Sistema descentralizado para la transferencia segura de datos

Trabajo Terminal No. 2020-A103

Alumno: Dávila Méndez Juan Manuel

Directores: Agustín Domínguez Verónica, Moreno Cervantes Axel Ernesto

*e-mail: *jdavilam1603@alumno.ipn.mx*

Resumen: El presente trabajo terminal tiene como objetivo ofrecer una alternativa a las soluciones de mensajería existentes que, por lo general, utilizan blockchain usando un modelo semejante a aquel utilizado en aplicaciones P2P de *torrents*, además de incorporar una arquitectura basada en pares de llaves públicas y privadas, así como algoritmos que frustran los ataques de análisis por repetición.

Palabras clave: Criptografía híbrida, Peer-to-peer, Sistemas descentralizado, Kademlia

1. Introducción

En la actualidad, y en pleno auge de numerosas herramientas tecnológicas que hacen de nuestra comunicación algo sumamente simple, se ha dejado en gran medida la privacidad como un privilegio: como algo de lo cual no deberíamos preocuparnos y simplemente asumir que está ahí, confiando nuestros datos en terceros.

Esto conlleva varios problemas, mismos que son ignorados por la mayoría de usuarios que utilizan dichas soluciones no auditadas y centralizadas, como son:

1) Fugas de datos o información
2) Backdoors, presentes en plataformas de comunicación *privativas*, por ejemplo Skype, o productos de Microsoft. [1]
3) Descubrimiento de vulnerabilidades a través de ingeniería inversa.
4) Desconocimiento del funcionamiento del software de comunicación: el usuario final no sabe si sus datos son desviados a otras entidades, ni cuánto tiempo se almacenan, o si el receptor y propietario son los únicos que conocen los datos que han intercambiado.
5) El modelo de seguridad a través de la oscuridad ha demostrado tener múltiples fallas [2], como es el caso de Windows y su problema con los virus.

Este trabajo terminal pretende solucionar la gran mayoría de estos problemas al utilizar un acercamiento distinto al de una arquitectura centralizada o de blockchain, pues, a pesar de que esta última tecnología es innovadora, sufre de algunos problemas primordiales cuando la red es pequeña o apenas está en sus inicios [3]. Se planea utilizar un enfoque similar al que utiliza el protocolo BitTorrent: existen nodos *trackers* (o seguidores), descubrimiento de nodos a través de la red utilizando algoritmos como Kademlia así como un factor que no se incorpora en el protocolo BitTorrent: cifrado utilizando una arquitectura de llaves públicas y privadas (cifradores asimétricos), y cifrando los datos con un algoritmo de cifrado simétrico (AES por ejemplo), es decir, cifrado híbrido. A diferencia de las soluciones de intercambio de datos seguras, como el email de Lavabit, o Whatsapp (ambas soluciones han sido vulneradas o los datos que poseían se han visto vulnerados, hasta cierto punto[4,5]) no se utilizarán intermediarios para el almacenamiento de datos o intercambio de datos: estas se darán directamente de nodo a nodo, evitando así los fallos inherentes a una arquitectura centralizada.

A continuación, se muestra software similar, así como la solución propuesta:

SOFTWARE	CARACTERÍSTICAS	PRECIO EN EL MERCADO	Código abierto
Fopnu	Es un software creado específicamente para compartir archivos, estilo Ares. Es descentralizado y completamente Peer-to-peer. [6] No es de código abierto.	Gratuito	✗
Dust	Cliente de mensajería basado en blockchain. Posee eliminación de mensajes. [7] No es de código abierto.	Gratuito	✗
Crypviser	Mensajería cifrada de punto a punto (se desconoce la infraestructura), basado en blockchain. [8] No es de código abierto.	43 € por año.	✗
Solución propuesta	Completamente descentralizado. Funcionamiento inspirado en BitTorrent, incluyendo implementación de *trackers* y algoritmos de descubrimiento de nodos. Infraestructura de cifrado para intercambio de información. Debido a que es resultado de un TT, su uso estará restringido a las aplicaciones que convenga el Instituto Politécnico Nacional.	Gratuito	✓ (Para miembros del Instituto Politécnico Nacional)

