# Comunicaciones Anónimas

Todas las redes de comunicaciones usan direcciones para realizar el enrutamiento de forma que los datos pueden ser transmitidos del origen al destino. Dichas direcciones suelen ser visibles para cualquiera que observer la red. Normalmente estas direcciones son identificadores únicos de forma que todas las transacciones relacionadas con un usuario puedan ser trazadas. Estas direcciones pueden ser asociadas con personas, lo que puede comprometer la privacidad.

Teniendo este en cuenta, anonimizar los canales de comunicación se vuelve algo necesario para poder salvaguardar la privacidad de los usuarios y proteger las comunicaciones contra el análisis de tráfico. Para ello puede ser necesaria la aplicación de técnicas de anonimización sobre la capa de aplicación como autenticación anónima, protocolos de voto anónimos o divisas anónimas.

Un sistema de comunicación anónimo oculta quien se está comunicando con quien y se pueden aplicar diferentes escenarios:

- El remitente debe ser ocultado para todo el mundo, incluyendo el receptor.
- El receptor debe ser ocultado para todo el mundo, incluyendo el remitente.
- Tanto el receptor como el remitente deben ser ocultados de terceras partes, puede que hasta tengan que autenticarse el uno al otro.

Esot significa que es necesario proveer la capacidad para que los usuarios puedan usar internet sin revelar sus identidades mientras operan con normalidad. El problema es que hay una contradicción entre privacidad personal y la aplicación de las leyes, incluyendo seguridad nacional. Se pueden establecer varios niveles de anonimidad:

- La **privacidad total** puede ser garantizada, de forma que todo el mundo es anónimo:
  - Todo el mundo usa los servicios para su propio beneficio
  - Esto incluye criminales y actores maliciosos, lo que facilita las actividades ilegales sin control alguno.
- **Privacidad parcial**, las agencias que aplican la lay pueden deshacer la anonimidad para todo el mundo.
  - Esto significa que las actividades de todo el mundo pueden ser monitorizadas
  - Hay una versión ligera en la cual las agencias solo pueden deshacer la anonimidad con una orden judicial debido a la sospecha de acciones ilegales.
- **Ausencia de Privacidad**: Todo el mundo puede observar todo
  - No hay privacidad, todas las actividades son públicas.

Existen varias definiciones que pueden ser aplicadas a los sistemas de comunicación:

- **Anonimidad**: El estado de no ser identificable en un conjunto de sujetos. Requiere un conjunto de anonimidad, donde varios sujetos tienen potencialmente los mismos atributos. En cuanto a comunicaciones, los conjuntos de anonimidad consisten en sujetos que pueden ser localizados para enviar o recibir transmisiones.
- **No-Enlazabilidad**: Significa que un usuario puede usar cualquier recurso o servicio sin que sea posible enlazar los usos juntos. No se puede determinar si un usuario realizó una operación o no. Si dos elementos o acciones de interés son observadas e inspeccionadas por un atacante, no están más o menos relacionadas que las acciones anteriores.
- **No-Observabilidad**: El estado de un objeto de interés es indistinguible de cualquier otro elemento de interés. Un mensaje no puede ser diferenciado de ruido aleatorio. No se puede

identificar cuando se ha intercambiado un mensaje.
- **Pseudoanonimidad**: Se usa un pseudónimo como identificador. Puede ser asociado a un individuo. Permite reclamar responsabildiades en caso de mal comportamiento.

