Modelización del protocolo Tor y extracción de características de servicios ocultos

Jorge García de Quirós

Ricardo J. Rodríguez

Universidad Zaragoza

Universidad de Zaragoza, Spain

28 de mayo de 2019

Agenda

Introducción

¿Qué es Tor?

- **Servicio para mejorar el anonimato y la privacidad en Internet**
- **Tor: The Onion Router**
  - Aplicación Tor
  - La red Tor, formada por dispositivos de voluntarios
  - Organización Tor Project

Introducción

¿Cómo funciona?

- **Comunicación anónima de baja latencia basada en circuitos**
  - **Circuitos virtuales**: garantizan que no hay conexión directa entre cliente y servidor.
  - **Saltos intermedios**: cada salto es un nodo. Cada router sólo conoce al siguiente y al anterior
  - **Tráfico cifrado por capas** → **encaminamiento cebolla**

Introducción

¿Cómo funciona?

- **Comunicación anónima de baja latencia basada en circuitos**
  - **Circuitos virtuales**: garantizan que no hay conexión directa entre cliente y servidor.
  - **Saltos intermedios**: cada salto es un nodo. Cada router sólo conoce al siguiente y al anterior
  - **Tráfico cifrado por capas** → **encaminamiento cebolla**

Introducción

¿Cómo funciona?

- **Comunicación anónima de baja latencia basada en circuitos**
  - **Circuitos virtuales**: garantizan que no hay conexión directa entre cliente y servidor.
  - **Saltos intermedios**: cada salto es un nodo. Cada router sólo conoce al siguiente y al anterior
  - **Tráfico cifrado por capas** → **encaminamiento cebolla**

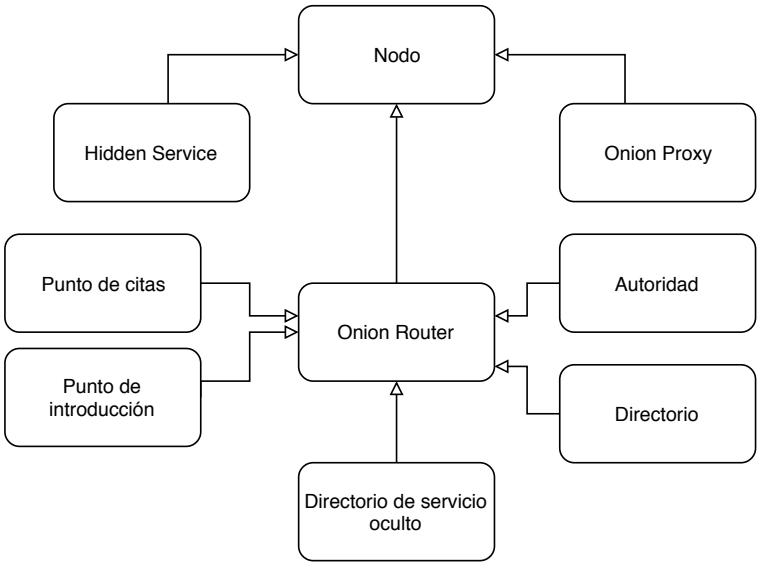
Introducción

¿Cómo funciona?

- **Comunicación anónima de baja latencia basada en circuitos**
  - **Circuitos virtuales**: garantizan que no hay conexión directa entre cliente y servidor.
  - **Saltos intermedios**: cada salto es un nodo. Cada router sólo conoce al siguiente y al anterior
  - **Tráfico cifrado por capas** → **encaminamiento cebolla**

Introducción

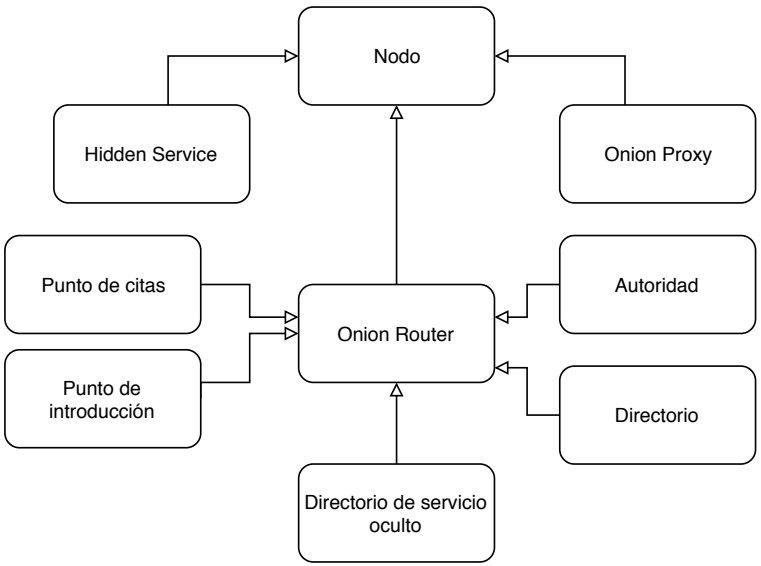
Tipos de nodos



- **Onion Proxy** (OP): **cliente Tor**
- **Onion Router** (OR)
  - **Elemento básico de la red Tor**
  - Mantenidos por voluntarios, establecen los circuitos para conectarse
- **Hidden Service** (HS)
  - **Proveen servicios únicamente accesibles a través de Tor**