Como se puede observar, todo el software de la tabla anterior sirve para un mismo propósito: comunicación. Sin embargo, los mecanismos que utilizan a nivel interno son desconocidos y no han sido auditados por la comunidad que hace uso del software. En la justificación se detallan las razones por las cuales este TT es innovador y ofrece algo nuevo a diferencia del software mencionado anteriormente.

Cabe destacar que este software es más enfocado al usuario intermedio y destinado al uso en computadoras, pues su uso en el teléfono implicaría tiempos de desarrollo sumamente largos, además de ser necesario tomar en cuenta que los teléfonos celulares generalmente no son actualizados de manera frecuente a comparación de una computadora de escritorio con Windows, o alguna distribución popular de Linux, mismas que reciben actualizaciones de manera automática.

Siendo que la solución propuesta es resultado de un TT, como trabajo a futuro se pretende hacer una solución inspirada en el TT actual con un modelo de licenciamiento que ofrezca más permisos, para permitir el uso público ya sea con o sin fines de lucro.

2. Objetivo

Desarrollar un sistema que ofrezca una arquitectura descentralizada, privada y segura para la transferencia de datos, inspirándose en el funcionamiento del protocolo BitTorrent.

Objetivos específicos:

- Utilizar el algoritmo Kademlia, mismo que se utiliza en el protocolo BitTorrent, pero con un enfoque para comunicación.
- Agregar seguridad al mecanismo de intercambio de datos utilizando infraestructura de cifrado híbrida (RSA + AES, por ejemplo).
- Diseñar la interfaz en una plataforma de widgets, ya sea Wx o QT.

3. Justificación

Como se mencionó anteriormente, la privacidad es un tema que debería ser sumamente importante para todas las personas que utilicen servicios de comunicación, sin embargo, es un tema que se ha dejado atrás, quizá por razones de facilidad, o por la simplicidad que ofrece el implementar una arquitectura centralizada y menos segura.

Una de las varias soluciones podría ser un sistema utilizando una infraestructura de llaves públicas y privadas, sin embargo, esto por sí mismo no ofrece resiliencia ante fallos, ataques, así como seguridad en caso de que el servidor central (si es que se utiliza una arquitectura centralizada) caiga. Es aquí donde toma importancia que el sistema sea descentralizado.

Este sistema da solución a los problemas anteriormente mencionados, pues los datos no se almacenan en ningún lugar antes de llegar al destinatario, no pasan por ningún intermediario (algo que es común en una arquitectura centralizada) y al usar una infraestructura de cifrado híbrida, aún si los datos son interceptados, serían inútiles para el interceptor.

Dicho sistema compartirá ciertas características con otros sistemas destinados para otros usos (por ejemplo, con cualquier cliente de *torrent*) en el sentido que se planea utilizar el algoritmo Kademlia para descubrir otros nodos, o la idea de que se tienen entidades llamadas *trackers* para asistir con el descubrimiento de nodos de manera más rápida y eficiente. Sin embargo, una de las características novedosas y diferenciadoras, es que: 1) este sistema está diseñado para transferir mensajes y datos de nodo a nodo de manera directa, no para compartir datos de manera masiva, y 2) el tráfico entre nodo y nodo estará fuertemente asegurado con cifradores confiables, como RSA + AES (conocido como cifrado híbrido), esto con el fin de que, incluso si el tráfico es *sniffeado* o capturado, no será de gran utilidad a menos que dicho atacante conozca la llave privada del destinatario.

