★CACM 中国版★

计算机协会通讯

CACM.ACM.ORG

2016 年 4 月第 59 卷第 4 期

押宝比特币

病毒式传播之外

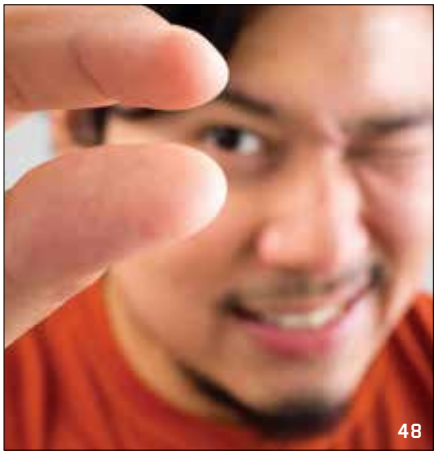
利用多模态生物特征识别技术来提高移动设备的安全性

后缀树的 40 年历程

给系统管理员的建议

Association for Computing Machinery

观点



实践

投稿文章

评论文章

研究亮点

观看此独家《通讯》视频中作者对本研究的讨论：
http://cacm.acm.org/videos/secure-multiparty-computations-on-bitcoin

关于封面：
本月的两篇研究亮点文章聚焦了比特币与这一数字货币背后的技术以及其安全问题。S. Meiklejohn 等人分析了比特币网络，侧重于潜在匿名与实际匿名之间越来越大的间隙。M. Andrychowicz 等人提出利用比特币来设计协议，即使没有受信第三方也能确保安全。封面插图照片由 Kollected Studio 提供。

从左至右图片分别由 RACHATA TEYPARSIT、PAPUCHALKA 和 KOLLECTED STUDIO 提供。

Association for Computing Machinery
Advancing Computing as a Science & Profession

ACM通讯
(ISSN 0001-0782) 由计算机协会（2 Penn Plaza, Suite 701, New York, NY 10121-0701）按月发行。

Association for Computing Machinery

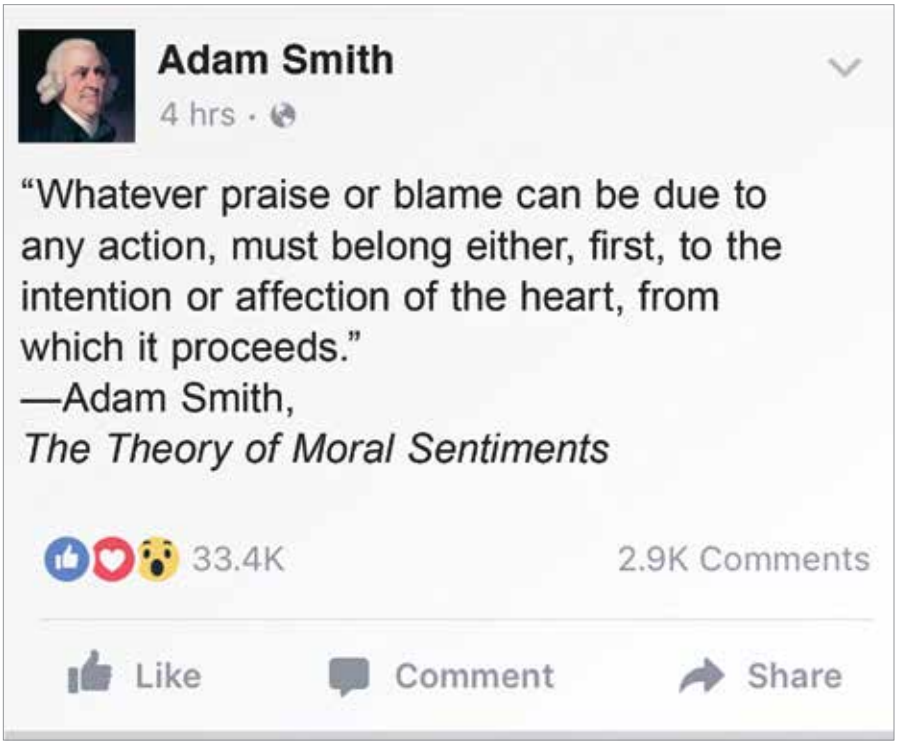
DOI:10.1145/2818992

Manuel Cebrian, Iyad Rahwan, 和 Alex “Sandy” Pentland

观点
病毒式传播之外

社交媒体的广泛使用
并没有引起重大的社会变化。

与社交媒体黄金时代相伴的是全球性的领导危机。近年来我们对任何主要的全球问题都无能为力，由此可见一斑。[32] 当前，没有人无论是充满魅力的领导者还是无名英雄，都看似无法使问题的热度保持足够长的时间，从而动员社会采取行动。由于这种领导力真空，各种社会发展似乎都身陷囹圄、停滞不前。社交媒体却被赞扬为提高集体意识、动员社会的终极工具，且已成熟并被广泛使用。赞誉之下，又如何会导致上述情形？在此，我们争辩道，社交媒体技术与“权力终结”[18] 的共存绝不是一种巧合，而这展现出了 21 世纪的第一个技术 - 社会悖论。



最近几年，我们目睹了社交媒体在具有历史意义的社会动员事件中发挥了重大作用，例如阿拉伯之春、占领华尔街运动、乌克兰的亲欧盟示威以及英国骚乱和波士顿马拉松爆炸案追捕所引发的混乱。电子社交平台，特别是 Facebook 和 Twitter 被着重强调为这些动员的催化剂。数据可用性使我们首次可以详细观察这些事件的演变。[10,11,13,33] 这些事件的分析显示政治积极分子发现难以利用社交媒体进行大规模动员；即便他们成功了，除非他们能够广泛动员政治家、机构和社会，否则很难维持抗议的聚焦点。结果，大多数事件突然爆发，使我们关注了几天时间，然后便逐渐被人遗忘，并没有造成什么实质性影响。基于我们对社会动员的所有了解，为什么社交媒体不能是建设性社会变革的更可靠渠道呢？

一个相关的观察显示即使国家情报机构密切监控个人社交媒体网络，它们仍不能预测社会动乱。最近来自爱德华斯诺登及其他人的全球监控泄密也没提供哪怕一个这样的案例：社交媒体的分析能预测一个社会动乱事件或公众运动。社交媒体一直以来更擅于为无法预测、突发性的动员提供燃料，而不是按部就班、细致周密地构建持续性的社会变革。

相互协作的集体行动是所有集体智慧和社会决策过程的一个重要方面。然而，虽然我们在了解社会动员过程方面已取得了一些成就，我们离研发出一种可靠、量化的理论还有很长的路要走。换言之，我们已研发出可以预测观点和新闻网

图片由 ALICIA KUBISTA/ANDRIJ BORYS ASSOCIATES 提供

络传播的模型，但我们仍缺乏模型来预测由同一传播活动产生的行为变化。我们争辩道，这些使用和预测的失败并不是由于缺乏数据分析的技能，而是由于对潜在激励结构的关注不足所导致的，这个结构指能为集体决策和行动提供动力的人际关系激励的潜藏网络。

许多大规模社会动员实验显示了激励结构在现实的、对抗性场景中具有重要作用。这些全球范围的实验包括美国国防高级研究项目局 (DARPA) 发起的 DARPA 网络挑战 (DARPA Network Challenge)，要求参赛队伍定位出放在美国大陆随机位置的十只气象气球，最终我们团队获得了胜利。我们使用的是一种递归的激励方案，在 48 小时内雇用了约 200 万名研究员。还有 DARPA 碎纸恢复挑战，在该挑战中我们雇用了逾 3500 名人员，集体协作，共同拼接实际的碎片文档。最近的一次是美国国务院发起的标签挑战 (Tag Challenge)。在这次挑战中我们雇用了志愿者在 12 小时内定位在偏远城市在逃的人员，我们通过使用同一激励方案再次取得了胜利。在每个挑战中，所有的竞争团队都有相同类型的信息（即找到气球、组成碎片、找到目标人员），许多团队成功地制造了病毒营销，扩散至大规模人群并引起了关注。但这些策略的效率差别很大，并与他们的激励设计和参与者动机的匹配度非常相关。即使是在找气球这样的简单任务中，我们发现团队将人们的动机理解为个人收益、慈善、利益互惠或娱乐，获得了不同程度的成功。由竞争力强的团队构建的一些激励结构与个人内在的激励结构相吻合，从而能够激发参与者，制造出行动的网络级联。而另一些团队则无法做到。

我们认为激励网络扮演重要的中间层角色，连接着诸如意识形态和文化的高阶概念，和社会运动在

为什么社交媒体不能是建设性社会变革的更可靠渠道呢?

网络电子平台如 Twitter 和 Facebook 上留下的电子指纹。意识形态和文化塑造了个人在日常生活中想要取得的目标，影响着他们如何将自己与别人的福祉相关联，以及他们如何通过互相帮助达到这些目标。这可以反映在激励网络上：每个个人的收获取决于他人的收获。虽然激励结构受到更抽象的潜在过程的影响，但还是可以通过这些大规模的集体行为实验来量化映射。

突发的社会动员行为无法被维持并转移来创造持续性的社会变革，这种现象正是根源于当今电子社交媒体的设计。今天的社交媒体设计目标是将信息传播和病毒式扩散最大化（通过优化点击和分享），而牺牲受众的参与和共识构建。例如，Onnela 和 Reed-Tsochas[19] 显示，即使在外部信号缺失的情况下，电子社交影响自发地带有不稳定的不全则无的性质。结果变为“一时的狂热”，无休止的开端，竞争，新的狂热为赢取人们注意互相搏杀，灭亡，没有留下持久的影响。[31] 有效的社会动员是信息扩散和行为动员激励的产物，但社交媒体商业的压力过分集中在扩散，而忽视了动员人们行动的激励措施。即使商业角度看，以吸引行动为目的时，社交媒体是极其无效的。例如通过点击广告购买。我们这一代最聪明的大脑可能不再思考如何使人们点击广告（正如 Hammerbacher 在 2013 年出名的言论所提到的那样），[15] 但是他们也只是进阶为思考如何使人们点击“分享”和“赞”。

事件日历

4 月 3-6 日
ISPD’16: 物理设计国际研讨会，
Santa Rosa, CA,
赞助商：ACM/SIG，
联系人：Fung Yu Young,
Email: fyyoung@cse.cuhk.edu.hk

4 月 4–8 日
SAC 2016: 应用计算研讨会
比萨，意大利，
赞助商：ACM/SIG，
联系人：Sascha Ossowski,
Email: sascha.ossowski@urjc.es

4 月 11–14 日
CPS 16 周：2016 信息物理系统周
维也纳，奥地利
联系人：Radu Grosu,
Email: grosu@cs.sunysb.edu

4 月 12–14 日
HSCC’16: 第 19 届混合系统国际会议：计算和控制（CPS 周的一部分），
维也纳，奥地利
联系人：Alessandro Abate,
Email: a.abate@tudelft.nl

4 月 12–14 日
ICCPS‘16:ACM/IEEE 第 7 届网络物理系统国际会议（和 2016CPS 周同期），
维也纳，奥地利
联系人：Ian Mitchell,
Email: mitchell@cs.ubc.ca

4 月 12–14 日
IPSN‘16: 第 14 届传感网络信息处理国际会议（和 2016CPS 周同期），
维也纳，奥地利
联系人：George J. Pappas,
Email: pappasg@seas.upenn.edu

4 月 18–21 日
EuroSys‘16:2016 年第 11 届欧洲计算机系统专业协会会议
英国伦敦
赞助商：ACM/SIG，
联系人：Peter R Pietzuch,
Email: prp@doc.ic.ac.uk

商业社交媒体对病毒式传播的偏见使大多数研究人员和实践者将社会行动的研究聚焦在信息扩散的动力学上，特别是引起病毒式信息传播的条件。但是关于什么样的内容会引起病毒式扩散的可靠事先预测看似并不可及。顶尖的网络科学学者，如 Duncan Watts，[30] Jon Kleinberg，[17] 以及 Matthew Jackson[12] 长期以来一直争辩道：病毒式传播是高度不可预测的，而我们选择性的观察成功活动使我们对其潜在原因表述不当。

此外，尽管有可能通过工程使产品具备“病毒特征”，[2] 相比传播信息本身，病毒传播通常与信息传播背后的激励更相关，特别是对于政治等有争议的领域而言。通过内容创作招徕人群自其诞生以来就是一门手艺，[4] 而考虑到其对当前个性化的社会文化背景的依赖性，在可预见的将来，它仍很可能继续作为一门手艺存在。相比之下，如果我们将努力的方向转移至激励手段的筹划上，那么我们可能会更好地决定内容是否适于激发行为，并创造持久的社会变革。

除了商业社交媒体对于病毒式传播的偏好，出于两点实用主义的考虑，研究可能过于强调病毒式传播。首先，将社会动员等同于病毒式信息传播使这一现象可以通过使用来自流行病学和公共健康的工具进行分析。[9,24] 不过，这种流行病学视角仅适用于拥有建设性社会政治动机的人群，即一个已“开窍”的社会。强调信息病毒式传播的第二个原因是一种我们可以称之为*网络可测性偏向*的现象，即倾向于关注在电子社交网络上易于观察的过程（例如“赞”和“转发”），而忽略重要的潜在过程，如意识、文化及经济刺激因素。社交媒体是一种神奇的新工具，可以使社会科学家实时测量社交信息的传播，然而它却几乎无视其它相关因素，[33] 例如构建过程，[6] 反应，[5] 共识形成，或论证过程，[23,25] 这些对于将内容连接至持续激励至关重要。

经济、社会和政治科学中激励理解方面的研究已取得了很大进展。Hurwicz，Maskin 和 Myerson 因其开发出机制设计，获得了 2007 年的经济学诺贝尔奖。机制设计是一种数学工具箱，用于发现并利用个人参与策略互动的真实倾向。此外，通过程式化的重复合作游戏例如囚犯的困境（the Prisoner’s Dilemma），最后通牒游戏（Ultimatum Games）等，近来的一些实验室试验已能够识别社会结构和动力学如何塑造激励。[21,29] 这些策略场景可能与现实世界中刺激人们行动的因素有很大差距，[27] 但是它们可以作为一种初始“探针”，用实验方法发现动态的激励网络，并为旨在促成行为改变的大规模社交网络实验提供补充。[8,14,26]

信息传播是形成集体信念、观点和态度的关键。但激励扮演着同等重要的角色。让人们接受一种想法是一件事。动员他们投入大量时间，精力和风险来支持一项事业需要更多条件。需要的是新的实验范式，观察工具，它们不仅可以引出传播动力还有在社会动员过程中发生作用的潜在的个人、社会、文化动机的动力学。来自这些实验的结果可以帮助我们构建新一代的社交媒体，超越昙花一现的热潮和病毒式模因，趋向持续变革的共识构建。

个人并不是原子。没有正确的激励结构，一群个体无法变身为一个成熟的问题解决群体，更不用说改变社会。这正是完全开放且平等互联社会的悲剧：当人们在线讨论社会问题时，很难可靠地量化被提及的不同问题的重要性。意识（多少人在意某一件事）以及持续性（他们对这件事的关注度持续多久）均呈重尾分布。[3,7,16,20] 这使得民众，包括研究这一现象的科学家很难构建重要性的清晰阈值，以对大量潜在事件的重要性进行优先级排序。没有有意义的行动阈值，可替代的事件则互相排挤，导致“懒人行动主义”，使得军事或经济力量成为改变的唯一途径。

个人和集体的关注是有限的，社交平台的性能及它们算法推断、操控及获取关注的能力似乎在不断地改进。但是倘若社交媒体没有改进复杂的合作和制度建设，那么到头来将不会取得任何成果。我们需要更加深入地了解如何挖掘网络激励，以及通过信息筛选和共识构建来触发正确的激励。

然而，与信息内容和社交网络结构不同的是，激励的可见性远远低于前者。它们通过个体们的行动得以体现，而且通常一个特定的行动来自多种激励。在我们产生社会动员的一种实际理论之前，我们需要研究出在网络中测量、影响及塑

造激励的新方式，以此解读个人行为。我们在大规模动员挑战中做出的努力只是朝这一方向迈出的一小步。

亚当斯密被很多人认为是“可见行动决定一切”观点的先父：人们在市场中行动，无形的手在不了解私人信息和人们背后动机的情况下产生了高效的结果。但是在他的*道德情操论*中，斯密很清楚地说明对社会现象的真实认知必须包含大量的心理和文化动机。我们将关注点从可观察的病毒式传播过程转移至塑造他们的潜在动机动态上，这要感谢斯密对人性的微妙理解。以及，也许接下来我们能够设计下一代的社交媒体。 C

参考资料

1. Alstott, J. et al. Homophily and the speed of social mobilization:The effect of acquired and ascribed traits. *PLOS ONE 9*, 4 (2014), e95140.
2. Aral, S, and Walker, D. Forget viral marketing—Make the product itself viral.*Harvard Business Review* (2011), 34–35.
3. Bakshy, E. et al. Everyone’s an influencer:Quantifying influence on Twitter.In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*.ACM, 2011.
4. Bartels, R. *The History of Marketing Thought*. Publishing Horizons, Columbus, OH, 1988.
5. Baumer, E.P. et al. Reviewing reflection:On the use of reflection in interactive system design.In *Proceedings of the 2014 Conference on Designing Interactive Systems* (2014), ACM, 93–102.
6. Benford, R.D. and Snow, D.A. Framing processes and social movements:An overview and assessment. *Annual Review of Sociology*, (2000), 611–639.
7. Blumm, N. et al. Dynamics of ranking processes in complex systems.*Physical Review Letters 109*, 12 (2012), 128701.
8. Bond, R.M. et al. A 61-million-person experiment in social influence and political mobilization.*Nature 489*, 7415 (2012), 295–298.
9. Braha, D. Global civil unrest:Contagion, self-organization, and prediction.*PLOS ONE 7*, 10 (2012), e48596.
10. Conover, M.D. et al. The digital evolution of occupy Wall street.*PLOS ONE 8*, 5 (2013), e64679.
11. Conover, M.D. et al. The geospatial characteristics of a social movement communication network.*PLOS ONE 8*, 3 (2013), e55957.
12. Golub, B. and Jackson, M.O. Using selection bias to explain the observed structure of Internet diffusions. In *Proceedings of the National Academy of Sciences 107*, 24 (2010), 10833–10836.
13. González-Bailón, S. et al. The dynamics of protest recruitment through an online network.*Scientific Reports 1* (2011).
14. Gutiérrez-Roig, M. et al. Transition from reciprocal cooperation to persistent behaviour in social dilemmas at the end of adolescence.*Nature Communications 5* (2014).
15. Hammerbacher, J. Charlie Rose and Jeff Hammerbacher talk Data Science in Healthcare; http://www.cloudera.com/content/cloudera/en/resources/library/aboutcloudera/jeff-hammerbacher-charlie-rose.html
16. Karsai, M. et al. Small but slow world:How network topology and burstiness slow down spreading.*Physical Review E 83*, 2 (2011), 025102
17. Liben-Nowell, D. and Kleinberg, J. Tracing information flow on a global scale using Internet chain-letter data. In *Proceedings of the National Academy of Sciences, 105*, 12 (2008), 4633–4638.
18. Naim, M. *The End of Power:From Boardrooms to Battlefields and Churches to States, Why Being In Charge Isn’t What It Used to Be*.Basic Books, 2014.
19. Onnela, J.P. and Reed-Tsochas, F. Spontaneous emergence of social influence in online systems.In *Proceedings of the National Academy of Sciences 107*, 43 (2010), 18375–18380.
20. Papadopoulos, F. et al. Popularity versus similarity in growing networks.*Nature 489*, 7417 (2012), 537–540.
21. Peysakhovich, A. et al. Humans display a ‘cooperative phenotype’ that is domain general and temporally stable.*Nature Communications 5* (2014).
22. Pickard, G. et al. Time-critical social mobilization. *Science 334*, 6055 (2011), 509–512.
23. Rahwan, I. et al. Laying the foundations for a world wide argument web.*Artificial Intelligence 171*, 10 (Oct. 2007), 897–921.
24. Rutherford, A. Limits of social mobilization.In *Proceedings of the National Academy of Sciences 110*, 16 (2013), 6281–6286.
25. Schneider, J. et al. A review of argumentation for the social semantic Web.*Semantic Web 4*, 2 (Feb. 2013), 159–218.
26. Shirado, H. et al. Quality versus quantity of social ties in experimental cooperative networks.*Nature Communications 4* (2013).
27. Stefanovitch, N. et al. Error and attack tolerance of collective problem solving:The DARPA shredder challenge.*EPJ Data Science 3*, 13 (2014).
28. Tang, M. et al. Reflecting on the DARPA red balloon challenge.*Commun.ACM 54*, 4 (Apr. 2011), 78–85.
29. Tsvetkova, M. and Macy, M.W. The social contagion of generosity.*PLOS ONE 9*, 2 (2014), e87275.
30. Watts, D.J. *Everything Is Obvious:How Common Sense Fails Us*.Random House LLC, 2012.
31. Weng, L. et al. Competition among memes in a world with limited attention.*Scientific Reports 2* (2012).
32. World Economic Forum (WEF).*Outlook on the Global Agenda, 2014*; http://reports.weforum.org/outlook-global-agenda-2015/
33. Zuckerman, E. The first Twitter revolution?*Foreign Policy 14* (2011).

