

RedBee

Sistema de análisis inteligente de ciberataques en honeypots

Marta Sestelo

Galician Research Center in Advanced Telecommunications
VII Xornadas de usuarios de R Galicia
15 October 2020

Índice

1. Introducción

2. Sistema RedBee

- Módulo de preprocesado
- Módulo de análisis inteligente
- Módulo predictivo
- Módulo de visualización

3. Demo

¿Qué es Gradient?

Proveedor de innovación, con más de 10 años de experiencia en incubación de tecnología

100
Personas

5M€
Ingresos en 2019

68%
Facturación con
empresas

32%
Financiación
pública competitiva

23
Proyectos
europeos

2
Proyectos
Red Cervera

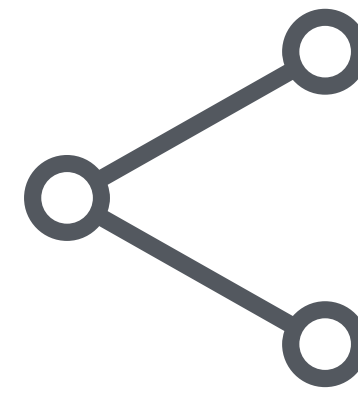


¿Qué es Gradient?



conectividad

- Subsistemas de Comunicaciones (digitales y analógicos)
- IoT (*Internet of Things*)
- Sistemas Integrados y Onboard
- Redes



inteligencia

- *Data Analytics y Big Data*
- Análisis inteligente de vídeo
- Gestión y despliegue de infraestructura IT avanzada
- *Learning Analytics y Adaptive Learning*
- Bioinformática