Introducción

Tipos de nodos

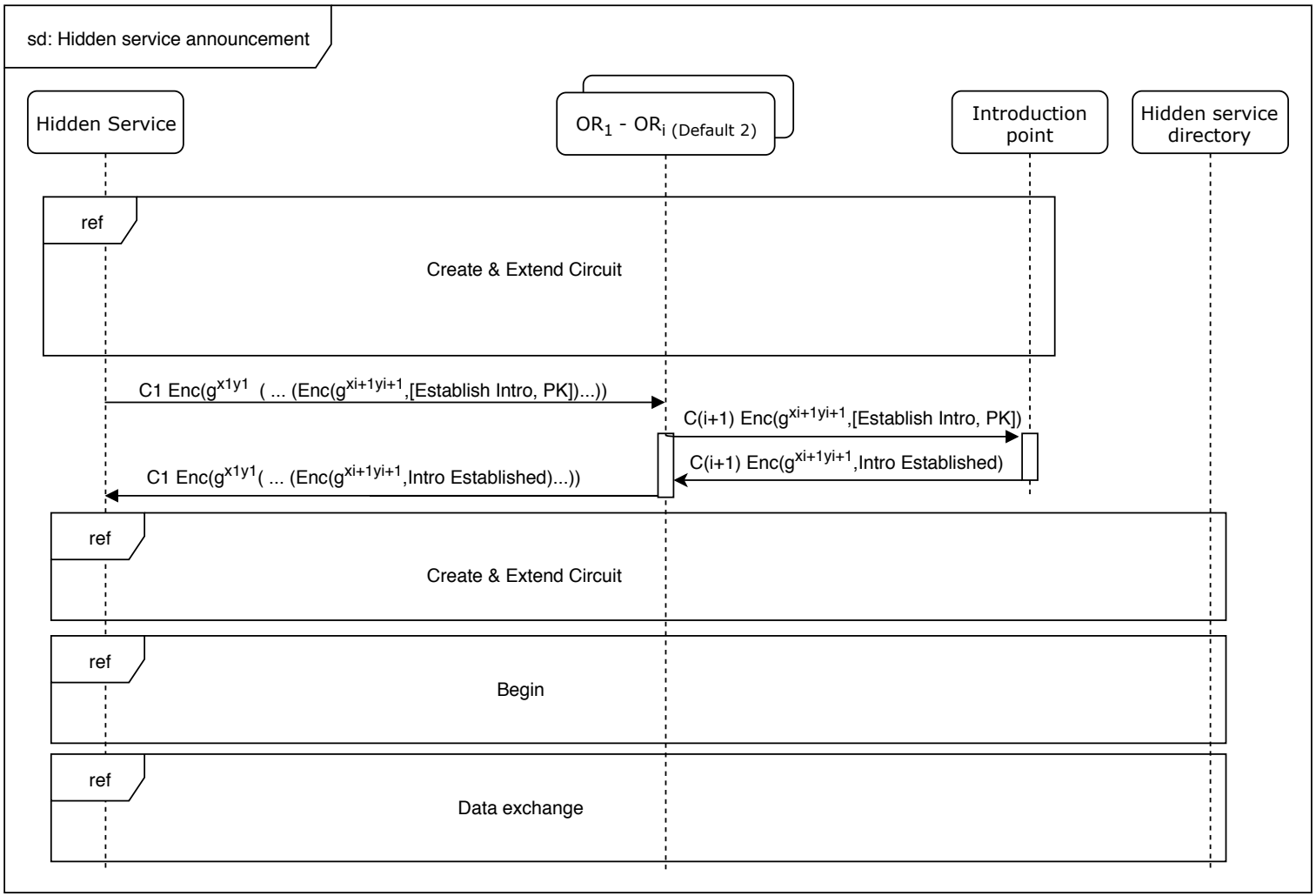


- **Directorio** (Dir): Suministra información del estado de red
- **Autoridad** (Auth): Genera el estado consensuado de la red
- **Directorio de servicio oculto** (HsDir): Almacena descriptores de HS
- **Punto de Introducción** (IP) y **Citas** (RV)
- **Puente**
- **Puente autoridad**