El uso de tecnologías como cifrado híbrido (RSA+AES, por ejemplo) tiene como objetivo ofrecer mayor privacidad y seguridad a costa de un consumo mayor consumo de recursos, y el hecho de que sea descentralizado es para ofrecer mayor disponibilidad, además de evitar que grandes corporaciones, entidades externas o gobiernos puedan desestabilizar el sistema, modificar o espiar los datos en tránsito, además de que ofrece un sistema cuya infraestructura no depende de un proveedor, como es el caso de Facebook, Whatsapp, entre otros.

Este sistema es viable en los periodos de tiempo asignados a TT1, TT2, así como protocolo. Cabe destacar que este sistema ya se tenía pensado con anterioridad y se implementaron unos cuantos módulos para la materia de criptografía, por lo cual el tiempo de desarrollo podría reducirse ligeramente.

Con este sistema se pretende innovar la manera en que se comunican aquellas personas que necesitan de un sistema que sea verdaderamente privado, eficiente y seguro, a diferencia de los sistemas mencionados en la tabla anterior, mismos que anuncian y hacen publicidad de las tecnologías que incorporan, pero que sin embargo jamás han sido auditadas o no se sabe cómo funcionan en realidad.

Al realizar este proyecto, se hará uso extensivo de los conocimientos adquiridos en la carrera de Ingeniería en Sistemas Computacionales, especialmente aquello visto en las materias de redes, programación, así como criptografía.

Este sistema va dirigido a las personas dentro del Instituto Politécnico Nacional que deseen hacer uso de un medio de comunicación rápido, verdaderamente seguro y eficiente, o bien, aquellos que deseen tener certeza de que los datos que intercambian no están siendo desviados o tratados por terceros sin su consentimiento. Por último, pero no menos importante, el sistema podría tener infinidad de aplicaciones, tanto para comunicación entre áreas administrativas dentro del Instituto Politécnico Nacional como para simple intercambio de mensajes entre los miembros de la comunidad.

4. Productos o resultados esperados

Al final del TT, como resultado general se espera tener el sistema descentralizado para intercambio de datos de manera segura funcionando, así como software para computadora de escritorio tanto para el lado del cliente (nodos) como servidor (trackers), sin olvidar las pruebas de tráfico para comprobar que, efectivamente, los mensajes están cifrados de manera correcta sin mostrar patrones repetitivos, además de los resultados y pruebas que se especifiquen en TT1.

A continuación se muestra un diagrama que muestra a grandes rasgos los principales elementos que compondrán el sistema, incluyendo el mecanismo de intercambio de datos.

