Detección de anomalías de tráfico en servidores web

Propuesta de Valor

GESTIÓN DE PROYECTOS DE CIENCIA DE DATOS

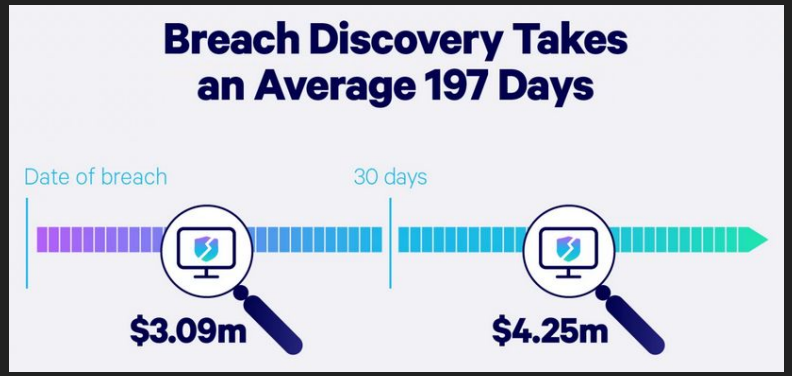
Mitsiu Alejandro Carreño Sarabia - E23S-18014

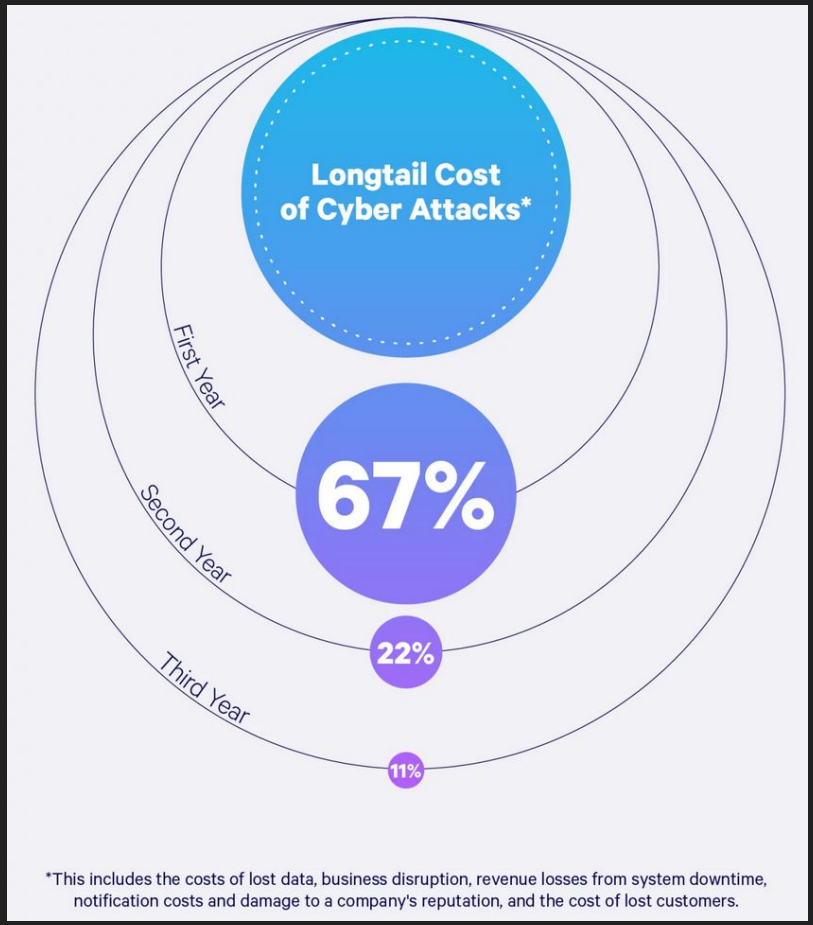
Contenido

1. Introducción
2. Solución propuesta
3. Descripción del problema
4. Propuesta de valor
5. Viabilidad técnica
6. Diferenciadores competitivos
7. Próximos pasos

1. Introducción

En un mundo interconectado, los riesgos se distribuyen igual para todos.