Manuel Cebrian (manuel.cebrian@data61.csiro.au) 是澳大利亚联邦科学和行业研究组织（CSIRO）Data61 组的研究组领导。

Iyad Rahwan (irahwan@mit.edu) 是麻省理工学院媒体实验室媒体艺术与科学的副教授。

Alex “Sandy” Pentland (pentand@mit.edu) 主管麻省理工连接科学以及人类动态实验室，此前曾协助创建并管理 MIT 媒体实验室以及印度的亚洲媒体实验室。

译文责任编辑：唐杰

DOI:10.1145/2844546

文章编写主导者 acmqueue
queue.acm.org

失友怒僚之道

THOMAS A. LIMONCELLI

系统管理员的自我贬值之道

问：亲爱的汤姆：怎样才能降低工作的价值？最近我感觉，每个人都欣赏我。但实际上，我拿的工资过高，工作却不紧。你能帮我降低我在工作中的价值吗？

答：亲爱的读者：绝对没问题！我了解把高薪拖回家的那种痛苦。如果一直有人拍自己的背，那也确实让人分心。哎唷！不仅如此，受大家追捧之后，情况堪比与著名乐手和电影明星的会面（只要问问 Taylor Swift 或 Leonardo DiCaprio 这样的大明星）。想玩超级棒的视频游戏的时候，谁会想被那么骚扰？

下面列举了一些久经考验的，人人应知的技巧。

每周工作 40 小时以上

这是系统管理员采取的最为简单，可能也是最为常见的技巧。每周工作 80 小时，就可以轻松地把每小时的价值降低一半。首先每周一或两个晚上加班很晚。然后周末再加班。很快，你就可以顺利地迈向满满的 80 个小时。

从公司的角度看，超出薪酬之外的工时是免费劳动。这降低了你的平均每小时工资。如果你乐意利用闲暇时光，为什么还要招更多的系统管理员呢？薪酬过高的 CEO 会说这么说，“我的薪酬与我的责任相称。”注意，他们并没有把薪酬与工作时长关联起来。你也不该这样。

嘲笑不喜欢的事物。

如果你不喜欢微软 (Microsoft)，把它叫成微扁（Microsquish）。别说开源 (open source)；说“开疮”(open sores)。听起来越像 12 岁的小孩越好。

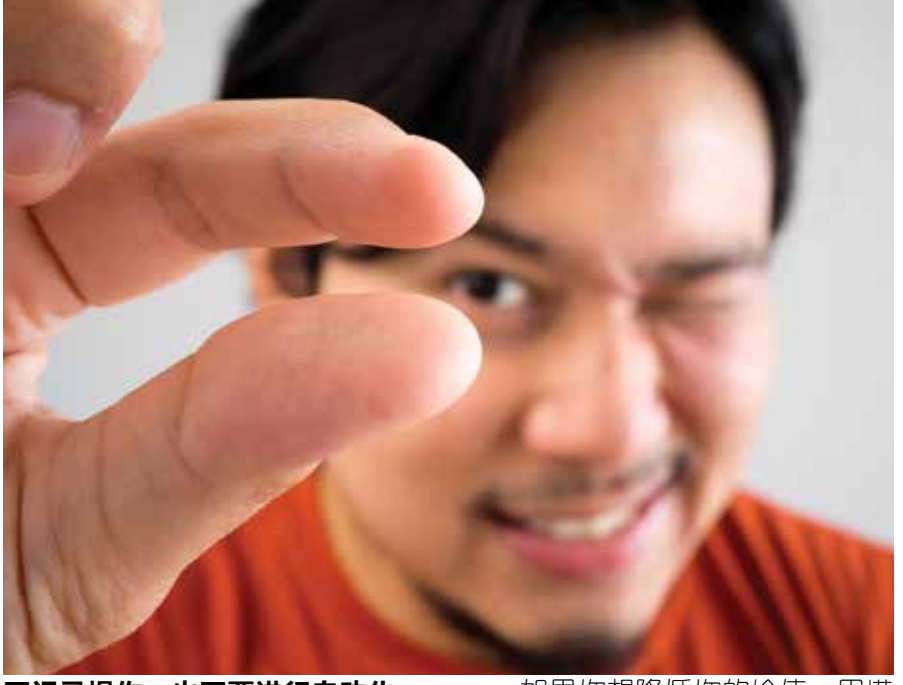
如果你想降低你的价值，就把职业精神抛到九霄云外吧。用最幼稚的方式表现你的不敬。我知道有这样一位工程师，他把不喜欢的操作系统写成“Windoze”的工程师。即使在给使用该系统的同事和客户发邮件时，他也这么写。就该这么干！

打断别人。

“别尊敬我”的最佳宣言莫过于对他人不敬。不能让人把事情说完很重要，这就是原因。一旦你听懂了大概，马上就要开始嚷出自己的答复。这说明你不关心别人正在说的东西，他们也会回以不敬。

别让人把话说完，这能让他们感到你非常聪明。你的大脑超强，已经有了超能力（ESP）。在他们告诉你问题之前，回答他们的问题。用这种方式证明你的超能力。

尊敬是相互的。就像回旋镖。你对人不敬，他人也会用相同的方式回敬你。



图片由 RACHATA TEYPARSIT 提供

不记录操作，也不要进行自动化。
这一点存在一些争议。有些人认为，拒绝记录任何东西后，可以增加他们的价值。这样公司就无法解雇他。真相是，经理们高度重视那些时刻更新操作诀窍和其他文档的雇员。

与此类似，某些系统管理员担心，如果自动化脚本写得太多，那么他们就会失业。真相是，如果你通过自动化消除了某项任务，那么还会有更多的任务等着你去自动化。做到这一点的人成了工作倍增器：一个人支持多人做很多工作。这是非常有价值的。

因此，如果你希望降低自己的价值，不要写文档，也不要搞自动化。让每个人确信，稍后你会把某个东西记录下来：抵制那种在执行任务时更新维基条目的冲动。当有人要求你自动化某个任务时，就看着那个人的眼睛，叹口气，然后说，“我太忙了，没办法通过自动化来节省时间”。

关注技术，而不是商业效益。
你想买的新服务器超酷，如果业务部门不理解这点，叫得再凶一点。

某些人可能不同意这一做法。他们认为，每次购买的技术应该从它能给业务带来多少（涉及金钱或时间的）收益的方面进行论证——例如，服务器应整合所有的销售信息，能让销售人员在需要某些信息的时候找到那些信息。太无聊了。它配有 20T 的 SSD 加速存储、Intel 5655 CPU 和三倍冗余的电源，这样解释要有趣得多。

如果你想降低你的价值，用模糊的方式描述项目的商业价值。使用最详细的技术术语，让人们猜测业务原因。像业务要为技术服务，而不是技术为业务服务那样行事。

只雇佣与你相似的人。
多样化是指重视不同背景的人们能够为工作带来不同的技能这一事实。研究发现，团队中加入一个背景不同的人之后，会提高团队的生产率。

生产率？听起来好像与降低自我价值相悖。为了真正地降低自己的价值，需要确保团队中的每个人想法一致，拥有相同的技能和类似的背景，并且都犯相同的错误。

正如我之前写的那样，尊敬是相互的。如果你想降低自己的价值，那就不要重视差异。

成为怪人。
成为公司里的“怪人”。如果你的同事并不理解你晦涩地说了 *Dune*（*科幻作品沙丘*），*Animaniacs*（*电影狂欢三宝*）和 *LOTR*（*电影指环王*）中的东西，也没什么关系。在我们前往末日火山（Mount Doom）之前，香料必须流淌，这样我们能做出腊肠，放进宽松长裤（Animaniacs 中的台词，意指找出问题的原因）中。假装你没有注意到大家疑惑的表情。每个人肯定都读过 *Dune*（*沙丘*）。不要解释你的文化用语，也不要仅仅因为没人理解就不用这些词语。

很多人可能认为不同于自己的人很奇怪，但这是两种不同的概念。多样化是指重视差异。成为怪人则是指忽视他人的反应。多样化要求投身于指导他人和向他人学习。成为怪人正好相反。

每个人都应该可以自由地挥舞着怪人的大旗。如果你想降低自己的价值，永远都别解释。

把服务器机房（server room）称作郡 (shire) 不要说“生日快乐”，说“孵化日快乐”。每次碰到某样东西有红色按钮时，问问它是不是像糖果。

让人找到你很难。
如果你不存在，那就不可能有价值。如果很难找到你，或者需要你的时候你不在，那么你就没有为任何人提供价值。

让工作时间变得与众不同。中午之前别到岗——除非公司文化是中午到岗；如果那样，就早点到。

无论用哪种方式，确保你的工作时间与需要你的其他人的工作时间错开一点。

总结寄语
如果我们同心协力，那么对于作为一个社区的所有系统管理员而言，大家能够保证人们会长期低估系统管理员的作用。 C

queue.acm.org 上的相关文章

Innovation and Inclusion
Telle Whitney, Elizabeth Ames
http://dx.doi.org/10.1145/2676861

Are You Invisible?
Jack Rosenberger
http://cacm.acm.org/blogs/blog-cacm/94307-are-you-invisible/fulltext

Automation Should Be Like Iron Man, Not Ultron
Thomas A. Limoncelli
http://queue.acm.org/detail.cfm?id=2841313

Thomas A. Limoncelli 是纽约市 Stack Overflow 公司的网站可靠性工程师。他的著作包括 *The Complete April Fools' Day RFC* (www.rfchumor.com)、*The Practice of Cloud Administration* (the-cloud-book.com) 以及 *Time Management for System Administrators* (O'Reilly)。他在 EverythingSysadmin.com 发表博客。

译文责任编辑：陈文光

DOI:10.1145/2818990

利用来自多个生物特征的融合信息能够改善移动设备的身份认证机制。

作者：MIKHAIL I. GOFMAN 和 SINJINI MITRA

利用多模态生物特征识别技术来提高移动设备的安全性

每年都有上百万手机设备被盗，这些设备中存储的信用卡号码、密码以及其他安全和个人信息都会随之失窃。多年来，犯罪分子已经学会了破解密码和伪造生物特征，几乎可以攻破各种类型的用户身份认证机制，让这些机制形同虚设，进而获取设备数据。显然，设计更安全有效的移动设备身份认证机制已是迫在眉睫。

在本文中，我们展示了多模态生物特征识别这一蕴藏巨大潜力和前景的技术，它是一种基于人脸和声音等多种生理和行为特征的身份认证方式，可以保护消费者的移动设备抵御非法入侵。尽管多模态生物特征识别技术已经在国土安全、军事以及法律执行领域得到应用，15，18 但是它们尚未广泛地整合到消费品移动设备中。这是因为这种技术实现起来并不容易，而且消费者可能会担心这会带来使用上的不便。

我们还展示了，多模态生物特征识别技术能以用户友好的方式整合到移动设备中，并大大提高移动设备的安全性。2015 年，我们在加州州

» 重要见解

- **多模态生物特征识别技术，即根据用户的多个生理和行为特征识别用户身份，将成为打造更加安全而可靠的基于生物特征的移动设备身份认证机制的新解决方案。**
- **本文所讨论的基于人脸和声音的生物特征识别系统成功地应用在 Samsung Galaxy S5 智能手机上，并且可以在非受控条件下实现更高的身份认证准确性，即使在处理昏暗的人脸图像和嘈杂的声音样本时也不例外，性能优于单模态人脸和声音识别系统。**
- **移动设备的多模态生物特征识别技术可以让普通消费者轻而易举的使用。**

照片由 ANDRIJ BORYS ASSOCIATES/SHUTTERSTOCK 提供

立大学富勒顿分校实现了一种多模态生物特征识别系统，并将其命名为为“Proteus”。该系统基于 Samsung Galaxy S5 智能手机的人脸和声音识别技术，并且整合了全新的多模态生物特征身份认证算法（针对消费品移动设备进行了优化）以及一个可以让用户轻松采集多模态生物特征的界面。我们的实验证实了以下结果：与仅基于人脸或声音的系统相比，Proteus 的身份认证准确性显著提高。下一步就是将包括指纹和虹膜扫描在内的其他生物特征识别技术集成到这一系统。希望我们的实验能够抛砖引玉，激励研究人员和移动设备制造商推进这一领域的创新。

生物特征识别技术

基于生物特征识别技术的身份认证机制根据人类的生理和行为特征（例如人脸和声音）来建立身份信息，让用户不必再创建或者记住他们的安全口令。同时，它也让攻击者伪造人类特征变得可望而不可即。[21] 基于生物特征识别技术的身份认证机制具有的这些优势，将继续推进它在智能手机和平板电脑上的应用。

尽管这种身份认证机制能否在移动设备上取得成功仍有待商榷，但一些严重的问题仍然不容忽视，例如，能够攻破 iPhone TouchID 和 Samsung Galaxy S5 的指纹识别系统的技术已经出现。[2,26] 此外，现代生物特征识别系统缺乏可靠性，经常无法识别授权用户，这些问题仍然会受到消费者的诟病。[4] 为了解多模态生物特征识别技术如何帮助解决这些问题，我们先来调查造成这些问题的根本原因。

移动设备的世界

移动设备的生物特征身份认证机制面临的一个主要难题就是样本质量无法得到保障。无论是人脸照相、录音还是指纹扫描，高质量的样本对确保身份认证的准确性至关重要。例如，低分辨率的人脸照片或嘈杂的声音录音会导致生物特征识

别算法将冒充者识别为合法用户，即出现“误接受”的问题。同样地，低质量的样本也可能导致算法将合法用户识别为冒充者，即出现“误拒绝”的问题。在移动设备上难以捕获高质量的样本，主要原因有两个：移动用户在不同环境条件下捕获生物特征样本，环境条件可能会受到多方面因素的影响，包括光照不足、姿势不同、摄像头角度变化以及存在背景噪音等。消费品移动设备上的生物特征识别传感器通常会以牺牲样本质量为代价，来提高设备的便携性和降低成本。例如，Apple iPhone 的 TouchID 指纹扫描器受尺寸限制，无法捕获整个手指的指纹，因此更容易被破解。[4]

另一个难题在于训练生物特征识别系统，使之能够识别设备用户。训练过程的主要内容就是从用户提供的生物特征样本提取差异特征。增加训练样本的数量和多样性可以提高身份认证的准确性。但实际上，大多数消费者为了图方便，只用少且单一的样本来训练身份认证系统。多模态生物特征识别技术是解决这些难题的关键。

多模态生物特征识别技术的前景

多模态生物特征识别系统采用多种完全独立的识别信息（如人脸和声音），而目前的移动设备所采用的单模态系统仅根据单个生物特征来认证用户身份，所以前者可以克服后者存在的安全性和可靠性问题[13,18]。此外，在现有的移动设备上部署多模态生物特征识别技术是切实可行的，大多移动设备都带有人脸、声音和指纹识别功能。如何以一种可靠且用户友好的方式将这些技术结合起来，是当前需要解决的问题。在消费品移动设备上应用多模态生物特征识别技术可以带来很多好处。

*提高移动设备的安全性。*攻击者要想攻破单模态生物特征识别系统，只需伪造该系统采用的单个生物特征模态即可。如果系统采用多个模态建立身份信息，攻击者就难以同时伪造多个相互独立的人类特征，让破解变得难上加难。[21]

*更加可靠的移动设备身份认证机制。*在使用多个生物特征时，不同生物特征模态之间可以实现多样性和质量的相互补足。例如，Proteus 会评估人脸图像和录音的质量；在认证身份时，质量较高的特征将起到更高的决策作用。

同理，多模态生物特征识别技术可以简化设备训练流程。用户如果提供多个模态的样本，那么每个模态的样本数量就较少；不像在使用单模态系统时，必须提供来自单个模态的大量训练样本。这些身份识别信息可以结合使用，为训练可靠的身份认证机制提供充足的数据。

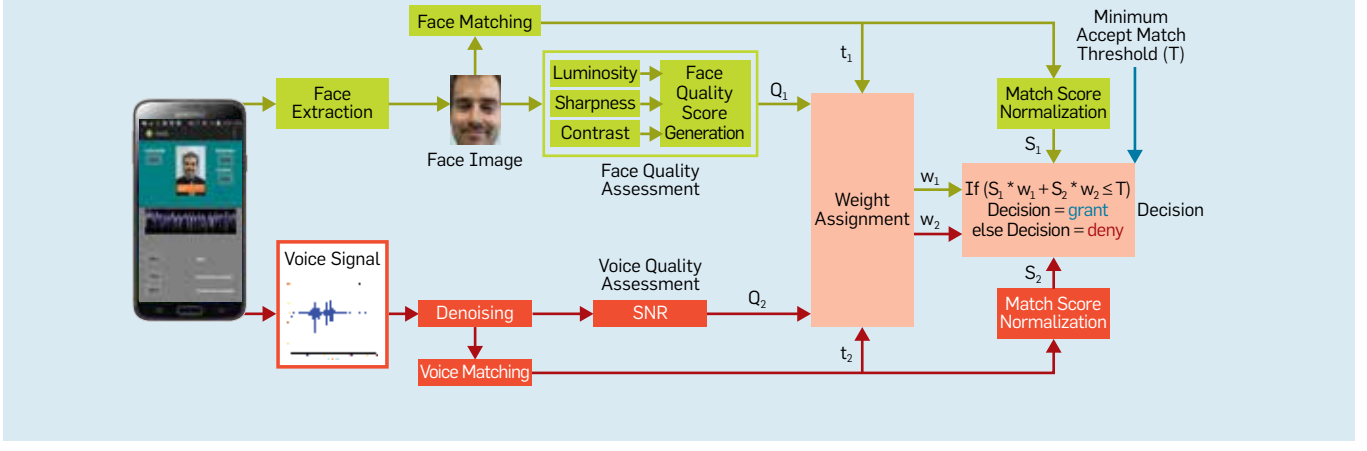
*市场商机遍布。*尽管生物特征身份认证机制已经在消费品移动设备上广泛使用，但多模态生物特征识别技术在相应市场的普及度仍很有限。[1,15] 这可能是因为用户担心采集多种生物特征会很麻烦。多模态系统的设计和实现难度也比单模态系统大。

但正如我们所说的那样，这些问题都是可以解决的。为了将生物特征识别传感器（如摄像头和指纹读取器）集成到产品中，Apple 和 Samsung 等公司已经投入大量资金。因此，它们无需大幅增加产品成本就能部署多模态生物特征识别技术，而产品会因安全性和可靠性的提升而变得获取更多的利润，让厂商更加盈利。在以下的章节中，我们将讨论如何来实现可带来盈利的安全性。

融合人脸和声音生物特征

为展示多模态生物特征可为消费品移动设备带来哪些好处，我们实现了基于人脸和声音生物特征的 Proteus 系统，我们之所以选择这两种模态，是因为大多数移动设备都带有摄像头和话筒，可以捕获这些信息。在本章节，我们将先介绍人脸

图 1. 示意图演示了 Proteus 基于质量的分数层面融合方案。



和声音识别技术，然后再探讨我们采用哪些技术来融合这些信息。

人脸和声音识别技术。我们在 Proteus 中使用的是名为 FisherFaces[3] 的人脸识别技术，它可以较好地处理在不同环境下获取的图片；这正好适用于通过移动设备获取的人脸图片。FisherFaces 使用人脸图片中的像素强度作为识别特征。今后，我们计划研究其他人脸识别技术，包括 Gabor 小波[6] 和“方向梯度直方图”(HOG)。[5]

我们采用两种方法识别声音：将基于梅尔倒频谱系数 (Mel-Frequency Cepstral Coefficients) 的隐马尔可夫模型 (Hidden Markov Models) 用作声音特征，[10] 也就是我们的分数级融合方案的基础；将线性判别分析 (Linear Discriminant Analysis)[14] 用作我们的特征级融合方案的基础。这两种方法都可以识别用户声音，不受用户所说词组的影响。

