# EL DIARIO DEL TFG

## Anteproyecto

Bien, lo primero a decir es que cuando se me propuso la idea inicial no me llamaba mucho la atención principalmente por las expectativas que tenía sobre el TFG y el enfoque que había ido modelando durante años. Todo esto sumado a que no tenía ni puta idea del tema y no sabía como enfocarlo me desmotivaba bastante. Pero esto no deja de ser una asignatura más, así que tiempo, esfuerzo y me guste más o menos toca hacer las cosas bien por lo que dando mi mejor cara con los tutores empecé a hablar con ellos y a informarme.

Al principio no yo sabía buscar muy bien información ni tenía muy claro el enfoque, solo contaba con papers largos en inglés que hablaban de temas que no conocía. Pero poco a poco iba haciendo preguntas, informándome más y montando una idea preliminar de lo que había que hacer, pero siempre con la sensación de que por mucho que podía entender qué pretendíamos crear, con qué objetivos y demás, yo no sabía cómo materializarlo.

A todo esto había que hacer el anteproyecto, y aunque realmente tardé en ponerme debido a mi desmotivación y a que tampoco tenía muy claro qué debía poner y qué no, al final, apurando las fechas casi a lo máximo, hice una primera versión que pareció ser bastante top según mis tutores, y tras unas breves correcciones se entregó todo a tiempo en el segundo plazo de entrega de anteproyectos.

Después de esto admito que estuve una temporada tocándome los cojones, primera tarea hecha, más o menos bien, tocaba lo gordo y no sabía cómo enfocarlo, además que faltaban cosas por definir en el proyecto, así que hasta que no llegó el paper de “Preliminary Document TFG” no tenía ganas de ponerme.

Pero tocaba mini hostia de realidad, EMPEZABA EL DESARROLLO, así que a ponerse las pilas de nuevo.

## Desarrollo del TFG

El paper era escasillo y no se explicaba demasiado bien. Sumando esto a que hacía referencia a otros de los que había visto y que no me acordaba de nada de ellos, supuso necesitar tiempo, pausas y leerlo bastantes veces para plantearme posibles preguntas que me sirvieran para obtener información de cómo empezar esta mierda.

Las preguntas surgieron, las apunté en el repo de git, pero entre que yo no me explico cuando no controlo del tema (y es el caso) y que los tutores son un poco meh, pues llegó la tutoría, pregunté, entendí algo más algunas cosas, pero muchas otras siguen siendo un misterio. Pero bueno, son problemas del futuro, el caso es que hay que empezar YA a programar y redactar y para eso está este documento y las primeras pruebas que voy haciendo en base al framework que me pasó Davide y la información que estoy buscando.

Lo primero que destaco es que hay muy poca información sobre brokers y sobre MQTT, porque o te aprovechas de los servicios ya construidos o si quieres buscarte la vida o indagar en cómo se construyeron es o demasiado difícil o no hay nada. Sigo sin tener del todo claro todo lo que debo hacer, pero al menos entendí mejor el tema del gestor de anomalías y la comunicación con nuestro gestor de confianza y por ahora lo que creo que debo hacer es conseguir una comunicación bidireccional entre estos módulos y en base a la anomalía del dispositivo/sujeto X recalcular para este la confianza de alguna forma y reenviarla a anomalías.

Para hacer esto estoy probando a desglosar tareas, ver cómo puedo mandar datos con java entre distintos puertos y cómo funciona el protocolo MQTT, con sus librerías. Me he hecho un esquema a mano que quiero pasar a ordenador, para tenerlo más claro y poder apuntar las preguntas de forma visible y así también compartirlas con los tutores sobre algo que puedan ver, quizás así me explique mejor, pero antes de eso tengo que llevar algo, no puedo ir sin nada, así que ya sea desarrollo de la memoria o pruebas de implementación toca empezar de verdad.

He encontrado Paho, pero la verdad que la documentación es escasa y pese a que tenía una idea inicial de cómo comunicarme con MQTT haciendo algo similar a los sockets, no parece funcionar porque todo lo que he visto online para MQTT requiere de un bróker externo y la gracia es que realmente nosotros creo que no queremos usarlo porque precisamente queremos ser ese bróker, así que veremos cómo me apaño. Por ahora sigo buscando información sobre esta librería, indagando si hay más o probando de mientras con sockets a mandar, recibir información e ir modelando un poco qué datos requerimos, qué clases, qué funciones…

Mi idea primero es conseguir comunicar bidireccionalmente el módulo de anomalías con el de confianza, estudiar qué datos (variables) definir en cada uno y qué funciones hacen falta. Después seguramente debería definir los objetos/dispositivos/cosas para poder comprobar si el ID que envía anomalías existe entre lo que tenemos y entrar entonces o no en el cálculo de confianza.

