TIAA – Toolkit for Intrusion Alert Analysis

* Tất cả mọi cảnh báo dạng thô và báo cáo sẽ được lưu trong database.
* TIAA chuyển từ cảnh báo dạng thô sang hyper-alert thông qua knowledgebase. KB này có thể tùy chỉnh đc.
* KB còn là từ điển chứa các thông tin về hyper-alert type. KB còn chứa các mối quan hệ giữa các predicate.
* Hyper-alert được tập hợp thành các hyper-alert collection dùng để làm đầu vào cho việc phân tích cũng như đầu ra cho các báo cáo.
* Các chức năng phân tích : bao gồm 2 loại: chức năng tạo hyper-alert và chức năng trích xuất thông tin
  + Chức năng tạo hyper-alert
    - Focused Analysis: tập trung lấy các alert thỏa mãng 1 ràng buộc nhất định. Ví dụ như IP= 192.168.1.2
    - Clustering Analysis: chia 1 tập các hyper-alert có tính chất giống nhau vào các nhóm dựa trên 1 ràng buộc trên các thuộc tính. Ví dụ: chia tập các hyper-alert thành các tập có cùng 2 thuộc tính là SourceIP và DestIP
    - Aggregation/Disaggregation Analysis:
      * Aggregation: tất cả alert cùng 1 loại xuất hiện trong 1 khoảng thời gian sẽ được gọp lại tạo thành 1 hyper-alert chung (của cùng 1 hyper-alert type). (đọc paper để hiểu rõ)
      * Disaggregation: ngược lại của aggregation
      * Nhận xét: Clustering trên các alert trong cùng 1 thời điểm ??? Aggregation trong 1 khoảng thời gian
    - Link analysis: dùng để phân tích mối quan hệ 2 chiều giữa các thuộc tính của hyper-alert
      * Link analysis sẽ phân tích dựa trên các cảnh báo thô
      * Có 2 chế độ là count mode và weighted mode.
        + Count mode: đếm số lần xuất hiện. Càng nhiều thì hiển thị càng rõ (xem hình)
        + Weighted mode: có 1 số thuộc tính sẽ quan trọng hơn.
      * Nếu các thuộc tính có cùng domain thì có thể phân tích dựa trên 2 cách:
        + Uni-domain: ví dụ 1 địa chỉ IP xuất hiện cả ở src và dest thì sẽ được xem là 1
        + Dual-domain: ví dụ 1 địa chỉ IP xuất hiện cả ở src và dest thì được xem là 2 địa chỉ IP khác nhau.
    - Association analysis: tìm kiếm sự xuất hiện đồng thời của các thuộc tính ở các hyper-alert khác nhau. Ví dụ: có nhiều hyper-alert có chứa srcIP là 172.16.1.10 đến destIP/port:172.16.1.31/80. Có nhiều kết quả như vậy nhưng chỉ những kết quả nào vượt qua 1 mức (threshold) t thì mới được cảnh báo. Có 2 phương pháp: count mode và weighted mode.
      * Count mode: t% alert trong tập hợp có thuộc tính X.
      * Weighted mode: t% alert trong tập hợp có thuộc tính X với t được tính toán có trọng số.
    - Attack strategy extraction: dùng để tìm các cuộc tấn công dựa trên đồ thị.
    - Missed attack hypotheses: dùng để xác định các alert mà IDS bỏ sót. Thướng dựa vào ràng buộc giống nhau. 2 alert xuất hiện liên tiếp nhau thường có cùng 1 giá trị thuộc tính (ví dụ như cùng destIP)
* Knowledge base
  + Không có phép toán negation trên các pre-con
  + Chứa các Predicates, Implications, HyperAlertTypes
  + Predicates
    - <Predicate Name="MSSSL">
    - <Arg id="520" Pos="1" Attr="varchar(15)"></Arg>
    - <Arg id="521" Pos="2" Attr="int"></Arg>
    - </Predicate>
    - Mỗi Arg có 1 ID duy nhất trên tất cả predicates
    - Pos xác định vị trí của arg. Ở đây char là arg đầu tiên, int là arg thứ 2.
    - Các predicate này sẽ đc dùng cho phần Implications và đc lưu trong database
  + Implications: chỉ mối quan hệ giữa các predicate thông qua ImplyingName và ImpliedName.
    - Có 2 loại: normal và phantom
      * Phantom chỉ những loại implication ko rõ với expert nhưng rõ với attacker.
      * Sử dụng phatom implication sẽ correlate đc nhiều hơn nhưng ngược lại cũng tạo nhiều false positive hơn. (trong paper cũng như report của Thi có nói phần này)
  + Hyper-alert Type
    - Chứa Fact Protocol Prerequisite Consequence
    - Fact là thành phần bắt buộc phải có.
    - Prerequisite và Consequence là optional
    - Protocol chứa thông tin về protocol mà cuộc tấn công tiến hành thông qua nó.
    - Các Fact Pre Con Protocol sẽ được lưu vào database thành từng table.

1 số nhận xét:

* Có thể dùng Ontology thay cho Implication ?
* Các chức năng analysis của nó mình có thể áp dụng. Các chức năng analysis này ko hoàn toàn mang ý nghĩa aggregation hay correlation như các paper đã đọc.
* Hệ thống này chưa có chức năng ngăn chặn.