Existen varios modelos de ataque sobre redes de comunicaciones:

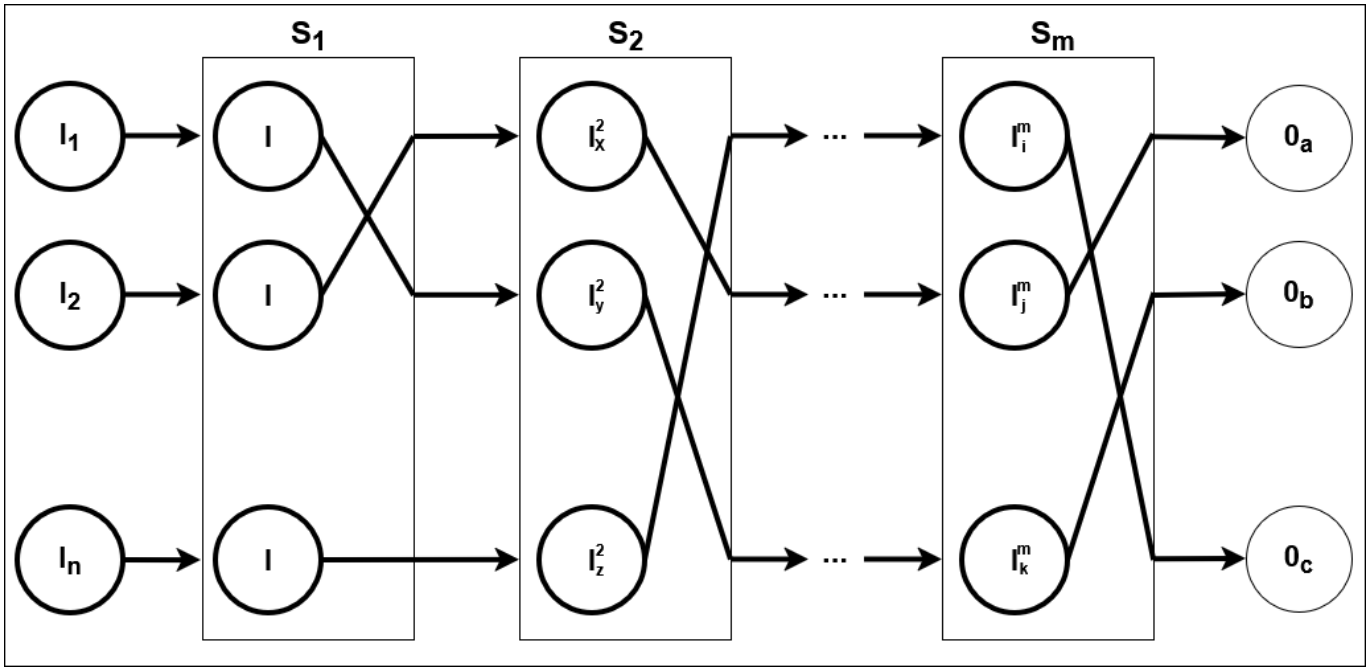
- **Tipo 1 (Atacante pasivo)**: Observalas comunicaciones y enlaces
- **Tipo 2 (Atacante pasivo con capacidades de envío)**: Además de observar las comunicaciones, el atacante puede tomar parte en el proceso emitiendo mensajes.
- **Tipo 3 (Atacante activo)**: Puede controlar todos los enlaces de comunicación, eliminar, responder, enviar o retrasar mensajes.

Requerimientos para la anonimidad en redes de comunicaciones:

- **Tráfico de cobertura**: Se envía tráfico adicional con el mensaje de una persona para enmascarar la transmisión. Si un atacante controla el tráfico de cobertura no se puede asegurar la anonimidad.
- **Tráfico embebido**: El tráfico generado por un usuario debe ser introducido de forma silenciosa y adecuada dentro del tráfico de cobertura de forma que un atacante no lo pueda distinguir. Esta función suele ser realizada por una tercera parte que agrupa varios mensajes y los embebe en una transmisión con otros. Solo se alcanza la anonimidad si hay al menos 2 partes honestas que trabajen juntas. Si hay N participantes y N-1 son deshonestos, no se puede asegurar la anonimidad.
- **Efectividad**: Si tenemos N mensajes de diferentes usuarios. K mensajes son reales, mientras que $M=N-K$ son tráfico de cobertura. La efectividad del sistema es definida por $K/N$ donde 1 es el caso óptimo, donde todos los mensajes son reales y no es necesario tráfico de cobertura. Para alcanzar el sistema más óptimo tienen que existir N usuarios diferentes que transmitan N mensajes diferentes, para ello es necesario esperar hasta que hayan suficientes mensajes, lo que puede demorar las comunicaciones de la red.

