卫生信息化国际发展动态

北京市卫生健康大数据与政策研究中心 2024.11.20

（十一）数据所有权

1. **标题：** AI 驱动的综合医疗保健数据的所有权

来源： JMIR Med Inform.

时间： 2024 年 11 月.

链接： https://doi.org/10.2196/57754.

概要： 在综合医疗保健（IHC）和人工智能（AI）快速发展的当下，数据所有权问题已成为必须面对的重要问题。在 IHC 实践中，利用 AI 处理的大量患者数据有病历、健康的社会决定因素、生活方式因素和治疗反应等，这些敏感数据的所有权和控制权的明确至关重要，现有的个人保留健康数据所有权等不同法律规定和伦理观已导致医疗保健数据所有权的一定困境，比如，在美国等国家/地区，医疗保健提供者通常拥有物理意义上的患者记录，而非患者。同样，政府机构（如疾病登记管理员）拥有存储在其数据库中的患者数据。而患者虽然根据隐私或信息自由法（以及某些地区的健康信息法）享有访问权限，但这些并不等同于拥有所有权。基于此，本研究将聚焦探讨 IHC 和 AI 时代背景下的数据所有权的复杂动态，阐述数据所有权涉及患者、提供者、研究人员和 AI 开发人员等多方面性质，解决诸如模棱两可的同意、见解的归因和国际不一致等挑战，并评估适用于 IHC 环境的常见数据所有权模型，根据法律和伦理方面的适当性审查，提出新颖的协作医疗保健数据所有权（CHDO）框架，包括私有化和社区化假设，以及分布式访问控制、数据信任和区块链技术，并评估了它们的潜力和局限性，最后还就其在美国的影响和挑战提出相应的建议，以确保在综合健康（IH）中负责任地、透明地实施 AI 集成，为数据驱动型医疗保健的伦理与可持续发展铺平道路。

2. **标题：基于二维混沌映射和区块链的安全医疗数据共享方案**

来源： Scientific Reports.

时间： 2024 年 10 月.

链接： https://doi.org/10.1038/s41598-024-73554-x.

概要： 健康保险是国民基本医疗保健的重要保障，对社会稳定和公共利益至关重要，但是当下存在因数据提供者的隐私、权利和安全等原因，致使参保人员产生的大量医疗保健数据无法有效共享和利用。虽然现已有许多学者提出了基于密码学的混沌映射、区块链支持等加密技术，用于在数据共享中保护隐私，但是现有的医疗保健数据共享解决方案仍存在“复杂密码系统在实时通信和大规模加密场景中需要更多的计算成本”和“未能很好地集中数据完整性验证和身份验证”的问题。为此，本研究提出了一种基于二维混沌映射的加密方法，以增强医疗数据共享的数据安全性。其主要贡献如下：1）提出一种新的二维混沌映射——二维无限逻辑映射（2D-ILM），并分析其动态行为；2）基于 2D-ILM 设计了一种称为二维混沌映射加密（2CME）的数据加密方法；3）设计了一个基于二维混沌映射的数据共享方案 2DCM-DS，利用智能合约实现数据完整性验证和身份验证功能。研究最后的性能评估和安全性分析表明，所提方案具有鲁棒性、高效耗时和较强的抗攻击能力，满足医疗数据共享的通信和数据隐私保护要求，证明其在医疗保险数据共享场景中的应用潜力。但是，目前的方案仍然存在一些限制，比如区块链本身的可扩展性将限制一些基于区块链的数据共享方案的发展；当一些邪恶节点提供低质量数据或网络遭受拜占庭式攻击时，系统是否还能保证数据安全；将加密算法封装到智能合约中是否足够安全等，未来的工作将考虑扩展该计划内的区块链安全环境。

（徐健编辑）

译文一：

AI 驱动的综合医疗保健数据的所有权

Shuimei Liu，L Raymond Guo，徐健（译）

简介

综合医疗保健（IHC）强调的是对患者福祉的综合措施，所以人工智能（AI）被越来越多地应用于医疗保健服务的增强，两者的结合，使得相关数据所有权的问题变得至关重要。在 IHC 实践中创建并由 AI 处理的大量患者数据包括病史、健康的社会决定因素（SDOH）、生活方式因素和治疗反应。这些数据都强调了澄清对这些敏感信息的所有权和控制的重要性。特别是在 AI 时代，关于包括个人保留对数据的某些权益在内的健康数据所有权的不同法律和伦理观已经出现，导致了医疗保健数据所有权的困境。例如，在美国等国家/地区，医疗保健提供者通常拥有物理意义上的患者记录，而不是患者。虽然患者根据隐私或信息自由法（以及某些地区的健康信息法）享有访问权限，但这些并不等同于所有权。同样，政府机构（如疾病登记管理员）拥有存储在其数据库中的患者数据。基于此，我们将聚焦于对 IHC 和 AI 时代数据所有权的详细探索，并讨论其在美国的影响和挑战。

许多国家都在努力解决数据所有权框架问题，尤其是在医疗保健领域。有些，例如那些拥有国家电子健康记录的，优先考虑集中可访问性。其他一些国家则拥有联邦化的医疗保健系统，在国家和地区政府之间的责任划分中游刃有余。这种复杂性凸显了研究美国背景的必要性。在美国本土，数据所有权也可能有很大差异。就像医疗保健一样，一些区域可能具有更集中的数据访问，而另一些区域则使用更加分散的系统运行。这种内部多样性更加强了在美国各地探索此类框架的重要性。

本研究旨在评估适用于 IHC 环境的常见数据所有权模型，根据法律和伦理方面的适当性审查，我们计划通过这项探索为 AI 在 IHC 中负责任地整合的可持续交流做出贡献。该研究阐明了以患者为中心的数据治理的重要性，提供了对法律和伦理影响的见解，以及确保在综合健康（IH）中负责任地、透明地实施 AI。

IH 模型

IHC 采用整体健康方法，以思想、身体和精神的相互联系为中心，倡导解决个人健康各个方面的全面康复。IHC 通常采用跨学科模式，促进来自不同领域的从业者之间的合作，以提供最佳护理。这个综合团队可能包括医生、护士、针灸师、脊椎按摩师和其他医疗保健专业人员。近年来，IHC 模型发展势头强劲。美国国立卫生研究院（United States National Institutes of Health）建立了一个专门的中心，即国家补充与综合健康中心（National Center for Complementary and Integrative Health），用于 IHC 研究。此外，退伍军人健康管理局的整个卫生系统是其在大型医疗保健系统中实施的例证。这种日益增长的许可将 IHC 定位为学习健康系统的关键组成部分，旨在通过数据驱动的洞察不断改善以患者为中心的医疗护理。

IHC 的核心是整合传统和互补方法，形成一个协调的医疗保健生态系统。这种方法强调多模式干预，将常规医疗保健实践（药物治疗、身体康复和心理治疗）与互补健康方法（如针灸、瑜伽、按摩、生活方式指导等）相结合。这些量身定制的干预措施针对整个个体，而非仅仅关注特定器官的系统。IHC 旨在通过促进不同提供者和机构之间良好协作的护理，提供全面的医疗保健，解决个人福祉问题。

在美国的医疗保健环境中，将综合方法纳入健康和保健的势头越来越大。研究人员正在积极研究 IHC 在不同情况下的潜在益处，例如军事人员和退伍军人的疼痛管理、癌症患者和幸存者的症状缓解以及促进健康行为的计划。这些正在进行的调查旨在阐明 IHC 在提高患者整体健康状况方面的变革潜力，患者和提供者生成的数据通常会通过这些过程来体现。

IH 实践过程通常包括以下步骤：

1. 患者咨询 IH 医生，讨论他们的健康问题和目标。

2. IH 从业者评估患者的健康状况并制定治疗计划，该计划可能结合传统和 IH 实践的混合。

3. IH 从业者可能会将患者推荐给其他提供者，例如脊椎按摩师、针灸师或营养师。

4. 患者与这些额外的提供者合作，以制定更全面的治疗计划。

5. 医疗保健提供者共同努力，为患者提供最好的护理，重点关注具体情况和整体健康状况以及长期健康结果的改善。

例如，患有慢性疼痛的患者可能会咨询 IH 从业者，后者会制定包括物理治

疗、针灸和草药补充剂在内的治疗计划。医生还可能推荐营养师指导患者改变饮食，支持他们的康复，并将食物作为药物以缓解对身体变化的心理反应。IH 模型在医疗保健中越来越受欢迎，这反映出人们对解决健康问题的整体方法越来越感兴趣。IHC 从业者之间的跨学科合作使患者能够获得全面有效的护理。然而，值得注意的是，患者、提供者和其他利益相关者在整个过程中生成的数据在大规模上可能是大量的，涵盖了广泛的数据字段和类别（表 1）。

除了前面提到的 3 个主要类别外，IH 临床实践还可以生成以下方面的数据：

1. 患者报告结果（PRO）：包括疼痛量表、情绪评估和睡眠日记等评估

2. 临床实验室检查：涉及血液检查和影像学检查等程序

3. 患者参与和遵守治疗计划：监测患者积极参与和遵循治疗计划的积极性和一致性

4. IH 干预的成本效益：评估 IH 方法的经济效率和价值

5. 不良事件或副作用：跟踪 IH 疗法导致的任何不良反应或不良影响

IHC 代表了医疗保健模式的范式转变，从以疾病为中心的方法转变为整体方法，通过整合更广泛的数据源、促进跨学科合作以及可能利用 AI 获得个性化见解来优先考虑患者的健康。IHC 的优势在于其跨学科性质，各种提供者（医生、治疗师和营养师）在这里合作并不断从患者数据中学习。这种协作方法在促进创新和新干预措施的同时，创造了一个复杂的数据环境。由于 IHC 收集的数据范围很广，包括病史、行为、SDOH，甚至患者制定的健康计划（并且经常使用 AI 进行分析），因此这些组合数据集的所有权以及从中得出的潜在新知识变得不清楚。此外，AI 分析的过程本身就产生了额外的复杂性。当 AI 在这些丰富的患者数据中识别模式和趋势时，可能会产生新的知识。谁拥有这些“衍生数据”？所有权是属于提供原始信息的患者，还是属于开发 AI 创建见解的平台？

数据所有权的这些不确定性可能会阻止患者充分参与 IH 计划，因为他们担心失去对个人健康信息的控制。明确的数据所有权政策和法律框架对于驾驭这些复杂性至关重要。医疗保健利益相关者（包括患者、护理人员、提供商和医疗保健系统）需要了解他们的数据是如何被使用的，谁可以访问这些数据，以及用于什么目的。只有这样，IHC 才能在确保患者信任和隐私的同时，释放其整体和个性化医学潜力，改善健康结果。接下来我们将从实践的角度剖析数据所有权问题。

表 1. 临床实践中生成的数据类型

类别	例子
患者信息	此类数据可以包括人口统计详细信息、病史、家族史、生活方式因素和 SDOH[a]。
治疗详情	这包括提供的综合健康疗法，包括针灸、按摩疗法、草药、营养咨询和正念练习。它还包括有关这些治疗的频率、持续时间和强度的信息。
患者结局	这涉及评估综合健康干预对患者健康和福祉的影响。相关结局指标包括症状管理、生活质量、功能状态和对护理的总体满意度。

[a] SDOH：健康的社会决定因素。

植根于 IHC 实践的数据所有权问题

数据所有权问题可能出现在医疗保健的所有阶段，包括 IHC 实践过程，从初始评估到制定治疗计划再到监测患者的进展。

评估阶段

在评估阶段，IHC 从业者可能会从患者的电子病历（EMR）、诊断测试和问卷中收集数据。与其他医学领域一样，从 EMR 和测试收集的数据通常用于规划治疗方案。在 IHC 设置中，PRO 数据或真实世界数据比传统模型具有更重要的作用，因为 PRO 数据可能有助于共同决策，并与 IHC 方法的入组有关。但是，从数据所有权的角度来看，这些数据可能包含有关患者健康状况和生活方式的敏感信息。

