Background

- Intrusion detection forms the second line of defense
- But current Intrusion Detection Systems (IDSs) are still not perfect
  - False negatives (missed attacks)
  - False positives (false alerts)
  - Large numbers of alerts
- It is challenging to understand
  - What are the intrusions
  - What the intruders have done
Background (Cont’d)

- A common problem of most existing alert correlation methods
  - Cannot handle missed attacks.
- Abductive correlation [Cuppens and Miege 2002]
  - Hypothesis of missed attacks are guided by known attack scenarios specified in LABMDA

Our Contributions

- A framework to **automatically** hypothesize and reason about missed attacks based on **knowledge about individual attacks**
  - Hypotheses of missed attacks
  - Inference of attack attributes
  - Validation of hypothesized attacks
  - Consolidation of hypothesized attacks
- Prototype implementation and initial experimental evaluation of this approach
Outline

• A series of techniques to hypothesize and reason about attacks missed by IDSs
  – A naïve approach
  – Type graph guided hypothesis
  – Inferring the attribute values of hypothesized attacks
  – Validation/pruning through raw audit data
  – Consolidation of hypothesized attacks

• Experimental results
• Conclusion and future work

Previous Work: Correlation Based on Prerequisites and Consequences of Attacks

• Model
  – Hyper-alert type: represent our knowledge as prerequisites and consequences of attacks
  – Hyper-alert: an instance of a hyper-alert type; instantiated from IDS alerts
  – An earlier hyper-alert prepares for a later one if the former makes the later easier to be successful

\[
C(h_1) = \{\text{VulnerableSadmind}(152.1.19.5)\} \quad P(h_2) = \{\text{ExistHost}(152.1.19.5), \text{VulnerableSadmind}(152.1.19.5)\}
\]
Previous Work (Cont’d)

- An example hyper-alert correlation graph
  - Would be split into multiple graphs if critical attacks are missed by the IDSs

![Hyper-Alert Correlation Graph]

Naïve Approach

- Integrate complementary correlation methods
  - Clustering correlation methods
    - Based on the similarity between alert attribute values
    - May still cluster related alerts even if critical attacks are missed
    - Unable to discover the causal relationships between alerts
  - Causal correlation methods
    - Based on prerequisites and consequences of attacks
    - May discover the causal relationships between alerts
    - Don’t work if critical attack steps are missed
Naïve Approach (Cont’d)

• Put multiple attack scenarios together if the clustering correlation method says they are similar
• But …
  – How about the possible causal relationships between these alerts?

Naïve Approach (Cont’d)

• Given attack types T and T’, T may prepare for T’ if
  – Informally, a type T attack may contribute to a type T’ attack

• T may indirectly prepare for T’ if
  – Informally, a type T attack may indirectly contribute to a type T’ attack through other intermediate attacks
Naïve Approach (Cont’d)

- *May-indirectly-prepare-for* relations can help hypothesize missed attacks
  - More complete attack scenarios

```
ICMP_PING_NMAP1
  -->
SCAN_NMAP_TCP2
  -->
Rsh3
  -->
Mstream_Zombie4
```

Type Graph Guided Approach

- *May-prepare-for* and *may-indirectly-prepare-for* relations give us more opportunities
  - We may use them to hypothesize about what have been missed by the IDSs

```
SCAN_NMAP_TCP
  -->
T

FTP_Glob_Expansion
T

Rsh
T'
```
Type Graph Guided Approach (Cont’d)

A type graph over a set of known attacks

- Note: A type graph is computed automatically over a given set of attack types

A Type Graph Guided Approach (Cont’d)

Hypotheses of Missed Attacks
Reasoning about the Hypotheses

• How do we know these are good hypotheses?
• Equality constraint
  – Represent dependency between adjacent attack steps
  – An equality constraint for two hyper-alert types (or attack types) \( T_1 \) and \( T_2 \) is a conjunction of equalities \( u_1 = v_1 \land \ldots \land u_n = v_n \),
  – where \( u_1, \ldots, u_n \) are attributes of \( T_1 \), and \( v_1, \ldots, v_n \) are attributes of \( T_2 \),
  – such that if a type \( T_1 \) hyper-alert \( h_1 \) and a type \( T_2 \) hyper-alert \( h_2 \) satisfy this condition, then \( h_1 \) prepares for \( h_2 \)

\[ T_1 \text{ prepares for } T_2 \] if and only if they satisfy at least one equality constraint

\[
\text{ExistService}(\text{DestIP}, \text{DestPort}) \quad \text{ExistService}(\text{DestIP}, \text{DestPort})
\]

\[ T_1.\text{DestIP}=T_2.\text{DestIP} \land T_1.\text{DestPort}=T_2.\text{DestPort} \]