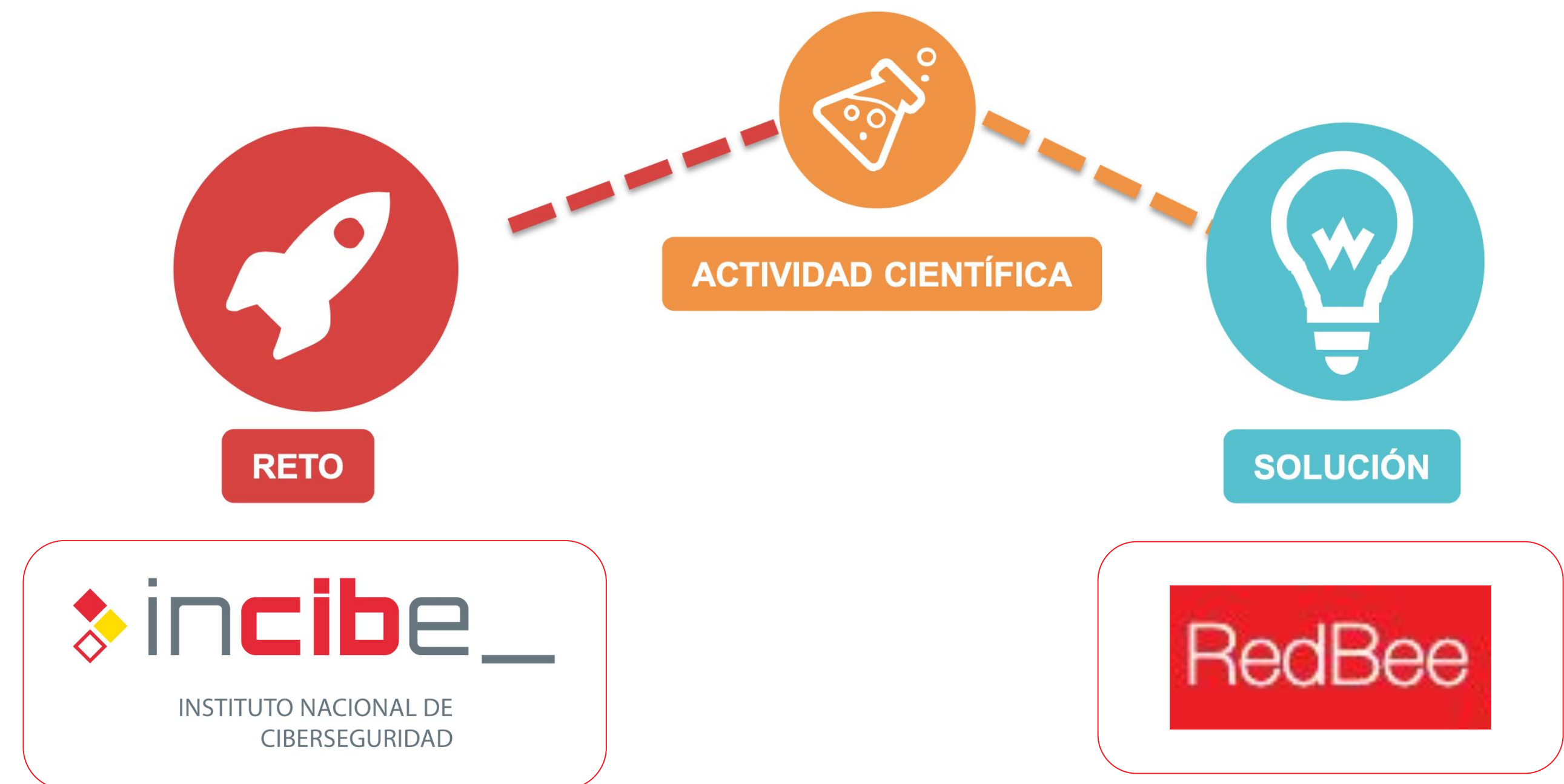
seguridad

- Seguridad de la información
- Seguridad en Cloud
- Sistemas de protección de la Privacidad
- *Privacy by Design*
- Sistemas biométricos



Introducción

- Programa de Transferencia JNIC
(Jornadas Nacionales de Investigación en Ciberseguridad)
- Certamen de Retos Científicos de 1 año de Duración (2018-2019)
- Reto: Caracterización de ataques a honeypots SSH



Honeypots, ¿qué son y para qué sirven?

