Kleros

White Paper V. 1.0.7

Clément Lesaege, Federico Ast y William George

Septiembre 2019

Resumen

Kleros es una aplicación descentralizada construida sobre Ethereum, que funciona como una tercera parte descentralizada para el arbitraje de disputas en cualquier tipo de contratos, ya sean simples o muy complejos. Se basa en incentivos de teoría de los juegos para lograr que los jurados juzguen los casos correctamente. El resultado es un sistema de resolución de disputas que alcanza resultados de manera rápida, económica, confiable y descentralizada.

1 Introducción

"Quien controla las cortes, controla el estado." Aristóteles.

El mundo está experimentando un proceso acelerado de globalización y digitalización. Un número exponencialmente creciente de transacciones son realizadas online entre personas ubicadas en distintas jurisdicciones. Si la promesa de blockchain se hace realidad, en un futuro no muy distante, la mayoría de los bienes, el trabajo y el capital serán asignados a través de plataformas globales descentralizadas. Las disputas seguramente estarán a la orden del día. Los usuarios del eBay descentralizado reclamarán que los vendedores no les enviaron los bienes especificados en el contrato, los huéspedes del AirBnb descentralizado reclamarán que la casa alquilada no era como se mostraba en las fotografías, e inversores en una plataforma de financiamiento colectivos van a reclamar reembolsos en la medida en que los emprendimientos financiados no obtengan los resultados prometidos.

Los contratos inteligentes son lo suficientemente inteligentes como para ejecutarse automáticamente tal como fueron programados, pero no pueden realizar juicios subjetivos o incluir elementos que provengan de fuera del blockchain. Las tecnologías existentes de resolución de disputas son demasiado lentas, costosas y poco confiables para una economía global que funciona en tiempo real. Por ello, una institución clave en la era de blockchain es un mecanismo de resolución de disputas rápido, de bajo costo, confiable y descentralizado que permita resolver disputas en contratos inteligentes.

Kleros es un protocolo de toma de decisiones capaz de resolver cualquier tipo de disputa. Es una organización autónoma construida sobre el blockchain de Ethereum que funciona como tercera parte descentralizada para arbitrar disputas en cualquier tipo de contrato, ya sea simple o muy complejo. Cada paso del proceso de arbitraje (asegurar la evidencia, seleccionar jurados, etc.) se encuentra completamente automatizado y codificado en contratos inteligentes. Kleros no depende de la honestidad de unos pocos individuos que actúan como jurados, sino en incentivos económicos provenientes de la teoría de juegos.

Las raíces de Kleros se basan en un concepto fundamental de la epistemología legal: una corte es un motor epistémico, una herramienta para descubrir la verdad acerca de los hechos a partir de un conjunto de pistas confusas. Un agente (*jurado*) sigue un procedimiento donde un conjunto de datos (*evidencia*) se utiliza para producir un resultado (*decisión*) (20). Kleros se apoya en tecnologías de crowdsourcing, blockchain y teoría de juegos para desarrollar un sistema de justicia que produce decisiones verdaderas de una manera segura y a bajo costo.

2 Antecedentes: el Mecanismo de SchellingCoin

El experto en teoría de los juegos Thomas Schelling desarrolló el concepto de Punto Focal (también conocido como Schelling Point) (25) como una solución natural que las personas tienden a utilizar para coordinar su comportamiento en ausencia de comunicación.

Schellling ilustró el concepto con el siguiente ejemplo: “Si mañana te debes encontrar con un extraño en Nueva York. ¿Dónde y cuándo se encuentran?”

Cualquier momento y lugar es una solución posible, pero la respuesta más común es “al mediodía en el kiosco de información de la Grand Central Terminal”. No hay nada que haga que el mediodía en la Grand Central Terminal sea un momento y lugar mejor que otro (cualquier otro momento y lugar serían buenos si ambas personas se pusieran de acuerdo), pero su tradición como punto de encuentro lo convierte en un punto focal natural.

Los puntos focales típicamente surgen cuando la comunicación es imposible, pero también cuando, aunque la comunicación sea posible, ninguna de las partes puede dar una razón para justificar lo que dice (16). Basado en este concepto, el fundador de Ethereum, Vitalik Buterin, propuso la creación de la denominada SchellingCoin (9), una ficha criptográfica (*token*) que genera incentivos económicos para que un grupo de agentes diga la verdad.

Si queremos saber si llovió en París esta mañana, podemos preguntar a un conjunto de tenedores de la moneda SchellingCoin: “¿Llovió esta mañana en París? Sí o no”.

Cada tenedor vota en una elección secreta y después de que todos lo han hecho se revelan los resultados. Los que votaron como la mayoría son recompensados con un premio equivalente al 10% de sus monedas. Los que votaron de manera diferente a la de la mayoría pierden el 10% de sus tenencias.

Thomas Schelling escribió (25) “el punto focal es la expectativa de cada persona acerca de lo que los demás esperan que ella espera que que los demás hagan.” La SchellingCoin usa este principio para generar incentivos en un número de agentes que no se conocen o no confían en que los demás digan la verdad. Esperamos que estos agentes elijan la respuesta verdadera porque ellos esperan que los demás elijan una respuesta verdadera, a su vez porque aquellos esperan que otros elijan una respuesta verdadera… En este simple caso, el punto focal es la honestidad.

Mecanismos similares a la SchellingCoin han sido utilizados por oráculos descentralizados y en mercados de predicción (26) (23) (3). La idea fundamental es que votar coherentemente con los otros es un comportamiento deseable que debe ser incentivado.

El diseño de incentivos de Kleros está basado en un mecanismo similar al de la SchellingCoin, aunque ligeramente modificado para dar respuesta a desafíos específicos vinculados a la escalabilidad, subjetividad y privacidad del ámbito de la resolución de disputas.

La mayor parte de los votos \ tu voto	Sí	No
Sí	+0.1	-0.1
No	-0.1	+0.1

Figura 1. Matriz de pagos en un juego de Schelling sencillo.

3 Un Caso de Disputa

Alice es una emprendedora que vive en Francia. Ella contrata a Bob, un programador de Guatemala a través de una plataforma online para construir un nuevo sitio web para su empresa. Acuerdan el precio y los términos, y Bob se pone a trabajar. Dos semanas más tarde, Bob entrega el producto. Pero Alice no está satisfecha. Ella argumenta que la calidad del trabajo es considerablemente más baja de lo esperado. Bob responde que él hizo exactamente lo que estipulaba el contrato. Alice se siente frustrada. No puede

contratar a un abogado para entablar una demanda por un par de cientos de dólares contra alguien que está al otro lado del mundo.

¿Qué ocurriría si en el contrato hubiera una cláusula estableciendo que, en caso de desacuerdo, la disputa se resolvería mediante una corte de Kleros?

En este caso, Alice y Bob hubiesen instrumentado su acuerdo en una plataforma que permitiera a ella enviar un pago en criptomoneda que quedara en garantía hasta la finalización del trabajo (escrow.kleros.io). Bob puede ver que el dinero está en el fondo de garantía, con lo que no tiene que temer que hará su trabajo y luego no recibirá el pago. Alice, por su parte, se asegura que, en caso de que el trabajo sea de menor calidad, habrá un tercero neutral al que recurrir.

Al ver que Bob no reconoce los problemas en el trabajo realizado, Alice presiona un botón que dice "Enviar a Kleros" y completa un formulario explicando su reclamo.

