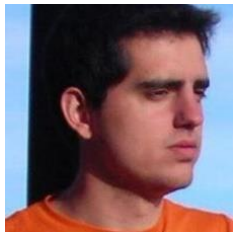
VIII JORNADAS STIC CCN-CERT

La defensa del patrimonio tecnológico frente a los ciberataques

10 y 11 de diciembre de 2014

Pre-fetching of suspicious mobile apps based on Big Data techniques and correlation

Equipo de investigación



Dr. Alfonso Muñoz
Senior Cyber Security Researcher
(Co)Editor Criptored, CISA, CEH…
alfonso.munoz@11paths.com @mindcrypt

Dr. Antonio Guzmán
Senior Cyber Security Researcher
Head of the research
antonio.guzman@11paths.com

Sergio de los Santos
Senior Cyber Security Researcher
Head of Labs
Sergio.delosantos@11paths.com

Pre-fetching of suspicious mobile apps based on Big Data techniques and correlation

Amenazas en el mundo móvil: ¿Estamos preparados?

Pre-fetching of suspicious mobile apps based on Big Data techniques and correlation

Objetivo: ¿Detección de "aplicaciones maliciosas"?

Detección: "Análisis estático/dinámico de código, sandboxing, …"

el dinamismo con el que se modifican las aplicaciones para móvil, la necesidad de completar la instalación de las aplicaciones para poder completar el análisis, unido al coste en recursos y tiempo hacen que la eficiencia de estas soluciones para contener el avance del código malicioso no sea suficiente…

Necesitamos más ayuda para minimizar el número de ataques creciente

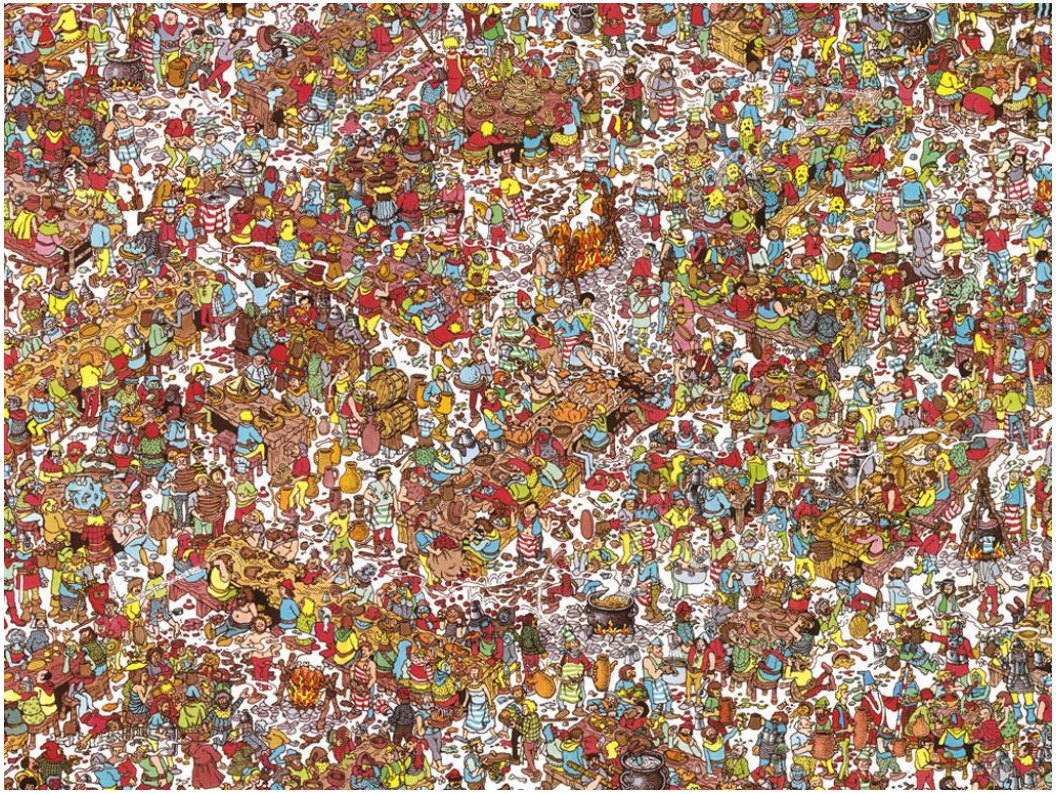
¡Sigamos el consejo de nuestras abuelas!