例如，IHC 从业者可能会询问患者的饮食、运动习惯、睡眠模式、压力水平以及当地政策或法律可能没有规定的草药和其他补充剂的使用情况。当某些患者遇到可能不想与其他利益相关者分享以避免潜在麻烦的情况时，这些信息的敏感性可能会给一些提供者带来负担。例如，患者可能不愿透露药物滥用史或心理健康问题，担心受到雇主或保险公司的歧视。这在 IHC 的背景下尤其令人担忧，因为扩大对数据的访问增加了患者的风险，这些患者可能会因过去的医疗决定而面临负面后果，例如拒绝寻求治疗或不依从。这些数据对于从业者为患者制定准确有效的治疗计划至关重要确定这些数据如何用于医疗诊断、患者如何授权提供者访问这些数据以及患者如何控制或参与数据传输和使用提出了重大挑战。但是，同样重要的是要注意

这些数据是敏感和机密的。患者有权知道他们的数据是如何被使用的，并有权控制谁可以访问这些数据。

治疗计划制定

一旦医生评估了患者的健康状况，他们将制定治疗计划。如上所述，该计划可能包括传统实践和 IH 实践的组合。例如，IHC 从业者可能会建议慢性疼痛患者联合服用非处方止痛药、针灸和瑜伽。医生还可能建议饮食改变和压力管理技巧。

治疗计划可能要求患者与医生分享其他数据，例如他们的治疗反应或进展。其中许多数据可能没有分类，或者可能是其他医学领域没有接触过的创新探索性工作，这反映了整个“以人为本”的护理模式。这些数据可以通过新的干预措施生成，而患者和跨学科提供者之间制定的治疗计划结果可能没有终极标准。这里另一个具有挑战性的部分是了解与一个人的生活相关的不同数据级别。除此之外，许多数据可能与 SDOH 相关联以及其他尚未发明或发现的测量方法，这可能会在未来带来隐私挑战。例如，基因数据的整合或持续的环境监测可以创建更详细的档案，从而引发关于谁可以访问、如何使用以及潜在的歧视的新问题。

监控阶段

在监测阶段，医生将跟踪患者的进展并根据需要调整治疗计划。这可能涉及从患者那里收集额外的数据，例如他们的症状、生活质量和对治疗的满意度。

这些数据对于从业者确保患者得到最好的护理至关重要。但是，同样重要的是要注意这些数据是个人的、敏感的和机密的;与治疗阶段生成的数据一样，它们可以由多个提供商收集。它还可能包含从患者家属和照护者那里收集的健康数据，包括可能包含患者家属个人信息的定性数据。因此，患者有权知道他们的数据是如何被使用的，并有权控制谁可以访问这些数据。

IH 设置中的 AI

事实证明，AI 技术是许多医疗保健领域的强大工具，引领了医疗保健服务的转变，重点关注患者及其整体福祉。AI 具有通过改善患者治疗效果、提高效率和转

变医疗保健服务而彻底改变 IHC 的巨大潜力。AI 的主要贡献之一在于实现个性化医疗。通过分析包括病史、遗传学、生活方式和环境因素的患者数据，AI 可以帮助 IHC 提供者根据个人需求定制治疗计划，确保最佳护理。此外，AI 可以分析患者数据以识别表明潜在健康状况的模式或标志物，从而促进更早、更准确的诊断。AI 的功能扩展到通过提供患者数据的实时分析和提供相关的治疗方案来改善治疗结果，使 IHC 从业者能够做出明智的决策。此外，AI 可以自动执行日常任务，例如安排预约和管理患者记录，减轻从业者的管理负担，并允许他们将更多时间投入到患者护理中。

AI 和 IHC 的重大突破主要集中在优化治疗模型，包括 AI 辅助针灸，中医通过唇舌分析进行诊断和中医证候鉴定。此外，该研究还探讨了人工智能在正念练习中的应用和药物依从性，利用 EMR 和自然语言处理改善中西医结合中肺病的证候模式诊断。虽然这些临床试验和应用显示出可喜的进展，但除了改进现有模型（例如患者教育和 AI 驱动的症状分析）之外，其他尝试还有充分利用 AI 在 IHC 中的优势。总而言之，AI 在 IHC 中的潜在应用之后可以进行 IHC 的 3 个阶段——评估、治疗和监测。

评估阶段

个性化风险评估

AI 可以分析大量患者数据，包括遗传、生活方式和环境因素，以识别患慢性病或不良健康结果风险较高的个体。这种个性化的风险评估可以指导预防性医疗保健措施和早期干预。

症状分析和模式识别

AI 驱动的工具可以分析患者报告的症状、病史和临床数据，以识别模式和潜在的潜在疾病。这可以帮助临床医生做出更准确的诊断并相应地定制治疗计划。

心理健康评估和筛查

基于 AI 的聊天机器人和虚拟助手可以与患者进行对话，以评估他们的健康状况并识别抑郁、焦虑或其他心理健康问题的潜在迹象。这可以促进早期干预和支

持。

治疗阶段

个性化治疗计划

AI 可以分析患者数据和临床指南，以根据个体因素（包括遗传易感性、既往治疗和共存条件）生成个性化治疗计划。这可以优化治疗效果并最大限度地减少副作用。

药物剂量优化

AI 可以分析患者数据和药物概况，以确定处方药的最佳剂量，降低药物不良反应的风险并改善治疗结果。

营养指导和膳食计划

AI 驱动的工具可以分析个人饮食需求、偏好和健康目标，以提供个性化的营养指导和膳食计划建议，支持健康的生活方式和疾病管理。

监控阶段

实时远程监控

支持 AI 的可穿戴设备和传感器可以持续收集患者数据，例如生命体征、活动水平和睡眠模式，并将其传输给医疗保健提供者进行实时监测。这允许及早发现潜在的健康问题并及时干预。

疾病恶化的预测分析

AI 可以分析患者数据并识别预测潜在疾病恶化或不良健康事件的模式，从而实现主动干预并预防并发症。

患者参与和依从性支持

AI 驱动的聊天机器人和虚拟助手可以与患者互动、提供提醒并提供个性化支持，以提高药物依从性和生活方式的改变，提高治疗结果。

在 AI 结合的 IHC 中生成的数据

随着 AI 在 IH 临床实践中的潜在应用，生成的数据类型可以扩展到传统的患者信息和治疗细节之外。以下是可通过 AI 集成创建的数据的一些示例：

1. 患者生成的健康数据：AI 可以分析来自可穿戴设备、健身追踪器和患者报告的症状追踪器的数据，以提供对患者生活方式、睡眠模式和整体健康状况的见解。这些数据可用于个性化治疗计划和监测患者进展。

2. 实时生物反馈数据：AI 可以分析来自测量心率变异性、皮肤电导和其他生理信号的设备的生物反馈数据。这些数据可用于评估患者的压力水平、焦虑和疼痛，从而允许实时调整 IH 干预措施。

3. 基因组和蛋白质组学数据：AI 可以分析遗传和蛋白质表达数据，以确定药物代谢、疾病易感性和对 IH 疗法反应的个体差异。这些信息可以定制治疗计划并预测潜在的不良反应。

4. 预测分析：AI 可以分析历史数据和患者特征，以预测未来健康事件或治疗结果的可能性。此信息可用于主动识别有风险的患者并定制预防保健或治疗计划。

5. AI 驱动的治疗建议：AI 可以分析患者数据和临床指南，以提供个性化的治疗建议，包括 IH 疗法的类型、剂量和频率。这可以简化治疗计划并提高患者的依从性。

6. AI 驱动的临床决策支持：AI 可以为医疗保健提供者提供实时临床决策支持，根据患者数据和循证指南建议适当的 IH 疗法。这可以增强临床决策并改善患者护理。

7. AI 驱动的研究和临床试验：AI 可以促进 IH 临床试验的设计、分析和解释，从而加快循证实践的进步。

IHC 或 AI 设置中的数据所有权问题

由于多种因素，IHC 和 AI 数据收集引发了复杂的所有权问题。首先，个人对这些系统的贡献往往交织在一起，因此不清楚谁真正“拥有”结果数据。其次，机器生成的数据和 AI 衍生的见解引入了关于谁拥有这些智力创作的权利的新问题。最后，版权和隐私等传统法律框架难以适应 IHC 和 AI 的独特动态，导致所有权模糊，并可能引发争议。此外，还有一些其他挑战：

1. 同意和控制的模糊性：与 IHC 和 AI 系统交互的人可能难以理解他们的数据是如何被收集、使用和共享的。同意机制可能不透明，使用户不确定他们是否保留了对其信息的任何控制权。

2. 难以归因作者和创造力：随着 AI 系统越来越多地为数据生成和分析做出贡献，确定谁应该为生成的见解获得荣誉变得具有挑战性，是提供初始数据的人、创建 AI 的开发人员还是 AI 本身？

3. 平衡个人权利与集体利益：虽然通过 IHC 和 AI 收集的数据可以提供社会效益，例如改善医疗保健或个性化服务，但这些优势可能会以牺牲个人隐私和自主权为代价。在这些相互竞争的利益之间取得平衡仍然是一项重大挑战。

4. 利用和偏见风险：不伦理的行为者可能会利用数据所有权的模糊性来操纵或歧视个人。在扭曲的数据集上训练的有偏见的算法会进一步延续这种不公正。

5. 国际复杂性：不同司法管辖区的数据所有权法律和法规差异很大，给全球 IHC 和 AI 项目带来了挑战。这可能会导致混乱并阻碍负责任的数据治理。

解决这些复杂问题需要技术开发人员、政策制定者、法律专家和公众之间的持续合作。我们可以通过公开对话和创新解决方案确保公平的数据所有权和负责任的 AI 开发，使所有人受益。

谁有权拥有这些数据？

在 AI 和机器学习时代，数据比以往任何时候都更加重要。在 IHC 中尤其如此，其中 AI 或机器学习可用于开发新疗法、改善患者护理和进行研究。然而，IHC 数据的所有权是一个复杂的问题。有几个利益相关者可能会声称拥有 IHC 数据的所有权。

首先，患者可能会争辩说他们拥有自己的数据，包括来自他们的医疗记录和诊断测试的数据。传统上，一旦与更广泛的受众共享去标识化数据，患者对数据的控制就有限了。由于立法强制要求患者隐私，医疗保健提供者、机构和管理机构制定了确定患者访问和控制其个人健康信息的能力的政策和做法。

其次，IHC 提供商可能会争辩说，他们拥有从患者互动中获得的数据，例如来自临床记录和患者门户的数据。此外，如果从新的干预或临床试验中收集，提供商可以要求数据所有权。患者和提供者之间的互动也很有意义，因为患者可以拒绝以

任何研究身份共享数据。虽然数据隐私法允许患者控制谁拥有他们的数据以及如何使用这些数据，但对现有数据持有的有限了解造成了信息不对称，阻碍了他们充分行使这些权利的能力。因此，医疗保健提供者可能会在一开始就主动让患者退出广泛的研究计划，以解决患者对数据重用的有限控制问题。虽然选择退出可以防止将来的数据收集，但它不一定会删除研究人员、政府机构或私人实体持有的现有数据。这造成了这样一种情况，即由于对哪些实体持有其数据的了解有限，患者可能难以行使其数据隐私权。

第三，研究人员可能会争辩说他们拥有来自研究的数据，包括来自 IHC 患者数据和干预措施的数据。许多临床提供者参与线性研究活动，这些活动通常遵循从识别疾病到开发干预措施的医疗保健流。在考虑人工智能在临床实践中的应用时，研究人员可以要求对开发的算法拥有所有权（通过专利）;但是，在某些情况下，他们可能会要求对研究试验和项目中使用的数据拥有所有权。

第四，AI 和机器学习开发人员可能会争辩说，他们拥有用于训练 AI 或机器学习算法的数据，包括 IHC 患者数据，因为根据其可用性和审批流程，可以购买或许可一些去标识化的个人数据用于研究。但是，对其他敏感健康数据的访问，尤其是从 IHC 设置的访问，通常受到更多限制。这些数据集可能需要研究伦理委员会的特别批准才能获得访问权限。

最后，科技公司可能会争辩说，他们拥有通过其可穿戴设备和其他健康跟踪应用程序收集的数据，包括 IHC 患者数据。多家公司可以同时对收集的相同数据提出索赔。

与 IHC 相关的数据所有权模型

医疗保健领域的数据所有权模型是一个复杂且不断发展的话题。有多种不同的模型，每种模型都有其优点和缺点。本节介绍了医疗保健中一些最常见的数据所有权模型。

私有化假设

源自约翰·洛克（John Locke）的自然权利理论，医疗保健数据所有权的私有化假设数据是由个人或组织拥有和控制的有价值的私人资产。在 IHC 和 AI 合作的

背景下，这种模型引发了人们对私营实体可能将 IHC 数据货币化的担忧。这种做法可能会阻碍跨学科合作，因为对数据保护的担忧可能会限制医疗保健提供者之间的信息共享。此外，在私有化假设下开发 AI 可能会导致专有算法，限制其可访问性并阻碍 IHC 治疗的集体进步。对个人或组织所有权的关注可能会对无缝共享见解和创新造成障碍，阻碍 IHC 在 AI 时代的协作潜力。