Chief es un desarrollador de software que vive en Nairobi, a miles de kilómetros de distancia. En sus "tiempos muertos" en el bus que lo lleva a su trabajo, revisa el sitio web de las cortes de Kleros (court.kleros.io) buscando trabajo como jurado. Bob gana algunos cientos de dólares extra al año, participando como jurado en disputas de desarrollo de software, entre trabajadores independientes y clientes. Usualmente trabaja en casos dentro de la corte *Calidad de Sitios Web*. Este tribunal requiere conocimiento de html, javascript y diseño web para resolver los casos. Chief deposita una garantía de 2000 Pinakion (PNK). Ese depósito le da la posibilidad de ser elegido al azar como jurado y, gracias a ello, ganar dinero por honorarios de arbitraje. Mientras más tokens deposite, mayores son sus probabilidades de salir sorteado.

Unas horas después, llega un email a la casilla de Chief: "Has sido seleccionado como jurado en una disputa en la corte Calidad de Sitios Web. Descarga aquí la evidencia. Dispones de tres días para enviar tu decisión". Un email similar fue recibido por Benito, un programador de Cusco, y Alexandru, de Rumania, quienes también depositaron PNK en la corte de Calidad de Sitios Web. Fueron seleccionados aleatoriamente entre un conjunto de aproximadamente 3000 candidatos. Nunca se conocerán, pero colaborarán para resolver la disputa entre Alice y Bob. En el bus de regreso a su casa, Chief analiza la evidencia y emite su voto.

Pocos días después, cuando todos los jurados votaron, Alice y Bob reciben un email: "El jurado ha resuelto a favor de Alice. El sitio web desarrollado no cumple con los términos y condiciones acordados por las partes. Los fondos han sido reembolsados a Alice." Los jurados son recompensados por su trabajo y el caso queda cerrado.

4 Descripción del Proyecto

4.1. Contratos arbitrados

Kleros es un sistema de cortes digitales al cual las partes deben optar explícitamente para participar (opt-in). El sistema debe estar designado dentro de un contrato inteligente para actuar como árbitro de ese mismo contrato. Los creadores del mismo definen el tipo de corte y el número de jurados que actuarán, llegado el caso[1], y la idea es que elijan un tipo de corte especializada en el tema del contrato. Por ejemplo, un contrato de desarrollo de software seleccionará una corte de desarrollo de software, un contrato de seguros seleccionará una corte de seguros, etc.

La Figura 2 ejemplifica un árbol de cortes entre las que un usuario puede elegir. El equipo de Kleros ha desarrollado una serie de contratos estándar usando Kleros como mecanismo de resolución de disputas. Además, se han propuesto estándares (21) (27) que permitirían el desarrollo de contratos de manera tal que no sea necesario anticipar el mecanismo de resolución de disputas que será usado, en caso de surgir una disputa.

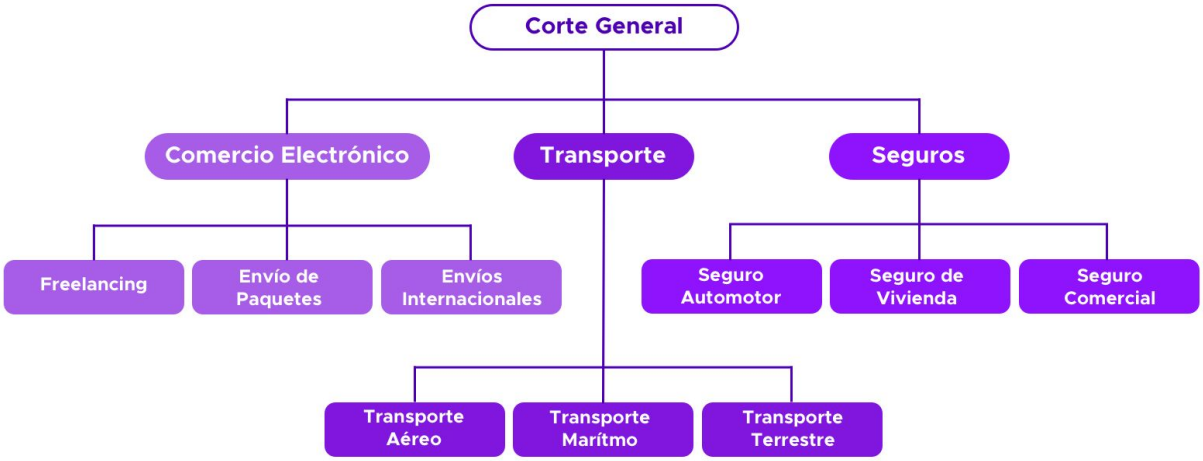

Figura 2. Ejemplo de árbol de cortes en las que se resuelven las disputas dentro de Kleros.

4.1.1. Opciones para los jurados

Los contratos especifican qué opciones de voto tienen los jurados. En el ejemplo anterior, las opciones serían: “Reembolsar a Alice”, “Asignar una semana adicional a Bob para que termine el trabajo” y “Pagar a Bob”.

El contrato inteligente también especifica el comportamiento del mismo posterior a la toma de decisión. En el ejemplo:

- “Reembolsar a Alice” transfiere los fondos a la dirección de Alice.

- “Asignar una semana adicional a Bob para que termine el trabajo” modifica en una semana el tiempo remanente del que dispone Bob para finalizar su trabajo. Además, el contrato inteligente podría estar escrito de manera tal que esta opción sólo pueda ser utilizada una vez. Es decir, ya no estaría disponible en caso de una futura disputa sobre ese contrato.
- “Pagar a Bob” transfiere fondos a la dirección de Bob.

4.2 Selección de Jurados

4.2.1. El Token: Pinakion (PNK)

Los usuarios persiguen su propio interés económico cuando optan por servir como jurados en Kleros: cobrar honorarios de arbitraje a cambio de su trabajo. Los usuarios se postulan a sí mismos como jurados depositando como garantía un token denominado Pinakion (PNK)[2] en la corte donde quieren trabajar.

La probabilidad que tiene un usuario de ser elegido como jurado para una disputa específica es proporcional a la cantidad de tokens que deposite en garantía. Mientras mayor sea dicho monto, más alta será la probabilidad de ser seleccionado. Aquellos jurados que no depositan PNKs en garantía, no tienen posibilidades de ser seleccionados. Esto evita que sean elegidos jurados inactivos.

El PNK tiene dos funciones claves en el diseño de Kleros. En primer lugar, protege al sistema contra un ataque de tipo Sybil (14). Si los jurados fueran elegidos aleatoriamente, un usuario malicioso podría crear un gran número de cuentas para ser seleccionado varias veces en cada disputa. Si consiguiera ser elegido más veces que los jurados honestos, el usuario deshonesto controlaría el sistema.

En segundo lugar, el PNK brinda a los jurados un incentivo para votar honestamente[3]: hace que los jurados incoherentes, o sea aquellos cuyos votos no coinciden con la mayoría, paguen parte de su depósito de garantía a los jurados coherentes.

4.2.2 Selección del Jurado

Una vez que los candidatos se han postulado depositando sus tokens en la corte para la que quieren trabajar, la selección final se realiza por sorteo. La probabilidad de ser seleccionado como jurado es proporcional a la cantidad de tokens depositados. Teóricamente, un candidato podría ser elegido más de una vez para una disputa específica, pero en la práctica esto es improbable. La cantidad de veces que un candidato es elegido para una disputa (su *peso*) determina el número de votos que tendrá en dicha disputa y la cantidad de tokens que ganará o perderá durante la redistribución de los mismos.

