基于风险的认证系统的隐私考量

Stephan Wiefling*、Jan Tolsdorf 和 Luigi Lo Iacono*H-BRS*应用科学大学，德国圣奥古斯丁
*波鸿鲁尔大学，德国波鸿{
stephan.wiefling,jan.tolsdorf,luigi.lo iacono}@h-brs.de

arXiv:2301.01505v1 [cs.CR] 4 Jan 2023

摘要——**基于风险的身份验证（RBA）扩展了认证机制，使其能够更有效地防范账户接管攻击，例如利用被盗密码实施的攻击。NIST 和 NCSC 建议采用基于风险的身份验证来强化基于密码的认证，许多主流在线服务也已在使用这种方式。此外，用户认为基于风险的身份验证比双因素认证更易于使用，且安全性不逊于后者。然而，目前用户在享受基于风险的身份验证带来的高安全性和高可用性优势时，需要付出暴露潜在敏感个人数据（例如 IP地址或浏览器信息）的代价。这与用户隐私相冲突，因此必须考虑用户在处理个人数据方面的权利。**

我们概述了针对不同攻击者模型的潜在隐私挑战，并提出了在 RBA 系统中平衡隐私的改进措施。为评估保护隐私的基于风险认证增强措施在实际环境中的特性，我们使用某真实在线服务的 780 名用户的长期数据，对其中一部分措施进行了评估。结果表明，基于风险的认证解决方案有望提升隐私保护水平。不过，这种提升仅限于某些参数，因此在设计基于风险的认证时，应以这些参数为指导来保护隐私。我们还概述了需要考虑的研究方向，以推动保护隐私的基于风险的认证得到广泛采用，并获得较高的用户接受度。

索引词—**密码、基于风险的身份验证、可用安全与隐私、大数据分析**

1. 引言

尽管新威胁不断出现，密码仍是在线服务认证的主要方式 [25], 。凭证填充和密码喷洒攻击 [14]使用从数据泄露事件中获取的登录凭据（用户名和密码），并以某种方式尝试将其用于其他在线服务。这些攻击如今十分普遍 [2] ，因为攻击者只需付出很少努力便可将其自动化。为应对这一威胁，主要在线服务采用了基于风险的身份验证（RBA） [36], ，旨在加强基于密码的认证，同时尽量减少对用户的影响。

基于风险的身份验证（RBA）。RBA 用于判断登录尝试是合法操作还是账户接管尝试。为此，RBA 会在用户提交登录凭据时监控其他特征。常见特征包括网络信息（例如 IP地址）、设备信息（例如智能手机型号和操作系统）、客户端信息（例如浏览器供应商及其版本），以及（行为）生物特征信息（例如登录时间） [34], [36]。RBA 会根据特征值及以往登录的特征值计算风险评分。访问阈值通常将评分划分为低、中、高风险 [12], [15], [21]。风险较低时（例如使用常用设备并位于常见位置），RBA 系统无需进一步干预即可授予访问权限。风险中等或较高时（例如使用陌生设备并位于陌生位置），RBA 会要求用户提供其他信息，例如验证电子邮件地址。提供正确的证明后，即可获得访问权限。

在许多情况下，如果无法轻易用更安全的身份验证方法取代密码，基于风险的身份验证（RBA）便被视为一种可扩展的过渡性解决方案 [34], [35]。美国国家标准与技术研究院（NIST，美国）和国家网络安全中心（NCSC，英国）建议使用基于风险的身份验证来缓解涉及密码被盗的攻击 [13], [23]。除此之外，用户认为基于风险的身份验证比同等的双因素认证（2FA）方案更易于使用，且安全性相当 [35]。此外，与双因素认证不同，基于风险的身份验证既能提供良好的安全性，在实际使用中也很少要求额外身份验证 [34], ，从而减轻了用户的负担。

研究问题。 然而，用户获得基于风险的身份验证在安全性和可用性方面的提升，却要以披露更多可能敏感且可关联到个人的数据为代价，例如 IP 地址和浏览器标识符。因此，如果基于风险的身份验证数据库遭到转发或泄露，用户隐私便会面临风险，因为除用户名之外的额外数据可能会使个人身份得以识别。

越来越多的数据保护法旨在保护用户免受在线服务大规模数据收集的影响。鉴于此，我们想知道，集成 RBA 系统是否符合现代数据保护原则，以及在多大程度上符合这些原则。我们还想知道，可以采取哪些权衡措施来平衡安全性和隐私目标。

为进一步探讨基于风险的身份验证的隐私方面，我们提出了以下研究问题：

研究问题1：a) 基于风险的认证特征可以通过哪些方式存储，以提高用户隐私？**b)** 如何存储基于风险的认证特征，才能在发生数据泄露时保护用户隐私？

研究问题2： 在实际应用中，基于风险的认证特征在保护隐私的同时，能够在多大程度上维持良好的安全性？

贡献。 我们提出并讨论了五种隐私增强措施，可供实际部署中大多数基于风险的认证模型使用。

本研究得到由北莱茵-威斯特法伦州资助的研究生培养小组“以人为中心的系统安全”（*NERD.NRW*）的支持。

© 2022年 Stephan Wiefling、Jan Tolsdorf、Luigi Lo Iacono。本文发表于 IWPE ’21 的开放获取版本。数字对象标识符：10.1109/EuroSPW54576.2021.00040

为了评估这些增强措施在实际中的效用，我们针对一项与安全性和隐私高度相关的基于风险的认证特征——IP地址，对其中一部分增强措施进行了评估。我们使用一个数据集进行评估，该数据集包含某真实在线服务中780名用户超过1.8年的登录历史。

我们的结果首次表明，在维持基于风险的身份验证的安全性和可用性特征的同时，提高特征隐私是可行的。不过，提高隐私水平受到某些条件的限制，设计基于风险的身份验证系统时需要考虑这些条件。我们还指出了未来广泛采用基于风险的身份验证可能带来的挑战和研究方向。