评估人脸和声音样本的质量。对基于生物特征的身份认证系统来说，评估生物特征样本的质量是确保此类系统的准确性的关键。如前所述，这条结论尤其适用于移动设备。因此，Proteus 会根据亮度、锐度和对比度来评估面部图像的质量，而声音录音的质量则根据信噪比 (SNR) 来确定。这些经典的质量指标在一些生物特征研究文献中都有详述。[1,17,24] 今后，我们还希望研究一些其他的有价值的指标。

Proteus 根据成分像素的强度来计算人脸图像的平均亮度、锐利和对比度，具体的方法可以参见 Nasrolli 和 Moeslund 的著述[17]。然后，Proteus 使用最小 - 最大归一化方法，将每个质量指标归一化到 [0, 1] 的范围内，最后通过计算平均值来获得每个人脸图像的质量评分。这里有一个有趣的问题，就是确定每个质量指标对最终人脸图片质量评分的影响。例如，如果人脸图像太暗了，那么亮度较差就会成为影响最大的因素，因为光照不足会严重妨碍图像的识别。同样，如果明亮的图像因为动态模糊而发生失真，那么锐度就会成为影响最大的因素。

为了使算法适应设备的有限资源，Proteus 需要能够缩减人脸图像的尺寸，防止算法耗尽设备的可用内存。

SNR 是指声音信号水平与背景噪声信号水平之间的比率。为获得声音质量评分，Proteus 采用了 Vondrasek 和 Pollak[25] 所提及的概率处理方法来评估声音和噪声信号，然后使用最小 - 最大归一化方法，将 SNR 值归一化到 [0, 1] 的范围内。

多模态生物特征融合。在多模态生物特征识别系统中，来自不同模态的信息可以在以下层面结合起来，或者融合起来：[21]

特征。来自多个传感器和 / 或来源的数据或特征集都可以融合；

匹配分数。适用于不同生物特征模态的多个特征匹配算法生成的匹配分数都可以结合起来，以及

决策。由多种匹配算法提供的判定结果可以整合为单个判定结果，可以使用的技术包括多数表决法等。

生物特征研究人员认为，在处理流程的早期（例如特征层面）整合信息比在后期（比如分数层面）整合信息更加高效。[20]

移动设备多模态生物特征识别技术的框架

Proteus 在分数或特征层面融合人脸和声音生物特征。由于决策层的融合不会起到太大作用，[21] 所以我们在开发 Proteus 时忽略了这点。

用户使用手机摄像头拍摄面部视频，同时说出特定词组，这时 Proteus 就会利用这些信息来训练和测试。该系统使用 Viola-Jones 算法[24] 来检测每个视频中的人脸，同时提取其中的声道。系统会对所有音频帧降噪，滤除超出人类音域 (85Hz–255Hz) 的频率，并且剔除

不存在声音活动的帧。所得到的结果在我们的融合方案中作为输入。

分数层面的融合方案。图 1 概括了我们在分数层面的融合方法，融合的对象是人脸和声音生物特征。每个模态的匹配分数对用户身份认证的最终判定结果的影响程度取决于相应的样本质量。Proteus 的工作原理在以下段落中介绍。

用 t_1 和 t_2 分别代表来自用户设备的人脸和声音训练样本的质量评分。然后，从测试视频序列中，Proteus 分别计算这两个生物特征的质量得分 Q_1 和 Q_2。这四个参数随后被发送至系统的权值分配模型，由其分别计算人脸和声音模态的权值 w_1 和 w_2。每个 w_i 的计算方式为 $w_i = \frac{v_t}{p_2 + p_2}$，其中 p_1 和 p_2 分别是 Q_1 到 t_1 和 Q_2 到 t_2 的百分比近似值。该系统要求用户使用优质样本作为主要训练素材，我们随后会予以讨论。因此，测试样本质量与训练样本质量极为接近，则代表测试图像质量高。模态的样本质量越高，被分配的权值就越大，这样就能确保系统在认证身份的最后流程充分考虑质量因素。

随后，系统会计算和归一化匹配分数 S_1 和 S_2，这两个数字是由相应的人脸和声音识别算法采用 Z-score 归一化方法针对测试图像计算出来的。我们之所以选择这一方法，是因为它是一种广泛应用的归一化方法，不但易于实现，而且非常高效。[11] 不过，我们希望在今后测试一些更可靠的方法（例如 Tanh 函数和 Sigmoid 函数）。随后，系统使用加权求和公式计算融合方案的匹配总分数：$M = S_1w_1 + S_2w_2$。如果 $M \geq T$（T 是预设的阈值），系统就会认为用户身份是真实的，否则就会将用户认定为冒充者。

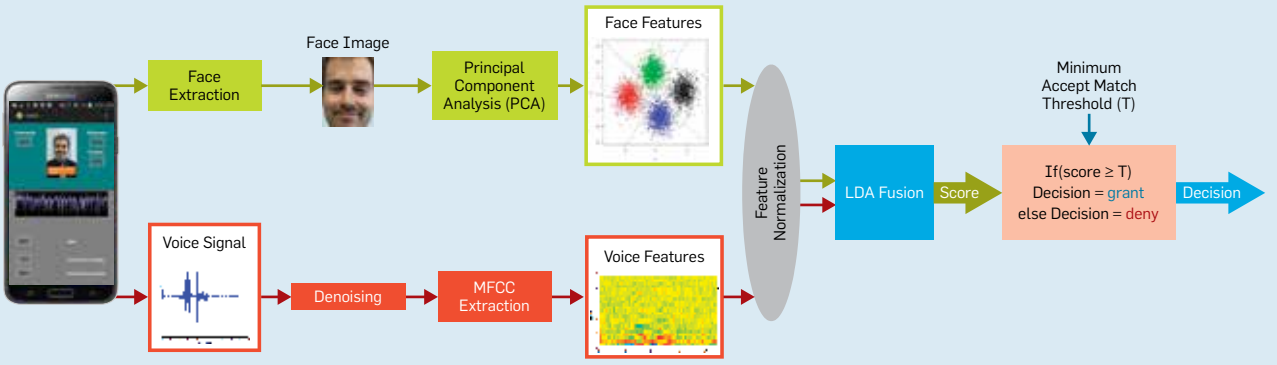
讨论。当 $t_1 = Q_1$ 和 $t_2 = Q_2$ 时，该方案的有效性预计会达到最高。不过，系统这时必须谨慎，确保这两种模态在融合过程中都具有代表性。例如，如果 Q_2 与 t_2 差别很大，而 Q_1 与 t_1 十分接近，那么身份认证流程就会以人脸模态为主，整个过程因此削弱为基于人脸生物特征的单一模态方案。因此，必须为每个质量分规定基准，确保在达不到每个基准分数时，基于融合的身份认证流程不会授予用户访问权限。如果没有此类基准，那么整个身份认证流程都可能受到潜在欺诈行为的威胁，包括蓄意修改某个生物特征模态质量评分的行为。因此，系统必须确保每个模态的权值不会低于特定阈值，否则多模态方案就无法发挥作用。

在移动设备上存储和处理生物特征数据，而不是将这些任务交给远程服务器处理，这样就避免了在有潜在威胁的网络中，如何安全传递生物特征数据和认证结果的问题。

2014 年，IBM 的研究人员提出了一种适用于 iPhone 和 iPad 的基于人脸、声音和签名生物特征的分数层面融合方案。[1] 他们在实践中仅考虑录音的质量，忽略了人脸图像的质量；而我们的方法则完全不同，会同时考虑这两种模态的质量。此外，由于他们的目标是设计远程服务器的安全登录方式，所以大部分的计算任务都由目标服务器承担；而 Proteus 则自己在移动设备上直接执行所有计算工作。为了使算法适应设备的有限资源，Proteus 需要能够缩减人脸图像的尺寸，防止算法耗尽设备的可用内存。最后，Aronowitz 等人 [1] 使用的多个人脸特征（例如 HOG 和 LBP）尽管无疑比 FisherFaces 更加可靠，但它在移动设备上运行速度极其缓慢以至于无法接受，我们希望以后有机会在研究中使用多个人脸特征。

特征层面的融合方案。大多数多模态特征层面的融合方案都假设，要融合的模态是相互兼容的（参见 Kisku 等人 [12] 以及 Ross 和 Govindarajan[20] 的著述），也就是说这些模态的特征能以类似方式（例如基于距离）计算。在特征层面融合人脸和声音模态充满挑战性，因为这两种生物特征不兼容：人脸特征来源于像素强度，而声音特征来源

图 2：基于线性判别分析的特征层面融合。



于梅尔倒频谱系数。当融合特征向量变得过大时，就会发生维度灾难，这是特征层面融合的另一个难题。我们采用 LDA 方法解决了这两项难题。此外，根据我们的实验结果，我们观察到，LDA 与神经网络和 HMM 相比，所需的训练数据更少。

这一流程如下所示（参见图 2）：

第 1 阶段（人脸特征提取）。 Proteus 算法将主成分分析方法 (PCA) 应用到人脸特征集，执行特征选择任务；

第 2 阶段（声音特征提取）。 Proteus 从每个经过预处理的的音频帧中提取一组 MFCC，然后将它们以矩阵形式表示出来，一行对应一个帧，一列对应一个 MFCC 指数。为减少 MFCC 矩阵的维度，Proteus 使用矩阵的列均值作为声音特征向量；

第 3 阶段（融合人脸和声音特征）。 由于该算法使用不同单位衡量人脸和声音特征，所以只能采用 Z-score 归一化方法单独对这些特征进行标准化，与在分数层面融合采用的方式相同。随后，该算法将这些归一化特征连接起来，形成一个大特征向量。如果存在人脸特征 N 个和声音特征 M 个，那么连接或融合集的总特征为 $N + M$ 个。接下来，该算法使用 LDA 从融合特征集执行特征选择。这样就可以从合并的集中移除不相关的特征，有助于避免维度灾难的发生。

第 4 阶段（身份认证）。 该算法使用欧几里得距离，来确定训练数据中的融合特征集与每个测试样本之间的相似度。如果距离值小于等于预设阈值，则判定测试对象为合法用户。否则，就会判定测试对象为冒充者。

实现

我们在随机选取的 Samsung Galaxy S5 智能手机上实现了基于质量的分数层面和特征层面的融合方法。用户友好性和执行速度是我们的指导原则。

用户界面。 在设计用户界面时，我们的首要任务就是确保用户可以同时无缝地捕获人脸和声音生物特征。所以，我们的解决方案会要求用户录制一段简短的视频，既要拍摄下整个面部，同时还要说出一个简单的词组。我们的图形用户界面 (GUI) 的原型（参见图 3）可以向用户实时反馈人脸和声音质量指标，指导用户捕获质量最佳的样本。例如，如果视频的明暗度与训练数据库中图像的平均亮度相差过大，则用户可能会收到以下提示，建议：增加光照。除了用户友好之外，拍摄视频还有利于集成其他的安全认证功能（例如：认证对象是否是活着的 [7]），并且可以将唇动与语言相互关联起来。[8]

为保持较快的身份认证速度，Proteus 的人脸和声音特征提取算法在不同处理器内核上并行执行；Galaxy S5 拥有 4 个内核。Proteus 使用了类似的并行编程技术，以确保 GUI 的响应能力。

生物特征数据的安全性。 在移动设备上存储生物特征数据（Pro-

图 3.Proteus 的用户交互界面。

teus 存储多种生物特征数据），最大的风险就是攻击者可能会窃取并使用这些数据来冒充合法用户。因此，Proteus 必须安全地存储和处理生物特征数据。

当前的方法是仅将 MFCC 和 PCA 系数（而不是原始生物特征数据）存储在设备的永久内存中，因此要从中获取有用的生物特征数据并不容易。[16] Proteus 通过使用可撤销的生物特征模板 [19]，并且在可信的运行环境中（通常是与设备中其他软硬件完全隔离的防篡改硬件）对生物特征数据进行加密、存储和处理，来大幅度提高它的安全性。Galaxy S5 使用这种方法来保护指纹数据。[22]

在移动设备上存储和处理生物特征数据，而不是将这些任务交给远程服务器处理，这样就避免了在有潜在威胁的网络中，如何安全传递生物特征数据和认证结果的问题。此外，这种方法让生物特征数据免于传输到远程系统进行存储，因此减轻了消费者对于数据安全性、保密性和可能被误用的顾虑。

性能评估

我们对比了 Proteus 与基于人脸和声音特征的单模态系统的识别准确性。我们在测量准确性时，使用的是标准的等错误率 (EER) 指标，即误接受率 (FAR) 与误拒绝率 (FRR) 相等时的值。我们必须建立安全存储和处理生物特征数据的机制。

数据库。为开展实验，我们创建了一个名为“CSUF-SG5”的自用多模态数据库，其中包含使用三星 Galaxy S 5 智能手机采集的人脸和声音样本。这些样本来自于加州州立大学富尔顿分校的学生、职工，还有部分校外人员（这就是这个数据库名字的来源）。为了在样本中包含各种类型、各种程度的变形和失真，我们在不同的现实环境中取样。考虑到市面上还没有这种多样化的多模态生物特征数据库，我们计划将自己的数据库公开。目前，数据库中已经包含了 54 个不同性别和种族的用户的视频录像，他们使用手机摄像头拍摄面部，同时说出特定词组。

这些视频中的人脸反映出以下变化类型：

四种表情。平常、高兴、悲伤、愤怒以及恐惧。

三种姿势。正面和侧面（左侧及右侧）；以及

两种亮度条件。均匀光照和部分阴影。

声音样本反映出不同等级的背景噪音，从汽车的轰鸣声到音乐，再到人们的聊天声，外加声音自身的失真（如刺耳声）。我们使用 20 个不同的常用词组，包括“玫瑰是红色的”、“足球”和“13”等。

结果。在我们的实验中，我们使用来自数据库中的一半的数据（在 54 个对象中选取 27 个）来训练 Proteus 人脸、声音和融合算法，同时我们考虑所有实验对象都接受测试。大多数训练视频是在光照充足、背景噪音低的受控条件下拍摄的，而且摄像头正对着实验对象的面部。对于这些实验对象，我们还添加了一些来自质量不佳的视频的人脸和声音样本，以便模拟普通消费者可能会提供的训练样本的有限变化，提高算法在类似条件下正确识别用户身份的几率。总体来说，我们将每个实验对象的 3 个人脸帧和 5 个声音录音（从视频提取）用作训练样本。我们的测试方法是：从数据库中的 54 个实验对象随机选出 1 个实验对象，然后随机选择该实验对象的 1 个人脸和声音样本。总体来说，我们的实验对象创建和使用了 480 个训练和测试集的组合，我们计算出平均 EER 和测试时间。我们采用这种统计交叉验证方法，根据现有数据库中的 54 个潜在实验对象来评估和验证我们所提出方法的有效性。

基于质量的分数层面融合。表 1 列出了单模态和多模态方案的平均 EER 和测试时间。我们认为 HMM 声音识别算法的 EER 较高是因为很多样本中都存在复杂的噪声信号，例如车辆行驶噪声、人们聊天的声音以及音乐等，它们都难以检测和清除。我们的基于质量的分数层面融合方案检测到较低的 SNR 水平，并且通过增加质量明显较好的人脸图像的权值来补偿。通过增加人脸图像的权值，在最终判定用户是否合法时，人脸生物特征就会比声音生物特征起到更大的决策作用。

在相反的情况下，如表 1 所示，当声音样本的质量相对较高时，单模态声音方案和分数层面融合方案的 EER 分别是 21.25% 和 20.83%。

表 1. 分数层面融合的 EER 结果。

模态	EER	测试时间（秒）
人脸	27.17%	0.065
声音	41.44%	0.045
分数层面的融合	25.70%	0.108

表 2. 特征层面融合的 EER 结果。

模态	EER	测试时间（秒）
人脸	4.29%	0.13
声音	34.72%	1.42
特征层面的融合	2.14%	1.57

这些结果令人鼓舞，因为这表明不同模态的质量可以随着用户所处的环境不同而变化。这还表明，Proteus 可以通过合理地调整质量权值来适应不同条件。经进一步优化后（如采用更可靠的归一化方法），多模态方法的精确度可以进一步提升。

特征层面的融合。表 2 概括了特征层面融合方案的效果，表明特征层面融合方案与单模态方案相比，前者的身份认证准确性明显更高。

我们的实验清楚地反映了，多模态生物特征识别技术可以提高当前移动设备上的基于单模态生物特征身份认证机制的准确性；此外，根据系统识别合法用户的速度，Proteus 方法可伸缩以适应消费品移动设备。这是人类首次尝试在现代消费品移动设备上实现两种类型的融合方案，并解决用户友好性的实际问题。但相关研究仅处于起步阶段。我们正在努力改善这两种融合方案的效果和效率，未来的研究将充满无限可能。

结论

消费品移动设备的生物特征身份认证机制，下一步将向多模态生物特征识别技术发展。如何让多模态生物特征技术适用于主流消费品移动设备，这仍然是一项挑战，但是人们已经开始尝试向这些设备添加多模态生物特征识别技术。我们的研究朝着这个方向迈出了第一步。

如果您可以使用人脸、声音、指纹、耳朵、虹膜和视网膜这些特征的组合来解锁移动设备，而且移动设备的身份验证机制可以像 iPhone 的 TouchID 指纹系统一样只需一步就能识别这些生物特征，那么您会有何感想？这用户友好的方式利用基于生物特征样本质量的可靠的底层融合逻辑，可以让设备尽可能正确地识别其用户。手指被弄脏、环境昏暗或嘈杂、生物特征传感器损坏，这些都不会妨碍身份认证系统；如果一个生物特征识别功能发生了故障，可以由其他功能替代。黑客必须获取多种模态才能解锁设备，因为这些生物特征模态仅被设备合法所有者持有。设备还使用可撤销生物特征模板、强加密算法以及可信运行环境来安全地存储和处理所有生物特征数据。

Proteus 的多模态生物特征方案利用移动设备硬件的现有功能（如视频录制），但移动设备的硬件和软件无法处理更加复杂的生物特征组合；例如，主流的消费品移动设备缺少能以用户友好方式来获取虹膜和视网膜生物特征的传感器。因此，我们会努力设计和开发包含能够支持这类组合的高效、用户友好和廉价软件和硬件的设备。我们计划向当前的融合方案中加入新的生物特征，开发新的、更可靠的融合方案，并且设计能够无缝同时捕获多种生物特征的用户界面。将用户友好的界面与可靠的多模态融合算法相结合，这标志着消费品移动设备身份认证机制迎来了新时代。 C