Imaginemos que 6 candidatos se anotaron en una disputa y depositaron 10.000 PNK en total con la siguiente distribución:

Candidato a jurado	Depositados	Inicio	Fin	Peso
A	1000	0	999	0
B	1500	1000	2499	1
C	500	2500	2999	1
D	3000	3000	5999	2
E	1500	6000	7499	0
F	2500	7500	9999	1

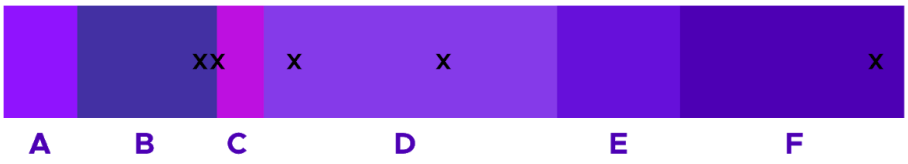

Figura 3:

Ejemplo de tokens depositados y elección de miembros del jurado.

En una disputa que requiere 5 votos, se extraen 5 tokens de los 10.000 que han sido depositados. Los tokens extraídos (representados en la Figura 3) son los identificados por los números 2519, 4953, 2264, 3342 y 9531. Los candidatos B, C y F son elegidos con un peso igual a 1. El candidato D es elegido con un peso igual a 2. Los PNK depositados (excepto aquellos perdidos por jurados incoherentes) pueden ser retirados una vez que el tribunal haya alcanzado una decisión final.

4.2.3. Generación de números aleatorios

La selección de jurados debe realizarse a través de un proceso de generación de números aleatorios resistente a la manipulación. El protocolo de generación de números aleatorios entre dos partes (5) no es una buena alternativa. Un atacante podría crear disputas contra sí mismo, elegirse como jurado

varias veces y seleccionar a otro jurado como víctima. Luego, podría coordinar sus propios votos de manera tal de ser considerado coherente, dejando a la víctima (el otro jurado) como incoherente y robando así sus tokens. (ver la sección Sistema de Incentivos).

En la versión actual de Kleros, los números aleatorios son tomados de los *blockhashes* de los bloques de la red Ethereum. Aunque tales números resultan imposibles de predecir de antemano, los mineros podrían optar por no liberar un bloque que resulte en números aleatorios desfavorables para ellos (al costo de perder la recompensa ligada al bloque). En el futuro, para producir números aleatorios que no puedan ser manipulados ni siquiera por mineros de gran capacidad, los números serán generados mediante el método Sequential-Proof-of-Work (12) usando un esquema similar al de Bünz et al. (13) (Véase la Sección Trabajo a Futuro).

4.3 Votos

Después de evaluar la evidencia, los jurados emiten (7) su voto. Esto se hace mediante el envío de un hash(vote, salt, address)[4]. Salt es un valor aleatorio generado localmente para aumentar la entropía. Su objetivo es evitar el uso de tablas rainbow. Address es la dirección Ethereum del jurado. Es necesaria para que el hash enviado por cada jurado sea diferente, de manera tal que un jurado no pueda copiar el voto de otro.

Cuando la votación finaliza, los jurados revelan su {vote, salt} y un contrato inteligente verifica que coinciden con los enviados previamente. Aquellos jurados que no revelan su voto son penalizados (véase Sección Sistema de Incentivos).

Una vez que un jurado ha comprometido su voto, el mismo no puede ser modificado. Pero éste no es aún visible para los otros jurados ni para las partes en litigio. Esto evita que un jurado pueda influir en los votos de los demás.

Aunque un jurado pueda declarar haber votado de determinada manera, es difícil que pueda dar razones para que los demás crean que está diciendo la verdad. Esta es una propiedad clave del Punto Focal que debe ser destacada. Si los jurados supieran el voto de los demás, podrían copiar su voto en lugar de votar pensando en el Punto Focal. (Véase la Sección Trabajo a Futuro para una discusión más profunda de estas ideas.)

Una vez que todos los jurados han votado (o cuando el tiempo para hacerlo ha finalizado), los votos son revelados. Aquellos que no hayan revelado su voto son penalizados. Finalmente, los votos se suman, y el contrato inteligente es ejecutado. La opción con la mayor cantidad de votos es considerada ganadora [5].

Este procedimiento, conocido en la jerga criptográfica como *commit and reveal*, requiere interacciones adicionales. En algunas cortes, en las cuales hay escasos aportes de PNK, puede ser preferible que los votos sean emitidos públicamente para simplificar la experiencia del usuario [6]. El sistema a utilizar

puede ser definido a través de un parámetro relativo a la corte. Algunas cortes utilizarán el sistema de commit and reveal y otras no (ver la Sección Mecanismo de Gobierno).

4.4. Honorarios de Arbitraje

Para compensar a los jurados por su trabajo y evitar que un atacante sature el sistema, la creación de disputas y la apelación requiere el pago de una tarifa de arbitraje. Los jurados que voten de manera coherente con el resultado final recibirán un pago determinado por la corte donde se trata la disputa. Un contrato inteligente especificará qué parte debe pagar dicha tasa.

Las reglas pueden ser simples. Por ejemplo, podrían requerir que la tasa sea pagada por la parte que crea la disputa o la que apela. Pero también podría haber reglas más complejas que permitan mejorar los incentivos. Por ejemplo:

- En primera instancia, cada parte debe depositar un monto igual a la tarifa de arbitraje en un contrato inteligente. Si una de las partes no lo hace, el contrato considerará que la corte decidió en favor de la parte que sí depositó la tarifa (sin que sea necesario crear una disputa efectivamente). Si ambas parte depositan los fondos, la parte ganadora será reembolsada cuando finalice el proceso.
- En instancia de apelación, ambas partes deben depositar las tarifas de arbitraje. La parte que apela debe además depositar una cantidad adicional proporcional a la tarifa de apelación que será entregada a la parte ganadora de la disputa. De esta manera, si una parte realiza apelaciones infundadas con el objetivo de dañar al oponente o demorar la decisión final, éste obtendrá una compensación por el tiempo perdido. En cambio, si la apelación resulta finalmente válida, dicho monto será devuelto a la parte que realizó la apelación [7].

Una discusión más en profundidad de las tarifas será objeto de futuras investigaciones.

4.5 Apelaciones

Si, después que el jurado ha resuelto en un caso, una de las partes no está satisfecha (porque cree que el resultado no fue justo), tiene la posibilidad de apelar la decisión. Cada nueva instancia de apelación tiene el doble de jurados que la instancia anterior, más uno. Debido al mayor número de jurados, las tarifas de apelación que tendrán que ser pagadas se calculan como:

*Tasa de Apelación = Cantidad de Jurados Nuevos * Tasa Medida por Jurado.*

Dado que la tarifa de apelación se paga a un número de jurados que crece exponencialmente cada vez que uno apela, la misma también crece exponencialmente en función del número de apelaciones. Esto significa que, en la mayoría de los casos, las partes no apelarán, o solo lo harán un número acotado de

veces. De cualquier forma, la posibilidad de apelar muchas veces es importante para prevenir que los jurados sean sobornados (Véase la sección Resistencia al Soborno).

4.6 Sistema de Incentivos

Recuérdese que se considera que un miembro del jurado ha votado de forma coherente si su voto coincide con el de la mayoría. Los jurados trabajan en disputas con el objetivo de recolectar tarifas de arbitraje. Son incentivados para votar honestamente ya que, una vez que la disputa ha finalizado, aquellos cuyo voto no es coherente no recibirá su tarifa de arbitraje y además perderá los tokens bloqueados como garantía. En tal caso, tanto las tarifas de arbitraje que hubiera recibido como el depósito de tokens son entregados a los jurados que han votado de manera coherente.

