



一个基于区块链，由信任连接你我
的口碑平台

— —
coomi.io

白皮书

版本V1.0



目录

1. 摘要.....	3
1.1. 关于Coomi	3
1.2. 引言.....	3
1.3. 口碑市场.....	4
2. 口碑生态的核心问题	5
3.解决方案.....	6
4.商业模式	12
5.产品的落地.....	14
6.Token分配	16
7. 数学模型	19
8. 基于人工智能的排序	25
9.代币的发行和分配.....	27
10发展计划.....	28
11团队简介	29
12代码部分	32
13免责声明	32
14风险提示	33

1. 摘要

1.1. 关于Coomi



Coomi是一个基于区块链的去中心化口碑共享与搜索平台；是一个用户发表经验见解评论可以获得赏金的社区；也是一个以六度空间理论为基础的熟人口碑库；更是一个用户消费决策行为发生前的重要商业入口。朋友或者陌生人对于某事或者某物的判断评论，以一条朋友圈的方式，透过数据库里的熟人关系链传递给你，帮助到你。本文论述了**Coomi**如何将区块链技术和六度空间理论相结合，运用基于网络末梢神经原则的排序机制来挖掘有价值的口碑，并且实现口碑信息的生产、传播和变现。

1.2. 引言

故事开始于我们的团队成员**Nami**，她在买化妆品之前，经常会去问肤质相近的朋友们，化妆品用下来的实际效果和感觉。然而，这些口碑在网上是找不到的。类似的需求广泛存在我们身边，它的商业价值

也有待我们挖掘。这让我们认识到了周边朋友们口碑的价值所在，以及基于信任关系的口碑所蕴藏的极大价值。

另一个故事是，团队成员**Tsuki**在美食口碑平台和**Google**上搜索对比了半天后，去了一家网上口碑很不错的店，结果体验却非常差。这让我们意识到目前互联网上的口碑信息过剩和虚假带来的问题，它不光影响到用户查询信息的效率，甚至影响到了最终的消费体验。

所有这些和口碑相关的体验，让我们开始思考真正的口碑是什么？下一个时代的互联网口碑应该是个什么形态？也正好是在这个时候，我们的成员**Erichin**参与了**Alis**（日本版**Steem**）的创建工作，他设计智能合约的过程，将我们吸引到了区块链和信息社区结合的这片神奇的领域。

※**Steem**(<https://steem.io/>)是一个去中心化的社交媒体平台，用户可以通过发文或为平台上热门内容投票获得奖励。

1.3. 口碑市场

众所周知，口碑生态是互联网经济的基础性设施之一，它作为消费者的决策行为的辅助工具，扮演着重要的角色。去哪里，吃什么，住哪家？越来越多的人参考着网上的口碑信息。在口碑的商业价值挖掘方面，近年**Instagram**，小红书等优秀产品，已经将口碑的角色从幕后推到“入口”的位置，用户从口碑库里面直接搜索口碑，然后选择商品进而消费，而不是以往的先看商品再看口碑。

就其价值而言，2017年国内市场在大众点评和美团的推动下规模已近万亿，日本单单饮食市场也达上百亿规模。其市场潜力和作为入口的价值，我们在这里就不过多着墨赘述。

2. 口碑生态的核心问题

正如马云曾在世界互联网大会上所说“今天的网络空间需要更加客观、理性、真实的声音和内容。”地球的另一侧，饱受假新闻困扰的马斯克则在发推特说要成立新闻评级网站。如今的互联网信息虽然发展迅速，但是存在很多结构性的问题显而易见。这些这里所提的互联网信息当中，当然也包括我们的聚焦点：用户对于万物的评论见解判断。我们通过定性调查，总结了口碑的6大核心问题，罗列如下。

如何让用户30分钟的搜索决策过程缩短到30秒。

如何解决中心化口碑平台的各种乱象。

如何实现口碑的民主化，让口碑不再是大V和KOL的专属，让大众的口碑帮助他人的同时可以变现，获得利益回馈。

如何改变现有利益关系，改变商家为鱼肉，平台为刀俎的零和局面。

如何改进建立在用户流量之上的广告模式和用户不愿被广告干扰的自相矛盾的商业模式。

在区块链诞生之前，这些问题始终没有答案。

3.解决方案

上文中，我们列出一些当前互联网口碑领域一些较大的问题，这些问题将是**Coomi**优先解决的对象。解决问题的钥匙在哪里？我们的答案是口碑平台、区块链、熟人关系链这三把钥匙。



问题1 如何让用户30分钟的搜索决策过程缩短到30秒？

互联网上口碑从无到有的过程惠泽了很多消费者。但是，伴随着互联网的普及，受利益驱动的带路党们（SEO，Listing等等）经常将我们带到并非用户想去地方。甚至，我们经常苦恼，这个口碑可信吗？这么多信息，我该信谁？要不要再多看看其他家的口碑？于是整个搜索甄别口碑信息的过程耗费用户大量的时间。

Coomi在经过定性调查和需求分析后，从两方面着手解决这个问题。首先，改变信息展示方式和搜索顺序。以找一家美食店为例，我们先看一下用户目前的整个搜索流程。

输入条件→进入店铺1→看口碑1→决策1→进入店铺2→· · ·→看口碑N→决策N，

而我们采用了已被Instagram，小红书验证过的另一套方法来大大缩短整个流程。

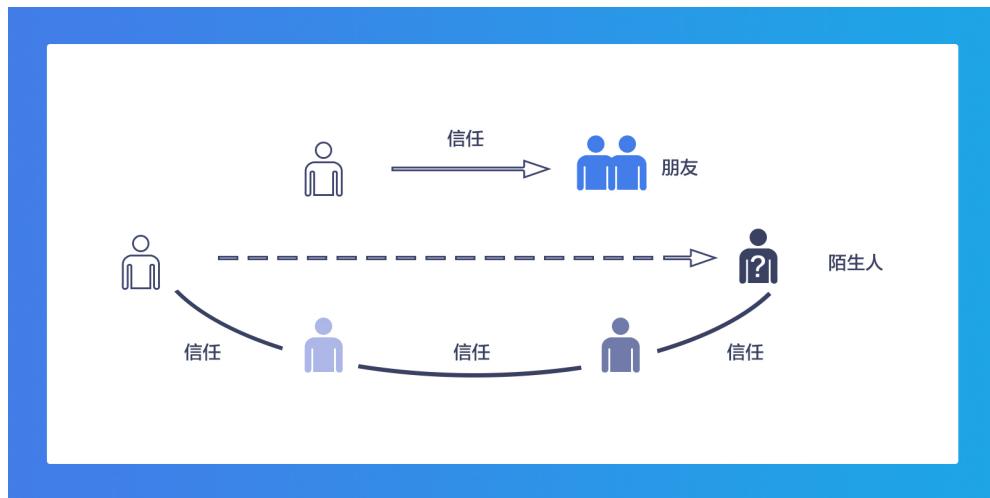
输入条件→看口碑1→看口碑2→决策。

具体概念，请参照下图。



其次，我们从朋友圈挖掘有参考价值的口碑信息推给用户，过滤掉那些没有参考价值的口碑。一条口碑有没有参考价值，Coomi从两个点判断，一个是我信任这条信息的发布者（下图中的信任力），另一个是客观来讲这条口碑评价真的写得不错，即有绝对价值。详细介绍请参见第六章节。第一点很好理解，正如大家所知，我们在决策时愿意参考一条信息，不仅是因为我们相信这条信息，更是因为我们相信发布这条信息的人。所以发掘并量化口碑的信任属性的时候，Coomi通过导入用户的人际关系链，基于六度空间理论，即便是一个遥远的陌生人的信息，也能沿着这张“信任”之网在6人之内传递到你的搜索结果里。好比，如果我信任你，而你给我推荐你信任的朋友，那么比