研究结果有助于服务所有者提供符合数据保护要求的基于风险的认证解决方案，也能协助开发者设计隐私保护能力更强的基于风险的身份验证实现。研究人员则能深入了解如何提升基于风险的身份验证的隐私友好性，以及未来的研究方向。

2. 背景

在以下章节中，我们将简要介绍基于风险的身份验证，并说明其使用与行业标准和法律法规所定义的多项隐私原则之间的关联。

2.1. 基于风险的认证模型

由于基于风险的身份验证并非标准化流程，实践中存在多种解决方案。我们重点介绍 Freeman 等人提出的实现 [12], ，因为它在先前的一项研究中表现最佳 [34]。此外，这一基于风险的认证模型据悉已被广泛采用，例如亚马逊、谷歌和领英等热门在线服务都在使用 [34], [36]。

模型针对用户 u 和一组特征值 $(FV^1, ..., FV^d)$，根据 d 个特征计算风险评分 S，如下所示：

$$S_u(FV) = \left(\prod_{k=1}^{d} \frac{p(FV^k)}{p(FV^k|u, legit)} \right) \frac{p(u|attack)}{p(u|legit)} \quad (1)$$

S 表示特征值在所有用户的全局登录历史中出现的概率 $p(FV^k)$，而 $p(FV^k|u, legit)$ 表示合法用户的个人登录历史中包含该特征值的概率。概率 $p(u|attack)$ 描述用户正在遭受攻击的可能性，而 $p(u|legit)$ 描述合法用户正在登录的可能性。

2.2. 监管基础

过去几年里，新数据保护法的出台，例如《通用数据保护条例》（GDPR） [8] 和《加州消费者隐私法》（CCPA） [30],，极大地改变了在线服务（即数据控制者）处理用户数据的方式。过去对用户数据处理方式较为宽松的建议，如今已被明确且具有约束力的数据保护原则取代，数据控制者必须遵守这些原则。然而，这些原则的具体内容和适用范围因司法管辖区而异。对于在国际范围内运营的数据控制者而言，这意味着其数据处理操作必须设计为兼容不同的要求。所幸，ISO 29100: 2011 [16] 中规定的隐私框架已汇集了全球各地数据保护法中共同的隐私原则。因此，它为数据控制者设计符合法律要求的数据处理操作提供了坚实基础，并可根据不同司法管辖区的具体要求进行调整。我们依据 ISO 29100:2011 中定义的隐私原则，概述了 RBA系统的设计要求，旨在使其兼容不同司法管辖区的要求。

隐私原则的适用性。一般来说，既定隐私法律和框架中定义的隐私原则旨在保护个人的隐私。因此，这些原则仅适用于具有个人关联的数据。这类数据例如称为“个人数据”（《通用数据保护条例》 [8]）、“个人信息”（CCPA [30]）或“个人身份信息”（PII）（ISO [16]）。这些定义非常相似，通常指“任何 *(a)* 可用于识别 *[*某个人*]* （与该信息相关的人），或 *(b)* 已经或可能与 *[*某个人*]*直接或间接相关的任何信息” [16]。

RBA处理的数据无疑符合这一界定，因为相关实现依赖的特征本身就可作为（唯一）标识符（例如IP地址） [36]。此外，RBA计算出的风险评分本身也是一种标识符，因为它由一组能够唯一识别个人的特征构成。因此，RBA必须遵守下文讨论的ISO 29100:2011隐私原则。

同意与选择。一般而言，数据控制者必须确保数据处理的合法性。虽然大多数司法管辖区承认用户同意是合法依据，但适用法律也可能允许在未经同意的情况下处理数据。根据与用户账户相关联的资产，数据控制者可能会主张，使用RBA是履行实施适当技术保障措施以防止未经授权的访问这一义务所必需的。尽管如此，为确保合规，提供者应在设计RBA机制时考虑用户同意，并向用户提供清晰易懂的说明。

收集限制与数据最小化。数据控制者必须将个人身份信息的收集和处理限制在实现特定目的所必需的范围内。因此，应评估RBA特征集是否适用，并移除冗余或不适当的特征 [34]。这包括考虑在RBA中使用假名化数据，并在特征值不再有助于实现RBA目的时将其销毁。在实践中，这带来的挑战是不能降低风险评分的可靠性。

使用、保留和披露限制。 数据处理必须限于数据控制者指定的目的，且不得向指定对象以外的接收方披露数据。基于风险的身份验证应确保相关特征不能用于计算风险评分以外的目的。此外，特征值过时后，应将其安全销毁或匿名化。需要指出的是，隐私法律不适用于匿名化数据，因此数据控制者可将其用于开发和测试，且不受隐私声明中规定的保留期限限制。

准确性与质量。 数据控制者必须确保

所处理的数据准确且质量可靠。这不仅出于其自身的业务利益，也是因为数据主体有权期待其数据准确无误。这会直接影响基于风险的身份验证，因为它可能拒绝用户一项重要权益（即访问其用户账户），并可能造成重大损害。因此，数据控制者必须采取适当措施，确保存储的特征值正确且有效。

个人参与和访问。数据控制者必须允许数据主体访问和查看其个人身份信息。对于基于风险的身份验证而言，这意味着应允许用户获取所用特征值的副本。

信息安全。数据控制者有义务在运营、职能和战略层面采取适当的控制措施，保护个人身份信息免受风险影响。这些风险包括但不限于与未经授权的访问或处理以及拒绝服务相关的风险。隐私法律对此要求提供全面保护，“考虑到现有技术水平、实施成本、处理的性质、范围、背景和目的，以及对自然人的权利和自由造成风险的可能性和严重程度”（《通用数据保护条例》第32条第1款）。由于 RBA 风险评分不一定依赖于评估明文特征值 [34]，因此应根据基于风险的身份验证的具体实现，以适当的假名化、屏蔽、截断或加密形式存储所收集的数据。此外，数据控制者应视需要实施额外的技术和组织措施，并确保基于风险的身份验证的完整性、可用性和韧性。