# Redes MIX

El diseño de redes mix fue creado para implemetar sistemas de correo electrónico anónimos.




- Si todos los nodos son honestos y siguen el protocolo, las salidas son permutaciones de las entradas y el contenido del mensaje se mantiene intacto.
- Si al menos uno de los nodos oculta el intercambio que hace, este es secreto, demodo que la correspondencia entre las entradas y salidas es desconocida.
- La honestidad de los nodos puede ser verificada públicamente de forma que se puede garantizar que las salidas son permutaciones de las entradas.
- Las redes mix deben funcionar correctamente incluso si falla uno de los nodos o si alguno de los nodos está comprometido.

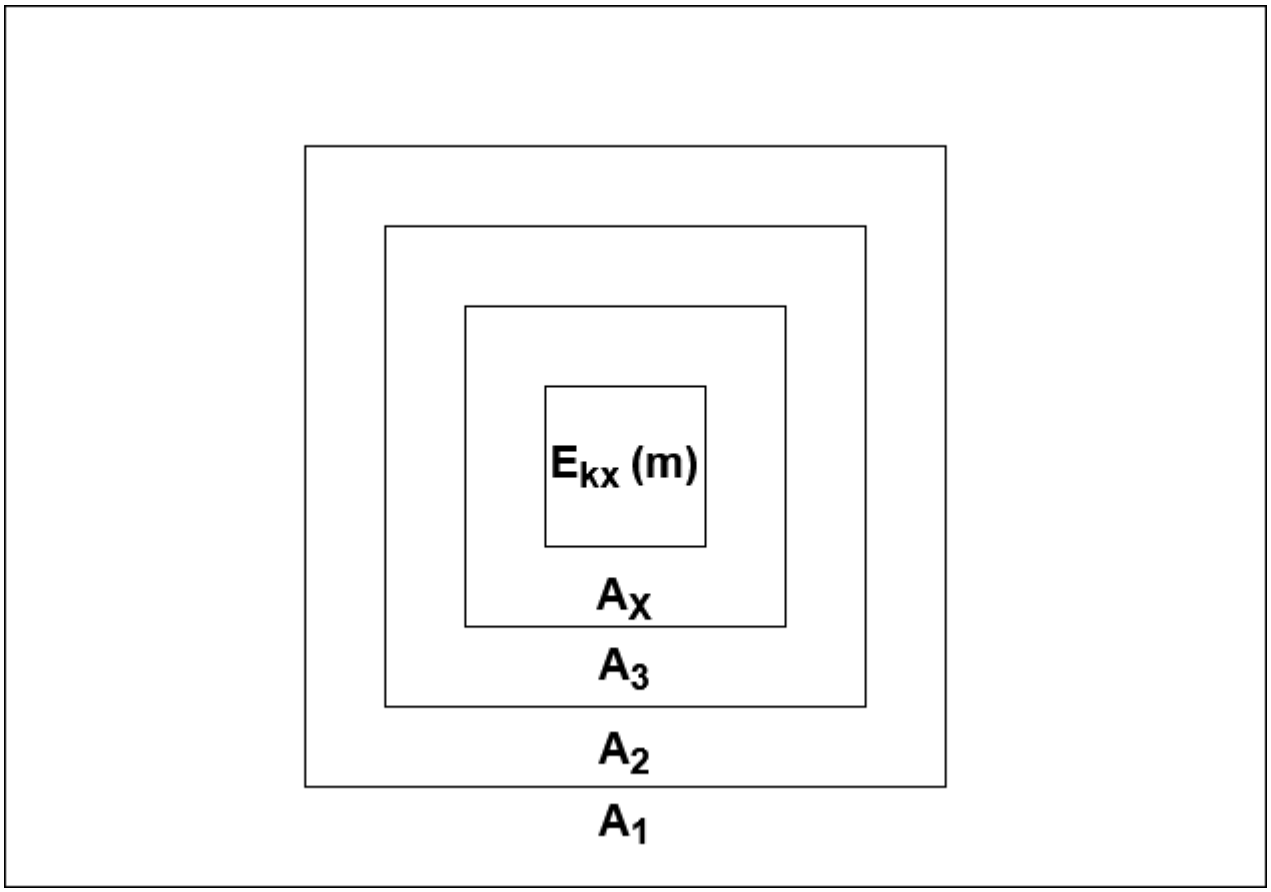
# Modos de procesado

## Cadena de descifrado

Esquema basado en criptografía RSA. Cada entrada ($m$) es secuencialmente cifrada usando la clave pública ($K_i$) de cada nodo.

$E_k(m,r) = A_1 || E_{k1}(A_1 || E_{k2}(A3 || ... E_{kn}(A_x || E_{kx}(m) || r_n) ... || r_2) || r_1)$

Donde $A_i$ es la dirección de la etapa $i$ y $r_i$ es una cadena de caracteres aleatoria usada para aleatorizar el cifrado de la capa $i$




- A través de la red mix se realizan 2 operaciones en cada etapa:
  - Cada nodo $i$ usa su clave privada $K_i^{-1}$ para eliminar una clave de cifrado de cada una de sus entradas
  - Estos mensajes parcialmente descifrados se permutan en la etapa $i$ antes de enviarse a la siguiente etapa con un orden aleatorio
- La comunicación en dos direcciones es posible incluyendo un RPI (Return Path Information) y claves asimétricas compartidas $K^s_s$ junto con el mensaje $m$ donde $A_i$ es la dirección de la etapa $i$, $r_i$ es una cadena aleatoria usada para aleatorizar el cifrado de la capa $i$ y $K_i^s$ son claves simétricas compartidas entre el remitente y la etapa $i$.