Agenda

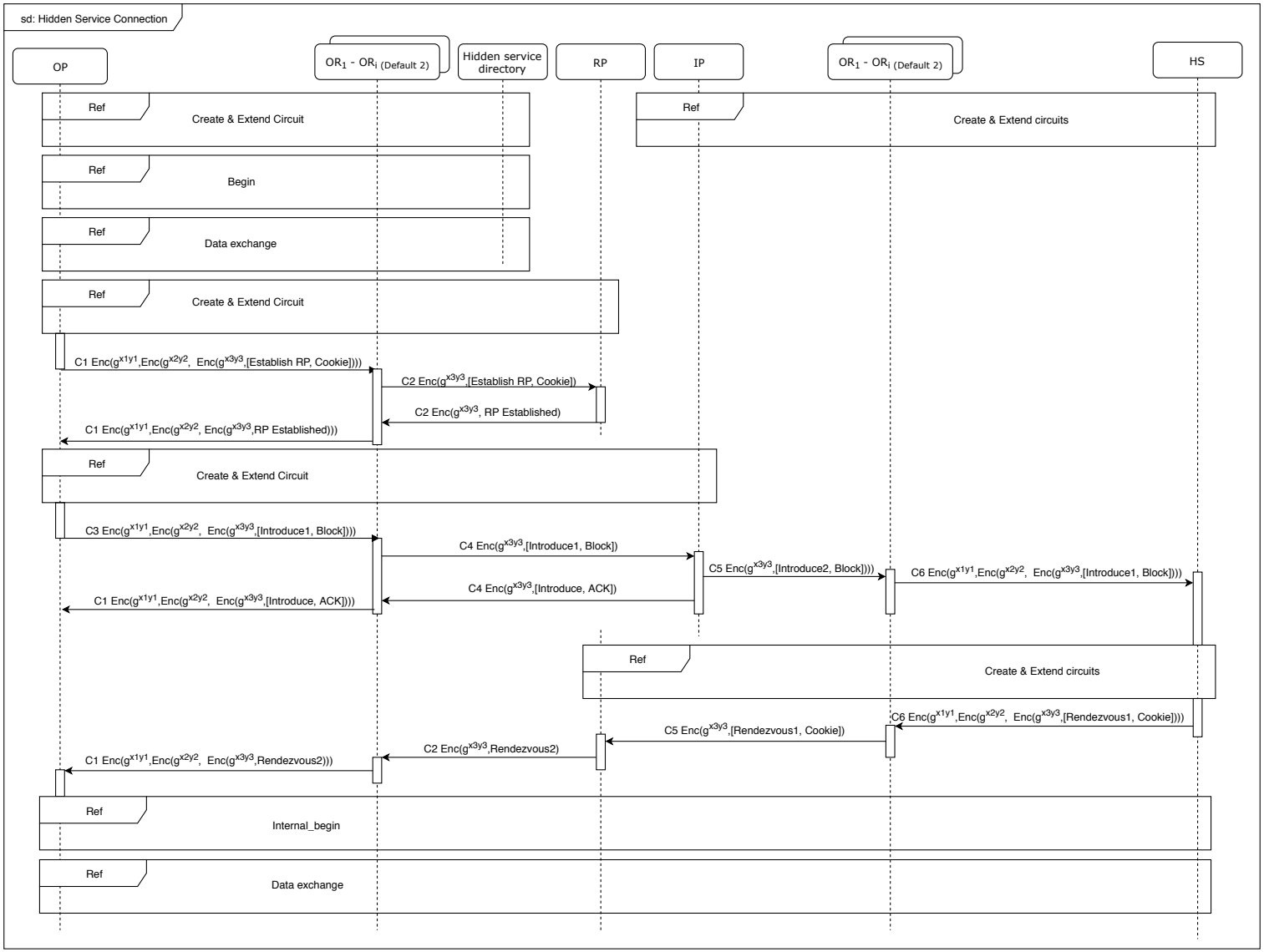
Diagramas de secuencia

Conexión a servicios ocultos – anuncio del HS

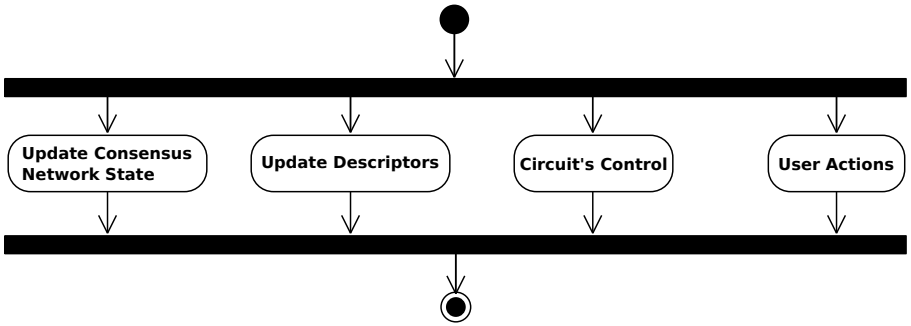


