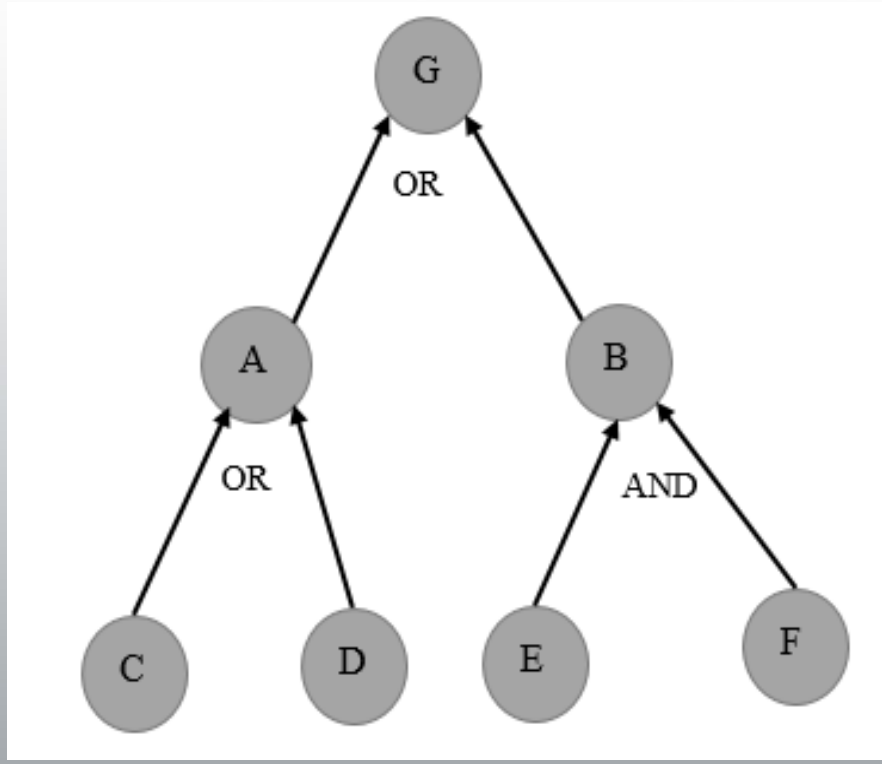
Attack Graph

- Detail view of system security
- Determine attacker's path from initial state to final state
- Nodes represent attack states
- Edges represent transition of different state
- Post and pre-condition are defined between the states.



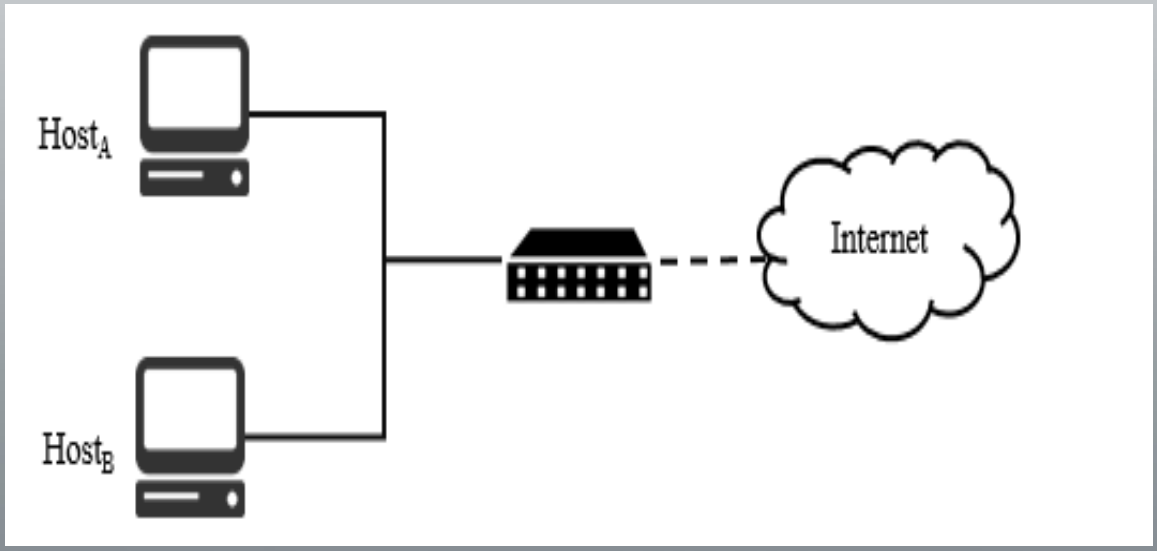
Attack Tree

- Basic Concept
  - Analyze different security threat
  - Identify different paths to goal
  - Design structure defining attacks intention
- Nodes represent attack goals.
  - Root of the tree refers final goal
  - Leaf nodes represents atomic attacks
  - Internal nodes are disjunctive or conjunctive refinement of their lower level nodes.
- Edges represent the path of attack