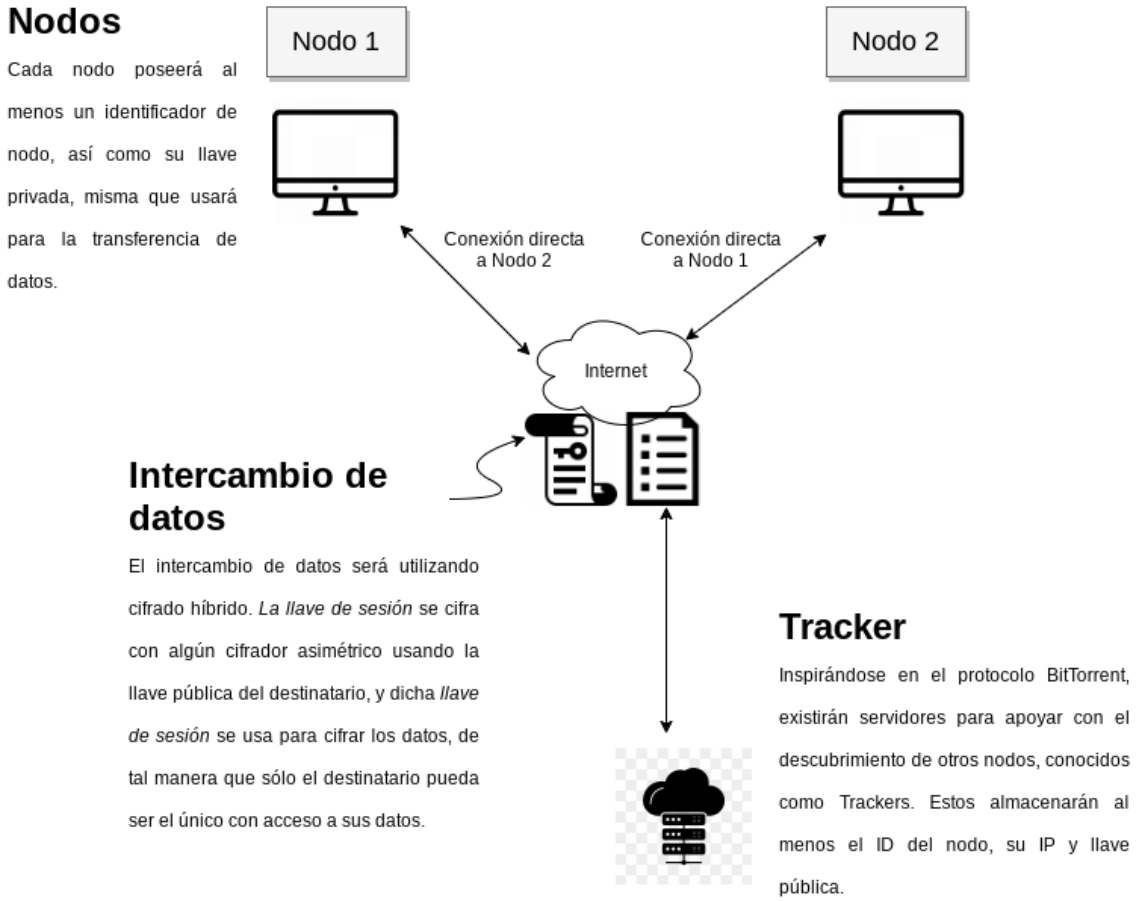

Figura 4.1: Arquitectura planeada para el sistema propuesto.

5. Metodología

Para el sistema en cuestión, se decidió utilizar una metodología en V, debido a que permite tener más control sobre las necesidades del proyecto a corto plazo, pues dada la naturaleza de este sistema, habrán componentes que tomarán más, o menos tiempo de implementar, por lo cual es necesario definir los requerimientos y delimitar cuánto tiempo se le asignará al desarrollo de cada uno de éstos.

Otras metodologías distintas a la ya elegida se tomaron en cuenta, como es el caso de Scrum. Sin embargo, esta última es más recomendada para equipos de trabajo que manejen procesos de manera sumamente rápida.

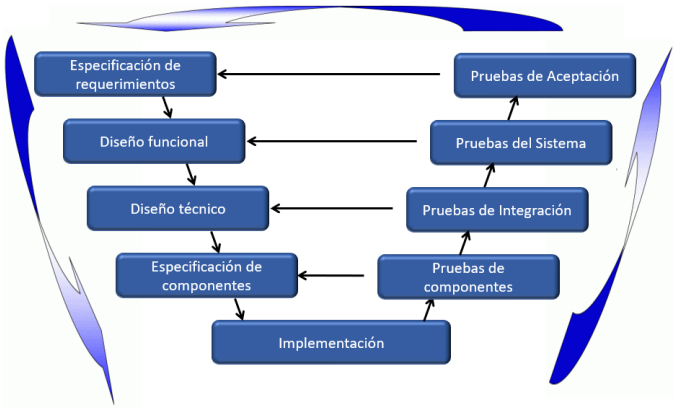

Figura 5.1: Diagrama de la metodología en V.

Para lograr todos los objetivos del sistema descentralizado en cuestión a realizar, se pretende trabajar por partes el proyecto así como definir las posibles subpartes o dificultades que pueda presentar cada una de las partes mencionadas anteriormente. La metodología en V se presta en gran parte para lograr estos objetivos, pues de esta manera es posible aplicar pruebas individuales a cada uno de los módulos y, en caso de que se considere que el módulo ya está listo, será posible seguir avanzando con los módulos o partes faltantes del sistema.

