Universidad Complutense de Madrid
Facultad de Informática
Departamento de Ingeniería del Software e Inteligencia Artificial

Proyecto Fin de Máster en Investigación en Informática

Director de proyecto: Luis Javier García Villalba

Sistemas Anónimos en Escenarios Globales

Rodolfo Leonardo Sumoza Matos

rlsumoza@fdi.ucm.es

Madrid, Septiembre 2008

Agradecimientos

...sabemos que la dosis de privacidad en la virtualidad se mide. Ocultarse detrás de un monitor para molestar a alguien, ya no es tan buena idea ya que esa unidad te delata... como si te viera desde mi propio oráculo.

Bettina Perroni

Tengo una deuda impagable con Dios, a quién le debo cada uno de los pasos que he dado en mi vida.

Estoy especialmente agradecido con mi familia, mi hija Sofía y mi esposa Rosa, por su invaluable comprensión y paciencia durante el desarrollo de este proyecto. Ellas son mi fuente de inspiración.

Agradezco a la Universidad Simón Bolívar de Venezuela, por darme la oportunidad de continuar mi formación académica a través de su apoyo institucional y financiero.

A la Universidad Complutense de Madrid le agradezco el permitirme realizar los estudios del Máster en Investigación en Informática aprovechando los recursos y servicios ofrecidos.

A mi tutor, el profesor Luis Javier García Villalba le agradezco su particular guía y soporte durante esta etapa de formación.

Este trabajo ha sido elaborado con el apoyo del Programa Alban, Programa de Becas de Alto Nivel de la Unión Europea para América Latina, beca no. E07D401826VE.

Índice general

Índice de figuras

Índice de tablas

Resumen

Este proyecto contribuye con el campo de la implementación de los sistemas anónimos a escala mundial. Al mencionar las bases conceptuales de las tecnologías que mejoran la privacidad (*Privacy enhancing Technologies*), se consigue afianzar el conocimiento sobre las técnicas hasta ahora utilizadas. Se propone una solución al problema del descubrimiento de nodos en la implementación de los sistemas anónimos en escenarios globales en zonas donde se censuran y bloquean las comunicaciones, para lo cual se integran los mecanismos de los sistemas de reputación y confianza, la teoría de los pequeños mundos y la subcultura carcelaria. Se propone un mecanismo para los sistemas anónimos basado en los procesos markovianos, el cual procura conseguir los niveles de anonimato y latencia deseados. Se plantea la utilización de un modelo matemático que permite optimizar estos sistemas en relación a estos índices.

Capítulo 1

Introducción

No es necesario apelar a los artículos de la carta sobre los Derechos Humanos establecida por la Organización de las Naciones Unidas para darse cuenta que cada una de las personas que habitan este planeta tiene el derecho de decidir sobre el destino de su información privada. Esto incluye no sólo decidir quién, cómo, dónde y cuándo terceras partes puedan tener acceso a sus datos en general, sino que se debe prestar una particular atención a los que están relacionados con la identidad, el perfil social, cultural, personal, etc.

Tanto en las organizaciones privadas, como en las públicas, y a nivel individual, la protección de la información no sólo debe incluir los aspectos típicamente enmarcados dentro de la integridad, confidencialidad y disponibilidad de los datos, sino que debe ampliarse al resguardo de la privacidad donde, entre otros, se procura evitar que se devele la identidad de las partes comunicantes. Se han desarrollado varias estrategias, mecanismos, técnicas y sistemas que tienen esto como objetivo, y que pueden enmarcarse en lo que se denomina las tecnologías que mejoran la privacidad (*Privacy Enhancing Technologies*).

Este tipo de tecnologías han tenido sus frutos en escenarios de diversa índole, que van desde aplicaciones militares, donde se procura evitar que el enemigo pueda descubrir las conexiones estratégicas, pasando por aplicaciones científicas/comerciales, que evitan revelar información sobre las comunicaciones hechas entre socios científico/comerciales, hasta las aplicaciones de particulares que le ayudan a mantener en privado sus datos personales: los referentes a su salud, su estado financiero, sus preferencias de consumo, etc. Uno de los puntos críticos de la privacidad

es el encubrimiento de la identidad de las partes comunicantes, es decir, es la procura de que las comunicaciones sean anónimas: *anonimato*.

Cada una de las técnicas y mecanismos utilizados tienen sus ventajas y desventajas en cuanto al perfil de ataque considerado. Es decir, dependiendo del tipo de atacante que se considere, cada una de éstas posee un conjunto de fortalezas y debilidades asociadas. Adicional al perfil del atacante, se debe incluir su radio de acción, esto quiere decir, que se debe considerar su capacidad para manejar ciertos volúmenes de usuarios, su heterogeneidad, su distribución y localización. Además se debe considerar el tipo de comunicación anónima que se desea o necesita entablar: mensajería instantánea, correos electrónicos, servicios web, etc.

1.1. Enfoque y propósito

En este trabajo se considera el estudio de los sistemas anónimos a escala global, es decir, se consideran el manejo de grandes volúmenes de usuarios, con necesidades muy diferentes, distribuidos irregularmente sobre el planeta, y localizados en sitios muy heterogéneos que van desde zonas (organizaciones, regiones, o países) muy permisivas y con muchas facilidades en cuanto a sus comunicaciones hasta zonas restrictivas por sus características políticas, religiosas, etc.

Como uno de los puntos centrales de este trabajo se trata el problema de las comunicaciones en zonas restrictivas donde se censura y bloquea la información que se considera contraria a cierta ideología política o religiosa, y se proponen algunas técnicas para hacerlas factibles. Especialmente, en estas zonas, el principal inconveniente está en el proceso del descubrimiento de los puntos de acceso a las redes anónimas, es decir, suponiendo que se tienen redes anónimas establecidas fuera de las fronteras de las zonas restringidas, y considerando que para poder comunicarse fuera de estas fronteras se requieren puntos o nodos específicos que permitan el envío de información hacia el exterior, el principal problema es establecer un proceso estable y seguro que le permita a los usuarios descubrir y comunicarse a través de estos nodos o puntos de acceso con las redes anónimas, esto es lo que se le denomina el *problema del descubrimiento*.

Dado que aún no se han diseñado y, por supuesto, no se han implementado sistemas para escenarios globales, otro de los problemas tratados en esta propuesta es el de tener un mecanismo que pueda permitir tener un nivel deseado de anonimato y un nivel mínimo de latencia, a

escala mundial. Para esto se pueden utilizar algunas de las técnicas hasta ahora propuestas por diversos autores, combinándolas con algunos mecanismos adicionales que procuren alcanzar estos objetivos. Para poder probar este planteamiento, junto con los que se utilizan en otros sistemas anónimos, se propone un modelo a través del cual se pueda calcular el óptimo relacionado con la eficiencia de los sistemas en cuanto a su nivel de anonimato con respecto a la latencia obtenida en las comunicaciones.

Los objetivos de este trabajo se pueden sintetizar de la siguiente manera:

- Proponer mecanismos que contrarresten los sistemas de bloqueo y censura. Esto implica resolver el problema de la dinámica del descubrimiento de nodos iniciales con los cuales se pueda acceder a las redes anónimas.
- Proponer mecanismos que proporcionen mayores niveles de anonimato y que no incrementen considerablemente la latencia.
- Proponer modelos para la optimización en relación al rendimiento y/o desempeño que permitan decidir cuál es la mejor combinación de mecanismos en la implementación de sistemas anónimos.

Para conseguir los objetivos planteados en el capítulo 2 se muestra un compendio de términos, técnicas, mecanismos y sistemas utilizados en las tecnologías que mejoran la privacidad. En el capítulo 3 se muestran los fundamentos teóricos o "building blocks" de la propuesta, sobre los cuales se basan la ideas para solventar los problemas antes indicados. Para el problema de los entornos restrictivos se propone utilizar los fundamentos de los sistemas de reputación y confianza, la teoría de los pequeños mundos y la subcultura carcelaria como modelo. Para el problema de la falta de mecanismos para sistemas anónimos globales estables (con latencia relativamente baja y nivel de anonimato relativamente alto) se propone uno basado en los procesos markovianos. En cuanto a la medición de la eficiencia se propone utilizar las técnicas de la optimización multiobjetivo. En el capítulo 4 se muestra un estado del arte en cuanto a los trabajos hechos sobre sistemas anónimos en grandes escalas, sistemas resistentes a la censura y al bloqueo, y sobre los sistemas que utilizan técnicas de reputación y confianza. El capítulo 5 se presenta la propuesta que indica el cómo utilizar los "building blocks" integrándolos en un mismo mecanismo.

Capítulo 2

Marco Conceptual

2.1. Terminología

Con el fin de sentar las bases conceptuales, en parte debido a que este campo del conocimiento tiene un inicio reciente como área de estudio, y en parte para enmarcar el contexto de este trabajo, en este capítulo se muestran las definiciones más importantes asociadas a las tecnologías dedicadas a la privacidad de la información.

Tal como se menciona en [50] las personas en general utilizan Internet para poder comunicarse, para el envío de correo electrónico, para la investigación en diversas áreas de interés, para la interacción con distintos organizamos públicos y/o privados, etc. Al mismo tiempo, gran cantidad de estos organismos públicos y privados en distintas regiones del planeta buscan maximizar la interacción electrónica en todos los niveles entre los usuarios y sus centros tecnológicos, intercambiando información a través del uso de bases de datos controladas por ellos mismos, buscando así utilizar el poder de la informática para tener el control de la información concerniente a innumerables aspectos relacionados a los individuos, tales como las preferencias en sus consumos diarios, la interacción con su alrededor, sus estilos de vida, sus opiniones, sus preferencias, y todo esto en niveles que en gran medida son desconocidos por los mismos usuarios.

En respuesta a lo anterior, y procurando minimizar este tipo de control, se han propuestos diferentes tecnologías que buscan reforzar o mejorar la privacidad (*Privacy Enhancing Technolo-*

gies) del individuo (visto en un contexto amplio, es decir, pudiéndose considerar como individuo a un conjunto de personas, e incluso a organizaciones completas). Este tipo de tecnologías pueden asistir a los organismos en su cumplimiento de los principios de protección de la privacidad establecido en los derechos humanos [25], dándole a los usuarios mayor poder para controlar su información, pudiendo éstos decidir cómo y cuándo puede ser utilizada por terceras partes. Existen sistemas tales como los navegadores web anónimos y servicios especiales de correo electrónico que le permiten comunicarse sin necesidad de revelar su verdadera identidad. Los sistemas para el manejo de la identidad potencialmente le permiten a los individuos acceder a los servicios y recursos sin tener que proveer información sobre ellos. Esto implica involucrar a una o varias organizaciones sobre las cuales se deba tener cierto grado de "confianza", y que puedan verificar la identidad de los usuarios, y además puedan generar cierto tipo de certificación electrónica que no contenga información sobre la identidad, pero que permita acceder a los recursos y servicios ofrecidos por terceras partes.

Las tecnologías que mejoran o refuerzan la privacidad no son sólo aquellas destinadas a proveer un cierto grado de anonimato, sino que se extienden a la protección y mejora de la privacidad en general del individuo, incluyendo el cumplimiento de sus derechos sobre la protección de sus datos, en este sentido se pueden mencionar como ejemplos de este tipo de tecnología los siguientes:

- Los sistemas de acceso biométrico encriptado, que permite el uso de las huellas dactilares como mecanismo para autenticar la identidad de un individuo sin necesidad de retener su huella dactilar actual.
- Los accesos a los datos personales de los usuarios en línea seguros.
- Programas que permiten a los navegadores detectar automáticamente las políticas de privacidad de los sitios web y las comparan con las preferencias expresadas por los usuarios.
- Sistemas de alertas y avisos que son anexados a la misma información y que previenen su uso en caso del no cumplimiento de las políticas de privacidad.

Tal como se menciona en [17, 60, 26, 11, 39, 20, 71], en la mayoría de investigaciones de reciente publicación se ha adoptado como base la terminología propuesta en [53], y dada su amplia aceptación, se considera de igual forma en este trabajo. A continuación se muestran

las definiciones de los términos más relevantes: anonimato, no-relacionabilidad (unlinkiability), no-observabilidad (unobservability), seudonimato (pseudonymity), manejo de la identidad, y se finaliza con la definición de sistemas anónimos en general, donde se incluyen los diseños y sistemas propuestos más importantes.

Estos términos se basan en la configuración de un sistema general tradicionalmente compuesto por un *emisor*, un *receptor*, quienes utilizan una *red de comunicación* para transmitir un *mensaje*. En la figura 2.1 se muestra el diagrama general de este modelo.

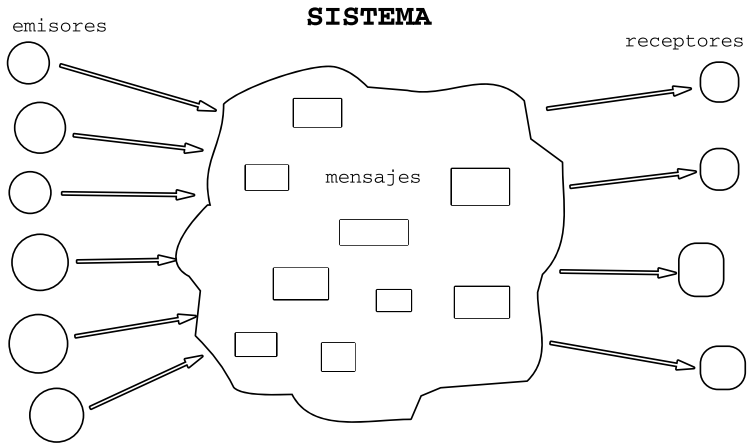

Figura 2.1: Configuración del Sistema General.

Este sistema está delimitado por los componentes antes mencionados, por lo cual los involucrados que se encuentren fuera de esta delimitación, en cada uno de los casos que se describen en este trabajo, se consideran participantes externos.

Cada uno de los casos de estudio presentados serán considerados desde la perspectiva del atacante, quien puede monitorear las comunicaciones, estudiar sus patrones, e incluso puede hacer cambios al manipular el contenido. El atacante puede estar dentro del sistema o puede ser uno de los participantes externos.

En todas las definiciones de los términos relacionados con las tecnologías asociadas a la mejora o refuerzo de la privacidad, se considera un *sujeto* (subject) a una entidad (ente o ser) que tiene la posibilidad de actuar en el sistema, por ejemplo, un ser humano, una persona jurídica, un

ordenador, etc.

2.1.1. Anonimato

Un sujeto es anónimo cuando no puede ser identificado dentro de un conjunto de sujetos, denominado el *conjunto anónimo*. Este conjunto está conformado por todos los posibles sujetos que pueden causar (o estar relacionados con) una acción. No ser identificado significa que ese sujeto no puede ser caracterizado de forma única o particular dentro de ese conjunto. Un sujeto actúa anónimamente cuando, desde el punto de vista del adversario, su acción no puede relacionarse con su identidad, dado que hay un conjunto de sujetos que podrían ser los causantes potenciales de la acción (y el adversario no puede distinguir a su verdadero causante). El anonimato debe permitirle a un sujeto utilizar un recurso o servicio sin revelar su identidad, esto implica que el anonimato por si mismo no procura proteger la identidad de un usuario en un ámbito general, lo que pretende es evitar que otros usuarios o sujetos no puedan determinar la identidad de un usuario cuando éste genera una acción u operación en particular.

Con respecto a las entidades que podrían generar una acción, el conjunto anónimo se conforma por los sujetos que pueden generar una acción en un instante de tiempo específico; desde el punto de vista de las direcciones o ubicaciones de las entidades, el conjunto anónimo está conformado por los sujetos que pueden estar relacionados a una ubicación o dirección. Lo anterior quiere decir que el anonimato se podría clasificar según las entidades involucradas o según la ubicación de las mismas.

De esta forma, para permitir el anonimato de un sujeto siempre tiene que existir un conjunto apropiado de sujetos que posean potencialmente los mismos atributos. Ser los emisores y los receptores de mensajes particulares son ejemplos de estos atributos. Un emisor de un mensaje puede ser anónimo sólo si constituye parte de un conjunto de emisores potenciales (con atributos similares), el cual es su conjunto anónimo, y puede ser un subconjunto de todos los sujetos a nivel global quienes pueden enviar un mensaje en un tiempo específico. Lo mismo aplica para los receptores de mensajes. Este esquema se representa en la figura 2.2. El conjunto anónimo es relativo al tiempo, es decir, puede variar según los cambios que se den en el sistema.

Con lo anterior se especifica que existe un conjunto anónimo para el emisor de un mensaje, existe otro conjunto anónimo para el receptor de ese mensaje, y estos conjuntos puedes ser

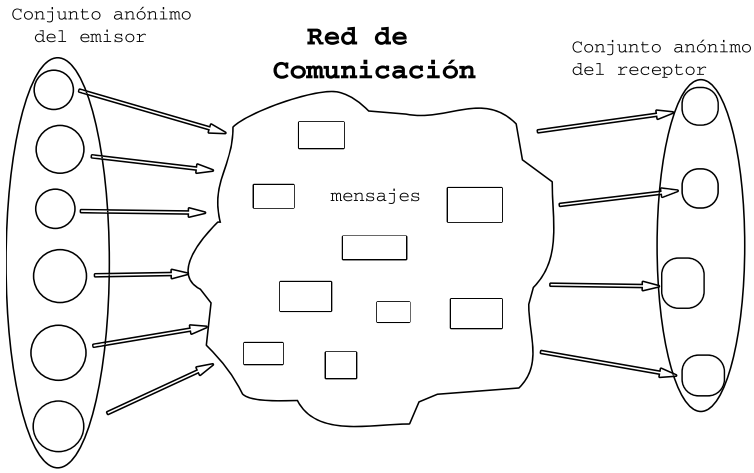

Figura 2.2: Conjuntos Anónimos.

disjuntos, pueden solaparse o pueden ser el mismo conjunto.

Por otro lado el anonimato además de estar relacionado al conjunto anónimo y al tiempo en el que se está ejecutando la acción, también tiene relación al contexto donde se aplica, es decir, un sujeto puede ser anónimo en relación al contexto *envío y recepción de correos electrónicos*, pero puede no serlo en ese mismo instante de tiempo para el contexto *interacción con una base de datos*. Esto se debe a que según el contexto de estudio existirán distintos atributos que caractericen al conjunto anónimo, y por ende al anonimato del sujeto.

Como se mencionó el conjunto anónimo está directamente relacionado con el atacante, esto quiere decir, que el conjunto anónimo se delimita según el grado de conocimiento que posee el atacante. De esta forma, el fin último del anonimato es procurar que el atacante posea la misma información antes y después de su ataque.

Dado que el anonimato es dependiente del contexto, definido por sus atributos, las variaciones del mismo podrían cambiar los niveles de anonimato. Si se pretende diferenciar entre "niveles" de anonimato, es necesario poder cuantificarlo (medirlo) con el fin de poder hacer distinciones entre distintos sistemas anónimos. Las métricas sobre anonimato son explicadas en la sección 2.4.

2.1.2. No-relacionabilidad (unlinkability)

En el estudio de los sistemas anónimos, la no relacionabilidad sólo tiene sentido práctico si previamente se han definido las propiedades del anonimato, seudonimato y no-observabilidad de dichos sistemas, y se han caracterizado las entidades o ítems de interés que se desean relacionar (por parte del atacante).

En [53] se menciona que las entidades o ítems de interés (IDI) son sujetos, mensajes, eventos, acciones, etc. La relacionabilidad se considera a la información que se tiene sobre la relación real que existe entre los IDIs, es decir, que en todos los casos reales existirá cierta relación entre los IDIs y la relacionabilidad es la información respecto a ésta. Por ende, la no relacionabilidad, desde la perspectiva del adversario, significa que la información obtenida después de un ataque es la misma información que se tenía antes del mismo, es decir, los IDIs no están ni más ni menos relacionados comparando el periodo anterior y posterior al ataque. Desde la perspectiva probabilística, la no relacionabilidad significa que la probabilidad de que dos o más ítems de interés estén relacionados (desde la perspectiva del atacante) es la misma antes y después del ataque.

La no relacionabilidad debe implicar que el usuario puede hacer múltiples usos de un recurso o servicio sin que esto conlleve a que el adversario pueda establecer una relación entre estos usos, es decir, requiere que los usuarios y/o los sujetos no tengan la disponibilidad de determinar que cierto usuario fue el causante de cierta acción en el sistema.

Si se considera que el envío y recepción de mensajes son los IDIs, el anonimato puede ser definido como la no relacionabilidad de un IDI con cualquier identificador de un sujeto (en este caso un emisor o un receptor). Específicamente, el anonimato de un IDI es la no relación con cualquier sujeto y el anonimato de un sujeto es su no relación con cualquier IDI. De esta manera se puede considerar *anonimato del emisor* como las propiedades que hacen que un emisor no pueda ser relacionado a cualquier mensaje, y que un mensaje no pueda ser relacionado a un emisor. De igual forma, el *anonimato del receptor* significa que un mensaje no puede ser relacionado a cualquier receptor, y un receptor no puede ser relacionado a cualquier mensaje.

La *relación de anonimato* es la imposibilidad de determinar quién se comunica con quién, es decir, el emisor y el receptor son no relacionables.

La no relacionabilidad es una condición suficiente para el anonimato, pero no es una condición

necesaria. Incluso en algunos casos se puede considerar que aún fallando la propiedad de la no relacionabilidad se puede conseguir un nivel de anonimato alto. Esto si se hace referencia a la definición estricta del anonimato: *no poder ser identificado dentro de un conjunto de sujetos.*

2.1.3. No-observabilidad (unobservability)

En contraste con la definición de anonimato, y la de no relacionabilidad, donde se considera la protección de la relación del IDIs con respecto a los sujetos o a otros IDIs, la no observabilidad considera la protección de los IDIs en sí mismos, ya que se define como el estado de que un IDI sea indistiguible entre un conjunto de IDIs del mismo tipo (el "tipo" lo definen la características de interés en cada caso, por ejemplo, considerar la emisión de mensajes como un IDI, tiene su características propias, que hacen que se pueda diferenciar de otro tipo de IDI). Por ejemplo, se podría decir que al ser no observable, no es posible distinguir entre un mensaje y el ruido aleatorio.

Semejante al caso del anonimato, también se tienen conjuntos de sujetos no observables con respecto a esta propiedad, es decir, el *conjunto de emisores no observables* tiene la propiedad que cada emisor no puede ser distinguido del resto, el *conjunto de receptores no observables* tiene la propiedad de no poderse distinguir, desde el punto de vista del atacante, a un receptor del resto de los receptores de ese conjunto. La *relación de no observabilidad* de esta forma significa que no es posible distinguir entre un emisor y un receptor que intercambian un mensaje, es decir, si se considera un mensaje en particular entonces el *conjunto de la relación de no observabilidad* está compuesto por todos los posibles pares emisor-receptor, entre los cuales no se podría diferenciar o determinar que existe una relación de envío-recepción. En la figura 2.3 se puede observar gráficamente la configuración de dichos conjuntos.

Desde una perspectiva de usuario, la no observabilidad se podría definir como la propiedad de que un usuario pueda utilizar un recurso o servicio sin que otros (terceras partes) tenga la posibilidad de determinar que el recurso o servicio está siendo utilizado.

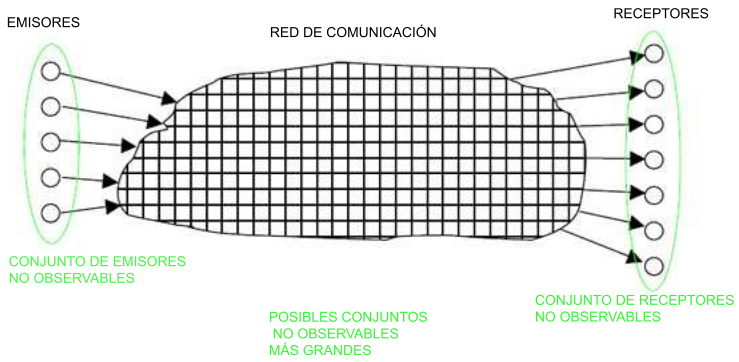

Figura 2.3: Conjuntos no observables.

2.1.4. Seudonimato (Pseudonymity)

Tal como se expresa en [53], los seudónimos son identificadores (nombres u otras cadenas de bits) de sujetos. Para el caso que se está tratando, son los identificadores de los emisores y receptores de mensajes. La palabra seudónimo proviene del griego ("pseudonumon") la cual significa "nombre falso", es decir, un nombre distinto al "nombre real". Sin embargo, como el "nombre real" tiene también un origen arbitrario, el nombre de seudónimo puede ser extendido a todos los identificadores o cualquier otra cadena de bits que identifique a un sujeto. Es posible que en algunos casos convenga agrupar los nombre reales en un conjunto distinto al de los "nombres falsos" (no considerados reales), pero en otros casos un "nombre real" puede ser considerado como un seudónimo que resulta de un selección inicial, para la identificación de un sujeto. Incluso, en gran parte de la literatura excluyen la fracción que le da el nombre de "falso" ("pseudo"), y sólo utilizan, para referirse a nombres de seudónimos, la palabra *nyms*.

Desde un punto de vista más básico (fundamental), un seudónimo puede ser considerado como un tipo de atributo del sujeto, que puede ser controlado por el diseñador de sistemas o por el mismo usuario.

Se pueden generalizar los seudónimos como identificadores de conjuntos de sujetos, pero en todos los casos que se tratan en este trabajo se refiere a un sólo sujeto, el cual representa el *contenedor* de ese seudónimo. En las configuraciones tradicionales se asume que cada seudónimo se refiere a un solo contenedor, invariante sobre el tiempo, el cual no puede ser transferido a otros contenedores. Solamente tipos especiales de seudónimos pueden ser extendidos y transferidos a otros contenedores, por ejemplo, un *seudónimo de grupo* se refiere a un conjunto de contenedores,

y un *seudónimo transferible* puede ser transferido de un contenedor a otro. Utilizando la noción de seudónimo de grupo, se podría inducir el conjunto anónimo: *Utilizando la información provista por el seudónimo solamente, un atacante no podría decidir si una acción fue ejecutada por una persona específica dentro del conjunto.*

En resumen, utilizando un criterio formal, un seudónimo se puede considerar como el estado de utilizar un "nombre falso" como identificación (ID), y *seudonimato* se puede definir como el proceso donde se utilizan los seudónimos como identificadores. En el contexto de este trabajo, el seudonimato tiene como uno de sus principales objetivos el permitirle a un usuario utilizar un recurso o servicio sin tener que revelar su verdadera identidad, y sin quitarle la responsabilidad sobre el uso del mismo.

Es evidente, por definición, que el anonimato y la responsabilidad son extremos opuestos con respectos a la relacionabilidad de los sujetos, sin embargo, a través del uso del seudonimato se procura establecer un punto intermedio para cumplir con ambas partes: Responsabilizar a los usuarios (sujetos) de sus acciones y evitar que se revele su identidad. Incluso utilizando el mismo seudónimo, le puede permitir al sujeto tener cierto nivel de reputación (esto suponiendo que existen los mecanismos adecuados para poder autenticar los mensajes enviados por el contenedor del seudónimo). Además, existen configuraciones de sistemas anónimos que permiten evitar el abuso en su utilización, pudiendo revelarse, en ciertos casos, la identidad del usuario que haya incurrido en cualquier tipo de acciones no permitidas (esto en el supuesto que sólo determinadas autoridades certificadas pueden revelar esta identidad).

El *seudonimato del emisor* se define como el uso de seudónimos por parte del emisor, y el *seudonimato del receptor* se define como el uso de seudónimos del receptor. En la figura 2.4 se representa esta noción de seudonimato (extraida de [53]).

Cuando se utilizan "nombres falsos" como etiquetas que identifican a un sujeto, es conveniente considerar cómo se trata o maneja la responsabilidad y la autorización. Es decir, los seudónimos digitales pueden ser utilizados para autenticar los mensajes, ya que un seudónimo digital es una cadena de bits que dentro de este contexto es un único ID del sujeto, el cual puede ser utilizado para autenticar el IDI del contendor relativo a su seudónimo, esto implica que la responsabilidad puede ser manejada a través del uso de seudónimos. Sin embargo, puede no ser suficiente para alcanzar este tipo de gestión de la responsabilidad, por tal motivo en algunos casos prácticos, es necesario acompañar al seudónimo con otro tipo de pruebas de validez (tales como firmas

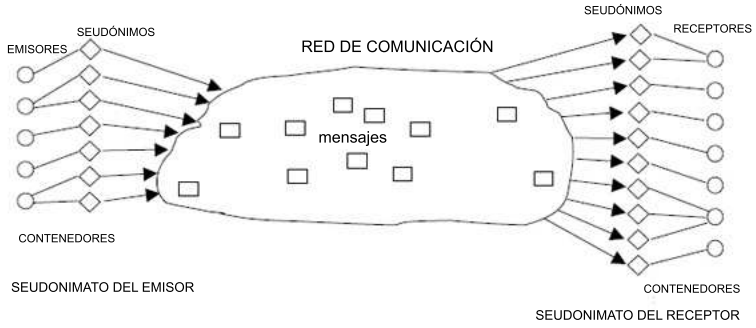

Figura 2.4: Pseudonimato.

digitales), y/o utilizar terceras partes (autoridades de certificación digital) que le den validez a los seudónimos.

Se pueden considerar las siguientes definiciones del seudonimato con respecto al anonimato:

- *Seudónimo público*: La relación entre un seudónimo y su contenedor puede ser de conocimiento público, por ejemplo podría estar en una lista o directorio público.
- *Seudónimo no público*: La relación inicial del seudónimo y su contenedor puede ser conocida por ciertas partes (entidades de control o administración centralizada o descentralizada), pero no son del conocimiento público.
- *Seudónimo no relacionado inicialmente*: La relación inicial entre un seudónimo y su contenedor es, al menos en el inicio, desconocida para cualquiera, con la posible excepción del contenedor mismo.