Example Scenario

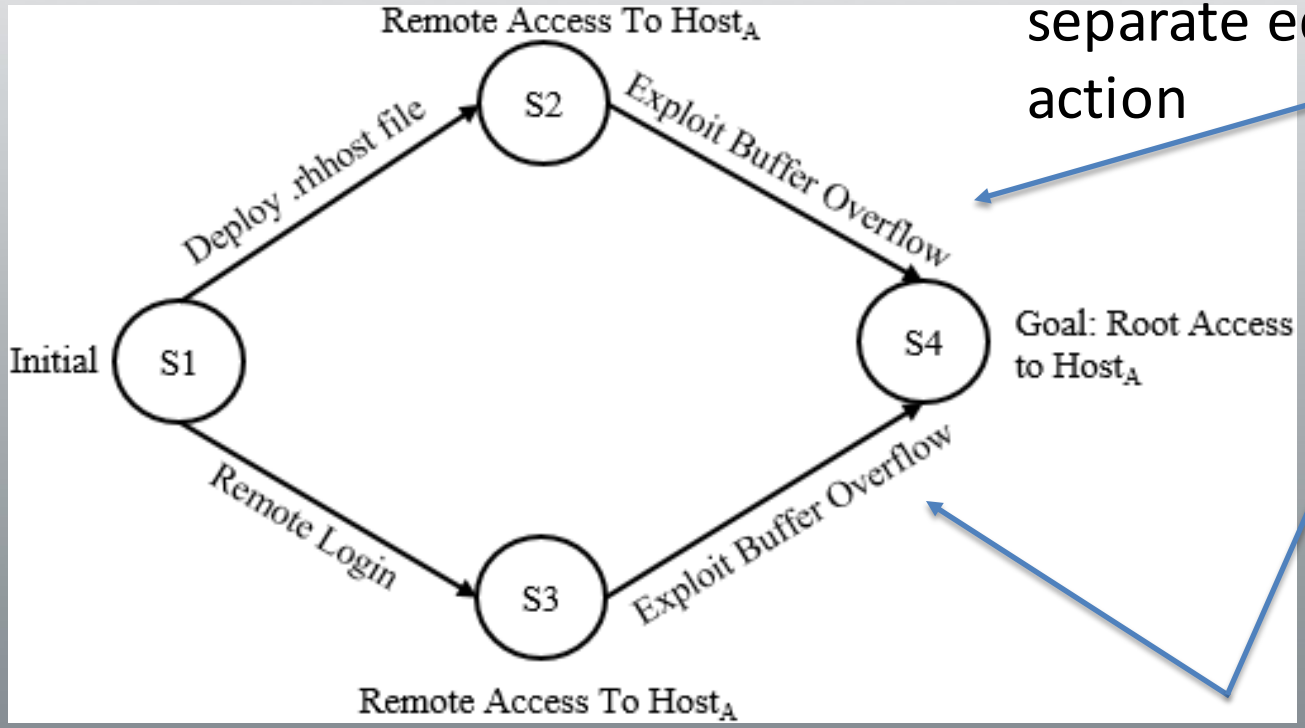
- *HostA* and *HostB* are connected in same network
- User of *HostB* wants root access in *HostA*
- *HostB* can deploy
  - *.rhhost file through FTP* or
  - use *remote login*
- Then *HostB* exploit *buffer overflow* to get root access



Attack Graph of Example Scenario

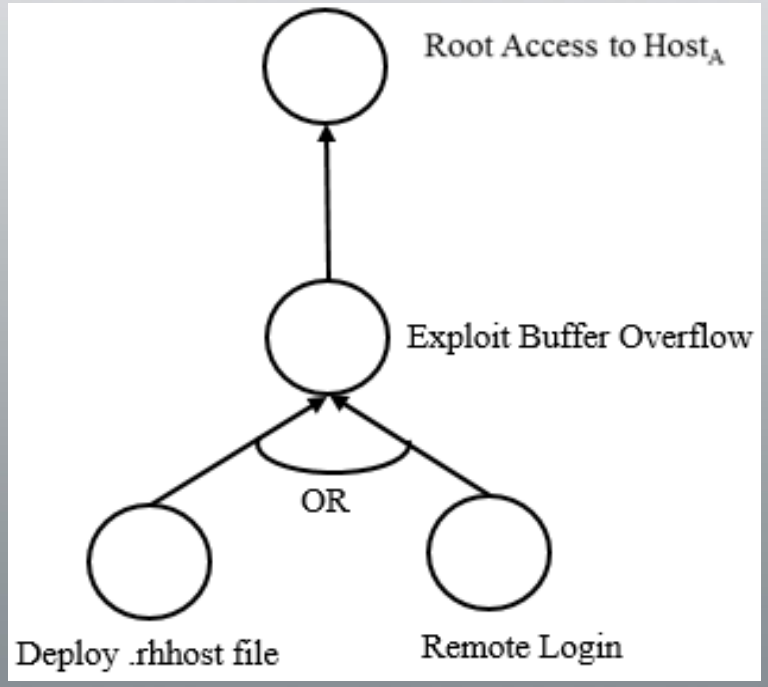
- HostB can choose to do 1 of 2 things
  - Once either has been done, exploiting the buffer overflow allows HostB to achieve their goal

- Notice how S2 and S3 have 2 separate edges to do the same action



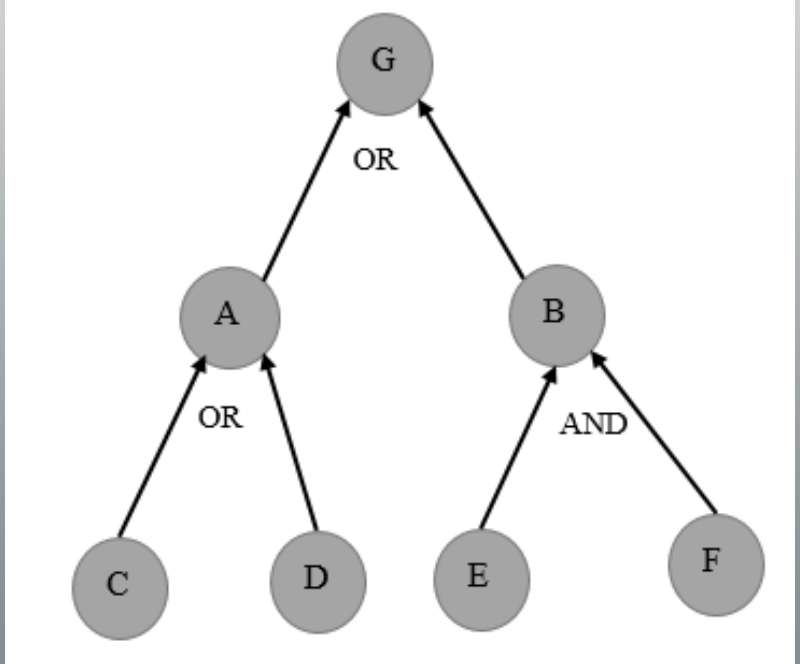
Attack Tree of Example Scenario

- Either a .rhhost file or remote login will allow HostB to exploit buffer overflow
  - By exploiting buffer overflow, HostB gains root access, and achieves their goal



Why Attack Tree?