起陌生人，我会更信任你的朋友，虽然我也不认识他。所以，通过这个关系链找出能够值得你信任的人和他发布的口碑信息。



使用关系链来提高用户看的口碑可参考性，这一点在商业角度已经被日本**Retty**这个口碑平台验证。**Coomi**将在突破“朋友”这个第一层级的关系，扩展关系层级，让有参考价值的口碑帮助到更多的人。这样即便是陌生人之间的信任问题都可以通过关系链得以解决。

你可能会问：即使是信任的人他的口碑也未必有参考价值啊。所以我们引入口碑的客观的绝对价值。如何量化口碑的绝对价值？这点我们参考同行的普遍做法，主要根据底层数据库里用户的行为数据，计算得出。这些数据包含不限于，用户对这条口碑的转发，收藏，屏幕位置停留时间，图片的打开次数等等。两者合一，就可以计算出这条口



碑对于这个用户的相对价值，并通过这个相对价值对搜索结果进行排序。这样，用户能快速获取“值得信赖”“值得参考”的口碑信息，那么消费决策的效率自然会更高更快。

所以，在我们的产品里面，搜索口碑功能将是一个主打功能。我们立志做口碑界的谷歌，算法将是我们的核心。搜索结果排序算法内部逻辑涉及到Token的分配算法，第六章节会做详细介绍（这套算法由毕业于清华大学和中科院的Coomi创始人共同创造，正在日本申请专利。）

问题2 如何解决中心化口碑平台的口碑乱象。

值得一提的是在这套以信任和客观参考价值为准绳的算法世界里，店铺不用担心恶意差评，也不用费心软文。因为在客观的算法下，我们

会将最真实的口碑还原给用户。



**问题3 如何实现口碑的民主化，让口碑不在是大V和KOL的专属，
让大众的口碑帮助他人的同时可以变现，获得利益回馈。**

Coomi理解的口碑民主化有三个要素，

一是算法的民主

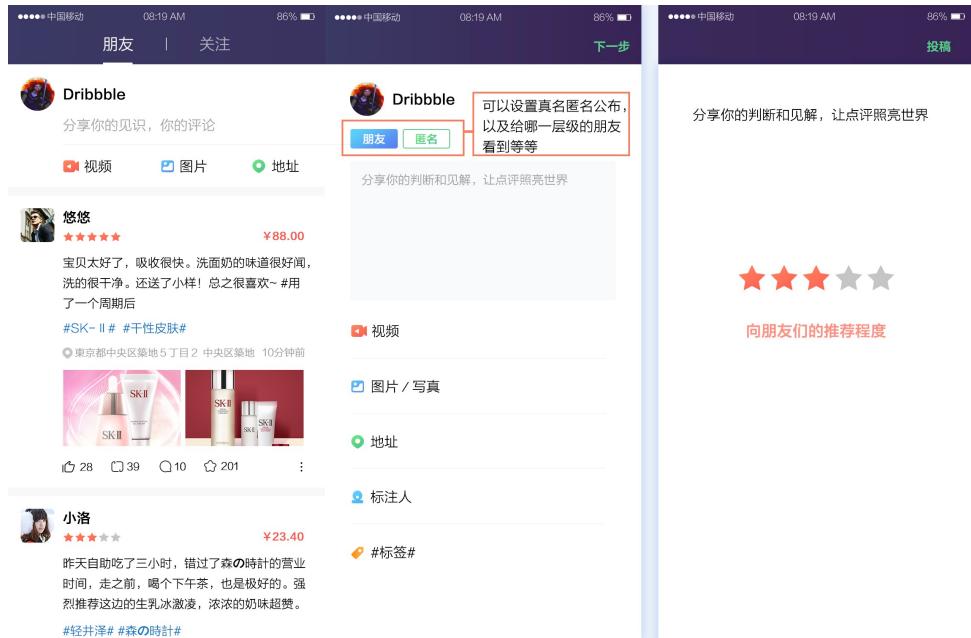
二是生产关系的民主

三是大众的参与

首先，不管这条口碑是来自自带流量的大V或者KOL，还是来自平民大众也好，我们算法排序的唯一准绳是，对于用户来说，这条口碑值得信赖，值得参考。

其次，在生产关系方面，目前即便是自带流量的KOL，利益分配也要看平台脸色，这种明显的上下级关系里谈何民主？但在Coomi这个强关系社区里，即便你是个默默无闻的人，也可以轻松简单地分享你的经验见识到朋友圈，让朋友或者陌生用户通过关键词可以搜到你的投

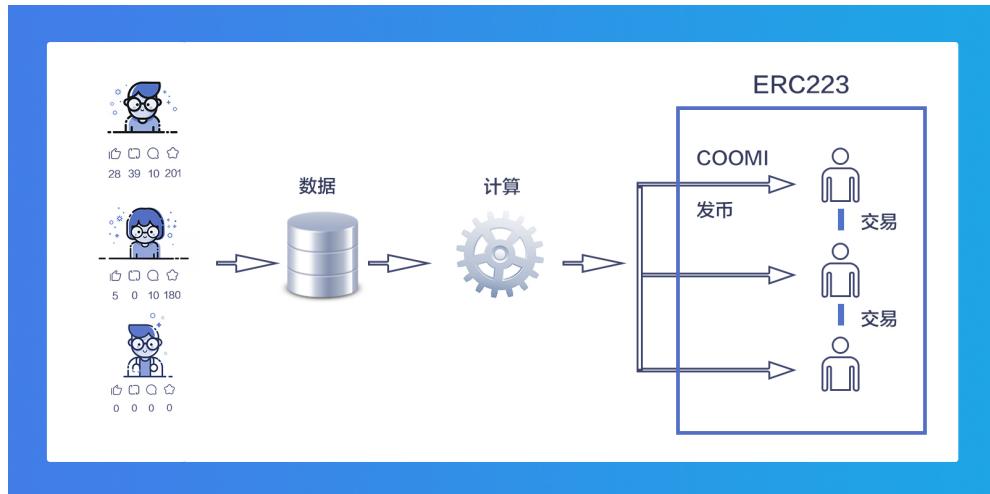
稿，并且通过第6章的Token分配原则将利益以最直接透明公开的方式回馈给你。



这里岔开来说一点，从搜索口碑者的角度，朋友圈这样SNS投稿的表现形式比起生硬冷冰的亚马逊式评论模式，会更有亲近感和信任感。我们不会像大众点评或者Yelp一样去规定用户投稿内容的类别，我们更倾向做口碑界的谷歌或者推特，不限口碑类别，包罗万象。用户操作就如同发朋友圈一样，类别可以自由地用#标记。而我们**独有的竞争壁垒正是基于这张网的口碑社区**。

最后，民主化还有个要素就是大众参与。关于口碑，很多用户会抱有疑问，“我为什么要去写评论？写了有什么好处？”是的，人们只有在有极好货极差体验的时候才会去写评论。没有足够的口碑，口碑的质量（值得信赖，值得参考）不够好，那么一个正真的口碑社区就很难建成。通证经济很完美地解决了这个问题。我们通过一套激励人心的算法去分配**Token**，用虚拟货币奖励那些帮助别人的口碑生产者，