Los *seudónimos públicos* y los *seudónimos no relacionados inicialmente* son los extremos que pudiesen ser análogos con la responsabilidad y el anonimato mencionados anteriormente, es decir, el anonimato es más fuerte mientras menos se relacione al contenedor con su seudónimo, y decrece en caso contrario. Esto quiere decir, que el nivel de anonimato dependerá de la relación relativa que el atacante le de al seudónimo con respecto a su contenedor.

También se pueden considerar varios tipos de seudónimos con respecto a la relacionabilidad:

- Seudónimo personal: Es un subtítulo para el nombre del contenedor el cual está relacionado

con su identidad civil. Este puede ser utilizado en todos los contextos, por ejemplo, el número de su tarjeta de identificación (número de pasaporte).

- Seudónimo por rol: Está limitado a roles específicos, es decir, su asignación está relacionada a los roles que desempeña el contenedor. Por ejemplo, los seudónimos que comúnmente se utilizan bajo el rol de "usuario de Internet" comúnmente son denominados "nombre de usuarios".
- Seudónimo por relación: Diferentes seudónimos pueden ser utilizados para establecer las asociaciones en pares dependiendo de los pares conformados para las comunicaciones.
- Seudónimo por relación y por rol: Son seudónimos asignados según el rol que desempeña cada par que se comunica. Es decir, que un mismo contenedor puede utilizar varios seudónimos, dependiendo del rol que desempeñe tanto él como su homólogo en la comunicación.
- Seudónimo por transacción: Un mismo contenedor podría tener un seudónimo por cada transacción que realice.

El nivel de anonimato en relación al tipo de estrategia de utilización de seudónimos se representa en la figura 2.5.

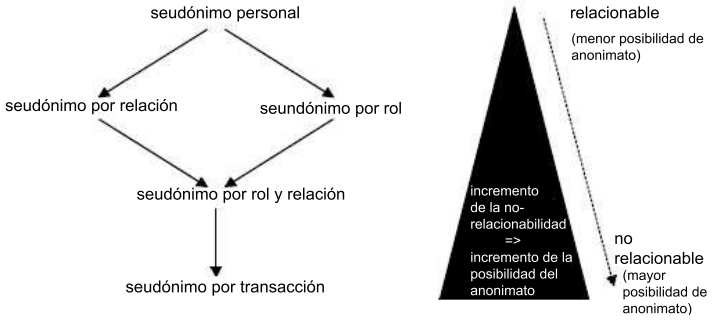

Figura 2.5: Anonimato respecto al Seudonimato.

En general, el anonimato del seudónimo por rol, y del seudónimo por relación es más fuerte que el anonimato debido al uso de seudónimos personales. La fortaleza del anonimato se incrementa con la aplicación de seudónimos por rol y relación. El último nivel de fortaleza se consigue utilizando los seudónimos por transacción dado que no existe información sobre la relación entre cada una de las transacciones. En ocasiones, se utiliza el nombre de *seudónimo de un solo uso* para denotar los seudónimos por transacción. La fortaleza del anonimato antes mencionada se

refiere a que desde un punto de partida se define que el mayor nivel (fortaleza) de anonimato se logra cuando no se proporciona información sobre la identidad en absoluto. En otras palabras, el anonimato es más fuerte mientras menos datos personales del contenedor puedan ser relacionados al seudónimo, y mientras los seudónimos sean utilizados con menor frecuencia (estos es cuando un mismo contenedor utiliza un mayor número de seudónimos).

Un seudónimo digital se puede conformar con el uso de algún mecanismo de clave pública utilizado para probar las firmas digitales, donde los contenedores del seudónimo pueden probar que son los verdaderos "dueños" del seudónimo solamente construyendo una firma digital creada con su clave privada. En la mayoría de los casos los seudónimos digitales son las mismas claves públicas generadas por los propios usuarios.

Un certificado de clave pública contiene una firma digital de una determinada autoridad certificadora que provee algún tipo de seguridad cuando un mismo sujeto utiliza la clave pública para otro de sus seudónimos. En el caso de que el seudónimo provenga del nombre real de un sujeto, se le llama certificado de identidad. Un certificado de atributo es un certificado digital que contiene más cantidad de información (atributos) y claramente se refiere a un certificado específico de clave pública. Independientemente de los certificados, los atributos también pueden ser utilizados como identificadores de conjuntos de sujetos, no sólo de sujetos particulares.

Además pueden existir algunas propiedades de los seudónimos que son utilizadas por aplicaciones específicas, por ejemplo la dirección de un seudónimo puede funcionar como la dirección de la comunicación, o pueden ser utilizadas en la participación de terceras partes (en los casos de abuso antes mencionado).

2.1.5. Manejo de la identidad

Para acotar mejor el contexto de trabajo se utiliza la siguiente configuración general:

- No es realista asumir que un atacante no puede obtener información sobre el emisor o el receptor del mensaje desde el contenido del mismo y/o del contexto del envío y recepción (hora, información local, etc.) del mensaje. Se debe considerar que el atacante puede utilizar estas propiedades para relacionar los mensajes y consecuentemente, los seudónimos utilizados con éstos.

- Además, se debe considerar que no son sólo seres humanos, personas jurídicas, o simples ordenadores enviando mensajes y utilizando seudónimos a su discreción, sino que existen aplicaciones involucradas, fabricadas por terceros, que influyen en gran medida en el envío y la recepción de mensajes, y pueden determinar y condicionar el uso de los seudónimos. En otras palabras, no se puede obviar la dependencia que se tiene de las aplicaciones ya creadas.

La identidad puede ser explicada como una percepción única o exclusiva de vida, con su integración a un grupo social, y con continuidad, la cual está acotada y formada por una sociedad. En este sentido se está considerando la identidad de seres humanos quienes son los principales interesados en la preservación de la privacidad. Sin embargo, desde un punto de vista estructural, la identidad puede estar ligada a cualquier sujeto, ya sea un ser humano, una persona jurídica, y un ordenador. Esto le da mayor generalidad a la terminología, pero para que tenga utilidad y motivación práctica, en la totalidad de los casos se considera asociada a seres humanos.

Con esta definición de identidad se logra distinguir entre dos instancias de los sujetos, la primera que les permite definirse asimismo, en un contexto ligado a su libertad y a su iniciativa (definición interior o personal), y la segunda que los define en un contexto social con sus respectivos atributos, y que se mantiene para darles la posibilidad del acceso a la comunicación y que los ata de cierta manera a un control y un grado de consistencia con respecto al resto.

Análogo al conjunto anónimo, también se puede trabajar con un *conjunto identificable* para definir la "identificabilidad" y la "identidad". Este conjunto está conformado por todos los posibles sujetos.

De esta forma, la *identificabilidad* es el estado de ser identificable dentro de una conjunto de sujetos, que es el *conjunto identificable*.

La identificabilidad es inversa al anonimato, es decir, mientras sea mayor el nivel o grado de identificabilidad menor será el grado o nivel de anonimato, y viceversa.

Según lo anterior, una identidad es un conjunto de atributos pertenecientes a un individuo que permiten diferenciarlo del resto de individuos que forman parte de un conjunto determinado. Por esta razón no existe una identidad única y universal, sino que pueden existir varias para un mismo individuo, según el conjunto y contexto al que se haga referencia. Incluso los valores de los atributos y los atributos mismos pueden cambiar en el tiempo.

Términos relacionados a la identidad

- Rol: Desde el punto de vista de la sociología, un rol o rol social constituye un conjunto de acciones conectadas o relacionadas, conceptualizadas por los actores en una situación social. Es frecuente definirlo como un comportamiento esperado en un contexto individual social dado.

- Identidad Parcial: La identidad de cada persona está compuesta por muchas identidades parciales, las cuales representan a la persona en un contexto o rol específico. De esta forma, una identidad parcial es un subconjunto de atributos del conjunto que compone la identidad completa, formada por la unión de todos los atributos de esta persona. Desde un punto de vista técnico, estos atributos constituyen datos. Como se mencionó anteriormente estos atributos y los valores de la identidad parcial pueden variar en el tiempo. Un seudónimo puede considerarse como una identidad parcial. A pesar de que una identidad parcial no permite caracterizar a un individuo de forma única dentro de un conjunto específico, si puede, según la cantidad de atributos contenidos en el subconjunto, hacer posible tener varios contextos de aplicación del anonimato, tal como se observa en la figura 2.6.

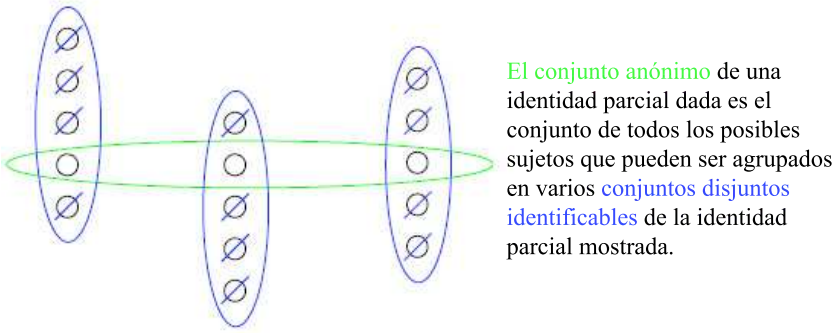

Figura 2.6: Conjunto Anónimo vs. identidades parciales.

- Identidad Digital: La identidad digital denota la atribución de propiedades a una persona, los cuales son, desde el punto de vista técnico, *operacionalmente accesible de forma inmediata*, por su característica digital. El identificador de una identidad parcial digital puede ser una dirección de correo electrónico en un grupo de noticias o en una lista de correos.

- Identidad Virtual: Algunas veces se utiliza como sinónimo de identidad digital, pero dada su connotación de no-real, aparente y no-existente, este término es utilizado principalmente

en ambientes multi-usuarios, o multi-jugadores masivos, es decir, en entornos de juegos virtuales.

Términos relacionados con el manejo de la identidad

- Manejo de la identidad: Significa el manejo de varias identidades parciales (usualmente denotadas por seudónimos) de un individuo. El establecimiento de la reputación es posible cuando un individuo re-utiliza las identidades parciales. Un pre-requisito para la selección de una identidad parcial es el de conocer el contexto en el que la persona está actuando.

- Manejo de la identidad para mejorar la privacidad: Dadas las restricciones de una aplicación, el manejo de la identidad es llamado *mejora de la privacidad* si al seleccionar los seudónimos y todas sus autorizaciones cuidadosamente, no se provee a un atacante más relacionabilidad entre las identidades parciales en comparación a darle los datos omitiendo la información de los seudónimos. En otras palabras, el manejo de la identidad es llamado *mejora de la privacidad (Privacy-enhancing)* si ésta no provee más relacionabilidad, esencialmente, entre las identidades parciales.

- Manejo de la identidad para mejorar la privacidad en el diseño de aplicaciones: Una aplicación está diseñada para permitir el manejo de la identidad para mejorar la privacidad si ningún patrón del envío/recepción de mensajes ni los atributos dados a las entidades implica más relacionabilidad que lo estrictamente necesario para los propósitos de la aplicación.

- Sistema de manejo de la identidad: La tecnología basada en el manejo de la identidad en su esencia más amplia se refiere a la administración y diseño de los atributos de las identidades. Se puede distinguir entre un sistema de manejo de identidad y una aplicación para el manejo de la identidad: La primera puede ser entendida como una infraestructura, y la segunda como un conjunto de componentes coordinados entre sí. Las aplicaciones para el manejo de la identidad son herramientas para manejar individualmente sus comunicaciones relevantes, las cuales pueden ser configuradas y operadas en el lado de los usuarios o en el lado de los servidores. El manejo de la identidad, soportado técnicamente, tiene que autorizar a los usuarios para reconocer diferentes tipos de comunicaciones o situaciones sociales, y acceder a ellas con respecto a su relevancia, funcionalidad y al nivel de riesgo de la privacidad y seguridad en función de hacer y asumir roles de forma adecuada. En general, las aplicaciones para el manejo de la identidad, específicamente en cuanto al manejo

de las identidades parciales, representan los diferentes seudónimos con su respectivos datos de acuerdo a los diferentes roles que el usuario ha asumido y de acuerdo a los diferentes patrones de comunicación.

- Sistema para el manejo de la identidad en el mejoramiento de la privacidad: En este caso se hace explícito el flujo de los datos personales, donde se le permite al usuario tener un mayor grado de control. La guía principal es la de "reconocer y escoger" su propia identidad, y se procura minimizar la cantidad de los datos utilizados. Esto significa que el usuario controla la relacionabilidad de sus datos personales. De acuerdo a una situación y contexto específico, tal sistema le da soporte al usuario en la selección de seudónimos que representen a sus identidades parciales.

2.2. Técnicas de Anonimato

Tal como se mencionó en los apartados anteriores el anonimato de un sujeto es el estado de no ser identificable dentro de un conjunto de sujetos, denominado el conjunto anónimo. También se ha mencionado, tal como se menciona en [1], que el emisor de un mensaje puede ser anónimo sólo dentro de un conjunto de potenciales emisores, lo que correspondería al conjunto anónimo del emisor, el cual a su vez puede ser un subconjunto de todos los sujetos a nivel mundial quienes podrían enviar mensajes en determinados instantes de tiempo. Este tipo de anonimato es llamado *anonimato del emisor*. Lo mimo ocurre para el receptor, quien puede ser anónimo sólo dentro de un conjunto de receptores posibles, llamado el conjunto anónimo del receptor, y a este tipo de anonimato es llamado *anonimato del receptor*. Además hay un tercer tipo de anonimato, el de relación, el cual consiste en tener la propiedad de no poder relacionar quién se comunica con quién. La no relacionabilidad significa que dentro del sistema las distintas entidades, aquí denominadas *ítems de interés o IDI* (mensajes, emisores, receptores, etc.) no están ni más ni menos relacionadas con respecto a la información que se tenía antes de que el adversario ejecute un ataque (*información a priori*). En otras palabras, el anonimato del emisor/receptor puede ser definido como las propiedades de que un mensaje particular no sea relacionado a cualquier emisor/receptor, y que cualquier mensaje no sea relacionado a un emisor/receptor en particular, entonces el *anonimato de relación* es la propiedad de no poder relacionar o determinar quién se comunica con quién.

El anonimato se fortalece mientras más grande sea su conjunto anónimo, y mientras más uniforme sea la distribución de la ejecución de las acciones por parte de los sujetos dentro del conjunto, es decir, el nivel de anonimato no sólo depende del tamaño del conjunto sino también de la probabilidad de que un sujeto en particular pueda generar cierta acción.

De esta forma se puede definir el entorno de acción que acota las técnicas de anonimato para las comunicaciones: Colectar un conjunto apropiado de usuarios para que un usuario en particular pueda ser anónimo cuando se comunica con los demás.

Los sujetos no pueden tener el mismo nivel de anonimato contra todos los tipos de ataques posibles generados por participantes internos o externos. El conjunto de los posibles sujetos y la probabilidad de que ellos puedan causar una acción puede variar dependiendo del conocimiento del atacante. Se asume que desde el punto de vista del atacante, el nivel de anonimato sólo puede decrementar, es decir, se asume que el atacante no olvida la información que tiene y que ha logrado recolectar durante su observación e influencia sobre la comunicación en el sistema.

Para definir las diferentes técnicas de anonimato se utilizan los siguientes criterios:

- *Objetivo de la protección*: Define cuál tipo de anonimato puede ser provisto (del emisor, del receptor, o de la relación).
- *Nivel de seguridad*: Se debe definir cuál es el nivel de seguridad alcanzado por el objetivo de la protección (la seguridad desde la perspectiva de la teoría de la información o incondicional y la seguridad criptográfica/computacional con los supuestos asociados a mecanismos como los de *clave pública*).
- Modelo de atacante: Contra qué tipo de atacantes protege el sistema (externos, participantes, proveedores de servicios).
- Modelo de confianza: En quién confía el usuario: en los proveedores de servicios, en los participantes externos, en otros usuarios, etc.

2.2.1. Proxies simples

Este corresponde a uno de los conceptos más populares en la utilización de las comunicaciones anónimas. La idea principal detrás de esta técnica es que las solicitudes que se pretendan hacer

a un servicio web, no se hagan directamente hacia el servidor, sino que un servidor intermedio (proxy) haga la solicitud desde un punto intermedio. Es decir, las solicitudes no son enviadas directamente desde el usuario hacia el servidor web, sino que el cliente se conecta a otro servidor, llamado proxy, y éste genera la solicitud web al servidor por el usuario (tal como se muestra en la figura 2.7). De esta manera el servidor web sólo puede "ver" la dirección IP del servidor proxy y no la del usuario. Además, el servidor proxy puede filtrar la información de la solicitud que podría ser utilizada para identificar al usuario, la cual incluye la información obtenida por "cookies", del sistema operativo, del navegador web, etc. Por ahora hay dos posibilidades de conectarse a un proxy: vía un sitio web o con un proxy local, y ambas pueden ser combinadas convenientemente.

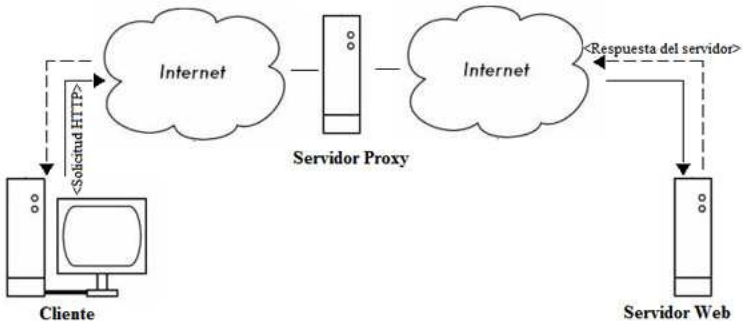

Figura 2.7: Proxies.

Sitio Web

Un sitio web proxy permite el uso de los servicios anónimos sin tener que instalar programas adicionales. En este sitio web el usuario normalmente llena un formulario con la dirección del sito adonde quiere conectarse. El proxy le envía una solicitud al servidor web con la dirección que indicó el usuario, luego, el servidor le envía su respuesta al proxy y la información obtenida le es enviada al cliente por parte del proxy. Además el proxy busca cualquier enlace adicional que se necesite. Si existen uno o varios enlaces adicionales, el servidor proxy automáticamente los transforma para que se utilicen también como la comunicación inicial, es decir, que la conexión con los enlaces adicionales también se hace a través del proxy, permitiéndole así al cliente permanecer anónimo aún cuando se monitoréan estos enlaces adicionales.

Proxies locales

Este método requiere la instalación de un programa adicional en el ordenador del cliente, quién se debe registrar y debe configurar su navegador para poder conectarse a los sitios web.

Este programa para proxies locales tiene un lista de los llamados *proxies abiertos* los cuales son proxies que han sido intencionalmente o por error dejados abiertos para el uso público. Esto significa que cualquier usuario puede utilizar estos proxies para esconder su identidad.

El programa de proxy local está diseñado para que busque y descubra automáticamente los nuevos proxies abiertos. Después, de forma aleatoria selecciona uno de ellos y lo utiliza de la misma forma que los sitios web proxy. Incluso algunos están configurados para hacer cambios cada cierto intervalo de tiempo.

Cadenas de proxies

Tal como su título lo expone, se refiere al uso de varios proxies conectados de forma secuencial o consecutiva conformando así cadenas. La combinación de distintos proxies, desde el punto de vista técnico, puede ser útil al considerar que cada tipo de proxy tiene ciertas ventajas y desventajas. Al combinar proxies locales en una cadena se puede obtener un mejor nivel de filtrado para las solicitudes web (http), ya que siendo locales, la velocidad de navegación no se verá afectada en gran medida. Por otro lado, si se utilizan varios proxies externos, la confianza en cada uno de ellos pasa a jugar un papel importante.

El objetivo de protección en este caso es el anonimato del emisor contra el receptor, y el anonimato de relación contra todo el resto (el resto lo constituyen los participantes externos). El modelo de ataque es el siguiente: protección contra el receptor de la solicitud, no hay protección contra un proveedor de proxy corrupto o varios proveedores en colusión, no hay protección contra atacantes externos quienes podrían relacionar los mensajes entrantes y salientes de un usuario a través de ataque del tipo *análisis de tiempo*. En cuanto al modelo de confianza, el usuario debe confiar en el proxy porque éste puede registrar toda la información transferida y observar las actividades del usuario.

2.2.2. Crowds (multitudes)

Crowds se basa en el hecho de que las actividades de cada usuario en particular pueden ocultarse dentro de las actividades de muchos otros usuarios. De esta forma, el sistema consiste en una colección dinámica de usuarios denominada "crowd" o multitud. Fue diseñada para proveer alto rendimiento y anonimato del emisor.

Tal como se describe en [63], una solicitud hacia un sitio web primero pasa a través de un número aleatorio de participantes del sistema antes de ser enviado al servidor web. Esto quiere decir, que cuando un miembro del sistema "crowds" recibe un mensaje decide aleatoriamente si éste se envía directamente al servidor destino o a otro miembro del sistema. Si el mensaje es enviado a otro miembro del sistema, ésta hace lo mismo hasta que el mensaje se envía a su destino final, tal como se muestra en la figura 2.8.

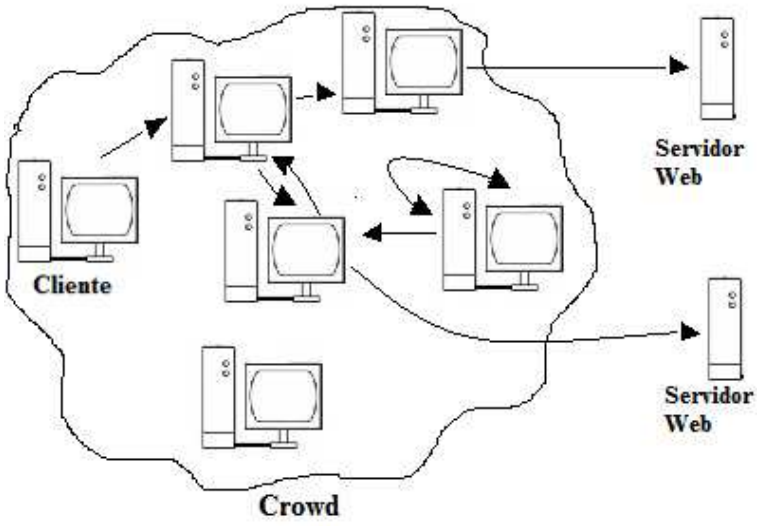

Figura 2.8: Crowds.

De esta forma el mecanismo permite que ni el servidor ni algún otro miembro del sistema pueda saber quién es el iniciador de la solicitud o mensaje. Esto es lo que se denomina *plausible deniability*.

Si el usuario quiere formar parte de la red, debe instalar un programa adicional en su ordenador local, que comúnmente se le denomina "jondo". Además, se debe registrar en un servidor

central, comúnmente denominado “blender”. Como parte del procedimiento de registro, el usuario debe crear un nombre de usuario (seudónimo) y una clave para su autenticación personal. El “jondo” es una especie de proxy local, así que se debe configurar al navegador para poder utilizarlo y navegar.

Cuando el sistema se inicia, un “jondo” primero contacta al “blender” para solicitarle acceso a la red Crodws. Para acceder, el servidor le envía una lista de los miembros de la red (la multitud de “jondos”) y sus claves de cifrado simétrico. Además, el servidor informa al resto de los miembros de la red sobre el nuevo integrante. Estas claves son necesarias porque toda la información (solicitudes y respuestas) son cifradas de un miembro a otro de la red. Cualquier solicitud web proveniente de un navegador se envía a un “jondo” local, el cual envía a su vez la solicitud aleatoriamente a un miembro de la red (puede ser otro miembro o asimismo).

Cuando un mensaje se envía o se recibe desde un miembro de la red, una identificación del camino (ID) se almacena en una base de datos local. Este ID hace posible el envío, en dirección opuesta, de las respuestas desde el servidor hacia el cliente solicitante. Cada “jondo” guarda los pares de ID del camino correspondientes a la entrada y salida de una solicitud. Si un “jondo” recibe un mensaje de otro “jondo”, éste verifica si el ID del camino recibido ya está almacenado en la tabla, si es así, el “jondo” reenvía el mensaje al próximo “jondo”, dependiendo del segundo ID del camino en la tabla. Si el ID del camino no está en la tabla, se selecciona un nuevo destino (“jondo” o el servidor final), se envía el mensaje , y se almacena un nuevo par en la tabla.

Esta técnica protege al anonimato del emisor contra los receptores (se suponen receptores corruptos), y al anonimato de relación contra todo el resto. El modelo no provee protección contra atacantes externos, quienes pueden monitorear toda la red. No hay protección para un “jondo” cuyos envíos y recepciones están siendo observadas. Dado que el sistema no utiliza criptografía asimétrica, no se basa en las suposiciones criptográficas de esta técnica, y de esta forma se puede decir, que puede ser alcanzado un nivel incondicional de seguridad (un nivel condicional de seguridad lo da el uso de técnicas asimétricas de cifrado, sobre las cuales existen ciertos supuestos para que se implementen). El modelo de confianza es el siguiente: Se debe confiar en una instancia central llamada “blender”, los reenvíos de los “jondos” deben confiar en la disponibilidad de la red.

2.2.3. Difusión o Broadcast

Es un técnica ya utilizada en las redes de comunicación para la distribución de información, por ejemplo, para la recepción de las señales de radio o televisión. Todos los participantes de la red reciben toda la información enviada y seleccionan localmente cuál información posee relevancia para ellos. Esto le imposibilita a un atacante pasivo determinar si la información va hacia un receptor en particular.

Si un mensaje está dirigido a un participante específico de la red de distribución, se debe utilizar una dirección implícita. Esto significa que no hay un enlace o relación entre esta dirección implícita y el receptor físico sino sólo para el receptor mismo. La dirección sólo puede ser interpretada por el receptor o recipiente, y de esta forma el emisor no puede obtener información concreta sobre este receptor. Una dirección implícita como, por ejemplo, un número aleatorio, tiene que ser enviada con el correspondiente mensaje y cada estación que recibe el mensaje compara ésta con su propia dirección implícita y verifica si es el verdadero receptor del mensaje. Si la dirección puede ser vista públicamente, se denomina *dirección implícita visible*.

Para evitar que diferentes mensajes destinados al mismo receptor puedan ser relacionados o enlazados con ésta, se deberían utilizar direcciones visibles distintas para cada mensaje. El mensaje también debería estar encriptado, para evitar que otros receptores en el broadcast puedan leer su contenido.

Por el contrario, si la dirección implícita está cifrada, entonces se le denomina *dirección implícita invisible*, pero esto obliga a cada estación a descifrar todos los mensajes que recibe para poder verificar si le corresponde alguno de ellos.

En una red conmutada, donde cada estación sólo recibe lo que le ha sido enviado por otro participante, o sólo recibe lo que ha solicitado, puede ser utilizado un sistema de comunicación múltiple o multicast. Este tipo de difusión parcial significa que cada uno de los participantes no recibe todos los mensajes sino un subconjunto de los mismos, reduciendo así el ancho de banda necesitado, sin embargo, también decrece en nivel de anonimato.

Los objetivos de protección de esta técnica son el anonimato del receptor utilizando las direcciones implícitas, la no observabilidad del receptor, contra los adversarios externos, si se utiliza un tipo de tráfico de relleno, denominado “tráfico dummy”. La no relacionabilidad de los diferentes mensajes pertenecientes al mismo receptor se consigue al utilizar distintas direcciones implíci-

tas para cada mensaje. El nivel de seguridad alcanzado es incondicional si no se utiliza cifrado asimétrico. En cuanto al modelo del atacante, en relación al anonimato y la no relacionabilidad, existe protección contra los atacantes pasivos (observadores) externos e internos. Con respecto a la no observabilidad, existe protección contra los atacantes externos. En lo que respecta al modelo de confianza, si se utiliza tráfico dummy y direcciones implícitas, no se necesita confiar en otro participante ni en ningún proveedor de servicios.

2.2.4. Red de anillo

En esta técnica, las estaciones están circularmente cableadas, tal como se observa en la figura 2.9a, lo que hace que sea válido sólo para configuraciones de redes locales o regionales.

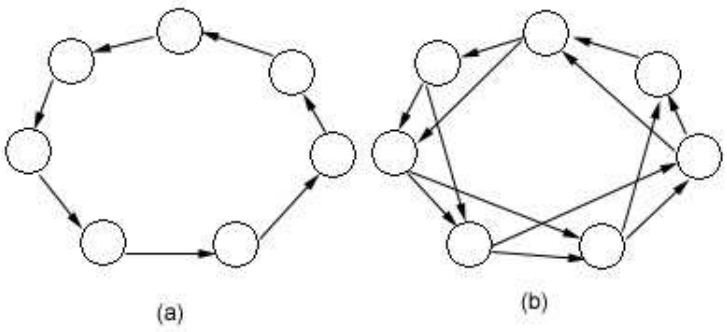

Figura 2.9: (a) Topología de anillo (b) Anillo Trenzado.

Si una estación necesita enviar un mensaje, éste se envía en sucesión en al menos una vez a cada una de las estaciones del anillo.

A través del uso de la regeneración de la señal digital en cada estación participante, cada mensaje es independiente del mensaje original (con respecto a las características analógicas). Cada estación regenera el mensaje de tal forma que ésta parece ser la estación inicial. Este método provee anonimato del emisor con respecto a adversarios quienes observan o controlan estaciones o conexiones de la red de anillos, siempre que no estén directamente antes ni después del emisor. El receptor también se convierte en anónimo y no observable debido al reenvío del mensaje sobre el anillo completo. Una precondición adicional necesaria para garantizar el anonimato del emisor, es que los permisos de envío estén apropiadamente configurados y garantizados.

