Kleros

白皮书 v1.0.7

Clément Lesaege, Federico Ast, William George

2019 年 9 月

摘要

Kleros 是一款搭建在以太坊之上的去中心化应用。无论非常简单还是十分复杂的合约，都可使用 Kleros 这一独立的第三方来仲裁产生的争议。它采用博弈论激励陪审员公正处理争议，是一套以快速、廉价、可靠和去中心化的方式实现高效判决的争议解决系统。

1 简介

”谁掌控了法院，就掌控了国家”亚里士多德

世界正在经历一场加速的全球化和数字化进程，跨越司法边界的人们之间的线上交易呈指数型增长。如果在不远的将来区块链愿景得以实现，大多数商品，劳动和资本将通过这些去中心化的全球平台进行分配，争议自然也会随之扩大。去中心化 eBay 用户将会投诉卖家未能如约发货，去中心化 Airbnb 客户将会投诉所租房屋与图片描述不符，以及众筹平台的支持者将会因为团队未能如约交付产品而索求赔偿。

智能合约在按照既定程序自动执行方面足够智能，却不能执行主观决断或者覆盖区块链之外的要素。对于一个实时运转的去中心化全球经济来说，现存的争议解决方案太慢，太贵而且太不可靠。一个快速、廉价、透明、可靠和去中心化的争议解决机制是区块链领域的关键组成部分，它将成为智能合约强制执行力的最终决断者。

Kleros 是一个决策协议，一个适用于处理各类争议的多目的庭审系统。作为一个去中心化的第三方自治组织，它可用于仲裁各类从十分简单到高度复杂的合约中产生的争议。仲裁过程的每一步（保管证据，选举陪审员等）全都自动化进行。Kleros 并不寄希望于一小部分个体的诚实而是依赖基于博弈论的经济激励机制。

它基于刑事诉讼认识论的一个基本观点：法院是一个认识引擎，一个从一系列令人困惑的线索中发现事件真相的工具。代理人（陪审员）遵循一份输入（证据）产生一份输出（决策）的程序 [20]。Kleros 结合众包、区块链和博弈论等技术开发出一套庭审系统，该系统以一种安全和廉价的方式做出公正决策。

2 前人工作：谢林币机制

博弈论学者托马·斯谢林提出了谢林点（也称为聚焦点）[25] 的概念，作为人们在缺乏沟通的情况下协调彼此习惯时倾向使用的解决方案，因为这对他们来说显得自然或者相关。谢林使用下面的例子阐释这个概念：“明天你要去面见一位在纽约的陌生人。你将选择何时何地与他见面?”虽然这座城

市的任何地点和时间都可能是一个答案，最常见的回答是“下午在中央火车站的咨询台“。并没有什么使得下午在中央火车站的咨询台成为更优解（假设两人达成共识，任何别的地点和时间都可以），但是中央火车站的咨询台作为见面地点的传统使得它成为一个自然的聚焦点。

当无法进行沟通的时候，或者即使可以沟通，但是没有一方可以说服另一方相信他所言为真的时候，谢林点自然出现 [16]。基于谢林点的概念，以太坊创始人维塔利克 · 布特林提议创造一种谢林币 [9]，一个使用经济激励鼓励人们讲真话的代币。如果我们想知道今天早上巴黎是否下雨，我们可以问每一个谢林币持有者：”今天早上巴黎下雨了吗？下或者没下“。每一个持币者通过秘密的无记名方式投票，等他们全部投票后，公布结果。投票结果占大多数的一方获得他们 10% 的代币奖励。与大多数投票不同的一方损失他们 10% 的代币。

托马斯 · 谢林 [25] 这样描述聚焦点”每个人关于别人希望他希望应该被期望做什么的预期“谢林币使用这个原则激励一群互不相识或者互不信任的经济人讲真话。我们预期经济人会投正确的答案因为他们预期别人会投正确的答案，别人预期他们会投正确的答案... 在这个简单的示例中，谢林点就是诚实。

谢林币机制已经应用于去中心化预言机和预测市场 [26] [23] [3]。基本的出发点是和别人投票一致是一个必须被激励的意愿行为。Kleros 底层的激励设计基于一个与谢林币相似的机制，并针对一些涉及扩展性，主观性和隐私性的挑战进行了些许修改以充分满足经济人的习惯。

The majority votes \ You vote	YES	NO
YES	+0.1	-0.1
NO	-0.1	+0.1

图 1: 一个简单谢林游戏的收益矩阵

3 一则用例

Alice 是一个位于法国的企业家。她在一个 P2P 自由工作平台上雇佣一位来自危地马拉的程序员 Bob 为她的公司开发一个新网站。他们在价钱、条款上达成协议后，Bob 开始工作。几周后，Bob 交付产品。但是 Alice 并不满意，她辩称 Bob 的工作质量比预期的要糟，Bob 答复称他确实按照合同约定开发网站。Alice 很沮丧，她没法和一个相距半个地球远的人因为几百美元的纠纷雇佣一名律师。

如果合同里声明要是产生纠纷，应由 Kleros 法院解决会怎么样？Kleros 是一个搭建在以太坊上的去中心化应用。Bob 不再答复她的邮件之后，Alice 点击一个“发给 Kleros”的按钮并且填写一张简单的表格表达她的诉求。

千里之外的内罗毕，Chief 是一名软件开发工程师。在上班大巴上的无聊时光中，他浏览 Kleros 法院网站 (court.kleros.io) 寻找一些仲裁工作。作为陪审员，他一年挣几千美元，主要负责处理自由职业者和他们的客户之间因软件开发产生的纠纷。他通常在网站质量子法庭处理案件，该法庭要求精通 html,javascript 和网站设计来处理自由职业者和他们的客户之间的纠纷。Chief 质押 2000PNK，这是