问责制与隐私合规。数据控制者应告知数据主体隐私相关政策、个人身份信息向其他国家的转移情况，以及数据泄露事件。数据控制者还应采取组织措施，帮助其核实并证明自己符合法律规定。这些措施包括但不限于风险评估和恢复程序。因此，基于风险的身份验证的实施应考虑RBA特征对攻击者和数据主体的价值，以及从数据泄露中恢复的问题。这对于避免损害用户账户及其相关资产的安全性至关重要。

3. 隐私增强措施（研究问题1）

为遵循隐私原则及由此衍生的数据保护要求，服务所有者应考虑采用相关机制，以增强其基于风险的身份验证实施中的隐私保护。下文将介绍威胁及其缓解措施，以增强基于风险的认证特征的隐私属性。

3.1. 特征敏感度与影响级别

RBA特征集旨在区分攻击者与合法用户。为此，特征中可能包含敏感的个人身份信息。然而，用户对这类个人身份信息敏感程度的认知各不相同 [28]。这些信息的（非预期）泄露也可能对用户隐私造成深远的负面影响。因此，开发者和提供者应评估 RBA 特征值的机密性丧失所造成的影响。具体而言，需考虑以下方面 [20]:

可识别性与可链接性。 应评估 RBA 特征集识别其背后自然人的能力。尤其是依赖侵入式在线跟踪方法（如浏览器指纹识别）的 RBA 系统，会存储敏感的浏览器特定信息，形成关联标识符。一旦机密性遭到破坏，即使用户在不同在线服务中使用不同的登录凭据或化名，这些特征也能将他们在不同服务上的个人资料明确关联起来。视服务类型而定，这可能给个人带来负面的社会或法律后果。它还可能使更广泛、非预期的活动跟踪成为可能，并导致与用户账户相关的信息被去匿名化。此前的研究发现，若重点是检测账户接管尝试，强大的 RBA 特征集并不需要唯一识别用户 [34]。此外，当用户确信敏感信息经过匿名化处理且无法识别其身份时，他们更愿意接受对这些信息的处理 [18], [29]。因此，使用非侵入式特征也可能提升用户对在线服务的信任。

特征值敏感性。除了RBA特征集能够识别个人之外，个体特征值本身也可能包含敏感的个人身份信息。在RBA范围内，敏感的个人身份信息可能是容易被伪造、且在发生数据泄露时可能被用于攻击其他在线服务的特征值。敏感的个人身份信息也可能指在线用户认为敏感的数据。例如，当前RBA方法中最重要的特征，即IP地址 [12], [15], [31], [34], [36]，被来自不同文化背景的在线用户视为高度敏感信息 [3], [19], [28]。由于用户通常不太愿意分享敏感度较高的数据，因此，RBA特征集应尽可能限制敏感数据的使用，以满足用户的利益。

3.2. 威胁

基于风险的认证特征可能包含个人敏感数据，因此必须防范攻击者。为了帮助在线服务做好防护工作，我们提出了三种隐私威胁类型。这些威胁类型基于文献中已有的分类以及我们在实践中的观察。

数据滥用。 在线服务可能会将自身的基于风险的认证特征数据用于非预期目的，例如用户跟踪、用户画像或广告 [5]。此前，曾发生过将为双因素认证用途而存储的电话号码用于此类滥用的情况 [33]。用户必须信任在线服务不会滥用其数据；与此同时，负责任的在线服务也应采取预防措施，尽可能降低滥用或非预期处理的风险，例如内部人员行为不当，或公司所有权发生变更后导致的此类情况。

数据转交。 在线服务可能会被要求或被迫交出存储的特征数据，例如交给国家机构、广告网络或其他第三方。尤其是 IP 地址，常常会被要求提供 [9]。将此类数据转交给第三方会侵犯用户的隐私。例如，IP地址可能会被用于揭示用户的地理位置，甚至其身份。

数据泄露。攻击者获取了包含特征值的数据库，例如通过入侵在线服务。结果，

在线服务失去了对其数据的控制。攻击者可以尝试根据特征值重新识别用户，例如将这些特征值与其他数据集结合起来。他们还可以进一步尝试重现特征值，并大规模发起账户接管攻击，类似于凭证填充攻击。成功后，他们可能访问存储在在线服务上的敏感用户数据，例如私人消息。

3.3. 缓解

在线服务可以采取多种措施，缓解上述隐私威胁。我们提出五项措施，这些措施基于相关研究领域的方法、隐私法规，以及我们在所选基于风险的认证模型（见第 2 节）中的观察。根据所介绍的基于风险的认证模型，我们将所有特征值视为分类数据，以便计算概率。满足这一条件时，所提出的措施也适用于其他基于风险的认证模型 [34]。

作为实际解决方案的示例，我们说明如何将这些考量应用于*IP*地址 特征，具体针对 IPv4地址。我们选择这一特征，是因为迄今为止，它既被认为是安全性方面最重要的基于风险的身份验证特征，也是隐私方面敏感的可重新关联数据（见第 3.1 节）。

3.3.1. 聚合。 基于风险的认证模型仅依赖特征值频率。为尽量减少数据并限制滥用 [16]，，我们可以对登录历史中的特征数据进行聚合或重新排序，而不影响结果。这样，数据集就会显示某种特征组合出现的频率，但不会显示其时间顺序。移除这些信息可以缓解登录序列中的重新识别风险。