传播者。以此，鼓励大众参与。**这也是口碑社区引入区块链的必然性和意义所在。**一般的内容社区，在解决标题党等问题上，也做了很多探索，比如**Steem**按照一定的规则，用**Token**奖励文章的发布者和阅览者。



此外，为了最大限度的降低用户的资产风险，相比与目前主流的ERC20，我们选用了更为安全的ERC223规格的以太坊代币。

4. 商业模式

接下来，结合**问题4**和**问题5**来谈一谈**Coomi**的商业模式。

问题4如何改变现有利益关系，改变商家为鱼肉，平台为刀俎的零和局面。

问题5如何改进建立在用户流量之上的广告模式和用户不愿被广告干扰的自相矛盾的商业模式。

上文谈到用户和平台的关系，这里谈谈商家和平台的关系。近年，发生了两则关于鱼肉和刀俎的故事。在日本，平台把不加入广告套餐的店铺评分统一归零，无情碾压商家辛苦经营的口碑。在中国，平台把加入竞争对手平台的店铺直接下架，显得更加残酷。平台的无情背后，更多是以广告收入为主的商业模式导致的无奈。**Coomi**摆脱传统的广告模式，我们把利益的分配法则写在智能合约里面，然后把合约放到区块链上。区块链上的合约具有公开透明不可更改的特性，以此形成共识，彻底颠覆原有的利益关系。生态的所有参与者不仅仅是参与者也是**Coomi**的共同建设者，是**Coomi**的主人。口碑平台现有的商业模型，不管是广告模型还是佣金模型，最后都是不平等关系下的剪羊毛过程。**这也是我们引入通证经济，引入区块链的一个重要原因。**

接下来，介绍下**Coomi**的商业模式。围绕口碑，**Coomi**有三驾马车带动我们的整个口碑生态。

第一 流量变现

当用户看完口碑后，点击#A商品#的时候，我们会和推特一样把用户导入到A商品的口碑目录，同页放会有预约购买平台的链接。方便用户消费的同时，我们也从平台那里获得流量的变现。

第二 基于口碑的附加价值变现

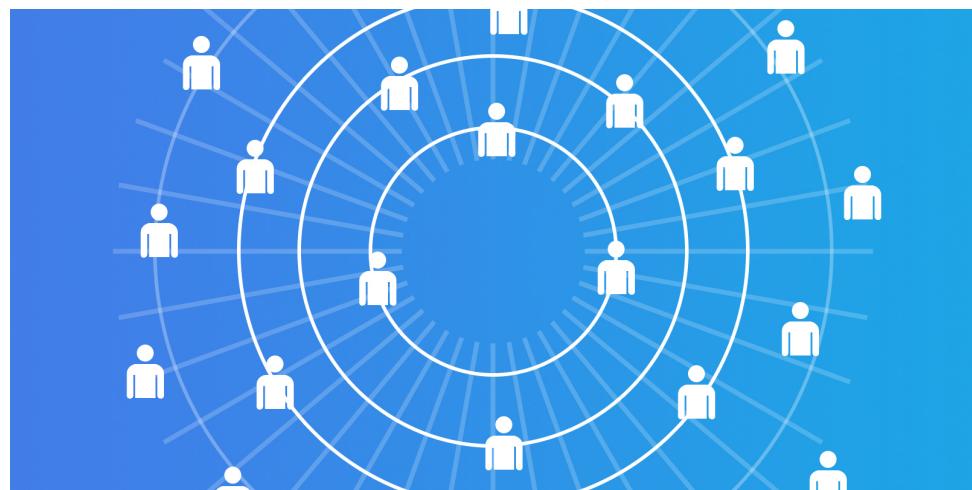
比如，我们积累了用户的浏览点赞转发收藏等行为数据，基于这些底层数据会生成每一个商品的口碑价值（基于前文介绍的口碑的绝对价值）。我们可以将这些口碑排名出售给商家或者它的竞争对手。我们也可以按照口碑价值，对一些品类的商品作出排名或评奖，然后将这些排名或者获奖名单的印刷权、使用权卖给商家。这个模式**cosume**在日本获得了极大的成功。

第三 繁荣Token生态，提升平台自身的价值

首先，打造**COOMI** (**Coomi Token**) 流通场景。比如，用**COOMI** 充值成为VIP会员之后，可以去看一些特殊的口碑或者排行榜。再比如，在一些指定商家那里，用**COOMI**可进行支付和折扣。同时，前两个商业模式的部分利润回购并销毁**COOMI**。

5.产品的落地

目前，各种区块链相关的信息社区产品落地程度参差不齐。比**Steem** 和**Alis**，**Coomi**对信息生产者的门槛更低，落地也更为容易，因为任何人都可以写出值得别人参考的口碑。



Coomi的落地分为四步。

第一步

100人范围的测试。以**Coomi**创始人的关系为基础，邀请100人作为我们的原始用户。通过测试，我们调整算法模型里的一些参数。同时对产品设计运营方面也做一些改善，为后面的第二步打下基础。

第二步

1000人范围的原始口碑积累。初期Coomi内的口碑是价值链的第一环，它的质量会影响到后期新增用户的留存，所以我们选择性地找一些人来做第一批生产者，生产出优秀的内容供后面的人参考。这里，我们会用Facebook 等API提供的关系链来完善Coomi用户的关系网。

参考**Steem**这些信息平台，最初的时候，首批用户应该是区块链参与者关心者。大家也更倾向于生产和区块链相关的内容。**所以，我们首先从“区块链相关口碑”这个品类切入，发表对某个区块链项目的见解，评价某一个币，推荐自己信任的钱包，甚至点评某个项目的创始人等等。**我们会邀请链圈的优秀推动者和区块链内容相关网站来驻扎，生产首批内容，去引导社区的初步形成，使Coomi先成为区块链行业的一个极具商业价值的战略性入口。

第三步

基于现有关系的逐步外扩。外扩的方式是通过现有用户介绍制和账号发行制，以保证整张关系网的完整性和用户增长速度的可控性。

第四步

我们将通过**TokenPocket**等战略合作伙伴的渠道，全方位获取用户。

6. Token分配

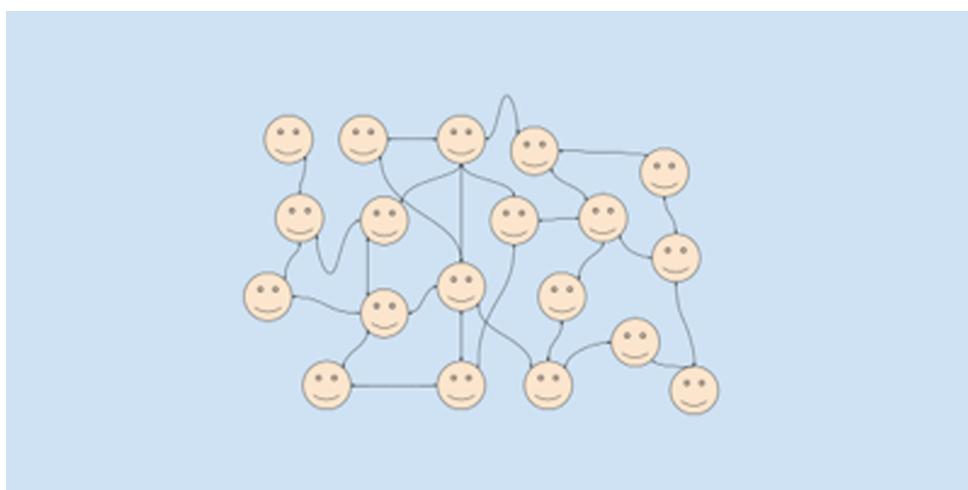
目前主流的区块链社交媒体**ALIS**, **STEEM**, **PRIMAS**基本都是通过间接或者直接限制用户活跃度的方式来防止刷币行为，