Servicios vulnerables que se despliegan como señuelo para recibir ciberataques



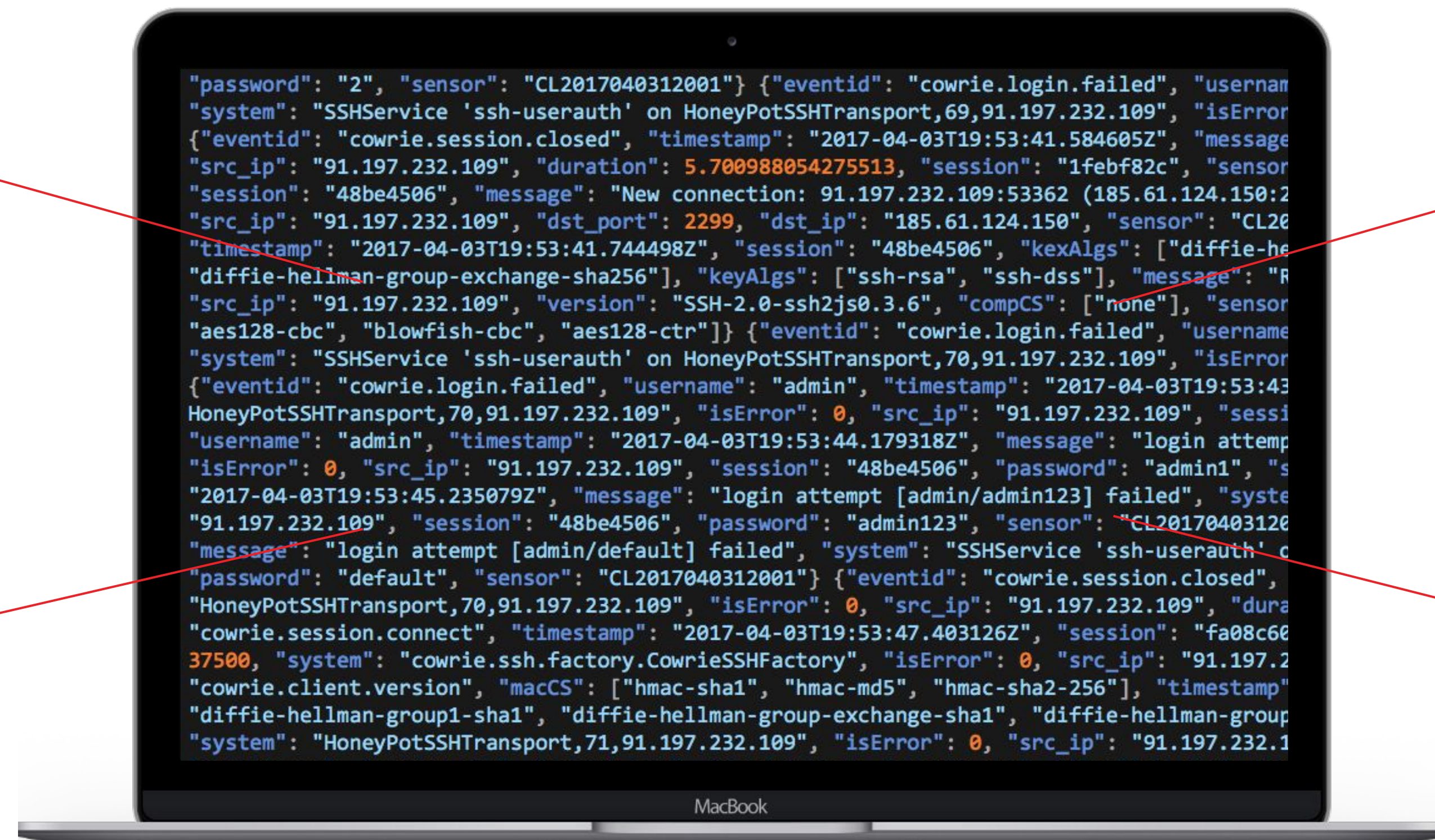
Versátiles

Pueden emular prácticamente cualquier servicio (bases de datos, web, SSH, SCADA...)



Vulnerables

Buscan ser objeto del mayor número posible de ataques



Registro en logs

La información de los ataques se almacena en ficheros (estructurados o no estructurados)



