



Huobi Research

相关报告

《【火量学派 4】 缠论在数字资产上的应用（一）：脉络梳理》2018-9-27

《【火量学派 3】 海龟交易法则在数字资产上的应用》2018-8-31

作者

袁煜明

胡智威

翁溢铭

乔晓峰

huobiresearch@huobi.com

【火量学派 5】 VPIN 在高波动市场中的应用

2019 年 4 月 10 日

摘要

作为市场流动性的提供者---做市商在日常经营中面临的主要风险为交易对手的信息优势导致的逆向选择风险。VPIN (Volume synchronized probability of informed trading) 是传统交易市场中较为流行的一个测算指标。

依据不同交易量的归类标准，VPIN 又可以划分为 TR-VPIN (Tick Rule VPIN) 及 BV-VPIN (Bulked Volume VPIN)等。其原理是知情交易者的加入会导致原本稳定的分布发生偏移。

火币研究院对 VPIN 模型进行分析并编写代码，以 1token 提供的多个交易所的实盘数据及回测平台进行回测。主要测试场景包括 4 月初比特币价格的快速拉升和 1 月份 EOS 价格的快速下跌等。

结果显示，无论是数字资产价格的暴涨或暴跌，VPIN 时常会出现大幅度的提升，具有一定的预测效果，可视为是一种波动率领先指标，可对期权交易、做市商提供流动性、交易所风控管理等具有指导意义。

目录

1. VPIN 简介3

2. 模型原理4

3. 算法步骤6

4. 实盘测试8

5. 总结 12

参考资料 13

报告正文

1. VPIN 简介

作为市场流动性的提供者——做市商在日常经营中面临的主要风险为交易对手的信息优势导致的逆向选择风险。为了测量做市商在交易中面临的逆向选择风险，行业里研究采用了各种指标，价格波动率是其中非常流行的一种。

但是在高频交易的环境下，波动率已经不是最可靠的预测指标。对于高频交易中的短期大幅波动，在传统金融市场中一个更为流行的指标是 VPIN。

VPIN 起源于做市商关注的主要市场风险来源——PIN(probability of informed trading)，也就是知情交易者概率。简单来说，当一个做市商在市场中提供报价时，需要考虑的一个核心问题就是交易对手是知情交易者 (informed trader) 的概率，因为和知情交易者或者有信息优势的交易者做交易时，该笔交易的预期收益肯定是负的。

VPIN (Volume synchronized probability of informed trading) 是基于 PIN 指标做的改进。因为直接计算 PIN 是很困难的，Easley (2015) 提出了用交易量时间来替代物理时间，可以实现对实时条件下的交易毒性的测算。依据不同交易量的归类标准，VPIN 又可以划分为 TR-VPIN (tick rule VPIN) 及 BV-VPIN (Bulked volume VPIN)。Easley (2015) 在后来的文章中提到了 BV-VPIN 是一种更优的算法，所以我们在本文中的算法为 BV-VPIN 的算法。

对于频繁发生短期大幅波动的数字资产二级市场，火币研究院对 VPIN 模型分析并以 Python 进行算法代码实现，以 1token 提供的多个交易所的实盘数据进行回测。对比特币、EOS 等交易历史数据的测试结果显示，VPIN 在数字资产交易中具有一定预测效果，可视为是一种波动率领先指标，可对期权交易、做市商提供流动性、交易所风控管理等具有指导意义。

2. 模型原理

在开始计算前，我们先回顾 VPIN 的模型逻辑。做市商在日常交易中的关注的核心指标之一就是交易对手有信息优势的概率，也就是 PIN (probability of informed trading)。这个模型对 PIN (probability of informed trading) 对做市商的报价的影响进行了建模。模型的推理过程比较复杂，我们在这主要陈述模型的核心假设及结论。

The Glosten-Milgrom Model

这个模型是市场微结构中关于 sequential model 的原型。模型中假设一个资产在 T 时会有两个价格分别为 $\{\bar{S}, \underline{S}\}$ ，其中 $\bar{S} > \underline{S}$ 。不知情的交易员只知道该资产的价格区间，但并不知道这个资产在 T 时间的最终价格。在这段时间的订单流中，有 μ 的比例的订单被假定为知情的交易单 (informed trades)，这些交易单的交易员了解该资产的真实价值。市场在交易开始前不知情的市场参与者对上涨和下跌的假设为 $\text{Prob}(S_T = \bar{S}) = \delta$ ， $\text{Prob}(S_T = \underline{S}) = 1 - \delta$ 。如果做市商通过贝叶斯概率不断调整报价，这个 δ 会随着市场价格的波动市场会不断调整该预期。最后在均衡状态下做市商报价的价差会停留在 $A - B = \frac{4(1-\delta_t)\delta_t\mu(\bar{S}-\underline{S})}{1-(1-2\delta_t)^2\mu^2}$ (A 为 ask price, B 为 bid price)。这个模型描述了对 informed trades 对做市商报价行为的影响。但结果为这个模型是有限时间模型，后人在此模型上提出了改进。