例如：**ALIS**是限制用户短时间内的大量点赞行为，**PRIMAS**则是对于用户采用HP值的方式限制其最大活跃度。而作为**Coomi**的优势之一，就是在法律范围之内给予用户绝对的自由，不用任何方式限制用户的行为，同时防止用户的刷币行为。下面来介绍我们的这种“自满”的独到之处：

首先罗列并解释**Coomi平台**的几个重要概念

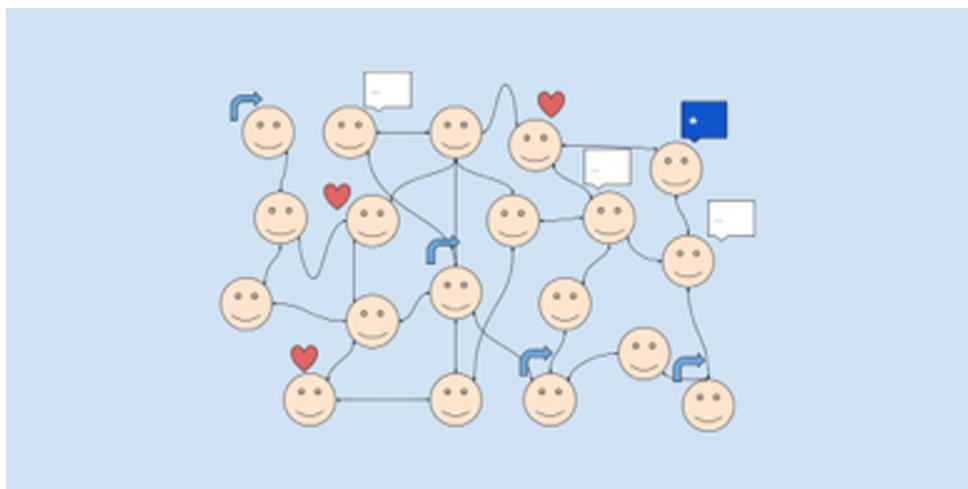
关系网

Coomi是基于关系型的信息平台，对于所有平台上的用户，**Coomi**的数据库中会维护所有用户之间的关系数据。**Coomi**的用户可以相互加好友也可以单向关注，也可以直接导入Facebook的好友关系数据。



信任力：

在Coomi的关系网中，每个用户之前的信任力是一个向量，也就是说A对于B的信任程度不同于B对于A的信任程度。这在现实中也是非常好理解的。而信任程度和向量的长度成反比。如果把每个用户作为一个节点，那么可以形象描述为哪个节点到A距离越近，则对A信任越强。而信任力决定于用户之间的互动频率，互动方式，共同好友数，等等。



信息的绝对价值

用户在Coomi上发布的每个信息都具有价值，其包含信息的绝对价值和相对价值。对于**每条信息的交互行为**（点赞，转发，留言，收藏，信息浏览时间，图片打开次数等等）和**参与者在关系链中的节点距离**，决定了这条信息的绝对价值。

信息的相对价值

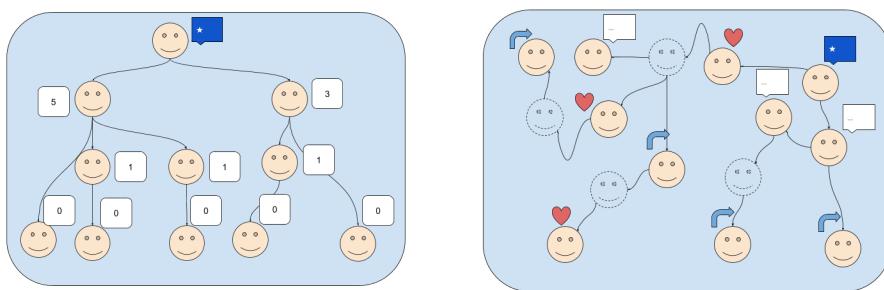
而对于每个用户而言，口碑信息本身的绝对价值固然重要，但口碑发布者对于参考者的“信任力”也很重要。因为人们总是倾向于参考和相信来自更信赖的人的信息。这样每条信息对于各个用户而言就有不同的相对价值。相对价值取决于信息的绝对价值和阅览者对于信息发布者的信任力。

奖励机制

对于每个信息的发布者，以及对这个信息进行互动者，Coomi都会给予**TOKEN奖励**。

贡献价值的计算

通过 Coomi的用户关系网数据，首先可以筛选出这条信息的价值贡献者的关系网数据，在这个子关系网数据中，以信息的发布者为顶点，可以构造一个最短路径树。这个树满足信息发布者到每个节点的路径都是关系网中的最短路径。这样的最短路径树可能存在多个，但是我们可以通过Dijkstra算法，在松弛操作遇到最短路径路径长度相等情况时，优先选ID小的节点，这样能唯一确定树。这个树可以理解为这条信息传播的最大可能路径。这个树的每个节点都是一个信息价值的贡献者。每个节点所拥有的字节数就是这个节点的贡献值。而贡献值决定了每个贡献者所能获得的**TOKEN奖励**。这样由于末节点没有子节点，所以末节点的贡献值为零。这就防止了用户通过大量的不具有信任力的互动行为的刷币行为，比如大量随机的点赞行为。



绝对价值的计算

每条信息的绝对价值在Coomi中，定位为对于这条信息作出贡献的每个人贡献价值的总和。也就是上述的最短路径树的每个节点的贡献价值之和这样不仅可以得到更合理的绝对价值，同时也可以防止用户通

过注册多个账号，互相大量互动，进行的刷币行为。

我们称这种基于最短路径树的算法原则为网络末梢神经原则。

相对价值的计算

正如上所描述，相对价值决定于每个信息的绝对价值和信息发布者对于自己的信任力。可以表示为：

$$\text{相对价值} = F(\text{绝对价值}, \text{信任力})$$

信息的排序

Coomi拥有一个朋友圈的功能，在这里可以浏览到朋友，朋友的朋友所发布的信息。**Coomi**希望能够让每个用户更容易发现对于自己更有价值的信息。这样朋友圈信息的排序方式正是居于每个信息对于这个用户的相对价值。相对价值越高的信息，排序越靠上。

可以表示为：

$$\text{信息的排序优先度} = F(\text{相对价值})$$

正是通过引入关系网，**Coomi**不但能够更合理的判断每个信息的价值，而且能够更合理的判断每个信息对于每个人的价值。正是由于对于价值更合理的判断，在不限制用户行为的基础上防止了无价值的刷币行为。

7. 数学模型

上一章的各个重要元素中，用户关系网是最为核心的要素。在计算用户信任力，口碑绝对价值，相对价值，TOKEN奖励的时候，发挥决定性作用。所以，用户关系网的合理和准确决定了**Coomi**平台能否合理的

奖励和驱动价值的生产和传播。这里我们重点介绍Coomi的关系网模型。

1. 模型基础

多模社交行为中的包含了多影响社交因素，但是起决定的因素是关系强度，它是隐变量，是不能直接观测的量，关系强度隐变量模型的基础是社会学中同质性理论：人们倾向于和与自己相似特点的人形成关系，也就是人们常说的“物以类聚，人以群分”，两者相似程度越高，所形成的社会关系就越强。关系强度隐变量模型是一种无监督式的学习模型，它的核心思想是将用户的个性化配置(喜好、关注的人等)和用户的交际活动(留言，点赞，转发等)的隔离，而联系两者的是关系强度。关系强度作为中间桥梁，受到个性化配置影响，同时又影响交际活动。