Una vez que Kleros alcanzó una decisión en una disputa, los tokens se descongelan y redistribuyen entre jurados. El mecanismo de redistribución está inspirado por la SchellingCoin[8], donde los jurados ganan o pierden tokens dependiendo de si su voto fue consistente con el de los otros jurados.

La cantidad de tokens perdidos por un jurado incoherente es: $\alpha \cdot \textit{depósito minimo} \cdot \textit{peso}$

El parámetro α determina el número de tokens que se redistribuirá después de un caso. Es una variable endógena que puede definirse a través del mecanismo de gobierno como consecuencia de la dinámica interna del ambiente de votación. El parámetro *depósito mínimo* es la mínima cantidad de tokens que pueden ser depositados como garantía en una corte.

La tasa de arbitraje y los tokens perdidos se reparten entre los jurados coherentes proporcionalmente a su peso[9]. Un ejemplo de redistribución de tokens se muestra en la Figura 4. Para desincentivar que algunos jurados no revelen su voto, la penalidad por no hacerlo es al menos tan grande como la penalidad por votar de manera incoherente. De esta manera, hay un incentivo para revelar siempre el voto. En caso de apelación, las tarifas de arbitraje y los tokens se redistribuyen a cada nivel de acuerdo al resultado final de la instancia de apelación[10].

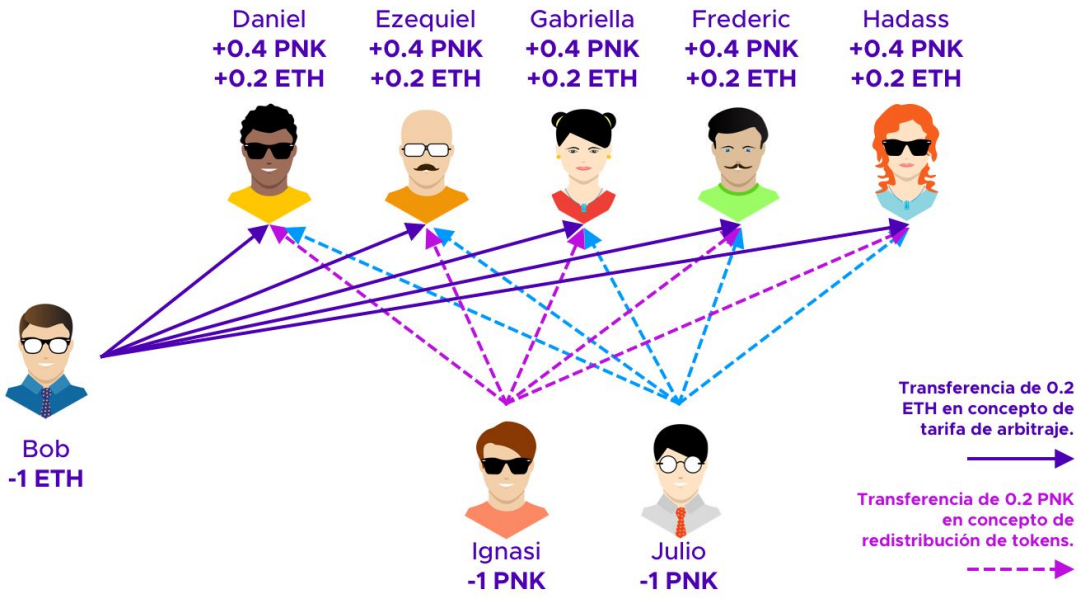

Figura 4. Redistribución de tokens después de una votación con 7 jurados. Los tokens de los jurados incoherentes se redistribuyen hacia los jurados que votaron de manera coherente. Bob perdió la disputa y paga la tarifa de arbitraje. Los demás depósitos son reembolsados.

Cuando no hay un ataque, los participantes son incentivados a votar como piensan que otros participantes piensan, que otros participantes piensan…. que es honesto y justo. En Kleros, el Punto Focal es la honestidad y la justicia.

Se podría argumentar que, siendo las decisiones subjetivas (al menos en comparación con el mecanismo SchellingCoin que se utiliza en mercados de predicción), no existirá un Punto Focal. En (25), los experimentos informales realizados por Thomas Schelling muestran que la mayoría de las situaciones no tienen un punto focal plebiscitado por la mayoría de los participantes.

Sin embargo, Schelling encontró que hay opciones que es más probable que sean elegidas que otras. Por lo tanto, incluso si una opción particularmente obvia no existiese, algunas opciones serán percibidas como más probables de ser elegidas por los otros participantes, y por lo tanto serán efectivamente seleccionadas.

Por supuesto, no es posible esperar que las respuestas de los jurados sean correctas el 100% de las veces. Ningún procedimiento de arbitraje puede alcanzar ese objetivo. A veces, jurados honestos perderán tokens. Pero, en la medida en que, al final, pierdan menos que lo que ganan como tarifas de arbitraje y depósitos de otros jurados incoherentes, el sistema funcionará.

4.7. Resistencia a Ataques

4.7.1. Ataque de 51%

Si un participante (o un grupo) comprara la mitad de los tokens, controlaría los resultados en la Corte General y, por lo tanto, podría decidir todos los resultados. De todos modos, que alguien compre más de la mitad de los tokens es altamente improbable si los mismos son distribuidos de manera justa.

Primero, la mitad de los tokens deberían estar a la venta, lo cual no está garantizado. Además, el hecho de que un participante pueda pagar por todos los tokens a precio de mercado no significa que pueda comprar la mitad. Los tokens, a diferencia de la mayoría de los activos físicos, tienen costo marginal creciente. Su precio es asignado dinámicamente en sitios de intercambio. Si un participante comprara una porción significativa de tokens, el precio subiría ya que la escasa profundidad de mercado haría cada vez más costoso adquirirlos.

4.7.2. Resistencia al Soborno

Las instancias de apelación constituyen un mecanismo importante contra sobornos. Sobornar a un jurado pequeño es relativamente fácil. Pero como la víctima siempre tiene el derecho de apelar, el atacante tendría que perpetrar sobornos cada vez más grandes con costos crecientes. El atacante tendría que gastar una enorme cantidad de dinero para sobornar jurados hasta llegar a la Corte General (e igualmente podría perder el caso). Para tener la certeza de controlar el veredicto, el atacante necesitaría sobornar a un conjunto de personas cuyas tenencias sean mayores del 50% de los PNK en circulación.

Este ataque no funciona en el modelo de mayoría honesta (donde más de la mitad de los tokens son controlados por participantes honestos que no aceptarían un soborno). Pero incluso si la mayoría fuese deshonesta (la mayor parte de los participantes solamente buscan optimizar su ganancia, incluso aceptando sobornos), el sistema puede soportar ataques por soborno bajo ciertas condiciones.

Un soborno exitoso de la Corte General provocaría un descenso dramático del valor del PNK (¿quién querría que sus contratos sean arbitrados por una corte deshonesta?). Entonces, para que una propuesta de soborno sea aceptada, el atacante debería ser capaz de ofrecer un valor mayor a la caída esperada en el precio (que en casi todos los casos excede el valor puesto en juego en la disputa). En la práctica, un caso en que una parte apele todas las decisiones hasta llegar a la Corte General es extremadamente improbable. Pero, aun así, la posibilidad debe existir para que los incentivos estén correctamente balanceados.