Análisis manual

No ofrecen una interpretación automática de los ataques



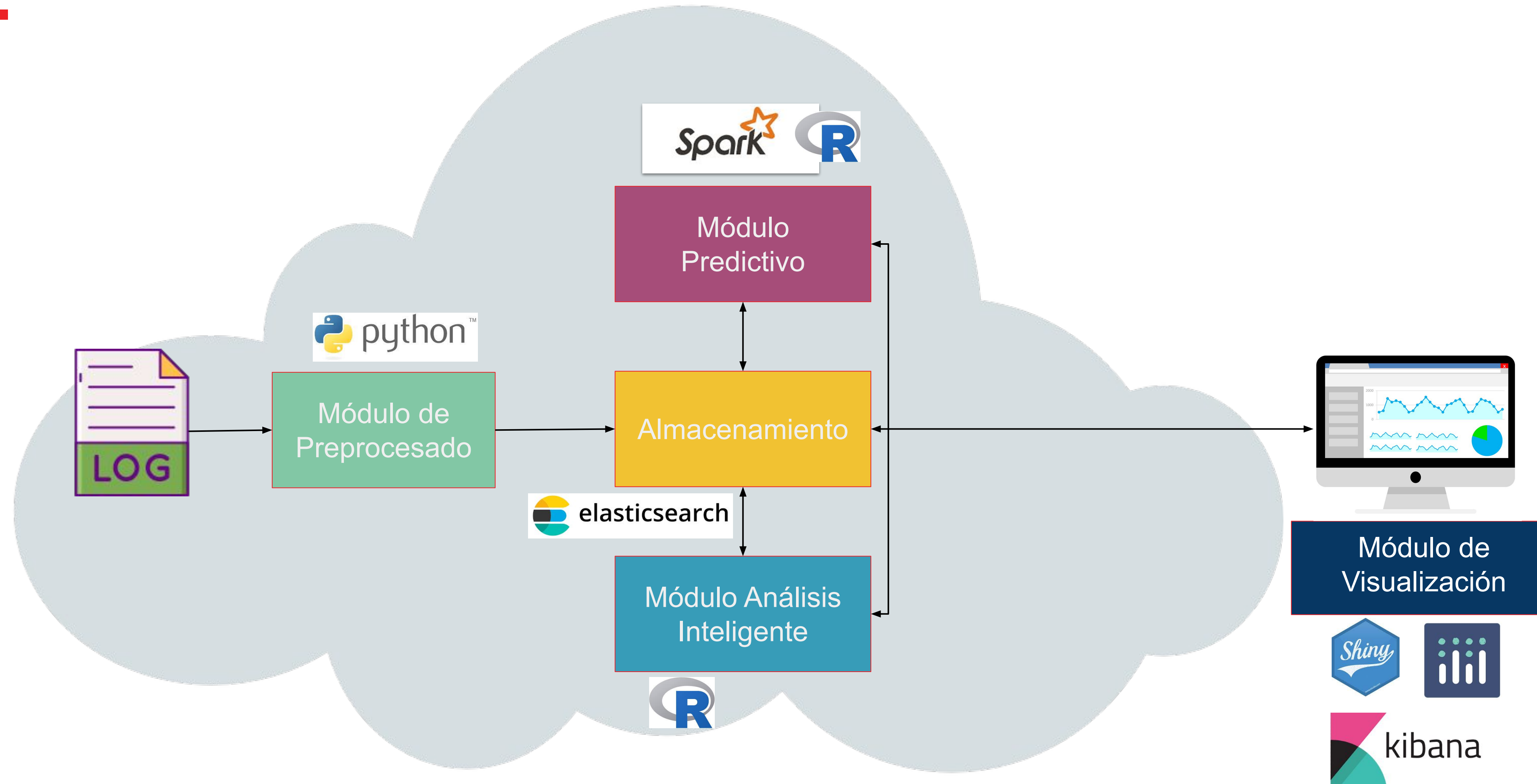


RedBee: sistema de análisis inteligente de ciberataques en honeypots

- Herramienta de seguridad basada en técnicas de **Machine Learning**
- Permite **caracterizar de forma automática** los ataques recibidos y **predecir su número**
 - ¿Desde dónde nos atacan?
 - ¿Con qué frecuencia?
 - ¿Qué objetivo tienen los ataques?
 - ...
- Ofrece una **interfaz sencilla** e intuitiva