Esto creo que será lo más difícil porque no tengo ni puta idea de cómo calcular esa confianza, de qué valores depende, si hace falta una reputación almacenada en una base de datos que haga las veces de reputación inicial… además si hiciera falta eso la reputación debería actualizarse cuando se calcule la confianza, pero cómo?? Qué valor se pondría, la nueva confianza calculada o en base a esta con alguna función se calcula la reputación??

Hola de nuevo, continúo contando. Primero debo decir que estoy bastante desmotivado con la vida ahora mismo así que no tengo ganas de nada, lo cual repercute de una manera u otra en el TFG. Pero soy consciente de que necesito al menos una tutoría o dos antes de Semana Santa para preguntar las dudas del Excalidraw y las que me vayan surgiendo y para presentar al menos un mínimo de lo que estoy haciendo. Por mucho que me joda, toca continuar.

Continuo mi narración y debo decir que el protocolo MQTT es un asco, necesito ayuda, porque si queremos hacer las veces de bróker para no utilizar uno no sé cómo cojones vamos a comunicarnos, además que siendo un protocolo de pub/sub qué gracia tiene entre anomalías y confianza?? Se quieren enviar datos, no seguir un sistema así. Obviando esa duda y para comenzar de alguna forma estoy jugando un poco con Sockets, comunicando un servidor que hace las veces del módulo de confianza con un cliente que sería anomalías. El cliente me envía un ID y una anomalía asociada a ese ID y en el servidor se verifica que el ID existe entre los objetos que se crean previamente de forma aleatoria y por ahora he llegado a poco más. Me surge ahí otra duda y es acerca de la anomalía 🡪 cómo va a venir ese dato? Y cómo lo utilizamos, me explico: viene un ID y sobre ese una anomalía, pero esta debe estar registrada o tener un valor o algo con lo que recomputar la reputación y confianza en base a lo conocido sobre ella o no. No es la mejor explicación que he dado, pero creo que me entiendo.

Dicho esto, también tengo mis dudas sobre los objetos a crear, tanto porque no sé qué variables y funciones deberían tener para que sean útiles para calcular la confianza, tampoco sé si HUMANOS cuenta como objeto a crear y tampoco sé si entre estos objetos debe haber comunicación, que en teoría sí, también con el MQTT, ¿¿pero con qué fin?? ¿¿¿De qué me sirve que se comuniquen entre sí y con el Broker???

Voy a hacer una pequeña lista de tareas que me gustaría completar antes de la tutoría que quiero pedir antes de las vacaciones y quizás también incluya el recopilatorio de preguntas para tenerlo todo centralizado.

|  |  |
| --- | --- |
| TAREA | DESCRIPCIÓN |
| Ordenar el código | Quiero que la estructura quede bien separada entre el main, las distinas funciones y variables, y los propios archivos bien ordenador en paquetes. |
| Investigar java + BD | Quizás sea conveniente usar una BD de cara a almacenar la reputación inicial de cada objeto o algo así, aunque sino se podría simular con una variable que se genere aleatoriamente. |
| Investigar cálculo de confianza | Cómo calculo la confianza, qué valores necesito, qué tiene que ver la reputación y como computo si es necesario la actualización de reputación. |
|  |  |
|  |  |

*19-3-24*

He estado revisando la clase dispositivos y nombrando las variables de forma que queden más claras. Me surge la duda de si Dispositivos Físicos y Virtuales deben ser subtipos de Dispositivo y tener las mismas propiedades o alguna diferente partiendo de lo mismo, o deben ser completamente distintas clases.

Fuera de esto lo que más me preocupa ahora que lo pienso es cómo cojones computar la confianza. Necesito que se me pase un ID y una anomalía, rescatar el dispositivo/objeto con ese ID, verificar cómo de grabe es la anomalía porque quizás ya se conocía y se corrigió o es nueva o se conocía, pero es más o menos grave… Esto junto con la reputación previa del ID si existe o si el dispositivo es nuevo sin este campo o algo así computar un nivel de confianza (cómo sé si es nuevo o cómo consigo que haya dispositivos nuevos sin reputación inicial y otros con reputación inicial que ya sean más antiguos).

*21-03-24*

Ayer miré algo sobre Docker MQTT, y al final no deja de ser un Docker compose y juegos de comandos, pero sigue necesitando de Mosquitto.