3.3.2. 哈希。 密码哈希函数（如 SHA-256）会将任意值的数据输入转换为固定长度的输出。从理论上讲，哈希函数无法逆向，因此攻击者需要重新计算所有可能的哈希值，才能还原输入值 [17]。假设这些哈希值在实际应用中无碰撞，那么使用哈希化特征值将产生与原始值相同的基于风险的身份验证结果。这是因为特征值只是被转换成了另一种表示形式。因此，这可以作为一种解决方案，保护特征数据免受上述威胁。

然而，IPv4地址的输入值仅有 32 位，其中一些地址具有特定语义和用途，不能分配给设备。因此，攻击者只需对所有 $2^{32}-1$ 值进行哈希，就能还原正确的 IP 地址。为解决这一问题，我们可以在输入值后附加一个较长的随机字符串（盐值）：

$$H(192.168.1.166 \mid\mid salt) = 243916...aad132 \quad (2)$$

攻击者必须正确猜出盐值；盐值越长，这项工作就越困难。因此，这种缓解策略会增加猜测每个特征值所需的时间。更进一步，我们甚至可以对结果进行多次哈希，以增加计算时间：

$$H(H(...H(192.168.1.166 \mid\mid salt))) = [hash] \quad (3)$$

这与密码数据库中使用的密钥派生策略类似 [22]。不过，我们只能使用一个全局盐值来处理数据库中的所有条目，因为基于风险的认证机制需要能够识别数据库中不同用户之间相同的特征值。通过提高计算成本，攻击者就无法像使用未哈希特征值那样大规模发起攻击。

3.3.3. 截断。一种更具破坏性、用于增强基于风险的认证特征隐私的方法，是更改或移除其数据值中的细节。这可以减少具有唯一准标识符的记录数量。由于特征数据因此对跟踪或重新识别等其他用途不再那么有用，我们将其视为缓解隐私威胁的一种措施。对于 IP 地址，我们可以将末尾的比特设为零。例如，截断末尾八比特后，结果如下：

$$Truncate(192.168.1.166, 8\ Bit) = 192.168.1.0 \quad (4)$$

这种机制在 IP 地址匿名化策略中已有应用 [6], [7]。不过，我们也可以将其应用于其他特征，例如降低时间精度，或对用户代理字符串中的浏览器版本号进行粗粒度处理 [24]。由于我们移除了可能用于识别个人的信息，例如设备的互联网连接，这有可能增强隐私保护。不过，这也可能影响基于风险的身份验证的结果，因为攻击者可猜测的特征值会减少。

3.3.4. K-匿名性。k-匿名性隐私概念 [32] 确保数据集中至少有 k 条记录具有相同的准标识符值。如果攻击者获取了该数据集并知道某个受害者的 IP 地址，他们将无法从 k 其他用户中辨别出此人。因此，它是应对数据转交和数据泄露情况下重新识别的有效对策。

要在基于风险的身份验证中实现 k-匿名性，至少需要有 k 个用户具有相同的特征值。为确保这一点，我们在每次成功登录后，都会向基于风险的身份验证登录历史中添加可能缺失的记录。我们将这些记录添加到随机用户的记录中，使其仅影响全局登录历史概率，从而维持较高的安全级别。创建这些用户仅仅是为了这个目的。为保留全局数据集的属性，用户数量逐步增加，使每位用户的平均登录尝试次数保持一致。

3.3.5. 登录历史最小化。另一种方法是限制登录历史中功能和记录的数量，可按记录数或固定时间段进行限制 [16]。一项研究已表明，只需少量记录即可实现较高的基于风险的身份验证保护水平 [34]。这样做可以减轻对用户进行长期跟踪的问题。不过，这可能会根据相应在线服务的使用模式影响基于风险的认证性能。特别是对于使用频率低于每月一次的在线服务，我们认为，与每日使用的网站相比，为实现相当的基于风险的认证性能，需要将特征存储更长时间。

4. 案例研究评估（RQ2）

在无碰撞的情况下，聚合和哈希不会影响基于风险的身份验证的结果，因为它们只会改变基于风险的身份验证模型的数据表示形式。其他方法则可能会产生影响。为了评估它们对

实际基于风险的身份验证行为的影响，我们使用真实世界登录数据研究了截断和 k-匿名性。由于数据集的属性和规模有限，我们无法可靠地测试登录历史最小化方法，因此将其留待未来研究。不过，为了确保内容完整，我们概述了这一相关的隐私考量。与其他示例一样，我们使用了 IP 地址特征。

4.1. 数据集

在评估中，我们使用了长期收集的基于风险的身份验证数据集，其中包括从真实世界的在线服务 [34]上收集的780名用户的特征。在线服务在每次成功登录后收集用户特征。从2018年8月至2020年6月，用户总共登录了9555次。他们大多每天登录（44.3%）或每周登录数次（39.2%），平均总登录次数为12.25次。为提高数据质量和有效性，我们移除了所有发现其账户存在非法登录的用户。该在线服务是一个电子学习网站，学生用它来练习学习课程和考试。由于用户大多位于同一座城市，因此这是一个对基于风险的身份验证颇具挑战性的数据集。他们更有可能获得相似的IP地址。因此，评估在这种极具挑战性的场景下基于风险的认证保护如何变化至关重要。

4.1.1. 法律与伦理考量。研究参与者 [34] 签署了同意书，同意为研究目的收集和使用数据。他们始终可以查看自己的数据副本，并可根据要求将其删除。收集的数据存储在加密硬盘上，只有研究人员能够访问。

我们所在的大学没有正式的机构审查委员会（IRB）流程。尽管如此，我们仍确保遵循德国社会学协会（DGS）的伦理规范和德国研究基金会（DFG）的良好科学实践标准，以尽量减少潜在危害。我们也确保遵守《通用数据保护条例》。