Si dos estaciones de la red de anillo tratan de observar una estación ubicada entre ellos sin una mutua colaboración no podrán observar algo significativo porque los mensajes de salida están cifrados. De esta forma un atacante debe rodear a una estación y comparar los mensajes de entrada y de salida. Si el atacante no puede hacer esto, sólo podría inferir si alguna estación envía un mensaje, dentro de un grupo de estaciones directamente conectadas, pero no podría establecer exactamente cuál estación lo hace.

Con el fin de asegurar que los mensajes sean recibidos por las estaciones adecuadas, es suficiente si el emisor obtiene el mensaje de regreso sin modificaciones después de dar una vuelta sobre el anillo.

Debido a la conexión serializada de las estaciones todas la conexiones y las estaciones tendrán que trabajar apropiadamente para que la comunicación entre cualquier par de estaciones sea posible. Las estaciones con defectos deberán ser removidas de la red de anillo.

Una posible solución para evitar las interferencias es el anillo trenzado. Tal como se presenta en la figura 2.9b, dos redes de anillos se interconectan mutuamente. El primer anillo conecta a las estaciones vecinas y el segundo anillo conecta dos estaciones separadas por otra intermedia. Esto no sólo duplica la capacidad de transmisión sino que compensa un mal funcionamiento o ruptura en una estación o conexión.

Una topología en anillo, de esta forma, utilizando la regeneración de la señal digital y una técnica para múltiples accesos anónimos proveen anonimato del emisor y del receptor contra un atacante que controla algunas estaciones.

El objetivo de protección es el anonimato del emisor, y el anonimato del receptor a través del envío del mensaje sobre el anillo completo. El nivel de seguridad que puede ser alcanzado será sólo a nivel computacional si el cifrado utilizado para los mensajes saliente esta basado en suposiciones criptográficas (cifrado asimétrico). Si no es así puede ser alcanzado un nivel de seguridad incondicional. En cuanto al modelo del adversario, esta técnica ofrece protección contra un atacante que controla algunas estaciones a excepción de las que se ubican antes y después del emisor. El modelo de confianza para este caso se refiere a que las estaciones vecinas de un usuario no deben colaborar contra éste.

2.2.5. Buses

En [5] propusieron un mecanismo para las comunicaciones anónimas basados en lo que llamaron "buses". En este enfoque cada usuario se modela como una estación de autobús, mientras que los mensajes de los usuarios son transferidos como pasajeros de los propios autobuses. La idea de base de esta premisa está en el de considerar que es muy difícil trazar (por parte de un observador) la trayectoria de una persona que se traslada en autobús en un centro urbano, y es aún más difícil si utiliza diferentes autobuses para completar su ruta. Si un usuario quiere enviarle un mensaje a otro, primero tiene que esperar a que el autobús llegue a su estación para poder hacer el envío (esta idea es semejante al mecanismo de "tokens" de la redes "token-ring"), luego procede a colocar el mensaje en los asientos del autobús.

En esta misma propuesta se plantean tres tipos de sistemas. El primero está basado en una topología de anillo y utiliza un autobús solamente. Como se muestra en la figura 2.10, el autobús siempre se mueve en una misma dirección. Además, el autobús tiene un asiento por cada par emisor/receptor que posee el sistema. Si, por ejemplo, una estación A quiere enviarle un mensaje a una estación B, A cifra el mensaje con la clave pública de B y lo coloca en el asiento AB del autobús. Para asegurarse que un atacante no pueda percatarse de que una estación quiere enviar un mensaje o no, cada una de ellas tiene que enviar mensajes a todas las demás estaciones. De este forma el atacante no podrá decidir si hay una comunicación real entre las estaciones. Para recibir los mensajes, una estación tiene que descifrar y verificar todos los mensajes en sus asientos, dado que otras estaciones también han podido enviarle mensajes.

Desde el punto de vista del rendimiento, si sólo se considera el número de mensajes se podría considerar que ésta es una comunicación óptima, pero los mensajes requieren de una gran cantidad de tiempo para ser transferidos desde el emisor hasta el receptor, en lo cual se debe considerar que la longitud de cada envío tiene un crecimiento cuadrático con respecto al número de estaciones en el sistema, además el mensaje tiene que pasar por cada una de las estaciones del anillo.

Una posible modificación es la de utilizar asientos variables, y no asientos fijos como en el caso anterior. En este caso, el emisor cifra los mensajes en forma de "cebolla" o por capas utilizando las claves públicas del resto de las estaciones, por las cuales el autobús pasará en el camino hacia el receptor. Se dice que es en forma de "cebolla" porque se cifra por capas, la primera capa consiste en el cifrado del mensaje con la clave pública del receptor, la segunda capa se obtiene cifrando el resultado anterior con la clave pública de la penúltima estación, y así sucesivamente

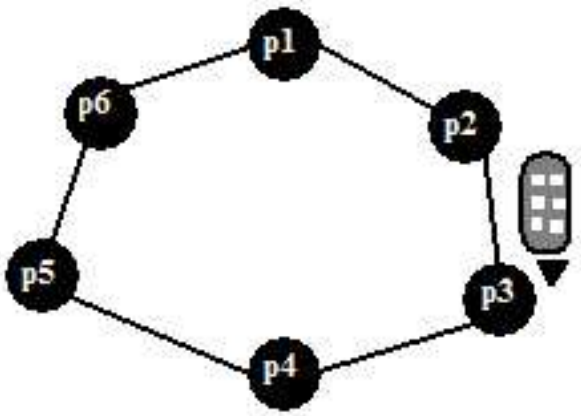

Figura 2.10: Red de Anillo con un solo autobús.

hasta cifrar la última capa con la clave pública de la primera estación. Al ser enviado el mensaje, cada estación descifra la capa que le corresponde y realiza el reenvío a la siguiente estación si se percata que el mensaje no le corresponde. Si al descifrar el mensaje puede ser leído, quiere decir que esa estación es el receptor del mensaje. Este esquema puede permitir incrementar la posibilidad de colusión, ya que no se tienen asientos confirmados, por ende se deben calcular con precisión el número de asientos provistos en el autobús.

El segundo tipo de sistema de buses que introdujeron en [5] utiliza dos autobuses, los cuales se mueven en direcciones opuestas en una ruta circular, que hace la conexión entre dos estaciones. Esto permite poseer mejores tiempos para los envíos, pero desmejora el proceso de comunicación. Para evitar esto, y conseguir optimizar la comunicación en cuanto a los tiempos de envío, se propuso un tercer tipo de esquema, el cual se base en el concepto de conglomerados o "clusters". Tal como se muestra en la figura 2.11, los nodos o estaciones están integrados en grupos ("clusters") de tamaño parecido. Cada "cluster" tiene su propio autobús para transportar los mensajes.

Los autobuses utilizan una técnica de rutas circulares (anillo), a un nivel de capas altas de comunicación (modelo OSI), considerando cualquier topología de red (a nivel de capa de red o de enlace). A pesar de que la idea apunta a tener un comportamiento óptimo, desde el punto de vista de la comunicación y del tiempo de envío, se han hecho implementaciones prácticas de la misma, que muestran que es sólo utilizable para implementaciones en redes pequeñas, y además

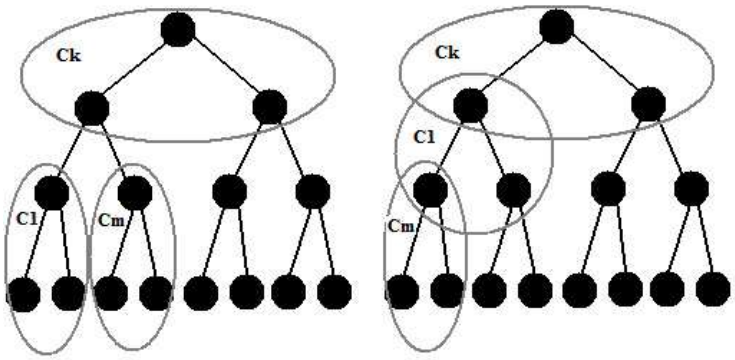

Figura 2.11: Red de Anillo dividida en “clusters”.

necesita una gran cantidad de tráfico de relleno (“dummy”) para lograr esconder las interacciones entre los usuarios.

Los objetivos de protección son el de proteger el anonimato del emisor, del receptor y el de relación. Sólo puede ser alcanzado un nivel de seguridad computacional, porque el cifrado asimétrico se utiliza para el envío de mensajes, el cual está basado en suposiciones criptográficas.

En cuanto al modelo del atacante pueden ser descrito dos tipos: un atacante que realmente lee los mensajes de la red y controla algunas de las estaciones, y los atacantes quienes pueden manipular, crear, o borrar mensajes. Este sistema no puede evitar un ataque de negación de servicio (DoS).

2.2.6. Red-DC

El término Red-DC significa red de criptógrafos de comedor (*D*ining *C*ryptographers network), acuñado por su inventor, David Chaum, en [13]. Su intención fue proporcionarle anonimato al emisor sobre una gran variedad de topologías de red.

Para comprender este tipo de técnica, se presenta el siguiente ejemplo: Tres criptógrafos están comiendo en su restaurante favorito, después de finalizar su cena el mesonero le informa a los tres que la cuenta ha sido pagada por alguien anónimamente. Ellos respetan esta acción, pero desean

saber si fue uno de ellos quién realizó el pago o fue la agencia de seguridad donde trabajan. Para procurar resolver esta incertidumbre, utilizaron el siguiente método: cada criptógrafo lanza una moneda y muestra el resultado ("cara o cruz") al compañero de su derecha. Esto quiere decir que cada resultado es conocido por dos de ellos solamente, y cada criptógrafo conoce dos resultados solamente. Cada uno compara los dos resultados y sólo revela a los demás si el resultado es igual o desigual (cara-cara y cruz-cruz para ser igual, o cara-cruz o cruz-cara para ser desigual). Si uno de los criptógrafos fue el que pagó la cena, él debería negar el resultado, es decir, si su resultado fue igual (cara-cara o cruz-cruz) debería decir que fue un resultado desigual (cara-cruz, cruz-cara), y viceversa. Cuando el número de resultados desiguales es impar significa que un criptógrafo pagó la cena, en caso contrario ninguno de ellos lo hizo.

Este mismo principio es aplicado a una red de comunicación, y es llamado *soporte de envío*. En esta técnica, cada una de las estaciones envía un mensaje real o uno ficticio en un punto determinado del tiempo, y la superposición (la suma dentro de un grupo conmutativo o Abelian) de estos mensajes serán recibidas por todas las estaciones.

Una primera estación genera claves aleatorias (tanto los caracteres de las claves como los de los mensajes pertenecerán al grupo Abelian), y le comunica sendas claves a las demás estaciones en la red. Estas claves deben ser transmitidas por un canal seguro. En el caso límite, esta procedimiento se repite para cada estación en la red. De esta forma, cada estación tendrá $n-1$ claves (donde n es el número de estaciones en la red), y las cuales mantiene en secreto.

Si una estación quiere enviar un mensaje, ésta superpone todas las claves conocidas junto con el mensaje. La superposición significa que todos los caracteres son la suma del mensaje, de las claves generadas, y el inverso de todas las claves recibidas. Esta es llamada una *superposición local*. Todas las estaciones que no desean enviar un mensaje, deben enviar un mensaje vacío, superpuesto con todas las claves conocidas.

Cada estación envía el resultado de su superposición local (su salida). Todas las salidas que se envían se superponen globalmente. Esto significa que son sumadas (la operación de grupo es aplicada a las salidas locales). La suma resultante se distribuye en cada una de las estaciones en la red. Dado que cada clave y su inversa fue sumada exactamente una vez, las claves se borran unas a otras en la superposición global. Además, el resultado de la superposición global es la suma de todos los mensajes enviados. Si ninguna estación quiere enviar un mensaje, la suma corresponderá al elemento neutral del grupo, que representará un resultado vacío. Si exactamente

un miembro envía un mensaje, la suma será igual a este mensaje.

Si para efectos prácticos son escogidos los dígitos binarios (0 ó 1) como los elementos del grupo Abelian, entonces éste mantiene el envío binario en cada caso, tal como se especificó antes, a través del operador $o - exclusivo$.

Cada clave se utiliza una sola vez, es decir, las claves deben cambiarse en cada una de las rondas. En caso contrario, la salida de una estación que envía un mensaje vacío será idéntica a la anterior, lo cual puede ser utilizado por el atacante para determinar este hecho.

Si dos o más estaciones envían simultáneamente su resultado superpuesto podrían causar colisiones, lo cual representa un problema en sistemas de distribución en canales con accesos múltiples.

Cada participante en el sistema obtiene el resultado de la suma global, y por ende el mensaje original. Debería ser utilizado un mecanismo de cifrado para mantener el contenido del mensaje en secreto (como se utiliza en otras técnicas de anonimato). El uso de las direcciones implícitas podrían preservar al anonimato del receptor.

Este tipo de red es susceptible a los ataques de denegación de servicio, ya que si una estación funciona mal o deja de funcionar, se transmitirían sólo mensajes sin significado.

Esta es una técnica costosa en lo que respecta al tráfico generado, y al ir agregando participantes en la red, se incrementará su tráfico linealmente.

Los objetivos de protección que se logran son el de anonimato del emisor, el anonimato del receptor cuando se utilizan difusión y direccionamiento implícito, el de anonimato de relación, y la no observabilidad del emisor y del receptor a través del uso de tráfico de relleno o "dummy".

Este modelo proporciona protección contra atacantes que sean participantes internos, pero es vulnerable a los ataque de denegación de servicio, sin embargo se pueden descubrir y excluir este tipo de atacantes. En cuanto al modelo de confianza, todos los participantes deberían comprometerse y tolerar el cumplimiento de las reglas.

2.2.7. Mezcladores o Mixes

Esta idea se describe en [12]. El método utiliza criptografía de clave pública y fue diseñado para que los sistemas de envío de correo electrónico proveyeran anonimato del emisor, del receptor y de relación sin necesitar un servicio de confianza central (por ejemplo una autoridad certificadora).

En general, los mezcladores o Mixes pueden ser entendidos como una cadena de proxies seguidos uno detrás del otro. Se considera que el atacante puede observar todas las comunicaciones y puede controlar todos los mixes a excepción de uno.

Topologías Mix

Este concepto funciona aun cuando se dispone de un solo mix. Pero en este caso el usuario debe confiar en este mix. Típicamente hay más de un mix en la red organizados en forma de cadena. Existen diferente métodos para organizar la cooperación dentro de la red. Uno de ellos puede ser que cada mix existe independientemente en la red y los participantes libremente deciden a través de cuál de ellos enrutarán sus mensajes. Así cada nodo puede comunicarse con el resto conformando lo que se denomina una topología de *red mix*.

Otra posibilidad es utilizar una cadena de mixes predefinida. A esta cadena se le denomina *mix en cascada*.

Además de los dos extremos antes mencionados, se pueden utilizar variaciones que resulten en diseños híbridos. Un análisis y comparación de ambas ideas se presenta en [17, 9].

En una red mix, el usuario puede decidir con cuáles mixes desea interactuar, proveyendo de esta manera un buen nivel de escalabilidad y flexibilidad. Además, debido a que los usuarios escogen aleatoriamente los mixes, un atacante no podría determinar cuáles de ellos debería controlar para poder observar un mensaje enviado, para esto debería controlar gran parte de la red.

Por otro lado, un atacante sabe con exactitud cuáles mixes debe controlar en una red en cascada (mix en cascada). Este diseño es vulnerable a los ataques de denegación de servicio, ya que al detener un solo mix en la red, lograría detener todo el sistema.

Por otro lado en [9] exponen que la red mix (pero no la red en cascada) es vulnerable a ciertos tipos de atacantes con altos niveles de control, es decir, que controlan a todos los mixes a excepción de uno. Mencionan que este tipo de red es vulnerable a los ataques $n-1$. Otra desventaja es que algunos mixes puede que no sean casi utilizados (se subutilizan) y otros se sobrecarguen.

Los objetivos de protección que se logran son el de anonimato del emisor, y el de relación. Provee protección contra atacantes que pueden observar toda la red y que pueden controlar muchos mixes. Es susceptible a ataques de denegación de servicio y ataques $n-1$. Desde el punto de vista de la confianza, se debe confiar en al menos un mix de la ruta seleccionada.

Funcionalidad Básica

En este enfoque los usuarios o clientes no envían sus solicitudes directamente al servidor (o a otro destino), sino que las envían a nodos (enrutadores) intermedios denominados mix. Para poder ocultar la comunicación de los participantes, los mixes no envían instantáneamente los mensajes que reciben, en vez de esto, los mixes almacenan varios mensajes de diferentes clientes por un tiempo definido, los transforman, y luego si los reenvían simultáneamente a los servidores de destino o a otros mixes en la red. Un observador que puede ver todos los mensajes entrantes y salientes de un mismo mix no podría determinar cuáles mensajes de entrada corresponden a cuáles mensajes de salida.

Los fundamentos que consolidan la seguridad de los mixes se muestran en la figura 2.12

Preprocesamiento: Transformación de los mensajes

El objetivo principal de la transformación de los mensajes es evitar que un atacante pueda trazar (descubrir su recorrido) un mensaje a través de la comparación de los patrones de bits correspondientes a los mensajes que entran y salen de un mix.

Para poder enviar un mensaje, el cliente primero lo debe preparar. Para esto, el primer paso que debe dar es escoger el camino por el cual se transmitirá el mensaje, este camino estará compuesto por los mixes que haya escogido, y debe incluir el orden específico de reenvíos antes de que llegue a su destino final. Para mejorar la seguridad del sistema, se recomienda utilizar más

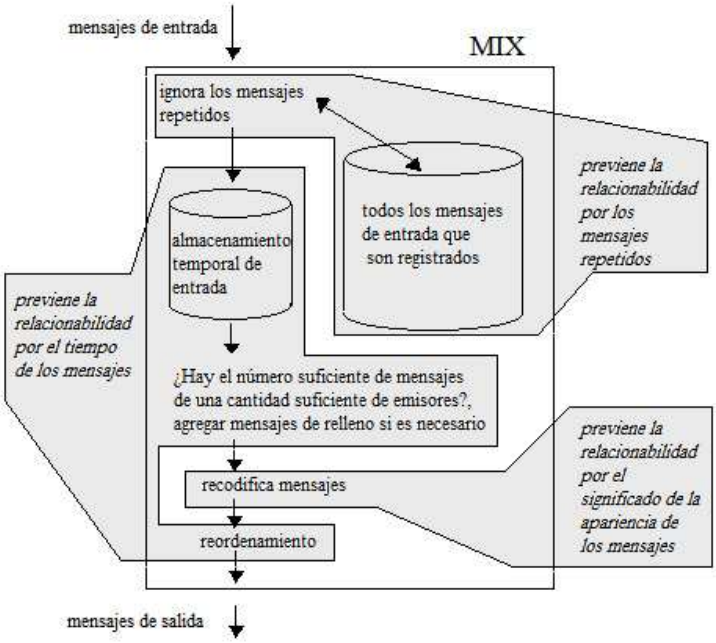

Figura 2.12: Fundamentos que sustentan las redes mix.

de un mix en cada camino. El siguiente paso, es utilizar las claves públicas de los mixes escogidos para cifrar el mensaje, en el orden inverso en el que fueron escogidos, es decir, el mensaje se cifra primero con la clave pública del último mix, luego con la del penúltimo, y así hasta cifrar por última vez con la clave pública del primer mix en el camino seleccionado. Cada vez que se cifra se construye una capa, y se incluye la dirección del siguiente nodo (ya sea el destino final u otro mix). Así cuando el primer mix obtiene el mensaje preparado, lo descifra con su clave privada, y obtiene la dirección del siguiente nodo al que debe reenviarle el resto del contenido que quedó después de su descifrado.

Este esquema puede ser descrito de la siguiente manera:

$A_1, ..., A_n$ pueden ser la secuencia de las direcciones y $c_1, ..., c_n$ la secuencia de la claves de cifrado conocidas públicamente y pertenecientes a la secuencia $Mix_1, ..., Mix_n$ escogidos por el emisor. Incluso c_1 puede ser una clave secreta en un sistema de cifrado simétrico. A_{n+1} puede ser la dirección del receptor o del destino final del mensaje, al cual se le denomina, por simplificación, Mix_{n+1}, y c_{n+1} sería su clave de cifrado. $z_1, ..., z_n$ puede ser una secuencia de bits aleatorias. El emisor crea los mensajes N_i que son recibidos por el Mix_i, y en la base del mensaje N es lo que

el receptor final debe recibir (Mix_{n+1}) supuestamente:

$$N_{n+1} = c_{n+1}(N) \tag{2.1}$$

$$N_i = c_i(z_i, A_{i+1}, N_{i+1}) \text{ para } i = n, ..., 1 \tag{2.2}$$

El emisor le envía N_1 al Mix_1. Después que se decodifica, cada mix recibe la dirección del siguiente mix y el mensaje que está destinado a ese siguiente mix. Debido a las implementaciones de los sistemas de clave pública o asimétrica se necesitan las cadenas aleatorias de bits.

Para asegurar que un atacante no pueda trazar un mensaje (seguir su trayectoria) a través de un mix, es necesario que todos los pares de entrada-salida de los mensajes no tengan características que permitan identificarlos, por ejemplo, el tamaño de los mismos. Una solución a esto es establecer tamaños fijos para los mensajes, y cuando los mensajes tengan un tamaño inferior al fijado, se deberán rellenar con información falsa, y cuando lo superan se deberán fragmentar en varias piezas.

Reordenamiento: Mezclas por grupos (pool) o mezclas por cantidad (batch): Cuando un mix opera en modo "por cantidad" o "batch", éste recolecta un número fijo n de mensajes, cifrándolos y reordenándolos antes de reenviarlos a todos en un solo envío. En contraste, un mix que opera en modo "por grupos" o "pool" tiene siempre un número n de mensajes almacenados en su memoria temporal o "buffer" denominado "pool". Si un nuevo mensaje llega al mix, entonces se escoge aleatoriamente y se reenvía uno de los mensajes almacenados. El número n representa al tamaño del "pool".

Prueba de reenvío: Uno de los tipos de ataques más frecuentes es el denominado ataque de reenvíos. Un atacante podría copiar un mensaje que desea monitorear y enviarle una o varias copias de éste al mix. Estas copias del mensaje podrían tomar el mismo camino en la red que el mensaje original, dado que los algoritmos de envío y descifrado trabajan determinísticamente. Así, puede ser encontrado un patrón característico del mensaje sólo con observar la red. Con el fin de evitar este tipo de ataque, las copias de los mensajes deben ser identificadas y eliminadas a través de un filtro. Una posibilidad para identificar los mensajes inválidos es a través del uso de estampas de tiempo. Cuando un mix obtiene un mensaje, también obtiene una etiqueta que le informa la franja de tiempo durante la

cual el mensaje es válido. Si el mensaje llega muy tarde (después de lo que la franja de tiempo le indica), el mix niega el reenvío del mensaje. Otra posibilidad es que el mix almacene una copia de los mensajes que ya haya enviado, y así los los mensajes nuevos que lleguen pueden ser comparados con esta base de datos. Por razones de seguridad y rendimiento, es conveniente restringir el tamaño de esta base de datos. Los mensajes deberán ser almacenados por un corto período de tiempo antes de que se borren.

Tráfico de relleno o dummy: Aun cuando ninguna información está siendo transmitida, es posible enviar información falsa en la red. Esto tendría el mismo efecto de no enviar ningún mensaje, pero un observador (atacante) no podría distinguir entre los mensajes reales de los que se envían como relleno. El envío de este tipo de mensajes de relleno es denominado *tráfico dummy*. Con respecto a la idea de los mixes, un mix podría aleatoriamente enviar tráfico dummy a otro en la red. Este mecanismo también beneficiaría a los mix que trabajan en el modo batch, ya que normalmente estos mixes tienen que esperar hasta que un número predefinido de mensajes hayan llegado antes de que todos los mensajes sean reenviados simultáneamente, y evitaría los posibles retrasos que podrían ocurrir cuando no hayan envíos suficientes de mensajes al mix, y éste puede hacer su respectivo reenvío, es decir, el tráfico dummy evitaría estos retrasos, ya que si no hay suficientes mensajes reales enviados, el número de mensajes necesarios para hacer el reenvío se pudiese alcanzar con los mensajes de relleno.

Anonimato del receptor (Direcciones de retorno no trazables): El hecho de permitir que un receptor pueda permanecer anónimo se le caracteriza por tener una dirección de retorno que no pueda ser registrada o trazada por un atacante. Esta dirección de retorno es un mensaje especial que tiene que ser creado por el receptor y tiene que ser utilizado por el emisor para el envío del mensaje al receptor anónimo. La idea de base de este tipo de direccionamiento es que el receptor, y no el emisor, define sobre cuáles mixes y el orden a ser utilizado para la entrega de cierto mensaje de respuesta. La dirección de retorno preparada por el receptor contiene una clave simétrica para cada mix en el camino que éste utilizará para cifrar el mensaje enviado por el emisor. Finalmente, el receptor recibirá un mensaje cifrado múltiples veces con claves simétricas como él mismo especificó. Dado que el receptor conoce todas las claves simétricas, para poder desarrollar esta técnica, éste puede descifrar el mensaje. Dado que la claves simétricas son desconocida por el emisor y la codificación del mensaje cambia en cada uno de los mixes (debido al cifrado), el emisor no

puede trazar el mensaje hacia el receptor.

Este esquema se explica de la siguiente forma: $A_1, ..., A_m$ pueden ser la secuencia de las direcciones y $c_1, ..., c_m$ pueden ser la secuencia de las claves públicas conocidas de la secuencia de mixes $Mix_1, ..., Mix_m$ escogida por el receptor, donde c_m puede ser una clave secreta de un sistema de cifrado simétrico. El mensaje añadido a la dirección de retorno pasará por estos mixes en orden ascendiente dependiendo de sus índices. A_{m+1} puede ser la dirección del receptor llamado Mix_{m+1}. De forma similar, al emisor se le llama Mix_0. El receptor crea una dirección de retorno no trazable (k_0, A_1, R_1) donde k_0 es una clave de un sistema de cifrado simétrico generada para este propósito. Mix_0 se supone que utiliza esta clave para codificar el contenido del mensaje con el fin de garantizar que el Mix_1 no sea capaz de leer este mensaje. R_1 es parte de la dirección de retorno, la cual se transmite a través del Mix_0 y contiene el mensaje generado y que ha sido cifrado utilizando k_0. R_1 inicialmente se crea escogiendo aleatoriamente un único nombre e de la dirección de retorno en un esquema recursivo como el que se muestra a continuación:

- R_j es la parte de la dirección de retorno que será recibida por el Mix_j.
- k_j es la clave de un sistema de cifrado simétrico, con el cual Mix_j codifica la parte legible del mensaje.

$$R_{m+1} = e \tag{2.3}$$

$$R_j = c_j(k_j, A_{j+1}, R_{j+1}) \text{ para } j = m, ..., 1. \tag{2.4}$$

El mensaje N_j está constituido por la parte de la dirección de retorno R_j y el contenido I del mensaje (codificado varias veces) generado por el emisor (también llamado *parte I_j del contenido*). Los mensajes N_j son creados por el Mix_{j-1} y son enviados al Mix_j de acuerdo al siguiente esquema recursivo. Estos son creados y enviados por el emisor Mix_0 y así, en secuencia, se pasan a través de los mixes $Mix_1, ..., Mix_m$

$$N_1 = R_1, I_1; I_1 = k_0(I) \tag{2.5}$$

$$N_j = R_j, I_j; I_j = k_{j-1}(I_{j-1}) \text{ para } j = 2, ..., m+1 \tag{2.6}$$

El receptor Mix_{m+1} recibe e, $N_{m+1} = e, (k_m(...k_1(i))...)$ y puede descifrar y extraer el contenido I ya que conoce todas las claves secretas k_j asignadas para el nombre e de la parte de la dirección de retorno en el orden correcto.

Verificación del tamaño del conjunto anónimo: Si un atacante bloquea el mensaje de un participante específico, este mensaje se aísla del conjunto anónimo. Lo mismo sucedería si un atacante rodea a un participante específico, manipulándolo a través de la generación de mensajes con fines ilícitos para el sistema. Este tipo de ataque es conocido como el ataque de mezcla o $n-1$. No existe una solución específica contra este tipo de ataques en ambientes abiertos, como por ejemplo en aquellos donde los participantes entran y salen del sistema a su discreción. Se podría utilizar una protección básica si el mix puede identificar a cada participante, así, de una forma confiable el mix puede verificar si los mensajes que tiene almacenados en su memoria temporal ("buffer") fueron enviados por un número relativamente adecuado de usuarios.