Pre-fetching of suspicious mobile apps based on Big Data techniques and correlation

Detectando aplicaciones sospechosas sin análisis de código ni ejecución…

Consejo de la abuela: ¿Dime con quién anda? ¿Dime donde anda? ¿Dime como viste? ¿su comportamiento es extraño? ¿es de “buena familia”?…

BigData: Procesando markets de aplicaciones móviles…



Plataforma cloud "Path5"

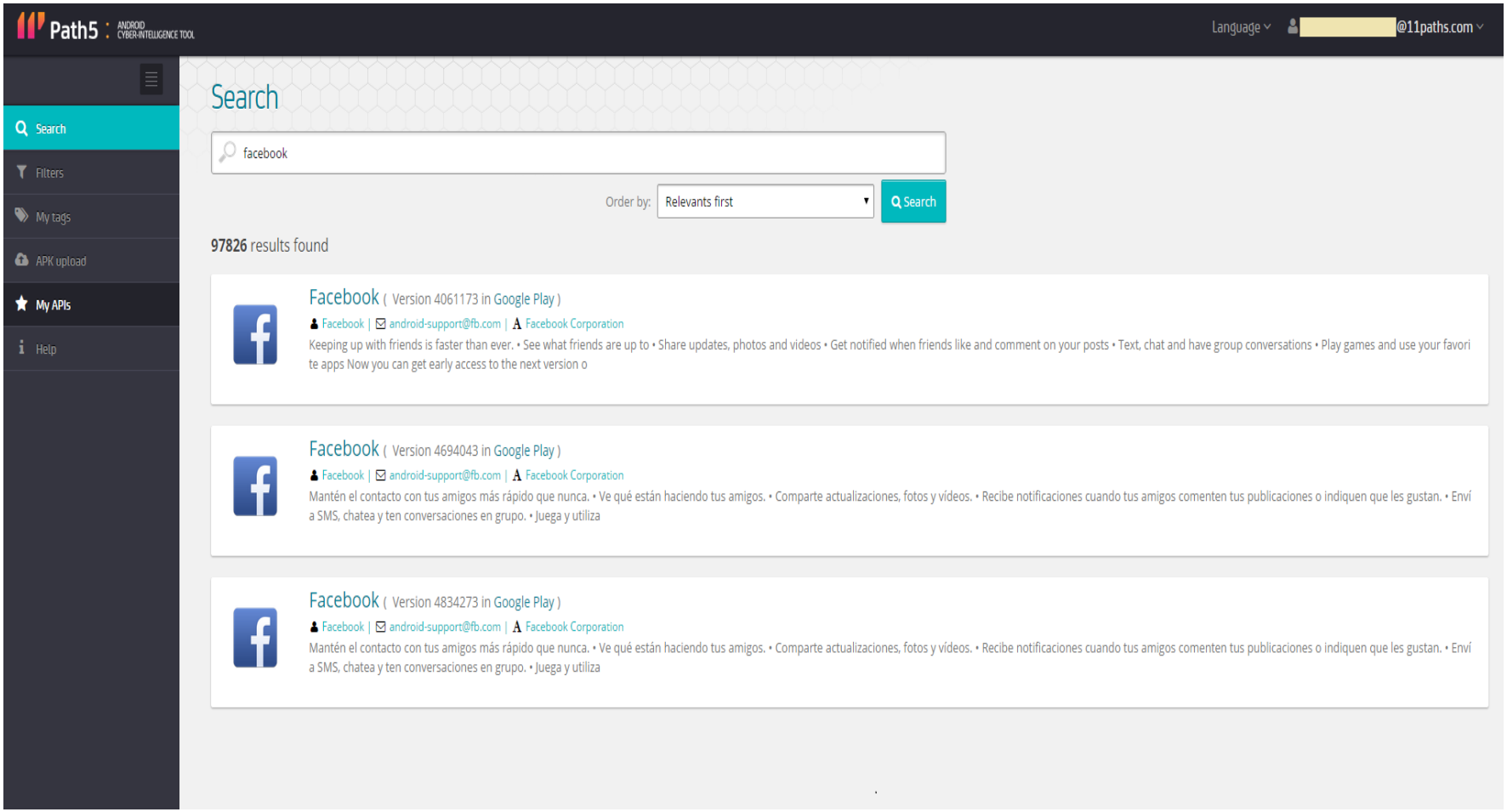
- Más de 1,5 millones de apks
- Google Play y *otros*…
- Detección de actualizaciones…