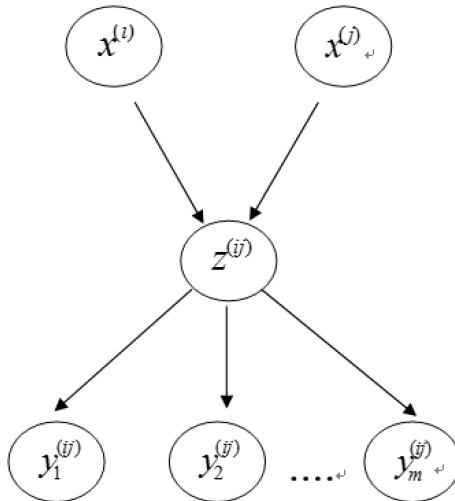
2. 模型建立

因为任何人的都只能用有限的资源来维持关系，比如说时间、信息，很自然地我们将会把更多的时间留给我们觉得更重要的人。我们假设关系强度直接影响好友间的社交行为的强度，即关系强度越强，社交行为强度越大，反之，社交行为强度越小。令 $x^{(i)}$ 和 $x^{(j)}$ 表示两个好友的 i 和 j 的配置， $y_t^{(ij)}, t = 1, 2, \dots, m$ 表示好友 i 和 j 间 m 次不同的社交行为。然后我们定义 $z^{(ij)}$ 表示 i 和 j 间的关系强度变量，它受 $x^{(i)}$ 和 $x^{(j)}$ 的影响，同时也影响 $y_t^{(ij)}, t = 1, 2, \dots, m$ 。各变量影响表示如下图，整个模型可以被看做是判别模型（条件概率模型）和生成模型（联合概率模型）的混合，上层结构是判别模型($p(Z|X)$)，下层结构是生成模型($p(Y|Z)$)。关系强度模型将变量间因果关系建模成条件依赖，故联合分布可以写成

$$P(z^{(ij)}, Y^{(ij)} | x^{(i)}, x^{(j)}) = P(z^{(ij)} | x^{(i)}, x^{(j)}) \prod_{t=1}^m P(y_t^{(ij)} | z^{(ij)}) \quad (1)$$

$$P(z^{(ij)} | x^{(i)}, x^{(j)}) = N(W^T S(x^{(i)}, x^{(j)}), V) \quad (2)$$

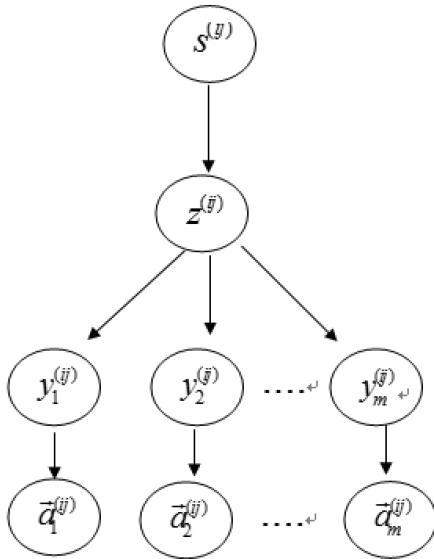
尽管关系强度变量 z 描述了好友间的相似性和社交行为间的关系，但是它的值不是能够直接观测得到，因此需要将 z 当做隐变量，增设一些辅助参数，再用观测的数据和最大似然估计法来求解 z 的值。我们采用被广泛认可的高斯分布来描述给定配置的相似度后，关系强度的条件概率。令 $s_k(z^{(ij)} | x^{(i)}, x^{(j)}) (k = 1, 2, \dots, n)$ 表示好友 i 和好友 j 配置中第 k 个分量的相似度，那么 $z^{(ij)}$ 和 $x^{(i)}, x^{(j)}$ 的依赖可以如下表示



S 是根据 x^i 和 x^j 计算的相似列向量， W 是一个要估计的 n 维权重行向量， V 是高斯模型中的方差。

在上面的模型中，给定了 $z^{(ij)}$ 每一个 $y_t^{(ij)}$ 都是条件独立的。在这里为了简化求解过程，我们用二值变量来表示社交活动，即好友间是否参与了社交活动，例如， y_1^{ij} 表示好友 i 是否和好友 j 互动。我们引入一组辅助变量 $\vec{a}_t = (a_{t1}^{(ij)}, a_{t2}^{(ij)}, \dots, a_{tl}^{(ij)}), t = 1, 2, \dots, m$ 分别表示 $y_1^{(ij)}, y_2^{(ij)}, \dots, y_m^{(ij)}$ 的参数，使得模型能有更好的准确性，

模型如下图所示



这些辅助变量用来描述每次社交行为的影响因素，他们和关系强度是相互独立的，例如如 y_1^{ij} 表示好友 i 是否和好友 j 互动，而 $\vec{a}_1 = (a_1^{ij}, a_2^{ij})$ 表示它的辅助变量， a_1^{ij} 表示 $i \rightarrow j$ 的互动总数， a_2^{ij} 表示 $j \rightarrow i$ 的互动总数。这些辅助变量能够表达好友 i 和好友 j 互动的趋势，很显然，如果互动越多，那么好友 i 是否和好友 j 互动的概率就越大，我们可以用 logistic 函数来描述给定了 $z^{(ij)}$ 和 $\vec{a}_t^{(ij)}$ 的条件概率：

$$P(y_t^{(ij)} = 1 | z^{(ij)}, \vec{a}_t^{(ij)}) = \frac{1}{1 + e^{-(\theta_{t1}a_{t1}^{(ij)} + \theta_{t2}a_{t2}^{(ij)} + \dots + \theta_{tl}a_{tl}^{(ij)} + \theta_{l+1}z^{(ij)} + b)}} \quad (3)$$

其中 $\theta_t = [\theta_{t1}, \theta_{t2}, \dots, \theta_{tl}, \theta_{l+1}]^T$ 是要估计的参数，为了简化公式，令

$$\vec{u}_t^{(ij)} = [\vec{a}_t^{(ij)}]_{z^{(ij)}}$$

$$P(y_t^{(ij)} = 1 | \vec{u}_t^{(ij)}) = \frac{1}{1 + e^{-(\theta_t^T \vec{u}_t^{(ij)} + b)}}$$

最后为了避免过拟合现象，在 w 和 θ 上作约束，认为他们满足高斯先验：

$$P(w) \propto e^{-\frac{\lambda_w}{2} \vec{w}^T \vec{w}}$$

$$P(\theta_t) \propto e^{-\frac{\lambda_\theta}{2} \vec{\theta}_t^T \vec{\theta}_t}, t = 1, 2, \dots, m$$