Canales Mix: Los canales mix son utilizados para manejar en tiempo real las cadenas continuas de datos o que contengan sólo pequeños retrasos a través de una cadena de mixes. Para este caso, es necesario que se divida el ancho de banda: una parte para la señalización y otra parte para el envío de los datos, ambos utilizados para la transmisión del mensaje. Se podría asumir que existe un sólo canal para la señalización, y varios canales para la transmisión de datos. Con el fin de establecer el canal, se envía un mensaje sobre el canal de señalización, el cual contiene la clave k_i que deberá ser utilizada entre el emisor y el Mix_i, la cual se cifra de forma asimétrica por la clave pública de dicho mix. Con esto, se define un canal de igual forma para todos los mixes, sobre el cual será transmitido el mensaje. Se podría utilizar un canal para el envío y otro canal para la recepción. Un canal de envío es análogo a un cifrado híbrido: el emisor establece un canal, y codifica continuamente su información N, transformándola en $k_1(k_2(...k_m(N)...))$ y enviándola al mix Mix_1. Cada mix Mix_i para $(i = 1, ..., n-1)$ decodifica los mensajes recibidos continuamente utilizando k_i y transmitiendo el resultado de la decodificación al mix Mix_{i+1}. El mix Mix_m crea el mensaje en texto plano en el fin de la cadena. Esto le permite al emisor enviar anónimamente los mensajes, pero en este caso el receptor no será anónimo. Un canal de recepción es en realidad un canal de envío el cual se utiliza en dirección opuesta, es decir, el receptor es el que establece el canal. El emisor le envía al mix Mix_m la cadena N de información que no está especialmente codificada por el mix Mix_m, luego lo codifica utilizando la clave k_m y conduce $k_m(N)$ un paso atrás, hacia el mix Mix_{m-1}. Los otros mixes hacen lo mismo, por ejemplo, el mix Mix_1 envía la cadena $k_1(...k_m(N)...)$ codificada. Dado que el receptor conoce todas la claves públicas k_i, tiene la disponibilidad de descifrar N. Esto le permite

al receptor recibir los mensajes anónimamente mientras que el emisor no es anónimo. Para alcanzar ambos niveles de anonimato, en [54] sugieren la creación de canales Mix como enlaces de los canales de envío y recepción. El emisor establece un canal de envío que finaliza en el mix Mix_m y el receptor establece un canal de recepción que inicia en el Mix_m. El mix Mix_m traspasa las cadenas de información que llegan por el canal de envío hacia el canal de recepción. Los canales que están supuestamente enlazados, se etiquetan con un marca común que se recibe consistentemente en ambos canales que establecen los mensajes asociados al mix Mix_m. Los datos transferidos están coordinados con un *mensaje de entrada al mix* cifrado asimétricamente, el cual contiene la información del mix que conecta a los dos canales, y el usuario emisor del *mensaje de entrada al mix* actúa como un emisor o un receptor. Cada mix en la cadena puede descifrar este *mensaje de entrada al mix* y en el último paso, el texto plano se difunde a todos los suscriptores. Ahora, los canales pueden ser establecidos utilizando los *mensajes de establecimiento* de ambos participantes. Estos escogen los mixes por el canal de transferencias de datos del mix Mix_m y los mantienen en secreto. Así todos conocen sólo la mitad del camino y el mix Mix_m reenvía los mensajes entrantes del *canal de envío del mix* al *canal de recepción del mix*. Cada emisor/receptor debe tener el mismo número de canales de envío/recepción, porque de lo contrario serían observables, por tal razón convendría utilizar canales "dummy".

2.3. Sistemas Anónimos

En los siguientes apartados se mencionan los principales sistemas hasta los momentos propuestos y que se han implementado en algunos casos.

2.3.1. Sistema de reenvío Anon.penet.fi

Tal como se menciona en [20], el autor de este sistema, el Sr. Johan Helsingius, inició la implementación y ejecución de un modelo confiable para el reenvío de correos, al cual llamó *anon.penet.fi*, el cual proveía cuentas de correo anónimas y seudónimas (1993). El principio técnico detrás del servicio fue la utilización de una tabla de correspondencias entre la dirección de correo real y las direcciones seudónimas, las cuales eran mantenidas por un servidor. Un correo electrónico enviado a un seudónimo le era reenviado al usuario real por medio de dicha tabla. Un correo

electrónico enviado a través de la utilización de un seudónimo, se le suprimía toda la información de identificación y luego se enviaba al receptor. Mientras los usuarios recibían o enviaban correos electrónicos a un seudónimo no tenían la posibilidad de encontrar la dirección real del correo. Sin embargo, esto fue superado trivialmente por los atacantes pasivos locales (o por el mismo servidor) los cuales podían revelar o inferir la verdadera identidad de los emisores de correos utilizando correlaciones de tiempo del trafico entrante y saliente. En 1996 fue clausurada la utilización de este sistema principalmente por causas de índole legal.

2.3.2. Anonymizer y SafeWeb

Anonymizer (http://www.anonymizer.com) es un compañía establecida por Lance Cottrell que procura proveer un servicio de navegación web anónima para los usuarios suscritos al sistema. Anonymizer, como producto de la empresa con el mismo nombre, actúa como un web proxy a través del cual son reenviadas las solicitudes y respuestas web. Los servidores web que son accedidos, no deberían tener la posibilidad de extraer información sobre la dirección del usuario solicitante. El sistema le proporciona un cuidado especial al filtrado del contenido "activo", tales como los javascript o los Java applet, que podrían ejecutar ciertos códigos en las máquinas locales de los usuarios, y enviar por la red la información de la identificación del usuario.

El anonimato depende críticamente de la integridad de la compañía y de sus miembros. El servicio es menos vulnerable a los ataques de compulsión legal, donde en determinados casos organismos públicos, legales, de seguridad, etc. le pueden exigir a la empresa el develado de ciertas identidades. Esto se debe a que no mantienen registros de larga duración que puedan relacionar a los usuarios con los recursos que han accedido. En otras palabras, cuando un usuario solicita una página web y ésta le responde, al finalizar la comunicación entre ambos, todos los registros de la misma son borrados.

Safeweb fue una empresa que proveyó un servicio similar al de Anonymizer, con la diferencia que Safeweb utilizaba para sus conexiones el cifrado bajo el esquema SSL, haciendo seguro el contenido de las páginas accedidas. También utilizó funciones especiales para la detección de contenido activo (como los mencionados con anterioridad). Lamentablemente, y como sucede en casi toda la historia de los sistemas orientados a la seguridad, el sistema no soportó ciertos tipos de ataque, y tuvo que salir del mercado.

En ausencia de cualquier tipo de tráfico de relleno o de sistemas de mezcla o mixes, un atacante pasivo, observador, podría tener la posibilidad de relacionar trivialmente a los usuarios con las páginas accedidas. Para este tipo de ataque se establece que un adversario es capaz de compilar una libraría con las "firmas" del tráfico del usuario y de todos los sitios web de interés que pueden ser accedidos. Estas "firmas" pueden ser comparadas con las características del tráfico de la conexión SSL cifrado para inferir las páginas que fueron accedidas. Esto es conocido comúnmente como *Análisis de Tráfico*, y ha sido estudiado en [21, 43, 64], el cual se detalla mejor en la sección 2.5.

2.3.3. Reenviadores de correos Tipo I Cypherpunk

Los reenviadores de correos Tipo I, tal como se menciona en [20], son nodos que reenvían correos electrónicos después de quitarle toda la información de identificación y después de cifrarlos con su clave privada. La primera implementación de este tipo de idea se hizo para una lista de correos denominada "cypherpunk". El cifrado se desarrolló utilizando las funciones de cifrado de clave pública del PGP (Pretty Good Privacy). También fueron diseñados esquemas de codificación para su ejecución manual, utilizando herramientas estándar para el manejo de texto y de correos. Varios de los reenviadores de correo se podrían encadenar juntos con el fin de evitar la dependencia y las ya conocidas debilidades de utilizar un sólo reenviador.

Soportaba bloques de reenvío para el envío de las direcciones anónimas. La dirección de correo del usuario podía ser cifrada utilizando la clave pública del reenviador, y luego podía ser insertada como una cabecera especial. Si un usuario quería enviar un correo anónimo, entonces el reenviador podría descifrarlo y reenviar el contenido. En este caso no es necesario el uso de bases de datos centrales que relacionen a los usuarios con sus direcciones reales, ya que la información de la dirección para el reenvión de mensajes está incluida en los mismos mensajes de una forma encriptada.

Sin embargo, a pesar de que el esquema de cifrado, que se utilizó cuando los mensajes se reenvían en la red, evita la mayoría de los ataque triviales basados en la observación pasiva del patrón de bits de los mensajes entrantes y su comparación y relación con los de salida, aun sigue dejando información que puede ser útil para cierto tipos de ataques, como el análisis de tráfico. Por ejemplo, el tamaño de los mensajes no se toma en cuenta en los mecanismos PGP, ya que sólo se considera el cifrado y la compresión de los mensajes, y no realiza ningún intento por

ocultar su tamaño, haciendo trivial la tarea de seguir un mensaje en la red sólo con observar sus dimensiones.

2.3.4. Crowds

Crowds fue desarrollado en los laboratorios AT&T, y su especificación completa puede ser vista en [57]. Esta es un implementación de la técnica descrita en la sección 2.2.2. Donde se intenta proporcionar una forma de preservar la privacidad de acceso a la web impidiéndole a los servidores determinar quién está solicitando sus servicios. Cada usuario contacta a un servidor central y éste le envía una lista de todos los participantes en el sistema. Un usuario envía su solicitud web enviándola a través de otros nodos en el sistema, seleccionados aleatoriamente. Después de que cada nodo intermedio recibe una solicitud, decide aleatoriamente también, si la envía a su destino final o a otro nodo intermedio. Finalmente, la respuesta del servidor retorna por la misma ruta establecida en el proceso de envío.

Este sistema es ampliamente conocido en el área de la investigación sobre anonimato, ya que su seguridad recae sobre la imposibilidad del adversario de observar los enlaces. Sólo se puede asumir que el adversario puede controlar una fracción de los nodos en cada "multitud" o "crowd", y al servidor final. Este es un de los primeros trabajos que proporciona un estudio cualitativo sobre el efecto que podría tener la colusión de los nodos sobre el sistema. Además introduce el concepto de *anonimato del iniciador*, el cual establece que cualquier nodo que recibe una solicitud no puede saber si el nodo previo a él era el iniciador de la comunicación (el usuario solicitante) u otro nodo intermedio. Este propiedad, para este sistema, fue contrarrestada utilizando ataques diseñados por investigadores que descubrieron que si un cliente solicita repetidas veces un recurso en particular, estas solicitudes puede ser relacionadas: el ataque recae sobre la intuición de que un iniciador real que genera solicitudes repetidas será el predecesor de un nodo corrupto con mayor frecuencia que un nodo cualquiera en el sistema.

2.3.5. Servidores Nym

Los servidores Nym, presentados en [47], almacenan un bloque de reenvío anónimo y lo "mapean" con una dirección de correo seudónima. Cuando se recibe un mensaje para esta dirección de correo, no se almacena sino que inmediatamente se reenvía anónimamente utilizando el bloque

de reenvío del propietario del seudónimo. En otras palabras, el servidor Nym funciona como una "puerta de enlace" o "gateway" entre el mundo convencional de correos y su mundo anónimo. Dado que este tipo de servidores no mantiene información sobre la identificación de los usuarios, ya que utiliza simplemente los bloques de reenvíos anónimos para efectos de enrutamiento solamente, no se necesita que los usuarios confíen en ellos para salvaguardar su identidad.

Es evidente que los servidores Nym están asociados con las comunicaciones anónimas, y dado que las identidades seudónimas tienen cierta persistencia en el tiempo, entonces hace posible la implementación de sistemas de reputación, u otras medidas que ayuden a evitar el abuso sobre estos sistemas.

2.3.6. El mix de Chaum

El primer trabajo y el más influyente dentro del campo de las comunicaciones anónimas fue presentado en [12], donde se introdujo el concepto de "mix" o mezclador, el cual, como se mencionó anteriormente, es un nodo que esconde la correspondencia (relación) entre los mensajes de entrada y los mensajes de salida, utilizando técnicas de criptografía.

Surgió a finales de la década de los 70, cuando el cifrado RSA de clave pública era relativamente nuevo, y por tal razón en esa época se utilizó RSA en su forma más antigua, basada en una aplicación directa de la exponenciación modular para cifrado y descifrado, solo con un esquema de aleatorización ad-hoc.

La idea principal es que los mensajes a ser "anonimizados" puedan ser reenviados a través de un nodo llamado "mix". Este mix tiene una clave pública RSA, y los mensajes son divididos en bloques y cifrados con esta clave. El primer bloque de mensajes (un grupo pequeño) es conceptualmente el encabezado del mensaje, y contiene la dirección del próximo mix. Al recibir un mensaje, un mix descifra todos los bloques, y le extrae el primer bloque que contiene la dirección del receptor (o del siguiente mix) y le añade un bloque de bits aleatorios (bloque chatarra o "junk") al final del mensaje. La longitud del junk se escoge de tal forma que el tamaño del mensaje permanezca invariable.

En [55] mostraron que el esquema de Chaum no provee necesariamente las propiedades de la no relacionabilidad. La estructura matemática del RSA puede permitir que ciertos ataques activos puedan obtener suficiente información durante el proceso de descifrado para relacionar

los textos cifrados con sus respectivos textos planos. Además es posible desarrollar otros ataques de marcación ("tagging attacks"), ya que los bloques cifrados utilizando RSA no son de ninguna forma dependientes de los otros, y por tal razón los bloques pueden ser duplicados o simplemente sustituidos por otros textos cifrados.

Además, han sido descubiertas otras debilidades de este sistema debido a su utilización de los esquemas RSA, por ejemplo, en los sistemas de firmas.

La segunda función de un mix es la de mezclar varios mensajes juntos con el fin de hacer más difícil la labor del adversario en cuanto a relacionar los mensajes de entrada y salida. De esta forma, un mix tipo batch almacena un número específico de mensajes, los recodifica en cada período, los reordena lexicográficamente y los reenvía.

En este tipo de sistemas también se han utilizado las técnicas de tráfico dummy.

2.3.7. Mixes ISDN, en tiempo real y web

En [54] proponen un sistema para hacer anónimas las conversaciones telefónicas que utilizan tecnologías ISDN. Este diseño puede ser considerado práctico desde el punto de vista de la ingeniería, ya que cumple con la especificaciones y requerimientos de las redes ISDN. Después este sistema se extendió y se generalizó para proveer un marco de trabajo para las comunicaciones en tiempo real, de baja latencia y con mixes. Finalmente varias de las ideas de ISDN y mixes en tiempo real fueron adaptadas para la navegación web anónima y se le dio el nombre de Web Mixes (ver [7]). Parte de este diseño ha sido implementado como un proxy web para anonimato, denominado JAP (http://anon.inf.tu-dresden.de). Estos tres diseños fueron el producto de lo que se podría llamar la comunidad anónima de Dresden.

El objetivo principal de utilizar estas tres propuestas juntas es el de conseguir comunicaciones anónimas seguras, aun en la presencia de adversarios con mucho poder. Este adversario, se asume, puede ser capaz de observar todas las comunicaciones, puede generar retrasos, creación y eliminación de mensajes, e incluso puede controlar cierto número de mixes. Estos tres diseños utilizan como topología de base las redes en cascada, con lo cual aseguran que todos los mensajes son procesados por todos los mixes en el mismo orden. Esto les permite remover la necesidad de pasar o enviar información sobre la ruta en los mensajes y además protege al sistema contra un conjunto de ataques presentados en [8]. El debate sobre las ventajas y las desventajas en la

utilización de la topología en cascada aun se ha mantenido a través de los años y ha permitido explorar las ventajas del uso de nuevas topologías (ver [27, 16, 32])

Además, también proveen cierta protección contra los denominados ataques "tagging", a través del uso del cifrado asimétrico del encabezado. Un cifrado de flujo junto con una clave globalmente conocida se utiliza para transformar el texto plano antes de desarrollar cualquier otra operación de cifrado.

Desde el punto de vista de los aspectos dinámicos de los mixes, ISDN, tiempo-real y web mixes también introducen un concepto novedoso: en la configuración de ruta los mensajes primero son separados del tráfico actual de datos en la red. En los mixes ISDN se utiliza el canal de señalización para transmitir el mensaje codificado en capas (tipo cebolla) el cual contiene las claves de sesión de cada mix intermedio. Cada mix entonces reconoce los mensajes pertenecientes a la misma cadena y utiliza la clave de sesión para preparar el cifrado de flujo y decodificar los mensajes. Es importante señalar que los mensajes para la configuración de ruta y los mensajes que contienen datos se mezclan con otros mensajes que son similares en su forma, es decir, todos los mensajes se incluyen en el proceso de mezcla de los mixes.

Se definieron los denominados *puntos de cita* con el fin de proveer anonimato del emisor y anonimato del receptor. El iniciador o emisor podría utilizar una etiqueta anónima adjuntada a un conmutador ISDN con el fin de conectarse anónimamente con el receptor actual. Este servicio quizás es el circuito equivalente de un servidor Nym, que puede ser utilizado en sistemas basados en mensajes. También se reconoce que para casos especiales como la parte del establecimiento y desconexión de la comunicación y utilizando las líneas ocupadas, un atacante podría ganar cierta información de la misma. En vista de esto, se utilizaron canales con ventanas deslizantes con el fin de sincronizar estos eventos, y hacerlos *no observables* por el atacante. De esta manera, el establecimiento y fin de la llamada suceden en tiempos particulares, en los cuales pueden ser mezclados con muchos otros eventos. Para crear este tipo de procesos (crear la ilusión de que estos eventos suceden en cualquier tiempo), se utiliza el envío de tráfico real o falso para lograr completar el tiempo de duración de la ventana deslizante. Como resultado se obtuvo un modelo de diseño que permitió la transmisión de grandes volúmenes de flujo de datos, preservando las características de seguridad, la no relacionabilidad de las entradas y las salidas, y el mezclado de toda la información relevante.

2.3.8. Babel y mixmaster

Babel fue propuesto en [40] y Mixmaster en [49], ambos diseños fueron presentados a mediados de los años 90, y han sido los sistemas de reenvío de correos electrónico más utilizados hasta el momento. Ambos siguen un *enfoque basado en los mensajes*, que, como se mencionó anteriormente, soportan el envío de mensajes individuales, típicamente constituyendo correos electrónicos, a través de una red de mixes.

Babel ofrece anonimato del emisor, llamado "camino de reenvío" y ofrece anonimato del receptor sobre el denominado "camino de retorno". La parte de envíos y reenvíos la construye el emisor quien envuelve un mensaje en capas de cifrado, tal como sucede en la mayoría de los otros sistemas. Sin embargo, en este caso, el mensaje también incluye la dirección de retorno. De esta forma soporta un anonimato bidireccional por medio del camino de reenvío y las dirección de retorno que ocultan la identidad de receptor.

Este tipo de sistemas no es eficiente contra los ataques de marcación o tagging, ya que a pesar de que provee un buen nivel de seguridad en el envío del mensaje, el proceso de retorno es mucho más débil, esto debido a que el proceso de envío está protegido contra los mensajes duplicados, utilizando estampas de tiempo e identificadores únicos, y funciones hash. Sin embargo, no se utiliza la misma estructura para el proceso de retorno, ya que no se incluye en la información de la dirección de retorno el hash del cuerpo del mensaje. Haciendo posible el retorno de mensajes duplicados a causa del ataque.

Babel también propone un sistema de desvíos intermixes. Esto quiere decir que los mensajes a ser mezclados podrían ser "reempaquetados" por mixes intermedios, y ser enviados a través de una ruta aleatoria sobre la red, distinta a la inicialmente concebida por el emisor. Esto hace que ni siquiera el mismo emisor pueda reconocer sus mensajes en la red.

Mixmaster ha estado evolucionando desde 1995, y soporta solamente anonimato del emisor, o "del camino de reenvío". Los mensajes son procesados a nivel de bit, es decir, que introduce la no relacionabilidad a nivel de bits utilizando un proceso híbrido que combina cifrado RSA, EDE y 3DES, mientras que el tamaño del mensaje se mantiene inalterado anexándole "ruido" aleatorio al final del mismo. En una segunda versión se incluyó un nivel de protección adicional al utilizar una función hash para proteger el encabezado cifrado con RSA, haciendo, de esta manera, imposible los ataque tipo tagging. En una siguiente versión, el "ruido" que se adjuntaba

al mensaje se generaba utilizando una clave secreta compartida entre el reenviador y el emisor del mensaje, y se incluía en el encabezado del mismo. Dado que el ruido es verificable o predecible por parte del emisor, es posible incluir en el encabezado un hash del cuerpo completo del mensaje, protegiendo así su integridad, lo que hace imposible la construcción de réplicas por parte del atacante, esto debido a las razones técnicas asociadas al uso de las funciones hash.

Mixmaster permite que una gran cantidad de mensajes sean divididos en pequeños pedazos o fragmentos, los cuales son enviados de forma independiente. La reconstrucción se lleva a cabo de forma transparente si todos los fragmentos finalizan en un mix común, de esta forma los mensajes de correo electrónico muy extensos podrían ser enviados si necesidad de utilizar programas especiales de envío. Este sistema también implementa una estructura contra los ataques de reputación, es decir, los ataques caracterizados por usuarios que abusan de la red de reenvíos de correos. Esto se logra etiquetando los mensajes y utilizando una "lista negra" con las direcciones desde las cuales no se desea recibir correos anónimos.

2.3.9. Mixminion y Minx: reenvío de correos tipo III

Este sistema fue presentado en [22] en el 2003, y recopiló todas las ideas anteriores para el reenvío de correo anónimo. Este sistema transporta anónimamente los mensajes, fijando su tamaño de envío en 28 kb. Soporta anonimato del emisor, del receptor, a través de lo que se denomina *bloques de reenvío de un único uso*, por ende, soporta el anonimato bidireccional al combinar los dos anteriores en una cadena de reenviadores de correo mixminion, los cuales hacen la mezcla de mensajes correspondiente. Los reenviadores intermedios no conocen su posición en el camino del mensaje ni la longitud total del mensaje, lo cual evita los ataque de partición descritos en [8]. Los reenviadores no pueden distinguir entre otros reenvidores y el emisor del mensaje.

La primera contribución de este sistema recae en el formato de cifrado utilizado. Los mensajes transportados son divididos en dos encabezados principales y un cuerpo. Cada encabezado principal es a su vez dividido en sub-encabezados cifrados con la clave pública de los mixes intermedios. El objetivo principal de esta transformación criptográfica es el de evitar el ataque de marcación o tagging descrito en [55], el cual consiste en que un adversario activo o un nodo corrupto puede modificar un mensaje, con el objetivo de poder detectar dicha modificación después de que un mensaje haya sido mezclado, lo que le permitiría al adversario trazar al mensaje y comprome-

ter el anonimato. El sistema descrito anteriormente, Mixmaster, resolvió este tipo de problema introduciendo una verificación de la integridad en el encabezado de lectura utilizado por cada intermediario: si un cambio fraudulento se detecta en el mensaje, entonces el mensaje se le envía al primer nodo honesto establecido en el camino de envío. Mixminion no puede utilizar este tipo de mecanismo porque depende en gran medida del soporte del enrutamiento indistinguible de los reenvíos anónimos. Por el contrario, este sistema implementa su nivel de seguridad a través de un proceso de cifrado a extremos (cifrado completo o sin cifrar) del segundo encabezado y del cuerpo del mensaje, lo que lo hace más débil.

El formato del paquete de Minx procura proveer las mismas propiedades que el Mixminion a un costo computacional más bajo y con menor exceso de carga (overhead). El sistema Minx fue propouesto en [23], el cual se implementa utilizando un sólo paso de cifrado en un modo que denominaron "IGE", en el cual se propagan, en el reenvío, los errores del texto cifrado. El modificar los mensajes tiene como resultado el no poder descubrir la dirección del receptor y además el mensaje se convierte en algo ilegible. Dado que todos los mensajes tienen un aspecto aleatorio, tampoco se pueden obtener informaciones parciales aun si existen escapes a través de fraudes en la red.

Mixminion utiliza mecanismos de transporte basados en TCP, lo cual permite incluir relleno adicional. Los mensajes se transferieren entre los reenviadores utilizando un túnel protegido TLS, con un intercambio de claves EDH (Ephemeral Diffie-Hellman) lo cual provee seguridad a los envíos. Esto tiene como resultado que cualquier material obtenido sea totalmente ininteligible por un adversario pasivo, además que permite la detección de adversarios activos que intentan corromper los datos en la red. En cualquier caso, para poder tener cierta oportunidad de éxito el adversario debe utilizar varios nodo corruptos en la red.

Después de las primeras versiones del Mixminion, se han propuesto dos ideas para fortalecer la seguridad en el reenvío y la resistencia a la compulsión. La primera, presentada en [22], asume que cualquier comunicación deja un rastro de las claves en los mixes intermedios el cual puede ser utilizado para descifrar las comunicaciones futuras, y propone que una vez que se utiliza la clave, se borre y se actualice utilizando una función de una dirección, dado que los mensajes subsecuentes pueden depender de los mensajes previos en cuanto a su codificación, y así un mix que honestamente borra sus claves no podría descifrar bajo compulsión los mensajes interceptados. La segunda técnica, presentada en [19], recae sobre el hecho de que un receptor

genuino de un sistema de reenvíos anónimos puede funcionar a la vez como un nodo intermedio, emulando una red punto a punto.

2.3.10. Onion routing (OR) o enrutamiento cebolla

Este sistema fue propuesto y estudiado en [46, 67, 68]. Es equivalente a una red de mixes, pero en el contexto de enrutamiento basado en circuitos. En vez de enrutar cada paquete separadamente, el primer mensaje lo que hace es abrir un circuito, etiquetando una ruta. Cada mensaje que tiene una etiqueta en particular se enruta por un camino predeterminado. Finalmente, un mensaje se envía para que cierre o clausure un camino. Con frecuencia se hace referencia a *flujo anónimo* como la información que viaja por estos circuitos.

Su objetivo es dificultarle la tarea al análisis de tráfico, uno de los tipos de ataques más conocidos. Este sistema procura proteger la no relacionabilidad de dos participantes que se comunican a través de terceras partes, y procura proteger la identidad de las partes comunicantes. En vista de que las redes ISDN son difíciles de implementar en Internet, lo que procuró OR es adaptar esta idea distribuyendo la red anónima y adaptándola para que se ejecute en el tope del modelo TCP/IP.

El primer mensaje enviado en la red se cifra en capas, que pueden ser descifradas en una cadena de enrutadores cebolla (onion routers) los cuales utilizan sus respectivas claves privadas. El primer mensaje tiene el material que debe ser compartido entre el emisor y los enrutadores, también las etiquetas y la información de direccionamiento del próximo nodo. Tal como sucede en los mixes de Chaum, se provee la no relacionabilidad a nivel de bits, de esta forma el camino que toma el primer mensaje no es trivial de seguir con sólo observar el patrón de bits de los mensajes. También se propuso un tipo de enrutamiento dinámico donde los enrutadores que reenvían el flujo a través del camino establecido no se especifican únicamente en el mensaje inicial, esto con el fin de incrementar el anonimato.

Los datos que circulan por la red en un circuito establecido están cifrados con claves las simétricas de los enrutadores. Las etiquetes se utilizan para indicar a cuál circuito pertenece cada paquete. Se utilizan etiquetas diferentes para los distintos enlaces, asegurando así la no relacionabilidad, y además las etiquetas de los enlaces también se cifran utilizando una clave que se comparte entre los pares de enrutadores OR. Lo anterior previene los ataques de observadores

pasivos que puedan determinar cuáles paquetes pertenecen al mismo flujo anónimo, pero no le oculta la información a un enrutador que pueda ser subversivo.

OR es susceptible a un conjunto de ataques, tal como el ataque de tiempo. Esto se debe a que los patrones pudiesen ser analizados por un atacante en ausencia de un gran volumen de tráfico pesado.

Para este sistema se afirma proveer anonimato en la navegación web la cual requiere comunicaciones con baja latencia, por tal razón se ha excluido toda la dinámica de los mezcladores o mixes, dado que pudiese incrementar demasiado los tiempos de respuesta. En ausencia de este tipo de características, lo hace vulnerable a distintos tipos de ataques superados por los mixes, por ejemplo el ataque de correlación del tráfico de mensajes, donde se pudiese determinar cuáles mensajes entrantes corresponden con los salientes, con respecto a un enrutador.

Los enrutadores se pueden configurar para que trabajen sólo con un determinado subconjunto de clientes, ya sea por zonas o de forma particularizada. Además se puede configurar para que trabajen sólo con un subconjunto de otros enrutadores.

2.3.11. Tor: la segunda generación de OR

El proyecto OR fue retomado en el año 2004, con el diseño e implementación de lo que se denominó la "segunda generación del onion router" o TOR, por sus siglas en inglés, la propuesta se muestra en [31]. Su política es la del reenvío de flujo TCP sobre una red de reenvíos, y junto con la ayuda de otra herramienta, el Privoxy (http://www.privoxy.org), está especialmente diseñada para el tráfico web.

Este sistema utiliza una arquitectura de red tradicional: una lista de servidores voluntarios se obtiene desde un servicio de directorio ofrecido por otro(s) servidor(es). De esta forma, los clientes crean caminos utilizando al menos tres nodos intermedios escogidos de forma aleatoria dentro de la lista, y sobre los cuales se hace la comunicación de la información. A diferencia de la arquitectura anterior, donde se enviaba y distribuía el material criptográfico, TOR utiliza un mecanismo interactivo: el cliente se conecta con el primer nodo, y le solicita a éste que se conecte con el segundo nodo, de esta forma un canal bidireccional se utiliza en cada paso para desarrollar un intercambio de claves autenticado DF (Diffie-Hellman). Éste garantiza el reenvío en forma secreta y la resistencia a la compulsión, debido principalmente a que solo son necesarias

claves de corta duración. Este mecanismo fue inicialmente propuesto en Cebolla (ver [10]), y no está cubierto en la patente de OR (ver [67]).

Otra notable diferencia entre TOR y los intentos anteriores por anonimizar el tráfico de flujo, es que TOR no ofrece seguridad contra los atacantes que pueden observar la red entera, es decir, contra atacantes pasivos globales. Un conjunto de técnicas de Análisis de Tráfico (ver [18, 42, 61, 73, 69]) han sido desarrolladas a través de los años para trazar el flujo de tráfico continuo viajando por redes de baja latencia como TOR. En estos estudios se ha demostrado que este tipo de ataques son muy difícil de contrarrestar, a menos que se utilicen técnicas que implicarían latencias elevadas, o que requieran la inyección de grandes cantidades de tráfico cubierto (tráfico inservible o "dummy"), los cuales representan soluciones muy costosas. Por esta razón en TOR se opta por obtener un nivel de seguridad que se pueda alcanzar en un sistema altamente utilizable y muy económico de utilizar (ver [3, 45]). Como resultado si un adversario puede observar el flujo entre dos puntos de la red, pudiese de forma trivial generar el mismo tráfico, y lograr ataques del tipo "tagging".