Es posible realizar un ataque más elaborado (el ataque p + épsilon), que promete pagar el soborno únicamente en caso de que el ataque no sea exitoso. Este tipo de ataque requiere un gran presupuesto, pero si es exitoso su costo es cero. En (11) se propone una respuesta teórica basada en teoría de juegos

contra este tipo de ataque, en la cual los jurados usan una estrategia mixta. Además, se han realizado experimentos en la prueba piloto “Doges on Trial”, en el cual se analizó el comportamiento de los usuarios en un evento de ataque del tipo p + épsilon. Véase en (17) el resultado de dichos experimentos, así como algunos comentarios sobre cómo el sistema de apelaciones de Kleros hace que dichos ataques sean menos viables.

4.8 Árbol de Cortes

Cuando alguien se registra como jurado, comienza en la Corte General y avanza hacia cortes especializadas de acuerdo a sus habilidades. Cada corte presenta rasgos específicos respecto de sus políticas, tiempo de las sesiones, costo, número de jurados a elegir y cantidad de tokens a depositar como garantía. Cuando un usuario deposita su token en una corte especializada, también puede ser elegido como jurado en todas las cortes que se encuentran por encima de ella.

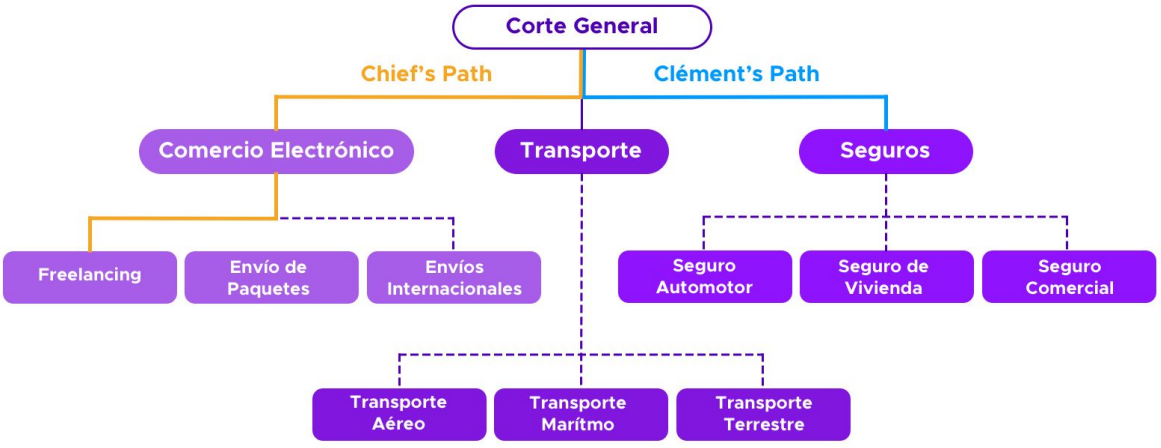

Figura 5. Ejemplo de los caminos elegidos por jurados en el sistema de cortes. Clément puede ser elegido como jurado en la Corte General y en la Corte de Seguros. Chief puede ser elegido como jurado en la Corte General, en la Corte de Comercio Electrónico y en la de Trabajo Freelance.

La necesidad de votar de manera coherente incentiva a los usuarios a elegir las cortes para las que tienen mayores habilidades. Participar en una corte para la que se tienen las habilidades adecuadas aumenta la probabilidad de votar de manera coherente con la mayoría y ganar honorarios de arbitraje.

4.9 Mecanismo de Gobierno

En la medida en que Kleros adquiera usuarios y casos de uso, será necesario crear nuevas cortes, hacer cambios en las políticas y parámetros de las mismas, y actualizar la plataforma mediante nuevas versiones que incorporen características adicionales. Tales decisiones serán realizadas por los

tenedores de tokens usando un mecanismo de votación líquida (15). Los mismos tendrán un número de votos igual a la cantidad de PNK que posean. El mecanismo de gobierno puede ser utilizado para:

1. Establecer políticas: éstas son indicaciones acerca de cómo arbitrar disputas. Son el equivalente a las leyes en un sistema de justicia tradicional. Determinan qué parte debería ganar una disputa cuando se dan ciertas condiciones particulares.

2. Crear nuevas cortes.

3. Modificar parámetros de las cortes:

(a) Tarifas de arbitraje

(b) Tiempo de sesión de cada corte.

(c) Mínima cantidad de tokens que deben ser depositados en garantía.

4. Realizar modificaciones en los contratos inteligentes de Kleros. Esto podría ser utilizado para realizar mejoras o en una emergencia si algún elemento de Kleros no está funcionando apropiadamente[12].

5. Trabajo a Futuro

En esta sección, discutiremos una serie de mejoras futuras al protocolo.

5.1. Privacidad de los Contratos

La resolución de disputas puede requerir que las partes compartan información sensible con los jurados. Para prevenir que observadores externos accedan a esta información, en el futuro, ni los contratos en lenguaje natural (inglés u otros) ni los rótulos de las opciones de voto serán publicados (y en particular no serán colocados en la blockchain). Cuando se cree el contrato, el creador enviará un hash(contract_text, option_list, salt), donde contract_text es el texto en inglés del contrato, option_list es la lista de rótulos de las opciones de voto disponibles para los jurados, y salt es un número aleatorio para evitar el uso de tablas rainbow.

El creador del contrato enviará {contract_text, option_list, salt} a cada parte usando encriptado asimétrico. De esta manera, las partes pueden verificar que el hash enviado se corresponde con lo que fue enviado a ellos. En caso de una disputa, cada parte puede revelar {contract_text, option_list, salt} a los jurados que podrán verificar que corresponde con el hash enviado. Esto lo pueden hacer usando encriptado asimétrico de manera tal que solo los jurados reciben el texto de los contratos y el de las

opciones. Estos pasos serán gestionados por la aplicación que los usuarios deberán correr cuando usen Kleros.

5.2. Mejoras en la Generación de Números Aleatorios

Como se explicó más arriba, actualmente los jurados son elegidos mediante el uso de un generador de números aleatorios basado en los blockhashes de bloques de Ethereum. Esto tiene el problema de que mineros con gran capacidad pueden sesgar la selección de jurados (al costo de perder su recompensa por la creación de bloques).

Esta sección explica en detalle los planes de crear una fuente más segura de aleatoriedad, basada en la Sequential-Proof-of-Work (12), usando un esquema similar al de Bünz et al. (13) adaptado para trabajar para blockchains Proof-Of-Stake de la siguiente manera[13]:

1. **Inicialización:** Se comienza con una semilla seed=blockhash y se permite a todos los participantes ingresar un valor aleatorio local (localRandom) para cambiar la semilla de manera tal que seed=hash(seed,localRandom). De esta manera, cualquier participante puede cambiar la semilla, pero ninguno puede elegirla, lo cual es deseable porque la elección de un seedAttack particular requiere que el atacante determine un localRandom tal que el hash(seed, localRandom)=seedAttack. Esto es difícil de lograr debido a la resistencia preimagen de las funciones criptográficas hash.

2. **Cálculo del valor aleatorio maestro**: cada participante que haya depositado fondos corre una sequential-proof-of-work con la semilla. Comenzando con h_0=seed, se calcula h_{n+1}=hash(h_n) hasta h_d donde d es un parámetro de dificultad. El cálculo de h_d n veces toma tiempo y asegura que una cantidad de tiempo transcurre entre el momento en que alguien toma conocimiento de la semilla y el momento en que obtiene el resultado. El valor de dificultad d está fijado de manera tal que ningún hardware puede calcular h_d durante el tiempo que tarda la fase de inicialización. Debido a que es necesario el resultado de la iteración previa antes de comenzar la siguiente, este proceso no puede ser paralelizado. Esto significa que nadie será capaz de obtener el resultado de manera significativamente más rápida que los demás.