RedBee Arquitectura del sistema



RedBee Módulo de Preprocesado

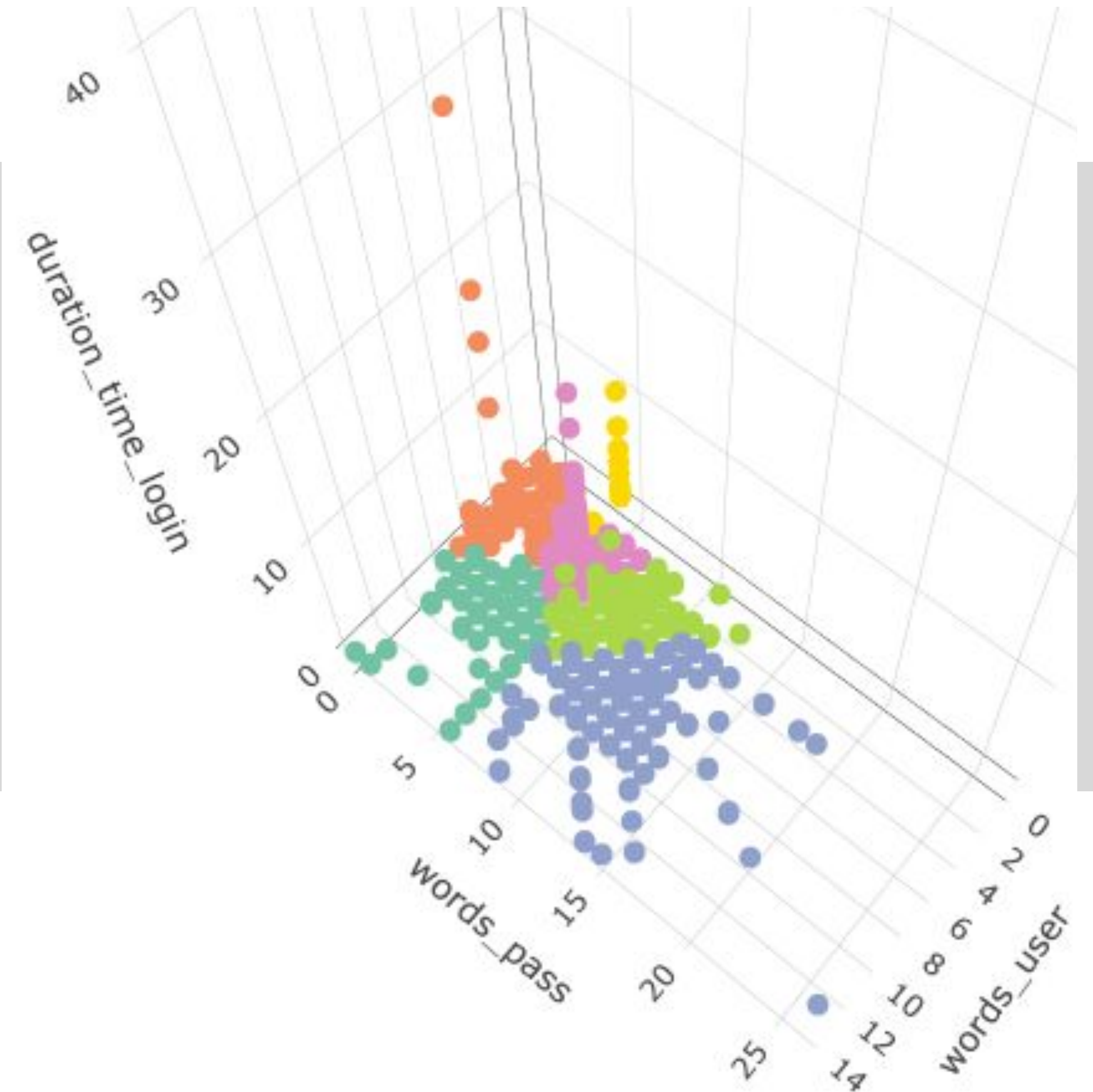
```
2018-03-27T10:06:24.381749+0200 [SSHChannel session (2) on SSHService 'ssh-connection' on
HoneyPotSSHTransport,52145,178.62.44.104] Terminal Size: 24 80
2018-03-27T10:06:24.993389+0200 [HTTPPageDownloader,client] Not storing duplicate content
8871bf7645c132b73f7d16c22ef5c643fe6c3de6537196ffe78f27329a2e3c94
2018-03-27T10:06:24.995200+0200 [HTTPPageDownloader,client] Command found: perl /tmp/y.txt
```

- **Extracción de variables:** versión SSH, login correcto/fallido, usuario, contraseña, IPs, dominios, hashes de los ficheros, etc.
- **Creación de variables:** número de intentos de login, tiempo entre intentos, tiempo total, geolocalización de IPs, distancia de Levenshtein, salida VirusTotal, etc.
- **Lenguaje de programación:** Python



RedBee Módulo de Análisis Inteligente

- Método no supervisado basado en técnicas **clustering**
- **Algoritmos:**
 - DBSCAN
 - Mixturas de Modelos Gaussianos
 - K-means
- **Lenguaje de programación:** R (integración con Spark)