Lista preguntas tutoría:

* **Requisitos 🡪 cálculo de confianza en base a lo que llegue de anomalía sobre el objeto que nos diga y teniendo en cuenta qué anomalía sucede para recomputar la confianza | comunicación MQTT confiable y segura entre objetos (esto de qué nos sirve)**
  + Respuesta 🡪 las anomalías pueden ser vulns, que se comuniquen con quien no puede y demás. El entorno del dispositivo.
  + Caso de uso: alguien que quiere entrar a su casa, teléfono, y cuando entra se abran las ventanas y se enciendan las luces.
* **Humano es objeto?? 🡪 Excalidraw**
  + Respuesta 🡪 el humano es objeto, pero las anomalías son intencionadas.
  + Mejor centrarse por ahora en dispositivos 🡪 heredan ambos de cosa.
* **pub/sub qué gracia tiene entre anomalías y confianza?? Se quieren enviar datos, no seguir un sistema así**
  + Respuesta 🡪 por ahora sockets y luego se podría extrapolar
* **El ID que se envía desde anomalía debe verificarse que exista entre los objetos/dispositivos existentes??**
  + Respuesta 🡪 sí
* **Me surge ahí otra duda y es acerca de la anomalía 🡪 cómo va a venir ese dato? Y cómo lo utilizamos, me explico: viene un ID y sobre ese una anomalía, pero esta debe estar registrada o tener un valor o algo con lo que recomputar la reputación y confianza en base a lo conocido sobre ella o no**
  + Respuesta 🡪 se pueden enviar más datos aparte. GRADO DE ANOMALÍA.
  + Si la anomalía no existe hay que añadirla a una base de datos con las anomalías contempladas y el grado de gravedad. Anomalías asociadas a un escenario concreto.
* **cómo calcular esa confianza, de qué valores depende 🡪 variables de los objetos?? Cuáles serían necesarias??**

**si hace falta una reputación almacenada en una base de datos que haga las veces de reputación inicial o simulada con un número de ramdom que exista o no en base a si el dispositivo/objeto tiene o no una reputación inicial… además si hiciera falta eso la reputación debería actualizarse cuando se calcule la confianza, pero cómo?? Qué valor se pondría, la nueva confianza calculada o en base a esta con alguna función se calcula la reputación??**

* + Respuesta 🡪 reputación necesaria, se actualiza con la confianza.
* **Paho**
  + Respuesta **🡪**
* **Docker MQTT**
  + Respuesta 🡪

TENGO QUE CENTRARME EN UN ESCENARIO CONCRETO. PARA CREAR EL MODELO DE CONFIANZA Y TODO LO DEMÁS.

*31-3-24*

Yepaaaaaaaaaaa, he vuelto. En semana santa no he tocado absolutamente nada, sigo bastante desmotivado con todo, pero bueno, hay que darle, la vida sigue.

Primero y principal, hay que avanzar con el código y quizás ir empezando con el chasis de la memoria, pero lo principal es el código. Dicho esto, vamos a hacer un breve resumen/recapitulación de los puntos a tener en cuenta hasta el momento para empezar a hacer algo.

Recapitulación desde la última tutoría. Cosas que hacer o avanzar YA DE YA:

1. Establecer un caso de uso en el que centrarme. (quizás como el del paper o algo así).
2. Es necesario crear una base de datos con diferentes tablas 🡪 para la reputación, para las incidencias/amenazas…
3. Hay que crear una clase COSA y de ahí dos hijos que sean los tipos de dispositivos.
4. Basarme en los papers porque el sistema será muy similar.

### Caso de uso.

Tengo que definir si o si el caso de uso en el que centrarme antes de nada para así definir la metodología con la que calcular la confianza, las variables de los diferentes agentes/objetos involucrados, las tablas de BD necesarias y las columnas de estas…

Red conformada por una bombilla, un móvil y una televisión. El móvil puede comunicarse con ambos dispositivos, la tele también, pero la bombilla solo con el móvil. (Se podría decir que el móvil y la tele están en la red externa, la bombilla en la interna??)  
  
Hay que conseguir esa comunicación segura y gestionar las incidencias que nos mande el gestor de anomalías, ya sean vulns, intentos de comunicaciones que no deberían darse…  
  
Hace falta una base de datos con información sobre la reputación inicial de estos dispositivos y las amenazas conocidas y su gravedad.

Pongamos ese caso de uso, aunque realmente tampoco sé si está bien definido del todo, pero bueno, por probar… Dicho esto qué toca ahora? Qué variables hacen falta en cada dispositivo:

|  |  |
| --- | --- |
| Dispositivo | |
| Virtual | Físico |
| Variables:   * Tipo (indique si un dispositivo es físico o virtual) * ID * Fecha Última supervisión * number of past vulnerabilities * Certificación de seguridad sí o no * Reputación inicial | Variables:   * Tipo (indique si un dispositivo es físico o virtual) * ID * Fecha Última supervisión * number of past vulnerabilities * Certificación de seguridad sí o no * Reputación inicial |

*2-4-24*

Ha costado bastante, pero al final he conseguido que funcione el putísimo JDBC en VSCode. He probado en todos los IDEs, pero finalmente haciendo uso de XAMPP y con el controlador de MySQL he conseguido incluirlo en el proyecto java y que funcione la conexión a la base de datos.