Sin embargo, dada esta vulnerabilidad, aun se necesita estimar la probabilidad de que un adversario pueda estar monitoreando la red en múltiples puntos sobre un camino o ruta establecida.

TOR también ofrece mecanismos para ocultar los servidores. Un servidor oculto abre una conexión anónima y la utiliza para publicar un punto de contacto. Si un cliente quiere contactar a un servidor, debe conectarse con un punto de contacto y negociar un canal anónimo separado del que se utiliza para el reenvío de la comunicación actual. Un ataque propuesto en [51] demuestra la vulnerabilidad de esta idea. La intuición detrás de este ataque está en el hecho de que un adversario puede abrir múltiples conexiones hacia un mismo servidor oculto, y secuencialmente o en paralelo podría controlar el flujo hacia ese servidor. Para esto, el atacante necesitaría controlar al menos un enrutador, y debe esperar a que el servidor escoja una de las conexiones de su enrutador como un primer nodo de un camino anónimo cualquiera.

2.3.12. Redes mix punto a punto

En el trabajo original de Chaum se asumió que cada participante en la red mix también actúa como un mix para otros, con lo cual se mejoraría la seguridad global de la red. El reciente interés

por las redes punto a punto ha influido en algunos investigadores para que examinen este tipo de redes como un gran número cambiante o dinámico de mixes.

En [37] se presenta un sistema denominado Tarzán, una red punto a punto en la cual cada nodo constituye un mix. Un nodo o punto que inicie el transporte de un flujo a través de la red crea un túnel cifrado hacia otro nodo, y le solicita a ese nodo que envíe ese flujo hacia otro. A través de la repetición de este proceso, es posible obtener una conexión cifrada tipo "cebolla" (OR), con reenvíos a través de un secuencia intermedia de nodos. Una de sus principales características es que la topología de red está restringida de cierta manera. Cada nodo mantiene conexiones persistentes con un conjunto pequeño de otros nodos, conformando un estructura denominada *mimics*. Las rutas de los mensajes anónimos son seleccionadas de tal forma que podrían ir a través de la red mimics sin poder ser relacionados aun teniendo un flujo con poco tráfico. Una de las grandes debilidades del esquema mimics es que el mecanismo de selección de los nodos vecinos se hace sobre una base de identificadores o direcciones de red básica, la cual, desafortunadamente, la hace vulnerable a los ataques de suplantación o spoof de las redes reales.

Tarzán fue diseñado originalmente con el requerimiento único de que cada nodo conociera de forma aleatoria al subconjunto de nodos con los cuales podría establecer comunicación. Esto es obvio, dada las características dinámicas de las redes punto a punto, y la volatilidad de sus nodos. Por otro lado, en [24] se proponen algunos ataques contra esta estrategia utilizada en sus primeras versiones, y se basan en el hecho de que la red es muy grande, y los nodos se comunican con un alto grado de entropía. Como resultado un sólo nodo podría conocer a un subconjunto pequeño de otros nodos, y un nodo adversario incluido en el camino anónimo podría determinar cuál es el nodo origen sólo con conocer tres nodos en la red: el nodo corrupto como tal, su sucesor, y su predecesor. Esto trae como consecuencia que con estos tres nodos se podría descubrir con una alta probabilidad al iniciador de un flujo de comunicación. La versión final de Tarzán evita este tipo de ataques exigiéndole a cada nodo que conozca a todos los demás.

En [58] plantean un sistema con una arquitectura muy similar denominado *MorphMix*. La diferencia principal radica en que no se especifica la ruta de la comunicación anónima en el origen (nodo iniciador) sino que lo hace un nodo intermedio en el cual confía el emisor, el cual se somete a la observación de terceros (testigos). Sin embargo, a pesar de que el ataque mencionado anteriormente no podría ser aplicado en la selección de la ruta, si se podría aplicar en el proceso de la selección de los testigos. Este sistema, por otro lado, incluye un mecanismo de detección

de colusión el cual monitorea cualquier ciclo en la selección de los nodos de la ruta.

2.4. Métricas de anonimato

En [17, 26] proponen y mencionan algunas formas de medir el anonimato. Se menciona que la necesidad de poseer métricas que permitan cuantificar o medir el rendimiento de las implementaciones del anonimato aparecieron cuando se desarrollaron las primeras aplicaciones para las transacciones electrónicas anónimas, tales como la votación electrónica, los sistemas de correo anónimos, los sistemas electrónicos monetarios y los navegadores anónimos.

Desde el inicio se buscaba dar respuesta a estas interrogantes: ¿cómo puede ser medido el anonimato?, ¿cómo se pueden diferenciar dos o más sistemas anónimos?, ¿cómo se puede evaluar la efectividad de los diferentes ataques sobre los sistemas anónimos?, ¿cómo se pueden cuantificar las pérdidas y las ganancias del anonimato?, ¿cómo se puede medir la información que puede ser obtenida por el adversario?,¿qué es un nivel adecuado de anonimato?.

Todas estas interrogantes se han procurado responder, en mayor o menor medida, utilizando varios tipos de métricas que pretenden medir el tamaño de un conjunto de sujetos potencialmente enlazados o relacionados entre sí y la distinción o diferenciación entre ellos, asociados a una transacción en particular, y bajo la influencia de un tipo específico de atacante. En otras palabras, las métricas están relacionadas al tamaño de ese conjunto, a las diferencias entre los miembros del conjunto y al tipo de ataque considerado. Por ende, para poder obtener una cierta idea del rendimiento de una implementación de un sistema anónimo bajo distintas condiciones, es necesario generar múltiples medidas del anonimato y ser analizadas en cada caso.

En [57] definen el *grado de anonimato* como una probabilidad $1-p$, donde p es la probabilidad asignada por un atacante a un potencial emisor. En este modelo, los usuarios son más anónimos si aparentan tener menor posibilidad de ser el emisor de un mensaje, ante un atacante en particular. Si se considera un caso donde existen sólo dos usuarios, con igual posibilidad de ser elegidos, esto implicaría que la probabilidad (el grado de anonimato) de cada uno es del 50 %, pero si en un segundo escenario se tienen 1000 usuarios, este modelo le asigna el 50 % de la probabilidad de ser el emisor al primer usuario, y el resto tendría una probabilidad inferior a 0,001 de haber enviado el mensaje, esto implica que para los dos escenarios el primer usuario tiene la misma probabilidad de ser elegido, pero si se comparan entre sí, en el primer caso los dos usuarios son indistinguibles

para el atacante, y en el segundo caso, el primer usuario tendría la mayor posibilidad de ser elegido, es decir, se distinguiría del resto.

En [6] definen el *grado de anonimato* como $A = log_2(N)$, donde N es el número de usuarios del sistema, por ende esta métrica sólo depende de la cantidad de usuarios, y no expresa las diferencias que existen entre los distintos sistemas anónimos con igual número de usuarios. Incluso el número total de usuarios podría ser desconocido. Además, un adversario pudiese sólo considerar un subconjunto de todos los posibles usuarios (en este caso emisores de un mensaje) tomando en cuenta información probabilística, evitando así analizar al conjunto completo.

En [17] y en [28] proponen utilizar métricas para el anonimato basadas en la teoría de la información, para lo cual utilizan la *entropía* como medida del tamaño efectivo del conjunto anónimo. En la segunda de estas propuestas dan un paso más allá y plantean normalizar dicha entropía con el fin de acotar sus valores en una escala $0 - 1$. Para la definición de estas métricas se utilizó es siguiente modelo:

Si se considera el envío y la recepción de mensajes como ítems de interés (IDI), y a un usuario (emisor o receptor de mensajes) en particular como un sujeto, el anonimato puede ser definido como la no relacionabilidad entre un ítem de interés y un sujeto. Específicamente se describe el anonimato de un IDI tal que no sea relacionado con cualquier sujeto, y que un sujeto no sea relacionado con cualquier IDI. De esta manera la no relacionabilidad se incrementa si se consiguen valores altos de entropía.

En la figura 2.13 se representa este modelo de anonimato de una forma simplificada (extraído de [28]).

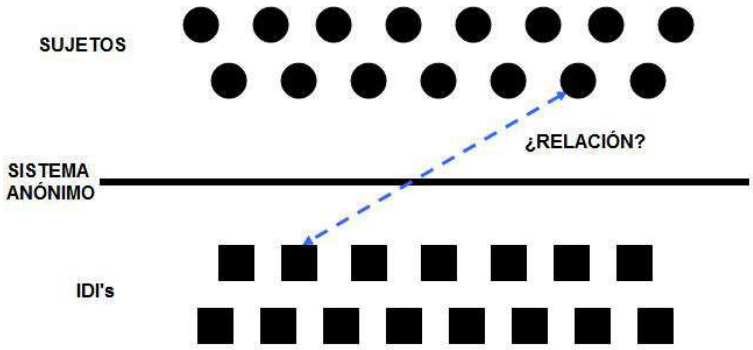

Figura 2.13: Sistemas Anónimos.

El objetivo de los sistemas anónimos es esconder u ocultar la relación que existe entre los sujetos y los ítems de interés, es decir, el ocultar estas relaciones o enlaces es el mecanismo básico de los sistemas anónimos. Un atacante pasivo (observador) puede ver a un conjunto de sujetos acceder al sistema anónimo, y en el otro lado del sistema ve a un conjunto de ítems de interés que son difícil de relacionar a un sujeto en particular. El conjunto de sujetos que pueden ser relacionados a un IDI se le denomina el conjunto anónimo (para este modelo). Mientras más grande sea este conjunto, mayor anonimato tendrá el sujeto. Por esta razón, la noción de este conjunto anónimo juega un rol fundamental en la definición de este tipo de métricas.

Se pueden considerar dos tipos de ataques, uno que atenta contra el anonimato en sí y el otro contra la disponibilidad del servicio que presta el sistema anónimo (comúnmente denominado denegación de servicio). Este último sólo puede ser generado por atacantes activos (que pueden generar cambios en la comunicación). La cuantificación del anonimato es dependiente del tipo de atacante considerado. Para un tipo de ataque pasivo (observador) el adversario puede desarrollar distintas estrategias con el fin de obtener información sobre la relación que existe entre los IDI y los sujetos. La mayoría de estos ataques se basan en asignarle una distribución de probabilidad a la asignación de estas relaciones IDI - sujeto.

La métricas propuestas considerando la teoría de la información (entropía) se basan en las probabilidades (su distribución) que le asigna el atacante a los usuarios para ser relacionados a los distintos IDI. Si el modelo de ataque varía, los resultados de las mediciones pueden que no sean válidos después de los cambios. Por esta razón conviene describir los tipos de ataques utilizando las siguientes etiquetas generales:

- Pasivo-activo: Un atacante pasivo escucha la comunicación y/o lee la información interna de las entidades participantes en los distintos protocolos de comunicación. Este tipo de adversario genera típicamente un Análisis de Tráfico. Un atacante activo puede agregar, modificar, mover, retrasar y/o eliminar mensajes y la información de las entidades participantes.
- Interno-externo: Un atacante interno controla una o varias entidades que son parte del sistema. Los atacantes externos sólo controlan los enlaces de comunicación.
- Parcial-global: Un atacante global puede observar todo el sistema, y un atacante parcial sólo puede ver a un conjunto limitado los recursos.

- Estático-adaptativo: Un atacante estático controla sólo un conjunto predefinido de recursos, y no le es posible generar cambios una vez que se inicia el proceso de ataque. Un atacante adaptativo puede generar cambios en el ataque durante la ejecución del proceso de ataque, según los resultados intermedios del mismo.

La entropía desde el punto de vista de la teoría de la información, propuesta en [62], propone una medida del nivel de incertidumbre de una variable aleatoria. Sea X una variable aleatoria discreta con función de probabilidad $p_i = Pr(X = i)$, donde i representa cada posible valor que X puede tomar con probabilidad $p_i > 0$. En este caso cada i representa a los sujetos dentro de un conjunto anónimo, y p_i es la probabilidad de que el sujeto i sea relacionado al IDI.

Se denota $H(X)$ como la entropía de una variable aleatoria X y por N al número de sujetos pertenecientes al conjunto anónimo. $H(X)$ puede ser calculada de la siguiente manera:

$$H(X) = -\sum_{i=1}^{N} p_i \log_2(p_i) \tag{2.7}$$

El *conjunto efectivo anónimo* se puede definir como una métrica del anonimato utilizando el valor de la entropía. Tal como se mencionó, después de que un adversario genera un ataque, éste obtiene una distribución de probabilidad que relaciona a los sujetos con el IDI del ataque. Tal como se muestra en la figura 2.14, extraída de [26], las probabilidades son mostradas como flechas que conectan al IDI con los distintos sujetos. El valor de cada probabilidad p_i puede variar según la información que haya obtenido el atacante. En este caso N denota al número total de sujetos que el atacante relaciona (potencialmente) al IDI, con probabilidad distinta de cero. Entonces el conjunto efectivo anónimo puede ser definido como la entropía $H(X)$ de la distribución de probabilidades de X que relacionan a los sujetos del conjunto anónimo con el IDI. Esta métrica significa que cuando el conjunto efectivo anónimo tiene un valor h, quiere decir que los sujetos anónimos son perfectamente indistinguibles dentro de un conjunto de 2^h sujetos. Esta métrica y de igual forma el anonimato, puede incrementar dependiendo de dos factores, uno es el tamaño del conjunto anónimo y el otro es la uniformidad de la distribución de probabilidad.

Con lo anterior, el *grado de anonimato*, propuesto en [28], se define como la versión normalizada del conjunto efectivo anónimo, el cual indica qué tan bueno es el sistema anónimo en una escala de $0 - 1$. El tamaño máximo del conjunto efectivo anónimo se alcanza cuando todos los sujetos han sido relacionados al IDI con igual probabilidad ($p_i = 1/N$), es decir, todos los

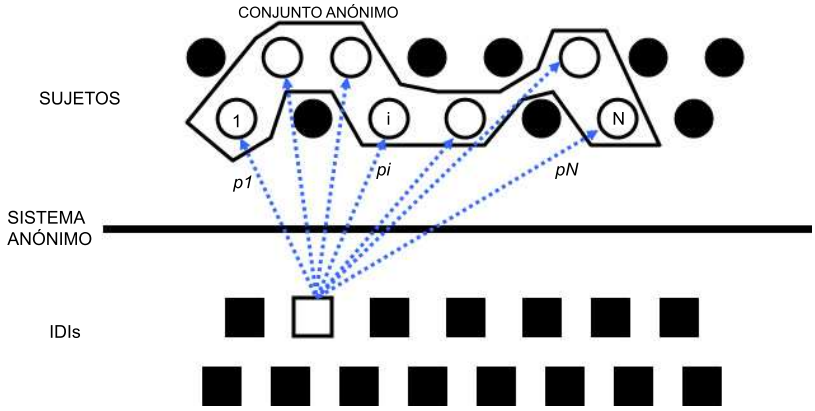

Figura 2.14: Conjunto Anónimo.

sujetos son perfectamente indistinguibles para el adversario con respecto al IDI. Para un número N de usuarios, el máximo nivel de anonimato se alcanza con una entropía que corresponde a una distribución de probabilidad uniforme, y este valor se denota como sigue:

$$H_M = \log_2(N) \tag{2.8}$$

Si se asume que el adversario no tiene información previa sobre las comunicaciones, entonces la cantidad de información que el adversario obtiene con un ataque es la diferencia de la entropía antes del ataque menos la entropía que resulta después del ataque.

El *grado de anonimato*, entonces, se define como esa diferencia normalizada que representa el grado de conocimiento del adversario después del ataque, la cual se puede representar de la siguiente manera:

$$d = 1 - \frac{H_M - H(X)}{H_M} = \frac{H_M}{H(X)} \tag{2.9}$$

Se puede observar que el grado de anonimato se obtiene dividiendo el valor del tamaño del conjunto anónimo por el grado de anonimato máximo para un número de sujetos dado. Este grado da una referencia del valor del anonimato independiente del número de sujetos que se considere en el conjunto anónimo. Dado un número específico de sujetos, el *grado de anonimato* da una idea de cuán cercano está el valor del anonimato después del ataque del valor máximo posible.

2.5. Tipos de Ataques

En las secciones anteriores se han mostrado, a grandes rasgos, la terminología más relevante en el área del anonimato. Constantemente han sido mencionados distintos tipos y estructuras de ataques, incluso se han definidos modelos de ataque para la definición de las métricas. Todo esto tiene una causa fundamental: en el área de la seguridad lo que se busca es proteger a algo o a alguien contra el "ataque" de algo o de alguien, si bien es claro que para cada rama de la seguridad en informática se tienen varios modelos o tendencias de ataques, la privacidad y específicamente el anonimato no son la excepción. Todos los diseños de sistemas, y las medidas de efectividad, rendimiento, etc., están basadas en el tipo y forma de los ataques. En está sección se presentan los tipos de ataque más comunes que se conocen y que se han diseñado contra los sistemas anónimos. Tal como se menciona en [1], cualquier sistema anónimo debe ser evaluado por su resistencia a varios tipos de ataque. Esto implica que una de las primeras decisiones a tomar en una fase de análisis es el tipo de adversario que se considera. Por ejemplo, un modelo de adversario muy débil implica un escenario de seguridad muy optimista, que implicaría un mayor nivel de susceptibilidad ante ataques reales. En la comunidad de la seguridad se prefiere en general un modelo conservador del adversario, ya que al considerar a un adversario muy poderoso se debe incurrir en inversiones y gatos muy onerosos que muy pocos usuarios y operadores están en la disponibilidad de cubrir.

Con la experiencia en la red TOR (ver [31]), se puede concluir que en algunos casos al considerar un modelo de adversario débil puede ayudar a mejorar la usabilidad y atraer mayor número de usuarios, lo cual de por sí mejora el nivel de anonimato del sistema. De cualquier forma, como mínimo debe ser considerado un modelo "realista y consciente" del adversario.

En la tabla 2.5 se listan los modelos de adversarios a considerar en los sistemas anónimos. El modelo más común que se ha considerado es el *adversario pasivo y global*, quien puede observar todos los mensajes enviados a todos los participantes. Los sistemas de alta latencia como Babel, Redes DC, y Mixminion están diseñados para resistir los ataques de este tipo de adversario. Por el contrario, los requerimientos de los sistemas de comunicación de baja latencia los hacen más vulnerables a los ataques que aplican técnicas de *análisis de tráfico*. Tal como se mencionó antes, en el caso de TOR, se asume un modelo de adversario más limitado que sólo puede observar una fracción c limitada de la red.

Tabla 2.1: Clasificación del Adversario en Redes Anónimas, su influencia en la red (donde c es la proporción de usuarios deshonestos en la red y $0 < c < 1$), y su comportamiento

Adversario	% de la red afectada	Comportamiento
Adversario pasivo global	100 %	Observa el tráfico
Adversario pasivo local	c	Inserta, borra y modifica el tráfico
Adversario activo global	100 %	Observa el tráfico
Adversario activo local	c	Inserta, borra y modifica el tráfico
Adversario participante	c	Participa en la red ejecutando un mix

En la tabla 2.5 se muestra una lista de los sistemas que han sido analizados desde la perspectiva de cada modelo.

Tabla 2.2: Clasificación del Adversario en Redes Anónimas y los sistemas que han sido analizados con esos modelos

Adversario	Systemas
Adversario pasivo global	Babel, Redes DC, Redes Mix, Mixmaster, Mixminion, Tarzan
Adversario pasivo local	Crowds, Redes Mix
Adversario activo global	Redes Mix
Adversario activo local	Mixminion, TOR
Adversario participante	Crowds, MorphMix, Tarzán, TOR

Un adversario activo puede alterar la comunicación al borrar, insertar, retrasar y modificar los mensajes. Algunos de los ataques más comunes que utilizan este modelo se caracterizan por incrementar o decrementar el volumen de tráfico que entra en la red o a un nodo de la red, y observar los efectos de dichas modificaciones en la salida de un nodo. Los adversarios también pueden modificar los mensajes, por ejemplo introduciendo ciertas clases de etiquetas (como un error en el patrón de bits), o pueden reenviar mensajes antiguos, y verifican en ambos casos las correspondencias de tales errores o duplicados en la otra parte de la red con el objetivo de establecer el camino de los mensajes. Este tipo de ataque se denomina ataque de etiquetado o "tagging".

En cuanto al modelo de un adversario que constituye un participante interno, se considera que la mayoría de los sistemas de comunicaciones anónimas están diseñados como una red de nodos. Este modelo supone que alguno de estos nodos pertenecen al atacante. Esto puede lograrse comprometiendo algunos de los nodos previamente existentes o introduciendo nuevos nodos al sistema. Dado que las redes anónimas típicamente se ejecutan con la colaboración de voluntarios, estas redes aceptan con gran facilidad la incorporación de nuevos participantes. Sin embargo, este tipo de atacante se puede modelar y configurar mejor que el adversario global y pasivo.

2.5.1. Ataque bizantino - Sybil

Se puede subclasificar al adversario participante interno en dos categorías: Uno con características de honestidad, pero curioso (HBC por sus siglas en inglés), quien se comporta tal cual como lo que se establece en los protocolos de anonimato, pero que intenta aprender tanto como le sea posible sobre la comunicación a través de la observación. Dado que es un participante pasivo se asume que puede observar algunos datos privados de otros participantes. La otra categoría considera un adversario quien puede violar al protocolo y tener un comportamiento arbitrario. Este tipo de adversario se le denomina *bizantino*, el cual puede mentir sobre sus capacidades y conocimientos en la red, y puede intervenir en las comunicaciones, pero también podría seguir las normas de un protocolo si así lo requiriese su ataque. Una particularidad de este tipo de adversario es que activamente intenta evitar que se le detecte para dificultar la tarea de aislar a los nodos maliciosos.

El ataque bizantino más prominente es el denominado ataque "Sybil", descrito en [34], en el cual los nodos violan al protocolo pretendiendo ser instancias de múltiples nodos dispersos con el fin de obtener influencias indebidas sobre la red.

Sin embargo, sobre este tipo de ataques bizantinos aun falta mucho por explorar, y podrían surgir ataques en base a este modelo que tengan un impacto significativo sobre las redes anónimas.

2.5.2. Análisis de Tráfico

Tal como se menciona en [56, 3, 1] el análisis de tráfico tuvo sus orígenes durante la Segunda Guerra Mundial, incluyendo su relación con el ataque que se hizo sobre Pearl Harbour.

Actualmente Google utiliza los enlaces de incidencia para evaluar la importancia de las páginas web, las compañías de tarjetas de crédito examinan las transacciones para descubrir patrones de gastos fraudulentos. La idea de fondo de esta técnica radica en que durante el tráfico de datos se pueden registrar el tiempo y la duración de la comunicación, y se examina esta información para determinar la forma detallada del flujo de datos, las identidades de las partes que se comunican, y lo que puede ser establecido sobre su ubicación. Incluso los datos pueden ser poco precisos o estar incompletos, y simplemente a través del conocimiento de patrones típicos de comunicación se podría inferir sobre una comunicación en particular que se esté observando.

Esta técnica a pesar de que obtiene información de menor calidad en comparación a la obtenida con las técnicas del criptoanálisis, es mucho más fácil, barata y viable en cuanto a la extracción y procesamiento del tráfico de datos.

El análisis de tráfico ha inspirado a otras técnicas utilizadas para la protección de sistemas, y para la construcción de sistemas de confianza.

En cuanto al análisis de tráfico del protocolo de seguridad SSH, a pesar que ofrece comunicaciones seguras para acceder a terminales remotos a través de un proceso de autenticación que utiliza mecanismos de clave pública, y que luego de este proceso toda la información viaja encriptada garantizando su confidencialidad e integridad, en [72] se muestra que aun existe gran cantidad de información que aun se deja pasar sin ocultarse. En su modo interactivo, el protocolo SSH transmite cada pulsación de tecla como un paquete distinto y de esta forma la longitud de la clave puede ser trivialmente descubierta. Además, dado que la distribución de los teclados no es aleatoria, y que las claves con frecuencia están basadas en palabras reales, el tiempo exacto de las pulsaciones de teclas está relacionado a qué tan rápido un caracter particular puede ser tipeado después de otro. Esto implica que al haber una suficiente variabilidad entre los patrones de tipeo de las personas entonces existe la posibilidad de identificarlos, particularmente después de observar una secuencia larga.

El análisis de tráfico del SSL y el TLS, protocolos introducidos para proveer accesos web privados, se basa en estudiar la información que aun deja escapar este protocolo cuando se establece una comunicación web hacia un servidor. Específicamente se estudia la forma del tráfico que ha sido rellenado y conciliado inadecuadamente. Los navegadores solicitan recursos, que son normalmente páginas HTML, y éstas a su vez están asociadas a otras recursos adicionales como imágenes, hojas de cálculos, etc., las cuales pueden ser descargadas a través de enlaces cifrados,

pero su tamaño aun puede ser determinado por el observador, quien puede inferir cuáles páginas están siendo accedidas (por ejemplo podría inferir cuáles reportes de una compañía están siendo descargados).

Incluso, estudiando el comportamiento del web cache puede inferirse cuáles sitios web se han accedido con sólo reconocer el patrón de almacenamiento en los cache de cada uno.

Se puede determinar la identificación de los dispositivos en la red sólo con estudiar el comportamiento y las características particulares de los cambios del reloj en cada dispositivo.

Capítulo 3

Fundamentos Teóricos

3.1. Sistemas de Reputación y Confianza

Tal como se menciona en [52], frecuentemente en las sociedades humanas se toman decisiones en base a información que está incompleta, es decir, con cierto grado de incertidumbre. Por ejemplo, se puede decidir comprar un producto desde un sitio web, lo que incluye evaluar la posibilidad (el riesgo) de recibir el producto de forma adecuada y en el tiempo previsto. Se podría determinar esta incertidumbre a través de la obtención y evaluación de la información que pueda ayudar a hacer predicciones sobre los aspectos relacionados a las decisiones con incertidumbre. En el ejemplo mencionado, se puede consultar el historial personal de compras de productos asociados a ese sitio web, o se le puede consultar a personas cercanas (amigos, familiares, etc.) sobre su experiencia al utilizar dicho sitio. En este sentido, cada vez que la compañía entrega un producto en buen estado y en el tiempo previsto se puede predecir que lo más probable es que esto vuelva a ocurrir en el futuro. Por otro lado, si la mayoría de las personas cercanas (amigos o familiares) comentan que cada vez que hicieron tratos con esa compañía quedaron satisfechos con el servicio, también se podría predecir que lo más probable es que se obtenga ese mismo comportamiento en el futuro. Como una opción adicional, se podrían examinar las conexiones sociales que el sitio web ha tenido con otras compañías, o las conexiones que se tienen con los empleados y líderes del sitio. De esta forma se podría decir, en términos generales, que en el proceso de toma de decisiones se evalúa la incertidumbre en las decisiones utilizando el concepto

de confianza. Así, la confianza es el grado de creencia en las acciones de otras personas que afectan el estado propio de otros (incluyendo el que toma las decisiones) al escoger una acción, y que forma un componente integral del razonamiento y racionalidad del decisor. En este contexto, al evaluar la honradez de los demás permite lidiar con la incertidumbre en el proceso de toma de decisiones. Según el ejemplo anterior, y considerando este concepto de confianza, se puede decir que se *confía* en que una compañía entregará el producto solicitado con las condiciones y en el tiempo previsto. Si la compañía cumple exitosamente lo establecido en el contrato, entonces se podría considerar en la evaluación como un incremento de la honradez de la compañía. En caso contrario, disminuiría.

Sin embargo, en una comunidad abierta, es probable que se tenga que interactuar con muchas entidades de las cuales no se tenga un historial de interacción. En estos casos, los común es recurrir a la opinión de terceros, incluyendo la evaluación de algunos entes reguladores de la industria. Típicamente, si no se tiene información suficiente para poder conformar una opinión sobre una entidad, se debe consultar a otros sobre su opinión. La opinión común de otros sobre una entidad se conoce con el nombre de *reputación de la entidad*, la cual puede ser utilizada en ausencia de la confianza formada en base a las opiniones personales.

Según lo anterior la confianza y la reputación están afectadas por las acciones pasadas de los individuos. Sin embargo, existe otro factor en la determinación del nivel de confianza que se le otorga a los demás: los seres humanos implícitamente evalúan la honradez de los demás frecuentemente examinando la estructura social circundante al individuo, por ejemplo, es frecuente utilizar el estatus social (gerente de una compañía), el rol dentro de un sociedad (doctor de una clínica) y las relaciones (amigo de un familiar) para establecer su nivel de honradez.

Desde un punto de vista informático, el uso de la confianza dentro del establecimiento de las relaciones incluye una selección optimizada de los patrones de comunicación, la delegación de funciones sobre los individuos y la posibilidad de establecer acuerdos entre dos o más miembros de una red antes que se inicie la comunicación. En [38] se define la confianza como un nivel particular de probabilidad subjetiva en el cual, en un sistema multi-agentes, un agente puede evaluar si otro agente (o grupo de agentes) realizarán una acción en particular, antes de que pueda monitorear tal acción en un contexto donde puede afectar sus propias acciones. El concepto de reputación representa una confianza indirecta e involucra el tener que preguntar la opinión de terceros quienes previamente han interactuado con el ente sobre el cual se quiere establecer el

nivel de confianza.

3.2. Pequeños Mundos

En [48, 70] se plantea el surgimiento de la teoría moderna de redes (denominado Ciencias de las Redes o "Science of Networks") como una evolución de los desarrollos teóricos de Leonard Euler sobre el estudio de los objetos formales denominados *grafos*, los cuales revolucionaron primero las matemáticas teóricas y luego se proyectaron sobre las ciencias sociales.