尽管存在这些挑战，但私有化假设确实提供了优势。它认识到医疗保健数据的经济价值，可能会激励个人和组织投资于数据收集和分析。这可能会导致个性化医疗保健解决方案和定制治疗计划的进步。然而，缺点在于对 IHC 领域的协作、数据可访问性和集体进步的潜在负面影响。在承认数据作为资产的价值和促进协作努力之间取得平衡，对于在这种所有权模式下成功将 AI 集成到医疗保健中至关重要。

公有化假设

与私有化假设不同，公有化假设将数据视为可以公开共享的公共产品，并且数据可以同时合法地使用。在 IHC 和 AI 的背景下，该模型强调跨学科医疗保健提供者、研究人员和患者之间的协作和协调。共享数据平台和开源 AI 的概念相结合，以改善资源使用，从而改善患者预后。当该模型设想一种更具包容性和集体性的医疗保健数据方法时，可能会出现挑战。对负责任和合乎伦理的 AI 使用以及公平分享利益的担忧需要仔细考虑。在开放合作和解决伦理问题之间取得平衡，对于实现在公有化假设下充分设想的积极成果来说，势在必行。

公有化假设的优势在于它有可能打破数据孤岛，促进医疗保健提供者之间的无缝数据共享和可访问性。这种协作环境可以促进创新，从而实现更有效的 IHC 治疗。然而，该模型也提出了伦理考虑，例如确保负责任地、公平地使用数据。实现这种平衡对于在 IHC 中成功实施公有化假设至关重要，确保共享数据的好处扩展到所有利益相关者，同时维护伦理标准。

知识产权

健康数据的所有权可以是有形和无形财产。关于有形财产权，答案有时是肯定的。例如，在美国，可能会说医疗服务提供者而不是患者通常拥有物理医疗记录。同时，健康数据是无形信息。相关利益相关者可以拥有基于知识产权领域不同类型法律的健康信息，包括专利法、版权法和数据库中的版权、商标法和商业秘密。但

是，此类所有权保护必须满足各种标准，从而导致健康数据的明确性和不完整或部分所有权保护。比如健康数据应符合专利条件才能享受专利保护。此外，商业秘密或相关机密信息法律适用于有限类型的健康数据，并且关于健康数据所有权的几个问题仍未解决。此外，在 AI 背景下，通过知识产权法授予所有权甚至更加复杂，例如 AI 主张知识产权的能力、确定人类与 AI 之间的贡献等。因此，在知识产权法的背景下回答有关 AI 生成的健康数据的所有权问题是高度复杂和不确定的。

接下来，我们将讨论当前与 IHC 相关的医疗保健数据中的数据所有权模型。了解这些模型的优缺点有助于解决 AI 时代 IHC 数据所有权中不断凸显的问题和冲突。

分布式访问控制模型

分布式访问控制（DAC）模型提供了一种去中心化的数据所有权方法，为个人医疗保健提供者或组织提供了对其数据的更多控制权，尤其是在 IHC 和 AI 的背景下。此模型解决了与医疗保健数据中的隐私和安全相关的关键问题。通过允许实体通过基于角色的访问控制或基于属性的访问控制等机制控制对其数据的访问，DAC 模型旨在保护敏感的患者信息。然而，强调个体控制可能会导致提供者之间数据共享的挑战，导致护理碎片化，并可能阻碍 IHC 跨学科领域的医学研究进展。

虽然 DAC 模型有助于缓解隐私和安全问题，但它引入了与数据孤岛和有效信息交换障碍相关的复杂性。DAC 下数据所有权的碎片化性质可能会阻碍 IHC 中的协作努力，限制对患者健康状况的全面了解，并可能损害治疗的有效性。此外，由于研究人员需要获得拥有数据的每个提供商或组织的许可，因此可能会出现研究访问方面的困难。平衡个人控制与无缝协作和研究访问的需求对于在 IHC 中有效实施 DAC 模型至关重要，尤其是在 AI 时代。

数据信任

数据信托作为为多个利益相关者持有数据的法人实体，在 IHC 和 AI 领域为 DAC 和社区化模型提供了替代方案。在这种情况下，数据信任为利益相关者提供了对数据的更多控制，促进了负责任和合乎伦理的使用。建立中立且值得信赖的第三方来管理数据有助于解决数据所有权、隐私和安全问题。然而，挑战仍然存在，尤其是在 IHC 中建立和维护数据信任的复杂性和成本方面。创建和维护这些法人实体所涉

及的复杂性可能不适用于所有 IHC 提供商，从而限制了该模型的普遍采用。

尽管具有改进数据共享和协作等潜在优势，但数据信托在协调信托和利益相关者的利益方面可能面临困难。可能会出现数据所有权和使用方面的冲突，这凸显了为 IHC 和 AI 领域数据信托的运作建立明确的指导方针和框架的重要性。此外，由于其复杂性和不透明性，让数据信托对其行为负责可能具有挑战性。在数据信任的好处和挑战之间取得平衡对于它们有效集成到 IHC 环境中至关重要，确保它们为 AI 时代的数据管理和使用做出积极贡献。

区块链技术

区块链技术成为一种很有前途的解决方案，用于安全透明的数据所有权记录，尤其是在 IHC 和 AI 的背景下。在医疗保健领域，区块链可以提高透明度、问责制和数据共享，从而降低违规和其他安全事件的风险。然而，对区块链技术的可扩展性和可靠性的担忧仍然存在，主要是当其应用到 IHC 固有的跨学科合作中管理大量医疗保健数据时。区块链相对较新的性质为其广泛实施和集成到现有医疗保健系统带来了不确定性。

区块链在 IHC 中的优势包括它有可能创建一个不可变和防篡改的账本，确保医疗保健数据的完整性。这对于维护准确的患者记录和支持医疗保健提供者之间的协作工作特别有益。然而，实施区块链技术的复杂性和成本可能会带来挑战，尤其是对于资源有限的小型 IHC 提供商。缺乏明确的监管框架增加了复杂性，在 IHC 环境中引入了数据所有权和使用的不确定性。此外，普通法和民法的数据隐私法通常与公共区块链不兼容，因为任何人都可以看到存储在其中的信息。这种透明度可能是敏感数据的主要问题。为了帮助开发人员应对这一挑战，美国国家标准与技术研究院（National Institute of Standards and Technology）创建了一个流程图来确定合适的区块链用例。在利用区块链的优势和应对挑战之间取得平衡，对于将其成功整合到不断发展的 IHC 和 AI 环境中至关重要。

重新思考 AI 时代 IHC 实践的数据所有权框架

虽然 IH 通过协作和 AI 驱动的洞察力为患者的整体健康提供了有希望的转变，但它创建的复杂数据环境需要一个强大的数据共享模型。目前对组合数据集和 AI

衍生知识的所有权缺乏明确性，这阻碍了患者的参与并阻碍了进展。解决这些问题不仅涉及技术解决方案;它需要一个数据所有权模型来保护数据所有权。

从法律角度进一步澄清和规范数据所有权对于考虑建议和法律行动至关重要。财产不是客体本身，而是通过 McGuire 等人的聚合法律利益来主张控制的能力指出。所有权包括合法权利，包括但不限于占有、访问和控制。尽管出于公共利益、缺乏市场失灵、基本权利和交易成本等原因，对在数据中建立财产权持相反观点，及时分配健康数据所有权至关重要。这一步对于积极激励高质量、高效地生成、传播和使用医疗数据，从而为 IH 环境中的 AI 发展提供动力是必要的。

同时，应限制健康数据的所有权，以平衡各利益相关者的利益并适当降低交易成本。限制不同实体的所有权符合该法律促进社会进步和保持平衡的宗旨。如前所述，患者、IHC 提供者、研究人员以及 AI 或机器学习开发人员都在 AI 时代为 IHC 实践数据做出了贡献。这些相关方之间的共同所有权是必不可少的，但向每个利益相关者授予完全所有权可能会显着增加交易成本，从而可能阻碍健康数据在临床环境、研究和 AI 领域的应用。由于涉及众多权利和利益持有人的复杂性，提供建议具有挑战性。

一个建议的框架是仅授予患者对其个人健康数据的所有权。由于没有个人信息的健康数据与患者的联系较少，并且与大量患者打交道会大大增加交易成本，因此将所有权限制为个人健康数据是一个谨慎的选择。此外，可以参考《通用数据保护条例》（GDPR）第 5-22 条来探讨有关患者个人数据所有权的具体法律规则。根据 GDPR，患者可以请求访问自己的数据并拥有一定程度的控制权（例如，删除数据或要求不与第三方共享）。但是，它们没有数据所有权，这类似于其他隐私框架。另一个关键方面是为患者、IHC 提供商、研究人员以及 AI 或机器学习开发人员等利益相关者建立有关健康数据所有权的限制或例外。这些限制与例外有助于平衡利益相关者之间的利益，从知识产权法中的各种合理使用模式中汲取灵感。

在 AI 时代的 IHC 背景下为利益相关者定义所有权和限制是具有挑战性的，尤其是在 AI 出现的情况下。识别和分配数据所有权（例如根据比例或主要贡献者确定所有权）是持续的挑战。必须考虑患者的权利和隐私、医生的努力、AI 从业者的参与和时间投入以及公共利益。只有通过这种全面的方法，才能促进医疗发展和各种利益的平衡，这与法律法规的目的和精神相一致。虽然这只是许多可能建议的一个方面，但它至关重要，因为这些问题的法律明确性将直接影响其他建议的建立

和实施，以解决 AI 时代 IHC 实践中的数据所有权问题。

总之，我们提出了协作医疗保健数据所有权（CHDO）框架。CHDO 强调利益相关者共同努力时数据的集体力量。它承认，从患者到提供商、研究人员和 AI 开发人员，各方都为医疗保健数据贡献了宝贵的见解。CHDO 框架通过提出三个关键功能来解决这个问题：

1. 共享所有权：CHDO 框架超越了传统的所有权模型，在这种模型中，一个实体拥有专有权。相反，它倡导共同所有权，根据利益相关者的贡献授予利益相关者对数据的特定权利和责任。这促进了信任并激励了协作，释放了数据在研究、开发和个性化护理方面的全部潜力。

2. 定义的访问和控制：CHDO 框架主张为访问和使用数据建立明确的指导方针。患者保留对其个人健康信息的控制权，而其他利益相关者可以访问匿名或汇总数据以用于批准的目的。这种平衡确保了个人隐私，同时促进了医疗保健的集体进步。

3. 公平透明的治理：CHDO 框架认识到需要强大的治理结构。透明的政策和程序可确保公平访问、防止滥用并解决潜在冲突。这促进了所有利益相关者之间的信任和问责制，为数据驱动的医疗保健进步创造了可持续的环境。

基于前几节中介绍的 IHC 和 AI 背景下的各种数据所有权模型的分析，CDHO 共同所有权模型与其他框架相比具有多项优势。它解决了私有化假设提出的担忧，即通过授予患者对其个人健康数据的所有权来影响公共利益。此外，它还减轻了社区化假设和独立 DAC 模型可能出现的伦理问题，例如隐私和安全问题。此外，CDHO 模型避免了在 IHC 和 AI 背景下基于专利法、商业秘密或相关机密信息法、版权法和商标法的有限所有权保护。需要注意的是，其他两种模型，即数据信任和区块链技术，主要涉及医疗保健数据的管理和存储，可能会产生大量成本。在这些模型中，利益相关者之间的冲突可能会在缺乏明确的所有权的情况下持续存在，并且没有明确的解决这些冲突的指导。无论使用信托还是区块链，建立它们的先决条件是明确识别拥有建立信托或区块链的所有权的一方。CHDO 模型通过明确定义和平衡各方利益、确保个人隐私和安全以及促进公共利益的实现，有效地解决了这个问题。

