Trustwave®

SpiderLabs®

An Innovative Obfuscated Code Analysis Algorithm

AppSec Latam 2011

Sobre mim..

- **Membro Trustwave SpiderLabs Research**
- **Um dos criadores do Suricata IDS/IPS**
- **Mantenedor Apache ModSecurity**
- **Membro IEEE (pesquisas publicadas – China, Inglaterra, EUA, Brasil, Portugal)**

Agenda

a. **Motivação**
b. **Objetivos**
c. **O que é um código ofuscado ?**
d. **Soluções atuais**
e. **Algorítmo**
f. **Resultados**
g. **Conclusões**

Motivação

- **Ausência de um modelo capaz de classificar a intenção do código ofuscado em IDS, IPS, WAFs, dentre outros devices.**

- **Provedores de conteúdo, Hosting... não são capazes de julgar a intenção de códigos ofuscados de seus clientes.**

Objetivos

- **A solução deve atender aos requisitos:**
  - Poder de classificação
  - Adaptabilidade
  - Generalização
  - Performance

O Que é um código ofuscado ??

- **Um código ofuscado é aquele que passou por um processo de transformação de forma que sua leitura seja de dificil entendimento.**

Clear text javascript	Obfuscated javascript
<script> alert('teste') </script>	eval(function(p,a,c,k,e,d) {e=function(c){return c};if (!''.replace(/^/,String)){while(c--) {d[c]=k[c]||c}k=[function(e) {return d[e]}];e=function() {return'\\w+'};c=1};while(c--){if(k [c]){p=p.replace(new RegExp('\ \b'+e(c)+'\\b','g'),k[c])}}return p} ('<0>1(\'2\')</0>',3,3,'script| alert|teste'.split('|'),0,{}))

Soluções atuais

- **Análise sintática dos códigos:**
  - Geralmente baseadas em string search e pattern matching.
  - Quase nenhum poder de adaptação e generalização.
  - Boa performance.
- **Análise semântica dos códigos:**
  - Geralmente necessita de emulação, execução, desofuscação do código.
  - Bom poder de adaptação e generalização.
  - Baixa performance.

Algorítmo

- **Processo estocástico com estados discretos.**
  - Cadeias de Markov

- **Carecterísticas de algorítmos para aprendizagem de máquina.**
  - Base de treinamento e validação
  - Supervisão e ajuste por correção de erro

Algorítmo

- **Definição dos estados**
  - Aqueles comumente encontrados em códigos ofuscados

States (S)
charCodeAt
String
fromCharCode
charAt
function
substr
unescape
replace
length
This
eval
call
u([0-9]){4}

Table II. Obfuscated code states (S)

Algorítmo

- Para cada estado (S), a ligação L(S) entre eles obedecendo a função

$$p\left(X_j = S_{i+1} | X_i = S_i \right)$$

- Origina uma cadeia

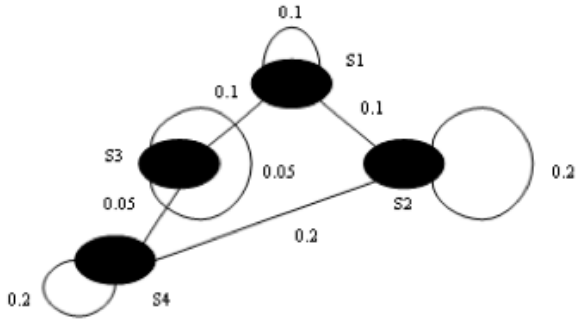


Fig 1. Example of a chain with four states S.

- Onde cada probabilidade de transição é dada:

$$D_{ij} = \frac{K}{\sum_{i=0}^{N} \sum_{j=0}^{N} L(S_{ij})}$$

Algorítmo

- Duas cadeias são criadas da base de treinamento, destas duas matrizes quadradas são criadas (modelos) :

$$M = \begin{bmatrix} 0.1 & 0.1 & 0.1 & 0 \\ 0 & 0.2 & 0 & 0.2 \\ 0 & 0 & 0.05 & 0.05 \\ 0 & 0 & 0 & 0.2 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0.1 & 0 & 0 & 0.1 \\ 0 & 0 & 0 & 0.5 \end{bmatrix} \quad (3)$$