  - Por lo tanto, el receptor recibirá la siguiente salida de la red mix:

$E_kx (m || RPI || K_s^s)$

- Tras el descifrado, el receptor envía la respuesta $m'$ a $A_n$ como:

$E_{k_s^s} [m'] || RPI$

- El receptor no conoce el camino hasta el receptor, por lo que no puede cifrar el mensaje de la forma típica, solo puede hacerlo usando la clave $K^s_s$
- Cada etapa retira una capa de RPI, obteniendo la siguiente dirección y la clave de la siguiente dirección $K^s_i$, la cual se usa para volver a cifrar el mensaje restante.

### Cadena de recifrado

Tipo de red mix basada en el sistema criptográfico ElGamal. El remitente cifra el mensaje $m$ usando la clave pública $K$ de la red mix:

$$E_k(m,r) = g^r || (A_x || m)K^r$$

Donde $g$ es un generador y $r$ es una cadena aleatoria. La clave pública K puede ser definida como:

$$K = ∏^n_{i=1} K_i = Π^n_{i=1} g^{d_i} = g^{Σ^n_{i=1}d_i}$$

Donde $K_i = g^{d_i}$ y $d_i$ so las claves públcia y privadas de la etapa i.

- Cada nodo i recifra sus entradas usando cadenas aleatorias para cambiar su apariencia.
- Como en la cadena de descifrado, todos los mensajes son procesados y permutados antes de ser enviados al siguiente nodo.
- Una vez las n etapas son completadas, entonces comienza una fase de descifrado donde $K = g^d$ es la clave pública y $d$ es la clave privada que se comparte entre las n etapas.
- La comunicación bidireccional también es posible enviando 3 cifrados ElGamal a la red mix donde $K$ es la clave pública de la red, $A_s$ y $A_x$ son las direcciones del remitente y el receptor, y $K_s$ y $K_x$ son las claves públicas para ElGamal del emisor y el receptor.