这些优势在 IHC 和 AI 的背景下尤为重要。IHC 的协作性质和对以患者为中心的护理的关注需要一个数据所有权模型来培养信任并激励协作（共享所有权）。此

外，平衡个人隐私与数据研发潜力的需求与 CHDO 框架定义的访问和控制机制非常一致。最后，CHDO 框架强调公平和透明的治理，这对于解决围绕医疗保健中使用 AI 的复杂伦理考虑至关重要。通过实施这些原则，CHDO 框架开启了协作的新时代，为利益相关者赋能并为所有人营造更健康的未来，最终推动医疗保健行业走向数据驱动和 AI 集成的未来，优先考虑个人权利和集体进步。

结论

虽然通用解决方案可能很难，但要应对 AI 驱动的 IHC 中的数据所有权挑战，还需要根据特定的利益相关者需求和法规定制方法。通过确保数据使用让个体患者受益、遵从法律框架并促进社会福祉，这种合作可以释放 AI 在 IHC 中的全部潜能，同时降低法律风险。从本质上讲，解决 IHC 中的数据所有权问题可以为 AI 在医疗保健中更加简化、有效和合乎伦理的整合铺平道路，最终扩大其对整个社会的好处。

译文二：

基于二维混沌映射和区块链的安全医疗数据共享方案

Zhigang Xu, Enda Zheng, Hongmu Han , Xinhua Dong,
Xiaohong Dang, Zhongpeng Wang,徐健（译）

简介

健康保险是国民基本医疗保健的重要保障，对社会稳定和公共利益至关重要。随着医疗保险参与者数量的逐年增加，这些参与者产生的大量医疗保健数据尚未得到有效利用。据报道，重要的是要解决大量医疗相关数据的有效流通问题，同时平衡数据提供者的隐私和权利，并为医学研究界提供更好的数据支持。因此，有必要开展安全高效的医疗保健数据共享研究工作。

数据共享过程中的数据隐私和安全问题是阻碍医疗保健数据共享发展的主要障碍。随着医疗大数据的发展，在线医疗保健结算系统、分布式健康管理服务以及基于联邦学习的医疗保健模型的训练等应用程序都需要大型数据集作为基础支持。这需要在不同机构之间共享本地数据集，这一过程面临许多数据安全问题。例如，犯罪分子利用数据共享中的漏洞非法访问患者的敏感私人数据。此外，他们可以从数据中推断出区域人口的健康状况和医疗财政支出，这可能会对社会安全构成威胁。共享数据被恶意篡改可能会导致研究工作因错误数据而无法达到预期结果，从而导致大量时间和经济损失。此外，以集中式聚合服务器为中心的数据共享网络带来了与隐私、所有权和法规相关的挑战。由于医疗保健数据共享涉及敏感信息，因此提出新的解决方案来补充传统方案中缺乏的数据验证机制（data verification mechanisms）和访客身份验证机制(visitor authentication mechanisms）非常重要。

目前，许多学者提出了基于密码学的加密技术，用于在数据共享中保护隐私。混沌映射是一种基于混沌理论的加密方法，为传统的公私钥系统提供了替代方案。由于混沌系统的伪随机性和高灵敏度，它可以生成高度复杂和随机的键序列。此外，它还表现出对差分和线性攻击的很强抵抗力。Naik，R.B.等人建议使用基于混沌映

射的伪随机数生成器，该生成器可用于构建更安全的加密算法，以防止未经授权的用户。这项工作将现有的一维逻辑图与一维混沌图相结合，提出了一种新的二维混沌图。基于此，设计了一种加密机制来增强数据共享的安全性。

区块链支持在多个参与者之间建立可靠且去中心化的数据存储和交换系统。它的去中心化性质减少了对第三方参与的需求，并降低了与数据共享相关的交易成本。使用区块链技术而不是传统的中心化网络显然更适合当前医疗保健数据共享的需求。目前，区块链还被应用于其他各种数据共享领域，例如车联网、交通流量预测和工业物联网。然而，现有的医疗保健数据共享解决方案仍然存在缺点：（1）复杂的密码系统在实时通信和大规模加密场景中需要更多的计算成本。（2）现有的解决方案未能很好地集中数据完整性验证和身份验证这两个问题。

针对上述问题，本研究提出了一种基于二维混沌映射的加密方法，以增强医疗数据共享的数据安全性。本文的主要贡献如下：

- 提出一种新的二维混沌映射，称为二维无限逻辑映射（2D-ILM），并分析其动态行为。
- 基于 2D-ILM 设计了一种称为二维混沌映射加密（2CME）的数据加密方法。它在数据所有者提供的生物声学信息和共享数据之间建立了很强的耦合相关性。
- 设计了一个基于二维混沌映射的数据共享方案，命名为 2DCM-DS。该方案利用智能合约实现数据完整性验证和身份验证功能。它具有出色的抗攻击性、低耗时和稳健性。

本文的其余部分如下。“相关文章”章节回顾了一些相关作品。建议方案内容在“系统模型”部分中定义。“方法”详细描述了方案的定义。“安全分析”分析了方案的安全性。方案的性能在“系统评价”中进行评估。最后，“结论”章节给出了整个工作的总结。

相关文章

医疗保健和数据共享

大数据在医疗保健领域的现有应用主要涉及疾病预测和欺诈性保险索赔检测。Chen 等人根据上海五年的医疗保健记录对癌症和心房颤动人群的特定部位癌症特

征进行了评估，揭示了癌症患者心房颤动患病率的增加，男性占主导地位，并且癌症发病率有年轻化趋势。Anand 等人用来自马萨诸塞州和 11 个美国医疗保健组织的医疗索赔数据比较了医疗保健患者的电子健康记录（EHR）数据的完整性，以减轻因 EHR 不连续性而导致的研究变量错误分类。这些研究侧重于医疗大数据在医疗保健领域的应用发展，但显然，这些研究并未关注底层应用的数据安全以及数据交互过程中缺乏隐私保护机制。

其他医疗保健大数据的研究侧重于开发更安全、更高效的数据共享方法。A.Awad Abdelatif 等人提出需要开发更快、更安全的数据共享，以将医疗保健系统内的各种医疗实体聚集在一起，从而提高医疗质量并实现对重大医疗事件的更准确控制。Rani 等人强调了医疗物联网（IoMT）设备和聚合服务器之间数据共享安全性在医疗领域联邦学习发展中的重要性。然而，集中式网络结构存在潜在的数据安全风险。方案利用基于点阵密码的安全关键字可搜索属性加密来实现安全高效的医疗保健大数据管理，但它需要解决区块链中电子病历管理和共享带来的安全问题。如何在基于区块链的数据共享网络中实现隐私保护和高效的数据共享，是医疗大数据应用发展的悬而未决的问题。

2D 混沌地图和隐私保护

混沌映射在行为上表现出随机性和不可预测性，对初始条件高度敏感，因此在加密音频、图像或时间序列等数据方面有广泛的应用。将参数扩展到二维空间，二维混沌映射可以更好地保持混沌状态下参数的连续性。Wu 等人提出了 2D-LNIC，基于该算法，AEA-NCA 在音频加密方面表现出出色的性能，显著降低了数据流相关性。Jasra 等人将 DNA 编码与二维混沌系统集成在一起，提出了一个全面的彩色图像加密框架，该框架具有高效、稳健和弹性，可抵御多种类型的攻击。这两种方案提出了有效的加密算法，但没有讨论所提出的加密算法与实际应用场景的结合。同时，许多学者探索了利用混沌映射的数据传输方案。Kumari 和 Punam 提出了一种轻量级加密（LWE）方案，用于保护物联网（IoT）环境中的图像数据传输，采用分段线性混沌映射（PWLCM）和粒度密钥流生成器（GKSG）。在方案中，Alsahafi、Y.S 等人通过一种新的二进制冠状病毒疾病优化算法获得了混沌映射的最佳初始序列，该算法为医学图像加密生成了最佳初始密钥。增强医疗物联网中医学影像传输的数据安全

性。虽然他们的方法直接将混沌映射应用于数据加密，增强了数据安全性，但还需要进一步探索，以加强用户与共享数据之间的连接，同时实现数据隐私保护。

区块链

随着大数据模型的飞速发展，业界对隐私保护数据共享方案的需求越来越大。在医疗保健的 IoT 领域，Yang 等人提出了一种基于动态共识委员会的安全数据共享方法，为公有云提供细粒度的数据安全隐私保障。后来，在方案中，他们设计了一种基于声誉激励机制的快速权限证明方法，用于物联网医疗保健中的匹配加密，但所提出的方案需要额外的计算来解决匹配加密的单个密钥暴露的潜在问题。Hong 等人提出了一种安全的点对点多方交易，通过在区块链方案中加入秘密共享机制来减轻多方交易场景中的通信负担，实现了用户身份的认证，但缺乏链上数据隐私保护机制。Li 等人注意到医疗数据共享中存在隐私泄露和共享效率低下的问题，并基于区块链构建了轻量级隐私保护模型，但这反过来又缺乏与用户的监管机制。因此，区块链技术的加入可以为医疗数据共享提供更好的数据安全性，如何提出一种安全有效的方法来集中解决医疗保健数据共享中身份验证和数据完整性验证的解决方案是一个值得研究的问题。

针对上述问题，我们提出了一种基于二维混沌映射的稳健耦合加密认证机制，利用区块链和智能合约的安全功能，旨在增强医疗保健数据共享的安全性。基于上述讨论，表 1 列出了与本文工作相关的一些方案比较。

表 1. 与其他医疗保健数据共享方案的比较

方案	隐私保护	数据完整性验证	身份验证
参考[27]， 参考[28]	✓	✗	✗
参考[32]， 参考[34]， 参考[35]	✓	✗	✓
参考[33]， 参考[36]， 参考[37]	✓	✓	✗
我们的	✓	✓	✓

系统模型

针对如何集中高效解决医疗数据共享中的数据完整性验证和身份验证问题，本

文设计了一种基于区块链的医疗数据共享方案 2DCM-DS，该方案在数据所有者和共享数据之间建立双向强耦合关联。链上数据完整性验证和身份验证功能以智能合约的形式实现。表 2 定义了 2DCM-DS 中使用的关键符号。

表 2. 2DCM-DS 中关键符号的定义

表示法	定义
DO	数据所有者
DR	数据请求者
CBC	联盟链
IVA	身份验证音频
SD	共享数据
PIVA, PSD	IVA 和 SD 的明文
CIVA, CSD	IVA 和 SD 的密文
Fa, Fb	PIVA 和 PSD 的干扰系数
Fc, Fd	来自 CSD 的两个干扰系数

2DCM-DS 的方案模型包括三种类型的实体：数据所有者（DO）、数据请求者（DR）和联盟链（CBC）。该方案以联盟链为网络载体，将设计好的加密、解密、数据完整性验证和身份验证功能封装到部署在区块链上的智能合约中。图 1 说明了 2DCM-DS 模型。

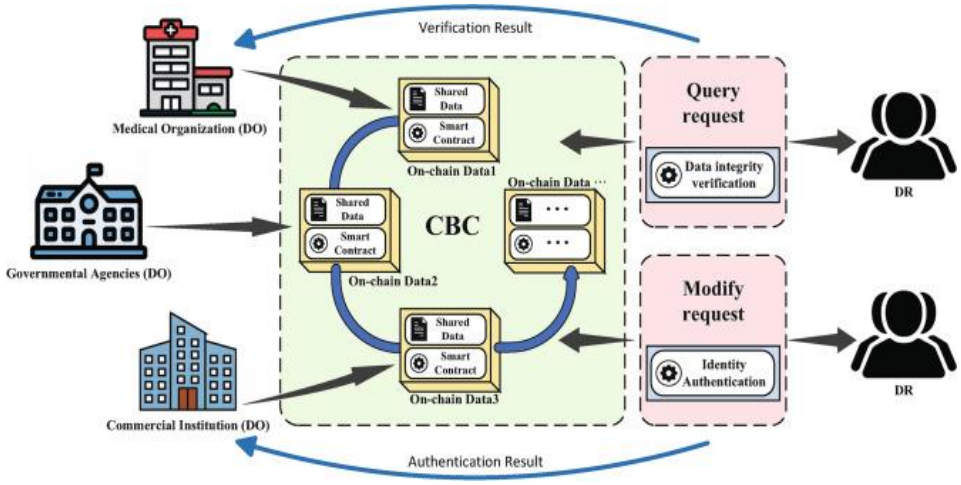

图 1. 2DCM-DS 模型的示意图

数据所有者（DO）是指拥有医疗保健数据集的组织，例如医疗保健机构、政府机构或商业实体。在 2DCM-DS 中，DO 作为医疗保健数据的所有者参与数据共享过程。

他们需要提供自己的身份验证音频（IVA）和他们拥有的作为共享数据（SD）的医疗数据集的数据摘要。DO 通过链上智能合约获取具有操作权限的用户的 SD 密文C_{SD}和 IVA 密文C_{IVA}，存储和$C_{IVA}C_{SD}$上链，并将 SD 发布到数据共享网络。此外，DO 还从用户那里接收 SD 的数据完整性和用户身份验证结果。