Kleros 用来选取处理纠纷的陪审员的代币。他质押的代币越多，越有可能被选为陪审员。

大约一小时后，Chief 收到一封邮件：”您已经被选为一个网站质量纠纷的陪审员，点这里 下载证据。您有三天的时间提交您的意见 “。Benito，一位来自库斯科的程序员和来自罗马尼亚的 Alexandru 也收到了相似的邮件。他们被从一个大约 3000 名候选人的池子里随机选出，彼此互不相识，不过他们将协作处理 Alice 和 Bob 的纠纷。在回家的巴士上，Chief 分析证据并投给正确的一方。

三位陪审员完成投票两天后，Alice 和 Bob 收到一封邮件：“陪审团判决支持 Alice，交付的网站与双方约定的条款不一致，一个智能合约把钱转给 Alice”。陪审员因为他们的工作得到了奖励，案子了结。

4 项目介绍

4.1 仲裁合约

Kleros 是一个选择性加入的庭审系统。智能合约必须指定 Kleros 作为他们的仲裁机。当他们选择 Kleros 后，如果产生纠纷，Kleros 的智能合约决定几位陪审员以及哪个法庭判决他们的合约[1]。这里的出发点是他们将会选择专门适用于合约主题的法庭。软件开发合约将选择软件开发法庭，保险合约将选择保险法庭等。图2展示了可供用户选择的法庭类别示例。Kleros 团队已经开发了一套使用 Kleros 作为纠纷解决机制的合约标准。此外，我们还提议两个 EIP 标准 [21] [27]，让开发者专注应用层合约的开发，而无需参与到解决纠纷的合约中。

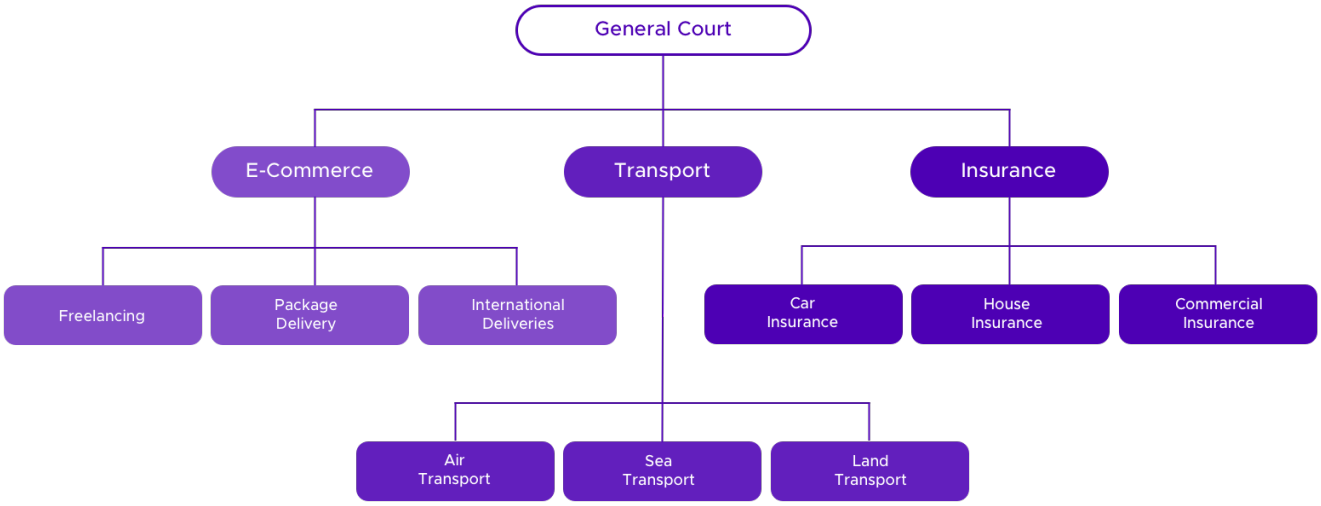

图 2: 法庭组织结构图—智能合约创建者必须从中选择一个

4.1.1 陪审员选项

合约将会设置供陪审员投票的特定选项。在上文介绍的用例中，选项可能是：“归还 Alice”，“再给 Bob 一周完成网站开发工作” 和 “支付 Bob”。

智能合约也将在判决结束后指定合约的行为。在上述用例中：

[1]更多关于法庭的信息，参见法庭部分。

- ”归还 Alice” 把钱转到 Alice 的地址上。
- “再给 Bob 一周时间完成网站开发工作” 设置 Bob 必须完成网站的时间为一周，例如：在这段时间里禁止 Alice 创建新的纠纷。此外，智能合约也可以编写为如果这个选项被选中，后续纠纷不能再选该选项。
- “支付 Bob” 把钱转到 Bob 的地址上。

4.2 选取陪审员

4.2.1 系统代币: pinakion (PNK)

用户有参选Kleros陪审员的经济动机：为他们的工作收取仲裁费。候选人使用一种称为pinakion(PNK)[2]的代币自荐成为陪审员。

被选为某场特定纠纷的陪审员几率与陪审员质押的代币数量正相关。质押的代币数量越多，当选为陪审员的几率越高。没有质押 PNK 的陪审员没有被选举的机会，这是为了防止不活跃的陪审员被选中。

PNK 在 Kleros 的设计中起到两个关键作用。

首先，它保护系统免受女巫攻击 [14]。如果陪审员只是被简单随机选取，恶意的一方可以生成大量地址以便在每场纠纷中被选取多次。通过被选取比所有诚实陪审员的数量还要多的方式，恶意的一方可以控制整个系统。

第二，PNK 通过让不一致的陪审员，例如那些投票结果与最终判决不一致的陪审员，支付一部分他们的抵押代币给一致的一方来激励陪审员诚实投票[3]。

4.2.2 选取陪审员

候选人自荐到某个特定的法庭并且质押代币之后，最终的选取随机产生。选为陪审员的几率正比于质押代币数量。理论上，一个候选人可能在一场特定的纠纷中被选不止一次（实际上不太可能发生）。用户在一场纠纷中被选中的次数（被称为它的权重）决定他在这场纠纷中拥有的票数以及在代币重新分发时他将挣得或损失的代币数量。