The Easley-O'Hara Model

这个模型是产生 PIN 这个关于信息流毒性指标的模型。在 Easley 的一系列文章中，交易被看成做市商和交易员之间的博弈。在这个模型中，交易被看成一系列的 trading periods $i = 1, \dots, I$ 。在每个 period 的开始，会有 α 概率发生影响资产价格的事件。如果这个事件发生，对于资产来说可能时

好消息，也可能是坏消息。在第 i 个 period 结束时，如果是个好消息，那么该资产的价值应为 \bar{S}_i ，如果是坏消息，该资产的价值应为 \underline{S}_i 。好消息发生的概率为 $(1-\delta)$ ，坏消息发生的概率为 δ ，可以将 δ 假设为一个先验概率。在消息发生后，订单流以泊松分布的进入交易所。知情交易者知道信息的好坏，他们将会在好消息时买入，坏消息时卖出。模型中假设知情交易者的到达率为 μ ，不知情交易者的到达率为 ε 。

这个模型的演绎结果就是做市商报价的价差 $A-B=[\bar{S}-\underline{S}]PIN$ （ A 为 ask price， B 为 bid price）。这个模型的演绎出 informed trading 的比例越大，做市商报价的价差越大。这点也和日常自觉相符。

从上面两个模型，我们可以得知如何从理论角度构建 PIN 这个指标。如果要将这个指标应用的到交易中，则必须对模型参数进行估计。标准的解决方法去求解这个模型并计算 PIN 需要对 $(\alpha, \delta, \mu, \varepsilon)$ 参数进行预测，这会使得实时预测变得十分困难。我们采用了 VPIN 的方法来预测 PIN，这样实时计算这个指标变得可能。在数字资产的二级市场中，这个想法也更加符合直觉。交易的发生通常和信息高度相关，和时间的相关性反而小，这也支持了用交易量时间而非物理时间对此事件进行建模。

在 VPIN 的模型中，会对交易进行分类。分类的分发是先将一段时间内的交易加总，然后用这段时间的起始价格和终止价格之差的分布将交易量分类。具体公式如下：

$$V_{\tau}^B = \sum_{i=(\tau-1)V+1}^{\tau V} Z\left(\frac{\Delta P_i}{\sigma \Delta P}\right)$$

$$V_{\tau}^S = \sum_{i=(\tau-1)V+1}^{\tau V} [1 - Z\left(\frac{\Delta P_i}{\sigma \Delta P}\right)]$$

将交易量分类后就可以进一步计算订单中的不平衡 (order imbalance)。 $E[OI_t^*] = E[|V_t^S - V_t^B|]$, 模型假设这种订单中的不平衡在调整后会变成一种均匀的分布。

文献中考虑了更复杂的订单流模型, 我们为了讨论的简单就提供了最简单的订单流模型。

从 Easley, Engle, O' Hara and Wu(2008) 中我们知道知情订单的到达速率为 $E[V_t^S - V_t^B] = \alpha\mu(2\delta - 1)$ and $E[|V_t^S - V_t^B|] \approx \alpha\mu$

文章中也假设了订单预计到达的速率为 $\frac{1}{n} \sum_{\tau=1}^n (V_\tau^S + V_\tau^B) = V = \alpha\mu + 2\varepsilon$

所以 VPIN 最终可以被简化为 $VPIN = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} \approx \frac{\sum_{\tau=1}^n |V_\tau^S - V_\tau^B|}{nV}$

这个指标和上文中的理论模型是否相符呢, 答案是肯定的。有兴趣的读者可以通过蒙特卡洛模拟验证模型中假定 $(\alpha, \delta, \mu, \varepsilon)$, 并用这些参数生成订单流。计算结果中的 PIN 与 VPIN 的计算结果十分接近。