3. **Obtención del resultado sobre blockchain:** Todos los participantes pueden postear el h_d realizando un depósito. Entonces, otros participantes pueden rechazar los resultados erróneos mediante verificación interactiva (24). La misma consiste en una búsqueda dicotómica de los resultados del atacante. Si un atacante envía un h_d falso, un participante honesto puede preguntarle su valor $h_{d/2}$. Si él provee un valor erróneo, existe un error en los valores del atacante entre h_0 y $h_{d/2}$. Si él devuelve el valor correcto, hay un error entre $h_{d/2}$ y h_d. De cualquier forma, el espacio de búsqueda está dividido en dos. El participante honesto continúa el proceso en un espacio reducido (donde está el error) hasta que quedan dos valores. Entonces, el participante honesto puede exhibir x tal que $h_{x+1} \neq$hash(h_x) en la respuesta del atacante, lo cual invalida su respuesta. Los participantes cuya respuesta es invalidada

pierden su depósito. Parte del mismo es destruido y la otra parte entregada al participante que lo invalidó. Nótese que el número de interacciones requeridas para invalidar un resultado falso es solamente O(log(d)).

4. **Obtención de todos los valores aleatorios:** Una vez que los participantes honestos han invalidado los resultados, solo hay un resultado correcto h_d restante. De este número aleatorio maestro se derivan todos los números aleatorios tales que r_n=hash(h_d,n).

El resultado de este proceso es un número aleatorio siempre que exista, al menos, un participante honesto. El cálculo de la sequential-proof-of-work, y el proceso de verificación interactiva toman tiempo. Sin embargo, para la mayoría de las disputas, no es un problema que haya una espera de algunas horas entre el momento en que la misma comienza y el momento en que los jurados son seleccionados.

En algunas cortes con tiempo de sesión particularmente bajo (por ejemplo, una corte que resuelva disputas en eventos deportivos casi en tiempo real) este método de generación de números aleatorios podría resultar demasiado lento. Otro posible mecanismo, que es menos seguro, pero que podría ser utilizado por dichas cortes es uno basado en threshold signatures (6).

5.3. Penalización de Jurados que Revelen su Voto Antes de Tiempo

Anteriormente se describió el mecanismo de *commit and reveal* que permite a los jurados mantener su voto oculto hasta que todos los votos han sido emitidos. Este esquema no evita que los jurados puedan revelar sus votos antes de tiempo con la intención de influir sobre los otros jurados. En esta sección se discute una potencial mejora futura que incentivaría a los jurados a no revelar sus votos antes del momento indicado.

Se propone permitir que cualquier participante que sea capaz de revelar a Kleros el voto emitido por un jurado antes de que la votación sea cerrada, pueda tomar los PNK de dicho jurado e invalidar su voto. Entonces, si un jurado quiere revelar su voto a otro participante, tiene dos opciones:

1. Revelar solo su voto. El otro participante no tiene ninguna prueba de que efectivamente votó de esa manera. El jurado podría mentir y el otro participante no tendría manera de verificarlo.

2. Revelar su voto y mostrar su compromiso (commit). El otro participante tendría prueba de su voto, pero también estaría en condiciones de robar sus PNK.

Este mecanismo impedirá que los jurados revelen sus votos[14]. Una potencial solución, más completa, que está bajo análisis es la detallada en (10).

5.4. Votación Líquida en Gobierno

En la sección acerca de mecanismos de gobierno, se describió cómo los tenedores de tokens pueden tomar decisiones sobre la plataforma. Aquí describimos un plan a futuro que permite a los tenedores de tokens la delegación de su voto si los mismos no desean votar directamente. Cuando un usuario no vota, su derecho a voto es automáticamente transferido a su delegado.

En la figura 6 puede verse una ilustración del mecanismo de votación líquida. La delegación del voto puede hacerse también específica de una subcorte. Los usuarios pueden elegir delegar su voto en algunas subcortes pero no en otras. Nótese que los delegados no precisan ser personas humanas. Pueden ser contratos inteligentes que implemente reglas de votación arbitrariamente complejas (por ejemplo votar una actualización de tarifas basada en datos de mercado).

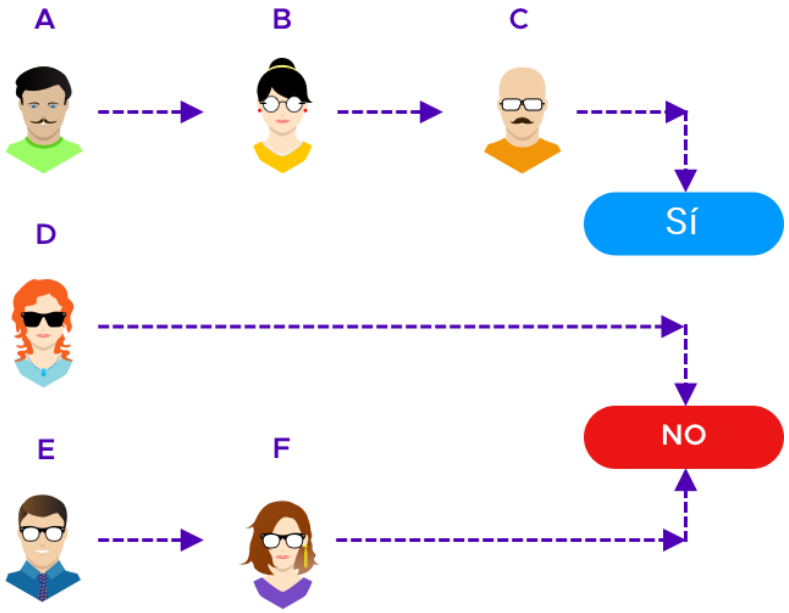

Figura 6. Ilustración del voto líquido.

6. Aplicaciones

Kleros es un sistema general y multipropósito que puede ser usado para resolver un gran número de situaciones. Aquí presentamos algunos ejemplos de posibles casos de uso:

- Depósito de garantía: el pago de un bien o servicio puede ser hecho mediante fondos depositados en un contrato inteligente. Una vez recibido el bien o servicio, el comprador desbloquea los fondos para que el vendedor pueda obtenerlos. En caso de desacuerdo, Kleros puede ser utilizado para que el contrato inteligente reembolse al comprador o pague al vendedor. Éste sistema de depósitos en garantía basado en Kleros ya está en funcionamiento, véase (18)