假设 6 位代币所有者参选一场纠纷的陪审员并且按照如下分布共质押 10000 代币：

[2]这个名字代指青铜牌饰 pinakion, 每位雅典市民用它作为身份的象征。pinakion 在雅典流行的审判中用作选取陪审员的代币。

[3]详情见激励系统部分。

Token Owner	Staked	Start	End	Weight
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

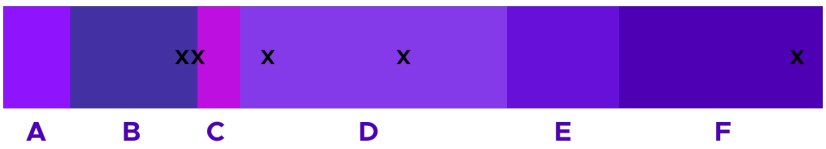

图 3: 代币质押和陪审员选取示例

对于一场需要 5 张投票的纠纷，将从 10000 枚质押代币中选出 5 枚。选取的代币（如图所示3）是数字 2519, 4953, 2264, 3342 和 9531。代币所有者 B，C 和 F 被选中，权重为 1。代币所有者 D 被选中，权重为 2。法院做出最终判决后可以取回质押的 PNK（除了那些投票不一致的陪审员支付的代币）。

4.2.3 生成随机数

为了选出陪审员，我们需要采用随机取数的方式来避免人为操纵。使用在双方之间创建一个随机数的协议 [5] 行不通。攻击者可以跟他自己创建纠纷，多次选举自己为陪审员，并选举另外一名受害陪审员。接着他将通过操纵自己的投票使得最终结果与他一致而与那些受害者不一致，借此达到当 PNK 重新分配的时候从受害者手中窃取代币的目的（参见激励体系部分）。

当前用于选取陪审员的随机数从以太坊区块的区块哈希中生成。尽管这些数值不可能提前预测，矿工可能选择以损失一个区块奖励的代价不发布区块，这将会导致随机数无法生成。将来，为了生成一个不受有很大算力矿工攻击者操纵的随机数，这些数字将使用与 Bünz et al[13] 类似机制的序列工作量证明 [12] 生成（参见将来工作部分。）

4.3 投票

陪审员评估证据后投给 [7] 其中一个选项，提交 $hash(vote, salt, address)$[4]。$salt$ 表示本地生成的熵增的随机值，用来阻止使用彩虹表破解加密值。$address$ 表示陪审员的以太坊地址。要求提交地址是

[4]本文中我们使用 $hash$ 表示密码学的哈希函数，以太坊使用的是 keccak256 函数。

用来使每个陪审员的提交内容不同，借此阻止一个陪审员复制另外一个的投票。投票结束后，他们解密 $\{vote, salt\}$，Kleros 的一个合约验证它与提交内容匹配。未能解密投票结果的陪审员将会受到惩罚（参见激励体系部分）。

陪审员提交结果后，他的投票结果无法更改。但是其他陪审员或者纠纷双方仍旧看不到该投票结果，用来避免一个陪审员的投票影响其他人的投票。

陪审员仍然可以通过别的方式宣布他们的投票结果，但是对他们来说让别的陪审员有理由相信他们所言非虚很有挑战。这也是谢林点得以出现的很重要的特点。如果陪审员知道别的陪审员的投票，他们可能像别人一样投票而不是投出谢林点。关于这些观点的深入讨论参见将来工作部分。

所有陪审员投票后（或者投票到期），陪审员公布投票结果，未能公布投票结果的陪审员将受到惩罚。最终，所有投票征集到一起，智能合约开始执行。获得最高投票数的选项被认为是胜利的一方[5]。

由于处理提交和公布结果这两步要求用户额外的交互，在某些低质押子法庭，陪审员可能想要公开投票来简化用户体验[6]，使用何种系统由子法庭参数决定（参见治理机制部分）。

4.4 仲裁费

为了补偿陪审员的工作，以及避免攻击者滥用这个系统，创建纠纷和申诉需要支付仲裁费。每个与最终判定结果投票一致的陪审员将从处理该纠纷的子法庭中获得一笔费用。仲裁智能合约将决定哪一方支付仲裁费。

规则可以很简单。例如，他们可能要求创建纠纷的一方或者上诉的一方支付费用。不过我们也许可以想出一些更复杂的规则实现更好的激励。譬如：

- 在第一个案例中，双方在智能合约里质押相等数量的仲裁费。如果一方不这样做，智能合约将会认为法庭判决支持质押仲裁费的一方（甚至不需要在该法庭创建纠纷）。如果双方都抵押保证金，当纠纷解决时获胜的一方将取回保证金。

- 在申诉环节，双方必须质押仲裁费。申诉人还需抵押一定比例的申诉费，这部分费用分给赢得这场纠纷的一方。这样一来如果一方恶意申诉来伤害对方，对方将会因浪费时间获得一定补偿，不过如果申诉最终判决是合理的，质押费将归还申诉人[7]。

关于由仲裁智能合约定义的费用结构的讨论将是未来工作的一部分。

4.5 申诉

如果陪审员做出决定后，一方不服（因为他觉得结果不公平），可以申诉要求再次审判。每轮新申诉将扩大陪审员数量为上轮的 2 倍再加 1。由于陪审员数量的增加，需要支付的申诉费为（申诉费 = 陪审员最新数量 * 每个陪审员平均费用）。

随着申诉进行，支付给陪审员的费用呈指数增长，仲裁费也将随申诉次数而呈指数增长。这就意味着，大多数情况下双方不会申诉，或者只申诉几次。不过，存在多次申诉的可能性对阻止攻击者贿赂陪审员很重要（参见抗贿选部分）。

[5]我们正在考虑比得票最多者获胜更复杂的机制，但是挑战是处理由此导致的激励矩阵的不对称。这种不对称可能会影响谢林点。譬如假如按照排序选取中间值作为结果可能导致偏向中心值。