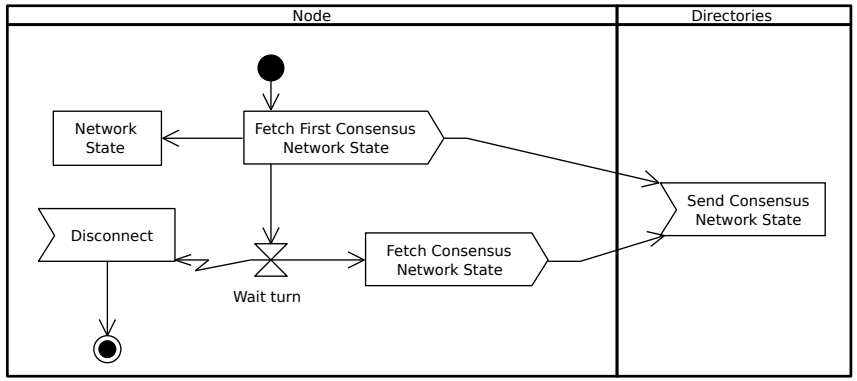
Diagramas de secuencia

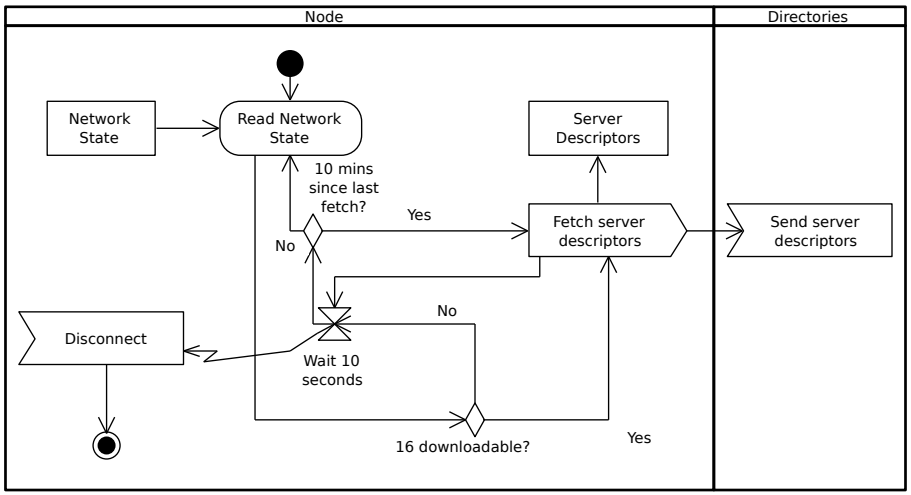
Conexión a servicios ocultos – conexión al HS