参考资料

1. Aronowitz, H., Min L., Toledo-Ronen, O., Harary, S., Geva, A., Ben-David, S., Rendel, A., Hoory, R., Ratha, N., Pankanti, S., and Nahamoo, D. Multimodal biometrics for mobile authentication.In *Proceedings of the 2014 IEEE International Joint Conference on Biometrics* (Clearwater, FL, Sept. 29–Oct. 2).IEEE Computer Society Press, 2014, 1–8.
2. Avila, C.S., Casanova, J.G., Ballesteros, F., Garcia, L.R.T., Gomez, M.F.A., and Sierra, D.S. *State of the Art of Mobile Biometrics, Liveness and Non-Coercion Detection*.Personalized Centralized Authentication System Project, Jan. 31, 2014; https://www.pcas-project.eu/images/Deliverables/PCAS-D3.1.pdf
3. Belhumeur, P.N., Hespanha, J.P., and Kriegman, D. Eigenfaces vector vs. FisherFaces:Recognition using class-specific linear projection.*Pattern Analysis and Machine Intelligence, IEEE Transactions on Pattern Analysis and Machine Intelligence 19*, 7 (July 1997), 711–720.
4. Bonnington, C. The trouble with Apple's Touch ID fingerprint reader.*Wired* (Dec. 3, 2013); http://www.wired.com/gadgetlab/2013/12/touch-id-issues-and-fixes/
5. Dalal, N. and Triggs, B. Histograms of oriented gradients for human detection.In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (San Diego, CA, June 20–25).IEEE Computer Society Press, 2005, 886–893.
6. Daugman, J.G. Two-dimensional spectral analysis of cortical receptive field profiles.*Vision Research 20*, 10 (Dec. 1980), 847–856.
7. Devine, R. Face Unlock in Jelly Bean gets a 'liveness check.' *AndroidCentral* (June 29, 2012); http://www.androidcentral.com/face-unlock-jelly-bean-gets-liveness-check
8. Duchnowski, P., Hunke, M., Busching, D., Meier, U., and Waibel, A. Toward movement-invariant automatic lip-reading and speech recognition.In *Proceedings of the 1995 International Conference on Acoustics, Speech, and Signal Processing* (Detroit, MI, May 9–12).IEEE Computer Society Press, 1995, 109–112.
9. Hansen, J.H.L. Analysis and compensation of speech under stress and noise for environmental robustness in speech recognition.*Speech Communication 20*, 1 (Nov. 1996), 151–173.
10. Hsu, D., Kakade, S.M., and Zhang, T. A spectral algorithm for learning hidden Markov models.*Journal of Computer and System Sciences 78*, 5 (Sept. 2012), 1460–1480.
11. Jain, A.K., Nandakumar, K., and Ross, A. Score normalization in multimodal biometric systems. *Pattern Recognition 38*, 12 (Dec. 2005), 2270–2285.
12. Kisku, D.R., Gupta, P., and Sing, J.K. Feature-level fusion of biometrics cues:Human identification with Doddingtons Caricature.*Security Technology* (2009), 157–164.
13. Kuncheva, L.I., Whitaker, C.J., Shipp, C.A., and Duin, R.P.W. Is independence good for combining classifiers?In *Proceedings of the 15th International Conference on Pattern Recognition* (Barcelona, Spain, Sept. 3–7).IEEE Computer Society Press, 2000, 168–171.
14. Lee, C. Automatic recognition of animal vocalizations using averaged MFCC and linear discriminant analysis. *Pattern Recognition Letters 27*, 2 (Jan. 2006), 93–101.
15. M2SYS Technology.SecuredPass AFIS/ABIS Immigration and Border Control System; http://www.m2sys.com/automated-fingerprint-identification-system-afis-border-control-and-border-protection/
16. Milner, B. and Xu, S. Speech reconstruction from mel-frequency cepstral coefficients using a source-filter model.In *Proceedings of the INTERSPEECH Conference* (Denver, CO, Sept. 16–20).International Speech Communication Association, Baixas, France, 2002.
17. Nasrollahi, K. and Moeslund, T.B. Face-quality assessment system in video sequences.In *Proceedings of the Workshop on Biometrics and Identity Management* (Roskilde, Denmark, May 7–9).Springer, 2008, 10–18.
18. Parala, A. UAE Airports get multimodal security. *FindBiometrics Global Identity Management* (Mar. 13, 2015); http://findbiometrics.com/uae-airports-get-multimodal-security-23132/
19. Rathgeb, C. and Andreas U. A survey on biometric cryptosystems and cancelable biometrics.*EURASIP Journal on Information Security* (Dec. 2011), 1–25.
20. Ross, A. and Govindarajan, R. Feature-level fusion of hand and face biometrics.In *Proceedings of the Conference on Biometric Technology for Human Identification* (Orlando, FL).International Society for Optics and Photonics, Bellingham , WA, 2005, 196–204.
21. Ross, A. and Jain, A. Multimodal biometrics:An overview.In *Proceedings of the 12th European Signal Processing Conference* (Sept. 6–10).IEEE Computer Society Press, 2004, 1221–1224.
22. Sacco, A. Fingerprint faceoff:Apple TouchID vs. Samsung Finger Scanner.*Chief Information Officer* (July 16, 2014); http://www.cio.com/article/2454883/consumer-technology/fingerprint-faceoffapple-touch-id-vs-samsung-finger-scanner.html
23. Tapellini, D.S. Phone thefts rose to 3.1 million last year.*Consumer Reports* finds industry solution falls short, while legislative efforts to curb theft continue. *Consumer Reports* (May 28, 2014); http://www.consumerreports.org/cro/news/2014/04/smart-phone-thefts-rose-to-3-1-million-last-year/index.htm
24. Viola, P. and Jones, M. Rapid object detection using a boosted cascade of simple features.In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Kauai, HI, Dec. 8–14). IEEE Computer Society Press, 2001.
25. Vondrasek, M. and Pollak, P. Methods for speech SNR estimation:Evaluation tool and analysis of VAD dependency.*Radioengineering 14*, 1 (Apr. 2005), 6–11.
26. Zorabedian, J. Samsung Galaxy S5 fingerprint reader hacked—It's the iPhone 5S all over again!*Naked Security* (Apr. 17, 2014); https://nakedsecurity.sophos.com/2014/04/17/samsung-galaxy-s5-fingerprint-hacked-iphone-5s-all-over-again/

Mikhail I. Gofman (mgofman@fullerton.edu) 是加州州立大学富勒顿分校计算机科学系的助理教授，也是该校网络安全中心的主任。

Sinjini Mitra (smitra@fullerton.edu) 是加州州立大学富勒顿分校的信息系统与决策科学专业的助理教授。

译文责任编辑：诸葛建伟

DOI:10.1145/2810036

追溯后缀树历史中头四十年的点点滴滴，它们的多种形式以及它们的各种应用。

ALBERTO APOSTOLICO，MAXIME CROCHEMORE，MARTIN FARACH-COLTON，ZVI GALIL，S. MUTHUKRISHNAN

后缀树的 40 年历程

当 WILLIAM LEGRAND 最终解密下列字符串时，与之前相比，它透露的意义并没有多增加多少。

53‡‡†305))6*,48264‡.)4z);806”,48†8P60))85;1‡(;:‡*8†83(88)5*†,46(;88*96*?;8)*‡(;485);5*†2:*‡(;4956*2(5*Ñ4)8P8*;4069285);)6‡8)4‡‡;1(‡9;48081;8:8‡1;4885;4)485†528806*81(ddag9;48;(88;4(‡?34;48)4‡;161;:188;‡?;

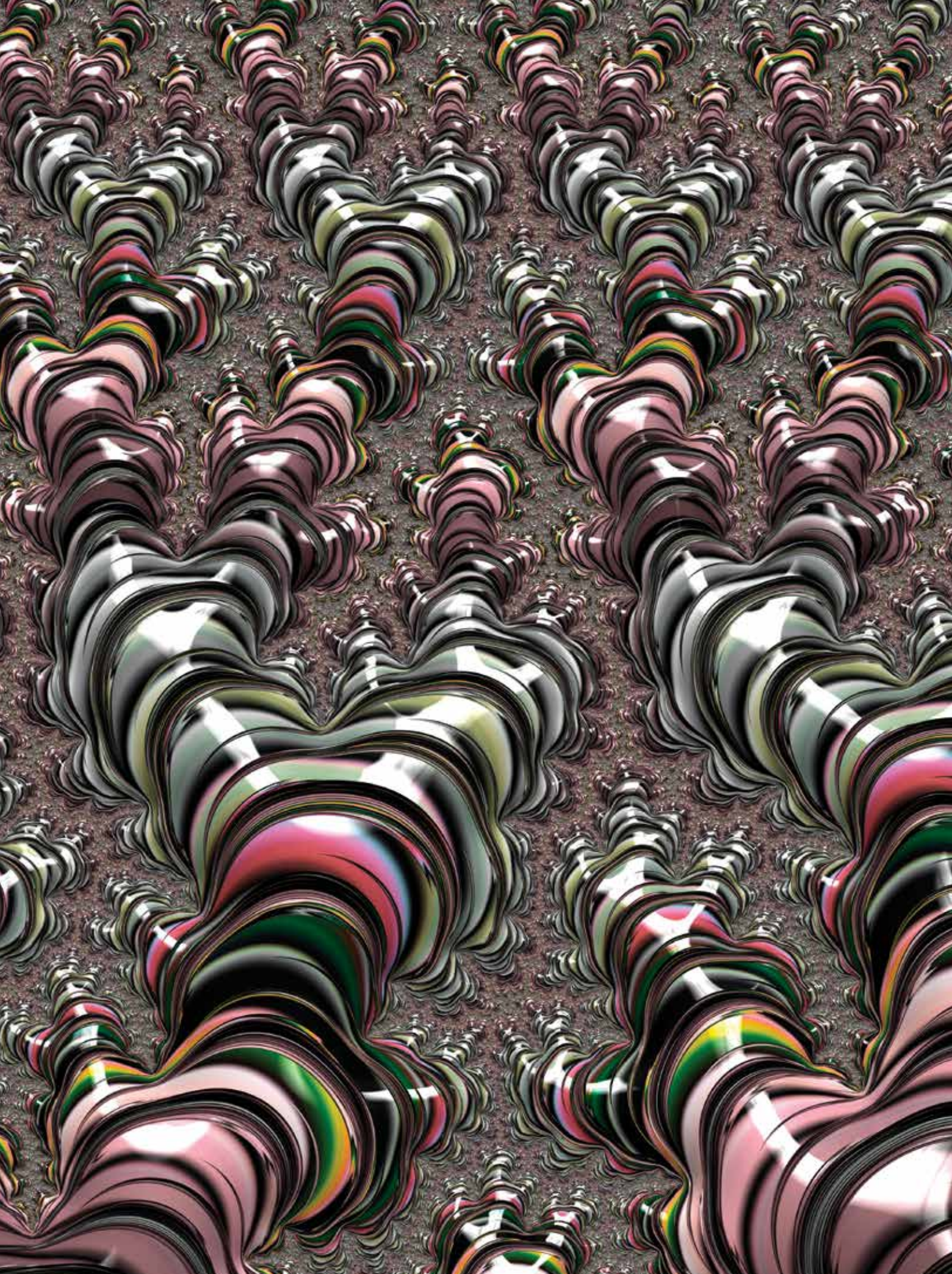
解密后的消息如下：“A good glass in the bishop’s hostel in the devil’s seat forty-one degrees and thirteen minutes northeast and by north main branch seventh limb east side shoot from the left eye of the death’s-head a bee line from the tree through the shot fifty feet out.”但是，至少它更像自然语言，而且它最终引导 Edgar Allan Poe 的“The Gold-Bug（金虫）”[36] 一文中的主角发现了他一直追求的财宝。Legrand 使用符号频率解开了替换式密码。他首先寻找频率最高的符号，然后把它替换成英语中最常见的字母，然后用相似的方式推断出频率最高的词，然后是标点符号等等。

1843 年前后，碰到某些神秘的消息时，人们自然地倾向于计算单独的标记或子块的频率，以找出线索。在此类详查和探索方面，最为深入、最激动人心的研究课题之一或许是生物序列。一旦得到了一些此类序列后，统计分析就能试图把字符或字符片段与相关的生物功能联系起来。早期的例子包括二十世纪九十年代中期出现的全基因组，其中似乎自然地就可以计数所有长度为 1、2 等直到任何期望长度的所有片段出现的次数，以从多个区域中找到编码区域、启动子区域的统计特性。

本文与密码学无关。它阐述了一种数据结构及其变体，以及该数据结构所带有的很多令人惊奇、有用的特性。其中，有一个事实是，对于一个由 n 个字符组成的文本字符串，建立其中任何长度的所有字串（也称为*因子*）的出现次数的统计表时，只需要与文本字符串的长度呈线性的时间和空间。虽然没人会蠢到想通过首先生成数量可能呈指数规模的所有字符串，然后依次计算它们的出现次数来求解这一问题，但是文本字符串可能仍然包含 $\Theta(n^2)$ 个不同的子串，所以，把所有这些子串在线性空间内用表格列出时，更不用说在线性时间内，这看起来仍然难以理解。

本文谨献给我们的朋友和同事 Alberto Apostolico (1948–2015)，他已于 7 月 20 日去世。在字符串算法的发展中，他是一位重要人物。

图片由 PAPUCHALKA 提供



多年以来，在文本搜索、索引、统计和压缩以及生物序列的拼接、比对和比较领域中，此类结构占据了中心位置。它们的范围还扩展到了各种截然不同的领域，如抄袭检测、在文本中找到不平常的子串以及测试某段编码的独特可解密性等等。它们对计算机科学以及更大的IT领域的影响，再怎么强调都不过分。如果没有它们，文本搜索和生物信息学将不会一样。2013年，组合模式匹配年会（Combinatorial Pattern Matching symposium）庆祝Weiner提出后缀树40周年[41]，并为此专门安排了一场特别研讨会。

图 1. 字符串 *x* = *abcabcaba* 的展开的后缀树。

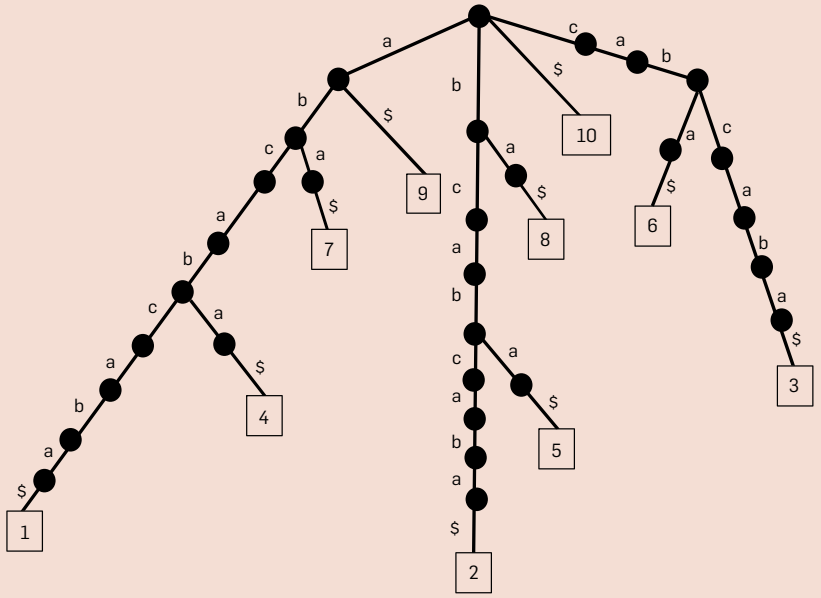
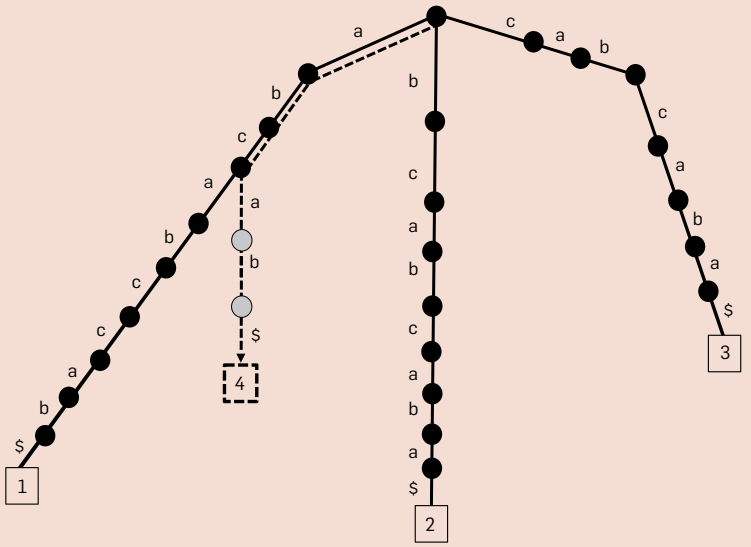


图 2. 通过连续插入后缀构建展开的后缀树（本图说明了插入 *abcaba$* 后的情况）。

后缀 suf_i (i = 1, 2, ⋯, n) 的插入过程包括两个阶段。在第一个阶段中，我们搜索 T_{i-1} 中的 suf_i。注意，$ 的存在保证每个后缀会终结于一个不同的节点。因此，不久之后，该搜索便会失败。此时，我们已经确定了在 T_{i-1} 中有位点（locus）（也就是终结节点）的 sufi 的最长前缀。设本例中的 headi abcab 为该后缀，且 α 为 headi 的位点（locus）。我们可以得出，$suf_i = head_i \cdot tail_i$，其中 $tail_i$（本例中的 *a$*）非空。在第二个阶段，我们需要向 T_{i-1} 添加一条离开节点 α 的路径，并把它标为 $tail_i$。这样就把 T_{i-1} 转换成了 T_i。



历史点滴

在“stringology（字符串学）”到来之前，Donald Knuth猜测，从两个总长度为 n 的长文本序列中找出最长公共子串需要 $(n \log n)$ 的时间。Karp, Miller 和 Rosenberg 提出了时间为 $O(n \log n)$ 的算法。[26] 该方法注定会在并行模式匹配中占据一席之地，不过Knuth的猜测很快就被打破了：1973年，Peter Weiner证明，只要字符串中的字母表是固定的，该问题有在线性时间内求解的优雅方法。[41] 实际上，该解法是他原本为实现其他目的（即，在不列举所有子串的情况下，识别一个文本文件中的任何子串）所创建的方法的副产品。为了达到这一目的，Weiner引入了一个文本倒排索引的概念（在40年里，这一概念诱导出了各种改进、分析和应用）以及计数，其他的数据结构都很少有这一特性。

Weiner最初的构造从右到左处理文本文件。随着每一个新字符的读入，他称为“bi-tree（二叉树）”的结构将被更新，以容纳该文本文件的越来越长的后缀。因此，这本质上是一个离线的方法，因为在该方法开始造之前，必须完全了解该文本。另一方面，也可以说该算法在线构建了该文本的倒排结构。大约三年之后，Ed McCreight提供了一个从左到右的算法，并把该结构的名称改成了“后缀树”，这一名称一直沿用至今。[32]

设 x 为由某个字符集合 Σ 上的 $n-1$ 个符号组成的字符串，$ 是不在 Σ 中的额外字符。与 x 关联的，展开的后缀树 T_x 为收集 $x$$ 的所有

后缀的数字搜索树。具体而言，T_x 定义如下：

1. T_x 有 n 片叶子，从 1 标到 n。

2. 每条弧采用 $\Sigma \cup \{\$\}$ 中的符号进行标记。对于任何 $i, 1 \le i \le n$，从 T_x 的根节点到叶子 i 的路径上的标记链恰好是下列后缀

$$sufi = x_i x_{i+1} \cdots x_{n-1} \$.$$

3. 对于 $x\$$ 的任何两个后缀 suf_i 和 suf_j，如果 w_{ij} 是 suf_i 和 suf_j 的最长公共前缀，那么对于 suf_i 和 suf_j 而言，在 Tx 中，它们关于 wij 的路径相同。

图 1 给出了展开的后缀树的样例。

可把该树解释为确定性有限自动机的状态转换图，其中所有的节点和叶子代表终结状态，根节点代表初始状态，有标记的弧（假定指向下方）代表部分状态转换函数。在该图中未说明的状态变换会让状态进入一个独特的，非终结的汇（*sink*）态。我们的自动机可识别由字符串 x 的所有子串组成的（有限）语言。这一观察还阐明了如何使用该树进行在线搜索的方法：设 y 为模式，我们在该树中沿着向下的路径按 y 中的符号依次搜索，一次一个符号。很明显，当且仅当该过程进入终结状态时，y 在 x 中出现。从 T_x 的角度来看，我们说字符串 y 的位点（locus）是节点 α（如果它存在），这样从 T_x 到 α 的路径就标记为 y。

直接构建展开的 T_x 树（通常称为的后缀 trie）的算法可以轻松地推导出来（见图 2）。我们从空树开始，然后每次向它加入 $x\$$ 的一个后缀。这一过程需要 $\Theta(n^2)$ 的时间和 $O(n^2)$ 的空间。不过，把空间降到 O(n) 相当容易，据此我们可以生成简凑的后缀树（图 3）。一旦这一步骤完成后，人们便有可能去追求一个意料之中不简单的 $O(n)$ 时间复杂度的方法。

在 2013 年的 CPM 会议中，McCreight 透露，他提出 $O(n)$ 时间的结构时，并没有想到作为 Weiner 的结构的替代品——他开发它的目的是为了理解 Weiner 的论文，但是当他向 Weiner 展示他的结构以确认他已经理解该论文，Weiner 的回答是“你并没理解，但是你已经提出了一个完全不同的，优雅的结构！”在未出版的 1975 年的讲义中，Vaughan Pratt 展示了该结构的双重性以及 Weiner 的“重复查找器”。[37]McCreight 的算法本质上仍然是离线的，而且它立刻激发人们寻找在线版本。研究人员对在线算法做出了部分尝试，但是直到约二十年之后，在 1995 年 Esko Ukkonen 的论文中，这一变体才出现。[39] 在所有这些线性时间复杂度方法中，线性性质基于有限数量的字母这一假设。如果这一假设不成立，将需要 $\Theta(n \log n)$ 的时间。1997 年，Martin Farach 提出了一种算法，该算法抛弃了在此之前流行的一次一个后缀的方法；该算法能给出从后缀树构造到字符排序一个线性时间降阶，因此对于所有的字母表，它都是最优的。[17] 具体而言，对于一大类字母表（例如，字母表大小为输入长度的多项式），它的运行时间为线性时间。