Preprocesado de información

- Despiezamos apks
- Información de contorno
- Developer (familia)
- Historial
- ¿Es malware? VIRUS TOTAL
- …

Pre-fetching of suspicious mobile apps based on Big Data techniques and correlation

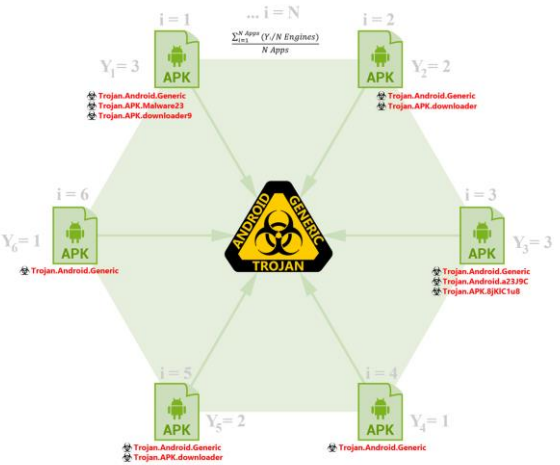
BigData: Procesando markets de aplicaciones móviles...

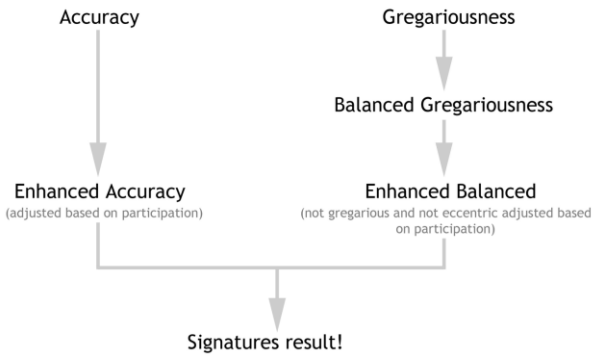


Pre-fetching of suspicious mobile apps based on Big Data techniques and correlation

Reflexión crítica: ¿Cómo detectar malware automáticamente?