RedBee Módulo Predictivo

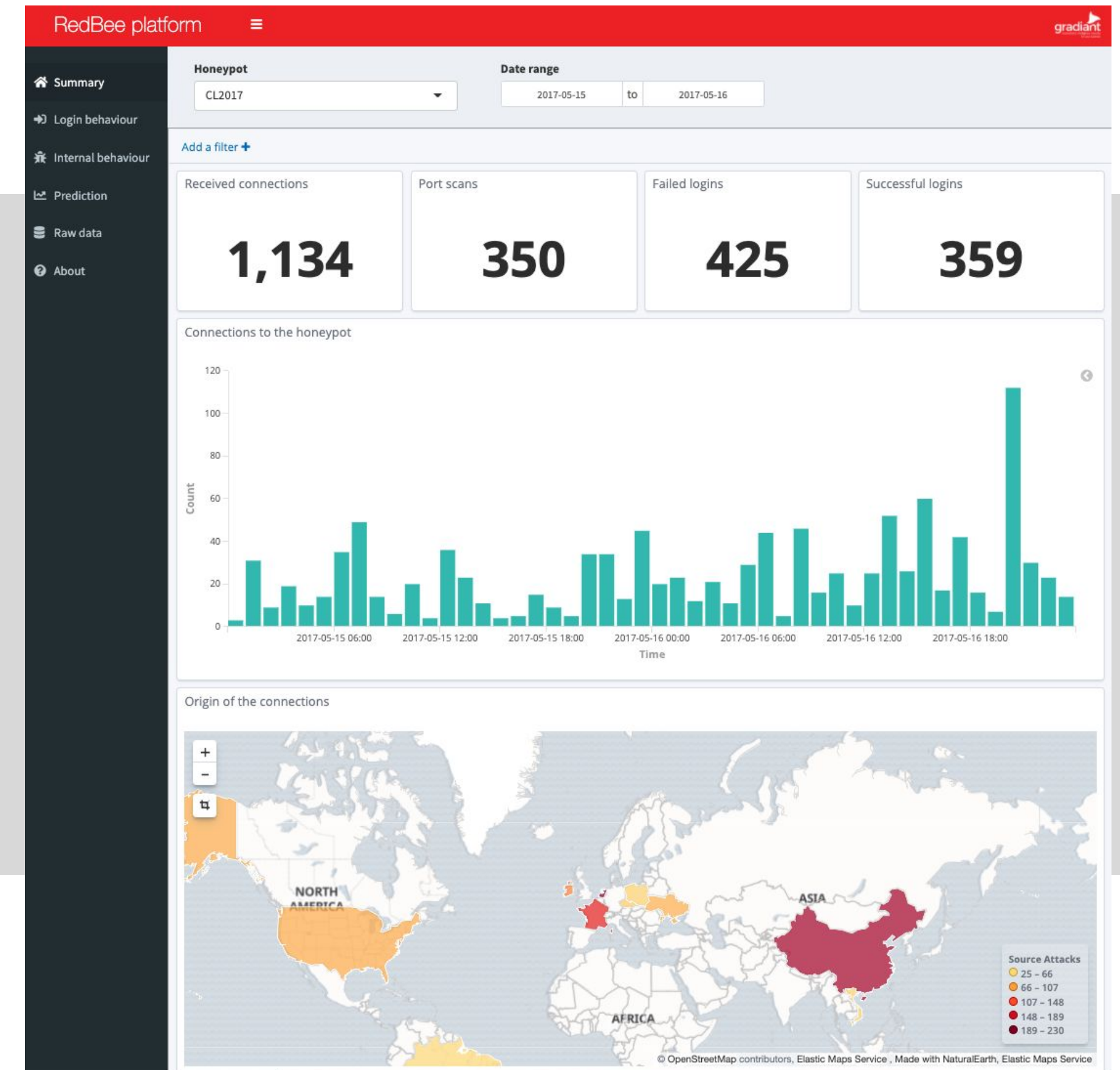


- Serie temporal: número de ataques por día, horizonte de predicción a 7 días
- Modelos:
 - Autoregressive Integrated Moving Average (ARIMA)
 - Autoregressive Random Forest
 - Generalized Additive Models (GAM)
- Lenguaje de programación: R