### Comparación de Cadena de Cifrado y Cadena de Recifrado

| Tipo | Ventajas | Limitaciones |
|---|---|---|
| Cadeba de descifrado | Se pueden incluir direcciones intermedias para el enrutado | El emisor debe realizar mútilples cifrados<br>Todas las etapas deben participar en un orden espcífico.<br>Las entradas pueden ser trazadas por apariencia/tamaño |
| Cadena de Recifrado | El emisor realiza solo un cifrado<br>Las entradas no pueden ser trazadas por tamaño/apariencia<br>No es necesario que todas las etapas participen y el orden no importa | No puede incluir direcciones intermedias para el enrutado<br>Las etapas tienen que colaborar en la fase de descifrado. |

## Topologías

### Cascada

Consiste en una secuencia fija de etapas las cuales son comunes para cada emisor o receptor.

- La primera etapa inizializa la mezcla de los mensajes de todos los emsiores, agrupándolos en grupos de tamaño $l$ los cuales se procesan de forma síncrona.
- Una etapa defectuosa puede comprometer toda la red
- Vulnerable a ataques pasivos mediante el trazado de mensajes a través de los grupos mezclados. Cuanto más ghrande sea el grupo, más difícil es seguir el mensaje.
- Vulnerable a ataques activos, pero el atacante debe tener cierto control sobre la red mix
    - Debe controlar múltiples etapas para poder trazar mensajes.
    - Tiene que manupular las entradas de la red para inyectar tráfico controlado.

### Enrutado Libre

Consiste en una serie de etapas interconectadas que no tienen por que ser dependientes.

- Cualquier etapa puede recibir entradas de los emisores, también pueden mandar una entrada directamente a un receptor.
- Una etapa debe también recibir las salidas producidas por cualquier otra etapa a la que este conectada.
- Cada etapa debe esperar hasta tener $l$ mensajes para agruparlos, pero solo por un período de tiempo fijo, pasado dicho período, los mensajes se pasan a la siguiente etapa.
- Este modo de operación es asíncrono ya que los mensajes pueden quedar estancados hasta que se cree un grupo o pueden ser enviados antes dependeindo de su ruta.
- Vulnerable a ataques pasivos, se pueden trazar los mensajes debido al tráfico no uniforme entre etapas, además de por el decremento de capas en cada etapa.
- Vulnerable a ataques activos, se puede inundar la red para aislar un objetivo.

## Esquema de verificación

Verificar una red ivolucra el análisis de como de correcto es el procedimiento de acuerdo a los siguientes criterios:

- El grupo de entrada ha sido procesado/transformado y permutado de forma apropiada
- El mensaje no ha sido corrompido
- No se han añadido/perdido entradas

La verificación se puede realizar a varios niveles:

- Salida de la red entera
- Salidas de cada una de las etapas

Estos mecanismos en general no se pueden aplicar a las redes mix de enrutado libre

### Red mix de emisor verificable

Este esquema solo detecta mensajes corruptos en la salida de la red. Para implementar este esquema de verificación el emisor incluye un cheksum con el mensaje:

- Tanto el checksum como el mensaje son cifrados con la clave de cifrado del emisor

- Cualquiera puede detectar si el mensaje ha sido alterado al descifrarlo con la clave pública del emisor.

No detecta la adición de mensajes. La eliminación de mensajes puede ser solo detectada si el emisor revisa si su propio mensaje está presente en la salida de la red. No se pueden identificar las etapas comprometidas.

