DISEÑO DE ENTREGA ANTI-SESGO PARA ARCHIVOS DIGITALES UTILIZANDO TECNOLOGÍAS DESCENTRALIZADAS

UNIVERSIDAD
COMPLUTENSE
MADRID

Componentes:

Alejandro Cancelo Correia
Alejandro Cilleros Garrudo
Tomás Golomb Durán
Roberto Portillo Torres
Raúl Sánchez Montaño
Mihaita Sorinel Tudor

Introducción

En clase de Ética, Legislación y Profesión hemos estudiado los privilegios y sesgos que se presentan como un tema preocupante de cara al desarrollo social. También se nos introdujeron las tecnologías blockchain y las redes descentralizadas. A raíz de esto se nos ocurrió la idea que vamos a presentar en este documento para paliar este problema.

La siguiente noticia fue debatida en clase y refuerza el sentido a desarrollar la solución que proponemos:
https://www.theguardian.com/technology/2016/feb/12/women-considered-better-coders-hide-gender-github

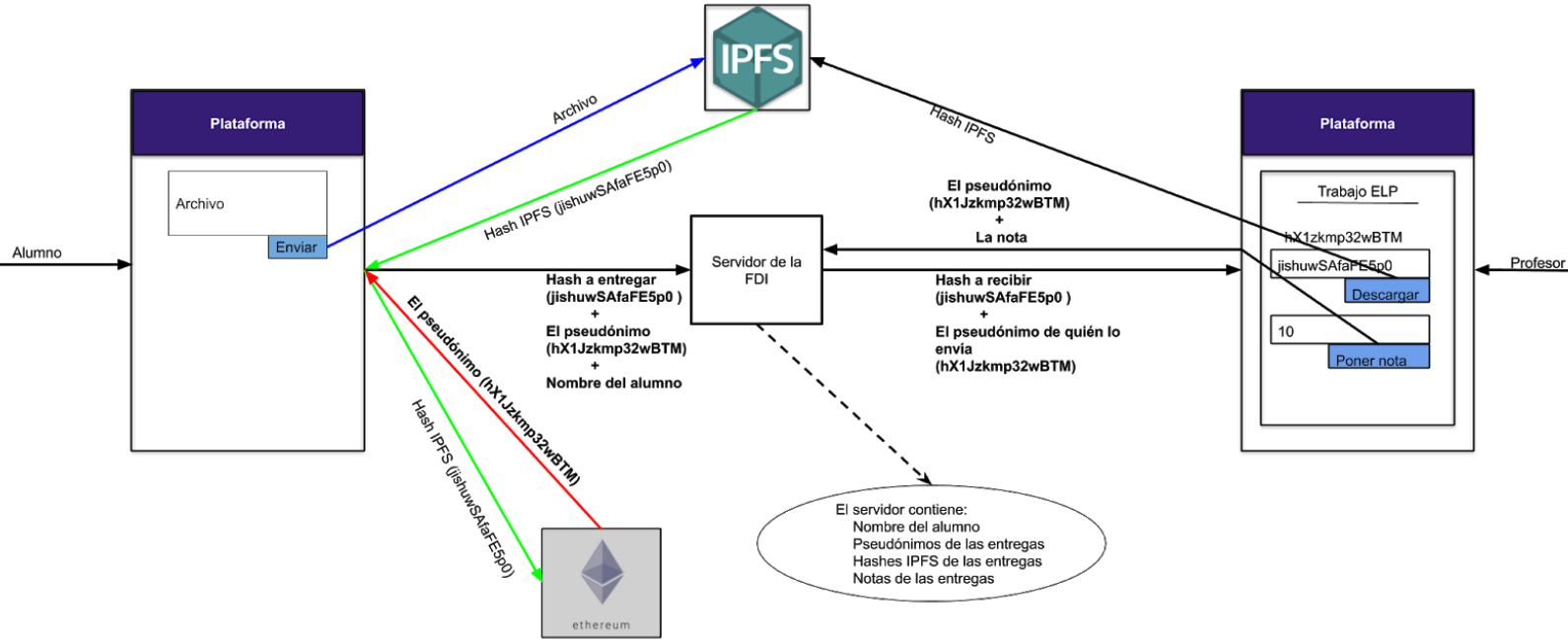
Nuestro trabajo consiste en el diseño, estudio económico y social de un sistema de corrección pseudo anónimo utilizando redes descentralizadas. Este diseño de entregas de prácticas, exámenes o cualquier otro documento que necesite una aprobación o corrección reduciría los sesgos utilizando los pseudónimos que proporcionan las tecnologías descentralizadas más todas las otras ventajas que ofrece. Además se añade la red IPFS para reducir la carga de servidor de ficheros.