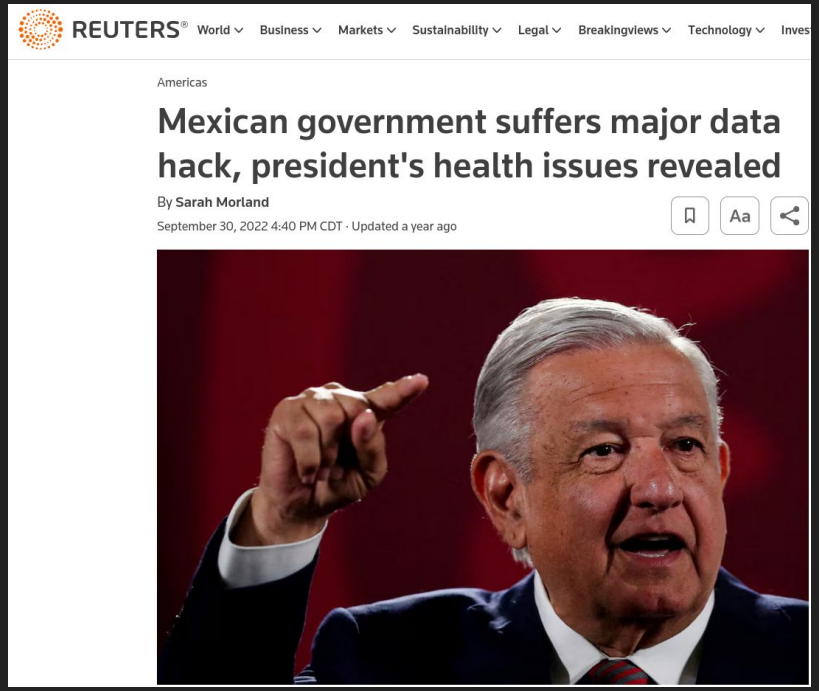


Y ser víctima es muy costoso

Fuente: 2024 Embroker Insurance Services LLC

1. Introducción

El riesgo puede estar mucho más cerca de lo que creemos



- https://www.vice.com/en/article/53d9k5/how-a-hacker-found-the-personal-information-of-all-mexican-voters
- https://www.reuters.com/world/americas/mexican-president-confirms-hack-government-files-2022-09-30/

1. Introducción

El riesgo puede estar mucho más cerca de lo que creemos

Cronista México • Actualidad • México

Ciberseguridad

Alerta por HACKEO a la Universidad Politécnica de Aguascalientes: filtraron miles de datos de alumnos y profesores

La base de datos fue vulnerada y circula libremente información de todos los miembros de la comunidad

- https://www.cronista.com/mexico/actualidad-mx/alerta-por-hackeo-a-la-universidad-politecnica-de-aguascalientes-filtraron-miles-de-datos-de-alumnos-y-profesores/

2. Solución propuesta

Tema:

Desarrollar una **solución integral** de **monitoreo y detección** de tráfico anómalo mediante la implementación de técnicas de análisis topológico y aprendizaje automático que permita la toma de **decisiones preventivas**

3. Descripción del problema

¿Por qué aprendizaje automático?

Cualquier contenido accesible desde internet está alojado en un servidor.

Un servidor de manera **predefinida registra todas las solicitudes,** pero al estar expuesto en internet, cualquier dispositivo, persona o programa puede solicitarle información.

3. Descripción del problema

Mucho se ha desarrollado en términos de escalabilidad de infraestructura así como adopción de soluciones distribuidas para dar servicio a la creciente demanda.

La cantidad de
información generada es tan grande que un análisis manual no es viable.

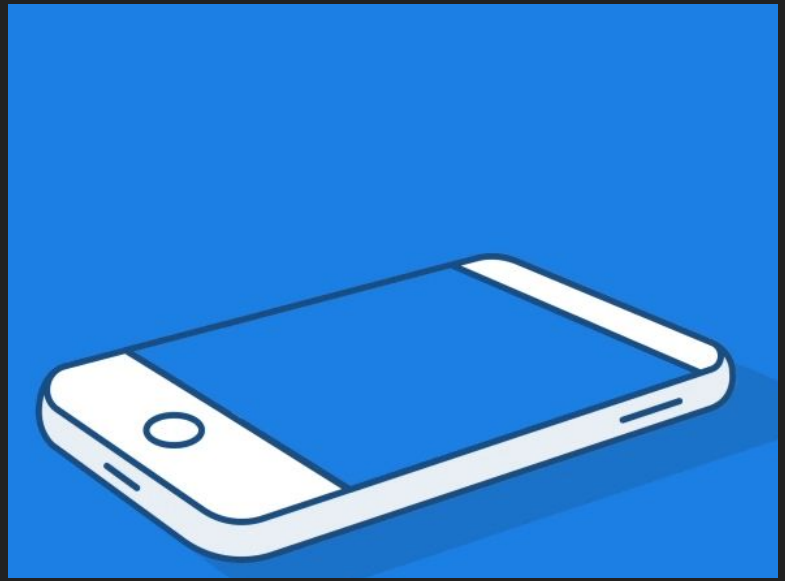
4. Propuesta de valor

De la misma manera que siempre queremos saber que nuestros bienes están seguros, el objetivo de este sistema es para darle la misma garantía a la información de tus clientes que alojas en internet.