数据请求者（DR）是指需要使用医疗保健数据集的组织或个人。在 2DCM-DS 中，DR 通过作为访问者对链上数据发起操作请求来参与数据共享过程。通过智能合约，灾备可以发起数据查询和修改请求。对应的智能合约会执行数据完整性校验，并根据 DR 的 IVA 确定用户权限。验证成功后，DR 会向对应的 DO 发起后续操作请求。

联盟链（CBC）作为 2DCM-DS 中的去中心化网络载体，区块链账本记录网络内的数据操作交易。2DCM-DS 选择 Hyperledger Fabric 作为 CBC 实体。Hyperledger Fabric 是一个许可区块链，只允许合格的参与者维护区块。在 2DCM-DS 中，所有 DO 和 DR 都作为 CBC 上的用户节点参与数据共享，但并非所有 DO 和 DR 都充当管理节点。2DCM-DS 中定义的功能是通过智能合约实现的。

模型定义

本节定义了 2DCM-DS 的每个功能，并在其中正式定义了 2CME 算法。在 2DCM-DS 中,提出了一种新的二维混沌映射 2D-ILM,提出了一种基于 2D-ILM 的密码算法 2CME。2DCM-DS 设计有基于 2CME 的数据完整性验证接口和认证接口。

1. *2D-ILM*（x_0, y_0, μ_1, μ_2）：2D-ILM 结合了力 1D 混沌映射和 1D 逻辑映射，将参数空间扩展到两个维度。

2. *2CME（IVA，SD）*：2CME 是一种以二维混沌映射 2D-ILM 为中心的加密算法，该算法对 IVA 和 SD 进行加密，生成的密C_{IVA}文之间存在很强的耦合相关性，C_{SD}可以在两个方向上进行验证。

- *获取*F_a*和*F_b：计算$F_a$and$F_b$fromP_{IVA}和P_{SD}。
- *IVA 预处理*：对 IVA 进行预处理以获取中间密文P_{A1}和P_{A2}.
- *获取*C_{SD}、F_C*和*F_D：计算 SD 密文C_{SD}，获取F_C和F_D。
- *密钥生成*：通过 2D-ILM 计算密钥流。

- *获取干扰矩阵*、*排序矩阵*：计算干扰矩阵和排序矩阵，对数据进行混淆和扩散。
- *获取*C_{IVA}：计算最终的 IVA 密文C_{IVA}。

3. *解码*_2CME（C_{IVA}，C_{SD}，K_A，K_D，P_{min}，）：P_{max}给出C_{IVA}，C_{SD}，的，是干扰因子F_b的 Hash256 值$K_A P_{A1} K_D$，P_{min}并且P_{max}是加密过程中的规范化参数。2CME 的解码是加密算法反向操作的完整。可以得到相应的P_{IVA}。

4. *Id_Verify（DR 的 IVA）*：DR 使用自己的 IVA 作为身份凭证，请求对链上 SD 的操作请求。智能合约获取链上公开可用的 SD 与 storedC_{SD}和传入的 IVA 之间的匹配，以验证访客的操作请求是否被允许。

5. *Intergrity_Verify（SD）*：DR 需要在链上获取 SD 之前验证 SD 数据的完整性。智能合约获取了存储在链上的，C_{SD}对其进行解密，并确定它是否与 SD 匹配。C_{IVA}通过这种方式，可以验证 SD 的数据完整性。

风险模型

2DCM-DS 的核心是 2CME。一些对手试图通过破解 2CME 来破坏正常的数据共享行为。在设计的方案中，可以通过确保所提出的 2CME 是安全的来证明 2DCM-DS 的安全性。其中 2CME 是基于 2D-ILM 设计的，是一种基于 2D 混沌映射的密码算法。具体来说，2CME 的安全性需要通过以下原语来保证：

引理 1

2CME 对按键敏感，这使得 2CME 能够抵抗暴力破解攻击。

引理 2

2CME 可抵抗统计攻击，包括差分攻击和选择性明文攻击。

引理 3

由 2CME 加密的密文消息应具有高度的不确定性。

2DCM-DS 利用 2CME 专注于数据完整性验证和身份验证。只要保证 2CME 能够满足上述引文，2DCM-DS 中数据完整性验证和鉴证的解决方案就可以被认为是安全和可行的。

方法

本节详细介绍 2DCM-DS 的功能。

2D-ILM

（1） 2D-ILM（x_0, y_0, μ_1, μ_2)

为了实现更好的加密效果，该工作通过将一维 logistic mapping 集成到一维无限坍缩映射的基础上，对传统的混沌映射进行了改进，将单参数空间系统扩展到二维空间。通过组合两个映射获得的新二维无限逻辑映射（2D-ILM） 在超混沌状态下具有更大且连续的参数空间。具体说来：

使用的一维逻辑映射的数学表达式为：

$$x_{n+1} = \mu_1 x_n (1 - x_n)$$

使用的一维无限折叠映射的数学表达式为：

$$y_{n+1} = \sin\left(\frac{\mu_2}{1 - y_n}\right)$$

所提出的 2D-ILM 的数学表达式为：

$$\begin{cases} x_{n+1} = \sin\left(\mu_2 (1 - x_n)\right) \sin\left(\frac{\mu_1}{y_n}\right) \\ y_{n+1} = \sin\left(\frac{\mu_1}{x_n}\right) \sin\left(\frac{\mu_2}{1 - y_n}\right) \end{cases}$$

根据等式（3），可以看出 2D-ILM 有四个输入参数，其中控制参数μ_1 和 μ_2 是实数。x_n和y_n表示第 n 次迭代时的系统状态，而x_0和y_0表示 2D-ILM 的初始值。

2CME

（1）2CME（IVA，SD）

身份验证音频（IVA）和要共享的数据（SD）用作 2CME 的输入。由于包含相互干扰因素，双方都需要确保数据的完整性，以便正确执行 IVA 和 SD 的加密和解密过程。值得注意的是，由于这种混合加密过程，每个 SD 都唯一地对应于一个 IVA。Method 要求提供正确且相同的身份验证音频数据，才能通过 SD 和 IVA 的匹配验证。综上所述，上述加密方法在 SD 和 IVA 之间建立了双向可验证的强耦合相关性。图 2 说明了 2CME 的加密过程。具体步骤如下：

- ***获取F_a和F_b***：计算F_a和F_b从和P_{SD}是$P_{IVA}$$K_D$干扰因子$F_b$的采集参数。

$$F_a = var(P_{IVA}), F_b = K_D[0:4], K_D = Hash256\,(P_{SD})\,[0:4]$$

- ***IVA 预处理：***计算 IVA 的中间密文P_{A1}和P_{A2}并通过归一化控制[0,255]中的值范围。P_{min}和P_{max}是的$P_{a'}$最小值和最大值。其中整数部分是P_{A1}，小数部分是P_{A2}。

$$\begin{cases} P_{a'} = F_b \cdot P_{IVA} \\ P_{min} = min(P_{a'}), P_{max} = max(P_{a'}) \\ P_A = \dfrac{P_{a'} - P_{min}}{P_{max} - P_{min}}\ 255 \\ P_{A1} = floor(P_A), P_{A2} = P_A - P_{A1} \end{cases}$$

- ***获取C_{SD}、F_C和F_D***：对P_{SD}执行逻辑映射以获取C_{SD}、并计算$F_c = var(C_{SD})$和$F_d = mean(C_{SD})$。其中需要根据实际P_{SD}元素p获取校正参数。具体来说：$p = var(P_{SD}), \lambda = p/100 * 255$。

$$\begin{cases} C_{SD} = P_{SD}[0] \\ k_1 = C_{SD}[i-1]^2\ mod\ F_a \\ k_2 = (2\ C_{SD}[i-1])\ \ mod\ F_a \\ C_{SD}[i] = P_{SD}[i] + ((k_1 + k_2)\ mod\ 3)(\lambda \dfrac{len(P_{SD})}{9}) \end{cases}$$

- ***密钥生成：***计算的 Hash256 值K_AP_{A1}，除以K_A8 组 32 位数据，并将每组转换为十进制。对每对数据执行数值处理，将生成的 4 个数据作为初始值x_0和y_0，以及控制参数μ_1和$\mu_2$2D-ILM。

$$\begin{cases} K_A = Hash256(P_{A1}) \to k[8] \\ x_0 = (k[0] \oplus k[4])/(e^{18} \cdot 100) \\ y_0 = (k[1] \oplus k[5])/(e^{18} \cdot 100) \\ \mu_1 = (k[2] \oplus k[6])/e^{18} \\ \mu_2 = (k[3] \oplus k[7])/e^{18} \end{cases}$$

- 使用x_0、$y_0\mu_1$和μ_2作为 2D-ILM 的输入参数来获取关键流X和Y。

$$\begin{cases} X:\ x_{n+1} = sin\,(\mu_2(1 - x_n))\,sin(\dfrac{\mu_1}{y_n}) \\ Y:\ y_{n+1} = sin(\dfrac{\mu_1}{x_n})sin(\dfrac{\mu_2}{1 - y_n}) \end{cases}$$

- ***获取干涉矩阵，排序矩阵：***通过X和F_c获取干涉矩阵H。

$$H = floor\left(X \cdot F_c \cdot 10^{10}\right) \ mod\ 256$$

•获取排序矩阵S。生成二维矩阵S_1，其中$S_1[0] = Y \times F_d$和$S_1[1] = i,\ i \in (1,2,3,\ldots,l)$，其中$l = Length(Y)$。根据的值进行排序$S_1$，并将排序$S_1[1]$后的结果视为$S$。

•***获取***C_{IVA}：使用H、S和P_{A1}进行计算C_{IVA}。C_A是的整数部分C_{IVA}，前三位的计算C_A如下：

$$\begin{cases} C_A[S[0]] = P_{A1}[0] + (H[0]\ mod\ 64) \\ C_A[S[1]] = P_{A1}[1] + H[0] + (C_A[S[0]]\ mod\ 128) \\ C_A[S[2]] = P_{A1}[2] + H[0] + C_A[S[0]] + (C_A[S[1]]\ mod\ 256) \end{cases}$$

•检索的C_A剩余数据：

$$\begin{cases} t_1 = floor\left(x_0 C_A[S[i-1]]\left(1 - C_A[S[i-1]]\right)\right)\ mod\ e^5 \\ t_2 = floor\left(y_0 C_A[S[i-1]]\left(1 - C_A[S[i-2]]\right)\right)\ mod\ e^5 \\ C_A[S[i]] = P_{A1} + t_1 + t_2 + (H[i]\ mod\ 3e^5) \end{cases}$$

•检索完整的签名密文C_{IVA}。

$$C_{IVA} = C_A + P_{A2}$$

（2）Decode_2CME（C_{IVA}，，C_{SD}，K_A，$K_D P_{min}$，P_{max}）

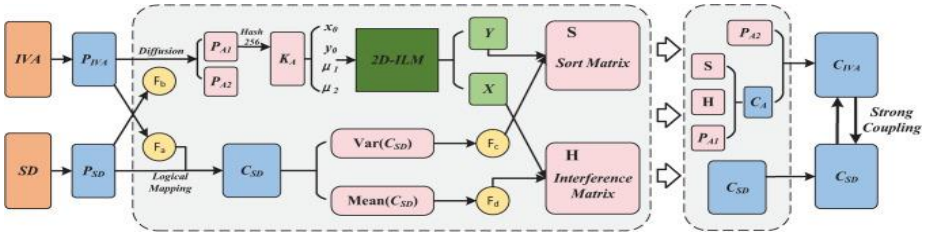

图 2. 2CME 的加密过程

C_{IVA}并C_{SD}用作解密算法的输入。需要保留加密过程的部分数据作为解密参数，具体来说：归一化参数P_{min}、P_{max}、的 Hash256 值$K_A P_{A1}$、的 acquisition 参数K_D F_b。2CME 的解密过程通常是加密过程的反面。由于空间原因，这部分通过伪代码进行描述。算法 1 是 2CME 的解密算法，如下所示：