Una de las características principales de este planteamiento es la idea de que las redes de naturaleza empírica no son estructuras estáticas sino que evolucionan en el tiempo en por lo menos dos vertientes: en primera instancia generando un "producto" (envío de información, toma de decisiones, generación de poder) y en segundo lugar produciendo modificaciones en su propia morfología. En otras palabras, lo importante es que lo que sucede y cómo sucede depende tanto de la estructura de la red como de su historia evolutiva.

El fenómeno de la sincronía es aquel atributo relacional que permite establecer procesos de coordinación, que en su mayoría son imprevistos, con respecto a distintos nodos que interactúan entre sí. Por ejemplo, se pueden mencionar las competencias de atletismo en las que los corredores tienen gran interés por permanecer a escasa distancia de sus competidores rivales, produciendo fenómenos de sincronización que contradicen en parte a lo que intuitivamente se esperaría en este tipo de actividades. Los estados de sincronización dependen del grado de atención que los participantes se conceden entre sí. Lo más importante en este proceso es que se realiza teniendo en cuenta a los vecinos más cercanos y no a toda la población implicada, lo cual conduce directamente al problema del pequeño mundo o "small world".

Stanley Milgram demostró a mediados del siglo 20 que, como máximo, se necesitan cerca de seis pasos para conectarse con cualquier persona en el planeta. Aunque parezca increíble y contra-intuitivo esto pone de manifiesto el alto grado de "clusterización" (clustering) o agrupamiento que gobierna al mundo social. La clusterización implica que, por supuesto, todos no conocen a todos, pero vinculando adecuadamente a los diferentes grupos en los que se mueven las personas pueden llegar a contactar a cualquiera.

Para formalizar la afirmación anterior se ha utilizado parcialmente el trabajo de Erdös y Ren-

yi, quienes inventaron la teoría de grafos aleatorios (random graphs). Uno de los descubrimientos más significativos de estos investigadores es que en un grafo aleatorio la conectividad aumenta drásticamente al incorporar más nodos al componente central de una red. Tal conectividad global, según estos autores, no se desarrolla incrementalmente y de manera regular, sino que crece incontrolablemente después que se ha sobrepasado determinado umbral. A pesar de la enorme utilidad de esta teoría para dar cuenta de la topología de muchas redes, Watts descubre que en el mundo real otras redes no tienen tal forma aleatoria en la que la distribución de la conectividad se ajusta a la figura de una campana de Gauss.

De manera que esta denominada "nueva ciencia de las redes" debe considerar también la dinámica o la evolución de las mismas, que no fueron incluidas por Erdös y Renyi. Las medidas tradicionales aplicadas al análisis de estas redes se despliegan alrededor del concepto de centralidad, pero son incapaces de ofrecer alguna explicación de la coordinación descentralizada, y, como inconveniente o insuficiencia adicional, tampoco ofrecen una consideración de las innovaciones que se producen en la periferia de estas formaciones.

Se puede decir, según lo mencionado en [70], que la dinámica del propio entramado en la red puede ser estudiada considerando otros enfoques. Se puede hablar de la dinámica *de* la red (cómo se desarrolla su estructura) o de la dinámica *sobre* la red, que es el producto generado por una red particular y más allá de su estructura. En las redes reales estas dos dinámicas están vinculadas estrechamente.

En el mundo de las cavernas los hombres vivían aislados y si compartían un amigo en común significaba que vivían en la misma comunidad. En un mundo opuesto, las relaciones previas no significan nada porque las posibilidades de conocer gente nueva y generar amistades no están determinadas por los amigos en común preexistentes. El espacio de posibilidades que se crea entre ambas opciones puede concebirse como un territorio de reglas de interacción factible de modelar. Si la mutualidad implica una elevada factibilidad de contraer relaciones estamos en el mundo de las cavernas, pero si no hay ninguna relación estamos en el mundo opuesto. Las redes de pequeño mundo (small-world networks) representan una instancia intermedia entre ambos extremos e inevitablemente existen en las situaciones en las que el promedio de longitud de pasos entre nodos es pequeño. En este estado de transiciones las regiones de la red están vinculadas por atajos (shortcuts) que vinculan a todos los nodos.

En 1997 una fraternidad universitaria inventó un desafío que sirvió para construir el llamado

"Juego de Kevin Bacon", que consiste en demostrar que este actor es el "centro" del universo fílmico a través de un simple proceso de mapeo o trazado de la conectividad del mundo de Hollywood. El resultado de este experimento es contundente: en un mundo formado por cientos de miles de individuos, cada actor está conectado con cualquier otro en un promedio de 4 pasos.

Watts y Steve Strogatz, a partir de estos indicios, procuraron buscar otras redes con características similares y lograron llegar a conclusiones idénticas. Tanto en fenómenos biológicos como en infraestructuras artificiales, y hasta en el mismo universo de Internet, las propiedades de las redes de pequeños mundos parecen estar presentes.

Siguiendo a la ley de Potencia formulada por Wilfredo Pareto en el siglo XIX, las redes que tienen la forma y las particularidades de las distribuciones libres de escala se diferencian enormemente de aquellos fenómenos que pueden ser descritos con el desarrollo de una curva normal. En primer lugar, las distribuciones de este tipo no tienen un pico que corresponda con su valor promedio y, en segunda instancia, comienzan con un valor máximo y decrecen lentamente hacia el infinito. Watts presenta la distribución de alturas de una población como caso típico que toma la forma de una curva normal y, como situación antagónica, la relación que puede establecerse entre la población de una ciudad grande y un pequeño poblado. Esas diferencias extremas son inconcebibles en una distribución de Poisson, pero se encuentran en muchos más fenómenos que los previstos por Watts y Strogatz. Al parecer la estructura física y virtual de Internet, por ejemplo, tiene esta forma, y lo mismo puede decirse de las redes metabólicas de muchos organismos biológicos.

Sin embargo, los descubrimientos de estos investigadores van más allá de esta corroboración, ya que postularon además un mecanismo por el cual estas redes pueden evolucionar en el tiempo. Para las redes que se ajustan a una distribución inequitativa de cualquier recurso la pregunta principal apuntaría a descubrir los motivos probables de estos resultados. Los autores apelaron al "efecto San Mateo" traído al ámbito de las ciencias sociales por Robert K. Merton. Este fenómeno, sustentado en el pasaje bíblico que plantea que *todo lo que tenemos en muestra vida terrenal nos será dado en abundancia en el cielo y todo aquello que poseemos en carácter de escaso nos será substraído aún más*, parece ser exactamente el tipo de proceso que rige la ampliación de las diferencias de cualquier tipo que los nodos de distintas redes parecen mostrar a largo plazo en las distribuciones libres de escala. Aplicando este principio al ámbito estrictamente económico, aquellos individuos que poseen riqueza se vuelven aún más ricos con sorprendente facilidad, y

aquellos que permanecen en la pobreza no solo no pueden salir de su condición aunque lo intenten, sino que es probable que se vuelvan aún más pobres en el futuro.

Los autores Bárabási y Albert tuvieron el acierto de postular el mecanismo por el cual este tipo de redes pueden configurarse como libres de escala de acuerdo a un principio no igualitario de acumulación de relaciones. Si se piensa en redes sociales en las cuales desde el inicio se cuenta con una cantidad diferencial de relaciones y tratamos de establecer una pauta de crecimiento realista de los contactos, se puede descubrir que aquellos nodos que se agreguen a la red en cada momento particular, con una alta probabilidad, lo hacen tratando de conectarse con aquellos nodos que están mejor relacionados. Este fenómeno es denominado por Bárabási y Albert como "enlace preferencial" (preferential attachment).

Herbert Simon, inventor del concepto de "racionalidad con límites" (bounded rationality) utilizó un modelo muy similar para explicar el tamaño desigual de las firmas comerciales, y George Kingsley Zipf postuló la famosa ley que lleva su apellido al describir la frecuencia desigual con la que aparecen las palabras en el idioma inglés. Estas distribuciones inequitativas tienen el sello exacto de las redes libres de escala.

En general, y cuando la gente no tiene suficiente información sobre un fenómeno, confía en la decisión de sus vecinos, y hasta es posible que actúe como ellos aún en contra de sus propias percepciones sobre la conveniencia de asumir o no determinada conducta. En rigor, lo que los economistas llaman "externalidades" no son elementos extraordinarios de la toma de decisiones colectivas e individuales sino sus componentes permanentes e imprescindibles. Este "contagio" de ideas responde a normas muy distintas a las del contagio de enfermedades. En el último caso los eventos de difusión de la enfermedad ocurren independientemente y la sumatoria de situaciones de contagio no implica un aumento en la posibilidad de contraer una dolencia determinada. En el caso de la difusión de ideas el dispositivo es distinto porque la factibilidad del contagio es directamente proporcional a la cantidad de vecinos que sean "portadores" de esa misma noción. El contagio social parece ser, en varios sentidos, un proceso mucho más dependiente de la vecindad de contactos que la expansión de enfermedades. Watts desarrolla, a partir de esta diferencia básica, el modelo de difusión de innovaciones introducido en la década del 60 por el investigador Everett Rogers, quien establece distinciones entre los umbrales variados de resistencia al cambio que los diversos nodos tienen en una red. Así Rogers distingue entre innovadores (innovators), adoptadores tempranos (early adopters) y nodos estables (stables) de acuerdo a la factibilidad

de difusión del contagio que cada nodo ofrece en una red.

Watts postula que los entramados sociales deben tener capacidad de lograr la cooperación sostenible a través de la retención de la conectividad y la capacidad de resolver problemas localmente y más allá de una coordinación centralizada, para lo cual se torna necesario un tratamiento efectivo de la ambigüedad en la información: ante descripciones del problema muy generales y hasta factibles de error, es la red quien tiene que proveerse de dispositivos de desambiguación y especificación provistos por sus propios recursos.

Las distancias sociales entre personas alejadas en la red son tremendamente cortas porque los lazos no se establecen al azar y los clusters o regiones pueden ser navegados a través de unos pocos atajos. Esta "navegabilidad" extrema, a pesar de los años ya transcurridos desde el primer experimento de Milgram, es un hallazgo experimental y teórico bastante opuesto al sentido común predominante. En las redes de pequeño mundo los eventos aparentemente insignificantes pueden generar acontecimientos incontrolables, pero también los eventos de alto poder disruptivo pueden ser tolerados por las redes si en su estructura cuenta con propiedades regeneradoras inferibles de antemano.

Aquí confluyen las teorías de los grafos aleatorios de los matemáticos, la de los modelos de percolación de la química y la de las transiciones de fase de los físicos. No todos los procesos se pueden estudiar con las mismas herramientas ni con los mismos procedimientos de formalización, pero todos son susceptibles de afrontarse con esta forma de interdisciplinariedad. No es sólo un conjunto de aplicaciones más o menos heterogéneas de un mismo conjunto de ideas, sino una auténtica y novedosa disciplina provista de dispositivos y técnicas de validación con un fuerte apoyo teórico y un buen pronóstico en el futuro.

3.3. La cultura carcelaria como modelo de ambientes restrictivos

Tal como se establece en [44], las cárceles son lugares típicamente oscuros, incluso durante el día cuando están cubiertas por los rayos solares. Son oscuras porque se vive en un mundo desconocido para la sociedad civil, lo que permite que se den tratos vejatorios y degradantes, debido en gran medida al aislamiento que facilita la formación de grupos o sectas de control en

los que frecuentemente se desarrollan negocios ilícitos. Las cárceles no sólo están ocultas para la sociedad civil, sino también están ocultas hacia el interior de ellas mismas, dado que las autoridades carcelarias desconocen cómo es la vida cotidiana de los internos, debido principalmente a la falta de interés que se tiene por conocer y participar en la vida de los presos, ya que su trabajo es más el de "vigilar y castigar", que el de re-educar, por ende la readaptación social no es un objetivo real.

Esta falta de conocimiento de la vida de los internos también se debe a la diferencia que existe entre los dos grupos que conforman la vida en el interior de estas instituciones: la sociedad de internos y la sociedad administrativa o del personal; donde se procura hacer patente la diferenciación principalmente por parte del personal, al grado que es común que los empleados de mantenimiento utilicen el uniforme del personal de seguridad, aun no perteneciendo a esta área, con el fin de hacer notar esta diferencia, lo que tiene como propósito que el uniforme les brinde seguridad y hacer notar la adscripción al grupo al que pertenecen.

El uso de la diferenciación conlleva el ser identificado como miembro del grupo de internos, pues se categoriza a las personas por medio de la atribución de características, supuestamente propias del grupo al que queda adscrito y que en el caso de los internos de una cárcel, sería el recibir la calificación de delincuentes, malvivientes, que son características profundamente desacreditadoras que los hacen indignos de confianza; atributos que a su vez afirman, por comparación, otra serie de características para quienes no pertenecen al grupo, dejando a los empleados dentro de una supuesta normalidad no deshonrosa, y es una referencia que, en términos éticos, les da ventaja respecto al grupo de internos y justifica, desde su perspectiva, el trato discriminatorio que se da dentro de este tipo de instituciones.

Los internos son un grupo marginal, dado que se es marginal en la medida en que se está al margen de la participación en determinadas áreas de la vida social, ya sea por decisión propia o por exclusión, al no formar parte de los roles que podrían desempeñar. El ser marginal también incide en no coincidir con los criterios de valor que rigen la homogeneización de la conducta de los seres humanos, dentro de la cultura institucional. Pudiendo ser marginal por decisión propia al ser producto del sostenimiento de los criterios a partir de los cuales se juzga la práctica social; o puede ser marginal, también, como resultado de la falta de participación de grupos o individuos en los roles que les corresponderían, de acuerdo a determinados criterios de la organización social. Esta falta de participación en ocasiones es forzada, como el caso de la reclusión que deja a los que

se encuentran en esta situación compartiendo las interpretaciones de realidad propias de estos grupos marginales, que recrean prácticas relacionadas a sus condiciones de existencia.

En el caso de la subcultura carcelaria, se encuentra la segregación de individuos, que debido a sus faltas, se encuentran en condición de marginalidad; lo que supone a su vez la existencia de formas de pertenencia desarrolladas como grupo marginal, que se explican en la medida en que se apoyan en la subcultura carcelaria, que se crea a partir de condiciones particulares de existencia. Estas formas de pertenencia generan a su vez mecanismos de autosostenimiento, que refuerzan algunas actitudes y algunos comportamientos. Así las situaciones de trato de inferioridad y aun el sentido de inferioridad con respecto al personal y la privación de derechos ayudan al mantenimiento de la marginalidad, en la medida en que cohesiona al grupo social.

La marginación está asociada a la falta de recursos y medios, es por esto que dentro de estos grupos sociales se obtienen los medios que permiten la subsistencia, la que logran por medio de la baja productividad y del aprovechamiento de los desperdicios de la civilización industrial. Así, los grupos marginales mantienen patrones de conducta que les impiden cumplir adecuadamente con los roles que les debían corresponder dentro de las estructuras institucionales.

En las cárceles se vive dentro de dos normativas: la de los internos y la de la parte administrativa, y el personal teniendo conocimiento de esto procura no entrometerse en los asuntos de los internos, a menos que sea por interés de la seguridad y de la institución. En el interior de las cárceles existen reglamentos, criterios y normas que rigen la conducta de los internos y del personal que ahí labora y que forman parte de la cultura dominante, y que es parte de la cultura homogeneizadora del comportamiento social: la cultura institucional; junto a esto existe también una cultura que subsiste al margen de lo permitido, y que se rige a partir de "contratos sociales informulados", la denominada *subcultura carcelaria*.

Esta subcultura carcelaria tiene sus propias reglas, que son válidas para los internos, y junto con la cultura institucional, regulan la forma cómo interaccionan los reclusos, permitiendo la integración o exclusión a los grupos. Exclusiones que podrían ser muy peligrosas en el interior de una cárcel.

Los contratos sociales informulados se apoyan en los intereses compartidos, que sólo se cumplen por la colaboración de los que comparten el espacio, y es a partir de estos acuerdos que vemos claramente estas normas, las que tienen un fundamento ético y práctico: ético en la medida

en que se fundan en un compromiso grupal de cohesión, al establecer obligaciones normativas entre los miembros; y práctico, al ser un tipo de estrategia adaptativa basada en la búsqueda de la seguridad.

El respeto a estos contratos sociales informulados facilita una interacción menos conflictiva entre los internos; son precisamente los conflictos y las tensiones las que nos revelan la importancia de este tipo de acuerdos normativos de la subcultura carcelaria, al atenuar en cierta medida las fricciones constantes que hay en estas instituciones. El no cumplir con este tipo de contratos puede ser muy grave, como sería el que un interno denunciara ante las autoridades de la institución alguna infracción cometida por otro interno. El denunciar le es permitido a un empleado, pero no a un interno, el que no sólo sería excluido del grupo, sino que además podría ser objeto de agresiones y violencia por parte de otros presos, pudiéndole costar hasta la vida.

Entre los internos se crea la conciencia de un "nosotros" que en principio se da a partir de la existencia de los dos grupos presentes en todas las instituciones; lo que crea una conciencia de grupo que no garantiza la cohesión en el caso de los internos. En general, dentro de las instituciones totales la lealtad que debiera generar la conciencia de grupo es algo poco común, y lo que priva sobre la lealtad es el egoísmo, sobre todo dentro de una cárcel donde la carencia es la característica predominante, de hecho uno de los recursos que los internos utilizan para no meterse en problemas con los demás internos, es tener claro que dentro de la prisión se está solo, que dentro de la cárcel cada quien vela por sí mismo y lo prioritario es salir de la mejor manera posible; lo que no se contrapone con la idea de un "nosotros" que se basa en la diferenciación que resulta muy aparente entre internos y el personal. Sin embargo la aspiración a que exista dicha fidelidad es una constante dentro de la subcultura carcelaria, sólo que está siendo continuamente quebrantada en la práctica, aunque reiteradamente se haga mención de un "nosotros" y un "ellos".

Una de las causas que hacen que se quebrante esta fidelidad está relacionada con el sistema de privilegios que existe dentro de los presidios, que hace que se fracture la homogeneidad que en principio se da en estas instituciones, ya que en estos lugares se da un trato masivo y por lo tanto despersonalizado a los internos; pero en un lugar donde la carencia es la norma, la más mínima comodidad resulta en un bien muy apreciado; privilegios que no le serían substraídos estando en libertad, pero dentro de estas instituciones, la posibilidad de hablar por teléfono, ver una película, tener acceso a comprar fruta, comer limpia y nutritivamente, recibir visitas de

familiares, tener un espacio privado como dormitorio, etc., resultan ser un privilegio.

Sin embargo, la constante referencia de un "nosotros" y un "ellos" nos muestra la existencia de una serie de características o atributos que comparte el grupo. Estas características se adquieren en principio por el hecho de su adscripción como preso, se refrende o no, al compartir o no estos elementos culturales propios de la subcultura carcelaria. La ubicación otorgada en principio por la adscripción que se tiene en la institución se obtiene en el mismo grupo, si no se toma la "conversión" como estrategia adaptativa, pasando a ser parte del otro grupo, como es el caso de los soplones, delatores, que nos muestra la no observancia de estos contratos sociales informulados, que son la base de la solidaridad grupal y la norma que rige, en gran medida, el comportamiento de los internos.

Es a partir de estos contratos sociales informulados que podemos observar la existencia de la subcultura carcelaria, que tiene como propósito la consecución de metas, en la medida que resuelve problemas específicos, mediante la observación de reglas de comportamiento, que inciden sobre las decisiones que se toman. La subcultura carcelaria se expresa como una conciencia grupal manteniendo los elementos culturales de la identidad colectiva; y se explica al observar la territorialidad como fuente generadora de conflictos. Es por esto que el ver, oír y callar es característica de la práctica de los hombres que ven constantemente expuesta su seguridad en el interior de las cárceles; el no participar en los conflictos que se generan en estos espacios les asegura una estancia menos peligrosa; se trata de pasar el menor tiempo posible y de la mejor manera, en el encierro.

Los códigos de la subcultura carcelaria, los que están basados en contratos sociales informulados, se pueden resumir en no denunciar; no inmiscuirse en asuntos ajenos, mostrar valentía en un momento determinado y en el carácter de cada persona, que está presente en la interacción que se da entre los internos.

Siendo la cárcel un espacio donde abunda la carencia, los internos que roban dentro de los penales son muy mal vistos e incluso corren gran peligro si son sorprendidos en el robo, esto debido también a que quien no se defienda de los robos y los abusos de otros internos muestra falta de valor, de "hombría"; es por esto que es muy común que las personas por medio de la violencia hagan respetar sus propiedades y derechos, delimitando de esta manera su territorialidad, dándose frecuentemente acuerdos tácitos sobre la pertenencia de los espacios.

Así la subcultura carcelaria se explica desde la adaptación al espacio y el medio en el que viven los internos.

3.4. Procesos Markovianos

Tal como se presenta en [4, 66] una cadena de Márkov, que recibe su nombre del matemático ruso Andréi Márkov, es una serie de eventos, en la cual la probabilidad de que ocurra un evento depende del evento inmediato anterior. En efecto, las cadenas de este tipo tienen memoria. "Recuerdan" el último evento y esto condiciona las posibilidades de los eventos futuros. Esta dependencia del evento anterior distingue a las cadenas de Márkov de las series de eventos independientes, como tirar una moneda al aire o un dado.

Este tipo de proceso, introducido por Márkov en un artículo publicado en 1907, presenta una forma de dependencia simple, pero muy útil en muchos modelos, entre las variables aleatorias que forman un proceso estocástico. En los negocios, las cadenas de Márkov se han utilizado para analizar los patrones de compra de los deudores morosos, para planear las necesidades de personal y para analizar el reemplazo de equipo.

En matemáticas, se define como un proceso estocástico discreto que cumple con la Propiedad de Márkov, es decir, si se conoce la historia del sistema hasta su instante actual, su estado presente resume toda la información relevante para describir en probabilidad su estado futuro.

Una cadena de Márkov es una secuencia $X_1, X_2, X_3, ...$ de variables aleatorias. El rango de estas variables, es llamado espacio de estados, el valor de X_n es el estado del proceso en el tiempo n. Si la distribución de probabilidad condicional de X_{n+1} en estados pasados es una función de X_n por sí sola, entonces:

$$P(X_{n+1} = x_{n+1}|X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_2 = x_2, X_1 = x_1) = P(X_{n+1} = x_{n+1}|X_n = x_n).$$

Donde x_i es el estado del proceso en el instante i. La identidad mostrada es la Propiedad de Márkov.

Las cadenas de Márkov en tiempo continuo, en lugar de considerar una secuencia discreta $X_1, X_2, ..., X_i, ..$ con i indexado en el conjunto $\mathbb{N}$ de números naturales, consideran las variables aleatorias X_t con t que varía en un intervalo continuo del conjunto $\mathbb{R}$ de números reales, con lo

cual se tiene una cadena en tiempo continuo. Para este tipo de cadenas en tiempo continuo la Propiedad de Márkov se expresa de la siguiente manera:

$$\begin{aligned} P(X(t_{n+1}) &= P(X(t_{n+1}|X(t_n) = x_n, \ldots, X(t_1) = x_1) \\ P(X(t_{n+1}) &= P(X(t_{n+1}) = x_{n+1}|X(t_n) = x_n) \text{ tal que } t_{n+1} > t_n > t_{n-1} > \ldots > t_1 \end{aligned}$$

3.5. Optimización multiobjetivo

Tal como se expresa en [65, 59], en un problema de optimización se trata de encontrar una solución que represente el valor óptimo (un máximo o un mínimo) para una función objetivo.

Para el caso más simple se tiene un único objetivo, que está representado por una función del tipo $f : M \rightarrow N$, donde $M \subset \mathbb{R}$ y $N \subset \mathbb{R}$. Tanto el dominio como la imagen de la función son números reales (escalares), y el valor óptimo corresponde a un valor mínimo o a un valor máximo.

Pero en las aplicaciones prácticas en ciencias y en ingeniería se presentan, con cierta regularidad, problemas que requieren la optimización simultánea de más de un objetivo, y es lo que se conoce como optimización multiobjetivo. Se optimiza por tanto una función de la forma $f : S \rightarrow T$, donde $S \subset \mathbb{R}^n$ y $T \subset \mathbb{R}^k$. Pero el problema está en que normalmente no existe un elemento de S que produzca un óptimo de forma simultánea para cada uno de los k objetivos que componen f. Esto se debe a la existencia de conflictos entre estos objetivos, que hacen que la mejora de uno de ellos dé lugar a un desmejora de algún otro. Se tiene que llegar, por lo tanto, a una situación de compromiso en la que todos los objetivos se satisfagan en un grado aceptable, desde el punto de vista de diseño.

A diferencia de los problemas de optimización con un único objetivo, el concepto de óptimo en el caso multiobjetivo es relativo y es necesario decidir de alguna forma cuál es la mejor solución (o cuáles son las mejores soluciones) al problema.

En términos matemáticos, el problema de optimización multiobjetivo, puede establecerse de la siguiente forma:

Encontrar un vector $x^* = [x_1^*, x_2^*, \cdots, x_n^*]^T$, que satisfaga las m restricciones:

$$g_i(x) \geq 0 \qquad i = 1, 2, \cdots, m \tag{3.1}$$

y las p restricciones:

$$h_i(x) = 0 \qquad i = 1, 2, \cdots, p \tag{3.2}$$

y optimice la función vectorial

$$f(x) = [f_1(x), f_2(x), \cdots, f_k(x)]^T \tag{3.3}$$

donde $x = [x_1, x_2, \cdots, x_n]^T$ es el vector de variables de decisión.

En otras palabras, se desea determinar la solución particular $x_1^*, x_2^*, \cdots, x_n^*$, del conjunto S formado por todos los valores que satisfacen a 3.1 y a 3.2, que de lugar a los valores óptimos para todas las funciones objetivo. Pero como ya se ha comentado, no existe normalmente una solución que optimice de forma simultánea todas las funciones objetivo.

Para tratar el problema comentado del conflicto entre objetivos se han utilizado diversos métodos:

* Métodos basados en el concepto de eficiencia de Pareto.

* Métodos basados en la combinación de objetivos. Dentro de estos métodos se puede mencionar el *método de la suma ponderada*, en el que se optimiza el valor obtenido mediante la suma de los valores correspondientes a los distintos objetivos, multiplicados cada uno por un coeficiente de peso. Estos coeficientes de peso establecen la importancia relativa de cada objetivo. El problema de optimización multiobjetivo se transforma así en otro problema de optimización escalar, que para el caso de la minimización se puede expresar de la siguiente forma:

$$\text{mín} \sum_{i=1}^{k} w_i f_i(x)$$

donde $w_i \geq 0$ es el coeficiente de peso correspondiente al objetivo i.

Existen variantes del método anterior, como el método de la programación por metas, en el que se establece una meta para cada objetivo y lo que se suma (multiplicado por el correspon-

diente coeficiente) es la distancia de cada objetivo a su meta. Para un caso de minimización su representación es:

$$\text{mín} \sum_{i=1}^{k} w_i \left| f_i(x) - M_i \right|$$

donde M_i representa la meta del $i-$ésimo objetivo.

* Métodos basados en la asignación de prioridades. Estos métodos establecen prioridades entre los distintos objetivos, teniendo en cuenta su importancia relativa durante el proceso de optimización.

Los métodos anteriores han sido utilizados en varias vertientes de investigación, e incluso han sido utilizados en combinación con los algoritmos evolutivos (ver [15]), que han demostrado ser una herramienta adecuada para resolver este tipo de problemas. Estos métodos se conocen como MOEA (Multi-Objective Evolutionary Algorithms o algoritmos evolutivos multiobjetivo).

Capítulo 4

Trabajos Relacionados

En términos generales, cuando se hacen estudios sobre la implementación de sistemas en un ámbito global se deben considerar un número incontable de variables, que por lo general, en su gran mayoría son desconocidas por los implementadores. La intención de este trabajo no es la de considerar y estudiar todas las variables involucradas en la implementación de los sistemas anónimos en escenarios globales, sino más bien es la de establecer un precedente donde se acoten algunos puntos estratégicos a considerar en este tipo de procesos. En otras palabras, el objetivo fundamental de este trabajo es el de establecer algunos de los temas que pueden ser de interés relacionados con la implementación de sistemas anónimos a escala mundial, e intentar proponer respuestas a algunas interrogantes que han surgido relacionadas con éstos.

Específicamente, la orientación global dada se enmarca en considerar los siguientes aspectos: la computación ubicua como referencia para la implementación a gran escala, los modelos matemáticos (de optimización) necesarios para procurar mejorar la eficiencia de las comunicaciones a gran escala, la heterogeneidad de las distintas regiones en cuanto a su nivel de restricciones sobre las comunicaciones. Este último aspecto se trata con mayor énfasis dado que actualmente uno de los principales problemas en cuanto a la implementación y uso de este tipo de sistemas es el de hacer factible las comunicaciones anónimas en zonas (regiones, países, organizaciones, etc.) donde se regulan las comunicaciones, censurando y bloqueando el tráfico en las redes. Como posible solución a los entornos restrictivos, se propone el uso de los sistemas de reputación y confianza, inspirándose también en la subcultura carcelaria y aplicando la teoría de los pequeños

mundos.

En las siguientes secciones se muestran algunos de los trabajos relaciones con estos temas, y que de alguna manera establecen precedentes a esta propuesta.

4.1. Computación Ubicua

La computación ubicua se refiera a los ambientes donde la mayoría de los objetos físicos poseen componentes digitales con el fin de mejorar sus prestaciones. Esto quiere decir, que esos objetos involucran pequeños dispositivos informáticos que le incorporan con el fin de incrementar el beneficio que ofrecen.

La computación ubicua se puede interpretar como un caso de estudio para la implementación de sistemas a escala global, ya que aunque no se refieren directamente a sistemas de comunicación o informáticos, tales como los encontrados en los perfiles cliente-servidor, P2P, etc. son importantes de considerar en el aspecto de la privacidad. Sin embargo, a pesar de que su estudio no está en el objetivo inicial de este trabajo, se incluye a manera de referencia.