En las siguientes páginas se explicarán dos diseños, uno en que se utiliza IPFS y Ethereum, y otro en el que se utiliza solo IPFS. Utilizando cualquiera de los diseños, se obtienen las siguientes **ventajas**, que se desarrollarán más adelante:

- Disminuye el coste de almacenamiento, ya que el servidor solo almacena hashes y no archivos enteros.
- Con los pseudónimos se evitan los sesgos en la corrección, tanto aquellos que aumentan la nota como los que la reducen.
- Con los hashes de IPFS se obtiene un sistema de anticopia simple, ya que si dos archivos tienen el mismo hash es el mismo archivo (aunque si se cambia una sola letra, ya se obtienen hashes distintos).
- IPFS es gratuito y garantiza integridad.
- Impacto social bueno para la FDI, ya que se incorporaría un sistema anti-sesgo, siendo una de las primeras universidades en utilizar tecnologías descentralizadas.

En relación a la parte de evitar sesgos, ya se utiliza en la actualidad. Por ejemplo, la EVAU es anónima para evitar favoritismos en la corrección. También se utiliza en la corrección de trabajos anónima en la Universidad de Oxford.

Diseño utilizando IPFS y Ethereum



Abstract

Comenzando por los datos que maneja el alumno, al hacer la entrega se le envía el archivo entregado a IPFS, obteniendo su hash correspondiente. Este hash deberá mostrarse al usuario para que pueda obtener el trabajo entregado.

A continuación, se hace una transacción en Ethereum con el hash IPFS, obteniendo así el pseudónimo utilizado en Ethereum, con lo que el alumno queda “enmascarado”. Este pseudónimo deberá mostrarse al alumno, de manera que pueda ir a una revisión si le resulta necesario. Además, la transacción en Ethereum se convierte en una prueba de que el trabajo se ha entregado. Para no aumentar más el coste de Ethereum, solo se permitirá una entrega, lo que supone una desventaja.

De esta forma, el servidor contiene el nombre real del alumno, los hashes IPFS entregados, los pseudónimos con los que se hicieron la entregas y, como veremos más adelante, las notas.

Con relación a los datos manejado por el profesor, obtendría los trabajos, que se han descargado de IPFS con los hashes de los alumnos, asociados al pseudónimo de la entrega, asegurando el pseudonimato del alumno. Tras corregir, deberá

subir las notas al servidor, donde se le asignará la nota al pseudónimo, y por transitividad, al nombre del alumno.

A la hora de calcular la nota final, todos los porcentajes deberán estar subido en el servidor a inicio de curso. De esta manera, cuando se haya realizado el examen y el profesor haya subido las notas (asociadas a un pseudónimo, si el examen es digital, o al nombre real, si es en físico), el propio servidor calculará la nota media de la asignatura. También se podría añadir una opción adicional de añadir notas que se suman a la media final, para trabajos opcionales hechos por el alumno.

En caso de que el sistema anti-copia encuentre una copia, se le podría pedir a alguien con acceso al servidor (nunca un profesor) el nombre real del alumno, para realizar las penalizaciones pertinentes.

Casos de uso

1. Alumno realiza entrega individual.

El alumno hace una entrega de una forma parecida a cuando se usa el Campus Virtual. Cuando el trabajo ha sido corregido, y la nota puesta, también se mostrará la nota.

2. Alumno realiza entrega de grupo.

En este caso el alumno accederá al nuevo tipo de entrega en el que tendrá que especificar el nombre de sus compañeros mediante un método para hacerlo cualquiera. Al hacer la entrega se devuelve un pseudónimo de Ethereum, el cual se utilizará de pseudónimo para cada uno de los componentes del grupo.

El resto del caso de uso es igual que en el punto anterior solo que la nota puesta por el profesor al pseudónimo del trabajo se le asigna a los compañeros del trabajo.

3. Profesor corrige y pone nota a la entrega.

El profesor accede a la entrega y ve los pseudónimos junto a sus archivos (no los hashes IPFS, eso lo hace la aplicación). Tras corregirlo, le pone una nota y lo envía. Esta nota puede ser modificada en cualquier momento, ya que puede ser que tras una revisión se vea un fallo en la corrección de la entrega.

4. Alumno va a revisar su entrega con el profesor.

El alumno necesitará llevar su pseudónimo a la revisión, por lo que en ese momento, para esa entrega, queda identificado. Una posible solución sería un sistema de corrección que preserve el anonimato online y dejar como última instancia la revisión en persona.

5. Alumno va a comprobar su entrega.

El alumno podrá descargarse en cualquier momento la entrega que estará subida a IPFS.