[6]注意 Kleros 采用一个申诉系统，即使在一轮投票中的大多数投票已经投向某个选项，并不能保证该选项和最终结果一致（参见激励体系部分）。这限制了复制投票策略的效用。即使某些情形下可以接受公开投票。

[7]这要求有人为上场纠纷没有足够的资金抵押申诉费和仲裁费的双方担保。担保人将垫付一方押金，作为交换如果该方获胜，担保人将获得一部分收益。所有这些都可以由智能合约强制执行。

4.6 激励体系

再说一遍，如果一位陪审员投了大多数人选择的选项我们称他投票一致。

陪审员为了获得仲裁费而判决纠纷。他们有动机诚实判决，因为判决结束后，那些与集体投票不一致的陪审员不会获得仲裁费，而且还会损失一部分代币。如果投票一致他们本应能拿到的仲裁费和他们损失的质押代币都给了投票一致的陪审员。

Kleros 对纠纷达成一致判决后，代币解锁并重新分配给陪审员。重新分配机制受谢林币启发[8]，由陪审员的投票与其他陪审员是否一致决定他们将获得还是损失代币。

每个投票不一致的陪审员损失的代币数量为：a · 最小质押 · 权重。参数 a 决定判决后分发的代币数量。它是一个由治理机制决定的变量，是投票环境下内部动态调整的结果。最小质押参数表示在一个子法庭中可以质押的最小数值。

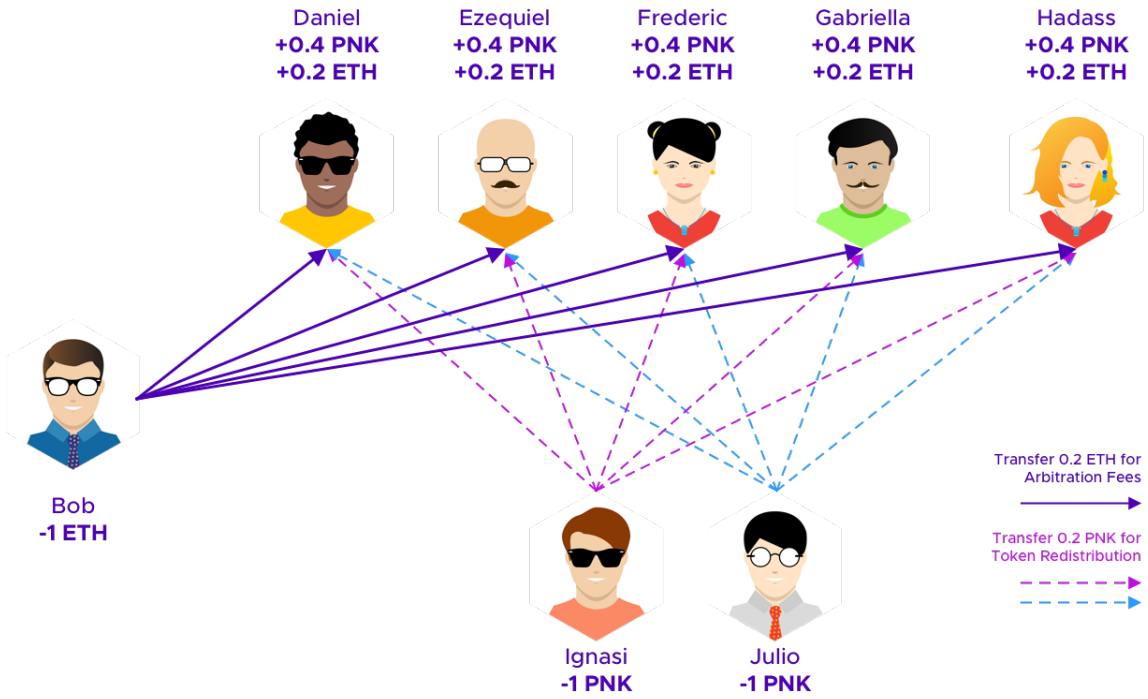

图 4: 七位陪审员投票后的代币分配情况。投票不一致的陪审员的代币被重新分配给投票一致的陪审员。Bob 败诉并支付仲裁费，胜诉方拿回质押的代币。

仲裁费和损失代币按照权重占比分配给投票一致的一方[9]。代币重新分配示例见图4。陪审员可能无法公布他们的投票结果。为了打击这种行为，对未能公布投票结果的惩罚将不低于投票结果不一致的惩罚。这是为了激励陪审员总是公布投票结果。就申诉而言，每一轮的仲裁费和代币将根据最终申诉结果重新分配[10]。

当没有攻击的时候，各方受到激励去投票他们认为诚实公平的，其他方认为诚实公平的，其他方认为他们认为诚实公平的等等。Kleros 体系内，谢林点是诚实和公平。有人可能会争辩那些决策可能

[8]参见前人工作部分：谢林币机制

[9]代币再分配机制仍在积极研究之中，未来可能提出一个更复杂的协议。

[10]如果在某一轮无人达成一致意见，那一场的代币和仲裁费将由治理程序决定如何分配；例如代币可能分给获胜的一方。

太过主观（例如，与预测市场的谢林点机制对比），不会产生谢林点。在 [25] 一文，托马斯 · 谢林进行的非正式实验表明大多数情形下被各方选出的谢林点并不存在。不过谢林发现一些选项比另外一些更可能被选中。因此即使一个特别明显的选项不存在，一些选项更有可能引起其他方的关注，更有可能被选中。我们不能期望陪审员任何时候 100% 正确，没有仲裁程序能做到这一点。有时候，诚实的陪审员也将损失代币。但是只要加总起来他们损失的远比他们获得的仲裁费和其他投票不一致方的代币少，该系统将正常运行[11]。