4.1.1. Escalas de privacidad para Computación Ubicua, RFID, servicios de localización, servicios ubicuos

En [1], la autora Sarah Spiekermann, propone tres escalas para medir la privacidad en la computación ubicua, estas escalas están basadas en el grado de control que posee la gente sobre la información personal. Específicamente este trabajo se concentra en el estudio de los sistemas basados en la identificación en las comunicaciones con radio-frecuencia (RFID). Esta tecnología consiste en introducir un circuito integrado o "chip" en los productos comerciales durante el proceso de su fabricación. Estos chips, a los que se les denomina *etiquetas* o *tags* emiten un número de producto único una vez que localizan o son localizados por un *lector*. Este lector le envía dicho número a un centro de información donde puede ser identificada la naturaleza del producto y potencialmente puede ser identificado su propietario.

La autora afirma que uno de los principales problemas que en la actualidad se deben enfrentar es el desconocimiento o el mal conocimiento que se tiene sobre la privacidad en este tipo de

ambientes. Afirma, que en este escenario, la privacidad está íntimamente relacionada con el control que el usuario tiene o puede tener sobre la información que se proporciona a través del RFID, es decir, definen a la privacidad como el control que tiene un individuo sobre sus datos y sobre sí mismo. Las escalas propuestas se relacionan con la *percepción del control* que las autoridades suponen tienen los usuarios, y se establecen tres enfoques para medir los niveles de control: el *nivel de impotencia percibido*, relacionado a la percepción de *no tener el control*. El segundo se denomina *el control de la información*, y es la sensación de poder sobre la información, y la última escala propone medir la privacidad como un nivel de competencia de la gente ante el control de su información, a lo que le denominó la escala de *fácil uso*.

En [1], los autores P. Najera y J. López, estudian la privacidad en la computación ubicua para los sistemas RFID. Plantean los problemas relacionados con la fuga de información que se puede presentar en el uso de este tipo de mecanismos, y proponen algunas soluciones para poder controlar los niveles de privacidad en el uso de los mismos. Los autores mencionan que las etiquetas RFID pueden estar construidas con tecnologías muy variadas, que van desde el uso de dispositivos "pasivos" los cuales no contienen fuentes de energía propias y dependen de la energía de los lectores para poder suministrar la información, hasta los dispositivos "activos", que al tener fuentes de energía propias, pueden generar acciones como las de actualización y modificación de la información contenida en ellos. En cualquiera de los casos, ya sean dispositivos activos o pasivos, es importante saber que cada vez que se encuentren bajo la influencia de un lector, les suministrarán la información relativa al producto, y potencialmente sobre la identidad de su dueño, a un centro de información. Es precisamente aquí donde el tema de la privacidad entra a jugar un papel importante en esta área, ya que si el usuario desconoce cómo, cuándo o dónde lo están identificando, se estarán violando sus derechos de privacidad. Para este punto los autores proponen utilizar herramientas que ayuden a mantener el control sobre la distribución de la información de estos mecanismos, las cuales van desde hacer conscientes a los usuarios sobre su existencia y sus implicaciones sobre la privacidad, hasta la de proponer métodos que le permitan al usuario tener el control en el envío de la información, por ejemplo, proponen el uso de las *cajas de Faraday* como aislantes eléctricos, para evitar que las etiquetas puedan "comunicarse" con los lectores sin autorización. Otra alternativa, es obligar, en la medida de lo posible, a que los instaladores de lectores coloquen avisos que le adviertan a los usuarios sobre su presencia.

Por otro lado, pero también con un matiz ubicuo, se encuentran los servicios de localización y la información que aportan, los cuales abanderizados por los servicios prestados por los Sistemas

de Posicionamiento Global o "GPS" por sus siglas en inglés, tienen la capacidad de revelar información sobre la ubicación de las personas, incluso sin su consentimiento. En este aspecto se deben involucrar a los responsables de la prestación del servicio y conseguir soluciones que puedan permitir que la información de cada usuario que desea mantener de forma privada se mantenga así. Se han propuesto soluciones que incluyen el uso de seudónimos, y el uso de técnicas de ofuscación que pretenden evitar que la información obtenida por terceros sea legible.

En general, los servicios que presta la computación ubicua tienen actualmente muchas interrogantes sin responder sobre el tema de la privacidad. Por ejemplo, si se hace referencia al caso médico a nivel de hogares, donde las personas reciben un tratamiento en sus propias casas, y están controladas por dispositivos que transmiten información hacia los centros hospitalarios o médicos, con el fin de llevar un registro y mantener informados a los profesionales del área médica sobre la evolución de la salud de cada paciente, es obvio darse cuenta que el escape de información privada puede ser mucho más crítico. Se deben aun hacer muchas consideraciones para poder llegar a un acuerdo entre el envío de la información necesaria al personal de la salud y el mantenimiento de la privacidad. En este caso se deben considerar aspecto sociales, tecnológicos, de seguridad, etc. ya que no pueden ignorarse los problemas potenciales que pueden presentarse cuando se confía en terceros (enfermeras, médicos, dispositivos de comunicación, etc.).

4.2. Sistemas resistentes al bloqueo y a la censura

En la mayoría de los entornos restrictivos utilizan dispositivos, como los corta-fuego o "firewalls", para lograr establecer y ejecutar sus políticas de censura y bloqueo sobre las comunicaciones. Estos dispositivos trabajan con diseños de mecanismos y estructuras específicas, que en mayor o menor medida filtran, bloquean y monitorean el tráfico de información dependiendo del entorno donde se encuentren. Es decir, dado que todos los entornos de comunicaciones tienen características específicas que los diferencias de los demás, algunas de políticas de filtrado, bloqueo y censura pueden tener éxito sobre unos entornos y no sobre otros.

Como ejemplo más representativo se encuentra el Firewall de China, que es un país con ideologías políticas, económicas, civiles, etc. particulares, donde acostumbran a monitorear y restringir las comunicaciones que consideren están en contra de sus ideas. Dado lo anterior, investigaciones recientes orientadas a conseguir evitar este tipo de bloqueo para permitirle a los

ciudadanos chinos disfrutar de su derecho a una libre comunicación, han estudiado el diseño y estructura de este dispositivo en particular, y tal como se menciona en [14], descubrieron que principalmente lo que hace este dispositivo es inspeccionar los paquetes TCP en búsqueda de palabras claves para generar el bloqueo de cada comunicación de forma particular. Si se detecta que una de las palabras clave se encuentra en los paquetes de transmisión, el "firewall" le envía a los dos puntos comunicantes *paquetes de reinicio de comunicación* (configura el campo RST del paquete marcándolo como activo), y esto hace que finalice la comunicación. Sin embargo, si se configuran a los nodos comunicantes (emisor-receptor) para que ignoren los paquetes RST enviados por el firewall, entonces podrían seguir comunicándose. Con esto se demuestra que el problema real no está en el hecho de cómo sobrepasar o ignorar a estos dispositivos en momentos determinados, sino que los entornos restrictivos hacen que constantemente evolucionen sus políticas de bloqueo, adaptándose con el fin de evitar las nuevas técnicas de evasión, lo que hace que los sistemas anti-censura deban ser reinventados constantemente.

En [41, 33, 35] estudian los aspectos relacionados con el diseño de sistemas anónimos resistentes al bloqueo. En cada caso muestran las ventajas y desventajas sobre los sus diseños, y en todos estos trabajos se concluye que el principal problema a enfrentar no tiene que ver con la evasión del bloqueo en sí, sino con la implantación, conexión y sostenibilidad de las comunicaciones anónimas. En este aspecto, comentan que la forma de conectarse a una red anónima es a través de puntos de acceso o nodos de salidas a los cuales han denominado *proxies* en unos casos, y *puentes* en otros, configurados en grandes o pequeños grupos y con servicios especializados de mensajería, o con diseños en forma de puentes (reenviadores de flujos). El problema crítico que tienen que enfrentar los usuarios que desean comunicarse anónimamente es el de poder descubrir cómo, cuándo y dónde conectarse, es decir, este problema está íntimamente relacionado con el proceso de la distribución de la información, pues es obvio que si la información se distribuye públicamente y ampliamente, los encargados de los sistemas restrictivos podrán obtener toda la información necesaria para bloquear por completo dichas conexiones.

Particularmente en [35] no se estudia una solución al problema del descubrimiento, pero en un trabajo más reciente (ver [36]) el mismo grupo de investigadores muestra una salida a dicho problema. Afirman que la solución está en evitar que se publiquen ampliamente los proxies (nodos de acceso) y éstos deberían se descubiertos en un proceso continúo, es decir, que mientras se están haciendo las comunicaciones también se desarrolla paralelamente el proceso de búsqueda en caso de que las conexiones actuales sean bloqueadas o fallen. Este proceso de búsqueda deben

hacerlo los mimos clientes ("in-band"). Proponen un mecanismo para el descubrimiento que denominaron "keyspace hopping", el cual está inspirado en la política de conexión de las redes inalámbricas donde los dispositivos pueden *saltar* entre distintos puntos de acceso a través de un sistema de descubrimiento continuo. Este tipo de descubrimiento sólo tiene efecto sobre una pequeña fracción del universo de puntos de acceso, y para el caso del descubrimiento en sistemas anónimos se hace sobre una pequeña fracción de proxies o puentes. Sin embargo al poder estar en presencia de innumerables proxies el sistema no podría garantizar la confianza que se puede tener sobre éstos, por tal razón proponen en su mecanismo separar el proxy en dos componentes: un mensajero a través del cual los clientes realizan la labor de descubrimiento aplicando el proceso de saltos ("keyspace hopping") y que actúa con una puerta de enlace ("gateway") hacia Internet, y un componente que denominan *portal* cuya identidad es ampliamente conocida y publicada y es el que se responsabiliza por las solicitudes de contenido censurado del cliente. De esta manera, cada proxy actuará como vía de enlace a un reducido número de clientes en un determinado momento, y en los siguientes períodos de tiempo es posible que atienda a otros clientes diferentes. El mecanismo propuesto no le permite a los clientes escoger un proxy en específico, así que éstos deben utilizar nodos en los cuales no confían para poder acceder a la zona no bloqueada, para lo cual los autores afirman que se soluciona con los mencionados mensajeros que no pueden establecer las diferencias entre una comunicación normal y una bloqueada. Sin embargo, en los ambientes reales, si se utiliza esta estructura de mensajeros no confiables, ya de por sí está indicándole a los entes reguladores que el cliente puede estar interesado en una comunicación no permitida, lo cual haría que llamase su atención sobre sus comunicaciones particulares, implicando una gran desventaja.

En resumen, tanto en [36] como en la propuesta presentada en [33] se menciona que para resolver el problema de la distribución de la información y del descubrimiento de los nodos de salida ya sean proxies o puentes se concluye, que como mejor práctica no se dé la información de todos los nodos de acceso simultáneamente, sino en pequeños grupos, y proponen mecanismos para poder mantener informados a los usuarios, tales como el uso de los servicios de mensajería instantánea, que puedan servir para compartir la información entre los distintos usuarios, de cómo podrían conectarse a la red no bloqueada. Se basan además en el hecho de que, incluso en un país como China, existe heterogeneidad en cuanto a los niveles de restricción en los distintos puntos de acceso a Internet, y de esta forma cada usuario podría servir de proxy o puente de otro usuario en otra zona donde no esté restringida su comunicación.

Sin embargo, en ninguno de estos trabajos proponen, de forma realista, cómo se establecería la conexión entre usuarios, cómo haría un usuario sin ningún conocimiento para poder establecer contacto con otro que le permita tener acceso a las redes no bloqueadas, cómo podría en realidad ser sostenible una red que constantemente necesita cambiar para poder sobrellevar los bloqueos, cómo podría dar respuesta al hecho de la existencia de "espías" que se hagan pasar por usuarios normales, con el fin de obtener información. Incluso considerando que sólo se dé información sobre pequeños grupos de nodos de salida, los espías podrían conformar grupos numerosos y dispersos, y actuar por largos períodos de tiempo, con lo cual podrían en gran medida comprometer toda la comunicación.

Una de las alternativas posibles, es utilizar los mecanismos propuestos en los sistemas de reputación y confianza, y por esto en las siguientes secciones se muestra parte del trabajo que han realizado sobre Internet (en un ámbito global) y en los sistemas anónimos en cuanto al uso de los mecanismos de reputación y confianza. Es importante mencionar, que en ningún caso han sido utilizados para resolver el problema del descubrimiento de nodos y el de la distribución de la información en ambientes restrictivos.

4.3. Sistemas de Reputación y Confianza

4.3.1. Sistemas de Reputación y Confianza a nivel global

En Internet se dispone de multitud de tecnologías que resuelven problemas concretos pero que no están integradas en un modelo global de seguridad. La red se utiliza para difundir elementos dañinos como los virus, lo que se hace con rapidez, y se le suma la forma creciente de otros comportamientos abusivos, como el envío masivo de correo electrónico con fines comerciales (spam) e incluso fraudulentos (phishing). Estos excesos en el envío de mensajes provocan el consumo innecesario de recursos y afectan la productividad de los usuarios finales, quienes, para solventar estos problemas, procuran conseguir e instalar herramientas localmente. La mayoría de las solicitudes de servicio se tratan por igual, sin más limitación que la propia capacidad de los equipos y enlaces de comunicaciones, sin tener en cuenta la reputación del interlocutor, aceptándose identidades que se usurpan fácilmente. En [52] se plantea cubrir el espacio entre la autenticación del cliente, que requiere despliegues costosos y no inmediatos, y la confianza ciega en cualquier usuario o equipo atribuida sólo al hecho de estar conectado a la red y "poseer"

una dirección. Cada servidor puede ser capaz de valorar la confianza que merecen sus clientes en función de las experiencias acumuladas directamente u obtenidas por terceros. Esa confianza dependerá de la fiabilidad con que se les pueda vincular a sus mensajes y su adecuada utilización del sistema, determinada por la exposición de un volumen suficiente de sus mensajes y solicitudes al análisis de los usuarios finales y de otras herramientas complementarias. En ese trabajo proponen un modelo dinámico y distribuido, descrito mediante la "lógica subjetiva" de Audum Jøsang, que permite representar los conceptos de confianza e incertidumbre en un valor t denominado opinión. La opinión, calculada en base al comportamiento reciente del cliente, resumen la confianza que merece, y se calculan las restricciones que se consideren pertinentes. Además se puede intercambiar con terceros, donde se constituyen formas de federaciones en las que se aplican políticas comunes de contención que acoten el consumo global de recursos. El álgebra asociada permite ponderar las opiniones ajenas y previas, y combinarlas mediante las operaciones denominadas recomendación y consenso bayesiano. Además se introducen en el modelo, también con este mecanismo, los resultados de herramientas externas como los filtros, listas negras, etc. Se propone su utilización como medida preventiva y complementaria en la guerra contra el spam en Internet, la priorización de la difusión de contramedidas o la prevención de ataques de denegación de servicio en web services, entre otras posibilidades.

4.3.2. Reputación en Sistemas Anónimos

Redes P2P

En [30] explican porqué los sistemas de reenvíos de correos y los sistemas para publicaciones anónimas pueden beneficiarse de la reputación, describen algunas experiencias en el uso de la reputación en los sistemas anónimos, y concluyen que es necesario un rediseño, desde el punto de vista conceptual, para que los sistemas anónimos puedan involucrar definitivamente la reputación como medio para hacer transacciones confiables.

En [29] se describe un diseño para un sistema de reputación en el cual se gestiona la confianza y la eficiencia en los servicios de reenvíos de correo electrónico. Se utiliza una red Mix en la cual cada Mix genera un "recibo" por cada recepción de mensajes. Estos recibos indican el resultado de cada una de ellas, es decir, indica si ha sido exitosa o no. Junto con la ayuda de un conjunto de testigos, estos recibos le permiten a los emisores verificar la correctitud de cada Mix y además

funciona como forma de prueba del mal o buen comportamiento ante los testigos y ante terceros.

En [2] presentan un esquema de reputación para un sistema punto a punto a través del uso de seudónimos en una red anónima. Se menciona que el uso abusivo o el mal comportamiento es uno de los más grandes problemas que se presenta en los esquemas de reputación de los sistemas punto a punto que utilizan seudónimos, debido a que existen muy pocos incentivos para tener un buen comportamiento. Este esquema utiliza *puntos* de reputación basado en los sistemas de dinero electrónico ("ecash"), y la reputación de cada usuario se relaciona más a la verdadera identidad y no a su seudónimo. Así, el esquema le permite a un usuario honesto utilizar varios seudónimos manteniendo su nivel de reputación, y le dificulta a un usuario malicioso la tarea de la eliminación de cualquier indicio que marque su mal comportamiento a través del uso de nuevos seudónimos. Para esto utilizan una entidad de control centralizada ("banco de reputación"), la cual maneja los niveles de confianza de los usuarios y genera los certificados pertinentes para la demostración de la reputación ante terceros. Cada vez que un usuario desea interactuar con otro utiliza un seudónimo por cada transacción que hace, es decir, cada punto o usuario utiliza dos niveles de seudónimos, el primero lo identifica como un punto o nodo en la red P2P, y el segundo nivel lo utiliza para identificar, con múltiples seudónimos, a cada una de las transacciones que hace. Cuando interactúa con el banco lo hace de dos formas, utilizando los dos niveles de seudónimos. El banco genera certificados "ciegos" de confianza inspirados en la idea de los billetes o monedas como formas de pago, que no están atados a la identidad de su portador. Para poder incrementar los niveles de reputación, cada usuario debe entregarle uno o varios puntos de reputación a otro usuario cuando lo crea conveniente (por ejemplo, después de haber ejecutado una interacción exitosa), estos puntos deben haber sido generados u otorgados por el banco previamente. Cuando un usuario le entrega a otro uno o varios de sus puntos de reputación, éste se los envía al banco para que los valide, y si lo considera adecuado, el banco le otorga un certificado ciego. Luego el usuario, utilizando su verdadera identidad (en este caso, su primer nivel de seudónimo) le entrega al banco este certificado ciego para que incremente su nivel de reputación. El banco también entrega certificados para demostrar los niveles de reputación que cada usuario tiene, esto sería el análogo a un estado de cuenta bancario.

Capítulo 5

Sistemas Anónimos Globales

En el diseño e implementación de sistemas anónimos, como primer paso, se deben considerar las necesidades específicas de comunicación que se presentan en cada caso en particular, tanto a nivel de infraestructura de red, como a nivel de programas, con el fin de establecer las características propias del sistema en cuanto a su estructura y políticas de transmisión. Por ejemplo, se debe decidir si se requiere un sistema de alta o baja latencia, si se requieren o no altos niveles de anonimato, si se dispone de una infraestructura centralizada o distribuida para el control de las comunicaciones, etc.

En el caso de los sistemas globalizados no existe una especificación concreta de necesidades debido a su heterogeneidad, encontrada en cada una de las partes del globo terrestre. En este sentido puede parecer una tarea infactible la de diseñar e implementar sistemas globales. Sin embargo, es posible establecer algunas pautas en común que puedan ser utilizadas como estructura de base (plantilla) y esquemas generales de trabajo que puedan ser aplicados adaptándolos a cada caso, por ejemplo, se pueden proponer varios tipos de mecanismos para el proceso de la distribución de la información, y varios mecanismos para el reenvío de mensajes, y hacer las combinaciones que se consideren pertinentes. Por esta razón el objetivo de este trabajo es proponer algunas ideas iniciales basadas en algunas características generales que han sido detectadas como patrones comunes en varios zonas del planeta. Estas ideas se engloban en tres necesidades generales:

- Implementar mecanismos que proporcionen mayores niveles de anonimato y que no incrementen considerablemente la latencia.
- Implementar mecanismos que contrarresten los sistemas de bloqueo y censura. Esto implica resolver el problema de la dinámica del descubrimiento de nodos iniciales con los cuales se pueda acceder a las redes anónimas.
- Establecer modelos para la evaluación del rendimiento y/o desempeño que permitan decidir cuál es la mejor combinación de mecanismos en cada uno los casos encontrados.

En los siguientes apartados se describen las soluciones propuestas para cada una de estas necesidades planteadas, la cuales se integran en un macro-diseño (macro-algoritmo). La idea para una implementación final es la de utilizar el modelo de optimización para evaluar un sistema concreto combinando las ideas propuestas con otras que puedan surgir o que ya se han presentado como soluciones parciales. Por ejemplo, se pueden combinar la ideas para solventar el problema en los entornos restrictivos, junto con algunos mecanismos de reenvío utilizados en otros modelos.

5.1. Diseño de un sistema anónimo markoviano

La necesidad aun no satisfecha de tener un sistema, o en su defecto, un mecanismo que pueda tener las características particulares de baja latencia y alto nivel de anonimato se ha presentado desde que surgió la idea de las comunicaciones anónimas. En otras palabras, no es trivial conseguir tal diseño. En este caso el objetivo fue el de encontrar un punto intermedio donde los niveles de latencia y anonimato sean aceptables para ciertos tipos de aplicaciones, lo que implica no tener niveles de anonimato bajos como los presentados en sistemas como Tor, y no poseer latencias elevadas como en los sistemas basados en redes mix.

Para esto se pensó que las técnicas utilizadas en las redes de mixes aún siguen presentando inconvenientes tal como están concebidas, pero permiten tener altos niveles de anonimato, incluso suponiendo un atacante con cobertura global, y con largos períodos de tiempo de observación.

Una posible solución está en conseguir un mecanismo de reenvío que no requiera los tiempos de los mixes por grupos o por cantidades, pero que no sea tan obvio establecer la relación de entrada salida de los mensajes. De tal forma, que la idea es acumular por cortos períodos de

tiempos algunos mensajes y luego reenviarlos a su siguiente destino en un orden distinto al de su llegada, lo que puede ser compatible con un mecanismo del tipo *cebolla*, es decir, se puede utilizar un mecanismo tipo cebolla, como el descrito para Tor y otros sistemas de OR (Onion Routing), pero que en cada nodo se aplique una política de reenvío distinta a la FIFO (primero en llegar, primero en salir). Esta política puede ser implementada siguiendo un proceso markoviano de la siguiente manera:

Al llegar un mensaje a un nodo intermedio, éste lo almacena en su memoria temporal. El envío del mensaje se determina probabilísticamente dependiendo de sus estados anteriores, por ejemplo, si el mensaje llega en el tiempo t, la probabilidad de envío en el tiempo $t+1$ será inferior a la probabilidad de envío en el tiempo $t+2$, y así sucesivamente en cada uno de los momentos de decisión se consideran los estados anteriores del mensaje, lo que se puede representar a través de una cadena de Markov. Al incrementar la probabilidad de envío cada vez que transcurre un instante de tiempo, se busca disminuir su tiempo de permanencia en el nodo.

Como parámetro de implementación se deberá decir cuál es la probabilidad de base, por ejemplo, para el envío en el tiempo t_0, momento en el cual el mensaje fue recibido. Luego, se incrementa esa probabilidad a una tasa cuyo valor se debe establecer como otro parámetro de entrada en la implementación. En otras palabras, para cada implementación se debe decidir cuál es la probabilidad inicial de envío, y cuál es la tasa de incremento para cada instante de tiempo transcurrido.

La probabilidad de envío se puede calcular recursivamente de la siguiente manera:

$$\begin{aligned} P(X(t_0) &= P(X(t_0 = x_0) \\ P(X(t_n) &= P(X(t_n)|X(t_{n-1}) = x_{n-1}) \text{ tal que } t_n > t_{n-1} > t_{n-2} > \ldots > t_0 \\ P(X(t_n) &= P(X(t_{n-1}) = x_{n-1}) * k \end{aligned}$$

Donde x_i determina la decisión de envío (0 ó 1). La probabilidad para x_0 y el valor de k se establecen como parámetros de entrada.

5.2. Modelo para la comunicación anónima en entornos restrictivos

Cómo solventar el problema relacionado a que un usuario en un entorno restringido pueda descubrir dinámicamente al menos uno de los nodos por medio del cual se pueda conectar con al menos una red anónima, es una pregunta aun sin respuestas satisfactorias a nivel de las implementaciones prácticas. Si un usuario dentro de un entorno censurado y bloqueado desea comunicarse anónimamente lo primero que debe hacer es conseguir la aplicación que le permita tener acceso a la red. Después, necesita saber con cuáles nodos se podría conectar. Saber cuáles son los nodos de conexión implica tener un mecanismo de descubrimiento eficiente, en la cual se le informe eficientemente al usuario sobre los nodos disponibles, y que evite en la mayor medida posible que éstos sean bloqueados por los entes reguladores, es decir, los entes reguladores, al monitorear la red, pueden darse cuenta de los nodos que están funcionando como proxies o puentes de salida y pueden así bloquearlos. Es por eso, que el proceso de descubrimiento debe desarrollarse dentro de un conjunto constante de cambios, ya que si ciertos nodos de salida son bloqueados, otros deberían estar disponibles y al alcance de los usuarios para poder darle continuidad a sus comunicaciones anónimas. En otras palabras, si se bloquean algunos nodos de salida, el mecanismo le debe proporcionar al usuario un listado de nuevos nodos de salida. Por lo anterior se deben considerar las siguientes premisas:

- No hay infinitos nodos de conexión en cada momento.
- Los nodos disponibles en un determinado momento, pueden ser bloqueados o no estar disponibles en los momentos posteriores.
- Los entes reguladores se pueden hacer pasar por usuarios normales, a través de los cuales pueden obtener información de la red, y hacer los bloqueos según la información obtenida (espías).

Dada estas premisas, en los siguientes apartados se muestra una forma de solventar el problema del descubrimiento de nodos, y la sostenibilidad del proceso en el tiempo. En la siguiente sección se propone combinar las ideas de los sistemas de reputación y confianza, junto con la teoría de los pequeños mundos, y las características de la subcultura carcelaria. En el apartado posterior se sintetizan todas estas ideas en un macro algoritmo.

5.2.1. Integración: Reputación, Pequeños Mundos y Subcultura carcelaria

Cada usuario debe conseguir una versión de la aplicación para la conexión a través de su entorno cercano (amigos, familiares), en quienes confíe, este tipo de confianza de primer nivel o implícita es necesaria para evitar que el programa sea remitido por uno de los espías o adversarios. Si en su entorno no los consigue, puede utilizar el entorno en un segundo nivel: los amigos y familiares de sus amigos y familiares, y así sucesivamente hasta alcanzar la obtención de la aplicación. Este proceso de obtención del "software" puede hacerse de forma física y directa, o por un envío personal cifrado (sftp, email con PGP, etc.) ya que este tipo de comunicación no está restringido desde un primer momento. Tal como se muestra en la figura 5.1, la estructura de solicitud y obtención tendrá forma de árbol, donde la comunicación la hace el usuario en su entorno social más cercano. Para esta primera fase los niveles de reputación y confianza están implícitos en cada usuario.

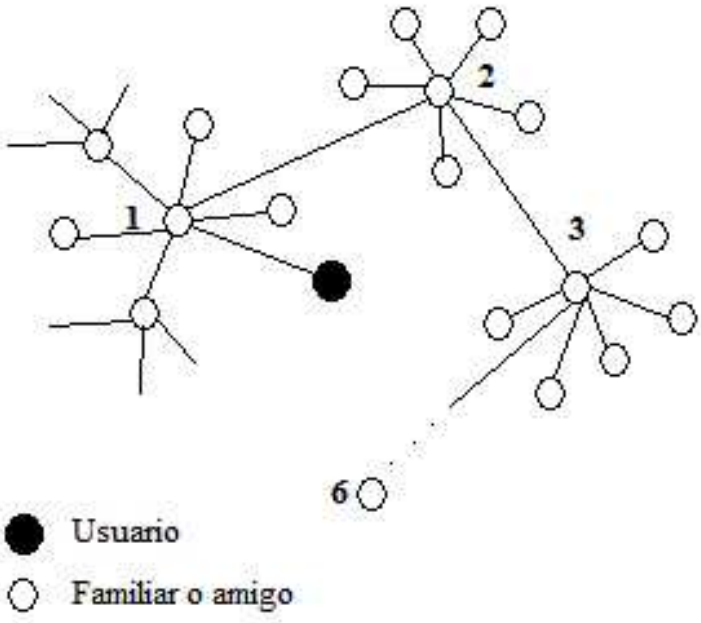

Figura 5.1: Fase I. Obtención de la aplicación

Basándose en la experiencia cotidiana y en la reflejada en la subcultura carcelaria, se supone que cada usuario conoce su entorno, sus necesidades específicas de comunicación, y la forma

posible de solventarlas, es decir, se supone que el usuario tiene la capacidad de comunicarse con su entorno para procurar satisfacer sus necesidades específicas de comunicación, y esto lo hace en base a los conocimientos previos del ambiente que le rodea, y del conocimiento de la tecnología disponible (si un usuario en particular no posee dichos conocimientos, puede utilizar a un familiar o amigo que le apoye en este proceso, pero se puede asumir, sin pérdida de generalidad, que es el mismo usuario quien los posee). Lo anterior implica que cada usuario al tener la necesidad de comunicarse anónimamente, sabe que existen sistemas y redes de comunicación anónima y que existe un proceso que debe seguir para establecer sus conexiones.

Basándose en la teoría de los pequeños mundos, el usuario no debería obtener la aplicación de contactos que estén a más de seis niveles de distancia, tal como se muestra en la figura 5.1.

En la siguiente fase, en la cual se supone que el usuario ya tiene la aplicación instalada, también es necesario hacer otra suposición: dado un entorno con restricciones no es aconsejable utilizar el listado de nodos de salida (puentes o proxies) que trae la aplicación por defecto, ya que es posible que todos los nodos en el listado por defecto ya hayan sido bloqueados. En otras palabras, considerando que la aplicación al ser instalada contiene un listado por defecto de nodos de salida, no debe ser considerado.