Coste

IPFS es gratis, a cambio de almacenamiento. Sin embargo, Ethereum sí que supondría un coste adicional. Ahora analizaremos ese coste:

Suponemos:

- 1700 alumnos (fdi)[1]
- 10 asignaturas por curso
- 1 entrega por semana
- 32 semanas por curso

Total de entregas en un año escolar =1700 x (32 x 10) = 544.000 entregas
Coste por byte (ethereum) = 0.00152 €[2]

coste por hash usado en IPFS= 46 Bytes * 0.00152 € = 0.07€
coste entregas en la FDI en un año escolar = 1700 x 32 x 10 x 0.07€ = **38.920€**

Conclusiones

Ventajas:

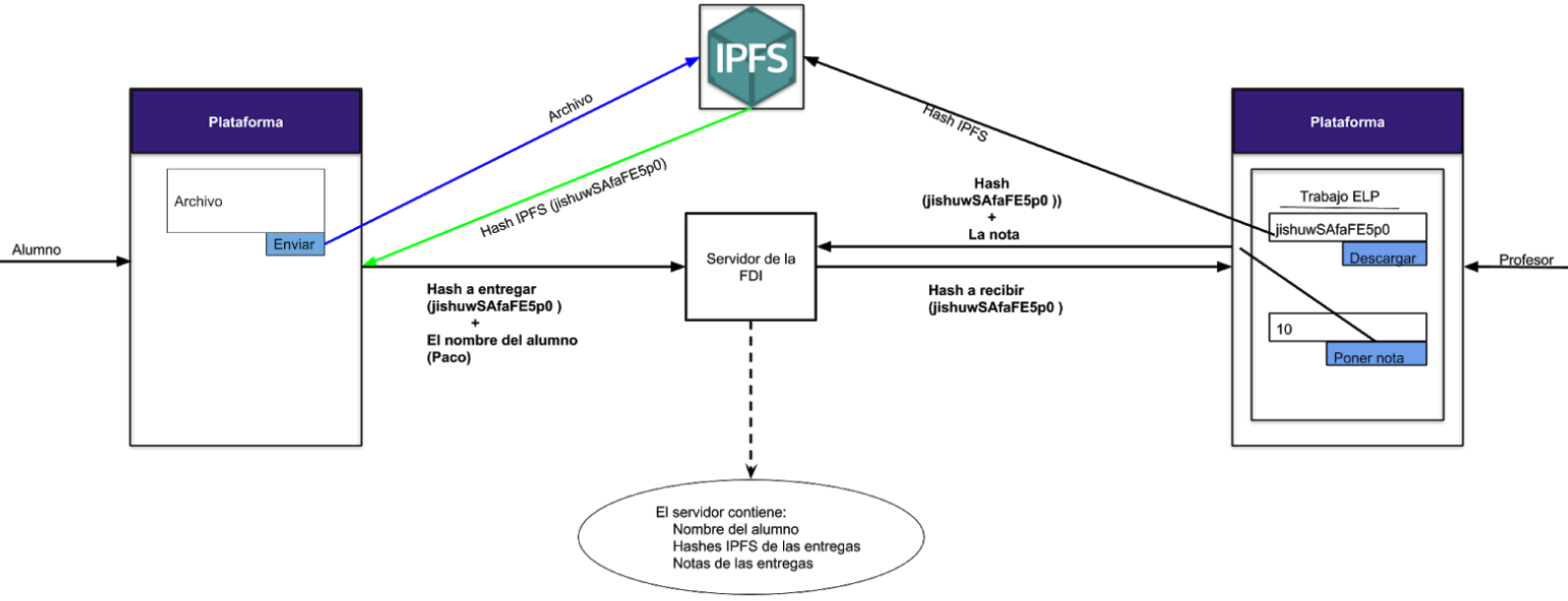
- El uso de Ethereum permite tener una prueba de que un trabajo ha sido entregado.

Inconvenientes:

- Coste anual de Ethereum (cota superior).
- Los alumnos no podrían modificar las entregas una vez hechas, para que no se dispare el coste de Ethereum.

[1] https://www.ucm.es/data/cont/media/www/pag-131427/E%2090.pdf
[2] https://medium.com/coinmonks/store-data-in-ethereum-by-logging-to-reduce-gas-cost-b70a13884485

Diseño de la plataforma con la red IPFS



Abstract

A diferencia con el diseño anterior, aquí se usarán los hashes devueltos por IPFS como entregas y como corrección. Debido al gran coste que tendría cambiar moodle para satisfacer esta condición, se creará otra plataforma ajena a este. Esta plataforma también dejará la opción de habilitar entregas por grupo o individual

El alumno subiría el archivo por medio de la nueva plataforma y este llegaría a una red descentralizada IPFS. La propia plataforma recibiría de la IPFS un hash del archivo que servirá para identificarlo unívocamente y la entregaría al servidor de la FDI junto al nombre de este. El servidor de la FDI tendrá registrado a qué alumno está asignado qué hash IPFS, el cual se utilizará como pseudónimo, pero el profesor que evalúe no será capaz de acceder a esta información.