4.1.2. 局限性。我们的研究结果仅适用于本数据集以及参与研究的用户，也仅适用于某个国家特定地区的人群。这些结果并不能代表大规模在线服务，但展示了一个典型案例：用户每天至每周使用某个网站。与类似研究一样，我们无法完全排除有高明的攻击者对该网站发起定向攻击的可能性。不过，多种对策降低了网站遭到入侵的可能性 [34]。

4.2. 攻击者模型

我们使用相关文献中提到的三种基于风险的身份验证攻击者模型，评估了隐私增强措施 [12], [34]。所有攻击者都拥有目标用户的登录凭据。

朴素攻击者会尝试从世界各地随机选择的互联网服务提供商（ISP）处登录。我们使用来源于针对在线服务的真实攻击的 IP 地址来模拟这些攻击者 [11]。

VPN 攻击者 知道受害者所在的国家。因此，我们使用来自现实世界中位于受害者所在国家的攻击者的 IP 地址来模拟这些攻击者 [11]。

定向攻击者 知道受害者所在的城市、浏览器和设备。因此，他们会选择相似的特征值，包括相似的互联网服务提供商。我们使用数据集模拟了这些攻击者，所用的是除受害者之外所有用户的独特特征组合。由于我们数据集中的 IP 地址彼此距离较近，因此模拟的攻击者了解这些情况，并以类似的方式选择了这些地址。

4.3. 方法论

为了从实际基于风险的认证解决方案的角度测试我们的隐私增强措施，我们定义了一组期望属性。我们的增强措施需要做到：（A）**保持被拦截攻击者的比例**：采用隐私增强措施时，拦截大量攻击者的能力不应降低。这对于维持基于风险的身份验证系统的安全属性至关重要。（B）**保留对合法用户与攻击者的区分**：应用这些措施后，合法用户与攻击者之间的风险评分差异应仅在极小范围内变化。否则，基于风险的身份验证系统的可用性和安全属性都会降低。

下面概述用于评估隐私增强措施的测试。我们根据 Wiefling 等人 [34], 提出的方法，通过重放用户会话，复现攻击者和合法用户的登录行为。我们在复现过程中集成了截断和 k-匿名性，以测试这些对策。

基于风险的认证模型使用 IP 地址和用户代理字符串作为特征，因为这可以被视为当前基于风险的身份验证的实践水平 [34], [36]。我们将 IP 地址截断 0 至 24 比特，以观察这对基于风险的认证性能的影响。我们假设，截断超过 25 比特将无法可靠地检测攻击者。我们还针对 IP 地址特征测试了 k-匿名性，直至 $k = 6$。由于美国政府机构认为少于五条记录属于敏感信息 [10], ，我们选择涵盖这一阈值。

4.3.1. 测试 A：被拦截攻击者的百分比。为了比较基于风险的身份验证在全部三种攻击者模型下的表现，我们计算了攻击者中会被拦截的比例。我们将这一比例称为真正率（TPR），正如先前研究所做的那样 [12], [34]。为进行公平比较，我们观察了以拦截 99.5% 的攻击者为目标时 TPR 的变化。我们选择这一 TPR 基线，是因为先前的一项研究表明，它在可用性和安全性方面表现良好 [34]。

为便于比较，我们将每个截断或 k-匿名性步骤对应的 TPR 调整为相对于未修改基线的百分比差异（相对 TPR）： x_i

$$TPR_{relative_{x_i}} = \frac{TPR_{x_i} - TPR_{baseline}}{TPR_{baseline}} \quad (5)$$

此后， $TPR_{relative_{x_i}} < 0.0$ 表示 TPR 相较于基线有所下降。

4.3.2. 测试 B：风险评分变化。为了确定基于风险的认证模型中攻击者与合法用户的可区分程度，我们计算了风险评分比率（RSR） [34]。它表示攻击者与合法用户的平均风险评分之间的比率：

$$RSR_{basic} = \frac{mean\ attacker\ risk\ score}{mean\ legitimate\ risk\ score} \quad (6)$$

为便于比较，我们将每个截断或 k-匿名性步骤 x_i 对应的 RSR 归一化为相对于基线的百分比差异（相对 RSR）。基线是未经修改的 IP地址：

$$RSR_{relative_{x_i}} = \frac{RSR_{basic_{x_i}} - RSR_{baseline}}{RSR_{baseline}} \quad (7)$$

因此，$RSR_{relative_{x_i}} < 0.0$ 表明，引入隐私增强措施后，攻击者与合法用户之间已无法像之前那样被准确区分。

4.3.3. 限制提取。对于每项测试，我们都设定了以下阈值，以提取不会使基于风险的认证性能降低到不可接受程度的限制。

（测试 A）我们要求基于风险的认证性能保持不变。因此，我们将合理限制设定为相对 TPR 相较于基线开始下降的临界点，即攻击者不再像以前那样能够被有效拦截。（测试 B）与跟踪不同，基于风险的身份验证除了使用已验证的标识符（例如密码）外，还会使用特征信息。因此，考虑到隐私，我们认为略微降低 RSR 是可行的。根据我们的观察，在案例研究评估中，RSR 的变化低于 0.01 是可以接受的。因此，我们将合理限制设定为相对 RSR 低于 0.01 的临界点。

4.4. 结果

下面我们将介绍所有攻击者模型的结果。本节之后将对结果进行讨论。我们使用了一个拥有 2000 多个核心的高性能计算集群进行评估。这是必要的，因为使用模拟攻击者计算结果的过程需要大量计算资源。

在统计检验方面，对于总体情况我们使用了 Kruskal-Wallis 检验，并使用带 Bonferroni 校正的 Dunn 多重比较检验进行事后分析。我们将小于 0.05 的 p 值视为具有统计显著性。

4.4.1. 截断。图 1 展示了所有攻击者的截断检验结果。定向攻击者与另外两种攻击者之间的 TPR 差异均具有统计显著性（定向/天真：p=0.0151，定向/VPN：p<0.0001）。对于天真攻击者，TPR 在超过 20 比特后超出限值；对于 VPN 攻击者，在超过 3 比特后超出限值；对于定向攻击者，在超过 14 比特后超出限值。