Decode_2CME$(C_{IVA}, C_{SD}, K_A, K_D, P_{min}, P_{max})$

$C' = floor(C_{IVA}), P'_{A2} = C_{IVA} - C'$

$K_A \leftarrow k[8]$

$x'_0 \leftarrow (k[0] \oplus k[4])/(e^{18} \cdot 100)$

$y'_0 \leftarrow (k[1] \oplus k[5])/(e^{18} \cdot 100)$

$\mu'_1 \leftarrow (k[2] \oplus k[6])/e^{18}$

$\mu'_2 \leftarrow (k[3] \oplus k[7])/e^{18}$

$X', Y' \leftarrow 2D - ILM(x'_0, y'_0, \mu'_1, \mu'_2)$

$F'_c \leftarrow var(C_{SD}), F'_d \leftarrow mean(C_{SD})$

$H' \leftarrow floor(X' \cdot F'_c \cdot 10^{10}) \ mod \ 256$

$S' \leftarrow S'[1].\ S'[0] = Y'F'_d,\ S'[1] = i,\ i \le l = \textbf{length}(Y'), S' = S'[S'[0].\textbf{argsort}()]$

$P_{A1}[0] \leftarrow C'[S'[0]] - (H'[0] \ mod \ 64)$

$P_{A1}[1] \leftarrow C'[S'[1]] - H'[0] - (C'[S'[0]] \ mod \ 128)$

$P_{A1}[2] \leftarrow C'[S'[2]] - H'[0] - C'[S'[0]] - (C'[S'[1]] \ mod \ 256)$

for i in range $(3, l)$

$t'_1 \leftarrow floor(x_0 C'[S'[i-1]](1 - C'[S'[i-1]])) \ mod \ e^5$

$t'_2 \leftarrow floor(y_0 C'[S'[i-1]](1 - C'[S'[i-2]])) \ mod \ e^5$

$P'_{A1} = C'[S'[i]] - T'_1 - T'_2 - (H'[i] \ mod \ 3e^5)$

$P'_A \leftarrow P'_{A1} + P'_{A2}$

$F'_b \leftarrow K_D[0:4]$

$A' \leftarrow (P'_A(P_{max} - P_{min})/255) + P_{min}/F'_b$

if $A' == P_{IVA}$; **continue**

else return $V = False,\ A' = [\,],\ D' = [\,]$

$F'_a \leftarrow var(A')$

$D'[0] = C_{SD}[0]$

for i in range $(3, len(C_{SD}))$

$k'_1 \leftarrow C_{SD}[i-1]^2 \ mod \ F'_a$

$k'_2 \leftarrow (C_{SD}[i-1] + C_{SD}[i-1]) \ mod \ F'_a$

$D'[i] \leftarrow C_{SD}[i] - ((k'_1 + k'_2) \ mod \ 3)(\lambda len(C_{SD})/9)$

if $D' == P_{SD}$ **return** $V = Ture,\ A',\ D'$

else return $V = False,\ A',\ D' = [\,]$

算法 1. 2CME 的解密算法

身份验证和数据完整性验证

1. Id_Verify（DR 的 IVA）

如算法 2 所示，对于未经授权的用户，2DCM-DS 通过设计基于 2CME 的身份认证机制来增强区块链访问控制。在我们的方案中，用户只能通过区块链中的智能合约来操纵链上的 SD。SD 的每个部分在发布时都会在C_{IVA}区块链上记录授权用户。DR 提供其 IVA 和链上 SD 作为 2CME 的输入参数。通过将获取到的 IVA 密文与记录C_{IVA}的 IVA 密文进行比对，智能合约可以判断请求者是否为授权用户。

Id_Verify$(IVA\ of\ DR)$
$P_{IVA} \leftarrow DR's\ IVA\ plaintext$
$C'_A \leftarrow 2CME(P_{IVA}, P_{SD})$
$C_{IVA} \leftarrow Obtain\ the\ corresponding\ C_{IVA}\ from\ CBC$
if $C'_A == C_{IVA}$ **return** $V = True$
else return $V = False$

算法 2.身份验证算法

2. Intergrity_Verify（标清）

如算法 3 所示，针对潜在的恶意篡改数据，2DCM-DS 设计了基于 2CME 的数据完整性验证机制。在我们的方案中，在访问区块链上的 SD 之前，DR 需要对 SD 进行数据完整性验证，以确保检索到的数据是正确的。发布 SD 时，对应的C_{IVA}和C_{SD}都记录在区块链上。解密C_{IVA}和C_{SD}后，DR 得到 SD′ 。通过比较 SD 和 SD′ ，它可以确定 SD 是否完整。

Intergrity_Verify(SD)
$P_{SD} \leftarrow Plaintext\ of\ SD$
$C_{IVA} \leftarrow Obtain\ the\ corresponding\ C_{IVA}\ from\ CBC$
$C_{SD} \leftarrow Obtain\ the\ corresponding\ C_{SD}\ from\ CBC$
$P'_D \leftarrow Decode\_2CME(C_{IVA}, C_{SD}, \_, \_, \_, \_)$
if $P'_D == P_{SD}$ **return** $V = True$
else return $V = False$

算法 3.数据完整性验证算法

安全分析

本部分将对系统模型中提出的风险模型进行安全分析。本工作中使用的数据集和系统环境如下：

系统环境为 Windows 11 系统，CPU：Intel i7-8700 3.20 GHz，GPU：NVIDIA GeForce RTX 3090,8GB RAM。MATLAB R2021b，Python 3.10.9，Go 1.20.4。使用的区块链是 Hyperledger Fabric 2.2.5。

该数据集使用了 ESC-50 数据集和本地医疗保健数据集。ESC-50 是一个广泛使用的音频数据集，其中包含 .wav 格式的音频文件。我们将从 ESC-50 中随机选择一些音频文件参与模拟实验。本地医疗保健数据集是 2022 年至 2023 年某个地区的医疗保健数据。所有数据在使用前都经过脱敏处理。

在本节中，我们将从 ESC-50 数据集中随机选择三个音频文件作为 IVA，从本地医疗保健数据集中随机选择三个数据片段作为 SD。

键敏感度

安全加密算法要求对密钥具有高度的敏感性。密钥的微小差异也会直接影响加密的结果。这反映了加密算法抵御暴力破解攻击的能力。关于 2CME，$K_A = Hash256(P_{A1})$作为 2D-ILM 的输入，2D-ILM X的输出，Y作为关键流。2D 混沌映射 2D-ILM 对初始条件的高灵敏度保证了 2CME 对密钥的敏感性。其中，由密钥差异引起的最终密文的更改定义为：

$$\frac{\Delta C}{C} = \frac{|2CME(2D-ILM(K_A)) - 2CME(2D-ILM(K_A + \Delta K))|}{C}$$

应用不同的 1 位差值 ΔK_i，ΔK_j以获得K_A不同的密文 C_i，C_j。密文的长度为 N，密文差异率 （）Cdr 的计算公式为：

$$Cdr = \frac{Diff(C,C_i) + Diff(C,C_j)}{2N}$$

$$Diff(A,B) = \sum_{i=0}^{N-1} Difp(A(i),B(i))$$

$$Difp(A(i),B(i)) = \begin{cases} 1, A(i) = B(i) \\ 0, A(i) \neq B(i) \end{cases}$$

A 和 B 表示两个相同长度的密文，$A(i)$表示第 i 个元素。所提议的 2CME 的关键敏感性通过获取Cdr该值来反映。原始键是K_o，稍作修改的新键是 K_1，K_2，K_3，K_4。

$$K_o = f8247deb44ff5d9581cfed6ff3fc02c1c53f8ddba45bc7bf24249a5f071b909f$$
$$K_1 = f9247deb44ff5d9581cfed6ff3fc02c1c53f8ddba45bc7bf24249a5f071b909f$$
$$K_2 = fa247deb44ff5d9581cfed6ff3fc02c1c53f8ddba45bc7bf24249a5f071b909f$$
$$K_3 = 78247deb44ff5d9581cfed6ff3fc02c1c53f8ddba45bc7bf24249a5f071b909f$$
$$K_4 = e8247deb44ff5d9581cfed6ff3fc02c1c53f8ddba45bc7bf24249a5f071b909f$$

Cdr测试结果如下表 3。从结果中可以看出，2CME 对密钥灵敏度保持了 99% 以上的水平，这可以说明基于 2D-ILM 的 2CME 算法具有优异的密钥灵敏度。

表 3.2CME 的关键敏感性测试

Key	K_o,K_1,K_2	K_o,K_1,K_3	K_o,K_1,K_4	K_o,K_2,K_3	K_o,K_2,K_4	K_o,K_3,K_4
$Cdr(\%)$	99.6109	99.5955	99.6200	99.6045	99.6382	99.6265

差分攻击分析

差分攻击属于统计攻击类别。在提出的 2CME 算法中，2D-ILM 的混淆与最终加密结果对统计攻击的抵抗力直接相关。如果 2D-ILM 的输出结果能够使 IVA 的密文数据在空间中随机分布，那么它就能够抵抗差分攻击。其中 2D-ILM 的混沌性能取决于算法的初始值灵敏度和遍历。关于 2D-ILM 的混沌性能，可以使用 Lyaponov 指数和吸引子来评估。

2D-ILM 的混沌行为反映了其安全特性。在这项工作中，它与三个现有的二维混沌图进行了比较：2D-LNIC、2D-LSM 和 2D-ICM。通过考察不同系统的动力学行为，比较了相同参数空间和控制参数下各种二维混沌图的混沌行为。

1. 李雅普诺夫指数（LE）

对于二维混沌系统，有两个 Lyapunov 指数对应于系统在两个正交方向上的指数增长或衰减速率。通常，具有正 Lyapunov 指数的系统被认为是混沌的，而具有多个正 Lyapunov 指数的系统被认为是超混沌的。使用 Jacobian 矩阵计算不同二维混沌映射在两个正交方向上的 Lyapunov 指数。如果两者兼而有之LE_x且LE_y均大于 0，则认为该映射为超混沌。根据所选音频文件的实际数据范围，我们以 control 参数μ_1为自变量，数据范围为 [10， 80]，在此范围内随机选取了 4 个μ_2值，带x_0和y_0作为初始值。2D-ILM 和其他比较的二维映射的 Lyapunov 指数如图 2 所示。仿真结果中可以观察到，在相同的参数范围内，二维 ILM 的 Lyapunov 指数均大于 0 并呈递增趋势。这表明 2D-ILM 在参数范围内处于超混沌状态，其轨迹极难推断。图 3 还将 2D-ILM 的 Lyapunov 指数与其他映射进行了比较，表明 2D-ILM 在两个方向上的 LE 值都大于其他映射的 LE 值，这表明与其他映射相比，2D-ILM 表现出更好的混沌行为。

2. 吸引

混沌地图的吸引子是指混沌地图中稳定且吸引人的轨迹的集合，在不同的初始

条件下，这些轨迹往往会向这些数值演变。吸引子相图说明了混沌系统在长期演化后的稳定状态。具有良好混沌行为的二维混沌图通常表现出形状复杂且不规则的吸引子相图，在相空间中均匀分布并占据较大面积。

在这部分，对 2D-ILM 和三个比较映射进行了吸引子相图的模拟。控制参数设置为μ_1= 22 和μ_2= 5.4，并在 2000 次迭代后生成吸引子相图。结果如图 1 所示。表示 2D-ILM 的吸引子均匀覆盖了[-1,1] 范围内的二维相平面空间。这表明 2D-ILM 表现出出色的混沌行为，能够在参数空间内更有效地生成不可预测的轨迹。

3. 差分攻击

在数据共享网络中，攻击者利用从原始数据和修改后的数据中获得的加密密文之间的差异来破译加密算法。这种类型的攻击称为差分攻击。在密码学中，NPCR 和 UACI 通常用于评估加密算法对此类攻击的抵抗力。计算公式如下，其中 C^1 表示原始数据，C^2表示修改后的数据，T 表示数据长度。

$$D(i)=\begin{cases}0, & C^1(i)=C^2(i)\\ 1, & C^1(i)\neq C^2(i)\end{cases}$$

$$NPCR=\sum_i \frac{D(i)}{T}100\%$$

$$UACI=\sum_i \frac{|C^1(i)-C^2(i)|}{T}100\%$$

在参考文献中，作者提供了 NPCR 和 UACI 的可接受范围。NPCR 的可接受范围为 99.5875-100%，而 UACI 的可接受范围为 33.3648-33.5623%。NPCR 和 UACI 值落在这些范围内表明该算法对差分攻击具有良好的抵抗力。如表 4 所示，本文计算了签名明文和数据明文经过 2CME 处理后的 NPCR 和 UACI 值。实验结果表明，2CME 对差分攻击表现出很强的抵抗力。