令 $D = \{(i_1, j_1), (i_2, j_2), \dots, (i_N, j_N)\}$ 表示N个样本。在训练过程中 $\vec{x}^{(ij)}, \vec{y}^{(ij)}$ 和 $\vec{a}^{(ij)}, ((ij) \in D, t = 1, 2, \dots, m)$ 都是可见的。属性的相似程度可以根据 x 计算出来，为了简化表示，令 $s^{(ij)} = s(\vec{x}^i, \vec{x}^j)$ ，将所有可观测量带入方程（1），得联合分布概率

$$\begin{aligned} & P(D | \vec{w}, \vec{\theta}) P(\vec{w}, \vec{\theta}) \\ &= \left(\prod_{(i,j) \in D} P(z^{(ij)}, \vec{y}^{(ij)} | x^{(i)}, x^{(j)}, \vec{w}, \vec{\theta}) \right) P(\vec{w}) P(\vec{\theta}) \\ &= \left(\prod_{(i,j) \in D} \left(P(z^{(ij)} | x^{(i)}, x^{(j)}, \vec{w}) \prod_{t=1}^m P(y_t^{(ij)} | z^{ij}, \vec{\theta}_t) \right) \right) P(\vec{w}) P(\vec{\theta}) \quad (4) \\ &\propto \prod_{(i,j) \in D} \left(e^{-\frac{1}{2v} (\vec{w}^T \vec{s}^{(ij)} - z^{(ij)})^2} \prod_{t=1}^m \frac{e^{-(\vec{\theta}_t^T \vec{u}_t^{(ij)} + b)(1 - y_t^{(ij)})}}{1 + e^{-(\vec{\theta}_t^T \vec{u}_t^{(ij)} + b)}} \right) e^{-\frac{\lambda_w}{2} \vec{w}^T \vec{w}} e^{-\frac{\lambda_\theta}{2} \vec{\theta}_t^T \vec{\theta}_t} \end{aligned}$$

3 模型求解

求解隐变量模型，一般有两种不同的方法。第一种方法，先推断隐变量 z 的联合分布概率，找到 $\vec{w}, \vec{\theta}$ 的点估计，这种方式经常使用需要迭代的EM算法。第二种方法，把隐变量当做一个参数，也就是构造似然函数 $P(\vec{y}, \vec{z}, \vec{w}, \vec{\theta} | \vec{x})$ ，用最大似然估计寻找点估计 $\vec{z}, \vec{w}, \vec{\theta}$ 。本文将使用第二种方法，因为第二种方法可以避免在EM算法中的迭代过程，迭代的过程参数往往难以控制。首先对式取对数后，得到下式

$$\begin{aligned} & L\left(\vec{z}^{\{(i,j) \in D\}}, \vec{w}, \vec{\theta}_t\right) \\ &= \sum_{(i,j) \in D} -\frac{1}{2v} (\vec{w}^T \vec{s}^{(ij)} - z^{(ij)})^2 + \sum_{(i,j) \in D} \sum_{t=1}^m -(1 - y_t^{(ij)}) (\vec{\theta}_t^T \vec{u}_t^{(ij)} + b) - \log(1 + e^{-(\vec{\theta}_t^T \vec{u}_t^{(ij)} + b)}) \quad (5) \\ &\quad - \frac{\lambda_w}{2} \vec{w}^T \vec{w} - \frac{\lambda_\theta}{2} \vec{\theta}_t^T \vec{\theta}_t + C \end{aligned}$$

注意到上式中，二次函数是凸函数，logistic函数也是凸函数。凸函数与

凸函数的和还是凸函数，所以函数 L 是凸函数，所以函数 L 的极值点只有一个。因此，我们用梯度下降算法来找到 $\vec{z}^{(ij)}, \vec{w}, \vec{\theta}_t (t=1, 2, \dots, m)$ ，使得函数 L 值最大。将 L 函数对 $\vec{z}^{(ij)}, \vec{w}, \vec{\theta}_t$ 分别求偏导，得到如下式

$$\frac{\partial L}{\partial z^{(ij)}} = \frac{1}{v} (\vec{w}^T \vec{s}^{(ij)} - z^{(ij)}) + \sum_{t=1}^m \left(y_t^{(ij)} - \frac{1}{1 + e^{-(\vec{\theta}_t^T \vec{u}^{(ij)} + b)}} \right) \theta_{t,l_{t+1}} \quad (6)$$

$$\frac{\partial L}{\partial \theta_t} = \sum_{(ij) \in D} \left(y_t^{(ij)} - \frac{1}{1 + e^{-(\vec{\theta}_t^T \vec{u}^{(ij)} + b)}} \right) \vec{u}_t^{(ij)} - \lambda_\theta \vec{\theta}_t \quad (7)$$

$$\frac{\partial L}{\partial w} = \left(\frac{1}{v} \sum_{(ij) \in D} (z^{(ij)} - \vec{w}^T \vec{s}^{(ij)}) \vec{s}^{(ij)} \right) - \lambda_w \vec{w} \quad (8)$$

梯度下降算法将会迭代 $\vec{z}^{(ij)}, \vec{w}, \vec{\theta}_t$ ，直到似然函数的值收敛。对于 $z^{(ij)}$ 和 θ_t ，应为(6)式和(7)式的根不能直接解得，这样结合梯度算法，在每一次迭代中用牛顿-拉夫森方法更新 $\vec{z}^{(ij)}, \vec{w}, \vec{\theta}_t$ ：

$$z^{(ij) new} = z^{(ij) old} - \frac{\partial L}{\partial z^{(ij)}} / \frac{\partial^2 L}{\partial (z^{(ij)})^2} \quad (9)$$

$$\vec{\theta}_t^{new} = \vec{\theta}_t^{old} - \frac{\partial L}{\partial \theta_t} / \frac{\partial^2 \Omega}{\partial \theta_t \partial \theta_t^T} \quad (10)$$

$$\frac{\partial^2 L}{\partial (z^{(ij)})^2} = -\frac{1}{v} - \sum_{t=1}^m \left(\frac{\theta_{t,l_{t+1}}^2 e^{-(\vec{\theta}_t^T \vec{u}^{(ij)} + b)}}{\left(1 + e^{-(\vec{\theta}_t^T \vec{u}^{(ij)} + b)}\right)^2} \right) \quad (11)$$

$$\frac{\partial^2 \Omega}{\partial \theta_t \partial \theta_t^T} = - \sum_{(ij) \in D} \frac{e^{-(\vec{\theta}_t^T \vec{u}^{(ij)} + b)}}{\left(1 + e^{-(\vec{\theta}_t^T \vec{u}^{(ij)} + b)}\right)^2} \vec{u}_t^{(ij)} \vec{u}_t^{(ij)T} - \lambda_\theta I \quad (12)$$

而 w 的根可以直接用脊状回归求得

$$w^{new} = (\lambda_w I + S^T S)^{-1} S^T \vec{z} \quad (13)$$

其中 $S = [s^{(i_1 j_1)} \ s^{(i_2 j_2)} \ \dots \ s^{(i_N j_N)}]^T$, $Z = [z^{(i_1 j_1)} \ z^{(i_2 j_2)} \ \dots \ z^{(i_N j_N)}]^T$, 算法

框图如下：

```

  While not converged:
  Step1:For each Newton-Raphson step:
    For  $t = 1, \dots, m$ :
      Update  $\vec{\theta}_t$  according to equation(10).
  Step2:For each Newton-Raphson Step:
    For  $(i, j) \in D$ :
      Update  $\vec{z}^{(ij)}$  according to equation(9)
  Step3:Update  $w$  according to equation(13)

```

在一次计算过程中，根据好友 (i, j) 的样本情况，在本模型中有两种不同的用法：第一种，如果好友的 $\vec{x}^{(i)}, \vec{x}^{(j)}$ 和他们的社交行为 $y_1^{(ij)}, \dots, y_t^{(ij)}$ 已知，那么我们可以用算法中的Step2来估计 $z^{(ij)}$ ；第二种，如果社交行为的数据没有获得，那么我们可以直接用等式（2）来推断 $z^{(ij)}$ ，往往由于社交行为的数据很稀少，或者都是短时间的或者难以获得，因此第二种情况在现实的在线社交网络中更加普遍。一旦模型学习好了（ w 参数确定），对于新的数据，隐变量可以直接用上层的变量来计算。