4.7 抗攻击性

4.7.1 买入一半代币

如果一方（或一群共谋方）要购买一半的代币，它将控制最高法院的结果，最终可以决定所有结果。然而，如果这些代币分配得当，一方很难购买超过一半的数量。首先，无法保证有一半的代币可供出售。此外，一方可以以当前市场价格买得起所有代币这一事实并不意味着它能够购买一半的代币。与大多数实物资产相反，代币的边际成本不断增加。它们在交易所的价格动态变化。要是一方购买了很大一部分代币，它的价格将因市场深度而上涨，从而使获得代币的成本越来越高。

4.7.2 抗贿赂

上诉是打击贿赂的重要机制。贿赂一小部分陪审员相对容易。但是，由于受害者始终有权提出上诉，攻击者将不得不继续以越来越高的成本贿赂越来越多的陪审员。一直到最高法院，攻击者必须准备花费巨额资金贿赂陪审团，而且最终很可能会败诉。为了控制整个法院的判决，攻击者将需要贿赂总共持有 PNK50% 以上的持有者。

此攻击在诚实占多数模型中不起作用（其中一半以上的代币是由不接受贿赂的诚实一方控制的）。但是即使有不诚实的多数（多数代币持有人都只追求自己的利益最大化），该系统仍可以在一定条件下经受贿赂攻击。

成功贿赂最高法院将大大降低代币的价值（谁希望他的合同由不诚实的法院仲裁?）。因此，为了使得贿赂得逞，攻击者要提供超过价格下跌预期损失 50% 的的价值（在几乎所有情况下，其价值都将超过争议的价值）。实际上一方对每场判决一直上诉到最高法院的可能性极小。但是，为了适当平衡激励该可能性仍有必要存在。

一方可以实施更精细的攻击（p + epsilon 攻击），并承诺仅在攻击失败时才行贿。这种攻击需要很高的预算，但如果成功，则成本为零。在 [11] 中已经提出了抵御这种攻击的博弈论应对，其中陪审员使用混合策略（陪审员仅以确定的概率接受贿赂，与接受贿赂相比增加了他们的预期报酬）。此外，我们在 Kleros“狗狗实验” 试点上测试用户在出现 p + epsilon 攻击时的行为表现。有关这些实验的结果，请参见 [17]，以及有关 Kleros 上诉系统如何降低此类攻击的可行性的分析。

4.8 树状法庭

当注册为陪审员时，用户会从最高法院开始，并参照自己的技能进入特定的分庭。每个分庭都有一些自己的特性，包括政策，庭审时间，费用，陪审团成员人数和所质押的代币。每个代币持有者都可以在其已质押代币的每个法院的至多一个分庭中登记一定数量的代币。图5展示了准许注册的示例。

[11]确实，要注意截止到目前该观点在实验（如考虑这些的 [17]）和实际应用（如 [19] 和 [23]）中与预期一致。

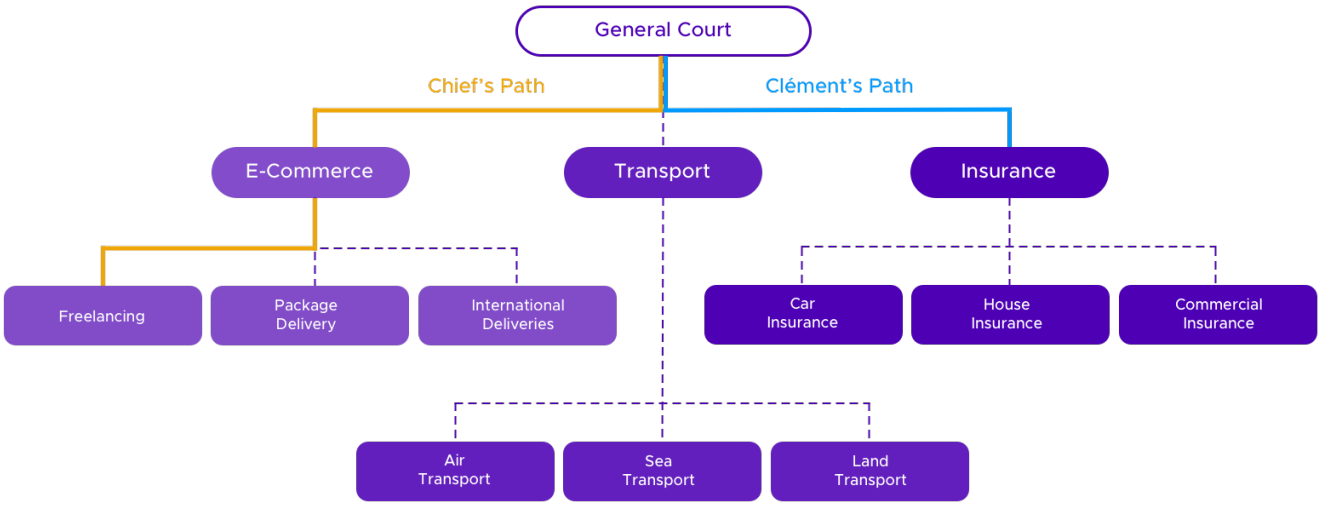

图 5: 分庭体系中可供陪审员选择的路线图。Clément 可以在最高法庭和保险分庭被选为陪审员。Chief 可以在最高法庭，电商分庭和自由工作分庭被选为陪审员。

要求陪审员在分庭之间进行选择会鼓励他们选择自己最擅长的分庭。如果他们能够选择每个分庭，那么一些人会选择所有分庭，以从其代币中获得最多的仲裁费。

4.9 治理机制

随着 Kleros 协议赢得用户和用例，将有必要创建新的分庭，更改分庭政策和参数以及将平台升级为具有更多功能的新版本。这些决策将由代币持有者使用流动投票机制 [15] 实现。代币持有者的票数将等于他们持有的 PNK 的数量。治理机制可用于：

1. 制定政策：政策涉及如何仲裁争议。它们等价于传统司法系统中的法律，决定在满足特定条件时哪一方胜诉，可以因特定分庭而异。
2. 增加新分庭。
3. 修改分庭参数：
   (a) 仲裁费。
   (b) 每场庭审时间
   (c) 最小质押代币数量
4. 更改 Kleros 依赖的智能合约。允许进行任意更改，用于改进优化或者如果 Kleros 的某些功能无法正常运行时进行紧急变更[12]。