- Dos modelos, temos duas definições: $g(x) < f(x)$ para códigos reconhecidamente benignos e $f(x) >= g(x)$ para maliciosos, sendo :

$$f(y) = \sum D_{ij} \rightarrow \forall L(S_{ij}) \mid \exists L(S_{ij}) \in M_{n,n} \quad (4)$$

$$g(y) = \sum D_{ij} \rightarrow \forall L(S_{ij}) \mid \exists L(S_{ij}) \in B_{n,n} \quad (5)$$

Algorítmo

- Em caso de erro de reconhecimento:

$$E = f(y) - g(y) \quad (6)$$

$$E = g(y) - f(y) \quad (7)$$

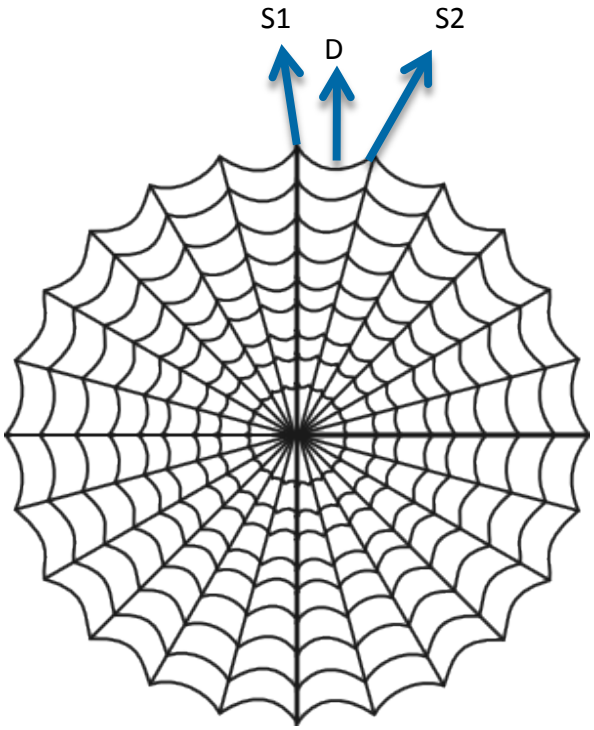
- E um ajuste na cadeia será feito:

$$D_{ij} = D_{ij} + E * \left[\frac{(D_{ij} * 100)}{f(y)} \right] * \upsilon \quad (8)$$

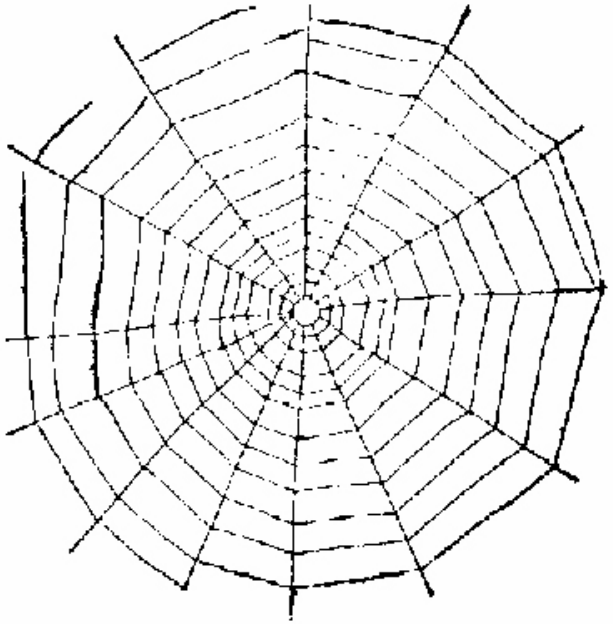
$$D_{ij} = D_{ij} + E * \left[\frac{(D_{ij} * 100)}{g(y)} \right] * \upsilon \quad (9)$$