Ahora toca crear las bases de datos e incluir esto en el código principal. La primera va a ser la de reputación que por ahora va a tener las siguientes columnas:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Reputación** | | | | | | |
| ID dispositivo | Nombre dispositivo | Tipo dispositivo | Reputación | Last\_update | Estado dispositivo | Comentarios/Descripción |

Creando la base de datos de reputación me ha surgido una duda 🡪 Sería más conveniente crear una base de datos de la reputación que enlace el ID\_Dispositivo con su reputación y algunos datos más, o una BD para los dispositivos y todos sus atributos que además incluya una columna reputación con un valor si el dispositivo tiene alguna antigüedad en la red o null si no. Respecto a esto creo que he llegado a la conclusión de que la tabla de reputación no va a ser como se explicaba, va a ser la que contenga los datos de los dispositivos con los que hacer el cálculo de confianza y una columna reputación que puede ser nullable

Creo que también sería óptimo ir haciendo un diagrama de flujo del MAIN para saber qué está ocurriendo, en qué orden y demás.

***Un caso: desde anomalías llega el ID y el grado de anomalía de una anomalía que consiste en que un dispositivo se quiere comunicar con otro 🡪 en este caso se debe evaluar la confianza de este dispositivo y determinar si pueden o no comunicarse, en caso de que sí, que se establezca la comunicación y en caso de que no, se bloque. Esto sería un caso de uso de la comunicación entre dispositivos junto con la comunicación confianza-anomalía.***

*5-4-24*

Tengo hasta el momento generación aleatoria de dispositivos, conexión con la base de datos y poder guardar estos dispositivos. Ahora bien, necesito conseguir relacionar anomalías y lo que me manda con lo que tengo de los dispositivos y reputación porque ahora mismo no le veo la relación. Además, tengo que desarrollar un modelo de cálculo de confianza aunque por ahora tenemos la idea de los pesos para cada parámetro del dispositivo por relevancia y la suma y un threshold.

Mi idea por el momento creo que va a ser la siguiente: desde amenazas nos van a llegar los datos para la gestión de riesgo, los cuales van a ser los del paper Segragated para riesgo, con sus niveles y sus mierdas. Con esto vamos a hacer una primera estimación del riesgo y en base a ese nivel se quedará el dispositivo en prohibido o su estado se tendrá que seguir evaluando si el riesgo no es demasiado alto, calculando la confianza con sus parámetros, reputación… Este estado debe quedar reflejado en la base de datos para saber si el dispositivo queda o no en “cuarentena”.

En cuanto deje la idea terminada y funcional explico el código función a función, clase a clase.

*7-4-24*

Vale, voy a hacer un poco diagrama de flujo de los mains, del cliente y del servidor y así vemos que está pasando, explicamos sus partes y qué consigo, dudas, etc.

Dudas:

* ID\_existente 🡪 Sería necesario obtener los datos de la BD antes de comparar el thingID que nos envía anomalía con el de los distintos dispositivos creados, o podemos aprovecharnos de la lista de dispositivos creada (operatorThing) ????
* Nivel\_riesgo 🡪 se evalúa el riesgo multiplicando los factores de detectabilidad, criticidad y probabilidad que nos envía anomalías sobre la anomalía asociada al dispositivo, es la forma correcta de hacerlo??? En caso de que sí, habría que actualizar el estado\_actual de los dispositivos en la BD pero al ser un método externo no tengo las variables de la conexión ni nada para actualizar esta data.
* Nivel\_confianza 🡪 la forma que he decidido por ahora es asignando un peso a las variables relevantes de un dispositivo que varía entre el 0-1 y multiplicando este por el valor del atributo. Después se suma todo y si pasa un umbral determinado pues es confiables y sino pues queda en cuarentena a la espera de algo, aunque no sabría decir de qué.
* Base de datos anomalías 🡪 hace falta?? Porque en mi planteamiento no. Si hace falta para qué?
* Caso de uso 🡪 no tengo muy claro cuál, definirlo mejor por favor y gracias.

Reputación inicial 🡪 o se cambia de nombre o hace falta una reputación actual.

Historial reputación para las estadísticas

Mqtt básico entre dispositivos e intentar incorporarlo a la comunicación hasta ahora

*21-04-2024*

Hola de nuevooo, cuánto tiempo… He vuelto y toca hacer cositas, mi idea de hoy es desarrollar texto y continuar con lo dicho en la última reunión. Para empezar, cuántos apartados va a tener mi TFG?

* Portadas
* Abstract and key words
* Resumen y palabras clave
* Índice
* Introducción
  + Motivación
  + Objetivos
  + Metodología
  + Estructura de la memoria
* Estudio del arte / Tecnologías a utilizar
* …
* Conclusiones y líneas futuras
* Manual de usuario
* Apéndice