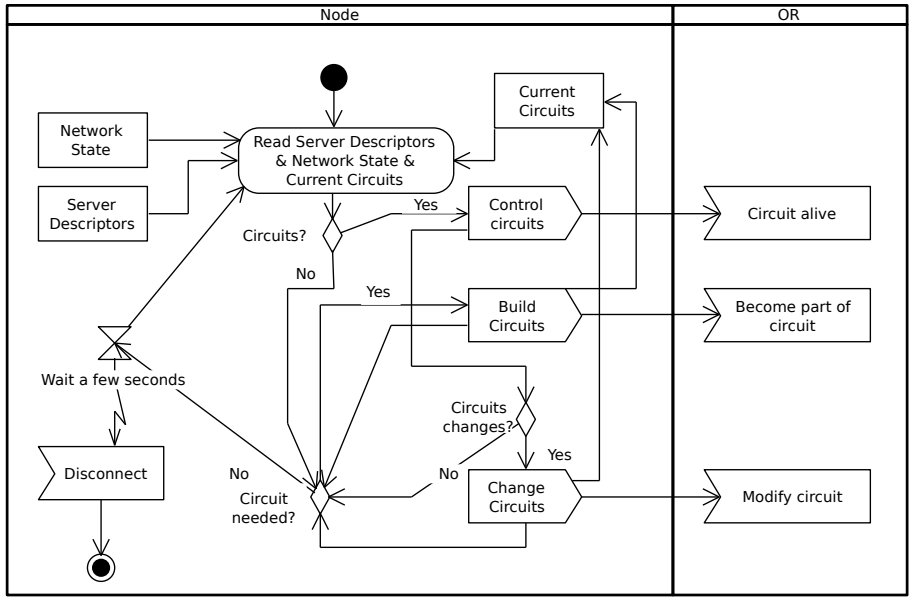


Diagramas de actividad









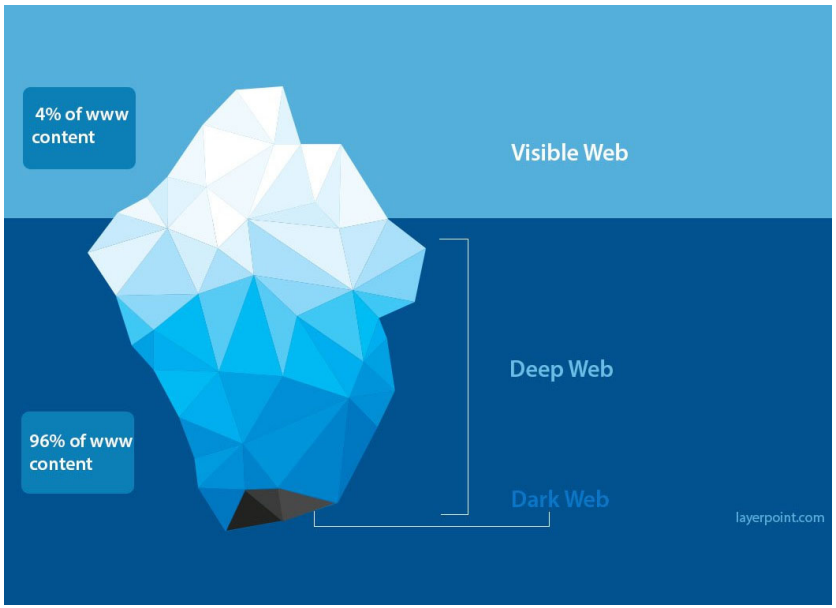
Agenda

Usos de Tor

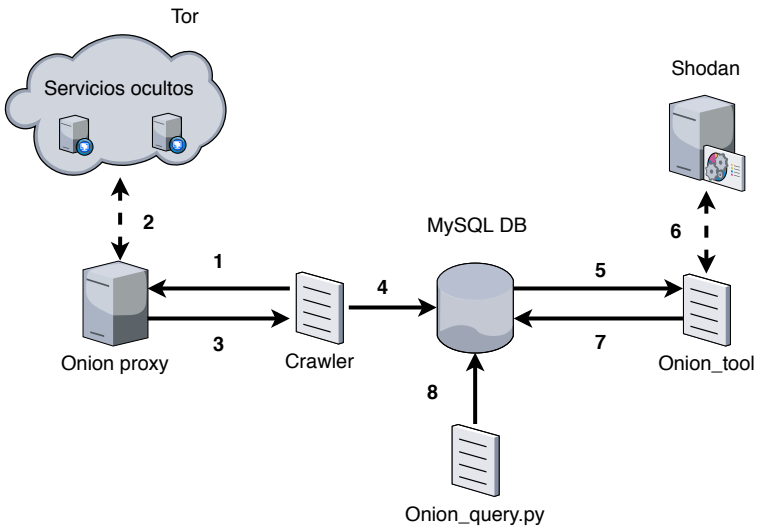
- **Usos legítimos: defensa de derechos y libertades del individuo**
  - Acceso a contenidos prohibidos en el país de origen del usuario
  - Denuncia anónima de violaciones de derechos por parte de empresas o estados

Usos de Tor

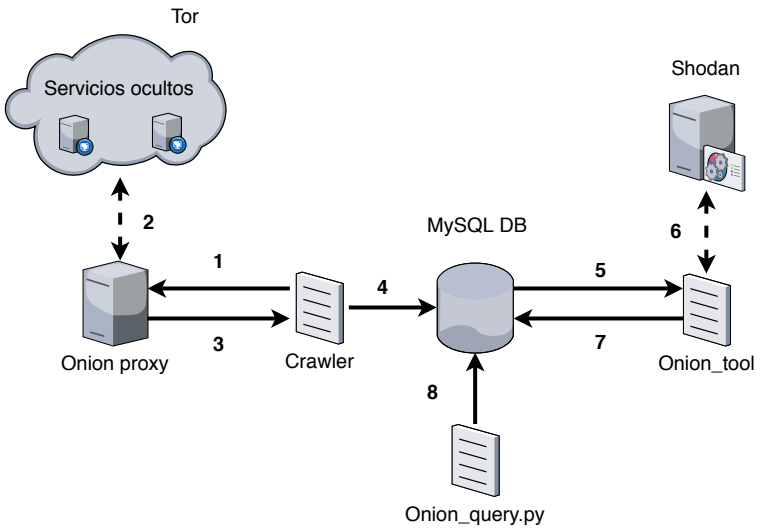
- **Usos legítimos: defensa de derechos y libertades del individuo**
  - Acceso a contenidos prohibidos en el país de origen del usuario
  - Denuncia anónima de violaciones de derechos por parte de empresas o estados
- **Usos ilegítimos: aprovechamiento por parte de delincuentes**
  - Servicios ocultos: *dark web*
  - Caso más famoso: *SilkRoad* (mercado online para compra de drogas). Creado en 2011, clausurado en 2013 por el FBI