- Micro tareas: las plataformas descentralizadas podrían pagar por la realización de microtareas (similar a la manera en que funciona Amazon Mechanical Turk (1)). Aquellos que realizan las tareas harían un depósito de seguridad y realizarían las mismas, que podrían ser replicadas por otros. Si los resultados de una misma microtarea son diferentes, los que la realizaron podrían admitir el error, lo cual haría que una parte de su depósito de seguridad sea transferido hacia quienes llevaron a cabo la tarea correctamente. En el caso en que varios de los trabajadores se mantengan en su posición, se aplica un procedimiento de resolución de disputas y se transfiere una parte del depósito de los perdedores hacia los ganadores.
- Seguros: el asegurado pagaría una tarifa al asegurador para obtener una compensación en caso de que ocurra un evento en particular. El asegurador debe depositar una cantidad, que podría ser común a varios asegurados (respetando reglas de gestión de riesgo). Cuando el evento ocurre, el asegurador puede validarlo y compensar al asegurado. Si el asegurador no valida el evento, se aplica un procedimiento de resolución de disputas. Si el asegurado gana la disputa, los fondos del depósito del asegurador se le transfieren. En el caso de que dicho depósito esté vinculado a varios asegurados que reclaman más de lo que hay en el mismo, se aplica un proceso de resolución de disputas para determinar cómo se distribuyen los fondos.
- Oráculo: una flujo de datos descentralizados a ser usada por contratos inteligentes fue uno de los primeros casos de uso previstos para Ethereum (8). Un participante (que puede ser un contrato inteligente) hace una pregunta. Cualquiera puede realizar un depósito y enviar una respuesta. Si todos dan la misma respuesta, la misma es devuelta por el oráculo. Si hay respuestas múltiples, se aplica un proceso de resolución de disputas. El oráculo devuelve la respuesta entregada por dicho mecanismo y los participantes que contestaron erróneamente pierden su depósito, que es transferido a quienes contestaron correctamente. Realitio brinda un servicio de oráculo que está basado en estos principios, con la opción de usar Kleros para realizar disputas (4). Además, otras aplicaciones que usen el oráculo de Realitio, tal como CryptoUnlocked (22), indirectamente dependen de la resolución de disputas.
- Curación de listas: Kleros puede utilizarse para organizar listas, blancas o negras. Por ejemplo, una lista blanca puede listar contratos inteligentes que han sido sometidos a procedimientos de auditoría apropiados. Una lista negra puede listar los ENS (Ethereum Name Service (2)) registrados por participantes que no tienen relación con el nombre. Por ejemplo, un participante malicioso podría registrar “kleros-token-sale.eth”, para engañar a la gente que envía fondos a esa dirección. Los participantes pueden enviar ítems a la lista realizando un depósito de seguridad. Si nadie impugna el ítem durante un tiempo, el nombre es agregado y el depósito reembolsado. Si algún participante impugna el ítem realizando un depósito de seguridad, se aplica un procedimiento de resolución de disputas. Si el ítem es considerado como perteneciente a la lista, se agrega a la misma y quien lo envió obtiene los depósitos de los oponentes. Si es al contrario, el depósito de quien envió el ítem se transfiere a los oponentes. Kleros está siendo usado para una lista curada de tokens que satisfacen varias propiedades (por ejemplo, que sean ERC20) (19)
- Redes Sociales: la prevención del spam, engaños y otros abusos son un desafío para las redes sociales descentralizadas. Los participantes pueden reportar violaciones a las normas de las redes y

realizar un depósito de seguridad. Si la violación es rechazada, se aplica un proceso de resolución de disputas. Si se resuelve que no hubo violación, quien la reportó pierde el depósito de seguridad que es transferido al participante acusado. Si la violación no es rechazada, o es confirmada por Kleros, pueden implementarse distintas soluciones: el contenido puede ser removido, quien lo envió puede perder un depósito de entrada y la jerarquía de sus otros envíos puede ser reducida.

7. Conclusión

Este artículo ha presentado Kleros, un sistema de cortes descentralizado que permite el arbitraje de contratos inteligentes mediante jurados seleccionados mediante crowdwourcing, que se apoya en incentivos económicos. Un resumen de cómo funciona Kleros puede verse en la Figura 7.

El ascenso de la economía digital ha creado mercados de trabajo, capital y productos que operan en tiempo real atravesando límites nacionales. La economía P2P requiere un sistema de resolución de disputas rápido, económico, descentralizado y confiable. Kleros utiliza teoría de juegos y blockchain en un protocolo de arbitraje multipropósito capaz de soportar un gran número de aplicaciones en comercio electrónico, finanzas, seguros, viajes, comercio internacional, protección al consumidor, propiedad intelectual y académicas, entre otras.

Las criptomonedas han permitido a muchas personas tener acceso a su primera cuenta bancaria, que les permite enviar y recibir dinero de manera segura. Están ayudando a millones de personas a alcanzar la inclusión financiera. Kleros tiene el mismo objetivo en cuanto al acceso a la justicia, permitiendo arbitrar una gran cantidad de contratos que sería muy costoso llevar a una corte. Así como el Bitcoin trajo “el banco para los que no están bancarizados”, Kleros tiene el potencial de traer “justicia para los que no tienen justicia”.

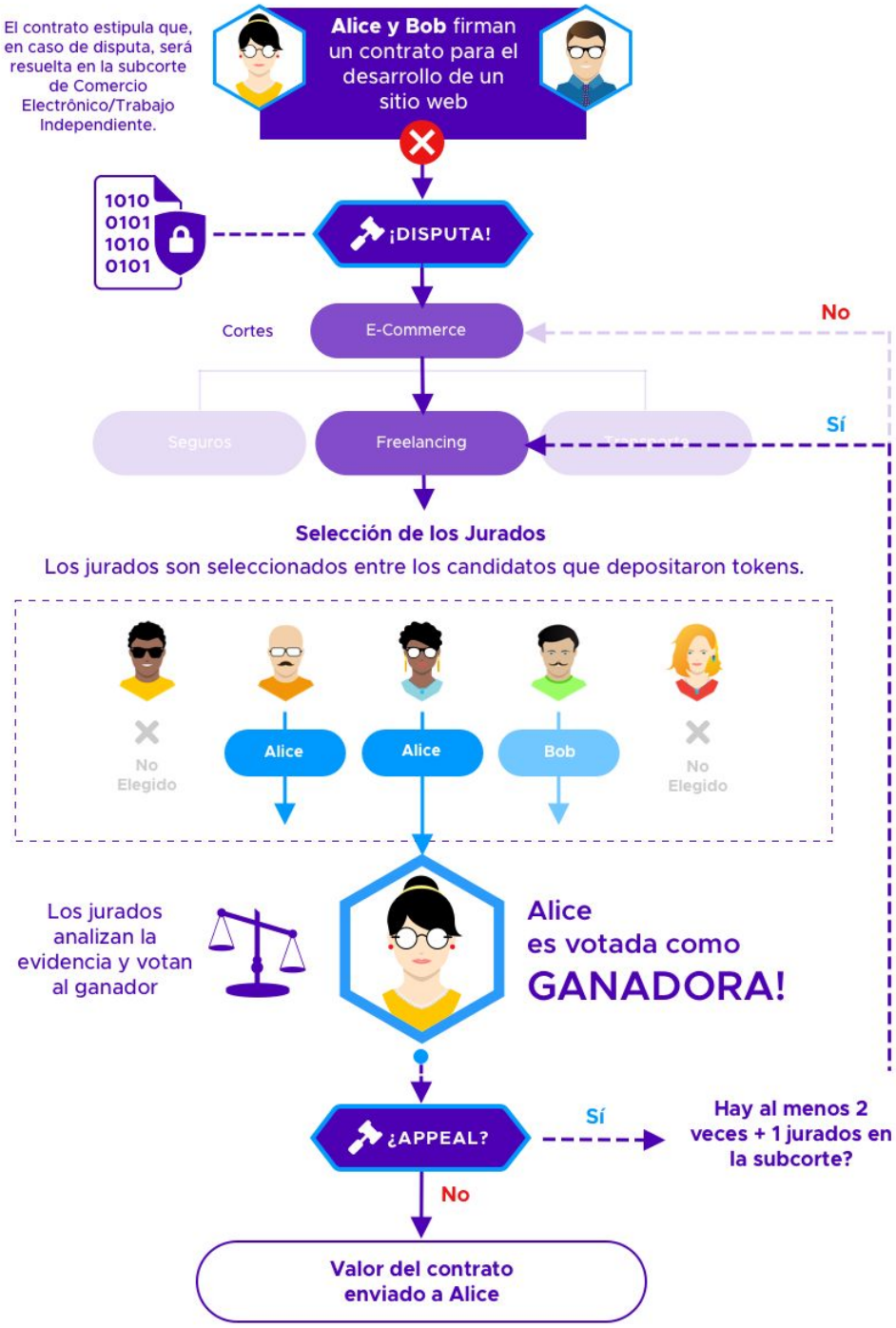

Figura 7. Ejemplo de una disputa que resume el funcionamiento de Kleros.