除此之外，由于学习算法中的附加参数 $\vec{\theta}_t$ 也会随 w 而确定，那么我们也可以用它来离线计算好友发生某些社交行为的规律，对社交行为做相关的预测。

8. 基于人工智能的排序

前面章节，我们提出了排序模型，是由核心的信任力，绝对价值，也即关系强度和其评论的内容本身质量来决定的，公示描述为

$$G=F(R(a,b), X)$$

其中R表示a,b之间的关系强度，X为与本身评论内容相关的特征向量，包含查询与评论内容结果的**COSINE**相似度，查询的**Proximity**值等，

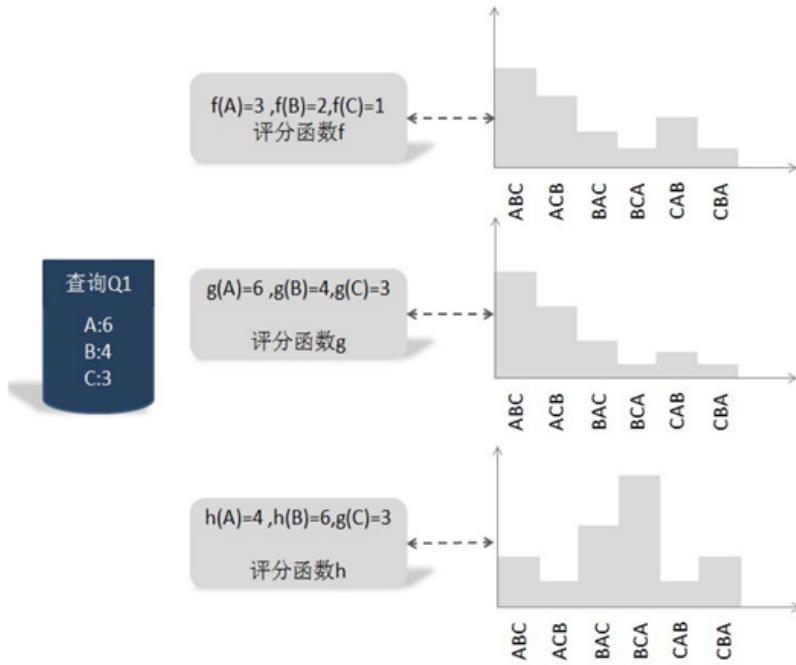
评论的引用数目（评论的出度），评论的被引用数目（评论的入度）。

排序算法的关键也就是通过机器学习的方式，学习到一个打分函数F，这个打分函数能够尽可能的接近人工的打分结果，通过该F函数对**QUERY**得到的每一个结果进行打分，从而排序得到最终结果。通过上节，我们建模中考虑的主要特征因素，为关系强度，含查询与评论内容结果的**COSINE**相似度，查询的**Proximity**值等，评论的引用数目（评论的出度），评论的被引用数目（评论的入度）。关键问题是，我们如何准备训练数据，以及如何训练得到最优的打分函数。

对于数据集问题，我们将相关程度{Perfect, Excellent, Good, Fair, Bad} 共五个类别，训练数据格式，类别标签，查询的query，特征值1，特征2，特征3，特征4等

Category	query	feature1	feature 2	feature 3	feature 4
Perfect	1	0.2	3	2	1
Bad	1	0.9	1	0	3

对于训练问题，我们采取经典的ListWise方法，它是基于搜索结果的排列组合概率来训练的.



首先解释一下搜索结果排列组合概率含义，假设我们对于某类相似query搜索结果有3个结果文档ABC，那么可能的搜索结果顺序有6种，ABC, ACB, BAC, BCA, CAB, CBA, 假设G为理想的打分模型（通过人工标注统计得到）。我们可以用贝叶斯学习，SVM等众多学习方法得到学习函数，衡量两个概率函数的相似程度，我们采用KL散度函数。

如上图，F, H是通过数据学习到两个函数，通过计算F,H函数对于G函数的KL散度，来选择最佳的打分函数。

9.代币的发行和分配

发行总量：65亿个COOMI，75000COOMI=1ETH

融资：30%（19亿个COOMI），计划募捐15000个ETH，硬顶20000个ETH。

10发展计划



11团队简介

团队背景

成员来自中日两国，大部分来自日本互联网巨头Recruit集团，我们拥有共同的工作背景和企业文化。（ Recruit是日本市值最高，规模最大的互联网公司）

Heming | Cofounder | CEO



曾在日本著名的Rotary奖学金资助下，攻读博士前期课程，研究互联网用户行为。2014年创立了日本版“饿了么”，引起一时轰动。现供职日本互联网巨头Recruit市场营销部门，在数字营销、活动运营、用户运营和团队管理，项目推进等方面拥有丰富的实战经验。梦想用技术和商业的力量让社会变得更加美好。区块链3.0让大家看到了希望，而我们要努力让这希望变成现实。

Nami | Cofounder | COO



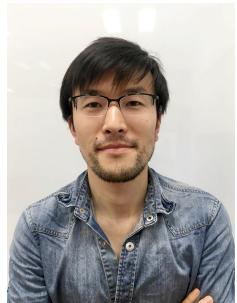
现任日本互联网巨头Recruit的口碑平台的最高负责人。深知口碑价值和行业痛点的同时，一直在探索解决方案。当遇到区块链和这个团队的时候，就知道改变时代的机会来了。深信我们的产品，能用区块链这个支点撬动整个行业，Coomi有潜力成为下一个谷歌，为用户带来全新的价值。

Tsukikawa Yuu | Cofounder | CMO



Ink联合创始人。北京大学、早稻田大学双学位学士。区块链早期投资人。2010年-12年担任湖南卫视「天天向上」节目主持人。此后在日本互联网巨头Recruit担任经营战略策划。2016年末作为日本第2大交易所Zaif的战略伙伴活动。2017年担任Sora Venture的Adovisor。目前主要负责Ink全球战略，纸贵科技战略研究顾问，以及纸贵亚洲市场战略研究活动。

Erichin | Founder | CTO



毕业于清华大学，日本著名区块链媒体平台ALIS元创始团队成员，日本最流行的以太坊钱包之一TokenPocket核心创始人。
目前作为一名工程师就职于日本Recruit公司。热衷于区块链技术的应用和开发，坚信区块链应用的普及将会带来多个领域的变革。

Koan | Founder | 算法工程师



毕业于中科院计算所，曾就职于新浪邮件反作弊部门，现Indeed (Recruit子公司) 高级工程师。熟悉算法及后端设计，开发，运维，在搜索引擎营销，邮件营销，用户增长策略方面经验丰富。

Haruchern | Founder | 全栈工程师



毕业于京都大学社会情报学研究科，曾就职于百度移动云事业部，现Indeed (Recruit子公司) 广告系统工程师，对前端技术，设计抱有极大热情，亦熟悉在线广告生态及广告投放业务。

Belle | Founder | 市场运营



曾任日本最大人力资源公司6年资深猎头经验。担当过Amazon,LINE,Metaps,Sony,SAMSUNG,GSK,Pfizer,L'ORÉAL等约500家国际国内知名及初创公司招聘，经手1000人以上成功转职。因此，在日本商业界和投资界拥有丰富的人脉。现任日本最大B2B猎头公司资源网，负责年商约20亿元发展战略及商品企划。

Alaric | Founder | UI/UX设计师



7年互联网从业经验，曾任职于亚信，暴风，荣联等多家上市企业，任相关产品主设计师，设计总监等职，涉及互联网，教育，OTO，VR/AR，通讯，车网，互金等行业，两次自主创业经历。2016年起关注区块链领域。