Conoce y explora cómo tus clientes usan los recursos de tu servidor, toma **acciones preventivas y correctivas basadas en información y métricas** en tiempo real.

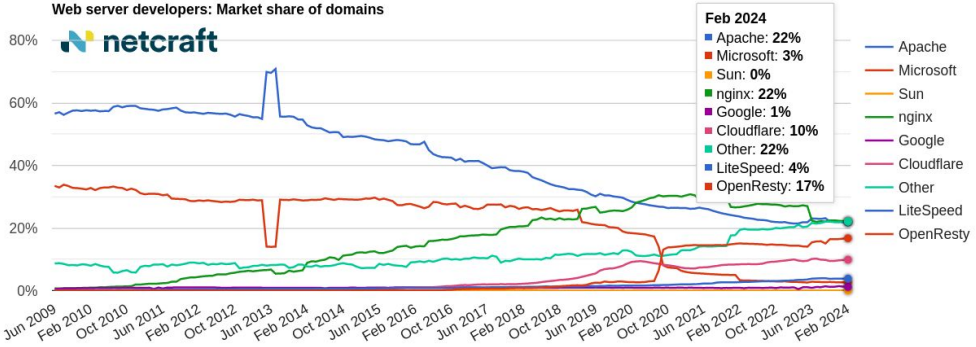
4. Propuesta de valor

Ajusta tu infraestructura a la demanda real de tu servicio y **responde a ciber-incidentes más rápido** que la competencia.



5. Viabilidad técnica

Actualmente el desarrollo está centrado en ser compatible con las bitácoras de registro generadas por NGINX, actual líder del mercado en tecnología reverse-proxy



Developer	January 2024	Percent	February 2024	Percent	Change
nginx	59,970,194	22.17%	59,707,340	22.02%	-0.15
Apache	58,971,448	21.81%	59,057,844	21.78%	-0.02
OpenResty	44,899,651	16.60%	45,325,578	16.72%	0.11
Cloudflare	26,361,454	9.75%	26,692,491	9.84%	0.10

https://www.netcraft.com/blog/february-2024-web-server-survey/

5. Viabilidad técnica

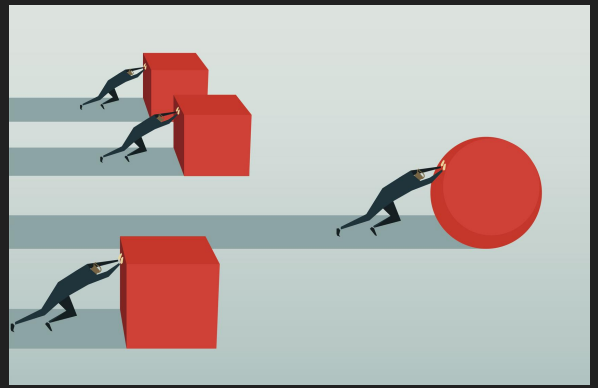
A pesar del vínculo actual con la tecnología NGINX, la n**aturaleza desacoplada del sistema** y del procesamiento vuelve **alcanzable realizar ajustes para trabajar con otras tecnologías como Apache**.

Sin necesidad de generar una nueva red de aprendizaje máquina.

6. Diferenciadores competitivos

Beneficios:

- Tener una **herramienta plug-n-play,**
- **Dar uso a las bitácoras de registro** que probablemente ya se generan de manera predefinida
- Basado en los datos históricos de tu propio servidor, **algoritmo custom-made**



7. Próximos pasos

1. Actualmente se está trabajando tanto en el **análisis topológico** de las bitácoras de entrenamiento como en el **diseño de la red neuronal**.
2. Una vez completado el análisis y entrenamiento se trabajará la integración para el **monitoreo en tiempo real** de las peticiones NGINX
3. Eventualmente es posible volver más robusto el sistema de monitoreo agregando **alertas y notificaciones push a administradores de sistema.**

Gracias