重要见解

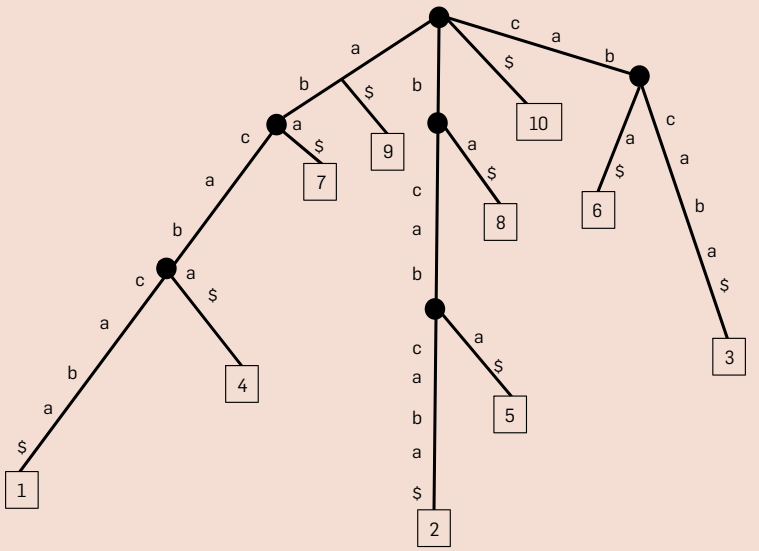
- **后缀树是字符串分析中的核心数据结构。**
- **它的历史厚重，涉及压缩、配对、自动机、数据结构及其他方方面面。**
- **在构建后缀树及在很多应用中有效使用后缀树方面，存在强大的技术。**

1984 年左右，Blumer 等人[9] 和 Crochemore[14] 揭露了令人惊喜的结果，即识别包含 n 个字符的字符串的所有后缀且只识别这些后缀的最小有限自动机拥有 $O(n)$ 个状态和边。在初始构建一个有向无环词图（DAWG）后，如果所有的状态都是终结状态，它甚至还能约减。[14] 然后，它接收该字符串的所有子串（称为因子——子串自动机）。当该字符串没有结束标记且它的后缀在树中都被标记为终结状态时，这些索引数据结构之间存在美妙的关联。

那么，该后缀树成了该树的边合并版本，且它的节点数量可

图 3. 简凑形式的后缀树

首先通过把由只有一个孩子的节点形成的每一条链合并成单弧，可得到此类后缀树。由此得出的，T_x 的简凑版本最多只有 n 个内部节点，因为总共只有 n + 1 片叶子，而每个内部节点都有分支。现在，该广义弧的标签是子串，而不是 $x\$$ 符号。不过，弧的标记可以通过指向 $x\$$ 的公用副本的，合适的指针对来表示，因此总体满足了 O(n) 的空间边界。



被最小化，这点和自动机有些相似，最后它得到了该字符串的紧凑的 DAWG。交换这两种操作的顺序（合并（compaction）和最小化（minimization））后，也会得到相同的结构。很明显，在与检测字符串重复有关的研究中，Anatoli Slissenko【ACM Digital Library 中 Source Material（原始资料）处的本文附录】最终也得到了类似的结构。当人们使用这些自动机作为模式匹配机匹配纹理（grain）时，它们提供了另一个与 Knuth 的猜想不符的，效率更高的反例（见图 4）。

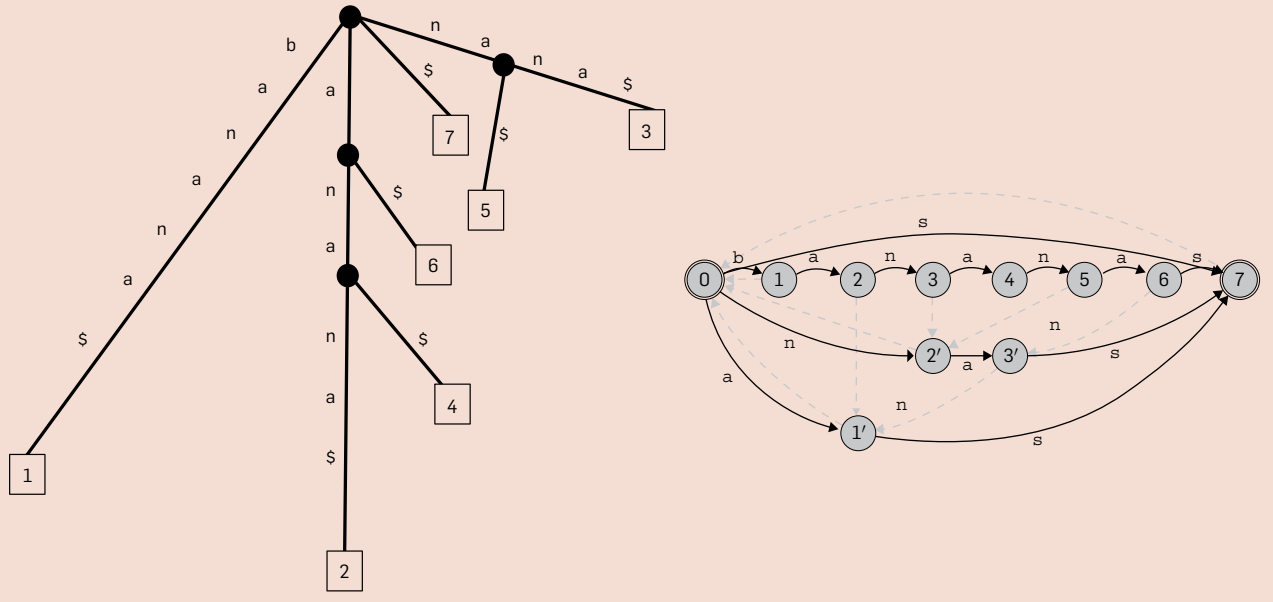
后缀树的出现恰好与信息论中某些有趣的和独立的进展相吻合。在他著名的信息概念方法中，Kolmogorov 认为字符串中的信息或结构等于通用图灵机生成该字符串所需的最小程序的长度。不幸的事情是，该度量是不可计算的。即便它可以计算，大多数长字符串也不可压缩（即，缺少生成它们的短程序），因为长字符串的数量越来越多，短程序的数量相对要少的多（它们本身也是字符串）。

Kolmogorov 的通用万能机可以利用任何人们能想象出的规则类型。但是，如果人们把它们限制在以重复子串的形式影响某个文本的句法冗余方面，那又会如何呢？如果某个字符串重复了多次，人们可以把所有的出现次数编码成指向同一份副本的指针，从而从中获益。该副本可以在文本的内部或外部。在前一种情况时，人们可以让这些指针拥有两个方向，或只有一个方向，或允许或禁止指针嵌套等等。在他的博士论文中，Jim Storer 说明，几乎所有的此类“宏观方案（macro schemes）”都是无法驾驭的，但有一个例外。在此前不久，在题为“论有限序列的复杂性（On the Complexity of Finite Sequences）”[30] 的标志性论文中，Abraham Lempel 和 Jacob Ziv 提出了块变量（variable-to-block）编码，其基于对文本的简单语法分析，在极限情况下，达到的压缩将与源概率相适配的压缩程序所能生成的压缩接近。因此，通过对这些星形进行巧妙的比对，Lempel 和 Ziv 提出的压缩方法不仅在信息论方面是最优的，而且通过后缀树，它还找出了一个最优的、线性时间的实现。Michael Rodeh，Vaugham Pratt 和 Shimon Even 接着就对此进行了详细的阐述。[38]

在开创性的论文中，Weiner 列出了“bi-tree（B 树）”的一些应用，其中包括最有名的离线字符搜索应用：预处理一个文本文件，以支持在与给定模式的长度呈线性时间内对给定模式的出现情况进行查询。当然，通过说明如何在线性时间内从由有限字母组成的两个文件中找

图 4. 字符串“bananas”的简凑的后缀树（左）和后缀自动机（右）

失败的连接用虚线表示。尽管事实上它是字符串的索引，但相同的自动机也可作为模式匹配机在另一个文本中定位“bananas”的子串，或计算它们的最长公共子串。在处理第二个字符串时，该过程是在线运行的。例如，假设刚刚从第二个字符串中扫描了“bana”，该自动机的当前状态为状态 4. 如果下一个字母是“n”，那么公共子串是长度为 5 的“banan”，新状态是 5。如果下一个字母是“s”，那么使用失败连接，从对应于状态 3 的，长度为 3 的公共子串“ana”中，我们得到了公共子串“ana”以及新状态 7。如果下一个字母是“b”，那么反复使用失败连接，最后到达状态 0，我们得到公共子串“b”和新状态 1。最终，接下来的任何其他字符将生成空的公共子串和状态 0。



到最长公共子串的方法，该“b 树”解决了 Knuth 的猜想。随后，Pratt 编写了题为“对 Weiner 的重复查找器的改进和应用”的讲义，但该讲义并未出版。[37] 十年之后，在题为“后缀树的多种优点”的论文中，Alberto Apostolico 列出了更多的应用。[2] 二十年之后，在 Crochemore 和 Rytter，Dan Gusfield，和 Crochemore，Hancart，以及 Lecroq 编著的参考书中，后缀树及其同类结构占据了几个章节（见 ACM Digital Library 中本文的附录）。

在最需要后缀树的应用中，后缀树所需的空间已经成了令人讨厌的问题。例如，对于基因组而言，所需空间达到了 GB 的数量级。所以，源数据的 20 倍的空间与元数据的 11 倍的空间之间的差异可能非常大。Stefan Kurtz 和他的同僚又用了十多年的时间，一直致力于研究如何巧妙地分配树及其某些同类结构。[28]2001 年，David R. Clark 和 J. Ian Munro 提出了在二级存储器上节省空间的最佳方法之一。[13]Clark 和 Munro 的“简洁后缀树”寻求尽可能地保持后缀树的结构。不过，Udi Manber 和 Eugene W. Myers 采用了不同的方法。1990 年，他们引入了“后缀数组”，[31] 它消除了后缀树的大部分结构，但仍能够实现很多相同的操作，它需要的空间为每个文本字符占 2 个整数的空间，且搜索时间为 $O(|P| + \log n)$（如果接受搜索时间为 $O(|P| + \log n)$，可降阶至 1）。后缀数组按字典顺序存储了输入的后缀，可视为在后缀树中按字典顺序展开每一个节点而进行先序遍历找到的叶子标记的序列。

虽然初看起来后缀数组似乎是不同于后缀树的数据结构，但它们的区别已经减弱。例如，对于任何字母表，Manber 和 Myers 最初构建的后缀数组需要 $O(n \log n)$ 的时间，但是对于任何字母表，可以在线性时间内从后缀树中构造出后缀数组。2001 年，Toru Kasai 等人[27] 证明，可以在线性时间内从后缀数组中构造出后缀树。因此，说明了后缀数组是后缀树的简洁表示。2003 年，对于 Farach 的后缀树构造算法，三个小组提出了三种不同的修改方式，首次给出了直接构造后缀数组的线性时间算法；也就是说，首次提出了在计算后缀数组时不先计算出完整的后缀树的线性时间算法。此后，又出现了很多快速构建后缀数组的算法，其中较著名的是由 Nong，Zhang 和 Chan 提出的算法[35]，它是线性时间的，在实践中相当快。由于拥有快速构建的算法且需要的空间较小，后缀数组成了在软件系统中使用最为广泛的后缀树变体。Grossi 和 Vitter 提出了更新的简洁后缀树和数组，对于一个二进制字母表，它需要 $O(n)$ 位表示（否则为 $O(n \log \sigma$ 位）。[21]

虽然初看起来后缀数组似乎是不同于后缀树的数据结构，但它们的区别较为模糊。

实际上，后缀树和压缩的历史紧密结合在一起。这并不在意料之外，因为模式发现试图揭示的冗余是压缩时要移除的理想备选项。1994 年，M. Burrows 和 D.J. Wheeler 提出了一个基于后缀排序的突破性的压缩方法。[11] 大约在 1995 年，Amihood Amir, Gary Benson，Gary Benson 和 Martin Farach 提出了在压缩的文本中进行搜索的问题。[1]2000 年，Paolo Ferragina 和 Giovanni Manzini 引入了 FM-index，这是一个基于 Burrows-Wheeler 变换的压缩后缀树。[19] 该结构可能会比源文件小，它支持在不解压的情况下执行搜索。使用 Burrows-Wheeler 变换的修改版本后，该方法可以扩展到 Ferragina 等人的论文中提出的压缩树索引问题。[18]

不良结果、扩展及挑战

正如篇首强调的那样，几乎所有的文本处理应用都在这个或那个地方需要这些索引。最著名的案例为有错误搜索，Gad Landau 在 1985 年的博士论文中首次有效地处理这一问题。[29] 在此类搜索中，人们寻找与某个模

式不同但错误的数量有限的文本子串，如删除、插入或替换了单个字符。为了有效地解决这一问题，Landau把后缀树与解决所谓最低共同祖先（LCA）问题的灵巧方案结合了起来。LCA问题假设给定了一颗有根的树，然后对于任何一对节点，它查找这两个节点在树中的最低祖先节点。[23] 人们发现，在对树进行线性时间的预处理后，任何LCA查询可以在常数时间内得到答复。Landau在后缀树上使用LCA查询，在常数时间内在确保可与模式匹配的文本片段间跳跃。当允许 k 个错误时，对某个给定位置的实例搜索可以在 k 次此类跳跃后放弃。这引出了一个算法，该算法可在 $O(nk)$ 步内在由 n 个字符组成的文本中搜索出与某模式相差 k 个错误的子串。

当然，除了后缀树和数组支持的基本类型之外，人们发现前文提到的，在文本中搜索某个模式所需的时间与模式的长度，而不是文本的长度成正比。事实上，甚至还能在与它们的数量成正比的时间内枚举各次出现的情况，而且，只要对该树稍加处理后，在与模式大小成正比的时间内能得出任意查询模式的总出现次数。此前，我们已经注意到了找出文本中出现两次的最长子串或两个文件均包含的最长子串的问题：这可能是所有问题的起源。与之密切相关的一个问题是检测文本中的块（squares）、重复和最大周期性，这个问题根源于Axel Thue一个多世纪之前的研究，在压缩和DNA分析中有多种当代的应用。块（square）是由同一字符串的两次相邻出现组成的一种模式。后缀树已经被用于在最优的 $O(n \log n)$ 时间内检测文本中的所有块（squares）（或重复），其中每一个都标出了它的起始位置集合，[5] 且后面可在线性时间内找出和保存文本中所有不同的块子串。在扩展后缀树的作用，使其适于报告任何查询模式的非重合出现次数方面，块发挥了重要的作用。[6,10]

在为文本字符串设置某种类型的签名以及测量其相似性或差异方面，后缀树可发挥多种作用。

在为文本字符串设置某种类型的签名以及测量其相似性或差异方面，后缀树可发挥多种作用。在测量相似性或差异方面，存在需要计算文本中的禁忌词或不包含的词的问题，它们是在文本中没有出现的最小字符串（但是，在文本中出现了它们所有严格意义上的子串）。[8,15] 除了其他方面之外，这些词还引出了文本压缩的新颖方法。[16] 一旦人们把后缀树当成文本的“词袋（bag-of-words）”的简凑表示后，便可使用后缀树评估两个文本文件的相似性，从而可支持聚类、文档分类、甚至是种系发生树（phylogeny）。[4,12,40] 从直觉上来说，通过评估两个输入序列的树之间有多少共同之处，可以做到这点。增加在每个节点终结的子串的概率信息后，树可用于检测过度表达达到惊人程度的，任何长度的子串。[3] 例如，在生物序列的启动子区域中查找过度表达的子串。

又如，把 $k \geq 2$ 的多个文本文件的连接后，得到的后缀树支持高效地解决多种问题，范围从抄袭检测到生物序列的motif发现。为了保持线性时间复杂度，由于需要 k 个独特的结束标记，造成了一些微妙的问题。有关这方面，读者可以阅读Gusfield的文章。[22] 在最早的时候，生成多个文本的索引的问题被称为“着色问题”。对于任何给定的查询字符串，在该查询的线性时间内，它设法报告在总共 k 个文件中在多少个文件中至少出现了一次该查询字符串。1992年，Lucas C.K. Hui[25] 给出了一个简单和优雅的解决方案。最近，由很多字符串组成的组合后缀树（combined suffix tree）（也被称为广义后缀树）被用于解决各种文件列表问题。此处，文本文档的集合被预处理为组合后缀树。现在的问题是在与此类文档的数量成正比的时间内返回含有查询模式的所有文档的列表，而不是在与总出现次数（occ）成正比的时间内，因为后者明显要大得多。通过把它约简为*范围最小查*

询（*range minimum queries*），Muthukrishnan[33] 解决了这一问题。自此以后，这一基本的文档列表问题被扩展到用于解决很多其他的问题，包括在多个字符串和信息距离内列出前 *-k* 个符合条件的结果。例如，在 Hon[24] 等人的论文中，他们讨论了在与模式大小接近线性的时间内，返回包含模式 *p* 的前 *k* 个频率最高的文件列表的线性机器 - 词数据结构，其中广义后缀树的结构是设计该数据结构的关键点。

Brenda Baker 介绍了该后缀树的一个精妙变体，用于检测学生报告中的抄袭情况以及软件开发中的优化工作。[7] 上述模式匹配的变体被称为“*参数化匹配*（*parameterized matching*）”，它支持人们找到除了系统性的重新命名参数以外相同的程序片段，或者相同的子串，从而让人们进行系统性地重标记或者系统性地改变字母表中字符的排列除了系统性的重新标记或者系统性的改变字母表中字符的排列以外相同的子串。对后缀树概念的明显扩展是把它扩展到多于一个维度，虽然扩展的机制本身还很模糊。[34] 在与后缀树联系更弱的结构中，有“*小波树*（*wavelet trees*）”。起初，小波树是作为压缩后的后缀数组表示提出的，[20] 它支持人们在广义的字母表（之前只能在位向量）上执行排序和选择等基本操作。

这一列表还可以继续往下写，但是本文的目的并不是穷举。事实上，经过 40 年不间断的发展，可以合理地假设该列表还会继续增长。也会出现大量待解决的问题。例如，在观察的序列中，有很多用数字表示，而不是字符表示。在上述两种情况，均会受到了各种类型的错误的影响。虽然两个字符的比较结果只有一位，但是两个数字的差别可大可小，具体取决于它们的差异或某个其他的度量尺度。与此类似，两个文本字符串的相似程度可大可小，具体取决于改变某个字符串使之成为另一个字符串所需的基本步骤的数量。在该框架中，最具颠覆性一面是丢失了传递性，而传递性促成了最高效的精确字符串匹配法。不过，如果索引能够支持刚刚着重提到的快速、优雅的近似模式查询类型，那么它们的作用也非常大。希望它们能够很快出现，而且有它们自己的 40 周年庆祝。

致谢在此感谢 Ed McCreight、Ronnie Martin、 Vaughan Pratt、Peter Weiner 和 Jacob Ziv 的讨论和帮助。特别感谢本文的推荐人，他们详细审阅了本文之前的版本，促成了本文的多处改进。