Precisión y gregarismo





SECURITY

Kaspersky defends false detection experiment

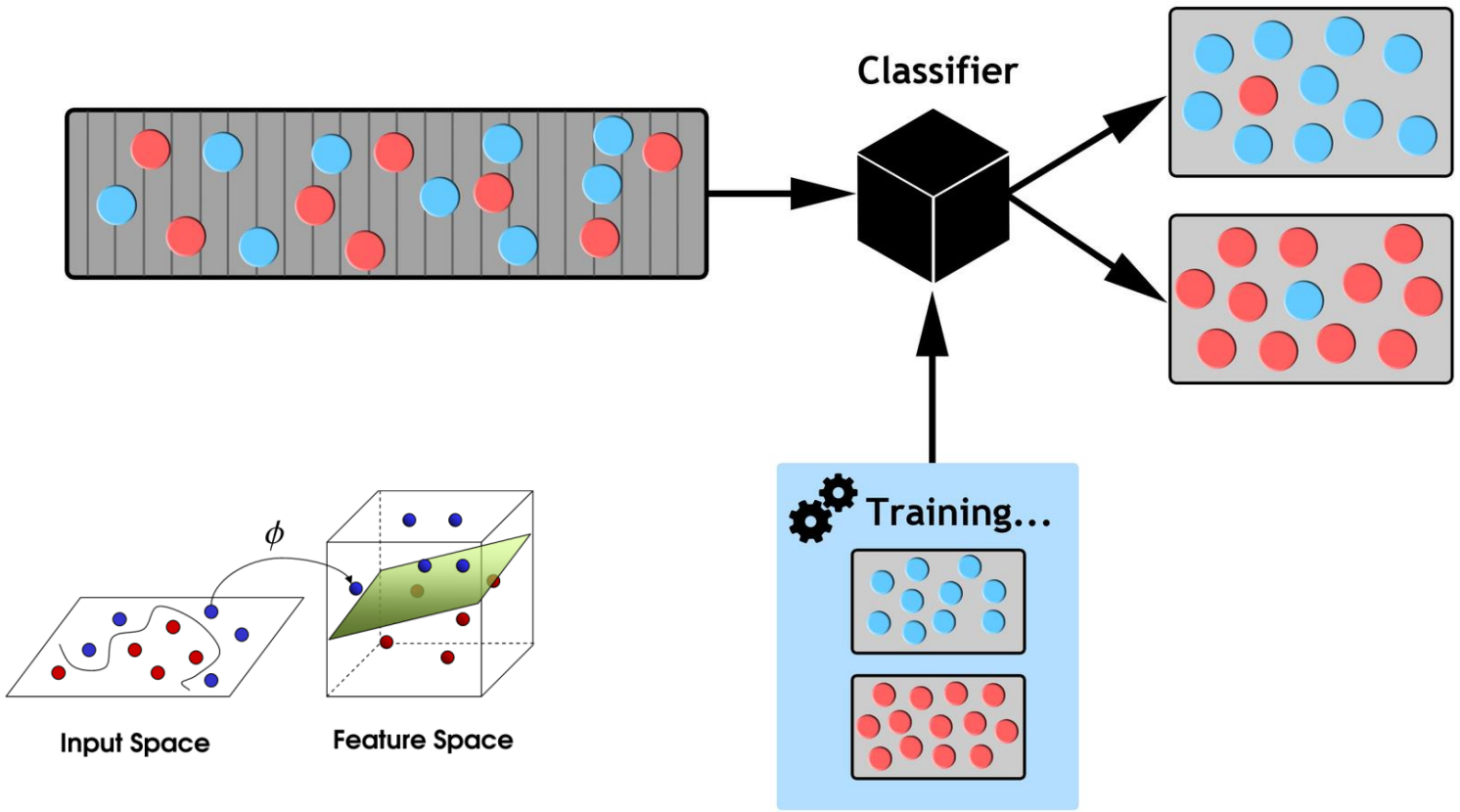
Claws in copy cat dust-up

By John Leyden, 10 Feb 2010

Pre-fetching of suspicious mobile apps based on Big Data techniques and correlation

Relacionando en base a estimadores. Priorizando esfuerzos...