- Comparatively powerful and methodological approach
- Can be represented through logical expression
  - (C OR D) OR (E AND F)
- Provides more insight(Cost and Benefit analysis) on the attack scenario (Defense Tree and Attack-Defense Tree)
- Helps in determining most probable attack goals (ROI, ROA)

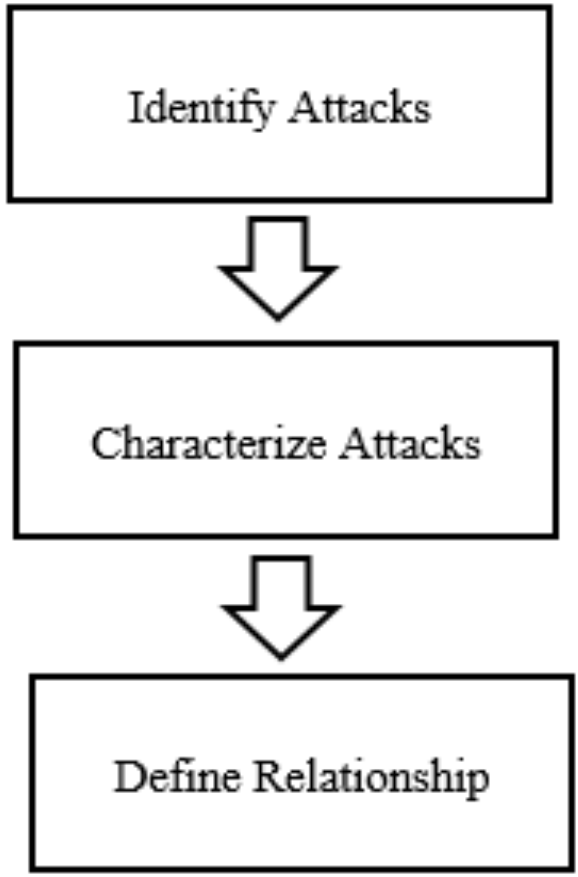


Attack Representation Modeling

- Attack Identification
- Identify Attacks
- Dividing attack into attack sub-goals

- Attack Characterization
- Assign attribute based on:
  - Observed events
  - System States

- Define Relationship
- Temporal (Sequence) relationship
- Attribute-value (Sources) relationship
- Prerequisite (Trigger) relationship

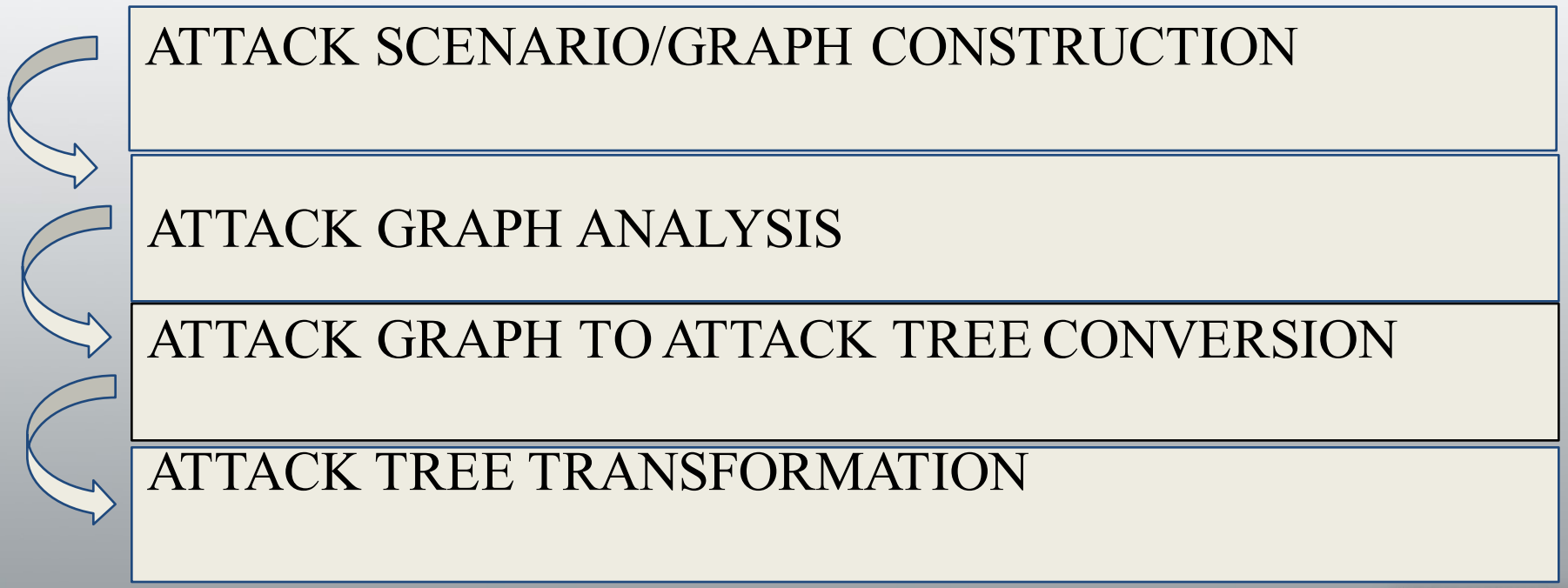


Presentation Outline

- Attack Representation Models
- **Research Goal**
- Analysis of selected papers
- Open Questions
- Conclusion

Research Goal

- Extend current Attack Representation Modeling process



Presentation Outline

- Attack Representation Models
- Research Goal
- **Analysis of selected papers**
- Open Questions
- Conclusion

Selected Papers

- Paper 1: X. Qin and W. Lee, “Statistical causality analysis of infosec alert data,” in International Workshop on Recent Advances in Intrusion Detection. Springer, 2003, pp. 73–93.
- Paper 2: X. Ou, W. F. Boyer, and M. A. McQueen, “A scalable approach to attack graph generation,” in Proceedings of the 13th ACM conference on Computer and communications security. ACM, 2006, pp. 336–345.
- Paper 3: B. Zhu and A. A. Ghorbani, “Alert correlation for extracting attack strategies,” IJ Network Security, vol. 3, no. 3, pp. 244–258, 2006.
- Paper 4: L. Muñoz-González, D. Sgandurra, A. Paudice, and E. C. Lupu, “Efficient attack graph analysis through approximate inference,”arXiv preprint arXiv:1606.07025, 2016.
- Paper 5: J. B. Hong, D. S. Kim, and T. Takaoka, “Scalable attack representation model using logic reduction techniques,” in 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications. IEEE, 2013, pp.404–411.