Reasoning about The Hypotheses (Cont’d)

• Indirect equality constraint

\[
\{n2.\text{DestIP}=n4.\text{DestIP}
\land n2.\text{DestPort}=n4.\text{DestPort}\} \quad \{n4.\text{DestIP}=n5.\text{DestIP}
\land n4.\text{DestPort}=n5.\text{SrcIP}\}
\]

\[ n2.\text{DestIP}=n5.\text{DestIP} \text{ or } n2.\text{DestIP}=n5.\text{SrcIP} \]

• Use indirect equality constraints to verify the hypothesized indirect causal relationships
Reasoning about Missed Attacks (Cont’d)

- Pre-computation of the indirect equality constraints
  - For each pair of attack types
    - For each path between these two attack types
      - For each combination of (direct) equality constraints between adjacent attack types in the path
        » Derive the equality conditions
    - Store the indirect equality constraint in a constraint matrix
- Check against the pre-computed indirect constraints when hypothesizing missed attacks

Infer Attribute Values of Hypothesized Attacks

- **SCAN_NMAP_TCP2**
  - DestIP = 10.10.10.2; DestPort = 21
- **Rsh3**
  - DestIP = 10.10.10.2
- **FTP_Glob_Expansion6**
  - DestIP = 10.10.10.2; DestPort = 21
  - Timestamp in [SCAN_NMAP_TCP2.end_time, Rsh3.begin_time]
Validating and Pruning via Raw Audit Data

- Filtering conditions for hypothesized attacks
  - Prior knowledge
    - protocol = ftp (FTP_Glob_Expansion)
  - Inferred attribute values
    - protocol = ftp \^ DestIP = 10.10.10.2
  - Possible range of Timestamp
    - protocol = ftp \^ DestIP = 10.10.10.2 \^ TS in [11:00AM, 11:10AM]

Again, have we hypothesized the right attacks?

An Example

- There doesn’t exist ftp traffic between SCAN_NMAP_TCP2 and Rsh3.

An Example

- There doesn’t exist ftp traffic between SCAN_NMAP_TCP2 and Rsh3.
Consolidate Hypothesized Attacks

- One missed attack may be hypothesized multiple times through different related alerts
- There may have been multiple instances of the missed attack, but
  - Introduce complexity into analysis
  
```plaintext
Sadmind_Ping1 ----> Sadmind_BufferOverflow4 ----> Rsh2
Sadmind_BufferOverflow5 ----> Rsh3
```

Consolidate Hypothesized Attacks (Cont’d)

- Consolidate two hypothesized attacks, if they possibly refer to the same attack
  - They have the same type
  - Their inferred attribute values do not conflict
  - The ranges of their timestamps overlap
  
```plaintext
Sadmind_Ping1 ----> Sadmind_BufferOverflow4 ----> Rsh2
Sadmind_BufferOverflow5 ----> Rsh3
```
Experiments

- Data Set
  - 2000 DARPA ID evaluation dataset: LLDOS 1.0
- IDS
  - ISS RealSecure Network Sensor 6.0
- Clustering
  - Same destination IP address
- Causal correlation
  - NCSU Intrusion Alert Correlator (Version 0.2)
- Network audit data process
  - Ethereal 0.9.14
- Type graph
  - All attacks detected by the IDS
- Drop all Sadmind_Amslverify_Overflow alerts
Conclusion and Future Work

- Integrates two complementary intrusion alert correlation methods
- Build attack scenarios based on type graphs and (indirect) equality constraints:
  - Hypothesize and reason about missed attacks
  - Infer about attack attribute values
  - Validate and consolidate hypothesized attacks
- Future Work
  - Additional techniques to validate and reason about hypothesized attacks
  - Large scale experiments
  - Quantitative evaluation