参考资料

1. Amir, A., Benson, G. and Farach, M. Let sleeping files lie:Pattern matching in Z-compressed files.In *Proceedings of the 5th ACM-SIAM Annual Symposium on Discrete Algorithms* (Arlington, VA, 1994), 705–714.
2. Apostolico, A. The myriad virtues of suffix trees. *Combinatorial Algorithms on Words, vol. 12 of NATO Advanced Science Institutes, Series F.* A. Apostolico and Z. Galil, Eds. Springer-Verlag, Berlin, 1985, 85–96.
3. Apostolico, A., Bock, M.E. and Lonardi, S. Monotony of surprise and large-scale quest for unusual words.*J. Computational Biology 10*, 3 / 4 (2003), 283–311.
4. Apostolico, A., Denas, O. and Dress, A. Efficient tools for comparative substring analysis.*J. Biotechnology 149*, 3 (2010), 120–126.
5. Apostolico, A. and Preparata, F.P. Optimal off-line detection of repetitions in a string.*Theor.Comput. Sci.22*, 3 (1983), 297–315.
6. Apostolico, A. and Preparata, F.P. Data structures and algorithms for the strings statistics problem. *Algorithmica 15*, 5 (May 1996), 481–494.
7. Baker, B.S. Parameterized duplication in strings:Algorithms and an application to software maintenance.*SIAM J. Comput.26*, 5 (1997), 1343–1362.
8. Béal, M.-P., Mignosi, F. and Restivo, A. Minimal forbidden words and symbolic dynamics.In *Proceedings of the 13th Annual Symposium on Theoretical Aspects of Computer Science, vol. 1046 of Lecture Notes in Computer Science* (Grenoble, France, Feb. 22–24, 1996).Springer, 555–566.
9. Blumer, A., Blumer, J., Ehrenfeucht, A., Haussler, D., Chen, M.T. and Seiferas, J. The smallest automaton recognizing the subwords of a text.*Theor.Comput. Sci.40*, 1 (1985), 31–55.
10. Brodal, G.S., Lyngsø, R.B., Östlin, A. and Pedersen, C.N.S. Solving the string statistics problem in time *O*(*n* log *n*).In *Proceedings of the 29th International Colloquium on Automata, Languages and Programming, vol. 2380 of Lecture Notes in Computer Science* (Malaga, Spain, July 8–13, 2002).Springer, 728–739.
11. Burrows, M. and Wheeler, D.J. A block-sorting lossless data compression algorithm.Technical Report 124, Digital Equipment Corp., May 1994.
12. Chairungsee, S. and Crochemore, M. Using minimal absent words to build phylogeny.*Theoretical Computer Science 450*, 1 (2012), 109–116.
13. Clark, D.R. and Munro, J.I. Efficient suffix trees on secondary storage.In *Proceedings of the 7th ACM-SIAM Annual Symposium on Discrete Algorithms*, (Atlanta, GA, 1996), 383–391.
14. Crochemore, M. Transducers and repetitions.*Theor. Comput.Sci., 45*, 1 (1986), 63–86.
15. Crochemore, M., Mignosi, F. and Restivo, A. Automata and forbidden words.*Information Processing Letters 67*, 3 (1998), 111–117.
16. Crochemore, M., Mignosi, F., Restivo, A and Salemi, S. Data compression using antidictonaries.In *Proceedings of the IEEE:Special Issue Lossless Data Compression 88*, 11 (2000).J. Storer, Ed., 1756–1768.
17. Farach, M. Optimal suffix tree construction with large alphabets.In *Proceedings of the 38th IEEE Annual Symposium on Foundations of Computer Science* (Miami Beach, FL, 1997), 137–143.
18. Ferragina, P., Luccio, F., Manzini, G. and Muthukrishnan, S. Compressing and indexing labeled trees with applications.*JACM 57*, 1 (2009).
19. Ferragina, P. and Manzini, G. Opportunistic data structures with applications.In *FOCS* (2000), 390–398.
20. Grossi, R., Gupta, A. and Vitter, J.S. High-order entropy-compressed text indexes.In *SODA* (2003), 841–850.
21. Grossi, R. and Vitter, J.S. Compressed suffix arrays and suffix trees with applications to text indexing and string matching.In *Proceedings ACM Symposium on the Theory of Computing* (Portland, OR, 2000).ACM Press, 397–406).
22. Gusfield, D. *Algorithms on Strings, Trees and Sequences:Computer Science and Computational Biology*.Cambridge University Press, Cambridge, U.K., 1997.
23. Harel, D. and Tarjan, R.E. Fast algorithms for finding nearest common ancestors.*SIAM J. Comput.13*, 2 (1984), 338–355.
24. Hon, W.-K., Shah, R. and Vitter, J.S. Space-efficient framework for top-*k* string retrieval problems.In *FOCS*.IEEE Computer Society, 2009, 713–722.
25. Hui, L.C.K. Color set size problem with applications to string matching.In *Proceedings of the 3rd Annual Symposium on Combinatorial Pattern Matching, no. 644 in Lecture Notes in Computer Science*, (Tucson, AZ, 1992).A. Apostolico, M. Crochemore, Z. Galil, and U. Manber, Eds. Springer-Verlag, Berlin, 230–243.
26. Karp, R.M., Miller, R.E., and Rosenberg, A.L. Rapid identification of repeated patterns in strings, trees and arrays.In *Proceedings of the 4th ACM Symposium on the Theory of Computing* (Denver, CO, 1972).ACM Press, 125–13.
27. Kasai, T., Lee, G., Arimura, H., Arikawa, S. and Park, K. Linear-time longest-common-prefix computation in suffix arrays and its applications.*CPM*.Springer-Verlag, 2001, 181–192.
28. Kurtz, S. Reducing the space requirements of suffix trees.*Softw.Pract.Exp.29*, 13 (1999), 1149–1171.
29. Landau, G.M. String matching in erroneus input.Ph.D. Thesis, Department of Computer Science, Tel-Aviv University, 1986.
30. Lempel, A. and Ziv, J. On the complexity of finite sequences.*IEEE Trans.Inf.Theory 22* (1976), 75–81.
31. Manber, U. and Myers, G. Suffix arrays:A new method for on-line string searches.In *Proceedings of the 1st ACM-SIAM Annual Symposium on Discrete Algorithms* (San Francisco, CA, 1990), 319–327.
32. McCreight, E.M. A space-economical suffix tree construction algorithm.*J. Algorithms 23*, 2 (1976), 262–272.
33. Muthukrishnan, S. Efficient algorithms for document listing problems.In *Proceedings of the 13th ACM-SIAM Annual Symposium on Discrete Algorithms* (2002), 657–666.
34. J. C. Na, P. Ferragina, R. Giancarlo, and K. Park.Two-dimensional pattern indexing.In Encyclopedia of Algorithms.2008.
35. Nong, G., Zhang, S. and Chan, W.H. Two efficient algorithms for linear time suffix array construction. *IEEE Trans.Comput.60*, 10 (2011), 1471–1484.
36. Poe, E.A. *The Gold-Bug and Other Tales*.Dover Thrift Editions Series.Dover, 1991.
37. Pratt, V. Improvements and applications for the Weiner repetition finder.Manuscript, 1975.
38. Rodeh, M., Pratt, V. and Even, S. Linear algorithm for data compression via string matching.*J. Assoc. Comput.Mach.28*, 1 (1981), 16–24.
39. Ukkonen, E. On-line construction of suffix trees. *Algorithmica 14*, 3 (1995), 249–260.
40. Ulitsky, I., Burstein, D., Tuller, T. and Chor, B. The average common substring approach to phylogenomic reconstruction.*J. Computational Biology 13*, 2 (2006), 336–350.
41. Weiner, P. Linear pattern matching algorithms.In *Proceedings of the 14th Annual IEEE Symposium on Switching and Automata Theory*, (Washington, D.C., 1973), 1–11.

Alberto Apostolico 同时在佐治亚理工学院计算科学学院以及交互计算工程学院任教授和研究员。他于 2015 年 7 月 20 日去世。

Maxime Crochemore (maxime.crochemore@kcl.ac.uk) 是伦敦国王学院和法国巴黎东区大学教授。

Martin Farach-Colton (farach@cs.rutgers.edu) 是新泽西州皮斯卡塔韦罗格斯大学计算机科学系教授。

Zvi Galil (galil@cc.gatech.edu) 是乔治亚州亚特兰大佐治亚理工学院计算学院院长。

S. Muthukrishnan (muthu@cs.rutgers.edu) 是新泽西州皮斯卡塔韦罗格斯大学计算机科学系教授。

译文责任编辑：陈恩红

DOI:10.1145/2898429

技术视角
公平性与掷硬币

撰稿人：David Wagner

如需查阅相关论文，
请访问 doi.acm.org/10.1145/2896386

爱丽丝和鲍勃一起吃了顿愉悦的晚餐，想要随机抽出负责洗碗的人。如何才能做到公平？有一种传统的做法，爱丽丝抛掷一个硬币（挡住不让鲍勃看见），鲍勃猜测哪一面朝上，然后爱丽丝展示硬币，从而公布谁是要洗碗的倒霉鬼。两人都可亲眼验证过程是否公平。

假如爱丽丝和鲍勃相隔万里，只能通过互联网沟通呢？三十多年前，密码学家为解决这种硬币抛掷问题设计了一个聪明的方案：大致如下，爱丽丝抛掷一个硬币，再向鲍勃发送结果的加密哈希；鲍勃把猜测结果发送给爱丽丝；接着爱丽丝公布硬币抛掷结果，这样爱丽丝和鲍勃便可检验谁胜谁负。在分布式环境中，如果互不信任的多方希望联合生成不受人影响或没有偏倚的随机值，此协议可以派上用场。

之后的论文为如何提供公平性引入了一种令人欣喜的新想法：将比特币的现有基础结构用于实现分布式一致性。

但遗憾的是，这种方案有个缺点。爱丽丝要比鲍勃先知道硬币抛掷结果。如果爱丽丝是个不诚实或输不起的人，她就能得到不公平的优势。在鲍勃发出猜测答案后，爱丽丝便知道自己的输赢；如果是赢家，她可以继续公布硬币抛掷结果并赢得钱；但如果是输家，她可以拒绝继续执行协议，中断与鲍勃的连线，并且（必要时）声称自己的电脑崩溃了。如此一来，不诚实的爱丽丝可以确保自己至少不输，而这对鲍勃而言不公。这被称为*公平性问题*。

在一些应用中，不公平可以容忍。比如，设有惩罚作弊者的机制，或者开始抛掷硬币前各方必须先在可信赖的保管人那里存上押金。而在其他情形中，这便是严重问题。研究人员考察了许多提供公平性的方法，但无一百分百令人满意。此外，还有负面结果：在没有可信赖第三方解决争议的一般环境中，公平性问题似乎无解。普遍的观点偏向于认为这基本上是无法避免的问题。

之后的论文为如何提供公平性引入了一种令人欣喜的新思路：将比特币的现有基础结构用于实现分布式系统一致性。比特币是一种精密的分布式系统，其设计甚至能够抵御高度复杂、资源丰富的攻击者。论文作者演示了如何构建安全性依赖于比特币所提供的基础的密码协议：破坏密码协议需要先攻破比特币，而这被认为是颇有难度的。

该论文利用了比特币技术的一个迷人特点。比特币提供交易审核日志，并且允许交易包含脚本，即确定交易是否发生的程序。作者利用比特币的这一方面来获得公平性：通过脚本实现其他情形中需要可信赖第三方保管服务提供的功能。

更广泛地说，分布式硬币抛掷不是我们想要在分布式世界中执行的唯一任务。数十年前，密码学家研究了多方安全计算的普遍问题，即爱丽丝和鲍勃希望一起对他们的数据进行某种计算，但不想将自己的数据透露给对方。硬币抛掷只是这种范式的一个例子。密码学家已经演示了一个非常有力的结果：这种形式的每一任务基本上都能安全地执行。然而，这些协议依然存在不可避免的公平性问题：一方在另一方之前获知计算结果，可以提前终止协议并阻止对方获得输出结果。该论文令人兴奋的一点是，它为获得普通多方安全计算的公平性指明了一个方向，只要各方愿意使用比特币。谁曾料到，比特币竟对安全分布式计算有这等影响？ C

David Wagner 是加州大学伯克利分校的计算机科学教授。

译文责任编辑：陈海波

DOI:10.1145/2896386

借助比特币进行安全多方计算

撰稿人：Marcin Andrychowicz、Stefan Dziembowski、Daniel Malinowski 和 Łukasz Mazurek

摘要

有没有可能设计出一种在线协议，以不依赖于受信第三方、彻底去中心化的方式进行抽奖？或者，是否有人能够构建一种完全去中心化的协议来销售秘密信息，确保卖家或买家都无法作弊？直到最近，似乎每一种在线协议都有财务上的后果，参与者需要依赖某种受信服务器，确保钱款在他们之间转移。在本文中，我们提出利用比特币（2008 年推出的一种数字货币）来设计此类完全去中心化的协议，即使没有受信第三方也能确保安全。作为这种创意的实现，我们构造利用比特币进行多方抽奖、且不依赖受信权威方的协议。我们的协议能为诚实的参与方保证公平性，无论输家行为如何。例如，如果有一方中断协议，其钱款将转给诚实的参与方。我们的协议比较实用（为进行演示，我们在实际的比特币系统中执行了交易），而且理论上可在现实生活中取代在线赌博网站。

1. 引言

互联网的一大引人之处在于它的去中心化：TCP/IP 协议以及在其基础上运行的其他协议，都不依赖单一的服务器，通常可以在不需要互相信任、甚至无需知道对方真实身份的多方之间执行。例如，这一类的协议有 SMTP 和 HTTP 协议、P2P 内容分发平台、信息收发系统。一个问题油然而生，即数字世界“去中心化”到底能走多远？换种问法，可以在互联网上实现的、无需受信第三方的现实应用有哪些？直到最近，比较著名的始终需要某种“受信服务器”的应用实例是在线金融交易，它必须依赖银行或信用卡公司。当中本聪[17,a]在 2009 年部署了称为比特币的第一种完全去中心化的数字货币，上述局面有了巨大的改观。比特币的巨大成功（当前市值约 50 亿美元）完全源自于它的分布式本质，并且无需中央权威机构控制比特币交易。我们将在第 2 节中更详细地阐述比特币。

比特币能够无需受信服务器汇钱转账，这提出了另一个发人深省的问题：我们能否更进一步“去中心化”金融系统？也就是说，我们能否以分布式方式实施一些更为高级的金融手段？比特币规范通过提供所谓的“非标准交易”，在一定程度上回答了此问题。我们将在第 2 节中详细说明此功能。但目前，我们先只说比特币允许各方就何时支出资金指定更加复杂的条件。这转而又允许他们创建所谓的“比特币合约”，也就是之后由比特币系统本身保证执行的、不需要受信第三方的参与协议形式。此类合约的示例包括迅速调整的微支付、保证合约和争议调解（详情见 https://en.bitcoin.it/wiki/Contracts）。

或许，能以数字方式执行的多方协议中最高级的一种是加密“安全多方计算 (MPC)”协议，它源自 Yao[20] 和 Goldreich[14] 等人的开创性研究。通俗而言，此类协议允许一组互不信任的参与方对他们的秘密输入进行联合函数 f 计算。比如，有爱丽丝和鲍勃这两方。爱丽丝输入 x，鲍勃输入 y，两人都想知道 $f(x,y)$，但不希望爱丽丝知道 y 或鲍勃知道 x。在本文中，我们展开了使用比特币执行 MPC 协议的研究。

硬币抛掷协议。此类协议的一个简单示例是*硬币抛掷问题*[6]，该协议在爱丽丝和鲍勃这两方之间执行，他们想要联合计算出等可能为 0 或 1 的一个比特值 b。换而言之，他们想要计算一个随机化函数 $f_{\text{rnd}} : \{\perp\} \times \{\perp\} \to \{0, 1\}$，此函数没有输入但输出一个均匀随机的比特值。此协议可以采用类似于石头剪子布游戏的理念来实施：爱丽丝向鲍勃发送比特 b_A，同时鲍勃向爱丽丝发送比特 b_B。输出 b 计算为 $b := b_A \oplus b_B$（其中“$\oplus$”表示 x 或函数）。显然，如果 b_A 和 b_B 这两个比特中至少有一个是均匀随机数，b 也会是均匀随机数。所以，只要自己行为诚实（均匀地选择其比特值），任一方都能确保游戏公平。如果要尝试在互联网上实

本文的较长版本发表于 2014 IEEE 安全与隐私研讨会论文集。本文的加长版也可在 Cryptology Eprint Archive (eprint.iacr.org/2013/784) 找到。本研究得到了 WELCOME/2010-4/2 基金（于 EU Innovative Economy (National Cohesion Strategy) Operational Programme 框架下授予）的支持。Łukasz Mazurek 是 Google Europe Fellowship in Security 奖学金获奖者，此研究在一定程度上受到此项谷歌奖学金的支持。

[a] 此名称被广泛认为是假名。

施此协议，主要挑战显然就是确保爱丽丝和鲍勃同时发送自己的比特值。这是因为，如果有一方（假设是爱丽丝）可以在知道 b_B 后选择自己的比特值 b_A，那就能使 b 等于自己想要的任何值 b'，只要选择 $b_A := b' \oplus b_B$ 便可。

Blum[6] 的文章中提出了解决方案，使用一种称为*加密承诺方案*的工具。通俗地来说，这种方案是在*承诺方*和*接收方*之间执行一个双方协议。最初，承诺方知道某个值 s，该值不为接收方所知。双方先执行*承诺阶段* (*Commit*)。此阶段完毕后，接收方依然不知道 s（此属性称为*隐藏*）。然后，双方执行公开阶段 (*Open*)，这一阶段接收方获知 s。承诺方案的关键属性在于，承诺方不能在承诺阶段之后“改变主意”。更准确地说，执行第一个阶段之后，便存在一个确切的 s 值，它可在第二阶段中公开。此属性称为*绑定*。在某种意义上，承诺阶段可以比作将消息 s 放在上锁的盒子中发送，公开阶段则可视为发送盒子的钥匙。显然，发送出盒子后承诺方就无法更改其内容，而在获得钥匙之前，接收方不清楚盒中有何物。

构建此类承诺的安全方法有许多。本文中我们采用基于密码学的哈希函数的方法（见第 3 节）。

现在，我们可以轻松了解如何使用承诺方案来解决硬币抛掷问题：爱丽丝不直接将自己的比特值 b_A 发送给鲍勃，而是进行承诺（也就是说，爱丽丝和鲍勃执行承诺方案，其中爱丽丝充当承诺方，鲍勃充当接收方，而 b_A 则是秘密值）。对应地，鲍勃也承诺其比特值 b_B。此承诺阶段结束后，双方执行公开阶段，并获悉对方的比特值。那么，其输入计算为 $b = b_A \oplus b_B$。承诺方案的安全性确保无任何一方在知道另一方的比特之后选择自己的比特，因此这一过程可以产生均匀随机的比特。

布尔运算。上述硬币抛掷问题是多方协议的一个极简情形，因为执行它的各个参与方不用取任何输入值。若要了解我们所说的各方取输入值的协议，可以思考这样的情形：爱丽丝和鲍勃计算的函数是交集 $f_\wedge(a, b) = a \wedge b$，其中 a、$b \in \{0, 1\}$ 是分别表示爱丽丝和鲍勃的输入值的布尔变量。这有时称为*求婚问题*，因为我们可以把每一方的输入解读为他 / 她是否想要与另一方结婚的声明。更确切地说，假设当且仅当爱丽丝想要嫁给鲍勃时 $a = 1$，当且仅当鲍勃想要迎娶爱丽丝时 $b = 1$。这种情形下，当且仅当双方都想与对方结婚时 $f_\wedge(a, b) = 1$；因此，如果 $b = 0$，那么鲍勃在获知函数输出结果时，并不知道关于爱丽丝的输入值信息。所以，爱丽丝的隐私得到了保护。

我们可以一般化这一示例，并思考*集合相交*问题。此处，爱丽丝和鲍勃把集合 A 和 B 作为他们的输入，其输出等于 $f_\cap(A, B) = A \cap B$。例如，假设 A 和 B 是爱丽丝和鲍勃的联系人名单的电子邮件地址集合，那么输出 $f_\cap(A, B)$ 就是他们共同拥有的联系人名单。此处的安全性意味着：(1) 各方对另一方输入掌握的信息不会超过从自身输入和输出结果中推断的信息；并且，(2) 有恶意的一方无法使结果不正确（比如，恶意的爱丽丝无法错误地使鲍勃认为某一电子邮件地址在她的联系人名单中）。在本例中，条件 (1) 意味着对于每一个 $a \notin A$，爱丽丝应当不会获得 a 是否在 B 中的信息（对鲍勃而言同样如此）。

一般结果和“公平性”缺乏。上述示例可以通过多种方式进行一般化。首先，我们可以思考在多于两个参与方的群体之间执行的协议（因此名为多方计算，而不是上例所说的双方）。比如一种除了参与者数目大于二之外，其他方面和两个参与方的硬币抛掷协议完全一样的多方硬币抛掷协议。

其次，我们可以思考复杂程度高于以上所述的函数。根据 Goldreich 等人 [14] 的文章，对于任何可有效计算的函数 f（包括硬币抛掷示例中那样的“随机化”函数），都存在可以安全执行计算的有效协议（假设存在陷门置换，而这是已被广泛接受的假设）。如果多方中有少数存在恶意（即不遵守协议），那么协议始终会终止，而所有诚实的参与方会知道输出结果。但是，如果超过半数的参与方存在恶意，那么恶意方可以在知道输出后终止协议，阻止诚实方获知结果。注意在双方协议的情形中，假定多数参与方会诚实，是没有任何意义的，因为那就表示任何一方都无恶意。此问题可见于上述硬币抛掷示例，因为每一方都可在知道另一方的比特值之后拒绝公开自己的承诺。在一些情形中，这不是问题，因为各方可以达成拒绝公开承诺等同于输掉赌局的一致意见。