Papers' Contribution in Research

ATTACK SCENARIO CONSTRUCTION

1. Statistical causality analysis of INFOSEC alert data
2. scalable approach to attack graph generation
3. Alert correlation for extracting attack strategies

ATTACK GRAPH ANALYSIS

4. Efficient attack graph analysis through approximate inference

ATTACK GRAPH TO ATTACK TREE CONVERSION

(Research Goal)

ATTACK TREE TRANSFORMATION

5. Scalable attack representation model using logic reduction techniques

Attack Sources (Paper 1,2, and 3)

- Alerts generated from IDS(Intrusion Detection System)
  - Host / Network Configuration Data

Graph Construction Approaches

- Alert Correlation approaches
- Time-series alert correlation(Paper 1)
- Feature based alert correlation(Paper 3)
- Logical Approach (Paper 2)

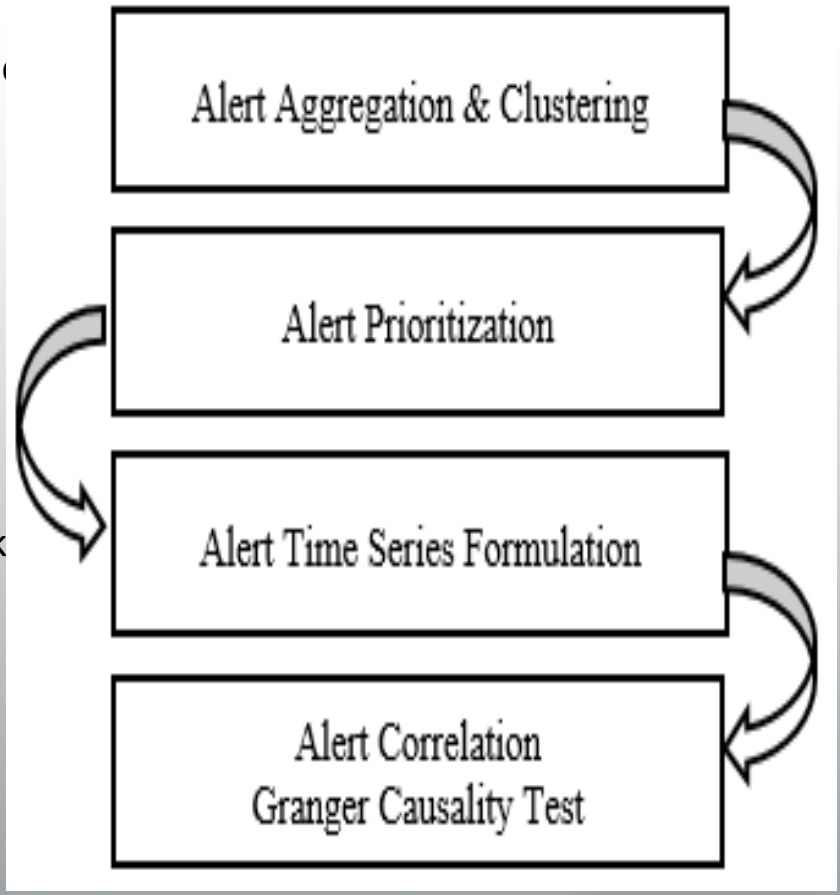
Alert Correlation Approach

- Research Goal
  - Reduce number of alerts
  - Applying causal analysis to define new relationship between alerts
    - I.e Generate Attack Graphs
  - Automatic extraction attack strategy
    - Otherwise time consuming and error prone

Time-Series Analysis (Paper 1)

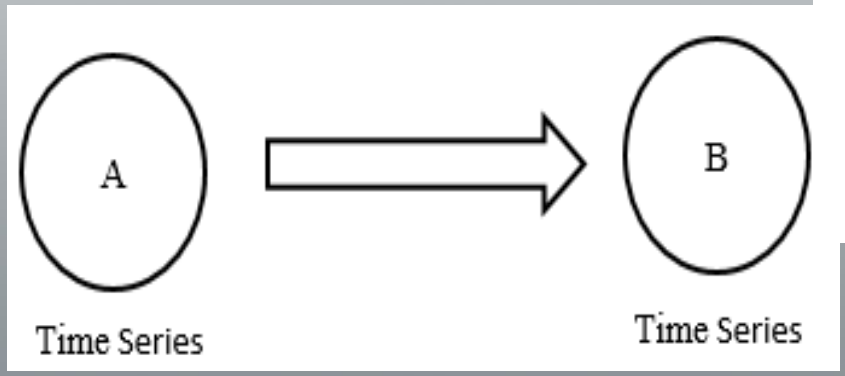
- Alert Aggregation and Clustering
  - Alerts with overlapping attributes are combine
    - Reduced number of alerts
  - Clustering algorithms are applied based on attribute similarity
  - Generates hyper alerts on correlated events

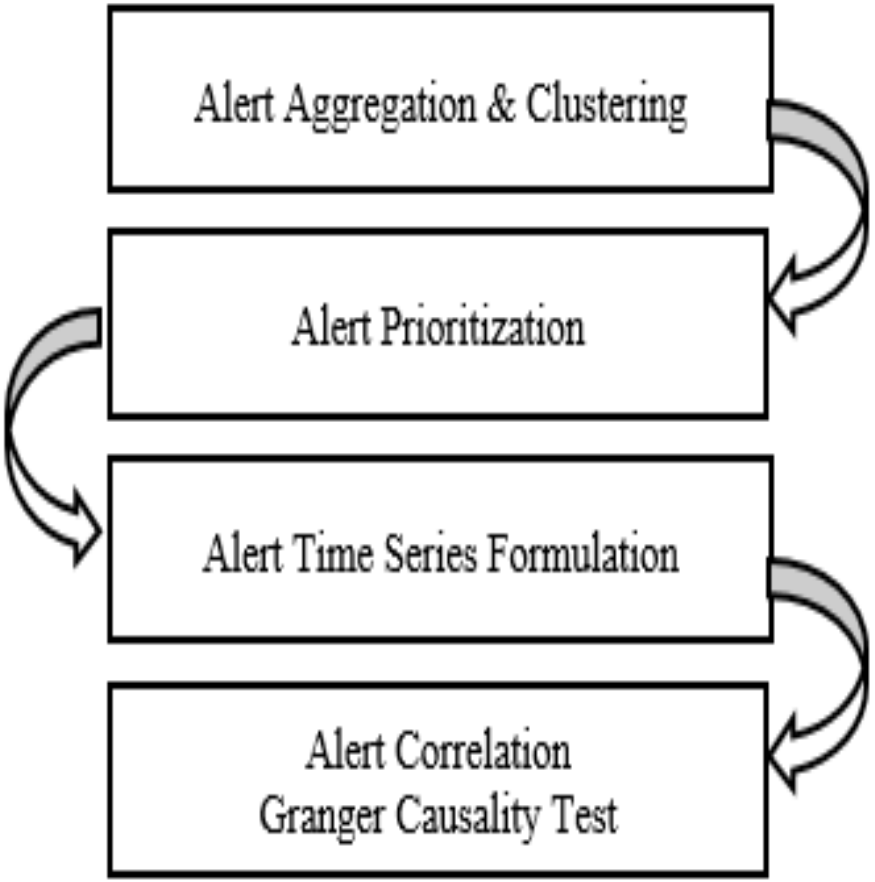
- Alert Prioritization
  - Rank alerts based on pertinence and severity score
  - Severity score calculated based on the network host configuration
  - Use concept of Bayesian Network



Time-Series Analysis (Paper 1)(Cont.)