Referencias

[1] Amazon mechanical turk. https://www.mturk.com/.

[2] Ethereum name service. https://ens.domains/.

[3] Gnosis. https://gnosis.pm/.

[4] Ast, F. Kleros-realitio oracle service - getting real information on-chain. https://blog.kleros.io/the-kleros-realit-io-oracle/, 2019.

[5] Blum, M. Coin flipping by telephone a protocol for solving impossible problems. SIGACT News 15, 1 (Jan. 1983), 23-27.

[6] Boneh, D., Lynn, B., and Shacham, H. Short signatures from the weil pairing. Journal of Cryptology 17, 4 (Sep 2004), 297-319.

[7] Brassard, G., Chaum, D., and Crepeau, C. ´ Minimum disclosure proofs of knowledge. J. Comput. Syst. Sci. 37, 2 (Oct. 1988), 156-189.

[8] Buterin, V. Ethereum, a next-generation smart contract and decentralized application platform. https://github.com/ethereum/wiki/wiki/White-Paper, 2014.

[9] Buterin, V. Schellingcoin: A minimal-trust universal data feed. https://blog.ethereum.org/2014/03/28/schellingcoin-a-minimal-trust-universal-data-feed/, 2014.

[10] Buterin, V. On anti-pre-revelation games. https://blog.ethereum.org/2015/08/28/on-anti-pre-revelation-games/, 2015.

[11] Buterin, V. The p + epsilon attack. https://blog.ethereum.org/2015/01/28/p-epsilon-attack/, 2015.

[12] Buterin, V. Introduction to cryptoeconomics. https://edcon.io/ppt/one/Vitalik%20Buterin_Introduction%20to%20Cryptoeconomics_EDCON.pdf, 2017.

[13] Bunzy, B., Goldfeder, S., and Bonneauy, J. ¨ Proofs-of-delay and randomness beacons inethereum.

[14] Douceur, J. R. The sybil attack. In Revised Papers from the First International Workshop onPeer-to-Peer Systems (London, UK, UK, 2002), IPTPS '01, Springer-Verlag, pp. 251{260.

[15] Ford, B. Delegative democracy. http://www.brynosaurus.com/deleg/deleg.pdf, 2002.

[16] Friedman, D. A positive account of property rights. Social Philosophy Policy 11 (1994).

[17] George, W. Doges on trial curated list observations part 2 - deep dive edition. https://blog.kleros.io/cryptoeconomic-deep-dive-doges-on-trial/, 2018.

[18] James, S. Kleros escrow explainer - secure your blockchain transactions today. https://blog.kleros.io/kleros-escrow-secure-your-blockchain-transactions-today/, 2019.

[19] James, S. Kleros TCR - a deep dive explainer. https://blog.kleros.io/kleros-ethfinex-tcr-an-explainer/, 2019.

[20] Laudan, L. Truth, Error, and Criminal Law: An Essay in Legal Epistemology. Cambridge Studies in Philosophy and Law. Cambridge University Press, 2006.

[21] Lesaege, C. ERC 792: Arbitration standard. https://github.com/ethereum/EIPs/issues/792, 2017.

[22] Long, P. Cryptounlocked oracle upgrade. https://blog.wetrust.io/cryptounlocked-oracle-upgrade-5c8b22e3375b, 2019.

[23] Peterson, J., and Krug, J. Augur: a decentralized, open-source platform for prediction markets. http://bravenewcoin.com/assets/Whitepapers/Augur-A-Decentralized-Open-Source-Platform-for-Prediction-Markets.pdf/, 2015.

[24] Reitwiessner, C. From smart contracts to courts with not so smart judges. https://blog.ethereum.org/2016/02/17/smart-contracts-courts-not-smart-judges/, 2016.

[25] Schelling, T. C. The strategy of conflict. Oxford University Press, 1960.

[26] Sztorc, P. Truthcoin, peer-to-peer oracle system and prediction marketplace. http://www.truthcoin.info/papers/truthcoin-whitepaper.pdf/, 2015.

[27] Vitello, S., Lesaege, C., and Piqueras, E. ERC 1497: Evidence standard. https://github.com/ethereum/EIPs/issues/1497, 2018.

Notas al pie

[1] Para más información acerca de las cortes, véase la sección Árbol de Cortes

[2]El nombre hace referencia al pinakion, una placa de bronce que cada ciudadano ateniense usaba como identificación personal. El pinakion fue utilizado como token para seleccionar jurados en los juicios populares atenienses.

[3] Véase la sección Sistema de Incentivos

[4] A lo largo de este artículo se utiliza hash en referencia a una función criptográfica hash, en Ethereum se utiliza la Keccak256.

[5] Estamos considerando métodos más complejos que "el que obtiene más votos gana". El desafío es lidiar con la asimetría en la matriz de incentivos que producen, ya que la misma puede afectar el Punto Focal. Por ejemplo, tomar la mediana de valores que pueden ser ordenados podría generar un sesgo hacia valores centrales.

[6] Nótese que como Kleros usa un sistema de apelaciones, incluso si una mayoría vota una de las opciones, votar por esa opción no garantiza coherencia con el resultado final (véase la Sección Incentivos del Sistema). Esto limita la efectividad de una estrategia basada en copiar los votos. Por lo tanto, los votos públicos serían aceptables en ciertos casos.

[7] Esto requiere un seguro para las partes que no dispongan de capital suficiente para depositar la tarifa de apelación y el monto adicional. El asegurador pagaría el depósito a cambio de una parte del mismo en caso de un resultado favorable. Todo esto puede ser incluido en un contrato inteligente.

[8] Véase Sección Trabajo Previo: mecanismo SchellingCoin

[9] Los mecanismos de redistribución de tokens están aún bajo análisis y es posible que se adopte un protocolo más sofisticado en el futuro.

[10] Si en un nivel determinado nadie votó coherentemente, qué se hace con las cantidades correspondientes a ese nivel puede ser determinado por el procedimiento de gobierno; por ejemplo, los tokens se podrían entregar a la parte ganadora.

[11] Nótese que esta idea ha funcionado tal que se esperaba en experimentos como el considerado en (17) y en aplicaciones prácticas como aquellas presentadas en (9) y (23).

[12] Se realizarán auditorías y revisiones antes de que el código sea desplegado. Pero no es posible garantizar al 100% que no haya un error (tanto en el código como en el sistema de incentivos). Este método a prueba de fallas brinda seguridad extra.

[13] En los blockchains Proff-Of-Work, como es imposible predecir exactamente el blockhash, se puede remover este paso y utilizer el blockhash como semilla. De cualquier manera, Ehtereum planea cambiar a Proof-Of-Stake.

[14] Es todavía posible para los jurados dar a conocer sus votos. Por ejemplo, haciendo un contrato inteligente consigo mismos comprometiéndose a votar de cierta manera y destruyendo un depósito si

votan de otra manera. La discusión acerca de este tipo de comportamiento será incluida en futuros trabajos.