整个模型的推理过程涉及到较多的理论证明, 有兴趣的读者可以直接阅读本报告的参考文献原文。而模型背后的思想并不复杂, 在日常交易策略开发中也值得借鉴。模型的假设是市场在均衡状态下订单流会有相对稳定的分布; 而知情交易者的加入会导致原本稳定的分布发生偏移。VPIN 就是对捕捉这种偏移做的尝试。

3. 算法步骤

在知道 VPIN 原理后我们可以尝试将 VPIN 进行算法化, 详见 Prado(2012)。

A. 输入

1. 某个币的交易数据的时间序列

- T: 交易发生时间 T_i
- P: 交易发生时的标的交易价格 P_i
- V: 交易量 V_i

2. V: 交易量

3. n: 用来预测 VPIN 使用的交易样本

B. 将交易量切成同大小的份数

1. 将交易单按时间顺序排列

2. 计算

3. 将 ΔP_i 扩展，使 ΔP_i 的数量与对应的 V_i 相同，在扩展后将获得 $I = \sum_i V_i$ 个 ΔP_i

4. 将 ΔP_i 从新排序 $i=1, \dots, I$.

5. 将 $\tau = 0$

6. $\tau = \tau + 1$

7. 如果 $I < \tau * V$, 那么跳到 11 步

8. 对于 i 属于 $[(\tau - 1)V + 1, \tau V]$ 将交易量进行买方驱动和卖方驱动的分类

- $V_\tau^B = \sum_{i=(\tau-1)V+1}^{\tau V} Z\left(\frac{\Delta P_i}{\sigma \Delta P}\right)$
- $V_\tau^S = \sum_{i=(\tau-1)V+1}^{\tau V} [1 - Z\left(\frac{\Delta P_i}{\sigma \Delta P}\right)]$

9. 对篮子中的交易量进行标签

10. 回到第六步

11. 将 $L = \tau - 1$

C. 使用 VPIN 计算公式

如果 L 大于 n ，将会有足够的信息用于计算

$$VPIN_L = \frac{\sum_{i=(\tau-1)V+1}^{\tau V} |V_{\tau}^S - V_{\tau}^B|}{nV}$$

D. VPIN 准确性的验证

可以使用蒙特卡洛算法来验证 VPIN 的有效性。方法可以是假设一个订单流背后的因子为 $(\alpha, \mu, \varepsilon)$ 。用模拟的订单流计算出的 VPIN 和真实的 PIN 将会非常接近。

4. 实盘测试

本报告中，我们用 1token (<https://1token.trade/>) 提供的多个交易市场的实盘逐笔成交数据，测试比较 VPIN 在资产价格暴涨暴跌时的表现。1token 是现在市场上为数不多的优质数据商，他们的数据具有很好的实时性及准群度，正好适合这种高频数据的回测。同时，1token 的回测平台也可减少回测代码的开发成本。

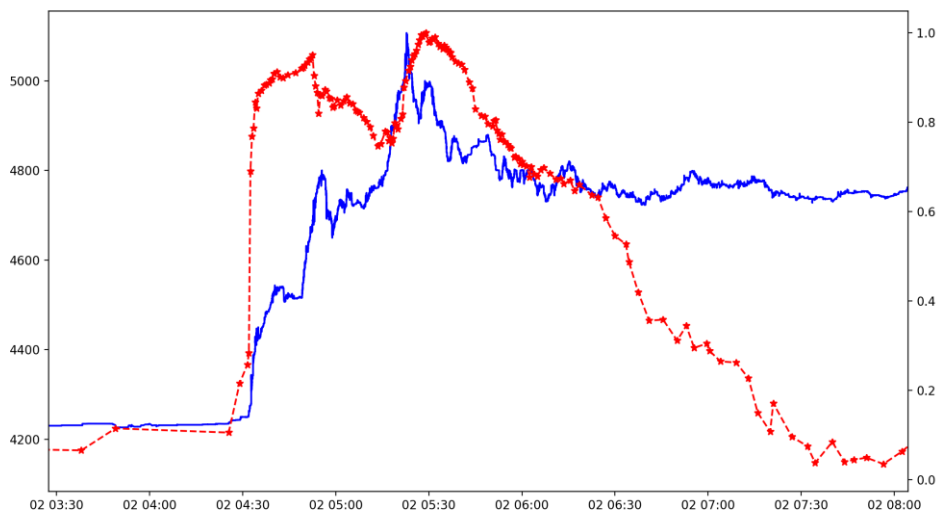
测试 1

比特币的价格在 2019 年 4 月初快速从 4100 美元附近一度快速飙升至 5000 美元及以上。在该测试场景中，我们选择多个交易所的 BTC/USDT 的逐笔数据进行测试。