然而，事实证明 [9] 这种名为缺乏*公平性*的问题通常不可避免。因此，双方协议通常不提供完全的公平性。

为何互联网上没有广泛采用 MPC？自 MPC 推出以来，已经有大量的工作提高这些协议的效用，[4, 10, 16] 有时甚至也在网上拍卖等现实应用中采用。[7] 而另一方面或许也令人诧异，许多似乎完美适合的其他领域中并未采用 MPC。一个明显的例子是互联网赌博：也许让人好奇，目前通过互联网进行的赌博几乎全都借助扮演“受信方”的网站，而不使用加密硬币抛掷协议来摒除对信任的需求。从安全性角度而言，这一局面显然令人不满，尤其是过去发生过不少这些网站运营方滥用其特权地位为自身牟利的案例。[18] 因此，剔除受信方需求的多方技术或可成为传统赌博网站的取代物。另一个好处是可以降低赌博成本，因为赌博网站通常要收取服务费用。

我们认为，至少有两大原因导致 MPC 没有运用于

网络赌博。其一，如果不满足大部分参与方诚实可信这一条件，多方协议不能提供公平性。例如，可以思考基于硬币抛掷协议的简单双方抽奖：双方先计算出随机比特值 b，如果 $b = 0$，则爱丽丝支付 1 美元给鲍勃，如果 $b = 1$，则鲍勃支付 1 美元给爱丽丝。假如协议没有正确终止的话，双方都不向对方付钱。在这种情形中，假设爱丽丝是恶意的一方，她可以在输出结果等于 0 时阻止鲍勃获知结果，从而使得 1 成为协议的唯一可能输出值。这意味着双方硬币抛掷在现实中并不安全。更广泛地说，只有大部分参与方都诚实时，多方硬币抛掷才可行，而这在完全分布式的互联网环境中并不是符合现实的假定。比如，*女巫攻击* [11] 允许恶意的一方创建并控制多个“伪造”身份，轻松获得参与方中“大多数”席位。

第二个原因更为基本，它直接源自 MPC 安全定义的固有局限：此类协议仅负责计算的安全性，而不“负责”确保用户向协议提供“真实”的输入并且尊重输出结果。

例如求婚问题：显然没有哪一种技术方法可以确保用户诚实地将其输入提供给受信方。没有手段可以阻止一方（比如鲍勃）对自己的情感说谎并设定比特值 $b = 1$ 来了解爱丽丝的输入 a。类似地，无法通过加密方式保证双方必须尊重协议结果并真正结婚。

此问题在赌博应用中尤为重要：即使在上文所述的最简单“双方抽奖”示例中，不存在任何加密方法来强制输家将钱转给赢家。

无论是数字和非数字世界，此问题存在一种务实的解决方案，它采用了“信誉”的概念：被发现作弊的一方（例如，提供错误的输入或者不尊重赌局的结果）将毁坏自己的信誉，下一次就可能难以找到愿意与其赌博的人了。有多篇论文构想和分析了信誉系统。[19] 然而，它们似乎都过于累赘，无法在许多应用中采用。其原因之一，无法确定如何去定义新用户的信誉，因为用户能够在任何时候随意选用新的名称。[12]

另一选择是利用金融交易以电子方式完成的事实。我们可以尝试将最终交易（将 1 美元从输家转给赢家）“融合”到协议之中，从而使得各方只有在此交易执行之后才知道究竟谁输谁赢。遗憾的是，没有一种显而易见的方式可以在现有电子现金系统的框架内如此操作。显然，由于各方互不信任，我们无法接受赢家知道输家信用卡号或帐户密码的解决方案。有一种可能的解决方案，即设计一种多方协议，以安全的方式模拟同时访问所有参与方的在线账户，并以他们的名义执行电子转账。即便在理论上可行，这种解决方案在现实中实施起来也非常困难，特别是该协议需要能够适应参与方使用的多家银行（并且需要在他们更改的任何时候予以更新）。

本文的主要贡献是引入一种新的范式，我们称之为“基于比特币的多方计算协议”，它为上文所说的两个问题提供了解决方案，即缺乏公平性的问题，以及“现实生活”和加密计算结果之间缺乏关联的问题。我们在第 1.1 节介绍此解决方案。

1.1. 我们的贡献

我们研究了如何进行“基于比特币的多方计算”。首先，我们演示比特币系统提供一种吸引人的方式来构建某一版本的“限时承诺”，[8, 13] 即承诺方必须在特定时限内公布其秘密值，否则将支付罚金。这进而能够被用于在特定多方协议中获得公平性。因此，它可被视为“将比特币应用于多方计算。”

更加有趣的或许是我们的第二种想法，它在某种程度上第一种想法颠倒过来，演示了“将多方计算应用于比特币”，也就是说我们推出了直接作用于比特币的多方协议概念。如上文所述，多方计算的标准定义仅保证其协议安全地执行计算，而确保输入正确并且参与方不会中断协议执行则超出了安全性定义的范畴。我们发现，比特币系统可以被用于超越这种标准定义，即构建将输入和输出与真实比特币交易关联的协议。这是可行的，因为比特币系统不存在中央权威机构，交易列表是公开的，而且其语法也允许比仅转账汇钱更为高级的交易。

作为这种想法的实现，我们构建了以利用比特币进行多方抽奖，并且不依赖受信权威方的协议。对于“抽奖”，我们的意思是这样的一种协议：一组参与方首先投入一笔钱，最后他们中被随机选出的一方获得所有投入的金钱（称为*赌资*）。我们的协议可以在纯粹的 P2P 环境中使用，并且可在匿名且互不信任的参与方之间执行。我们的构想附有非常强的安全保证：无论不诚实的参与方行为怎样，诚实的参与方都不会受骗。更确切地说，每一个诚实的参与方可以确信，赌局一旦开始，它始终会结束并且公平。

我们的主要构想将在第 4 节中阐述。其安全性通过*押*金来获得：每一用户需要先存上一定金额的钱，只要诚实地完成协议，这笔钱就会归还。否则，押金就会划给其他参与方，作为赌局提前结束的“补偿”。此协议使用了上文所述的限时承诺方案。此协议有一个缺点，押金的数额需要相对较大，尤其是协议在数量较多的参与方之间执行时。更确切地说，要达到安全性，每一参与方的押金需要是赌注的 $N(N-1)$ 倍，其中 N 是参与方数目。对于两方情形，这意味着押金是赌注的两倍。

在我们的协议中，参与方需要付出的成本是比特币交易费。大部分比特币交易目前是免费的。不过，协议的参与方需要进行少量的非标准交易（所谓的“奇

怪交易”，见第 2 节），通常存在一点费用（目前大约 0.0001 ฿ ≈ 0.04 美元）。[b] 为了让阐述简洁明了，我们在陈述结果时假设这些费用为零。为简单起见，我们也假定抽奖中的赌注等于 1 ฿。如何将我们的协议一般化到其他赌注值，这应当很明了。

我们的构想基于前文所述的硬币抛掷协议。我们设法将此协议改编为我们的模型，而无需修改当前的比特币系统。我们没有使用 MPC 或零知识编译器等任何通用方法，所以我们的协议非常高效。我们唯一使用的加密原语是承诺方案，利用哈希函数实施（标准的比特币原语）。我们的协议十分依赖比特币的高级功能（尤其是，所谓的“交易脚本”和“时间锁”）。由于篇幅的关系，我们只是粗略介绍正式的安全性定义。我们在实际的比特币系统中执行了交易。我们将提供这些交易的描述以及它们在比特币区块链中的引用。[c]

1.2. 独立和后续研究

运用比特币创建安全而公平的双方抽奖曾经由 Back 和 Bentov 独立提出。[3] 我们在本论文的加长版本中提供了他们的协议和我们的协议之间的详细对比。

在后续的研究中，[1,2] 我们演示了如何拓展本文中的思路，为任何功能性构想公平的双方协议，从而使本协议的执行具有“金融上的结果”。更确切地说，在第一篇论文中，[1] 我们演示了如何在假设比特币交易为非延展的前提下解决此问题（关于此注解的更多信息，请参见 Andrychowicz 等人的文章 [1,2]），而在 Andrychowicz 等人的文章 [2] 中我们演示了如何修改 Andrychowicz 等人 [1] 的协议来获得在当前版本比特币系统中安全的协议。在多方协议中获得公平性的一些其他方案由 Bentov 和 Kumaresan 独立发展而来。[5,15]

1.3. 应用和未来研究

尽管如本论文加长版中所述，在实践中运用我们的协议或许能有真正的经济意义，但我们大体上把赌博视为引入可称为“基于比特币的多方计算”概念的一个激励人心的例子，它也将会出现其他方面的应用。比如，可以运用我们的技术来实施的一个任务便是用于销售秘密信息来赚取比特币的协议。假设爱丽丝和鲍勃知道集合 X 的描述，它包括一些有价值的信息。例如，X 可以包含某些难以寻踪的敏感数据（譬如：通过某一公共权威密钥签名的个人数据。）爱丽丝知道 X 的子集 A，鲍勃知道 X 的子集 B。他们的目标是采用一种方法来互相出售 $A \cup B$ 中的元素，使得他们仅支付给对方自己事先不知道的元素的费用。换而言之，爱丽丝将支付给鲍勃 $(|B \setminus A| - |A \setminus B|)$ ฿（如果此值为负数，则鲍勃将支付给爱丽丝其取相反值）。如果没有 MPC 方法，就不清楚如何进行此操作：每当爱丽丝鲍勃透露某元素 $a \in A$，鲍勃始终都可声称他已知道 a。此外，即使采用了 MPC 方法，爱丽丝也无法强制让鲍勃付钱（反之亦然）。我们的工具（于第 1.2 节中所提后续论文中开发）解决了此问题：我们可以设计一种协议，使它转账给对方数额正好的比特币，而且当且仅当双方确实获知计算的输出结果时这才会发生！

上述示例可以通过多种方式进行一般化。例如，仅有一方（比如是爱丽丝）有输出结果，而爱丽丝愿意为其付钱的信息可能有非常复杂的条件。例如，爱丽丝可能是情报机构，通过一个特殊秘密函数 g 来指定给定信息的价值（对于某些输入集合，g 甚至可以输出 0）。接着，鲍勃可以试图将其信息 x“销售”给爱丽丝，同时设定自己认为值得的某一最小价值 v。该协议将计算 $g(x)$，并检查是否 $g(x) \geq v$ — 若是，则爱丽丝获知 x 并支付 v 给鲍勃，否则爱丽丝不获得任何信息（鲍勃则赚到 0）。

最后，要补充的一点是我们的协议有可能被用于恶意用途。例如，某一勒索软件对受害者的电脑硬盘进行加密，并且承诺只有在受害者支付赎金后才会提供解密密钥。目前，这样的恶意程序没有办法证明他们确实会在赎金支付后发送正确的密钥。利用我们的技术时，人们可以安全地交付此类密钥（也就是只有密钥确实解密了磁盘后才会发生付款）。另一潜在风险是对网上投票方案的攻击：众所周知，如果这些方案中包含了收据，对手可以购买选票。我们的技术可以让此类攻击更加容易，因为它们消除了选票卖家信任选票买家的需求。

2. 比特币概述

比特币 [17] 以 P2P 网络的方式运作，参与方在其中共同模拟中央服务器来控制交易正确性。在这个意义上，它类似于 MPC 协议的概念。回想上文所述，传统 MPC 存在一个基本问题，如果不满足大部分参与方诚实可信这一条件，它们就无法提供公平性，而这在有可能存在女巫攻击的 P2P 网络中特别难以保证。比特币系统通过以下方式攻克了这一问题：诚实多数的定义使用“多数计算能力”这一形式。换句话说，若要攻破该系统，攻击者需要控制的计算机的总计算能力要能与协议所有其他参与方的计算能力之和相匹敌。所以，女巫攻击无法奏效，因为在网络中创建大量伪造身份并不有助于攻击者。我们稍后会阐述如何实施这一点，但首先来说明由用户模拟的受信方的功用。

[b] 我们使用“฿”，作为比特币货币符号。

[c] 比如，三方抽奖的主要交易 (*Compute*) 位于此处：blockchain.info/tx/540d816bd57300209754dd36ffcec1d669-bd2068641844783451cd3ef-32c8aa4

数字货币的一个主要问题是可能存在重复消费：如果钱币仅仅是几串比特值，那么所有者就能多次花这些钱币。显然，如果用户可以访问列出所有交易的可信账本，此风险就能规避。在这一情形中，只有交易记录到账本，它才被视为有效。比如，假设交易的形式为“用户 A 将 x 比特币转给用户 B”。此时，每一用户可以验证 A 是否真的持有 x 比特币（比如在某些过往交易中收到的），而且还没有花掉。比特币网络模拟的受信方功能精确做到了这一点：它维护了一份系统中发生过的交易的完整列表。实际上，比特币交易的形式要比上述例子复杂得多。由于它对我们有着重要的意义，我们会在第 2.1 节中详细地阐述。但为了简单起见，我们省略了与我们研究无关的比特币功能，如交易费或如何造币。

比特币账本实际上是一种*区块链*（每个区块中包含多个交易），所有参与方都在尝试扩展这个链。系统的参数进行了特别的选择，使得系统平均每十分钟进行一次扩展。区块链的思路是，最长的链 C 被视为正确的链，向该链附加新的区块需要进行复杂的计算。因为扩展该区块链或者创建新链非常困难，所有用户将使用相同的原始区块链。更详细地说，这种构造可防止交易的重复消费。如果某一交易包含在区块 B_i 中，并且其后有多个新区块，那么攻击者若不具有比特币网络总计算能力一半的计算能力，就无法将它撤销 —他必须要开采新链 C'（从位于区块 B_{i-1}（或更早）的 C 分叉），并且 C' 必须比 C 长。其难度根据 B_i 上新区块数量呈指数级增长。在实践中，交易经过 10–20 分钟后可以得到合理的有力确认，而且 60 分钟（6 个区块）后几乎就能确信它们不可撤销。

总而言之，当用户希望使用比特币向别人付款时，他将创建一个交易并广播到网络上的其他节点。它们将验证此交易，将它进一步发送出去并添加到它们挖矿的区块中。当某一节点求解挖矿问题后，它将自己的区块广播到网络上。节点获得新区块，验证其中的交易及其哈希值，然后通过在其上挖矿予以接受。区块中存在交易便是对此交易的确认，但一些用户可以选择等待多个区块，从而获得更多保障。我们的协议中假定交易的广播和确认之间存在延迟 $T_{\max}$，并且每个交易在确认之后便不能撤销。

2.1. 比特币交易

与传统的银行系统相反，比特币基于*交易*，而非*账户*。如果用户 A 在系统中有他是接收方、并且尚未兑现的交易，则该用户持有一些比特币。每笔交易具有某一价值（转账的比特币数量），以及一个接收方*地址*。地址就是公钥 pk。通常，这种公钥分别都有对应的私钥 sk，其仅为一个用户所知，该用户便是地址 pk 的所有者。私钥用于为交易签名（授权），公钥则用于验证签名。系统的每个用户需要知道某地址的至少一个私钥，但这很容易实现，因为密钥对 (sk, pk) 可以轻松地离线生成。我们时常使用大写字母（如 A）来表示密钥对，并将 A 的私钥和公钥分别称为 $A.sk$ 和 $A.pk$。

简化版本。我们首先说明比特币的简化版本，然后再演示如何将它扩展，从而获得真实比特币的说明。假设 $(A.sk, A.pk)$ 和 $(B.sk, B.pk)$ 是分别属于用户 A 和用户 B 的密钥对。从我们简化后的角度来看，描述金额 v（称为交易的*价值*）从地址 $A.pk$ 转到地址 $B.pk$ 这一事实的交易表示为 $T_x = (y, v, B.pk, \text{sig})$，其中 y 是上一交易 T_y 的索引，sig 是利用发送方的密钥 $A.sk$ 对除签名本身外整个交易（即 $(y, v, B.pk)$）计算而得的签名。我们假设 $B.pk$ 是 T_x 的接收方，而交易 T_y 是交易 T_x 的*输入值*或者说 T_y 已由 T_x *兑现*。更确切地说，T_x 的含义是交易 T_y 中转给 $A.pk$ 的 v 数额的钱又进一步转给了 $B.pk$。只有满足以下条件时 T_x 才有效：(1) $A.pk$ 是交易 T_y 的接收方，(2) T_y 的价值等于 v，(3) 交易 T_y 之前没有兑现过，并且 (4) 签名 A 正确无误。所有这些条件都可公开检验。

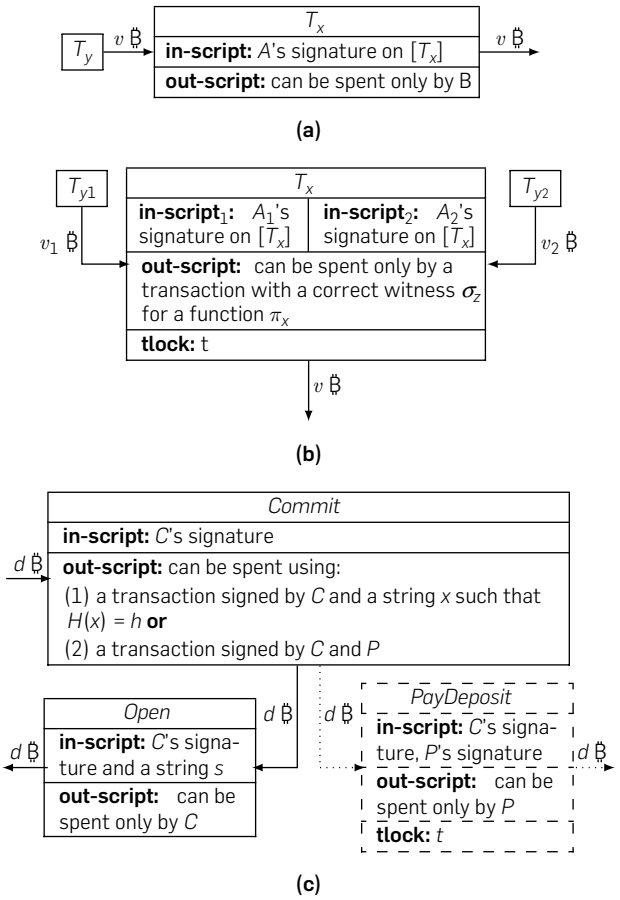
我们将以方框的形式来演示交易。交易的兑现使用注有交易价值的箭头来表示。例如，交易 $T_x = (y, v, B.pk, \text{sig})$，它将 v 比特币从 A 转给 B，如图 1(a) 所示。

此简化系统的第一个重要一般化是交易可以具有多个“输入”，也就是它可以累计来自多个过往交易 $T_{y1}, \ldots, T_y$ 的钱。假设 $A_1, \ldots, A$ 分别是这些交易的接收方密钥对。那么，多输入交易具有下列公式：$T_x = (y_1, \ldots, y, v, B.pk, \text{sig}_1, \ldots, \text{sig})$，其中每个 sig_i 是使用密钥 $A_i.sk$ 对除签名外整个消息计算而来的签名。此交易的结果是 $B.pk$ 获得金额 v，只要它等于交易 $T_{y1}, \ldots, T_y$ 的价值之和。而这只有在这些交易*都未*兑现过，并且*所有*签名全部有效时才会发生。每笔交易还可以有多个输出，这被用于在多名用户间分钱或者换零钱，但我们的协议中不使用此功能。

更加详细的版本。实际比特币系统的复杂度要远远高于以上所述。首先，存在一些语法上的区别，其中对我们最为重要的是每个交易 T_x 的标识使用的不是其索引，而是整个交易的哈希 $H(T_x)$。所以从现在起，我们假定 $x = H(T_x)$。此外，每个交易有一个*时间锁*，它告知该交易在哪一时刻生效。此时，我们便有了：$T_x = (y_1, \ldots, y, v, B.pk, t, \text{sig}_1, \ldots, \text{sig})$。只有到了时刻 t，并且前文中的所有条件均满足后，这一个交易才会生效。在时刻 t 之前，无法使用交易 T_x（它在时刻 t 之前不会包含到任何区块中）。