关于相对 RSR，VPN 攻击者与另外两种攻击者之间存在显著差异（p<0.0001）。对于天真攻击者，RSR 在超过 3 比特后超出限值；对于 VPN 攻击者，在超过 21 比特后超出限值；对于定向攻击者，在超过 3 比特后超出限值。

综合两项结果，根据我们的标准，所有攻击者模型可接受的截断限制均为 3 比特。

4.4.2. k-匿名性。图 2 展示了三种攻击者模型下 k-匿名性测试的综合结果。对于定向攻击者，相对 TPR 在 $k = 1$ 后下降；对于朴素攻击者，在 $k = 2$ 后下降；而 VPN 攻击者的相对 TPR 至少到 $k = 6$时都没有下降。朴素攻击者与 VPN 攻击者之间的 TPR 差异显著（p=0.0066）。

所有类型攻击者的相对 RSR 均未下降，且差异均不显著。

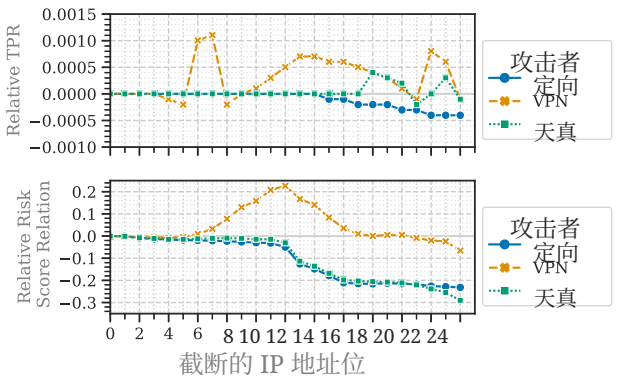

图 1. 截断 IP地址的结果。上：相对 TPR（测试 A）。定向攻击者与 VPN 攻击者和朴素攻击者之间的差异均显著。下：相对 RSR（测试 B）。VPN 攻击者与定向攻击者和朴素攻击者之间的差异均显著。

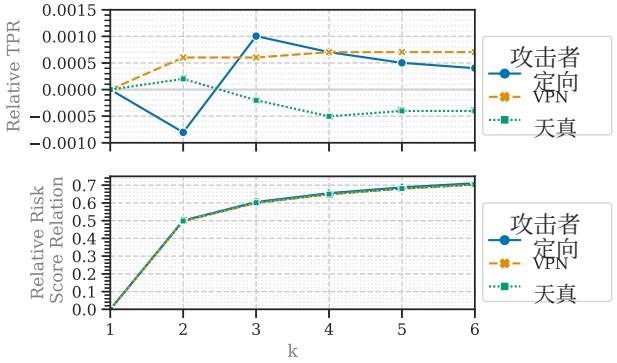

图 2. 关于 IP地址的 k-匿名性结果。上：相对 TPR（测试 A）。朴素攻击者与 VPN 攻击者之间的差异显著。下：相对 RSR（测试 B）。差异均不显著。

综合结果，根据我们的标准，可接受的 k 级别为：定向攻击者 $k = 1$，朴素攻击者 $k = 2$，VPN 攻击者至少为 $k = 6$。

5. 讨论

我们的结果表明，IP 地址截断会显著影响基于风险的身份验证的风险评分，并降低攻击检测的概率。对于 VPN 攻击者，截断使 RSR 在 12 比特处达到局部最大值，因而表面上改善了检测效果。然而，这是因为该 VPN 攻击者仅拥有一个局限于 VPN 服务服务器位置的 IP 地址范围。由于前若干 IP 地址位对应于节点的地理位置，因此它们与居住在不同区域的合法用户大多不同。于是，随着截断位数增加至 12 比特，VPN 攻击者的风险评分不断上升，因为全局登录历史 $p(FV^k)$ 的概率降低，而本地历史 $p(FV^k|u, legit)$ 的概率保持不变。相比之下，定向攻击者同样拥有有限的 IP 地址范围，但他们位于与合法用户相同的区域。此外，朴素攻击者拥有较大的 IP 地址范围。因此，在这两种情况下，$p(FV^k)$ 与 $p(FV^k|u, legit)$ 之间的差异在达到 12 比特之前都保持在相近水平。

基于上述内容以及我们的评估结果，我们

表 1. 为实现 K-匿名性而添加登录条目所产生的开销

k	额外条目	相对于基线的增幅
1	0	0.0
2	3928	0.41
3	7965	0.83
4	12013	1.26
5	16065	1.68
6	20120	2.11

不建议在我们的案例研究场景中截断超过三比特，以确保基于风险的认证性能稳定。

K-匿名性增强了合法用户与攻击者之间的可区分性，即 RSR。这是因为该机制向全局登录历史中添加了新条目。因此，全局登录历史中未知特征值的总体概率 $p(FV^k)$ </style}降低，使攻击者更难获得较低的风险评分。然而，在大多数情况下，这也降低了攻击者的检测率，即 TPR，因为 k 数据集中具有相似特征值的用户增多了。作为这些结果的副作用，原本独特的特征值整体上变得不再那么独特。因此，由于 $k = 1$ 的既定限制，只有牺牲基于风险的认证性能，才能为定向攻击者实现 k-匿名性。

额外条目产生的开销随着每个 k 而增加（见表 1）。当 k>3 时，开销甚至超过了数据集本身，这使得当前机制不适用于规模极大的在线服务。为缓解这一问题，可以引入相关机制，在经过一段时间后满足 k-匿名性时，删除部分额外登录条目。