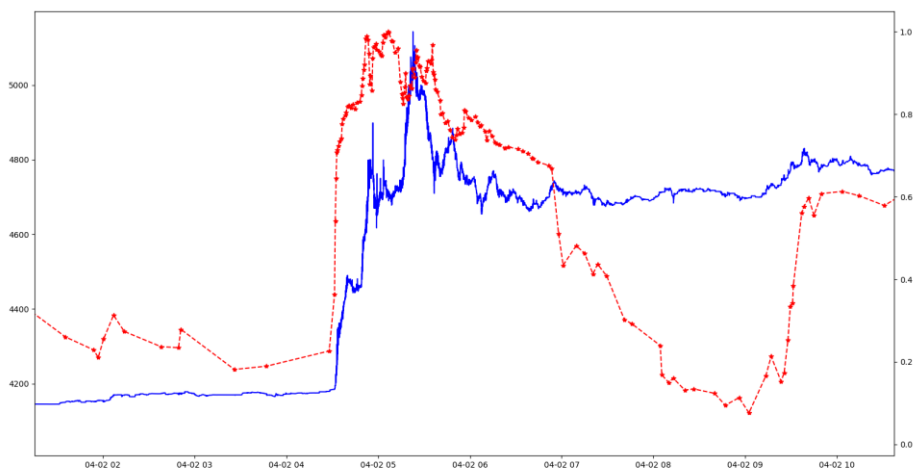
测试参数均设置为：日平均交易量的 1/50 设置“桶”的大小 (Bucket

Size), VPIN 的计算选取最近 25 个桶。

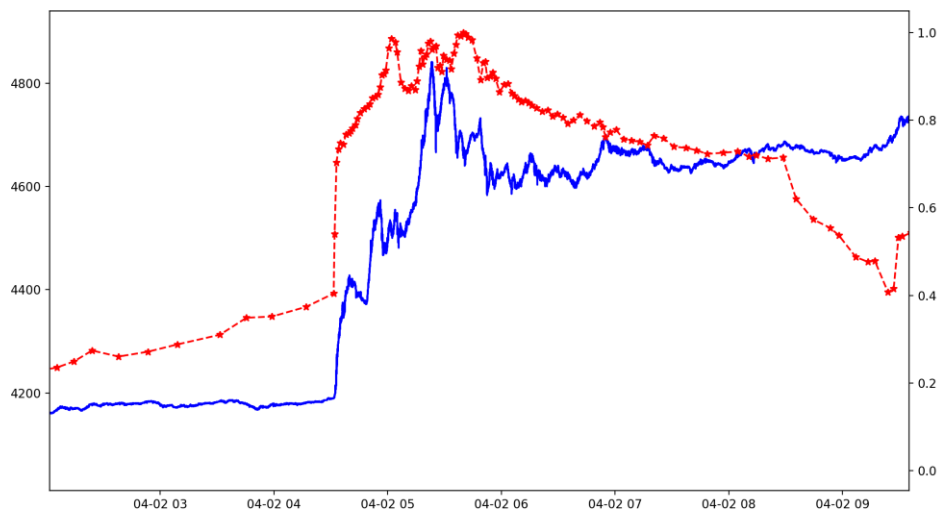
- 测试数据是 Bitfinex 的实盘交易数据结果为：



- 测试数据是 Kraken 的实盘交易数据结果为：



- 测试数据是币安的实盘交易数据结果为：



其中，上述各图的蓝线均表示 BTC/USDT 的交易价格，值对应在左侧纵坐标轴；红线为 VPIN 的累计概率分布值（CDF），值对应于右侧纵坐标轴上。

可以看到，在多个交易所的 4 月 2 日 4:30-5:00（UTC 时间）之间随着比特币价格飙升至 4500 美元附近，VPIN 的 CDF 值也一同快速上涨并在 5:00 以后的一小段时间内保持在很高的水平；而比特币价格随后也在 5:30 左右飙升至了 5000 美元以上的近期价格高点。