- Alert Time Series Formulation
  - Generate alert time series variables
  - Arrange hyper alerts in the time series
- Alert Correlation
  - Apply Granger Causality Test on the time series variables
  - Determine pairwise correlation between two hyper alerts.





Feature based Alert Correlation (Paper 3)

- Proposes 2 algorithms
  - Generate hyper alerts from correlated alerts
  - Generate attack graph from attack correlation matrix (ACM)

- Feature Selection
  - 6 features
    - Source IP, Target IP, Target Port
    - Consecutive alerts from Source IP and Target IP
    - **Backward Correlation** and Frequency

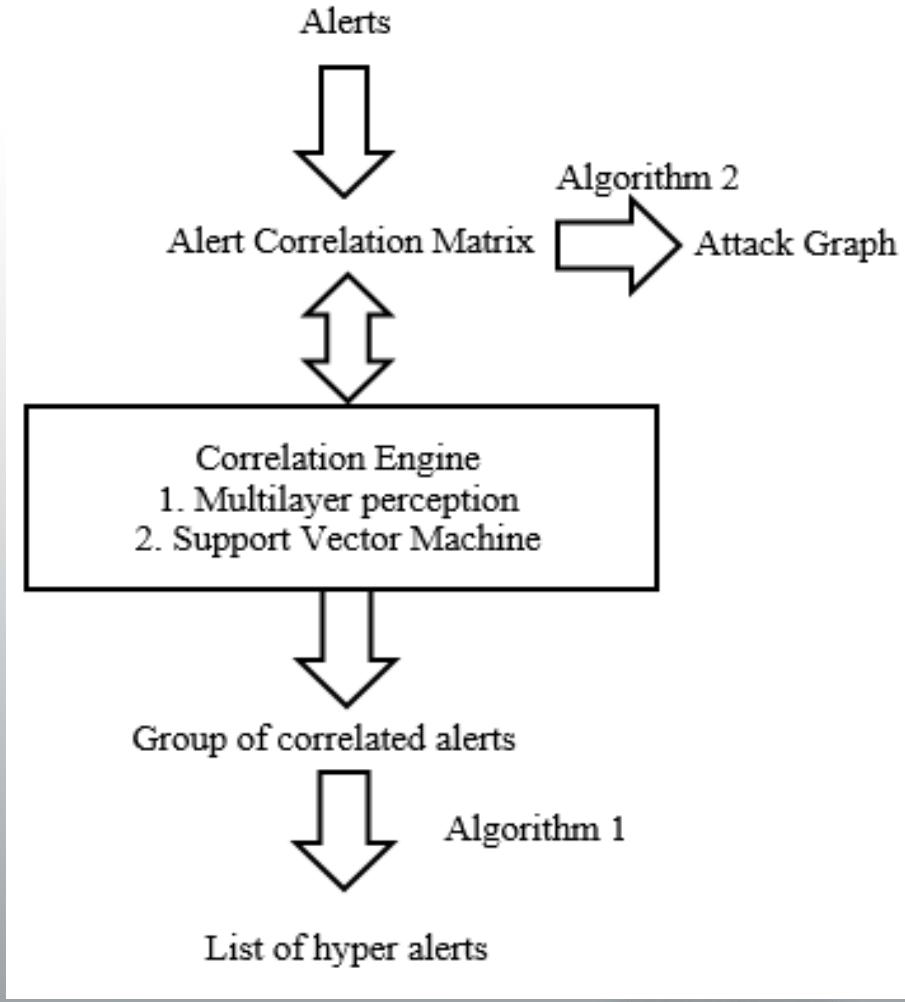
Feature based Alert Correlation (Paper 3) (Cont.)

- Alert Correlation Matrix (ACM)
  - Developed through a supervised learning process
  - Stores temporal relationship between two alerts
  - Each cell contain correlation weight calculated from backward and forward correlation
  - Contribute generating attack graph

	a1	a2	a3
a1	c(a1,a1)	c(a1,a2)	c(a1,a3)
a2	c(a2,a1)	c(a2,a2)	c(a2,a3)
a3	c(a3,a1)	c(a3,a2)	c(a3,a3)

Feature based Alert Correlation (Paper 3) (Cont.)

- Correlation Engine
  - Apply Multi Layer Perception or Support Vector machine (SVM) on the features
  - Calculate probability of alert correlation
  - Use both results for accuracy and precision
  - Result is used to generate hyper alerts.



Algorithm 1 = Generate Hyper Alerts
Algorithm 2 = Generate Attack Graphs from ACM

Logical Approach (Paper 2)

- Graphical representation of logically formulated attack scenario

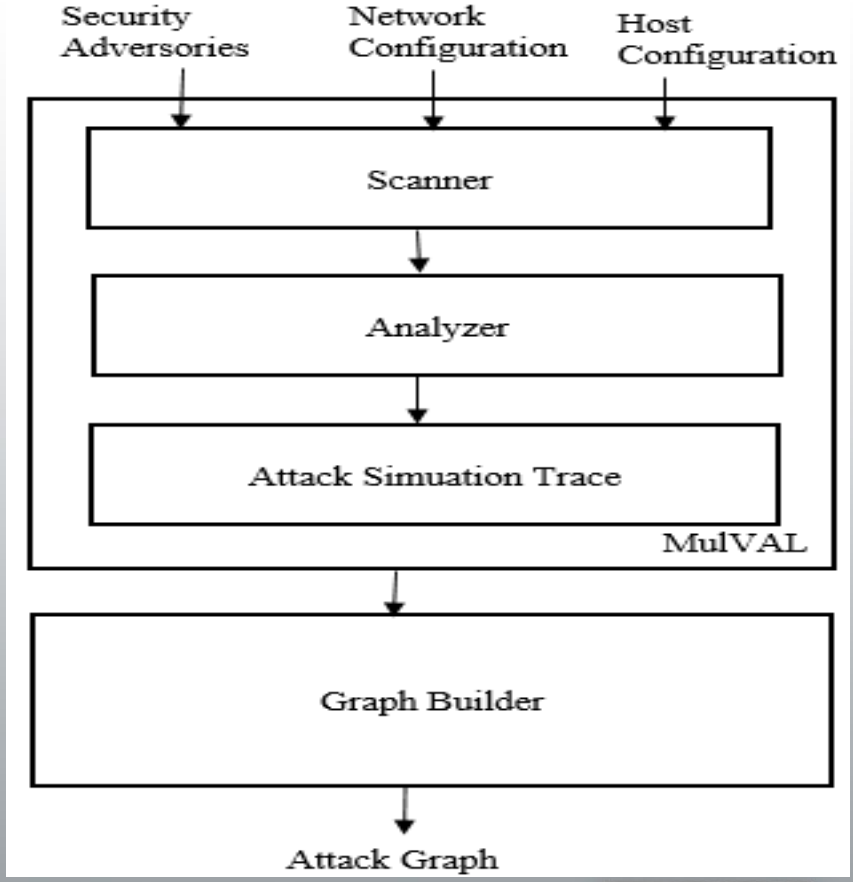
- Use MulVAL (Multihost, Multistage Vulnerability Analysis) tool
- Determine impact of software bugs
- Uses vulnerability specification reported by community