但主要区别在于，实际比特币中用户在定义如何兑现交易的条件上享有大得多的灵活性。我们暂时考

图 1：(a) 将 v 比特币从 *A* 转给 *B* 的标准交易，(b) 具有两个输入和一个时间锁的非标准交易，(c) CS 协议。



虑最简单的交易，它只有一个输入，而且没有时间锁。回想一下上文所述的简化系统，为了要兑现一个交易，接收方 $A.pk$ 必须生成另一个交易 T_x 并使用自己的私钥 $A.sk$ 进行签名。在实际的比特币中，这按照以下方式一般化：每个交易 T_y 附带一个函数（称为 *output-script*）π_y 的说明，其输出为布尔值。如果 π_y 对输入 T_x 计算结果真，则交易 T_x 兑现交易 T_y 有效。在标准交易的情形中，π_y 是将 T_x 视为值对（消息 m_x，签名 σ_x）的函数，根据公钥 $A.pk$ 检查 σ_x 是否为 m_x 上的有效签名。不过，也可能有更广义的函数 π_y。进一步探究细节，交易将是以下形式：$T_x = (y, \pi_x, v, \sigma_x)$，其中 $[T_x] = (y, \pi_x, v)$ 称为 T_x 的*正文*，而 σ_x 则是 *input-script* — 用于使脚本 π_y 对 T_x 计算结果为真的见证（在标准的交易中，σ_x 是 $[T_x]$ 上发送方的签名）。脚本采用比特币脚本语言编写而成，它是基于堆栈的非图灵完备语言（其中没有循环）。它提供基本的数字算术运算、堆栈运算、if-then-else 语句和一些加密函数，如哈希计算函数或签名验证。一般化到具有时间锁的多输入交易比较明了：交易具有公式 $T_x = (y_1, \ldots, y, \pi_x, v, t, \sigma_1, \ldots, \sigma)$，其中正文 $[T_x]$ 等于 $(y_1, \ldots, y, \pi_x, v, t)$，并且它只有满足以下条件才有效：(1) 达到时刻 t；(2) 每个 $\pi_i([T_x], \sigma_i)$ 计算结果为真，其中每个 π_i 是交易 T_{yi} 的输出；(3) 这些交易全都没有兑现过；并且，T_{yi} 等于 v。

图 1(b) 通过方框表示具有两个输入的普通交易 $T_x = (y_1, y_2, \pi_x, v, t, \sigma_1, \sigma_2)$。

最常见的交易类型是没有时间锁或任何特殊脚本的交易：输入脚本为签名，输出脚本则为签名验证算法。我们把它们称为*标准交易*，而执行验证时被检验的地址则称为交易的*接收方*。目前，一些矿主仅接受标准交易（虽然根据比特币说明，非标准交易也是正确的）。我们认为，接受非标准交易在未来也会变得常见。这对我们的应用而言很重要，因为我们的协议严重依赖于非标准交易。

3. 基于比特币的限时承诺方案

我们首先构想基于比特币的限时承诺方案。承诺方案之前已在第 1 节中介绍过。实施承诺的一种简单方式是使用加密哈希函数 H。为了承诺秘密 $s \in \{0,1\}^*$，承诺方选择随机字符串 $r \in \{0,1\}^{128}$ 并向接收方发送 $c = H(s\,r)$，其中 “ ” 表示字符串串联操作。要公开承诺，承诺方发送 (s, x)，而接收方验证 $H(s\,x) = c$。

标准承诺方案虽然在许多应用中都能大显身手，但也存在下列问题（引言中已有介绍）：无法强制承诺方公布其秘密 s，特别是如果他在公布阶段之前中止，则 s 仍会保持秘密。比特币系统为应对这一问题提供了引人注目的方法。即，我们可以利用比特币系统强制承诺方通过一笔钱（*押金*）来担保其承诺，如果他们拒绝在双方协定的某一时限 t 内公开承诺，这笔钱将付给接收方。更确切地说，承诺方在承诺阶段中存一笔押金到比特币系统中。如果在时刻 t 之前公开承诺，他将能够取回这笔押金。否则，押金将自动付给接收方。

3.1. 构想

我们基于比特币限时承诺方案 (CS) 的构想将以上文所述的简单承诺方案为基础。比特币中使用的哈希函数为 SHA256，我们的协议中也会用它，因为比特币脚本语言中可以采用。但为清晰起见，我们仍然在协议的描述中使用 H 表示它。此外，我们假设秘密已经填上了随机比特值，所以我们的描述中不会添加或剥离它们。实际上，我们稍后会使用 CS 协议来承诺长随机字符串，这种情况下不需要填补。

我们协议的基本思路如下所述。在承诺阶段，承诺方创建交易 *Commit*，它具有某一约定的价值 d，充当押金。兑现该押金的唯一途径是发布另一个交

易 *Open*，它将公布秘密 s。交易 *Commit* 经过了特别构造，使得 *Open* 交易必须要公开承诺（公布秘密值 s）。这意味着承诺方的钱将被“冻结”，直到他公布 s 为止。为了让接收方能够在承诺方未于特定时限内公开承诺时认领押金，我们也需要承诺方向接收方发送交易 *PayDeposit*，该交易可在时间 t 过后兑现 *Commit* 交易。

技术上而言，其实现可通过构造交易 *Commit* 的输出脚本，使得兑现交易必须要提供 C 的签名以及秘密 s（这将使它公开，因为所有交易都对公众可见）或者来自 C 和 R 的签名。在广播交易 *Commit* 后，承诺方创建交易 *PayDeposit*，该交易将押金发送给接收方并带有时间锁 t。承诺方对其签名，并发送给接收方。在收到 *PayDeposit* 后，接收方检查它是否正确并加上自己的签名。此后，他可以确信承诺方将在时刻 t 前公开其承诺，否则他可以使用交易 *PayDeposit* 来认领 d ฿ 押金。

图 1(c) 中通过图表描绘了此协议中的交易。该协议的完整描述可在本文的加长版中找到。

4. 抽奖协议

如第 1 节中所探讨的，作为“基于比特币的多方计算”概念的一个应用示例，我们为两个参与者之间执行的抽奖构建了一个协议，这两方为爱丽丝 (A) 和 鲍勃 (B)。如果协议*正确*且*安全*，则我们认为该协议是*公平的抽奖协议*。

为定义正确性，假设双方都遵守该协议并且两者之间的沟通渠道是安全的（即它能可靠地在双方之间传递消息，且没有延迟）。我们也假设在协议启动之前，双方都有参与抽奖的充足资金，包括他们的赌注（为简单起见，假设赌注等于 1 ฿）以及用作押金的钱，因为在协议中我们将使用第 3 节中所说的承诺方案。如果这些假设都成立，正确的协议必须确保协议结束时一方（通过均匀概率选择）必须获得包含双方赌注的所有赌资，另一方则失去其赌注。此外，双方也必须能收回自己的押金。

为定义安全性，我们从一方的视角来看待协议的执行，譬如 A（对于另一方是对称的），并假设她是诚实的。显然，A 无法确保该协议将成功终止，因为另一方可以在完成之前退出协议。重要的是 A 应当能够确保她不会因为协议终止而丢钱，比如，应当不允许另一方在获悉 A 是赢家后终止协议。按照如下所述对安全性进行形式化：我们将 A 在执行该协议过程中的*补偿*定义为等于 A 投入的钱和执行协议后拥有的钱之间的差值。如果对于任何控制网络并使一方变得恶意的攻击者的策略，另一方（诚实方）的补偿不是负值，那么我们可以说该协议是安全的。当然，我们也要注意，失信方始终都能在尚不清楚谁是赢家时提前终止协议，而这不会改变诚实方的补偿。

4.1. 协议

我们协议的构建基础是 Blum[6] 的经典硬币抛掷协议（见第 1 节）。如前文所述，此协议并不直接适用于我们的应用，所以我们需要使它适应比特币。特别是，在我们的解决方案中，创建和公开承诺是借助使用（双）SHA-256 哈希的交易的脚本执行的。在选择随机比特值 b_P 后，参与方 $P \in \{A, B\}$ 选择从均匀随机的 $\{0,1\}^{128+bP}$ 中取样而来的字符串 s_P，后者是长度为 128 或 129 位的字符串集合，具体根据 b_P 的值。参与方 P 而后使用限时承诺发出承诺 s_P。赢家由*赢家选取函数* f 决定，其定义如下：如果 $|s_A| = |s_B|$，$f(s_A, s_B) = A$，否则等于 B；其中，s_A 和 s_B 是双方选择的秘密字符串，而 $|s_P|$ 则是以比特表示的 s_P 的长度。显然，只要其中一方均匀选取其比特值 b_P，$f(s_A, s_B)$ 的输出也会是均匀随机的（条件是双方选择的字符串 s_A 和 s_B 的长度为 128 或 129）。

第一次尝试。我们首先展示不成熟且不安全的协议构想，然后演示如何进行修改来获得安全方案。双方互相公布其公钥。如前文所述，爱丽丝和鲍勃也（分别）随机抽取其秘密字符串 s_A 和 s_B，并且交换哈希 $h_A = H(s_A)$ 和 $h_B = H(s_B)$。如果 $h_A = h_B$，则参与方中止协议。[d] 双方广播其输入交易，并向对方发送它们在区块链中位置的链接。如果在任一点上，一方 $P \in \{A, B\}$ 意识到另一方在作弊，那么 P 要做的第一件事就是“拿钱走人”，即发布去兑现输入交易的交易。我们将它称为“停止执行”。这显然可以做到，只要输入交易尚未被某一其他交易兑现。在下一步中，其中一方构造交易 *Compute*，其定义如下：

Compute	
in-script$_1$: *A*'s signature	**in-script**$_2$: *B*'s signature
out-script: can be spent using: (1) strings x_A and x_B of length 128 or 129 s.t. $H(x_A) = h_A$, $H(x_B) = h_B$ **and** (2) *X*'s signature, where *X* is the winner (i.e., $X = f(x_A, x_B)$)	

请注意，*Compute* 的正文可以从公开提供的信息计算而得。因此，此构造可以按照以下方式实施：首先，其中一方（比如鲍勃）计算出 *Compute* 的正文并将自己对它的签名发送给爱丽丝。爱丽丝计算其正文，再将双方的签名添加到其中，然后广播完整的交易 *Compute*。

[d] 我们向 Iddo Bentov 和 Ranjit Kumaresan 以及独立研究的 David Wagner 致以诚挚谢意，感谢他们为我们指出此步骤的必要性。它可以抵御*拷贝攻击*：A 等待 B 提交其哈希 h_B，然后自己提交同样的哈希。在公开阶段，A 再次等待 B 公布其秘密 s_B，然后自己公布同样的秘密。这样，A 始终都能赢，因为 $f(s_A, s_B) = A$。

Compute 的输出脚本有些复杂。为了使它对 *body* 评估为真，我们需要提供作为“见证”的参与方 P 的签名以及字符串 x_A 和 x_B，其中 x_A 和 x_B 是 h_A 和 h_B 的前像（对于 H）。H 的抗碰撞性意味着 x_A 和 x_B 必须分别等于 s_A 和 s_B。因此，只有赢家选取函数 f 对输入 (s_A, s_B) 计算得到 P 时，才可满足条件。由于只有参与方 P 知道自己的私钥，只有她才能在稍后提供可将输出脚本计算为真的签名。

在 *Compute* 现身于区块链之前，每个参与方 P 都可“改变主意”并兑现自己的输入交易，而这会使交易 *Compute* 无效。如之前所说，其中一方中断硬币抛掷程序对我们来说也没关系，只是她必须在获知自己是输家之*前*做出这个决定。因此，爱丽丝和鲍勃将等待交易 *Compute* 变为确认状态，然后再继续决定赢家的步骤。最后一步很简单：爱丽丝和鲍勃只需分别广播 s_A 和 s_B。现在：如果 $f(s_A, s_B) = \mathsf{A}$，那么爱丽丝可以通过交易 $ClaimMoney_{\mathrm{A}}$ 兑现交易 *Compute*，其构造为：

$ClaimMoney_A$
in-script: strings s_A and s_B and *A*'s signature
out-script: can be spent only by *A*

另一方面，鲍勃无法兑现 *Compute*，因为条件 $f(s_A, s_B) = \mathsf{B}$ 计算结果为假。反过来：如果 $f(s_A, s_B) = \mathsf{B}$，则只有鲍勃可以通过类似的交易 $ClaimMoney_B$ 兑现交易 *Compute*。

此协议显然是正确的。它看上去也可能安全，因为它基本上与前文所述的 Blum 协议一致（将哈希函数用作承诺方案）。可惜它存在下列问题：无法保证参与方始终都会公布 s_A 和 s_B。尤其是，一方（比如鲍勃）可以在知道自己是输家（即 $f(s_A, s_B) = \mathsf{A}$）*之后*拒绝发送 s_B。由于他的钱已经“消失了”（其输入交易已在交易 *Compute* 中被兑现），他得不到任何东西，但他可以纯粹因为耍赖而这么做。遗憾的是，纯 P2P 环境中没有“信誉”之说，此类行为可以发生，而且没有惩罚的办法。这也正是我们需要使用基于比特币的承诺方案（见第 3 节）的原因所在。

方案的安全版本。支撑 SecureLottery 协议的一般想法是，每一方首先对其输入承诺，而且要利用基于比特币的限时承诺方案，而不是标准的承诺方案。回想之前所说，CS 协议可通过发送值 s 公布，而这种公布又通过检查 s 具有要求的长度（128 或 129）和承诺方在承诺阶段发送的值 h 的哈希加以验证。因此，爱丽丝作为承诺方执行 CS 协议，而鲍勃则充当接收方。假设 s_A 和 h_A 是如此创建的变量 s 和 h。与之相对称，鲍勃作为承诺方执行 CS 协议，爱丽丝则充当接收方，其对应的变量为 s_B 和 h_B。一旦两个承诺阶段都成功执行（这包括每一方接收经过签名的 *PayDeposit* 交易），双方便继续后续步骤，这和前文所述完全相同：首先，每一方广播输入交易。这些交易确认之后，他们像前面一样创建 *Compute* 交易，此交易现身于区块链时，他们便公开承诺。唯一的区别是，由于使用了 CS 承诺方案，他们现在可以“惩罚”对方（若对方没有在时刻 t 之前公开其承诺）并领走其押金。另一方面，诚实的一方始终都确保可领回其押金，所以在协议之初投入这一笔钱没有任何风险。图 2 中通过图表演示了此协议中的交易。

我们还需要说明一下参数的选择：t，押金可被接收方领取的时间，以及 d，押金的价值。我们的协议由四个“回合”的交易组成 — 在每一回合中，各方等待确认此回合中所有交易，然后进入下一回合。因此，正确执行协议始终在时限 $4 \cdot T_{\max}$ 内终止，其中 $T_{\max}$ 是交易确认所需的最长时间。因此，我们可以安全地将 t 设置为协议启动时间加上 $5 \cdot T_{\max}$。

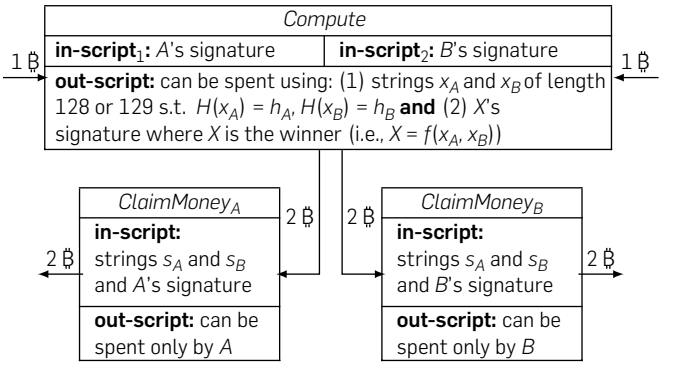
参数 d 的选择应当要确保它能够彻底补偿各方因另一方中途退出而招致的损失。对于包含两个参与方的抽奖，这意味着每一参与方应当存入等于赌注两倍的押金。这样，如果一方中止协议，那么另一方将失去价值 1 ฿ 的赌注，但获得价值 2 ฿ 的押金，所以协议执行的结果是她赚了 1 ฿，对她而言这并不比协议执行到最终时的情况差。

此协议的完整描述可在本文的加长版中找到，文中我们还会展示如何将它一般化到 N 个参与者的情形。在我们的多方解决方案中，每个参与方在押金中投入的钱款总数必须等于 $N(N-1)$ ฿。在现实中，对于小群体 $N = 2, 3$ 而言这或许可行，但更大的群体则不行。

鸣谢

感谢 Iddo Bentov 和 Ranjit Kumaresan 富有成果的讨论，也感谢他们指出我们上一版本抽奖描述中的错误。同时也感谢 David Wagner 仔细阅读我们的文章并提供了宝贵意见。 C

图 2：SecureLottery 协议。



参考资料

1. Andrychowicz, M., Dziembowski, S., Malinowski, D., Mazurek, Ł. Fair two-party computations via bitcoin deposits.In *1st Workshop on Bitcoin Research* (Christ Church, Barbados, March 7, 2014), Springer, Berlin, Germany, 105–121.
2. Andrychowicz, M., Dziembowski, S., Malinowski, D., Mazurek, Ł. On the malleability of bitcoin transactions. In *2nd Workshop on Bitcoin* Research (San Juan, Puerto Rico, January 30, 2015), Springer, Berlin, Germany.
3. Back, A., Bentov, I. Note on fair coin toss via bitcoin, 2013. http://www.cs.technion.ac.il/~idddo/cointossBitcoin.pdf.
4. Ben-David, A., Nisan, N., Pinkas, B. FairplayMP:A system for secure multi-party computation.In *ACM CCS 08:15th Conference on Computer and Communications Security* (Alexandria, VA, October 27–31, 2008), ACM, NY, 257–266.
5. Bentov, I., Kumaresan, R. How to use bitcoin to design fair protocols. In *Advances in Cryptology – CRYPTO*, 2014.Part II (Santa Barbara, CA, August 17–21, 2014), Springer, Berlin, Germany, 421–439.
6. Blum, M. Coin flipping by telephone.In *Advances in Cryptology – CRYPTO'81* (Santa Barbara, CA, 1981), U.C. Santa Barbara, Department of Electrical and Computer Engineering, 11–15.
7. Bogetoft, P., et al. Secure multiparty computation goes live.In *FC 2009:13th International Conference on Financial Cryptography and Data Security* (Accra Beach, Barbados, February 23–26, 2009), Springer, Berlin, Germany, 325–343.
8. Boneh, D., Naor, M. Timed commitments.In *Advances in Cryptology – CRYPTO 2000* (Santa Barbara, CA, August 20–24, 2000), Springer, Berlin, Germany, 236–254.
9. Cleve, R. Limits on the security of coin flips when half the processors are faulty.In *Proceedings of the 18th Annual ACM Symposium on Theory of Computing*, STOC '86 (Berkeley, CA, May 28–30, 1986), ACM, NY, 364–369.
10. Damgård, I., et al. Practical covertly secure MPC for dishonest majority — Or:Breaking the SPDZ limits. In *ESORICS 2013:18th European Symposium on Research in Computer Security* (Egham, UK, September 9–13, 2013), Springer, Berlin, Germany, 1–18.
11. Douceur, J.R. The sybil attack.In *First International Workshop on Peer-to-Peer Systems*, IPTPS '01, 2002.
12. Friedman, E.J., Resnick, P. The social cost of cheap pseudonyms.*J. Econ. Manage.Strat.10* (2000), 173–199.
13. Garay, J.A., Jakobsson, M. Timed release of standard digital signatures. In *FC 2002:6th International Conference on Financial Cryptography* (Southampton, Bermuda, March 11–14, 2003), Springer, Berlin, Germany, 168–182.
14. Goldreich, O., Micali, S., Wigderson, A. How to play any mental game or A completeness theorem for protocols with honest majority.In *19th Annual ACM Symposium on Theory of* Computing (New York City, NY, May 25–27, 1987), ACM, NY, 218–229.
15. Kumaresan, R., Bentov, I. How to use bitcoin to incentivize correct computations.In *ACM CCS 2014* (Scottsdale, AZ, November 3–7, 2014), ACM, NY, 30–41.
16. Malkhi, D., Nisan, N., Pinkas, B., Sella, Y. Fairplay – A secure two-party computation system.In *13th Conference on USENIX Security Symposium*, SSYM'04 (San Diego, CA, August 9–13, 2004), USENIX Association, 287–302.
17. Nakamoto, S. Bitcoin:A peer-to-peer electronic cash system. The Cryptography Mailing List, 2008.
18. Post, T.W. Cheating scandals raise new questions about honesty, security of internet gambling.*The Washington Post* November 30, 2008.
19. Resnick, P., Kuwabara, K., Zeckhauser, R., Friedman, E. Reputation systems. *Commun.ACM 43*, 12 (Dec. 2000) 45–48
20. Yao, A.C.-C. How to generate and exchange secrets (extended abstract). In *27th Annual Symposium on Foundations of Computer Science* (Toronto, ON, Canada, October 27–29, 1986), IEEE Computer Society Press, 162–167.

Marcin Andrychowicz、Stefan Dziembowski、Daniel Malinowski 和 Łukasz Mazurek ([marcin.andrychowicz, stefan.dziembowski, daniel.malinowski, lukasz.mazurek]@crypto.edu.pl) 来自位于波兰华沙的华沙大学信息学院。

译文责任编辑：陈海波

观看此独家《通讯》视频中作者对本研究的讨论：http://cacm.acm.org/videos/secure-multiparty-computations-on-bitcoin