Descripción de TorHSScanner

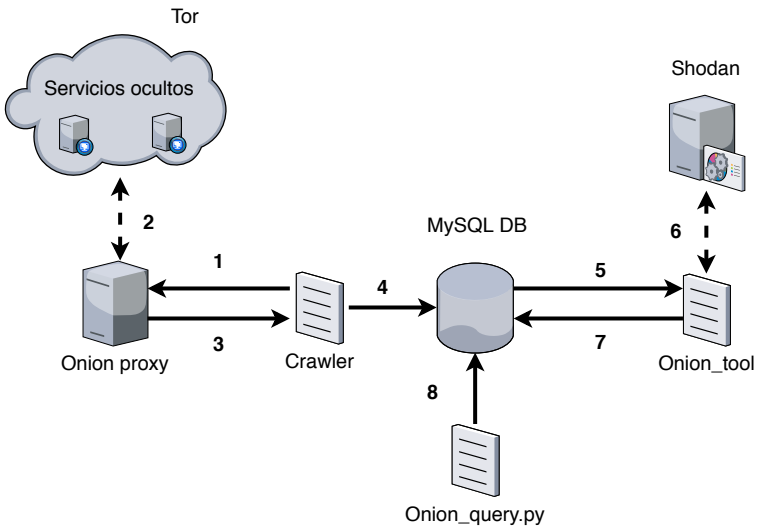


Descripción de TorHSScanner



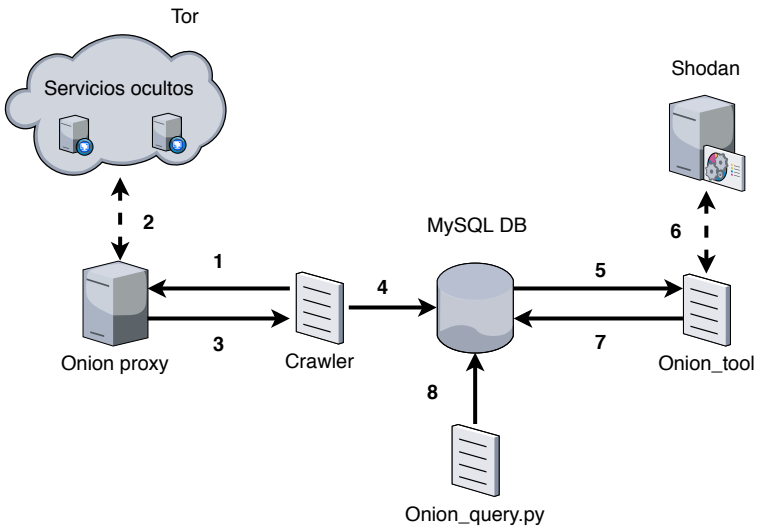
1. **Recopilación de direcciones .onion**
   - **HTTP + HTTPS requests**
   - **HTML del index y HTTP headers**

Descripción de TorHSScanner



1. **Recopilación de direcciones .onion**
   - **HTTP + HTTPS requests**
   - **HTML del index y HTTP headers**
2. **Desanonimización**
   - Metadatos de Internet (**Shodan)**
   - Algoritmo voraz para encontrar similitudes

Descripción de TorHSScanner



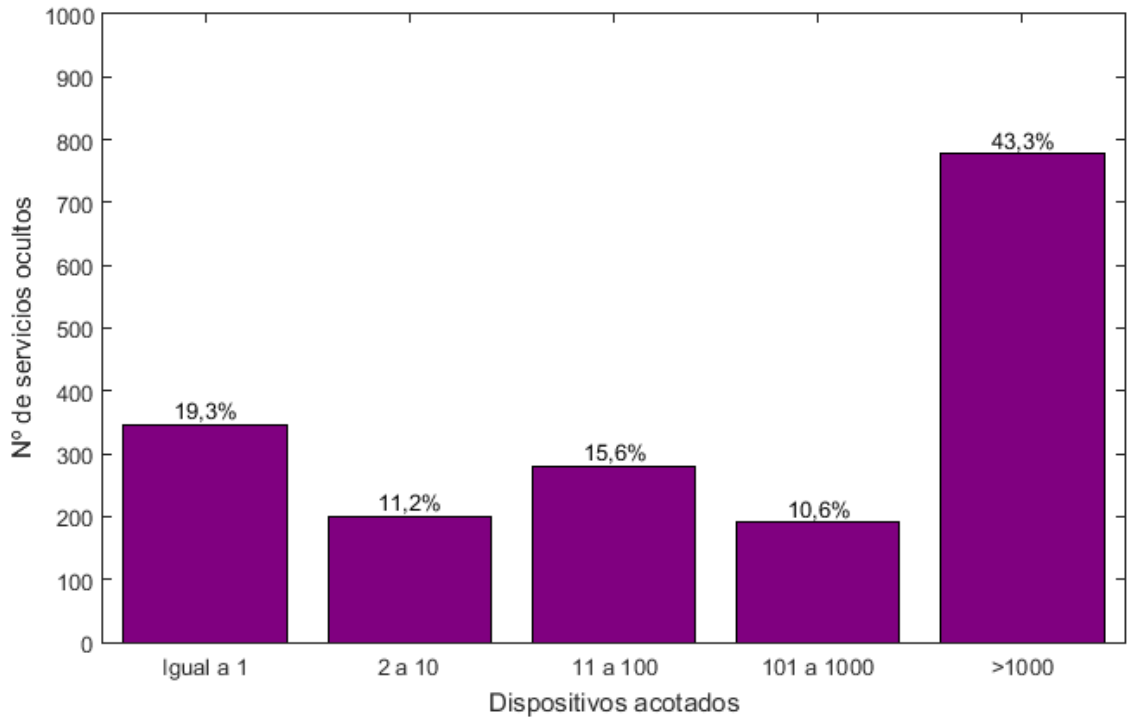
1. **Recopilación de direcciones .onion**
   - **HTTP + HTTPS requests**
   - **HTML del index** y **HTTP headers**
2. **Desanonimización**
   - Metadatos de Internet (**Shodan)**
   - Algoritmo voraz para encontrar similitudes
3. **Categorización**
   - **Categorías**: drogas, contenido-sexual, terrorismo y criptomonedas
   - **Diccionarios**