随着特征数量不断增加，k-匿名性无法扩展，而其他方法则可以。 [1], 因此，合理的 RBA 隐私增强措施可能是结合所有已概述的对策，以确保可扩展性。

基于我们的研究结果，下面将讨论隐私挑战和未来的研究方向。

5.1. 隐私挑战

将隐私纳入 RBA系统时，实践中应考虑若干挑战。我们将在下文加以说明。

IP 地址特征的作用。对于某些应用而言，结合多种针对 IP 地址的隐私增强措施可能已经足够。然而，这一特征仍属于敏感信息。因此，人们不禁要问，在线服务是否应考虑采用保护隐私的替代方案，而不是存储 IP 地址。一种替代方案是仅从 IP 地址中提取区域和自治系统编号，然后丢弃其余信息。其他方法甚至可以在完全不依赖 IP 地址的情况下识别网络异常，例如通过 VPN 连接进行 IP 欺骗。例如，服务器发起的往返时间（RTT） [34]可用于估算用户设备与服务器位置之间的距离，并可替代 IP 地址作为基于风险的认证特征。由于 RTT 会因服务器位置而异，因此对于利用泄露的数据库实施的大多数重新识别攻击而言，它们会变得毫无用处，因为实际应用中的服务器位置是分散的。还可以为 RTT 添加随机噪声，以进一步增强隐私保护。

特征填充的风险。随着未来基于风险的身份验证得到广泛采用，这些考量可能会变得越来越重要。我们认为，当包含基于风险的认证特征值的数据库被窃取时，可能会给其他使用基于风险的身份验证的服务带来严重后果。与密码不同，行为型基于风险的认证特征值一旦泄露便无法更改。攻击者可能会尝试在其他网站上自动重现这些特征值。因此，采用更难被攻击者欺骗的保护隐私的替代方案，或许对于缓解大规模的“特征填充”攻击至关重要。

处理数据删除请求。此外，还可能与数据保护法规产生冲突。法律赋予用户请求删除数据的权利。因此，当用户要求在线服务删除其基于风险的认证特征数据时，他们可能会失去其用户账户的 RBA 保护。

5.2. 研究方向

我们的案例研究评估首次揭示了截断特征值以增强隐私的效果。结果表明，在一定程度上可以做到这一点，同时维持基于风险的认证性能。未来的工作可以探讨将这种方法用于其他类型的特征，例如用户代理字符串。

所提出的 k-匿名机制可以增强数据集中唯一条目所涉及的隐私。不过，即使用户具有一组典型的特征值，例如家庭和工作场所的 IP 地址，仍可能被识别出来。动态数据库中的研究已探讨了这项复杂任务 [27], [37]。未来的工作可以研究此类机制是否也适用于基于风险的身份验证。

由于我们无法使用数据集可靠地测试登录历史最小化方法，未来的工作应在定期使用的中大型在线服务中对此进行研究。

6. 相关工作

Burkhard等人 [6] 研究了在异常检测系统中截断IP地址。他们发现，截断超过四个比特会降低这些系统的性能。Chew等人 [7] 进一步评估了在入侵检测系统中截断IP地址的做法。结果表明，在许多受测分类器中，移除超过 8 个比特后，检测准确率会下降。我们的研究表明，可以从IP地址中移除三个比特，同时维持基于风险的认证性能。

Safa等人 [26]，和Blanco-Justicia和Domingo-Ferrer [4]提出了使用移动设备进行隐式认证的隐私保护认证模型。这些模型依赖客户端生成的特征，前者还在客户端设备上计算风险评分。然而，这并不适用于我们的基于风险的认证用例，因为该用例依赖服务器生成的特征和风险评分来防止客户端欺骗。

据我们所知，尚无研究探讨 RBA系统中的隐私增强措施。不过，一些文献涉及与 RBA相关的隐私方面。Bonneau等人 [5] 讨论了使用额外特征进行认证所引发的隐私问题。他们发现，隐私保护技术或许可以缓解这些问题，但当时尚未在实践中部署。我们

首次在我们的案例研究中提出并测试了一些技术。Wiefling 等人 [35] 研究了 RBA的可用性和安全感知。结果显示，出于隐私原因，用户往往拒绝向在线服务提供电话号码。他们还在一个真实在线服务中研究了基于风险的身份验证特征 [34],，结果表明，要实现良好的基于风险的认证性能，特征集可以非常小。我们证明，通过不同的机制可以进一步增强隐私。

7. 结论

随着基于风险的身份验证（RBA）被广泛用于保护用户免受利用被盗凭据发起的攻击，越来越多的在线服务可能会长期存储用户的敏感特征数据，例如 IP 地址和浏览器标识符。此类信息一旦被转发或泄露，就会对用户隐私构成潜在威胁。为缓解这些威胁，RBA系统的设计必须在安全性与隐私之间取得平衡。

我们的研究结果初步表明，当前实际使用的基于风险的身份验证实现可以设计得更加注重隐私。不过，迄今为止的研究仍有一些尚未解决的挑战。例如，一个重要问题是如何用更能保护隐私的替代方案取代 IP 地址特征。一方面，我们认为 IP 地址与重新识别攻击密切相关 [9]。因此，将其从基于风险的身份验证登录历史中移除，可以增强隐私保护。另一方面，IP 地址是一项能够显著提升安全性的特征 [34]。未来研究必须仔细识别并分析此类权衡，以免基于风险的身份验证的用户接受度因首次数据泄露而下降。