因此可看到，VPIN 在各个交易所中均对价格的飙升会起到一定的预先提示效果。

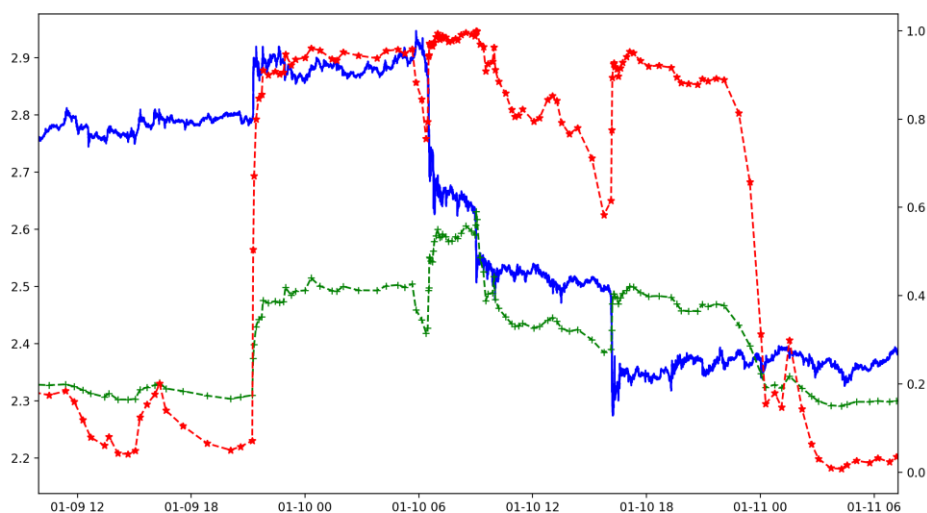
测试 2:

EOS 的价格近期经历了较大幅度的波动。EOS/USDT 的价格在 2019 年 1 月 10 日经历了大幅、快速下跌，从 2.9 左右下跌至 2.4 左右。在该测试场景中，我们选择 EOS/USDT 测试价格下跌的情况。具体测试数据如下：

- 测试数据：币安，Bitfinex 交易所的 EOS/USDT 交易对在 1 月份的逐笔成交数据

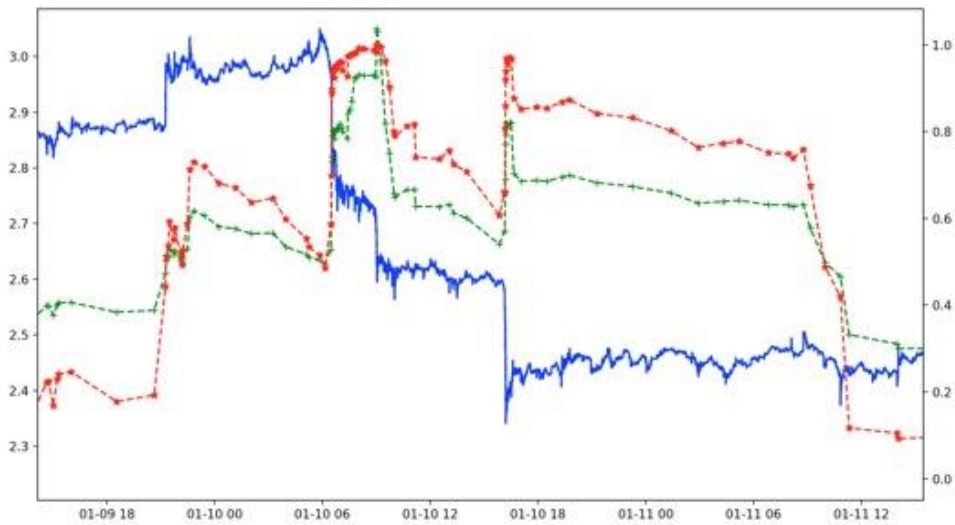
- 测试参数：日平均交易量的 1/50 设置“桶”的大小 (Bucket Size),
VPIN 的计算选取最近 25 个桶

测试数据集是币安交易所的结果：



其中，蓝线表示 EOS/USDT 的交易价格，值对应在左侧纵坐标轴；绿线为 VPIN 指标的数值、红线为 VPIN 的累计概率分布值 (CDF)，值对应于右侧纵坐标轴上。

测试数据集是 Bitfinex 的结果：



从图上可以看到 VPIN，特别是在币安交易所的 CDF，在 EOS 价格大幅下跌之前即保持在很高的水平（CDF>0.8）以上。说明在该数字资产价格剧烈波动期间及大幅波动前，订单的不平衡会有明显的统计异常，并且 VPIN、CDF 均有比较好的指示效果。VPIN 在 Bitfinex 交易所的水平会低于币安交易所，愿意在于 VPIN 在领先的交易所的效果会更强，在滞后交易所会相对更弱，其中的区别我们会在后面的文章中讨论。

