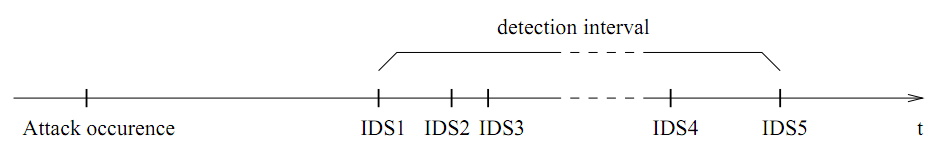
# Using an Intrusion Detection Alert Similarity Operator to Aggregate and Fuse Alerts

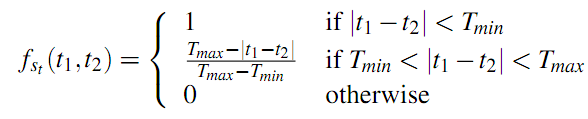
1. Introduction
   * Tại sao phải dùng nhiều IDS kết hợp?
     + Dùng 1 sẽ giảm số lượng tấn công có thể phát hiện đc
     + Dùng nhiều tăng số lượng cảnh báo
   * Tại sao phải kết hợp alert?
     + Lượng alert là rất lớn không phù hợp với human
     + Lượng alert nhiều nhưng thông tin thì ít
   * Kết hợp alert
     + Similarity scale từ [0;1]
2. Các công việc liên quan
   * Probabilistic Approach
   * [Julisch]
     + 1 dạng clustering
     + Cho rằng alert root cause là do config sai
   * Debar Wespi
     + Duplicate definition: 1 alert phải đc thấy nếu 1 alert trc đ1o xuất hiện
3. Độ tương tự của các Alert
   * Mô hình alert
     + Có nhiều cách mô hình alert
       - Dùng feature(ko rõ) [VS01]
       - Dùng 1 tập các attr (kiểu IDMEF)
         * Dùng IDMEF vì khả năng sử dụng với nhiều IDS
     + 1 tập E các event type ???
       - Có thể tạo các alert
       - 1 alert là 1 instance của 1 element của E
       - Associate an event and a set of properties linked to this event to an alert
   * So sánh các thuộc tính
     + Có 4 loại thuộc tính
       - Catergorical: IP,port
       - Numerical:counter, packet size
       - Temporal: numerical nhưng có tính tuần hoàn
       - Chuỗi string
     + So sánh thời gian (Time class: có 3 loại trong IDMEF)
       - 1 alert có thể đc detect bởi nhiều IDS. Mỗi IDS có khả năng xử lí khác nhau nên detect time cũng khác nhau.



* + - * + Sớm nhất
        + Lâu nhất

Cái này là train ???

* + - * Xét 2 trường hợp để xác định 2 temporal window cho từng traffic
        + Ko traffic => Tmin
        + Neutral trafiic => Tmax
      * Công thức



* + - * + Chỉ áp dụng đc nếu các IDS đc đồng bộ
    - So sánh IP Address
      * Dựa vào topology của mạng: đọc [Jul03] để rõ
        + Nếu 2 IP cũng subnet thì similarity cao hơn
        + Dựa vào cây IP (IP address taxonomy)
      * Cách tính
        + Dựa vào khoảng cách của 2 IP trong taxonomy tree
        + Similarity value= inverse of the maximal distance minus the calculated distance ???
    - So sánh dữ liệu của User (User data) (UserID của User class)
      * Do UserID chứa cả group và user nên ko thể so sánh => trc hết phải check type của userID
      * Cách so sánh
        + Kiểm tra 2 kiểu của name attr của UserID
        + Nếu là user

Trùng tên và group => 1

Trùng group =>0.5

Ko trùng =>0

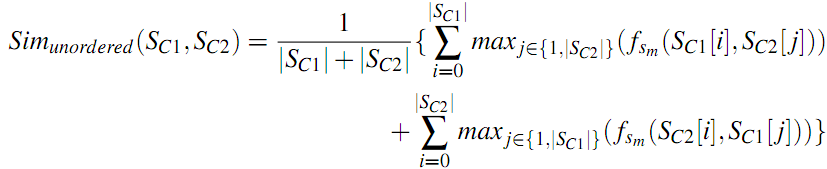
* + - * + Nếu là group

Trùng group => 1

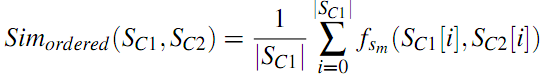
Ko trùng => 0

* + - So sánh 2 process (Process class của IDMEF)
      * So sánh PID and/or name để xác định similarity scale
      * So sánh argument của 2 process
        + Chúng đã ordered
    - So sánh 2 service
      * Dựa vào port,portlist,protocol
      * Sử dụng taxonomy tress như IP
      * So sánh webservice class
        + String compare cho url,cgi,http-method
        + Arg đc order như arg của process
      * So sánh SNMPService
        + Dùng string compare với oid,community và command

1. Kết hợp các similarity value đã tính đc
   * Áp dụng trọng số cho từng class khác nhau và tính kết quả. Đc thực hiện đồng thời.
   * So sánh 2 class unordered



* + So sánh 2 class ordered



* + Gộp các similarity value tính đc