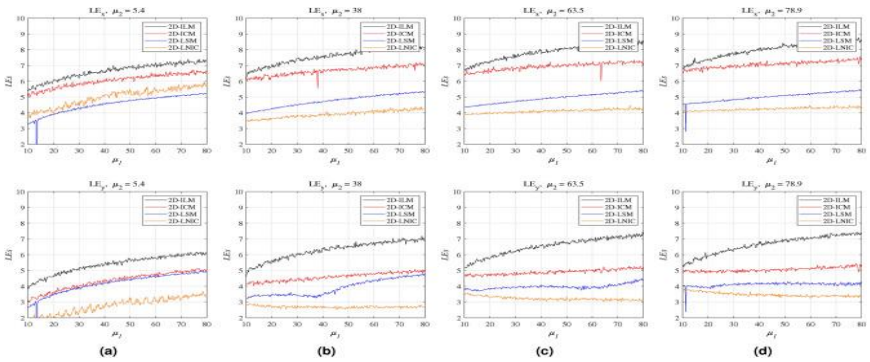

图 3.2D-ILM 的李雅普诺夫指数

（a）μ_2=5.4（序列 X，Y）。（b） μ_2=38（序列 X，Y）。（c） μ_2=63.5（序列 X，Y）。（d） μ_2=78.0（序列 X，Y）。

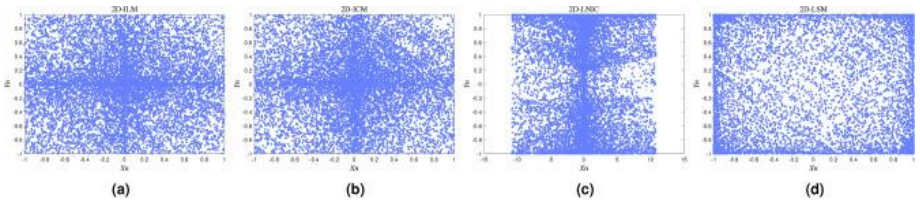

图 4.吸引子相位图

（a）μ_1=22， μ_2=5.4（2D-ILM）。（b） μ_1=22， μ_2=5.4 （2D-ICM）。（c） μ_1=22， μ_2=5.4 （2D-LNIC）。（d） μ_1=22，μ_2=5.4 （2D-LSM）。

表 4.2CME 的差分攻击分析

国际增值税	NPCR（%）	通过/失败	统一糖核酸（%）	通过/失败
1-12654-A-15.wav	99.6092	通过	33.4692	通过
1-110389-A-0.wav	99.6134	通过	33.4678	通过
1-208757-A-2.wav	99.6078	通过	33.4627	通过
1-13572-A-46.wav	99.6002	通过	33.4865	通过
1-116765-A-41.wav	99.6056	通过	33.4612	通过
1-101336-A-30.wav	99.6278	通过	33.4587	通过
1-160563-A-48.wav	99.6036	通过	33.4602	通过
平均	99.6096	通过	33.4667	通过
标清				
数据 1	99.6401	通过	33.4657	通过
数据 2	99.6121	通过	33.4639	通过
数据 3	99.6163	通过	33.4601	通过
平均	99.6228	通过	33.4632	通过

选择性明文攻击

选择性明文攻击通常分析差异较小的明文生成的密文之间的关系。为了保证算法的安全性，通常要求不同明文生成的密文之间有较大的差异。加密算法抵御所选纯文本攻击的能力通过下面描述的方法进行测试。

两段P_A，即A_1和A_2，两段P_D，即D_1和D_2，这些数据分别由 XOR 运算处理。生成

的输出用作输入以获取相应的密T_{A1}文、T_{A2}和$T_{D1}T_{D2}$。如果满足以下条件：

$$\begin{cases} A_1 \oplus A_2 \neq T_{A1} \oplus T_{A2} \\ D_1 \oplus D_2 \neq T_{D1} \oplus T_{D2} \end{cases}$$

这表明 2CME 可以抵御 chosen-plaintext 攻击。在选择明文抗攻击测试中，2CME 的明文和密文如图 2 所示。NPCR 表 5 可以评估明文和密文之间的区别。表中显示了所选数据的实验结果。从表 5 中可以看出，2CME 可以抵御 chosen-plaintext 攻击和四种经典攻击。

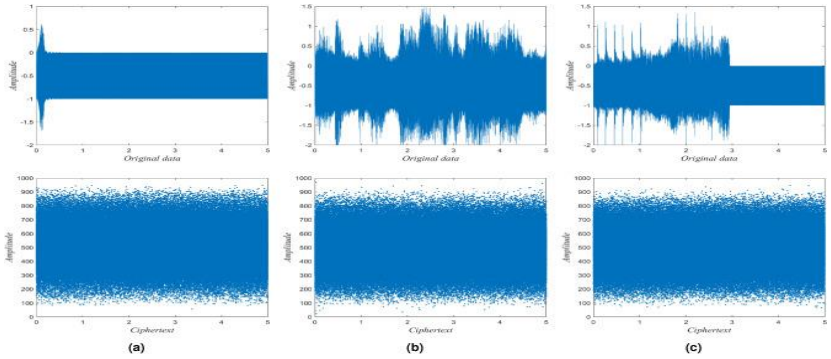

图 5. P_A2CME 的 Chosen Plaintext 攻击分析

（a） 1-12654-A-15.wav XOR 1-110389-A-0.wav 的音频。（b） 1-208757-A-2.wav XOR 1-13572-A-46.wav 的音频。（c） 1-116765-A-41.wav XOR 1-101336-A-30.wav 的音频。

表 5. 选择 2CME 的攻击分析

国际增值税	NPCR （%）	测试结果
图 5a	99.6463	通过
图 5b	99.6385	通过
图 5c	99.6347	通过
平均	99.6398	通过
标清		
数据 1	99.6403	通过
数据 2	99.6416	通过
数据 3	99.6221	通过
平均	99.6346	通过

相关分析

在 2CME 中，明文的相邻元素彼此之间存在某种关系，但安全密文的相邻元素应表现出较低的相关性，以避免统计攻击。相邻元素之间的相关性可以用相关系数来表示，相关系数的计算公式如下：

$$r_{(X,Y)} = \frac{Cov(X,Y)}{S_X S_Y}$$

其中Y表示 的X相邻元素，$Cov(X,Y)$表示两个样本的协方差，$S_X S_Y$表示两个样本的标准差。从 IVA 中随机提取 4000 对相邻元素作为 IVA 样品，并使用 SD 的完整数据作为 SD 样品。测试所选样本的明文和密文的相关系数。测试结果如下。

如图 6 和表 6 所示，P_D使用 2CME 处理后，不仅密文中相邻元素之间的相关性降低到较低的值，而且非相邻元素之间的相关性也被消除。这表明 2CME 对统计攻击具有良好的抵抗力。

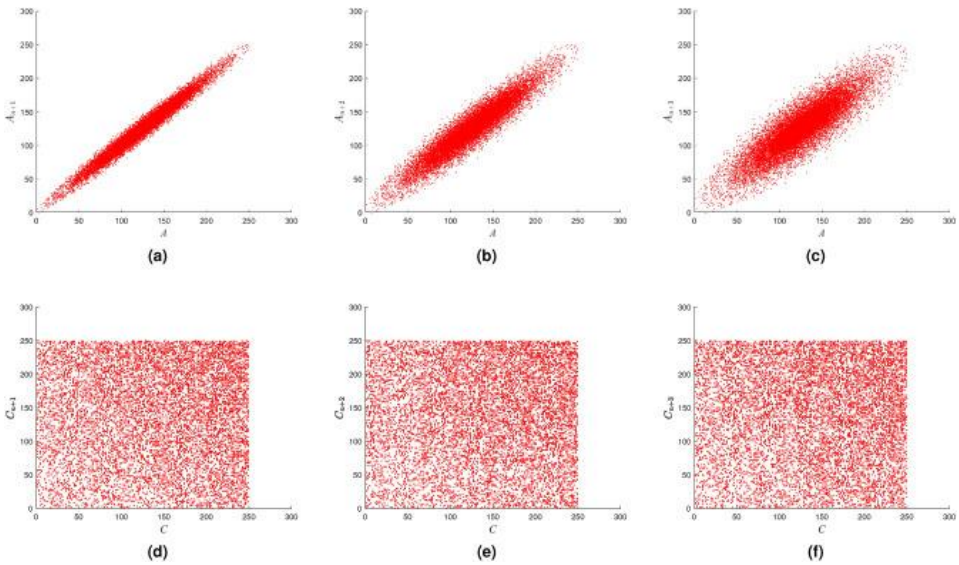

图 6. 1-34094-A-5.wav 的相关系数

（a） 和 之间的$A_n A_{n+1}$明文。

（b） 和 之间的$A_n A_{n+2}$明文。

（c） 和 之间的$A_n A_{n+3}$明文。

（d） 和 C_{n+1}之间的C_n密文。

（e） 和 之间的C_nC_{n+2}密文。

（f） 和 之间的C_nC_{n+3}密文。

表 6. 相关系数为 2CME

明文				Cipertexts			
IVA/SD	A_n, A_{n+1}	A_n, A_{n+2}	A_n, A_{n+3}	IVA/SD	C_n, C_{n+1}	C_n, C_{n+2}	C_n, C_{n+3}
1-12654-A-15.wav	0.9714	0.9535	0.9333	1-12654-A-15.wav	-0.0234	-0.0008	0.0079
1-110389-A-0.wav	0.0364	0.0376	0.0372	1-110389-A-0.wav	-0.0003	0.0101	-0.005
1-208757-A-2.wav	0.9829	0.9687	0.9377	1-208757-A-2.wav	0.0074	0.017	-0.0237
1-13572-A-46.wav	0.9844	0.9046	0.8651	1-13572-A-46.wav	0.0227	-0.0103	-0.0095
1-116765-A-41.wav	0.9851	0.9479	0.8824	1-116765-A-41.wav	0.0351	-0.016	-0.0154
1-101336-A-30.wav	0.9933	0.9961	0.9922	1-101336-A-30.wav	-0.0112	0.0041	0.0025
数据 1	0.5839	0.4732	0.4986	数据 1	-0.0522	-0.1044	0.0278
数据 2	0.6008	0.4701	0.4497	数据 2	-0.0529	-0.1084	0.0492
数据 3	0.5984	0.4735	0.475	数据 3	-0.0529	-0.116	0.0274

信息熵分析

信息熵可以反映数据密码中所含信息的随机性，分析信息熵可以显示信息分布的随机性程度。信息熵的计算方式如下：

$$H(s)=\sum_{i=1}^{2^n}P(S_i)log(\frac{1}{P(S_i)})$$

2^n表示数据源中状态的总数，S_i表示单个状态，$P(S_i)$并表示出现 S_i的概率。

在本文中进行的熵实验中，信息分布越混乱，熵值越接近理论值。2CME 的熵如表 7 所示。从实验结果中可以观察到，无论加密前的 IVA 和 SD 的熵是高还是低，经过 2CME 后得到的密文的熵都接近于理论值。这表明 2CME 表现出良好的加密效果，

使得攻击者难以从C_{IVA}或获取C_{SD}明文信息。

表 7. 2CME 的信息熵

国际增值税	熵P_A	熵C_A
1-12654-A-15.wav	7.1656	7.9896
1-110389-A-0.wav	2.3016	7.9981
1-208757-A-2.wav	7.6138	7.9936
1-13572-A-46.wav	7.6338	7.992
1-116765-A-41.wav	7.6368	7.9944
1-101336-A-30.wav	5.0435	7.9908
标清	**熵P_D**	**熵C_D**
数据 1	4.8769	7.9943
Data-2 （数据 2）	4.9711	7.9887
数据 3	4.9612	7.9954