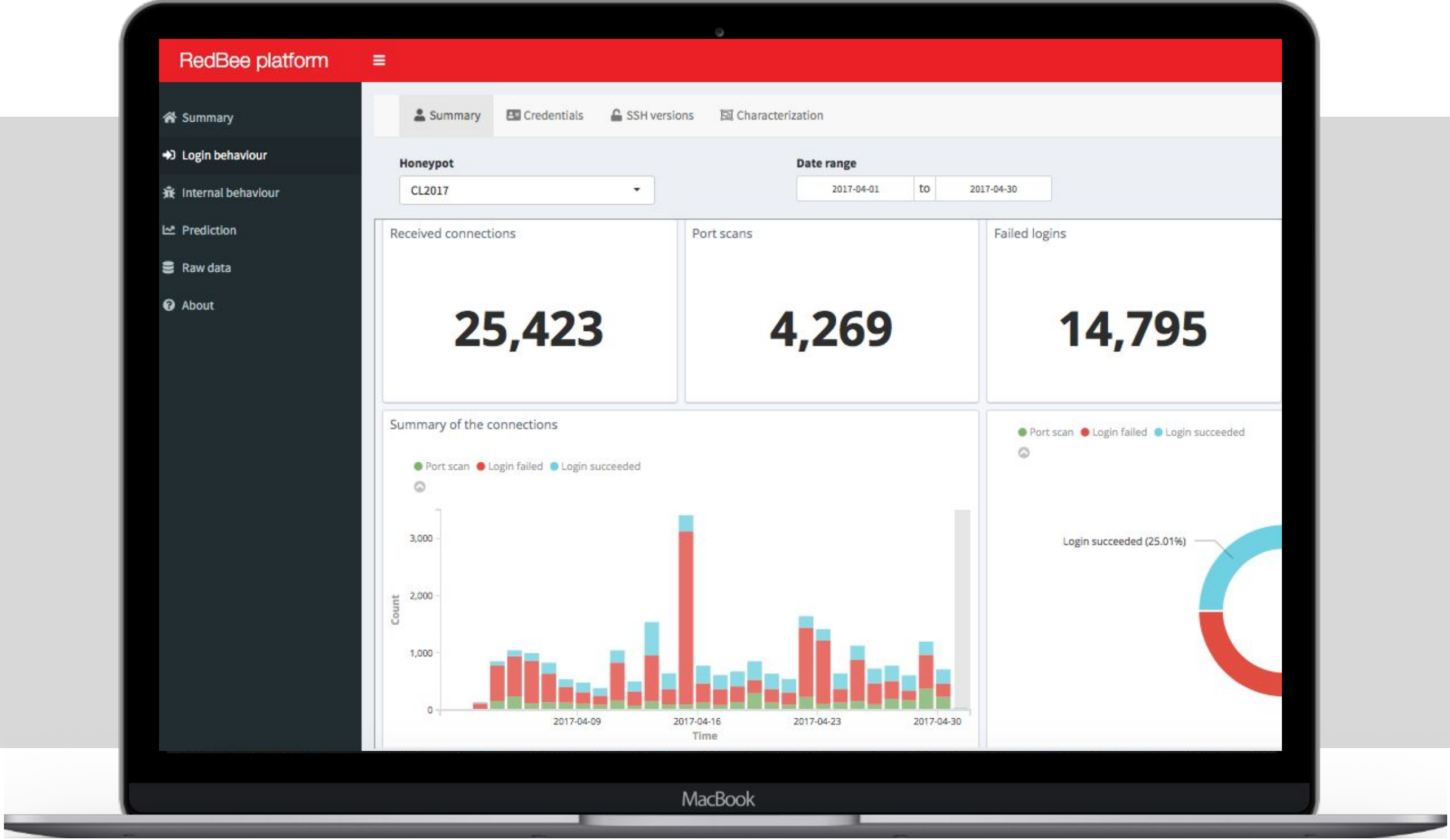


RedBee Módulo de Visualización

- **Interfaz gráfica** basada en la integración de dos herramientas:
 - **Shiny**: permite desarrollar aplicaciones interactivas basadas en R
 - **Kibana**: se utiliza para la visualización y exploración de datos almacenados en ElasticSearch



RedBee Demo





#FUTUREBEGINSTODAY



msestelo@gradiant.org



[@MartaSestelo](https://twitter.com/MartaSestelo)



[Marta Sestelo](https://www.linkedin.com/in/martasestelo)

(+34) 986 120 430 · gradiant@gradiant.org · www.gradiant.org