Agenda

Análisis de los servicios ocultos

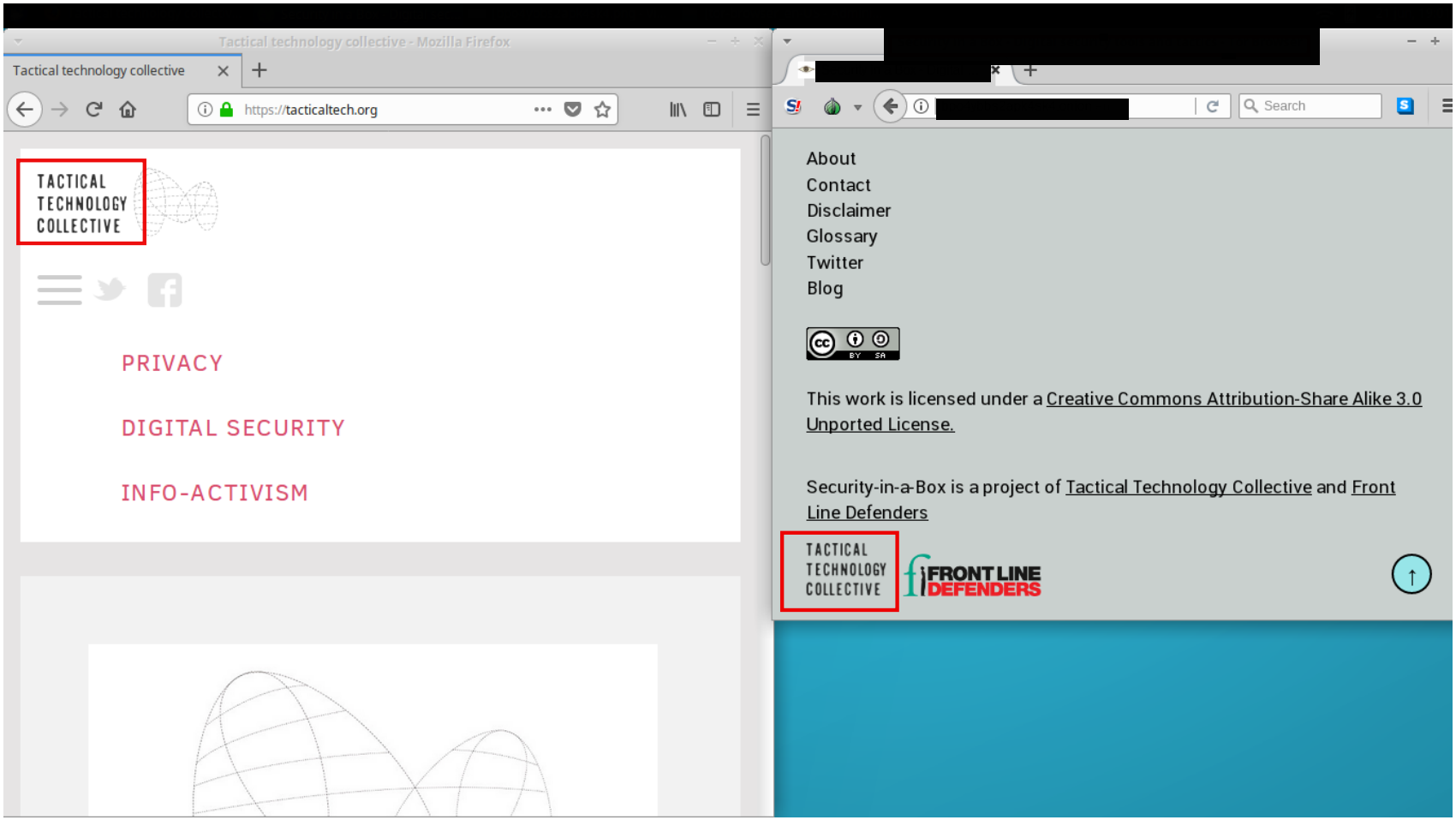
Resultados

- Resumen de resultados
  - **17328 direcciones `.onion`**
  - **1796 accesibles mediante HTTP/HTTPS**
- **Acotación de dispositivos**
  - **346 servicios acotados a un único dispositivo**