- Research Goal
  - Generate graph in a quadratic time
  - Downside: No customization of input information

Logical Approach (Paper 2) Cont.

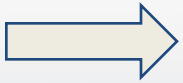
- Add Attack Simulation Trace with MulVAL
- Takes interaction rule as input
- Evaluate the interaction rule
- Generate Attack Simulation Trace from the given rule

- Interaction to Graph Transformation
  - Graph builder takes simulation trace as input
- Evaluate each item of the trace
- Generate Attack graph

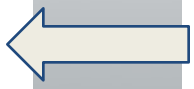


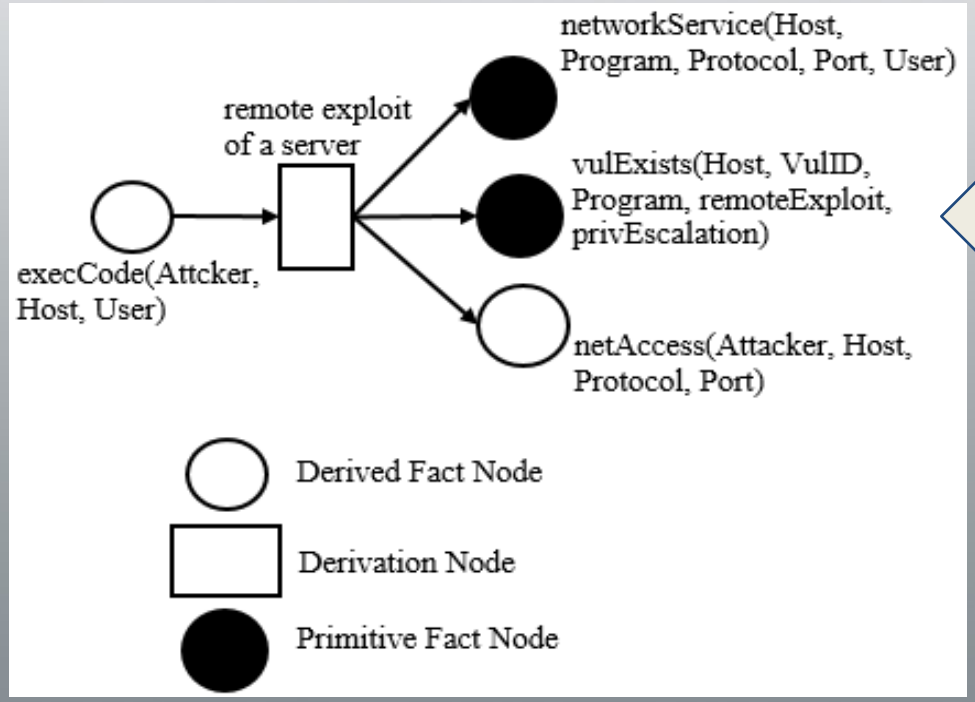
Logic-based Attack Scenario Construction Cont.

```
execCode(Attcker, Host, User) :-

      networkService(Host, Program,
           Protocol, Port, User)

      vulExists(Host, VulID, Program,
          remoteExploit, privEscalation)

      netAccess(Attacker, Host,
           Protocol, Port)
```



```
execCode(Attcker, Host, User) :-

      networkService(Host, Program,
           Protocol, Port, User)

      vulExists(Host, VulID, Program,
          remoteExploit, privEscalation)

      netAccess(Attacker, Host,
           Protocol, Port)

      assert_trace(because(
      'remote exploit of a server program',
      execCode(Attcker, Host, User)
      [networkService(Host, Program,
           Protocol, Port, User)

      vulExists(Host, VulID, Program,
          remoteExploit, privEscalation)

      netAccess(Attacker, Host,
           Protocol, Port)]))
```





Papers' Contribution in Research

ATTACK SCENARIO CONSTRUCTION

1. Statistical causality analysis of INFOSEC alert data
2. scalable approach to attack graph generation
3. Alert correlation for extracting attack strategies

ATTACK GRAPH ANALYSIS

4. Efficient attack graph analysis through approximate inference

ATTACK GRAPH TO ATTACK TREE CONVERSION

(Research Goal)