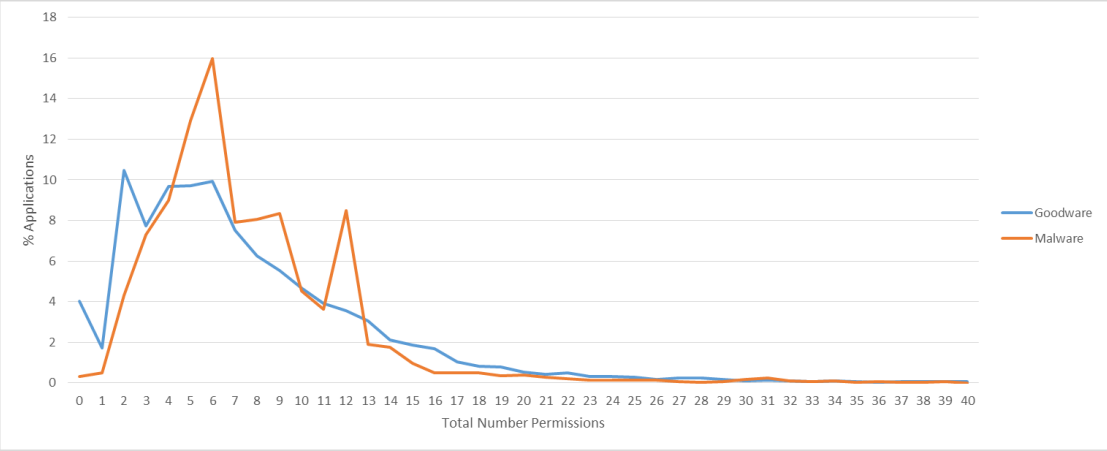
Machine learning

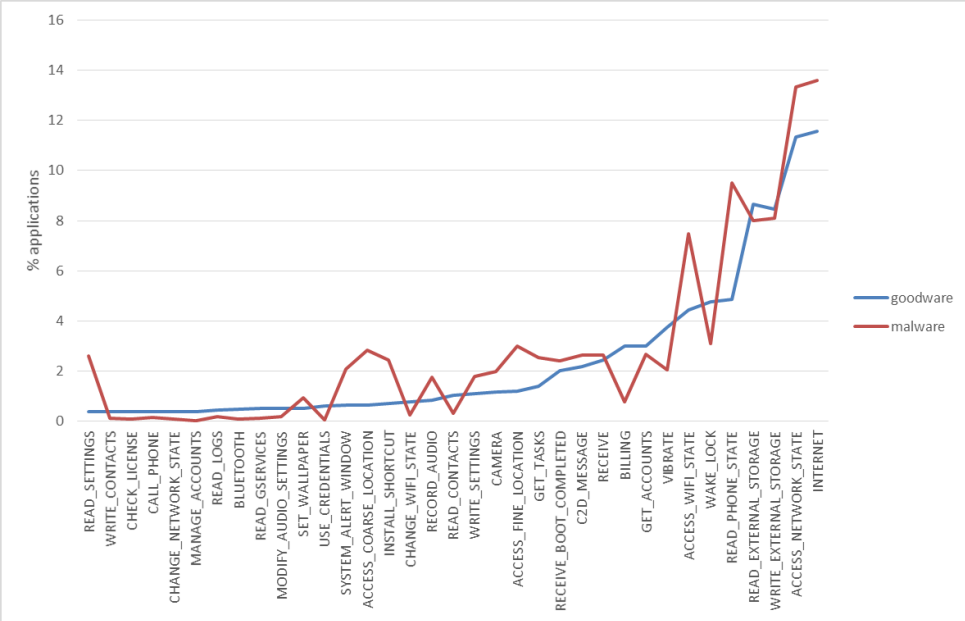


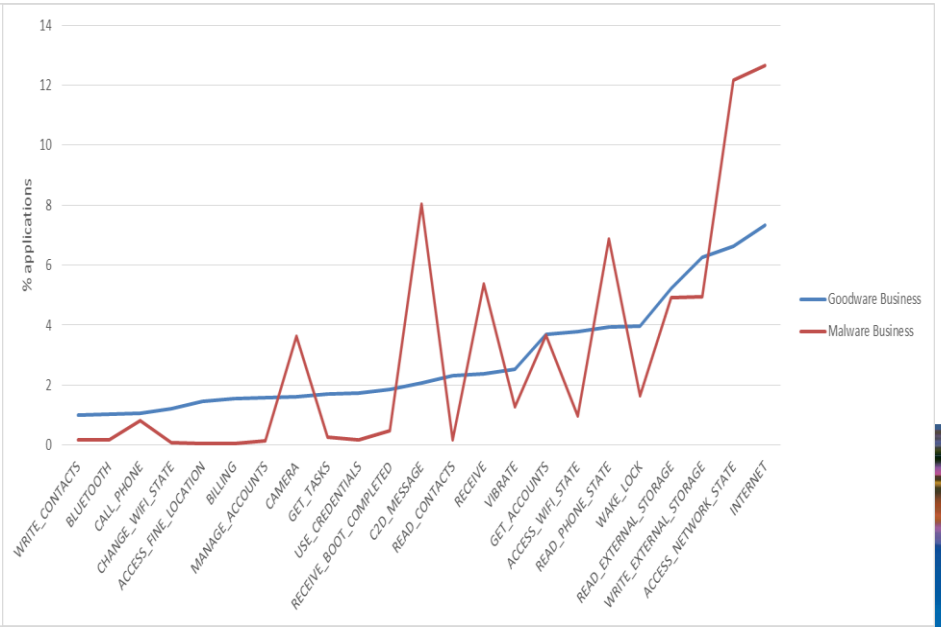
Pre-fetching of suspicious mobile apps based on Big Data techniques and correlation

Relacionando apks: ¿Qué es un estimador?

El caso del malware: Path5 tiene decenas de estimadores







Relacionando apks: ¿Qué es un estimador?

Estilometría y procesamiento de lenguaje natural en la detección de malware…

¿Cómo escriben los desarrolladores? ¿Cómo son los comentarios?