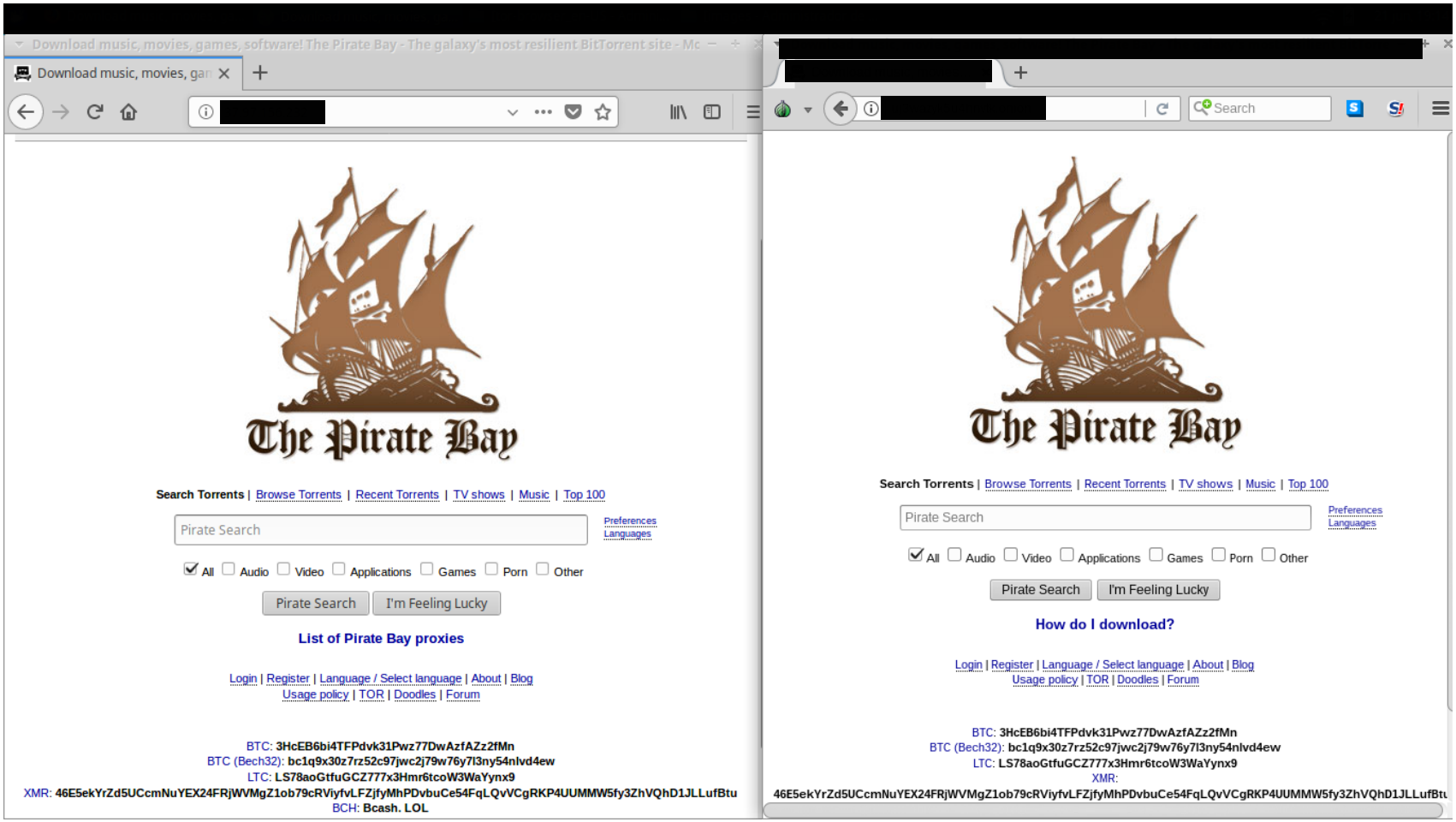
Análisis de los servicios ocultos

Ejemplos de desanonimización



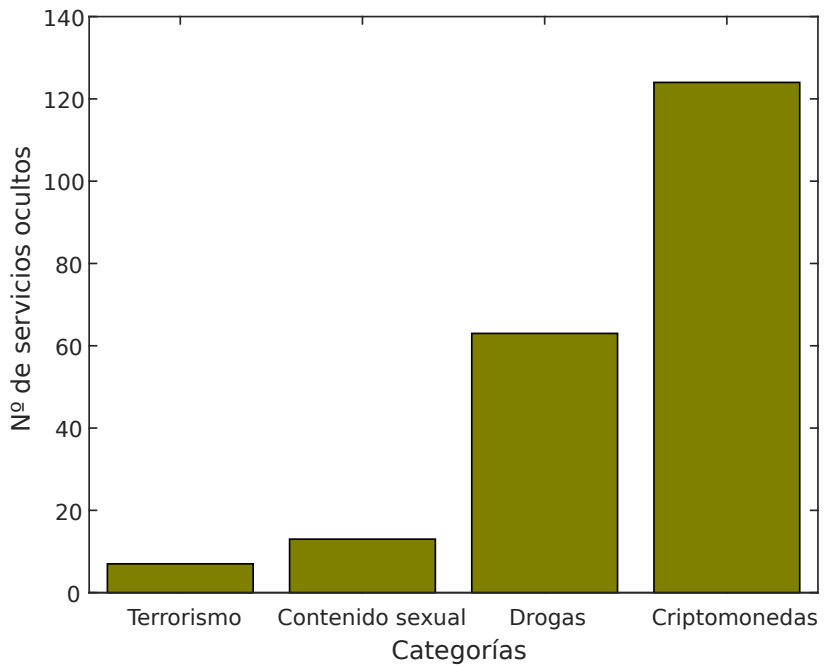
Análisis de los servicios ocultos

Ejemplos de desanonimización



Análisis de los servicios ocultos

Categorización



Idioma	# Num	Idioma	# Num
English	1313	Romanian	4
German	54	Turkish	4
Danish	38	Welsh	3
Portuguese	33	Slovak	3
Spanish	25	Swedish	3
French	19	Swahili	3
Italian	8	Tagalog	3
Norweigan	8	Vietnamese	3
Afrikaans	7	Indonesian	2
Dutch	7	Bulgarian	1
Somali	7	Estonian	1
Finish	6	Lithuanian	1
Polish	5	Albanian	1
Catalan	4	Unknown	229

Agenda

Conclusiones y trabajo futuro

- **Diagramas UML cubriendo diferentes aspectos de la red Tor**
  - **Elementos que la componen** (nodos, mensajes)
  - **Comunicación**
  - **Comportamiento de los nodos**
- **Importante cantidad de datos de servicios ocultos**
  - **Direcciones** `.onion`
  - **Información de su contenido, cabeceras y protocolo**
- **Herramienta de desanonimización y categorización**

Trabajo futuro

- **Evaluar el comportamiento con modelos formales** (i.e., Redes de Petri)
- **Sistema de evaluación para comprobar los resultados**
- **Mejora de la categorización**

Modelización del protocolo Tor y extracción de características de servicios ocultos

Jorge García de Quirós

Ricardo J. Rodríguez

Universidad Zaragoza

Universidad de Zaragoza, Spain

28 de mayo de 2019