En esta segunda fase se deben aplicar en rigor las estrategias de un sistema de reputación y confianza, por ejemplo como el presentado en la propuesta descrita en [30]. Cada usuario debe gestionar sus niveles de reputación ante un ente centralizado, pero a un nivel local, es decir, deben existir líderes de grupo que permitan gestionar la reputación de los usuarios de un grupo. A estos grupos se les denominará *cluster*, tal como se puede observar en la figura 5.2. Estos *clusters* también deben estar constituidos por usuarios que están en el entorno "conocido" del resto de usuarios, o por lo menos, al aplicar la teoría de los pequeños mundos, cada usuario en su grupo no debe estar a más de seis pasos o niveles de cualquier miembro del grupo.

Inicialmente, un usuario, al no tener reputación en su *cluster*, depende de las recomendaciones hechas por sus contactos más cercanos (familiares y amigos), hacia otros miembros. En el caso de que no posea ningún familiar o amigo, debe someterse a un período de pruebas. Se le puede informar al resto del grupo sobre su presencia, y nuevamente, al aplicar la teoría de los pequeños mundos, es factible obtener información sobre el usuario nuevo, dado que el resto de los usuarios no debería estar a más de seis pasos del usuario nuevo, ni a más de seis pasos de algún conocido del mismo. Mientras se obtiene información sobre éste, el líder del cluster puede informarle sobre

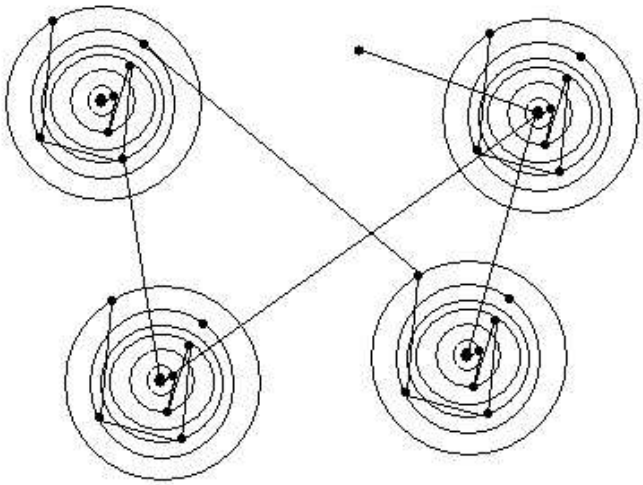

Figura 5.2: Fase II. Conformación de Cluster

uno o dos nodos de salida, dependiendo de la cantidad de nodos de salida que éste conozca, esto con la finalidad de que si es un adversario no pueda bloquear un número significativo de nodos de salida.

Para prevenir el ataque de un líder corrupto, cada usuario en el cluster debe conocer por lo menos un líder de otro grupo, el cual tendría la figura de *líder secundario*, pudiendo ser el mismo líder secundario para varios usuarios del mismo grupo. Si se sospecha alguna irregularidad del líder, se puede utilizar algún líder secundario para establecer alguna comunicación. Nuevamente, cada usuario no debería estar a más de seis niveles de distancia para poder contactar a cualquier otro líder.

Una vez conocidos los líderes de grupo y los secundarios, el proceso de comunicación se puede iniciar en el momento en que cada líder le informa sobre los posibles nodos de salida. A este nivel, es responsabilidad de cada líder mantenerse informado de (descubrir) los nodos de salida posibles en su entorno. Este entorno no tiene una connotación geográfica, sino técnica, ya que sería el entorno que le sea factible al líder en cuanto a sus comunicaciones.

Dado que gran parte de la responsabilidad recae sobre los líderes es importante establecer los mecanismos para la conformación de los mismos: Cada uno, en su fase inicial, puede conformar un cluster al tener la capacidad técnica de hacerlo (conexión a Internet, ordenador, etc.), y al

informar a su entorno cercano sobre su disposición de ser líder. Él puede obtener información de nodos de salida de dos formas: a través de otro líderes, quienes sólo deberían informarle sobre un fracción de sus nodos conocidos, y de otros tipos de participantes que en este caso se les denomina *informantes*. Estos pueden ser usuarios conocidos de forma directa y indirecta (ley de lo pequeños mundos), tal como se muestra en la figura 5.3

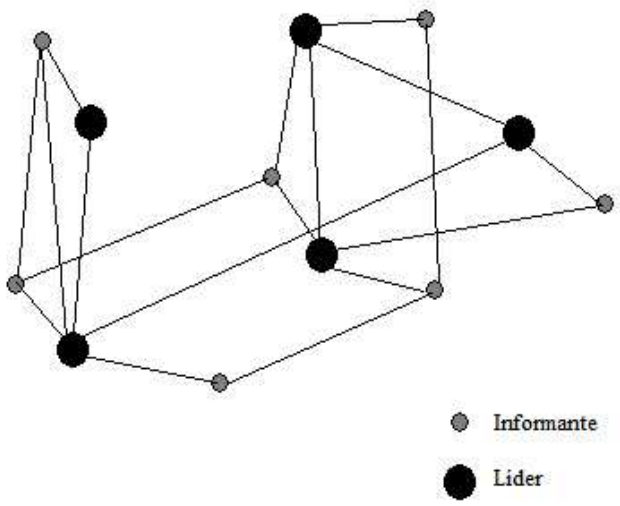

Figura 5.3: Conformación de Líderes

La reputación de los líderes y de los informantes debe ser manejada por los mismos líderes, quienes tendrán en sus bases de datos la información de la reputación de los usuarios, de otros líderes, y de algunos informantes. Los puntos de reputación se ganan en cada una de las transacciones exitosas que haga cada miembro. Se le suma un punto por cada transacción exitosa, y si existe alguna irregularidad por parte de alguno de ellos, los nodos finales, los informantes y los líderes pueden informar al respecto, para incluirlo en un mecanismo de *listas negras*. Esto quiere decir, que los líderes manejarán también un sistema de *listas negras* para controlar el mal comportamiento de los nodos: comportamientos positivos implican ganancia de puntos de reputación, y comportamientos negativos implican el registro temporal o permanente, en al menos una lista negra.

La estructura de informantes, líderes y usuarios está basada en las características de la subcultura carcelaria. Es decir, a nivel carcelario se forman grupos o cotos que tienen la capacidad de gestionar sus servicios y conseguir productos, de forma legal o ilegal, y a pesar del sistema

regulador son capaces de mantenerse comunicados. En estos grupos existe la figura de líderes, informantes y subalternos, quienes también se apoyan en las negociaciones (pasivas o no) que se hacen con grupos cercanos o del exterior.

Macro algoritmo

Para poder establecer un orden secuencial en cada uno de los procesos mencionados anteriormente, es conveniente sintetizarlos con la intención de un mejor entendimiento. El siguiente macro algoritmo muestra los pasos necesarios para generar las comunicaciones anónimas en entornos restrictivos, considerando la perspectiva de los usuarios, de los líderes y de los informantes.

1. Usuarios:
   - *a*) Si un usuario **no** posee la aplicación cliente debe ejecutar lo siguiente:
     1) Solicitar a un miembro de su entorno cercano la aplicación.
     2) Instalar la aplicación.
     3) Eliminar la lista por defecto de nodos de salida.
     4) Si en su entorno próximo no poseen la aplicación, solicitar a los miembros de su entorno que hagan una solicitud al siguiente nivel.
   - *b*) Si a un usuario le solicitan información sobre la aplicación cliente debe ejecutar lo siguiente:
     1) Si posee la aplicación, se la envía al solicitante.
     2) Si no la posee, le solicita a su entorno, que haga una solicitud a un siguiente nivel.
   - *c*) Si el usuario ya posee la aplicación, y no tiene la lista de nodos de salida, debe ejecutar lo siguiente:
     1) Solicitarle a su líder la lista de nodos actualizada.
     2) Establecer la comunicación a través de los nodos de la lista, utilizando un mecanismo de redes anónimas conocido (redes de proxies, de mixes, etc.)
     3) Actualizar la lista de nodos de salida a través de su solicitud al líder.
   - *d*) Si un usuario no posee un líder debe ejecutar lo siguiente:
     1) Preguntar a su entorno cercano (familiares y amigos) sobre la existencia de un líder conocido.

2) Si en su entorno cercano no poseen la información, solicitarle que hagan la petición en el nivel siguiente.

e) Si a un usuario le solicitan información sobre un líder debe ejecutar lo siguiente:

1) Verificar en su entorno la reputación del solicitante.
2) Si el solicitante tiene la reputación adecuada, debe enviarle la información sobre el líder conocido.
3) Si no posee información sobre su reputación, debe preguntar en su entorno cercano.

2. Líderes:

a) Si un usuario desea constituirse en líder debe ejecutar lo siguiente:

1) Conseguir la conexión a Internet, y el ordenador requerido.
2) Comunicarle a su entorno cercano sobre su disposición, para formar el cluster.
3) Formar el cluster con los miembros de su entorno interesados (formar una lista de interesados).
4) Obtener la aplicación de líder desde su entorno cercano. Incluyendo las bases de datos para el manejo de la reputación.

b) Si le es requerido una lista de nodos de salida, debe ejecutar los siguiente:

1) Si el usuario, líder o informante solicitante está en su base de datos con el suficiente nivel de reputación, entregar la lista parcial de los nodos de salida conocidos.
2) Si aun no posee una lista de nodos de salida, solicitarlo a otro líder cercano o a un informante.
3) Si el usuario, líder o informante solicitante no están en su base de datos, o no tiene puntos de reputación, debe solicitarle a su entorno información sobre él (puntos de reputación y listas negras).
4) Si el usuario, líder o informante solicitante están en su base de datos de lista negra, no le debe entregar información alguna.

c) Verificar el estado de los nodos de salida, si los nodos de salidas están activos por cierto período de tiempo, les asigna puntos de reputación a los usuarios, líderes o informantes a los que les haya enviado información sobre esos nodos.

d) Si en la verificación descubre que han bloqueado algún nodo de salida, no incrementa los puntos de reputación.

e) Si se entera por parte de un informante, otro usuario, otro líder o por sí mismo de que alguno a hecho acciones irregulares lo suscribe en su lista negra.

f) La lista negra se distribuye por medio de la estructura de entornos: inicialmente a su entorno cercano, luego le solicita a su entorno cercano que lo haga hacia un nivel superior (más lejano).

g) Si un informante le entrega información sobre los nodos de salida, ejecutar las siguientes acciones:

1) Primero debe verificar su nivel de reputación, si es el adecuado, establece su información como veraz.

2) Si no posee niveles de reputación debe preguntar en su entorno sobre la reputación del informante.

3) Si en su entorno no poseen dicha información, debe solicitarle que hagan la petición a un segundo nivel.

3. Informantes:

a) Si un usuario desea convertirse en informante debe ejecutar lo siguiente:

1) Obtener listados actualizados de nodos, a través de otros informantes, líderes o terceros en quien confíe (miembros de su entorno). Si en su entorno cercano no poseen la información, solicitar que hagan la petición en un segundo nivel.

2) Establecer conexión con uno o varios líderes para dar la información.

b) Si le desea entregar información a un líder, debe preguntarle a su entorno cercano, o a otro líder sobre los niveles de reputación.

c) Si le es solicitado la información, debe verificar con otros líderes la reputación del solicitante.

5.3. Modelo para optimizar sistemas anónimos

La optimización planteada en esta propuesta gira entorno a la maximización de los niveles de anonimato y la minimización de la latencia. Sin embargo se pueden plantear otras combinaciones, que puedan ser resueltas con estas mismas técnicas de optimización multi-objetivo.

Como se mencionó en el apartado 3.5, la idea de base para poder optimizar los dos objetivos planteados es crear el modelo matemático que permita establecer la búsqueda de los niveles óptimos. Este modelo matemático posee dos objetivos que entran en conflicto en las implementaciones prácticas: usualmente si se incrementan los niveles de anonimato se incurre en el aumento de la latencia, y si se desea disminuir la latencia se incurre en una disminución de los niveles de anonimato. Es por esto que se debe buscar un punto de donde confluyan niveles "aceptables" para cada caso práctico en particular (la decisión sobre un nivel "aceptable" la puede hacer el responsable o usuario solicitante), lo que implica que se deberán hacer los cálculos de óptimos en cada uno de ellos.

Para crear el modelo matemático se deben escoger las variables con las cuales se puedan generar las iteraciones en la optimización. Para el caso globalizado, el número de variables puede ser inmanejable, pero se puede simplificar esta tarea clasificándolas de la siguiente forma:

- Cantidad de nodos intermedios.
- Mecanismos de anonimato (mixes, cebolla, proxies, tráfico dummy, etc.)
- Tipos de ataques (tipos de atacantes).

Cada una de estas variables y sus combinaciones, deberán estar para cada uno de los casos donde se les quiera implementar: por ejemplo, en un sistema de servicio web, el tiempo de repuesta no debe sobre pasar cierto límite máximo.

En resumen, se deben buscar las variables, establecer las restricciones y crear las funciones objetivos dentro de las categorías planteadas.

Para un caso en particular, se podría suponer que se quiere modelar un ataque activo, global, adaptativo. Se desea establecer cuál es el mejor mecanismo ante las limitaciones de conexión que existen en China, por ejemplo, hay 40 puntos de salidas, monitoreadas por corta-fuegos. Cada corta-fuego restringe ciertos tipos de comunicaciones, por ejemplo, algunos restringen las comunicaciones ftp, otras la ssl/tls (https), otras las ssh, etc. Cada mecanismo tiene sus propias medidas de rendimiento por nodo intermedio de conexión. Para este ejemplo, cada uno de estas declaraciones constituye un requisito a satisfacer por el modelo. Se mide el rendimiento general utilizando cada uno de los mecanismos (OR, mix, etc.) y se escoge el que menor latencia y mayor anonimato ofrezca para ese escenario en particular, utilizando las técnica de optimización

matemática que corresponda al modelo: programación lineal, no lineal, entera, etc.

Capítulo 6

Conclusiones

Para delimitar el contexto de los sistemas anónimos en escenarios globales primero se presentaron un conjunto de términos y conceptos asociados a las tecnologías que mejoran la privacidad (Privacy enhancing Technologies), principalmente los relacionados al anonimato. Se plantearon algunos de los problemas que pueden surgir durante la implementación de dichos sistemas a escala mundial. Entre los cuales se estudió el problema relacionado al descubrimiento de nodos (puntos de acceso) en zonas donde se censuran y se bloquean las comunicaciones cuyo contenido se considera en contra de una ideología política, religiosa, etc. En este sentido, se propuso un mecanismo que puede resolver este tipo de problemas en estos escenarios, el cual utiliza una combinación de las ideas planteadas para los sistemas de reputación y confianza, la teoría de los pequeños mundos y las ideas extraídas de la subcultura carcelaria como modelo de un escenario donde se restringen las comunicaciones. El planteamiento propuesto sugiere la creación de grupos locales coordinados por líderes, y utilizando un participante adicional que permita apuntalar el tema del descubrimiento a través de un mecanismo de distribución de la información eficiente, denominado informante. Cada uno de los participantes en estos grupos interacciona con los demás gracias a las gestiones de los líderes quienes administran sus niveles de reputación y pueden certificar la sanidad en la comunicación con el resto de los participantes. Los informantes se encargan de buscar y distribuir la información asociada a los nodos (proxies o puentes) disponibles.

También se mencionó la carencia de un mecanismo o sistema que procure minimizar la latencia a la vez que incremente los niveles de anonimato en relación con los sistemas hasta ahora

propuestos. Para esto se propuso utilizar un mecanismo de reenvíos basado en los procesos markovianos, de esta manera, la decisión de cada reenvío en cada uno de los nodos intermedios de una red anónima se haría utilizando técnicas estocásticas basadas en los estados del sistema (cadenas de Markov).

En vista de que las propuestas planteadas involucran combinar varias ideas, mecanismos y sistemas, se plantea también la utilización de un modelo de optimización multi-objetivo, enfocado específicamente en la latencia y el nivel de anonimato, con el fin de encontrar la mejor solución (combinación) para un escenario global.

Toda la ideas planteadas tienen una connotación de propuesta, las cuales se desarrollarán en un trabajo futuro, donde se deberá plantear un mecanismo concreto de integración para los entornos restrictivos, se deberá construir con mayor nivel de detalle el mecanismo markoviano, y plantear el modelo matemático específico: función objetivo y restricciones, junto con sus variables y coeficientes.

Bibliografía

[1] A. Acquisti, S. Gritzalis, C. Lambrinoudakis, and S. D. C. di Vimercati. *Digital Privacy.* Auerbach Publications, 2008.

[2] E. Androulaki, S. G. Choi, S. M. Bellovin, and T. Malkin. Reputation systems for anonymous networks. In N. Borisov and I. Goldberg, editors, *Proceedings of the Eighth International Symposium on Privacy Enhancing Technologies (PETS 2008)*, pages 202–218, Leuven, Belgium, July 2008. Springer.

[3] A. Back, U. Möller, and A. Stiglic. Traffic analysis attacks and trade-offs in anonymity providing systems. In I. S. Moskowitz, editor, *Proceedings of Information Hiding Workshop (IH 2001)*, pages 245–257. Springer-Verlag, LNCS 2137, April 2001.

[4] G. Basharin, A. Langville, and V.Ñaumov. The life and work of A.A. Markov. *Linear Algebra and Applications*, 386:3–26, 2004.

[5] A. Beimel and S. Dolev. Buses for anonymous message delivery. *Journal of Cryptology*, 16(1):25–39, 2003.

[6] O. Berthold, H. Federrath, and M. Köhntopp. Project "Anonymity and Unobservability in the Internet". In *Workshop on Freedom and Privacy by Design / CFP2000*, 2000.

[7] O. Berthold, H. Federrath, and S. Köpsell. Web MIXes: A system for anonymous and unobservable Internet access. *Lecture Notes in Computer Science*, 2009:115–??, 2001.

[8] O. Berthold, A. Pfitzmann, and R. Standtke. The disadvantages of free MIX routes and how to overcome them. In H. Federrath, editor, *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, pages 30–45. Springer-Verlag, LNCS 2009, July 2000.

[9] O. Berthold, A. Pfitzmann, and R. Standtke. The disadvantages of free MIX routes and how to overcome them. *Lecture Notes in Computer Science*, 2009:30–??, 2001.

[10] Z. Brown. Cebolla: Pragmatic IP Anonymity. In *Proceedings of the 2002 Ottawa Linux Symposium*, June 2002.

[11] K. Chatzikokolakis. *Probabilistic and Information-Theoretic Approaches to Anonymity*. PhD thesis, Laboratoire d'Informatique (LIX), École Polytechnique, Paris, October 2007.

[12] D. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 4(2), February 1981.

[13] D. Chaum. Security without identification: Transaction systems to make big brother obsolete. *CACM*, 28(10), October 1985.

[14] R. Clayton, S. J. Murdoch, and R.Ñ. M. Watson. Ignoring the great firewall of china. In G. Danezis and P. Golle, editors, *Proceedings of the Sixth Workshop on Privacy Enhancing Technologies (PET 2006)*, pages 20–35, Cambridge, UK, June 2006. Springer.

[15] C. A. C. Coello. Introducción a la optimización evolutiva multiobjetivo. Technical report, CINVESTAV-IPN. Dpto. de Ingeniería Eléctrica, Sección de Computación, 2002.

[16] G. Danezis. Mix-networks with restricted routes, 2003.

[17] G. Danezis. *Better Anonymous Communications*. PhD thesis, University of Cambridge, July 2004.

[18] G. Danezis. The traffic analysis of continuous-time mixes. In *Proceedings of Privacy Enhancing Technologies workshop (PET 2004)*, volume 3424 of *LNCS*, pages 35–50, May 2004.

[19] G. Danezis and J. Clulow. Compulsion resistant anonymous communications. In *Proceedings of Information Hiding Workshop (IH 2005)*, pages 11–25, June 2005.

[20] G. Danezis and C. Diaz. A survey of anonymous communication channels. Technical Report MSR-TR-2008-35, Microsoft Research, January 2008.

[21] G. Danezis, C. Díaz, and C. Troncoso. Two-sided statistical disclosure attack. In N. Borisov and P. Golle, editors, *Proceedings of the Seventh Workshop on Privacy Enhancing Technologies (PET 2007)*, Ottawa, Canada, June 2007. Springer.

[22] G. Danezis, R. Dingledine, D. Hopwood, and N. Mathewson. Mixminion: Design of a type iii anonymous remailer protocol, 2002.

[23] G. Danezis and B. Laurie. Minx: A simple and efficient anonymous packet format.

[24] G. Danezis and B. Wittneben. The economics of mass surveillance and the questionable value of anonymous communications. In R. Anderson, editor, *Proceedings of the Fifth Workshop on the Economics of Information Security (WEIS 2006)*, Cambridge, UK, June 2006.

[25] O. de las Naciones Unidas. Derechos humanos para todos. *Declaración Universal de los Derechos humanos*, 1948.

[26] C. Díaz. *Anonymity and Privacy in Electronic Services*. PhD thesis, Katholieke Universiteit Leuven, Leuven, Belgium, December 2005.

[27] C. Díaz, G. Danezis, C. Grothoff, A. Pfitzmann, and P. F. Syverson. Panel discussion - mix cascades versus peer-to-peer: Is one concept superior? In *Privacy Enhancing Technologies*, pages 242–242, 2004.

[28] C. Díaz and A. Serjantov. Generalising mixes. In R. Dingledine, editor, *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*, pages 18–31. Springer-Verlag, LNCS 2760, March 2003.

[29] R. Dingledine, M. J. Freedman, D. Hopwood, and D. Molnar. A Reputation System to Increase MIX-net Reliability. In I. S. Moskowitz, editor, *Proceedings of Information Hiding Workshop (IH 2001)*, pages 126–141. Springer-Verlag, LNCS 2137, April 2001.

[30] R. Dingledine, N. Mathewson, and P. Syverson. Reputation in P2P Anonymity Systems. In *Proceedings of Workshop on Economics of Peer-to-Peer Systems*, June 2003.

[31] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. In *Proceedings of the 13th USENIX Security Symposium*, August 2004.

[32] R. Dingledine, V. Shmatikov, and P. Syverson. Synchronous batching: From cascades to free routes. In *Proceedings of Privacy Enhancing Technologies workshop (PET 2004)*, volume 3424 of *LNCS*, pages 186–206, May 2004.

[33] R. Dinglendine and N. Mathewson. Design of a blocking-resistant anonymity system. Technical report, The Tor Project, 2007.

[34] J. Douceur. The Sybil Attack. In *Proceedings of the 1st International Peer To Peer Systems Workshop (IPTPS 2002)*, March 2002.

[35] N. Feamster, M. Balazinska, G. Harfst, H. Balakrishnan, and D. Karger. Infranet: Circumventing web censorship and surveillance. In *Proceedings of the 11th USENIX Security Symposium*, August 2002.

[36] N. Feamster, M. Balazinska, W. Wang, H. Balakrishnan, and D. Karger. Thwarding Web Censorship with Untrusted Messenger Delivery. In R. Dingledine, editor, *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*, pages 125–140. Springer-Verlag, LNCS 2760, March 2003.

[37] M. J. Freedman and R. Morris. Tarzan: A peer-to-peer anonymizing network layer. In *Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS 2002)*, Washington, DC, November 2002.

[38] D. Gambetta. Trust: Making and Breaking Cooperative Relations. Technical Report 13, Oxford University, Oxford, UK, February 2000.

[39] I. Goldberg. *A Pseudonymous Communications Infrastructure for the Internet.* PhD thesis, UC Berkeley, December 2000.

[40] C. Gülcü and G. Tsudik. Mixing E-mail with Babel. In *Proceedings of the Network and Distributed Security Symposium - NDSS '96*, pages 2–16. IEEE, February 1996.

[41] S. Köpsell and U. Hilling. How to achieve blocking resistance for existing systems enabling anonymous web surfing. In *Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2004)*, Washington, DC, USA, October 2004.

[42] B.Ñ. Levine, M. K. Reiter, C. Wang, and M. K. Wright. Timing attacks in low-latency mix-based systems. In A. Juels, editor, *Proceedings of Financial Cryptography (FC '04)*, pages 251–265. Springer-Verlag, LNCS 3110, February 2004.

[43] M. Liberatore and B.Ñ. Levine. Inferring the Source of Encrypted HTTP Connections. In *Proceedings of the 13th ACM conference on Computer and Communications Security (CCS 2006)*, pages 255–263, October 2006.

[44] N. E. Álvarez Licona. Las islas marías y la subcultura carcelaria. *Editorial Letralia*, 1999.

[45] N. Mathewson and R. Dingledine. Practical traffic analysis: Extending and resisting statistical disclosure. In *Proceedings of Privacy Enhancing Technologies workshop (PET 2004)*, volume 3424 of *LNCS*, pages 17–34, May 2004.

[46] S. Mauw, J. Verschuren, and E. de Vink. A formalization of anonymity and onion routing. In D. G. P. Samarati, P. Ryan and R. Molva, editors, *Proceedings of ESORICS 2004*, pages 109–124, Sophia Antipolis, 2004. LNCS 3193.

[47] D. Mazières and M. F. Kaashoek. The Design, Implementation and Operation of an Email Pseudonym Server. In *Proceedings of the 5th ACM Conference on Computer and Communications Security (CCS 1998)*. ACM Press, November 1998.

[48] J. E. Miceli. La ciencia de las redes. *REDES- Revista hispana para el análisis de redes sociales*, 10(10), June 2006.

[49] U. Möller, L. Cottrell, P. Palfrader, and L. Sassaman. Mixmaster Protocol — Version 2. IETF Internet Draft, July 2003.

[50] I. C. Office. Data protection technical guidance note: Privacy enhancing technologies (pets). Technical report, Abr 2006.

[51] L. Øverlier and P. Syverson. Locating hidden servers. In *Proceedings of the 2006 IEEE Symposium on Security and Privacy*. IEEE CS, May 2006.

[52] J. Patel. *A Trust and Reputation Model for Agent-Based Virtual Organisations*. PhD thesis, University of Southampton, January 2007.

[53] A. Pfitzmann and M. Hansen. Anonymity, unobservability, and pseudonymity: A consolidated proposal for terminology. Draft, July 2000.

[54] A. Pfitzmann, B. Pfitzmann, and M. Waidner. ISDN-mixes: Untraceable communication with very small bandwidth overhead. In *Proceedings of the GI/ITG Conference on Communication in Distributed Systems*, pages 451–463, February 1991.

[55] B. Pfitzmann and A. Pfitzmann. How to Break the Direct RSA-Implementation of MIXes. In *Proceedings of EUROCRYPT 1989*. Springer-Verlag, LNCS 434, 1990.

[56] J.-F. Raymond. Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems. In H. Federrath, editor, *Proceedings of Designing Privacy Enhancing Technologies: Workshop*

on Design Issues in Anonymity and Unobservability, pages 10–29. Springer-Verlag, LNCS 2009, July 2000.

[57] M. Reiter and A. Rubin. Crowds: Anonymity for web transactions. *ACM Transactions on Information and System Security*, 1(1), June 1998.

[58] M. Rennhard and B. Plattner. Introducing MorphMix: Peer-to-Peer based Anonymous Internet Usage with Collusion Detection. In *Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2002)*, Washington, DC, USA, November 2002.

[59] Y. Sawaragi, H.Ñakayama, and T. Tanino. Theory of multiobjective optimization. *Mathematics in Science and Engineering - Academic Press*, 176, 1985.

[60] A. Serjantov. *On the Anonymity of Anonymity Systems*. PhD thesis, University of Cambridge, June 2004.

[61] A. Serjantov and P. Sewell. Passive attack analysis for connection-based anonymity systems. In *Proceedings of ESORICS 2003*, October 2003.

[62] C. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423:623–656, 1948.

[63] V. Shmatikov. Probabilistic model checking of an anonymity system. *Journal of Computer Security*, 12(3-4):355–377, 2004.

[64] V. Shmatikov and M.-H. Wang. Timing analysis in low-latency mix networks: Attacks and defenses. In *Proceedings of ESORICS 2006*, September 2006.

[65] R. E. Steuer. Multiple criteria optimization: Theory, computations, and application. *John Wiley & Sons*, 1986.

[66] W. J. Stewart. *Formal Methods for Performance Evaluation*, chapter Performance Modelling and Markov Chains, pages 1–33. Springer Berlin / Heidelberg, 2007.

[67] P. Syverson, M. Reed, and D. Goldschlag. Onion Routing access configurations. In *Proceedings of the DARPA Information Survivability Conference and Exposition (DISCEX 2000)*, volume 1, pages 34–40. IEEE CS Press, 2000.

[68] P. Syverson, G. Tsudik, M. Reed, and C. Landwehr. Towards an Analysis of Onion Routing Security. In H. Federrath, editor, *Proceedings of Designing Privacy Enhancing Technologies:*

Workshop on Design Issues in Anonymity and Unobservability, pages 96–114. Springer-Verlag, LNCS 2009, July 2000.

[69] X. Wang, S. Chen, and S. Jajodia. Tracking anonymous peer-to-peer voip calls on the internet. In *Proceedings of the ACM Conference on Computer and Communications Security*, pages 81–91, November 2005.

[70] D. J. Watts. *Six Degrees: The Science of a Connected Age*. W. W. Norton & Company, 1 edition, feb 2003.

[71] E. I. S. Work Package 13. D13.1 identity and impact of privacy enhancing technologies. Technical report, May 2007.

[72] D. Xiaodong, S. David, and W. X. Tian. Timing analysis of keystrokes and timing attacks on ssh.

[73] Y. Zhu, X. Fu, B. Graham, R. Bettati, and W. Zhao. On flow correlation attacks and countermeasures in mix networks. In *Proceedings of Privacy Enhancing Technologies workshop (PET 2004)*, volume 3424 of *LNCS*, pages 207–225, May 2004.