5. 总结

从上述结果可看到，VPIN 及其累积概率分布函数 CDF，在资产价格暴涨暴跌时常会出现大幅度的提升，具有一定的预测效果。因此，VPIN 在数字资产交易实践中，可视为是一种波动率领先指标。当出现 VPIN 的大幅度快速提升时，则提示交易者需注意该数字资产的波动。由于涨跌均有可能，因此 VPIN 在期权交易的场景中将会更有指导意义。此外，也可应用在做市商提供流动性、交易所风控管理等方面。

本报告由于篇幅所限，选取了近期有代表性的行情时间段及交易品种。如果想更严谨的测试 VPIN 的效果，需要进一步检验 VPIN 和价格波动的相关性及 VPIN 与价格波动间的条件概率。这些测试我们会在未来的研究报告中继续介

绍。值得注意的是，对于 VPIN 的效果，在学术圈也有争议。Anderson (2014) 中质疑了 VPIN 的有效性，Easley (2014) 中也对该质疑提出了反驳，有兴趣深入学习的读者也可以继续阅读该文章。

对于报告正文中的 VPIN 算法及 Python 代码的相关实现，欢迎联系火币研究院进一步探讨。

参考资料

- [1] Marcos M.Lopez De Prado. *Advances in High Frequency Strategies* (2012), 76-80
- [2] Glosten, L. R. and P. Milgrom(1985):"Bid, ask and transaction prices in a specialist market with heterogeneously informed traders", *Journal of Financial Economics*, 14, 71-100
- [3] Easley, D. and M. O'Hara(1992b): "Time and the process of security price adjustment", *Journal of Finance*, 47, 576-605
- [4] Andersen T G, and Bondarenko O. VPIN and the Flash Crash[J]. *Journal of Financial Markets*, 2014.17:1-46
- [5] Easley D, de Prado M M L, and O'Hara M. VPIN and the flash crash: A rejoinder[J]. *Journal of Financial Markets*, 2014. 17:47-52
- [6] Easley, D., R.F.Engle, M. O'Hara and L. Wu(2008):"Time-Varying Arrival Rates of Informed and Uninformed Traders", *Journal of Financial Econometrics*

火币区块链应用研究院

关于我们:

火币区块链应用研究院（简称“火币研究院”）成立于2016年4月，于2018年3月起全面拓展区块链各领域的研究与探索，主要研究内容包括区块链领域的技术研究、行业分析、应用创新、模式探索等。我们希望搭建涵盖区块链完整产业链的研究平台，为区块链产业人士提供坚实的理论基础与趋势判断，推动整个区块链行业的发展。

联系我们:

咨询邮箱:	huobiresearch@huobi.com
简书公众号:	火币区块链研究院 https://www.jianshu.com/u/a234cae11e43
Twitter:	Huobi_Research https://twitter.com/Huobi_Research
Medium:	Huobi Research https://medium.com/@huobiresearch
Facebook:	Huobi Research https://www.facebook.com/Huobi-Research-655657764773922
Website:	https://research.huobi.cn

免责声明:

1. 火币区块链研究院与本报告中所涉及的数字资产或其他第三方不存在任何影响报告客观性、独立性、公正性的关联关系。

2. 本报告所引用的资料及数据均来自合规渠道，资料及数据的出处皆被火币区块链研究院认为可靠，且已对其真实性、准确性及完整性进行了必要的核查，但火币区块链研究院不对其真实性、准确性或完整性做出任何保证。

3. 报告的内容仅供参考，报告中的事实和观点不构成相关数字资产的任何投资建议。火币区块链研究院不对因使用本报告内容而导致的损失承担任何责任，除非法律法规有明确规定。读者不应仅依据本报告作出投资决策，也不应依据本报告丧失独立判断的能力。

4. 本报告所载资料、意见及推测仅反映研究人员于定稿本报告当日的判断，未来基于行业变化和数据的更新，存在观点与判断更新的可能性。

5. 本报告版权仅为火币区块链研究院所有，如需引用本报告内容，请注明出处。如需大幅引用请事先告知，并在允许的范围内使用。在任何情况下不得对本报告进行任何有悖原意的引用、删节和修改。