### Red mix de etapa verificable

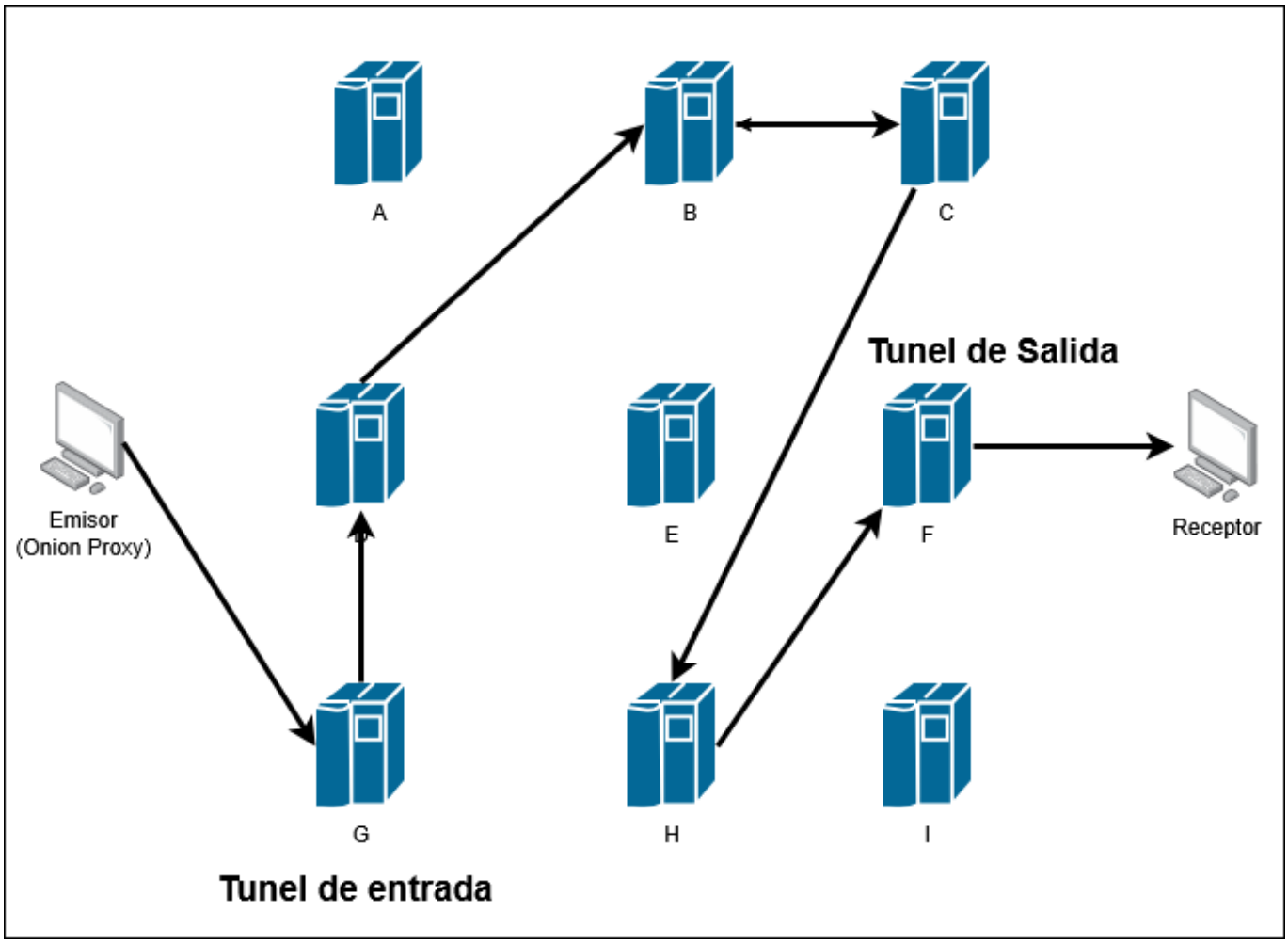
Cada estapa verifica las salidas de la red usando protocolos adicionales para asegurar que todo se hace de forma correcta.

- Se crean numerosas copias del grupo de entrada o se repite toda la mezcla para detectar etapas comprometidas en las cadenas de recifrado.
- Revelando secretos o usando pruebas de cero conocimiento de forma que las operaciones en ciertas etapas pueden ser verificadas por otros.
- Mecanismos de recuperación para reiniciar las operaciones en caso de que se detecten comportamientos extraños.