[12]在部署代码到主网之前会进行审计和复核，但是我们不能保证 100% 没有错误（无论是代码还是激励机制）。具备此故障保护功能可增加额外的安全性。

5 未来工作

本章将讨论大量的协议改进计划。

5.1 合约隐私性

解决纠纷可能需要当事方向陪审员展示特定的信息。为了防止外部观察者访问此信息，将来，自然语言合同（英语或其他语言）和陪审团投票选项的标签不会公布，尤其是不会放在区块链上。合约创建后，创建者将提交 $hash(contract\_text, option\_list, salt)$（$contract\_text$ 表示合同的纯英语文本，$option\_list$ 表示能被陪审员投票的可选项标签，$salt$ 表示一个用来避免使用彩虹表的随机数。）。

合约创建者使用非对称加密向各方发送 $\{contract\_text, option\_list, salt\}$。这样，各方可以验证提交的哈希与发送给他们的哈希是否一致。具体到一场纠纷中，双方可公布 $\{contract\_text, option\_list, salt\}$ 给陪审员，使用哈希比对来验证它与提交的内容一致。他们可以使用非对称加密实现这一点，保证只有陪审员收到合同文本和选项。当使用 Kleros 时所有这些步骤可由用户运行的程序处理。

5.2 改进随机数生成方式

正如上文所述，目前陪审员通过使用随机生成的以太坊区块的哈希来选出。但是，这样做的缺点是，大型矿工可能会以失去区块奖励为代价来左右陪审员选举。在本节中，我们将详细介绍使用基于顺序工作量证明 [12] 生成更加稳定来源随机数的协议，它使用一个类似于 Bünz 等人的方案 [13]，也适用于如下所示的权益证明区块链[13]。

1. **初始化**: 我们从 seed=blockhash 开始，允许所有各方输入一个 localRondom 值来改变 seed 使得 seed=hash(seed,localRandom)。它允许任意一方修改 seed。我们期望 seed 不由任意一方选出。这样各方都能修改 seed，却不能选择某个 seed，因为选择特定的 seedAttack 将要求攻击者决定 localRandom 使得 hash(seed,localRandom)=seedAttack。由于密码哈希函数的原像抵抗性，这难以做到。

2. **计算主随机值**: 随机数利益相关方对 seed 进行顺序的工作量证明。从 $h_0 = seed$ 开始，他们计算 $h_{n+1} = hash(h_n)$ 直到 h_d，d 表示难度系数。计算 h_d 需要花费时间，并确保在某人知道 seed 和得到结果之间经历了一段时间。难度系数 d 是固定值，使得在初始化阶段没有硬件可以计算 h_d。因为我们需要上一步的结果才能进行下一步，所以该过程无法并行进行。这意味着没有一方能够比其他方更快地获得结果。

3. **获得区块链上的结果**: 各方都可以通过保证金提交他们发现的 h_d。其他方可以使用交互验证 [24] 否决错误的结果。对攻击者的结果可使用二分法搜索。如果攻击者提交一个错误的 h_d，诚实的一方可以询问他 $h_{d/2}$ 值。如果他给出错误的值，在 h_0 和 $h_{d/2}$ 之间有一个错值误。如果他给出正确的值，在 $h_{d/2}$ 和 h_d 有一个错值。无论怎样，搜索空间都被除以 2。诚实的一方在缩小的空间里继续搜索（错值就在这里面）直到只剩下两个值。然后诚实的一方可以找出 x 使得攻击者的答案中 $h_{x+1} \neq hash(h_x)$，这就使得他的答案无效。答案无效方失去他们的保证金。一部分被销毁，另一部分分给判断他们无效的一方。注意验证一个错误的结果无效需要的交互次数仅为 $O(\log(d))$。

[13]在工作量证明区块链中，由于仍然无法精确预测区块哈希，因此我们可以略过此步骤，仅使用区块哈希作为种子。但是，以太坊已计划改用权益证明。

4. **获取所有随机值**: 诚实的一方验证错误的结果无效后，仅剩下正确的结果 h_d。从这个主随机数我们推导出满足 $r_n = hash(h_d, n)$ 的所有随机数。

只要有至少一方诚实，此过程的输出就是一个随机数。计算顺序工作证明和交互式验证需要时间。但是对于大多数纠纷，从纠纷开始到选出陪审员等待几个小时并不是问题。但是，对于某些庭审时间特别短的分庭（例如，解决从 Web 到区块链预言机的纠纷的分庭），这种随机数生成方法可能会太慢。这类分庭可能考虑使用另外一种不那么安全的基于门限签名的随机数生成器 [6]。

5.3 惩罚公布投票过早的陪审员

上面我们描述了一种提交和公布机制，该机制允许陪审员隐藏投票，直到所有陪审员的投票都提交为止。但是，该机制本身并不能阻止陪审员公开其投票以试图影响其他陪审员。在本节中，我们讨论了潜在的未来改进，这将激励陪审员在公布阶段之前不公开其投票。

我们提议允许任何能够在投票结束前向 Kleros 展示一位陪审员提交内容的一方窃取该陪审员的 PNK，并使该陪审员的投票无效。这样，如果陪审员希望将其投票透露给另一方，面临两种选择：

1. 仅公开其投票。另一方没有任何证据相信他以这种方式进行投票。陪审员可能会撒谎，而另一方无法核实。

2. 公开其投票和提交的内容。另一方将获得其投票证明，但也将能够窃取该陪审员的 PNK。

该机制将阻止陪审员以无需信任的方式公开投票[14]。[10] 中详细介绍了我们正在研究的更完整的潜在解决方案，最终将应用在 Kleros 上。

5.4 流动投票治理

在上面关于治理机制的部分，我们描述了代币持有者如何为平台做出许多决策。在本节中，我们描述了一个未来的计划。如果代币持有者选择不直接投票，那么允许他们委托别人行使投票权。用户未能投票时，其投票权将自动转移给其代理人。您可以在图6中看到流动投票机制的图示。投票委派也可以是分庭特定的。用户可以选择在某些分庭中委托投票，而在其他分庭中则不委托。请注意，代理人不必是人类。它们可以是执行任意复杂投票规则（例如，根据市场数据对更新费用进行投票）的智能合约。