From the spider point of view…

Algoritmo – pré-treinamento

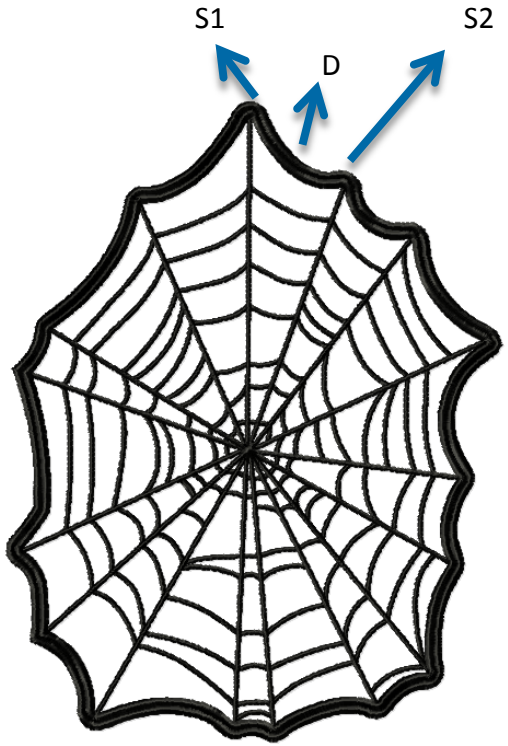


Modelo - Maliciosos



Modelo - Benignos

Algoritmo - pós-treinamento



Modelo - Maliciosos

Modelo - Benignos

Resultados

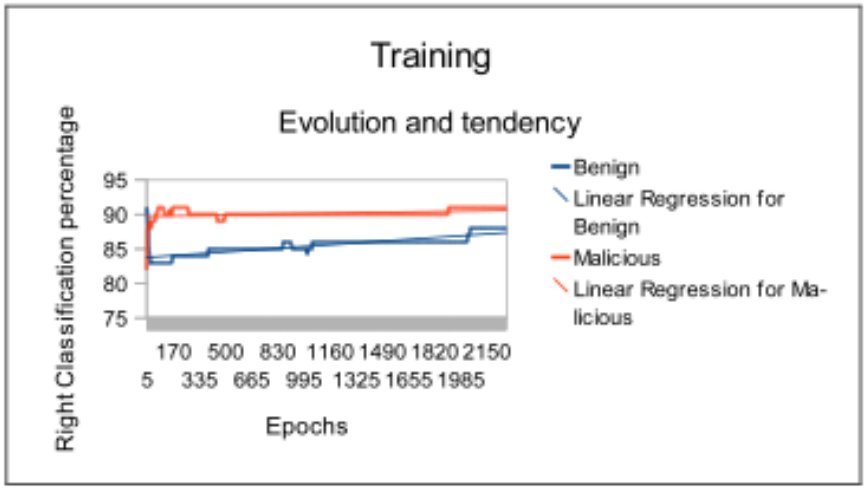


Fig 3. Evolution and tendency in training phase

- **90% de classificação correta**

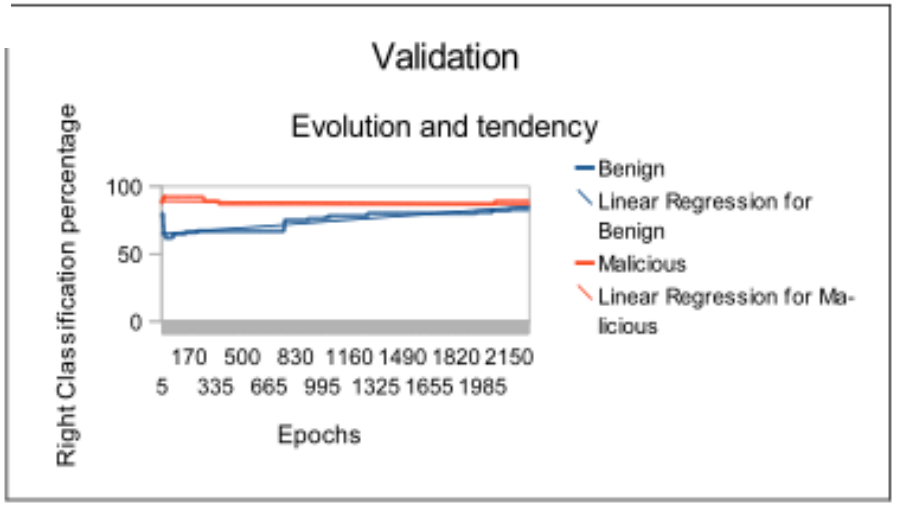


Fig 4. Evolution and tendency in validation phase.

- **90% de classificação correta**

Conclusões

- **O protótipo apresentou uma capacidade aceitável de classificar e de generalizar.**
- **Necessidade de aumento da base de treinamento e validação para deploy em ambiente real.**
- **Treinamento e validação ~95%**

Trustwave®

SpiderLabs®

Obrigado!