Yichen | Founder | 区块链研究员



毕业于清华大学能源与动力工程系，博士学位。现任职于剑桥大学CARES研究中心，常驻于新加坡。具有数学建模与科学计算的丰富经验。近年来开始深度关注区块链领域的技术发展与在东南亚地区的应用前景。

※顾问、部分投资人、投资机构、合作机构等信息在第二版白皮书中公布

12代码部分

下一版本对外公开

13免责声明

本文档只用于传达信息之用途，并不构成买卖项目股份或证券的相关意见。任何类似的提议或征价将在一个可信任的条款下并在可应用的证券法和其它相关法律允许下进行，以上信息或分析不构成投资决策，或具体建议。本文档不构成任何关于证券形式的投资建议，投资意向或唆唆投资。本文档不组成也不理解为提供任何买卖行为，或任何邀请买卖、任何形式证券的行为，也不是任何形式上的合约或者承诺。



Coomi明确表示相关意向用户明确了解**Coomi平台**的风险，投资者一旦参与投资即表示了解并接受该项目风险，并愿意个人为此承担一切相应结果或后果。

Coomi明确表示不承担任何参与**Coomi项目**造成的直接或间接的损失包括: (i) 本文档提供所有信息的可靠性(ii) 由此产生的任何错误，疏忽或者不准确信息 (iii)或由此导致的任何行为。

COOMI是以**Coomi平台**为其使用场景之一的数字**Token**。**COOMI**不是一种投资。我们无法保证**COOMI**将会增值，其也有可能在某种情况下出现价值下降。鉴于不可预知的情况，本白皮书列出的目标可能发生变化。虽然团队会尽力实现本白皮书的所有目标，所有购买**COOMI**的个人和团体将自担风险；

COOMI不是一种所有权或控制权。控制**COOMI**并不代表对**Coomi**或**Coomi应用**的所有权，**COOMI**并不授予任何个人任何参与、控制、或任何关于**Coomi**及**Coomi应用**决策的权利。

14风险提示

数字资产投资作为一种新的投资模式，存在各种不同的风险，潜在投资者需谨慎评估投资风险及自身风险的承受能力：

Token销售市场风险

由于**Token**销售市场环境与整个数字货币市场形势密不可分，如市场行情整体低靡，或存在其他不可控因素的影响，则可能造成**Token**本身即使具备良好的前景，但价格依然长期处于被低估的状态。

监管风险

由于区块链的发展尚处早期，包括我国在内全球都没有有关融资过程中的前置要求、交易要求、信息披露要求、锁定要求等相关的法规文件。并且目前政策会如何实施尚不明朗，这些因素均可能对项目的投资与流动性产生不确定影响。而区块链技术已经成为世界上各个主要国家的监管主要对象，如果监管主体插手或施加影响则**Coomi**应用或**COOMI**可能受到其影响，例如法令限制使用、销售**Token**诸如**COOMI**有可能受到限制、阻碍甚至直接终止**Coomi**应用和**COOMI**的发展。

竞争风险

随着信息技术和移动互联网的发展，以“比特币”为代表的数字资产逐渐兴起，各类去中心化的应用持续涌现，行业内竞争日趋激烈。但随着其他应用平台的层出不穷和不断扩张，社区将面临持续的运营压力和一定的市场竞争风险。

人员流失风险

Coomi集聚了一批在各自专业领域具有领先优势和丰富经验的技术团队和顾问专家，其中不乏长期从事区块链行业的专业人员以及有丰富互联网产品开发和运营经验的核心团队。核心团队的稳定和顾问资源对**Coomi**保持业内核心竞争力具有重要意义。核心人员或顾问团队的流失，可能会影晌平台的稳定运营或对未来发展带来一定的不利影响。

资金匮乏导致无法开发的风险

由于创始团队筹集的**Token**价格大幅度下跌或者开发时间超出预计等原因，都有可能造成团队开发资金匮乏，并由此可能会导致团队极度缺乏资金，从而无法实现原定开发目标的风险。

私钥丢失风险

购买者的**COOMI**在提取到自己的数字钱包地址后，操作地址内所包含内容的唯一方式就是购买者相关密钥(即私钥或是钱包密码)。用户个人负责保护相关密钥，用于签署证明资产所有权的交易。用户理解并接受,如果他的私钥文件或密码分别丢失或被盗,则获得的与用户帐户(地址)或密码相关的**COOMI**将不可恢复，并将永久丢失。最好的安全储存登录凭证的方式是购买者将密钥分开到一个或数个地方安全储存，且最好不要储存在公用电脑。

黑客或盗窃的风险

黑客或其它组织或国家均有以任何方法试图打断**Coomi**应用或**COOMI**功能的可能性，包括但不限于拒绝服务攻击、Sybil攻击、游袭、恶意软件攻击或一致性攻击等。

未保险损失的风险

不像银行账户或其它金融机构的账户，存储在**Coomi**账户或相关区块链网络上通常没有保险保障，任何情况下的损失，将不会有任何公开的个体组织为你的损失承保。

核心协议相关的风险

Coomi平台目前基于以太坊开发，因此任何以太坊发生的故障，不可预期的功能问题或遭受攻击都有可能导致**COOMI**或**Coomi**平台以难以预料的方式停止工作或功能缺失。

系统性风险

开源软件中被忽视的致命缺陷或全球网络基础设施大规模故障造成的风险。虽然其中部分风险将随着时间的推移大幅度减轻，比如修复漏洞和突破计算瓶颈，但其他部分风险依然不可预测，比如可能导致部分或全球互联网中断的政治因素或自然灾害。

漏洞风险或密码学加速发展的风险

密码学的加速发展或者科技的发展诸如量子计算机的发展，或将破解的风险带给**Coomi平台**，这可能导致**COOMI**的丢失。

应用缺少关注度的风险

Coomi应用存在没有被大量个人或组织使用的可能性，这意味着公众没有足够的兴趣去开发和发展这些相关分布式应用，这样一种缺少兴趣的现象可能对**COOMI**和**Coomi**应用造成负面影响。

不被认可或缺乏使用者的风险

首先**COOMI**不应该被当做一种投资，虽然**COOMI**在一定的时间后可能会有一定的价值，但如果**Coomi**不被市场所认可从而缺乏使用者的话，这种价值可能非常小。有可能发生的是，由于任何可能的原因，包括但不限于商业关系或营销战略的失败，**Coomi**平台和所有的众售资金支持的

后续营销将不能取得成功。如果这种情况发生，则可能没有这个平台就没有后续的跟进者或少有跟进者，显然，这对本项目而言是非常不利的。

应用存在的故障风险

Coomi平台可能因各方面可知或不可知的原因故障(如大规模节点宕机)，无法正常提供服务，严重时可能导致用户**COOMI**的丢失。

应用或产品达不到自身或购买者的预期的风险

Coomi应用当前正处于开发阶段，在发布正式版之前可能会进行比较大的改动，任何**COOMI**自身或购买者对**Coomi**应用或**COOMI**的功能或形式(包括参与者的行)的期望或想象均有可能达不到预期，任何错误地分析，一个设计的改变等均有可能导致这种情况的发生。

无法预料的其它风险

基于密码学的**Token**是一种全新且未经测试的技术，除了本白皮书内提及的风险外，此外还存在着一些创始团队尚未提及或尚未预料到的风险。此外，其它风险也有可能突然出现，或者以多种已经提及的风险的组合的方式出现。