6 应用

Kleros 是一个可以用在很多场合下的通用的多功能系统。下面我们展示一些潜在的用例：

- **托管**: 为了支付链下商品或服务，可以将资金放入智能合约中。买方在收到商品或服务后，向卖方解锁资金。如有争议，可以使用 Kleros 智能合约判断偿还给买方还是向卖方付款。这样的基于 Kleros 的托管系统已经成为现实，参见 [18]。

  托管形式也可以更复杂。例如，对于租赁协议，可以要求承租人支付押金。万一财产受损且承租人未就赔偿达成共识，出租者可以提出争议，起诉所求部分押金作为赔偿。

[14]陪审员仍然有可能提供有关投票的见解的信息。例如，通过他们自己之间创建一个合约约定按照某种行为提交投票，并且如果他们投票不同就销毁他们的保证金。有关此类行为的讨论包含在将来的工作中。

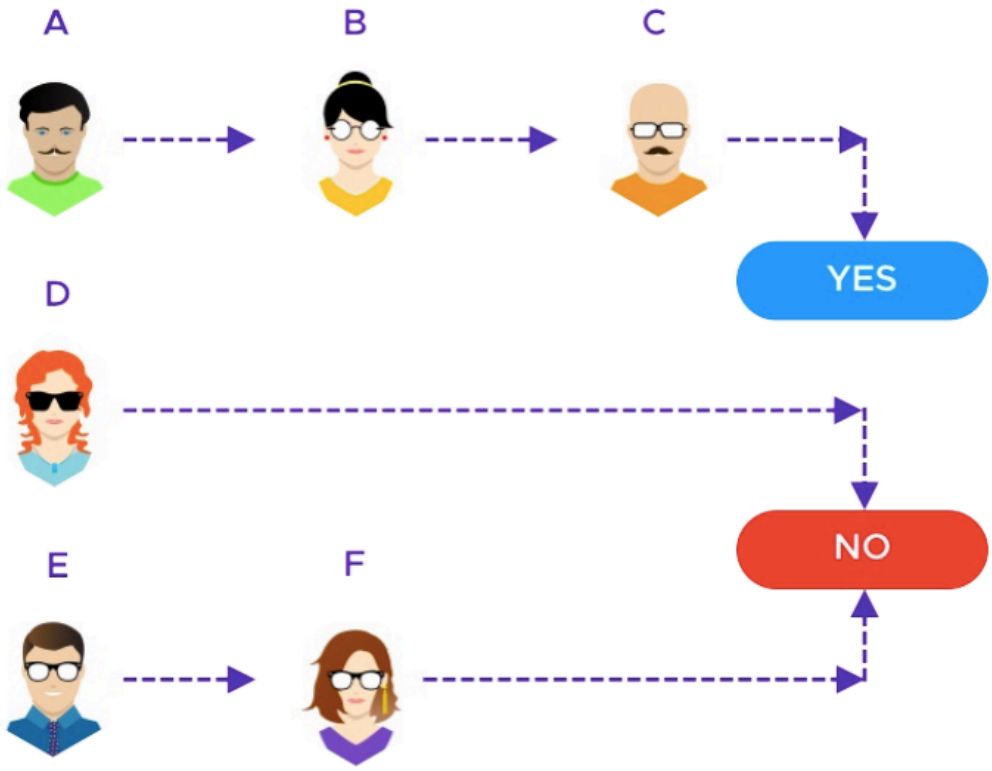

图 6: 流动性投票示意图

- **微任务:** 去中心化平台可以为微任务付费（按照亚马逊土耳其机器 [1] 的方式）。领取任务的用户会存入一笔保证金，并提交微任务的答案。任务可以重复。如果任务得到不同的答案，任务执行人可能会承认自己的错误，这会将部分保证金转移给正确执行任务的用户。如果有多个任务执行人坚持自己的答案，则会使用纠纷仲裁程序，而失败的任务执行人会将其部分保证金转给获胜者。

- **保险**: 投保人向承保人支付费用以期在发生特定事件后能获得赔偿。承保人必须存入一些保证金，这些保证金可能针对多个被保险人（遵守风险管理规则）。当需要保险的事件发生时，承保人可以对其进行验证并赔偿投保人。如果承保人未验证事件，则发起争议解决程序。如果投保人在争议解决过程中获胜，则承保人保证金中的资金将转给投保人。如果涉及索赔多于保证金的多名投保人，则还需要通过争议解决程序来确定投保人之间如何分配这些资金。

- **预言机**: 供智能合约使用的去中心化喂数机是以太坊 [8] 早期设想的用例之一。一方（可以是智能合约）提出一个问题。每个人都可以缴纳保证金并提交答案。如果每个人都给出相同的答案，则由预言机返回该答案。如果有多个答案，则将启动争议解决程序。预言机返回争议解决程序给出的答案，给出错误答案的当事方将损失其保证金，这些款项将被奖励给诚实的答案提交者。Realitio 提供了基于此类原则的预言机服务，提供使用 Kleros 解决随之而来的争议的选择 [4]。此外，使用 Realitio 预言机的其他应用程序（例如 CryptoUnlocked，[22]）间接依赖此争议解决方案。

- **精选列表**: 精选列表可以是白名单或黑名单。例如，白名单可以列出已采取适当审计程序的智能合约。黑名单可以列出与该名称无关的各方注册的 ENS（以太坊域名服务 [2]）（例如，恶意方可能注册 “kleros-token-sale.eth” 来骗人将资金发送到该地址），各方可以通过支付保证金的方式将项目提交到列表。如果在足够长的时间内没有人质疑该项目是否符合列表规则，则添加项目名称并退还保证金。如果有人通过缴纳保证金质疑该项目，将启动争议解决程序。如果该项目被认为