ATTACK TREE TRANSFORMATION

5. Scalable attack representation model using logic reduction techniques

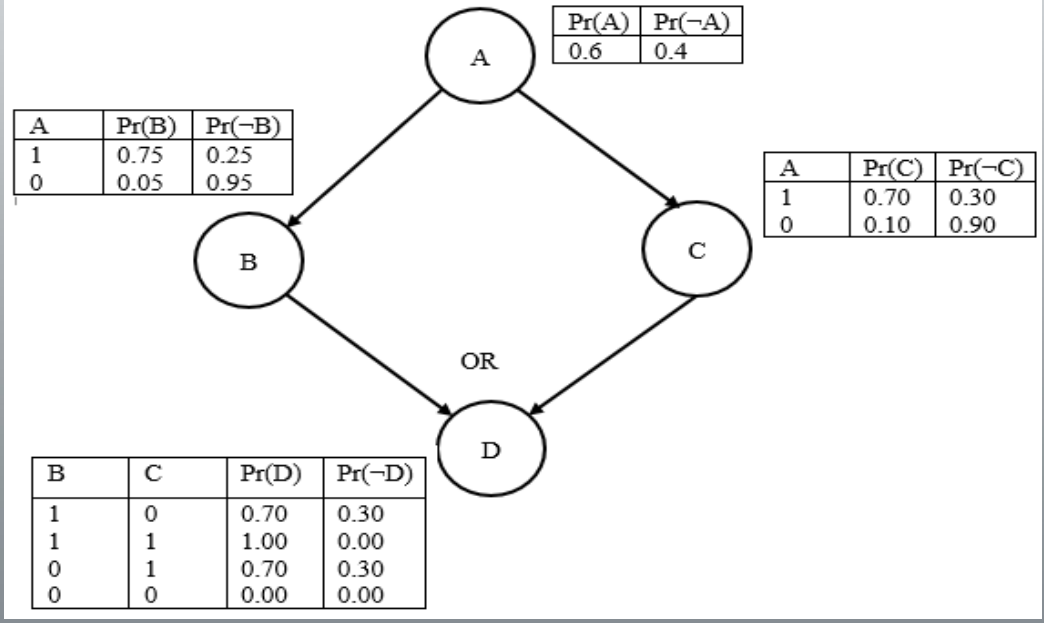
Research Motivation

- Apply Bayesian Inference to discover vulnerable points in a given attack graph

- Research Goal
  - Enable easy interpretation Interdependencies between attacks
  - Modeling uncertainty of attacker's behavior
  - Enable analysis for larger network

Attack Correlation Approaches

- Developed model based on Bayesian Attack Graph(BAG)
- BAGs are built on network topology or analysis on alerts
- CPTs are calculated based on CVSS (Common Vulnerability Scoring System)
- Static and Dynamic Analysis
- Static analysis is conducted based on given graph and CVSS (Common Vulnerability Scoring System)
- Attacks are recomputed in Dynamic analysis



Papers' Contribution in Research

ATTACK SCENARIO CONSTRUCTION

1. Statistical causality analysis of INFOSEC alert data
2. scalable approach to attack graph generation
3. Alert correlation for extracting attack strategies

ATTACK GRAPH ANALYSIS

4. Efficient attack graph analysis through approximate inference

ATTACK GRAPH TO ATTACK TREE CONVERSION

(Research Goal)

ATTACK TREE TRANSFORMATION

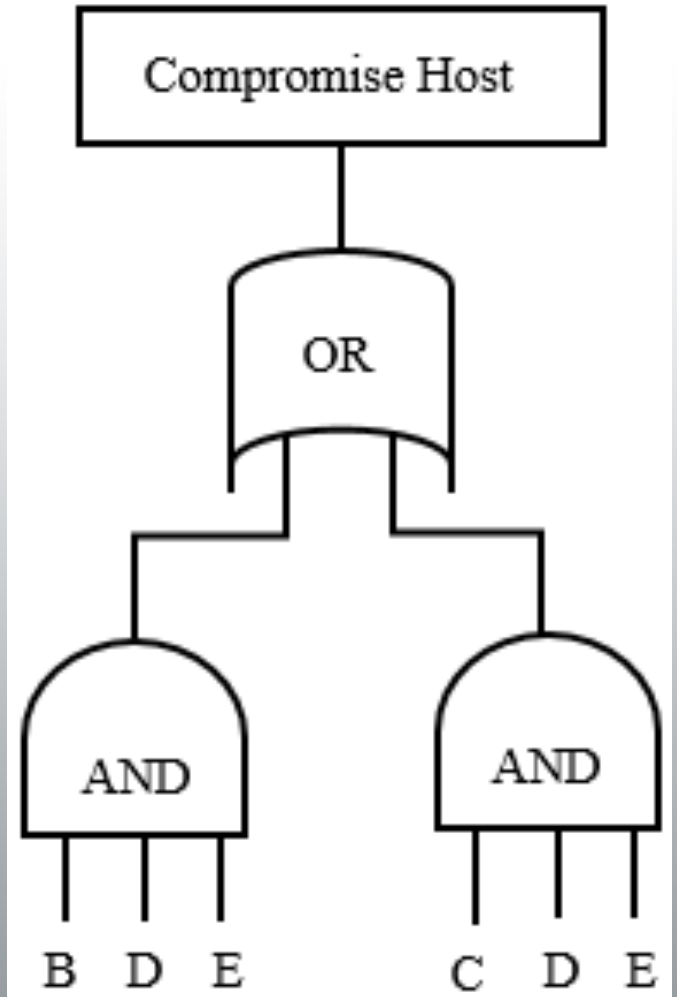
5. Scalable attack representation model using logic reduction techniques

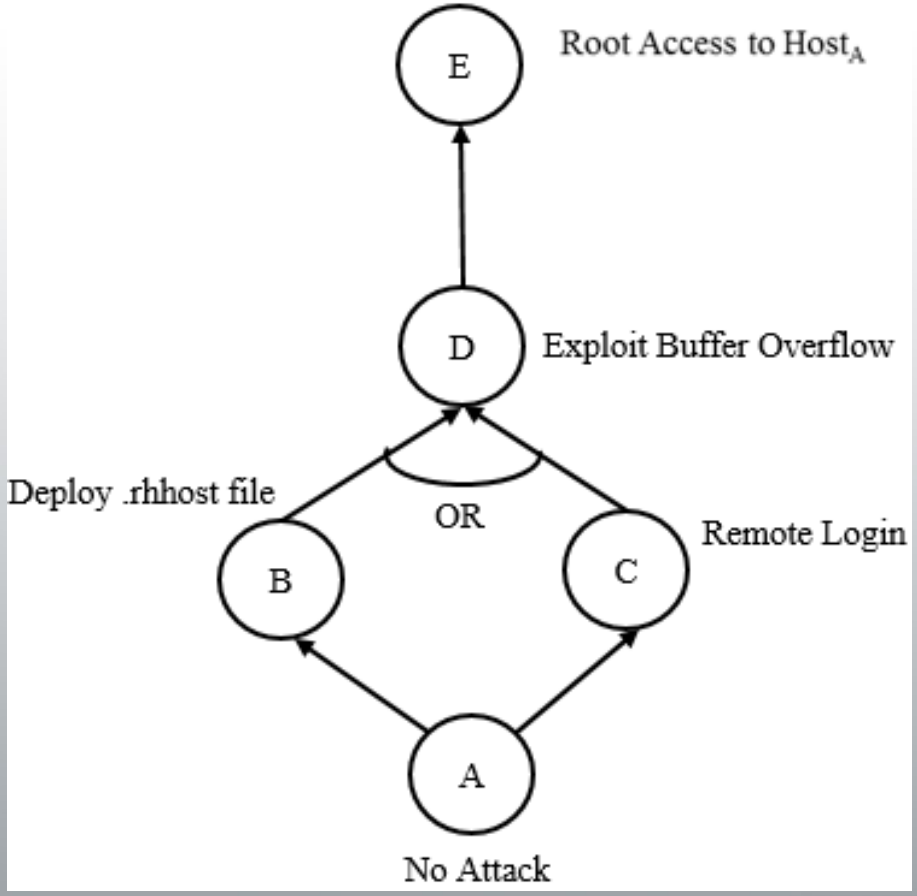
Research Motivation

- Research Goal
- Automatic transformation of attack tree
- Reduce the size of attack tree