- Ejemplo 1: Tamaño del campo descripción
- Ejemplo 2: Riqueza del lenguaje
- …

Campo Descripción: Goodware Vs Malware

SQL Reference will help you to understand the basics of SQL and become familiar with the advanced level. Content of the handbook is divided into four categories:- Basics - Advanced - Functions - Articles – Training. The description of each function includes some examples with comments. Reference is also shows major differences of the syntax of the various databases: MySQL, SQL Server, Oracle, Access. Training section allows you to practice in running sql queries. This section availible only in full version. If you have suggestions, ideas to improve the application and comments, please, use the feedback form. You can find it in the menu.

This App is a handy reference list of 140 browser-supported colors. which are sorted based on: - Color name - Hex code, and - Color family. These 140 color names can be defined under HTML and CSS color specification tags. All of the major browsers have included support for these colors.

- ¡Usemos decenas de estimadores y diferentes conjuntos para detectar similitudes!

Category	Acurracy	Category	Acurracy	Category	Acurracy	Category	Acurracy
Book and reference	83,53%	Game arcade	77,14%	Lifestyle	79,77%	Shopping	92,22%
Business	88%	Game educational	71,33%	Media and video	77%	Social	81%
Communication	79%	Game family	85,77%	Music and audio	83,11%	Sports	87,24%
Education	83%	Game puzzle	84,2%	News and magazines	86,24%	Tools	78,91%
Entertainment	78,4%	Game Racing	78,22%	Personalization	85,86%	Travel and local	83%
Game action	77,77%	Game sports	73,77%	Photography	85,16%		
Game adventure	80%	Health and fitness	83,55%	productivity	78,51%		

CASO DE ESTUDIO

Detectando una botnet de *Clickfraud en Google Play*

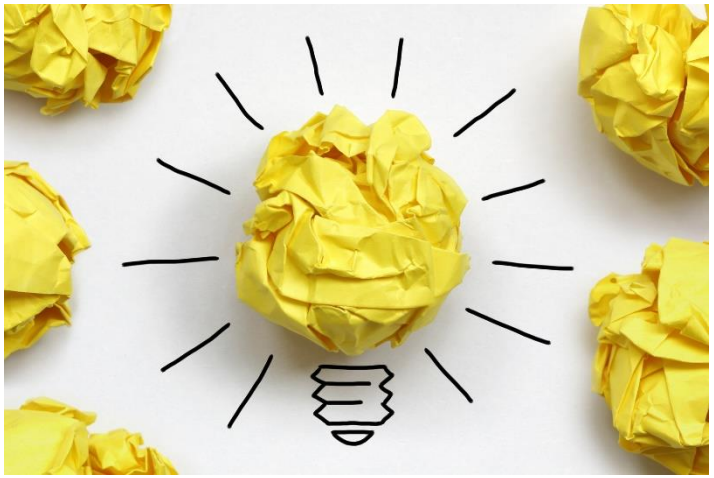
Path5+Escuela “malagueña” ☺

Conclusiones

- Detectar código malicioso es “caro y complejo”. Miles de apps son publicadas diariamente, es una buena idea priorizar esfuerzos antes de un análisis en profundidad.

- Relacionar y clasificar aplicaciones de forma no tradicional ofrece resultados con gran precisión. BG+ML

- Las posibilidades son enormes: detección de código malicioso, aplicaciones fraudulentas, predecir aplicaciones retirables de markets, seguimiento de vulneración de marca, detección de botnets, etc.



VIII JORNADAS STIC CCN-CERT

La defensa del patrimonio tecnológico frente a los ciberataques

10 y 11 de diciembre de 2014

Pre-fetching of suspicious mobile apps based on Big Data techniques and correlation

¿? || /* */

Dr. Alfonso Muñoz
alfonso.munoz@11paths.com | @mindcrypt

- **E-Mails**
  - ccn-cert@cni.es
  - info@ccn-cert.cni.es
  - ccn@cni.es
  - sondas@ccn-cert.cni.es
  - redsara@ccn-cert.cni.es
  - carmen@ccn-cert.cni.es
  - organismo.certificacion@cni.es
- **Websites**
  - www.ccn.cni.es
  - www.ccn-cert.cni.es
  - www.oc.ccn.cni.es

Síguenos en Linked in