符合列表，则将其添加，提交方获得质疑方的保证金；否则，提交方的押金将被分配给质疑方。Kleros 已被用于满足各种属性代币的代币精选列表（譬如，满足 ERC20 属性的代币 [19]）。

- **社交网络**: 对于去中心化社交网络而言，防止垃圾信息，欺诈和其他滥用行为是一项挑战。各方可以举报违反网络政策的行为，并支付押金。如果对违规提出质疑，则将采取争议解决程序。如果判决没有发生任何违规行为，则举报方将损失押金，押金分给被告方。如果 Kleros 未对违规行为提出异议或证实，可以进行其他操作：删除相关内容，罚没内容发布者的签约金，缩小其文章覆盖范围。

7 总结

我们已经介绍了 Kleros，一个去中心化的法院系统，基于经济学激励通过众包的陪审员进行智能合约的仲裁。您可以在图7查看 Kleros 如何运行的概述。

数字经济的兴起产生了跨越国界实时运转的劳动力，资本和商品市场。P2P 经济需要快速，廉价，去中心化和可靠的争议解决机制。Kleros 将博弈论和区块链技术应用在多用途仲裁协议中，能够支持电子商务，金融，保险，旅游，国际贸易，消费者保护，知识产权和学术界等众多领域。加密货币为许多人拥有第一个银行账户提供了可能性，用户可以安全的发送和接收货币。加密货币正在帮助数百万人实现金融包容。Kleros 在司法援助上也在做同样的事情，仲裁大量因成本太高无法在法庭上解决的合同。正如比特币带来了“无需传统银行许可的银行”，Kleros 也有可能带来“无需传统司法服务的司法”。

参考文献

[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 Crépeau, 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] Bünzy, B., Goldfeder, S., and Bonneauy, J. Proofs-of-delay and randomness beacons in ethereum.

[14] Douceur, J. R. The sybil attack. In *Revised Papers from the First International Workshop on Peer-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.

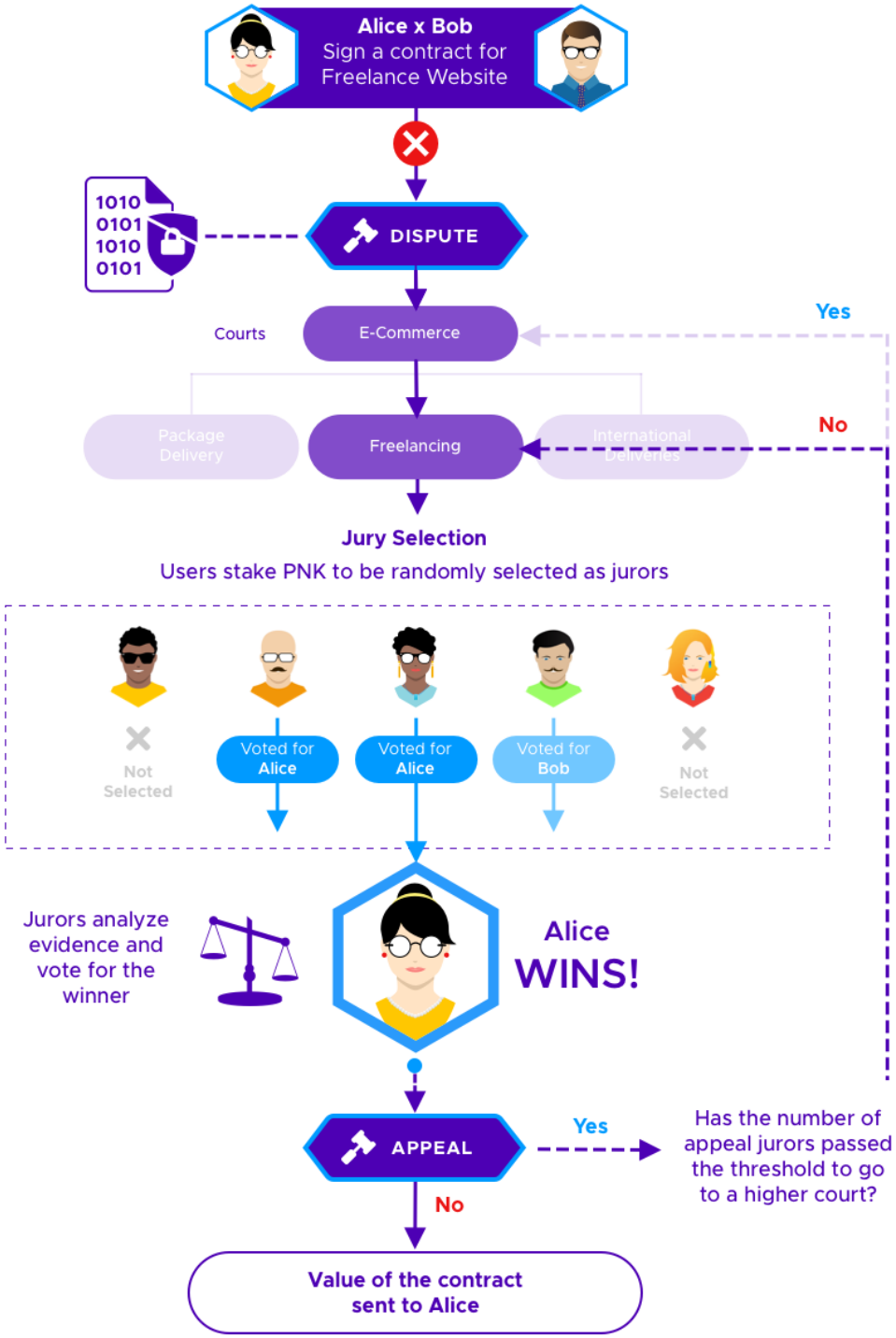

图 7: Kleros 如何运行示例