6. Cronograma

Nombre del alumno: Juan Manuel Dávila Méndez
Título del TT: Sistema descentralizado para la transferencia segura de datos

Actividades	Ag	Sep	Oct	Nov	Dic	Ene	Feb	Mar	Abril	May	Jun
Análisis de requerimientos											
Análisis profundo de Kademlia											
Diseño de algoritmo modificado usando Kademlia											
Diseño de algoritmo para ponchado de puertos en redes con NAT											
Implementación de Kademlia+Ponchado de puertos											
Implementación de software para trackers y clientes											
Creación de GUI											
Pruebas finales del sistema											
Evaluación de TT1											
Evaluación de TT2											

7. Referencias

[1] NSA Docs Detail Efforts To Collect Data From Microsoft's Skype, SkyDrive, And Outlook.com, Accedido el 10 de septiembre, 2020. [En línea]. Disponible: https://techcrunch.com/2014/05/13/nsa-docs-detail-efforts-to-collect-data-from-microsofts-skype-skydrive-and-outlook-com/

[2] Principle 6: Security Through Obscurity Is Not an Answer, Accedido el 5 de marzo, 2020. [En línea]. Disponible: https://www.pearsonitcertification.com/articles/article.aspx?p=2218577&seqNum=7

[3] What Is a 51% Attack?, Accedido el 10 de septiembre, 2020. [En línea]. Disponible: https://academy.binance.com/security/what-is-a-51-percent-attack

[4] WhatsApp sues spyware maker for allegedly hacking phones worldwide, Accedido el 5 de marzo, 2020. [En línea]. Disponible: https://nakedsecurity.sophos.com/2019/10/31/whatsapp-sues-spyware-maker-for-allegedly-hacking-phones-worldwide/

[5] Cheeky Lavabit *did* hand over crypto keys to US government after all. Accedido el 5 de marzo, 2020. [En línea]. Disponible: https://nakedsecurity.sophos.com/2013/10/04/cheeky-lavabit-did-hand-over-encryption-keys-to-us-government-after-all/

[6] Fopnu, Accedido el 5 de marzo, 2020. [En línea]. Disponible: https://www.mdpi.com/2076-3417/9/9/1788

[7] Dust, Accedido el 5 de marzo, 2020. [En línea]. Disponible: https://usedust.com/

[8] CrypViser, Accedido el 5 de marzo, 2020. [En línea]. Disponible: https://crypviser.network/

8. Alumnos y directores

Juan Manuel Dávila Méndez, Alumno de la carrera de Ingeniería en sistemas computacionales, Instituto Politécnico Nacional, Boleta: 2017630421, Tel: 5562880306, email: jdavilam1603@alumno.ipn.mx

Firma: _____

CARÁCTER: Confidencial FUNDAMENTO LEGAL: Artículo 11 Fracc. V y Artículos 108, 113 y 117 de la Ley Federal de Transparencia y Acceso a la Información Pública. PARTES CONFIDENCIALES: Número de boleta y teléfono.

Maestría en Administración de ESCA Unidad Santo Tomás IPN, Contadora Pública Certificada ESCA IPN
Profesora Titular de ESCOM-IPN (Departamento ISC) desde 2009
Áreas de interés: Gestión y asesoría empresarial, desarrollo de empresas, avagustin@ipn.mx, ext 52032
Administración de organizaciones y desarrollo de Start Ups.
Contacto: vagustin@ipn.mx, ext 52032

Firma: _____

Doctor en Educación del CUGS EN 2020, M.en C. D l CINVESTAV EN 2004, ISC de ESCOM en 2000. Profesor titular de ESCOM- IPN (Depto. ISC) desde 2004. Áreas de interés: Seguridad en redes. Contacto: axelernesto@gmail.com, ext: 52032

Firma: _____