参考文献

[1] C. C. Aggarwal，《k-匿名性与维度灾难》，载于*VLDB '05*。VLDB 基金会，2005年8月。[2] Akamai，《忠诚度计划被出售——零售与酒店业欺诈》，*[互联网现状] / 安全*，第6卷第3期，2020年10月。[3] K. Almotairi 和 B. Bataineh，《沙特阿拉伯互联网用户对信息敏感性的认知》，*AIP*，第9卷第2期，2020年。[4] A. Blanco-Justicia 和 J. Domingo-Ferrer，《高效的隐私保护型隐式认证》，《计算机通信》，第125卷，2018年7月。[5] J. Bonneau、E. W. Felten、P. Mittal 和 A. Narayanan，《用于网络身份验证的隐式辅助因素所引发的隐私问题》，载于*WAY'14*，2014年7月。[6] M. Burkhart、D. Brauckhoff、M. May 和 E. Boschi，《IP 地址截断中的风险—效用权衡》，载于*NDA '08*。ACM，2008年。[7] Y. J. Chew、S. Y. Ooi、K.-S. Wong 和 Y. H. Pang，《通过截断方法在基于网络的入侵检测系统中保护 IP 地址隐私》，载于*ICSCA '19*。ACM，2019年。[8] 欧盟，《通用数据保护条例》，2016年5月，欧盟条例 2016/679。[9] 欧洲刑警组织，《SIRIUS 欧盟数字证据情况报告 2019》，2019年12月。[10] 联邦统计方法委员会，《统计披露报告》，2005年12月。[11] FireHOL，《所有网络犯罪 IP 订阅源》，2020年8月。 [在线]。网址：http://iplists.firehol.org/?ipset=firehol level4

[12] D. Freeman、S. Jain、M. D¨urmuth、B. Biggio 和 G. Giacinto，《你是谁？一种衡量用户真实性的统计方法》，载于 NDSS '16。互联网协会，2016年2月。[13] P. A. Grassi 等，《数字身份指南：身份验证与生命周期管理》，美国国家标准与技术研究院，技术报告 NIST SP 800-63b，2017年6月。[14] M. J. Haber，《攻击向量》，载于《特权攻击向量：构建有效的网络防御策略以保护组织》。Apress，2020年。[15] A. Hurkała 和 J. Hurkała，《面向 Web 环境的上下文风险感知身份验证系统架构》，载于 ICIEIS '14，2014年。[16] ISO，《ISO/IEC 29100:2011(E)：信息技术——安全技术——隐私框架》。ISO/IEC，2011年。[17] D. Llewellyn-Jones 和 G. Rymer，《破解 PwdHash：针对客户端密码哈希的暴力攻击》。Apollo，2017年。[18] E. Markos、L. I. Labrecque 和 G. R. Milne，《新型信息透镜：自我概念与交换情境如何帮助理解匿名信息和个人身份信息的信息敏感性》，JIM，第42卷，2018年。[19] E. Markos、G. R. Milne 和 J. W. Peltier，《信息敏感性与提供意愿连续谱：美国与巴西的隐私比较研究》，JPP&M，2017年。[20] E. McCallister、T. Grance 和 K. A. Scarfone，《保护个人身份信息（PII）机密性的指南》，技术报告 NIST SP 800-122，2010 年。[21] I. Molloy 、 L. Dickens 、 C. Morisset、P.-C. Cheng、J. Lobo 和 A. Russo，《不确定性下基于风险的安全决策》，载于 CODASPY '12。ACM，2012年2月。[22] K. Moriarty、B. Kaliski 和 A. Rusch，《PKCS #5：基于密码的密码学规范 2.1 版》，RFC 8018，2017年1月。[23] 国家网络安全中心，《云安全指南：10，身份与身份验证》，技术报告，2018年11月。[24] G. Pugliese、C. Riess、F. Gassmann 和 Z. Benenson，《浏览器指纹识别的长期观察：用户的可追踪性与观点》，PoPETS，第2020卷第2期，2020年4月。[25] N. Quermann、M. Harbach 和 M. D¨urmuth，《真实环境中的用户身份验证现状》，载于 WAY '18，2018年8月。[26] N. A. Safa、R. Safavi-Naini 和 S. F. Shahandashti，《隐私保护型隐式认证》，载于 IFIP SEC '14。Springer，2014年。[27] J. Salas 和 V. Torra，《动态数据库上实现 k-匿名性的一般算法》，载于 DPM '18。Springer，2018年。[28] E.-M. Schomakers、C. Lidynia、D. M¨ullmann 和 M. Ziefle，《互联网用户对信息敏感性的认知——来自德国的见解》，IJIM，第46卷，2019年6月。[29] E.-M. Schomakers、C. Lidynia 和 M. Ziefle，《我的所有数据？用户对隐私保护型数据市场的偏好及匿名性的重要性》，Electronic Markets，第30卷第3期，2020年2月。[30] 加利福尼亚州，《加州消费者隐私法》，2018年6月，第375号议会法案。[31] R. H. Steinegger、D. Deckers、P. Giessler 和 S. Abeck，《适用于 Web 应用程序的基于风险的身份验证器》，载于 EuroPlop '16。ACM，2016年6月。[32] L. Sweeney，《k-匿名性：一种保护隐私的模型》，IJUFKS，第10卷第05期，2002年10月。[33] G. Venkatadri、E. Lucherini、P. Sapiezynski 和 A. Mislove，《调查 Facebook 定向广告中个人身份信息的来源》，PoPETS，第2019卷，2019年1月。[34] S. Wiefling、M. D¨urmuth 和 L. Lo Iacono，《网站用户能从中获得什么：一项关于基于风险的身份验证特征的数据驱动长期研究》，载于 FC '21。Springer，2021年3月。[35] S. Wiefling、M. D¨urmuth 和 L. Lo Iacono，《仅有好的密码就够了吗？一项关于基于风险的身份验证之可用性与安全感知的研究》，载于 ACSAC '20。ACM，2020年12月。[36] S. Wiefling、L. Lo Iacono 和 M. D¨urmuth，《这真的是你吗？一项关于真实环境中基于风险的身份验证的实证研究》，载于 IFIP SEC '19。Springer，2019年6月。[37] X. Xiao 和 Y. Tao，《M-不变性：迈向动态数据集的隐私保护再发布》，载于 SIGMOD '07。ACM，2007年。