- Proposed algorithm
  - Full Path Calculation(FPC)
  - Incremental Path Calculation(IPC)

Logical Attack Tree Reduction

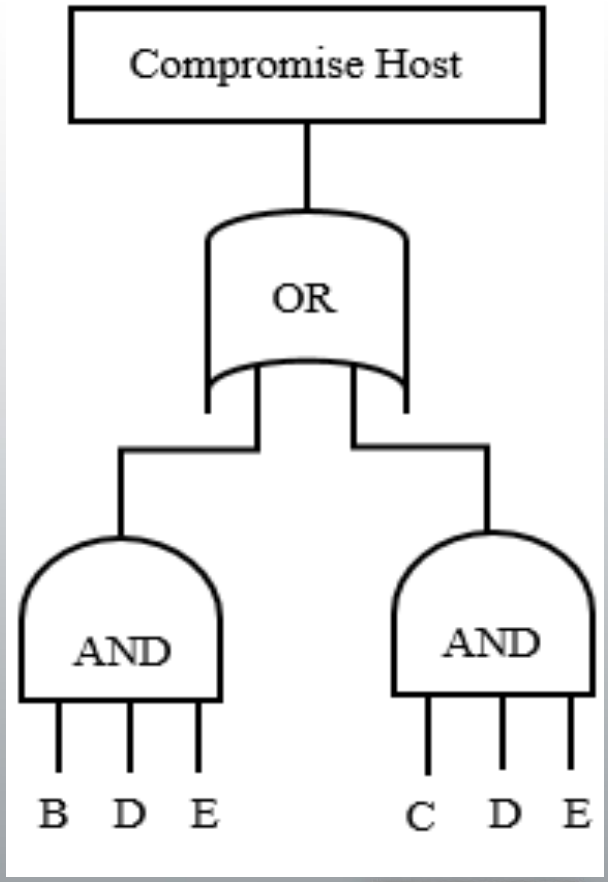




Nodes	A	B	C	D	E
Neighbors	B+C	D	D	E	-

Research Motivation

- Full Path Calculation
  - Represent the attack tree in logical expression
  - Eliminate attack sequence
  - Iteratively factorize the common element
- Incremental Path Calculation
  - Create a reachability table
  - For each node, check table and decide the next node to reach
  - Repeated nodes are eliminated to avoid cycle.

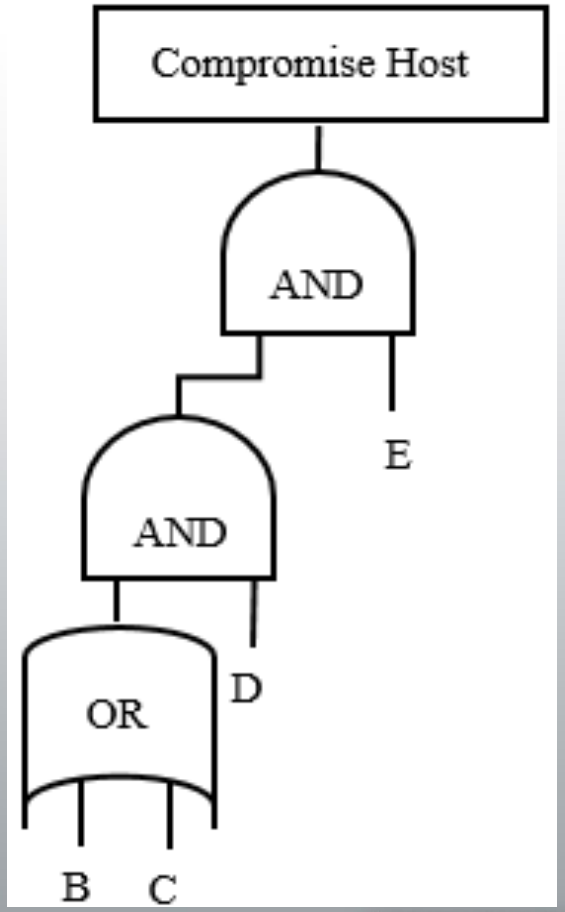


Research Motivation

- Full Path Calculation
  - BDE + CDE
  - BDE + CDE = DE(B + C)
  - = E(D(B + C))

Nodes	A	B	C	D	E
Neighbors	B+C	D	D	E	-

- Incremental Path Calculation
  - B + C
  - => B(D) + C(D)
  - => B(D)(E) + C(D)(E)
  - => E(D(B + C))



Paper	Contribution	Approach	Limitation
Paper 1	Attack Scenario Construction	Time-series Alert correlation	• CPTs in priority calculation depends on prior knowledge of the system administrator, not adaptive • Background alert raises chance of false causality alerts
Paper 2		Feature based alert correlation	• Supervised learning require manually generated and labelled training - increase chance of errors in training phase.
Paper 3		Logical Approach	• Cannot produce attack graph if a new attack is not expressed in propositional formula
Paper 4	Attack Graph Analysis	Approximate Inference on Bayesian Attack Graph	• The LBP algorithm cannot always guarantee convergence
Paper 5	Attack Tree Transformation	Logic Reduction	• With large attack tree, FPC suffers from exponential generation of nodes and IPC suffers from inefficient memory allocation.

Open Questions

- Can we develop a unified attack tree construction model?
  - Logical approach
  - Alert Correlation approach
  - Reducing System Administrator's effort
- Is it possible to extend the attack construction model in cloud environment?
  - Large volume of alert from multiple IDS and multiple layer of cloud infrastructure
- Can we introduce Attack Tree in Real Time Attack Analysis?
  - Reduce false positive and False negative alert data
  - Attack Representation Models have never used in this area

This Presentation has been made available online:

Thank You

Questions?