El profesor accederá a esa misma plataforma para corregir las entregas y las evaluará con el pseudónimo antes mencionado. En ese momento, el servidor de la FDI se encargará de asociar esa nota asignada al pseudónimo con el correspondiente alumno.

Por último al alumno le aparecerá junto a la entrega la calificación que ha obtenido.

Casos de uso

1. Alumno realiza entrega individual

El alumno hace una entrega de la una forma parecida a cuando usa el Campus Virtual. En este caso (a diferencia del anterior), el alumno no tendrá solamente un intento.

El alumno subirá su archivo por medio de la nueva plataforma a la red IPFS, esta le devolverá un identificador hash del archivo. Junto con ese identificador se enviará el nombre del alumno, que será lo que almacene el servidor de la FDI.

Cuando se realiza la entrega, la zona donde el trabajo ha sido entregado es sustituida por el hash IPFS y cuando el trabajo ha sido corregido, y la nota puesta, también se mostrará la nota.

Una vez finalizado el plazo de entrega, el alumno no podrá subir más archivos a la red IPFS por lo que únicamente constaría la última entrega que este realizó.

2. Alumno realiza entrega en grupo

En este caso, el alumno accedería al nuevo tipo de entrega en el que tendrá que especificar el nombre de sus compañeros mediante un método para hacerlo cualquiera.

El resto de caso de uso es igual que en el punto anterior solo que la nota puesta por el profesor al hash del archivo también se asigna a los compañeros del trabajo.

3. Alumno modifica su entrega.

Como en IPFS no se pueden modificar archivos, en el caso que el alumno necesite modificar la entrega, tendrá que realizar una nueva. Para esto, aparecerá una opción que permitirá esto. En el servidor se sustituiría el último hash por el nuevo.

4. Profesor corrige y pone nota a la entrega

El profesor accede a la entrega y ve los hashes de los archivos junto a una opción de descargar fichero. Para la descarga del fichero la nueva plataforma accederá a la red IPFS y buscará el archivo con el correspondiente hash.

Tras corregirlo, le pone una nota a través de la plataforma. Esto hace que la plataforma envíe al servidor la nota para que le pueda aparecer al alumno. Esta puede ser modificada en cualquier momento, ya que puede ser que tras una revisión se vea un fallo en la corrección de la entrega.

5. Alumno va a revisar su entrega con el profesor.

El alumno necesitará llevar el hash del archivo a la revisión. Revelar la identidad expondrá al usuario, una posible solución sería un sistema de corrección que preserve el anonimato online y dejar como última instancia la revisión en persona.

6. Alumno va a comprobar su entrega.

El alumno podrá descargarse en cualquier momento la entrega que estará subida a IPFS.

Coste

No hemos podido hacer una estimación más precisa de los posibles costes a la hora de implementar este diseño pero hemos estimado en lo que, a priori, se podría ahorrar y en lo que no:

- La UCM podría ahorrar bastante en el almacenamiento en servidores ya que la red IPFS es una red descentralizada y gratuita que permite almacenar y borrar ficheros.
- Se seguirá utilizando la infraestructura actual de servidores que tiene la UCM pero con menos sobrecarga ya que no tendría que almacenar archivos.
- La creación de la nueva plataforma supondría un gasto.

Conclusiones

Ventajas:

- Disminuye el coste anual frente a la anterior propuesta.
- En caso de fallo de moodle, no habría problemas a la hora de subir entregas.
- Se reduciría la carga de almacenamiento de ficheros en los servidores de la universidad drásticamente.

Inconvenientes:

- Si al servidor le ocurre algún problema, no habría forma de asegurar que la entrega se ha realizado.
- Habría que implementar una nueva plataforma.

Visiones de futuro

- Se podría implementar una IPFS privada entre todas las universidades de España, de manera que se podría implementar un sistema anticopia muy simple a nivel nacional. Además, esto facilitaría la cooperación entre universidades, ya que simplemente hay que compartir los hashes de los archivos. Estos hashes podrían ser compartidos en cualquier momento y plataforma sin necesidad de una tediosa subida de archivos, siendo tan sencillo como copiando y pegando el hash.

- El desarrollo de un sistema de revisiones online que mantenga el anonimato como se comenta en los casos de uso.

- Creemos que la digitalización la mayoría de lo evaluable en la educación es inminente, esto daría más utilidad e importancia al sistema.

- Extender el sistema de entregas no sólo para entregas en la educación sino para el sector profesional.