噪声攻击分析

在传输数据时，攻击者可能会使用噪声攻击来干扰数据传输，这可能会影响解密信息的质量。当密文受到干扰时，可以恢复的明文信息越多，算法对噪声攻击的抵抗力就越强。密文受到添加不同密度的椒盐卷饼噪声的干扰。均方误差 （MSE） 和峰值信噪比 （PSNR） 用于计算噪声对密文的影响，MSE 和 PSNR 由以下公式定义：

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(P_i - C_i)^2$$

$$PSNR = 10\log_{10}(\frac{Max_P^2}{MSE})$$

其中P_i表示原始明文，C_i是解密加扰密文得到的明文，N表示数据长度。Max_P是原始纯文本数据的最大值。

从数据集中选择两个音频数据片段作为 IVA。将 0.0005 、0.001 和 0.005 级别的饱和冲激噪声添加到密文中，然后进行解密。图 7 中 a、b、c、d 显示了相应的原始 IVA 和添加噪声后解密的 IVA。同样，图 7 中 e，f，g，h 显示解密前后的 IVA。

图 7 的 NPCR 值，a、b、c、d 分别为 0.3342%、0.4077% 和 0.5433%。图 7 的 NPCR 值，e、f、g、h 分别为 0.3574%、0.4549% 和 0.5279%。测试得到的 NPCR 值均小于 1%，说明添加噪声解密后的数据与原始数据差异不显著。表 8 中的数据显示，两个 IVA 段都呈现较低的 MSE 值和较高的 PSNR 值。实验结果表明，在噪声攻击后，2CME 在解密算法后仍能获得大部分原始信息，证明 2CME 具有优异的鲁棒性。

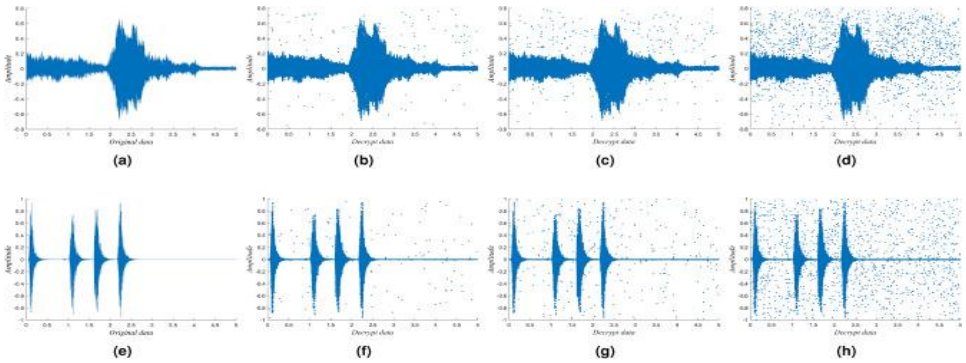

图 7. 2CME 的噪声攻击分析

（a） 1-34094-A-5.wav 的原始音频。

（b） 带有 0.0005 饱和冲激噪声的（a）。

（c） 具有 0.001 饱和冲激噪声的 （a）。

（d） 具有 0.005 饱和冲激噪声的（a）。

（e） 1-97392-A-0.wav 的原始音频。

（f） 带有 0.0005 饱和冲激噪声的（e）。

（g） 具有 0.001 饱和冲激噪声的（e）。

（h） 带有 0.005 饱和冲激噪声的 （e）。

表 8. 2CME 的 MSE 和 PSNR 测试

噪声强度		1-34094-A-5.wav	1-97392-A-0.wav
0.0005	小微电子	0.4114	1.1112
0.001		0.4313	1.15878
0.005		0.4644	1.2028
0.0005	PSNR	58.5636	55.201
0.001		59.3611	55.2564
0.005		60.682	56.2734

系统评估

2DCM-DS 的有效性

在本节中，所选的 IVA 和 SD 将用作 2DCM-DS 的输入。使用 2DCM-DS 处理前后的效果如下。如图 8 所示，经过 2DCM-DS 加密后，结果C_{IVA}呈现出视觉上令人满意的混沌性质，同时C_{SD}表现为在数据范围内随机分布的随机噪声。如表 9 所示，得到C_{SD}的表示十进制随机噪声数据，覆盖多个数量级并以数值形式呈现。解密后，两者都可以完全恢复到原始的明文数据。由此可见，2DCM-DS 可以将 IVA 加密成一段随机混沌噪声，并将 SD 加密成不同数量级的十进制随机噪声，通过适当的解密过程成功解密获得正确的明文数据。

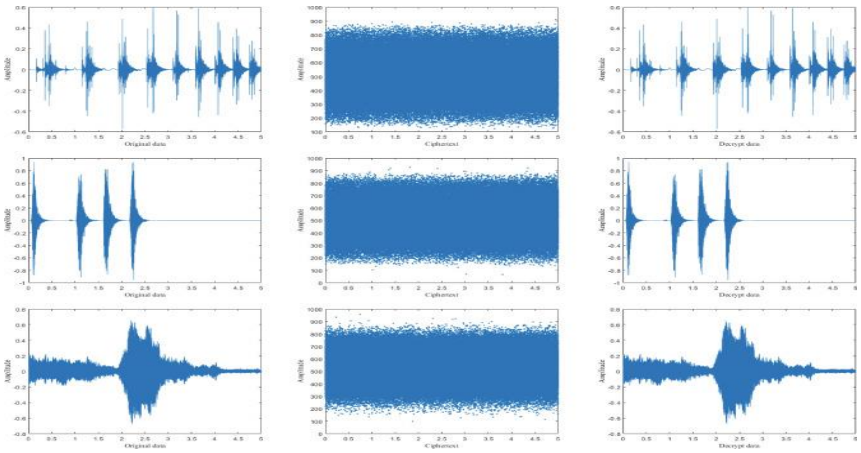

图 8. 2CME 对 IVA 的加密和解密效果

表 9. 2CME 处理后得到的 SD 密文和明文

数据	原始数据	密文	解密数据
数据 1	9***42	61525241909176.07	9***42
	1***46	52910012146873.93	1***46
	的医院	7.191190776348025E+42	的医院
	哈***ei	304358422440.738	哈***ei
	一般住院治疗	6.6879541176143E+54	一般住院治疗
	2022-**-*5	2.3188321528681803E+23	2022-**-*5
	3***.89	1.4094486445113918E+16	3***.89
Data-2 （数	1***35	519321652895.43353	1***35

数据	原始数据	密文	解密数据
据 2）	7***60	582859845929.0111	7***60
	C***l 医院	8.600154047157475E+35	C***l 医院
	哈***ei	6.564862937535509E+52	哈***ei
	一般住院治疗	6.6879541176143E+54	一般住院治疗
	2023-**-*3	8.889844084642599E+18	2023-**-*3
	14***.60	132957023118041.6	14***.60
数据 3	6***35	387525668719.83246	6***35
	8***93	401595655713.7395	8***93
	S***d 医院	1.2235318458355507E+43	S***d 医院
	哈***ei	2022832050.9865	哈***ei
	紧急处理	1.006382765994508E+43	紧急处理
	2023-**-*0	6.017760149706512E+18	2023-**-*0
	9***.38	104695389288680.67	9***.38

耗时

在本节中，我们将在 Hyperledger Fabric 22.5.0 中部署智能合约，以测试加密和验证、数据完整性验证和身份验证过程的时间消耗。如表 10 所示，平均加密时间为 0.14474s，平均解密时间为 0.12642s，平均数据完整性验证时间为 0.15092s，平均身份认证时间为 0.14969s。2DCM-DS 可以满足数据共享网络中的实时通信需求。

表 10. 2DCM-DS 加解密、完整性验证和身份验证耗时

国际增值税	标清	加密时间（s）	解密时间（s）	完整性验证（s）	身份验（s）
1-12654-A-15.wav	数据 1	0.15364	0.14319	0.14681	0.13011
1-110389-A-0.wav		0.15445	0.17677	0.14585	0.13291
1-208757-A-2.wav	数据 2	0.14668	0.14327	0.14927	0.12005
1-13572-A-46.wav		0.15866	0.14558	0.14193	0.12399

国际增值税	标清	加密时间（s）	解密时间（s）	完整性验证（s）	身份验（s）
1-116765-A-41.wav	数据 3	0.14118	0.14626	0.14517	0.13102
1-101336-A-30.wav		0.15091	0.14309	0.13940	0.12047
平均		0.15092	0.14969	0.14474	0.12642

与近似方案的比较

本文选取了四种相关的加密算法作为对照，替换了 2DCM-DS 的加密部分，并测试了一条大小为 431kb 的数据的加密时间。获得的结果如图 9 所示，表示所提算法在时间消耗方面优于其他算法。

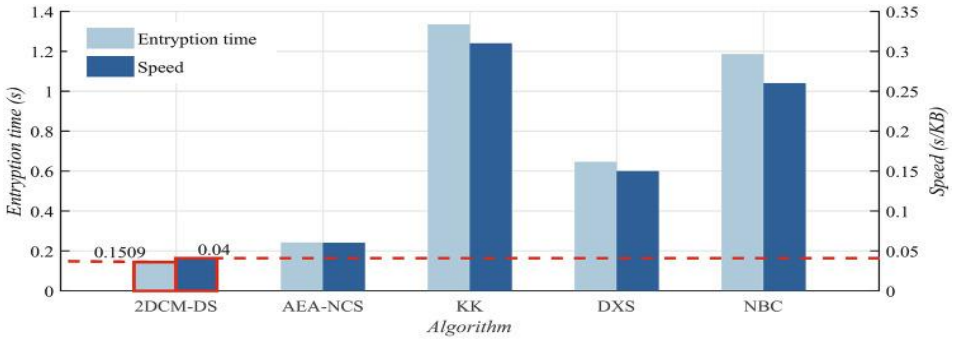

图 9. 算法的时间消耗比较：AEA-NCS[25]、KK[41]、DXS[42]、NBC[43]

在这部分，选择了一些基于区块链的数据完整性验证和身份验证功能解决方案，并在时间成本方面与 2DCM-DS 进行了比较。比较结果如下图 10，2DCM-DS 在数据完整性验证和认证方面都优于对比解决方案。

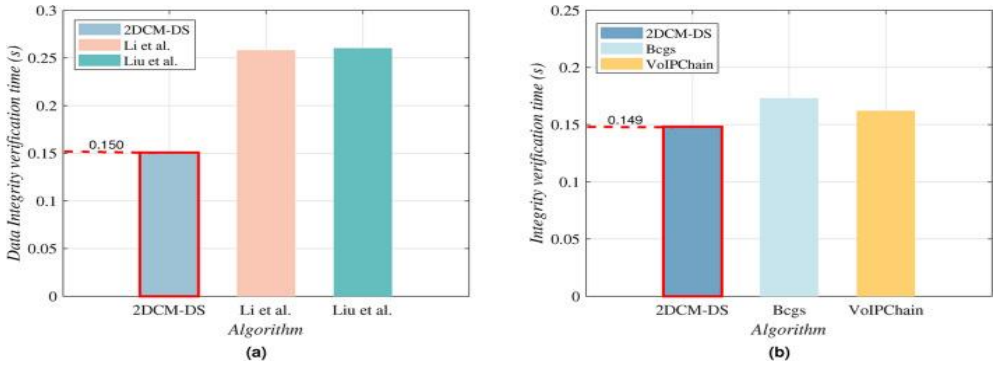

图 10.2DCM-DS 和近似方案之间的时间消耗比较

（a） 数据完整性验证时间比较，比较方案是 Li 等人，Liu 等人

（b） 认证时间比较，比较方案为 Bcgs、VoIPChain

结论

针对当前医疗健康数据共享中的数据隐私安全风险，本文提出了一种基于二维混沌映射和区块链的安全健壮的医疗健康数据共享方案。我们介绍了一种表现出卓越混沌性能的二维混沌映射（2D-ILM）。利用 2D-ILM 的混沌特性，结合生物识别音频信息，设计了一种能够双向验证的强耦合数据加密方法（2CME）。基于这些组件，我们开发了 2DCM-DS 并使用区块链作为底层网络实施了该方案。性能评估和安全性分析表明，所提方案具有鲁棒性、高效耗时和较强的抗攻击能力，满足医疗数据共享的通信和数据隐私保护要求。未来的工作将考虑扩展该计划内的区块链安全环境。目前的方案仍然存在一些限制：区块链本身的可扩展性将限制一些基于区块链的数据共享方案的发展。当一些邪恶节点提供低质量数据或网络遭受拜占庭式攻击时，系统是否还能保证数据安全。将加密算法封装到智能合约中是否足够安全。