# Onion Routing

Esquema basado en la idea de enrutado a través de varios nodos con varias capas de cifrado. Un mensaje es cifrado capa por capa usando las calves de todos los nodos que hay en el camino al receptor. Cada nodo deshace una capa descifrandola de forma que la dirección al siguiente salto es revelada y envía el resto del mensaje al siguiente nodo. Es el opuesto a las redes mix, no oculta el tráfico de red mezcalndo mensajes. Este mecanismo está compuesto por las siguientes entidades:

- **Aplicación cliente** (Envía el mensaje)
- **Onion Proxy**: Determina el camino desde el origen hasta el destino (n rutas diferentes), donde el primer nodo es llamado tunel de entrada y el último, tunel de salida.
- **Routers**: Deshacen una capa del mensaje y envían el resto del mensaje al siguiente nodo hasta que se alcanza el tunel de salida. En el caso de la respuesta se realiza la operación contraria, se van añadiendo capas de cifrado.
- **Tuneles de entrada y salida**: En tunel de entrada puede ver e interactuar con el emisor mientras que el de salida puede hacerlo con el receptor.




# La Idea Original

El procedimiento es similar al del esquema de cadena de descifrado de las redes mix con la excepción de que el número de nodos usados y su orden no es fijo:

$$O = A_1 || E_{k_1} (A_2 || E_{K_2}(A_3 || ... E_{k_n}(A_x || m)))$$

Donde $A_i$ es la dirección del router $i$. En este caso, si el receptor responde al mensaje, la red es capaz de mantener un estado que permita enrutar de vuelta la respuesta al mensaje a través del mismo camino. La respuesta es construida por todos los routers, que añaden una capa de cifrado en cada paso usando sus claves privadas. Una vez que todos los routers han completado sus operaciones, el tunel de entrada envía la entrada al onion proxy del emisor:

$$R0 = E_{sk_1}(E_{sk_2}(... E_{SK_n}(r)))$$

# TOR

La implementación más popular del Onion Routing es TOR, que es considerada una versión evolucionada de la idea original, incluyendo algunas mejoras y mecanismos que lo hacen utilizable en aplicaciones reales. No es una red P2P, los usuarios no actuan independientemente, son usuarios que se unen a la red usando una aplicación. La red opera a través de routers que son proveidos por organizaciones e individuos que donan su capacidad de procesado (nodos) al proyecto.

Algunos routers también pueden actuar como directorios que muestran el estado de la red. No es seguro contra ataques de punto a punto, si la misma entidad contral el primer y último nodo de un camino puede inferir el emisor, el receptor y el contenido del mensaje. Tampoco oculta la identidad del emisor a nivel de aplicación. Construye los circuitos de forma diferente a la idea original:

- Los circuitos de TOR son siempre de 3 nodos:
  - Primero se selecciona el tunel de salida ya que algunos pueden filtrar tráfico específico o puede no tener capacidad suficiente
  - Un nodo se selecciona solo una vez.
  - Para evitar que una entidad controle el túnel de entrada y el de salida, se aplican algunas reglas.
  - El primer nodo es un “Nodo Guardia” que ve la IP del emisor, el cual los routers consideran confiable.
  - Los circuítos son rotados periódicamente.

Cada router tiene una clave de identidad de largo plazo y una clave onion de corto plazo. La clave de largo plazo se usa para firmar certificados TLS que se usan para comunicarse con otros routers Onion y proxies. La clave de corto plazo es rotada periódicamente y se usa para construir circuitos y negociar claves para el cifrado y descifrado de los mensajes.

La red TOR construye los circuitos de una forma diferente a la idea original. El camino $A$ → $C$ es creado siguiendo una aproximación interactive e incremental.

- Cada segmento del camino es construido o destruido progresivamente usando mensajes de control para crear, extender o eliminar el circuito.
- Cada segmento es identificado de forma indemendiente a trawvés de un identificador de circuito: