分布式智能医疗系统中轻量级、匿名的医疗大数据互认证方案

解奥南 xxx

河海大学计算机与软件学院

2024 年 11 月 2 日

河海大学
HOHAI UNIVERSITY

大数据技术的发展背景

1. 大数据技术的发展背景：

- “大数据”是需要新的处理模式才能具有更强的决策力、洞察发现力和流程优化能力来适应海量、高增长率和多样化的信息资产。

特征	介绍
容量 (Volume)	数据的大小决定所考虑的数据的价值和潜在的信息
种类 (Variety)	数据类型的多样性
速度 (Velocity)	指获得数据的速度
可变性 (Variability)	妨碍了处理和有效地管理数据的过程
真实性 (Veracity)	数据的质量
复杂性 (Complexity)	数据量巨大，来源多渠道
价值 (Value)	合理运用大数据，以低成本创造高价值

表 1: 大数据的特征

大数据在智能医疗中的作用 (1/2)

2. 大数据在智能医疗中的作用：

分布式智能医疗是指通过集成各种智能医疗设备和通信网络，形成一个统一的基础设施，以实现系统的智能化、自动化和智能化管理。

分布式智能医疗中的大数据是指采用智能医疗技术所产生的海量医疗数据。医疗大数据的主要来源是组成分布式医疗网络的大量互联智能设备。

- 1. 医疗成像设备；
- 2. 电子健康记录 (EHR) 系统；
- 3. 医疗物联网 (IoMT) 设备；

大数据在智能医疗中的作用 (2/2)

- 4. 临床运营管理系统；
- 5. 供应链管理系统；
- 6. 患者参与平台；

这些设备和系统在医疗网络中生成和传输大量数据，为医疗专业人员提供了深入的洞察力，以改善患者的治疗效果和医疗系统的运营效率。随着技术的发展，更多的智能设备将被集成到医疗网络中，进一步扩大医疗大数据的来源。

医疗大数据技术的安全挑战 (1/2)

3. 医疗大数据技术的安全挑战:

分布式智能医疗网络主要由嵌入 IoT 设备、传感器、执行器、处理单元等的互联智能设备组成。分布式框架将大数据配置为多个系统，以减少单个系统的工作量。然而，更多的系统最终会带来更多的安全问题。

此外，**大多数用于大数据分析的工具都是在没有考虑数据安全问题的情况下开发的**。因此，对医疗大数据系统的攻击可能会破坏整个医疗系统的工作流程。未经授权的用户或设备可以访问和更改这些医疗保健数据，还可以控制连接到患者身体的 IoT 设备，这可能会给患者造成严重的医疗保健相关问题。

医疗大数据技术的安全挑战 (2/2)

4. 如何应对医疗大数据的安全挑战:

因此，需要依靠网络安全来保护医疗保健大数据，防止整个基础设施遭到破坏。医疗保健大数据的安全性可以通过维持 CIA 三要素来确保，即保密性、完整性和可用性。保密性意味着防止未经授权的数据访问或未经授权的数据披露。

完整性防止未授权的数据修改，可用性确保数据在需要时始终可用。这些方面可以通过控制对医疗保健大数据的访问并确保对每个通信实体进行身份验证来实现。因此，在开发医疗基础设施的同时，有必要同时预先考虑所有三个安全方面。

问题与相关工作

为了维护安全，管理员需要对所有用户和设备进行注册。在开始新的会话之前，医生（用户）和 IoT 传感器节点（设备）相互验证，然后进行密钥交换。

在支持批量的医疗保健系统中，有效的访问控制机制必须确保真实用户或设备不能访问医疗保健数据，除非他/她被授权访问。然而，访问控制模型的强度主要取决于身份验证过程的底层机制。近年来，开发了几种用于访问控制、身份验证和会话密钥生成的技术。

然而，由于计算和计算成本高，这些方案大多不适合资源受限的物联网环境，无法抵抗许多安全攻击，如节点捕获、被盗验证攻击、内部攻击、中间人（MITM）攻击等。

问题与相关工作

Zhang 等人设计了一种安全的个人医疗记录共享系统的访问控制机制。除了数据加密和访问控制外，该系统还具有匿名身份验证机制，用户可以匿名进行身份验证。因此，在认证过程中保留了用户的身份，并且通过采用离线在线流程，也减少了数据所有者的签名生成开销。但是，整体模型不适用于资源受限的设备。

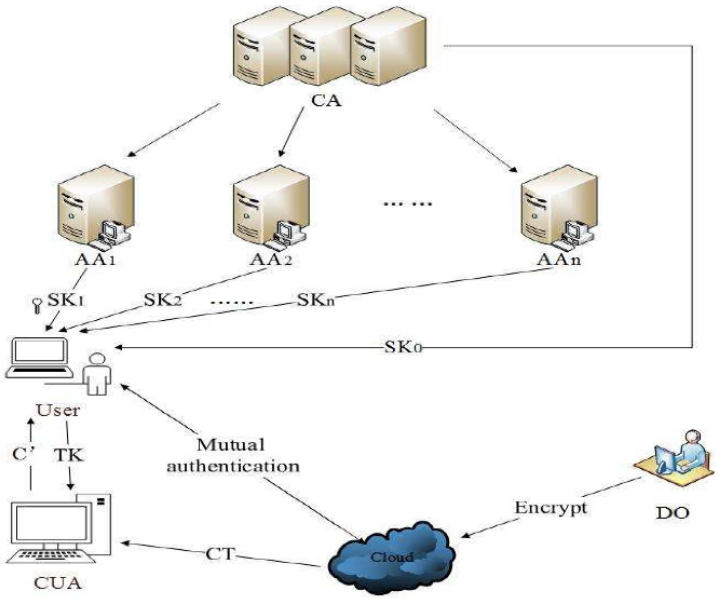

图 1: 系统模型

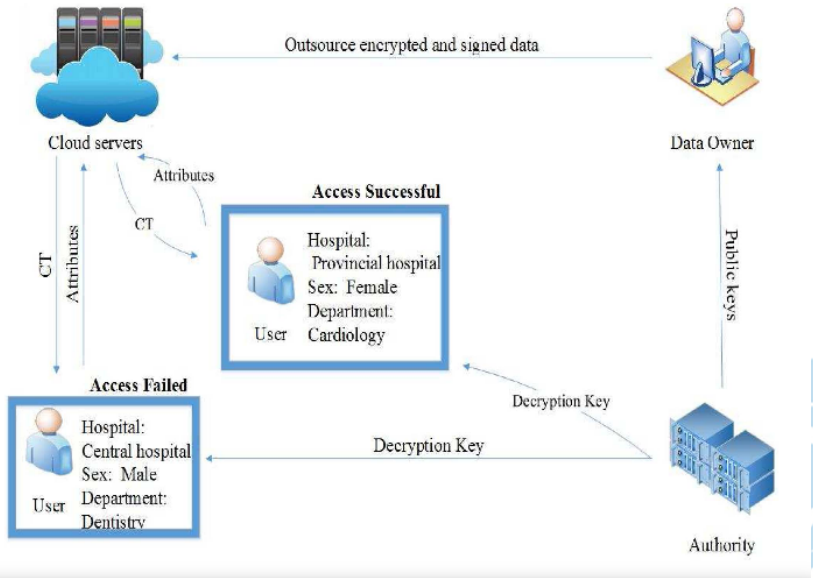

图 2: 访问控制示例

问题与相关工作

Koya 等人为国外网络节点开发了一种基于生物特征数据的身份验证模型。这是一种混合方案，将一种认证协议与患者的 ECG 信号相结合。为了提高认证过程，作者从 ECG 信号中提取生物密钥。这为医疗保健认证系统提供了更好的安全性。该方案可以抵御密钥托管问题和物联网设备模拟攻击。但是，它无法抵抗重放攻击和不可追溯性问题。

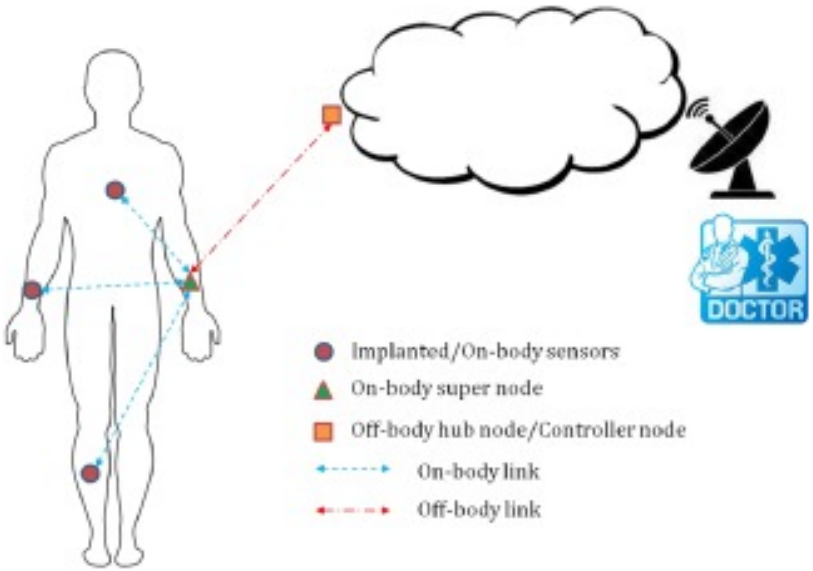

图 3: 系统模型

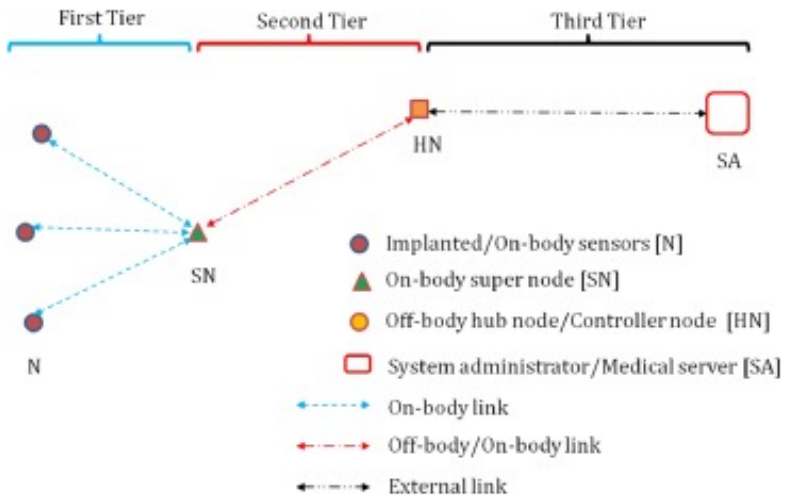

图 4: 访问控制示例

问题与相关工作

Gupta 等人提出了另一种认证模型来改进 Koya 等人的方案。在这个方案中，他们使用简单的 XOR 和哈希操作设计了一个安全的密钥协议和匿名身份验证协议。该方案还具有一个动态节点更新过程，可以动态地将新的传感器节点添加到网络中。但研究发现，该方案由于计算成本高、通信开销大，面临可扩展性问题，且无法抵抗去同步攻击。

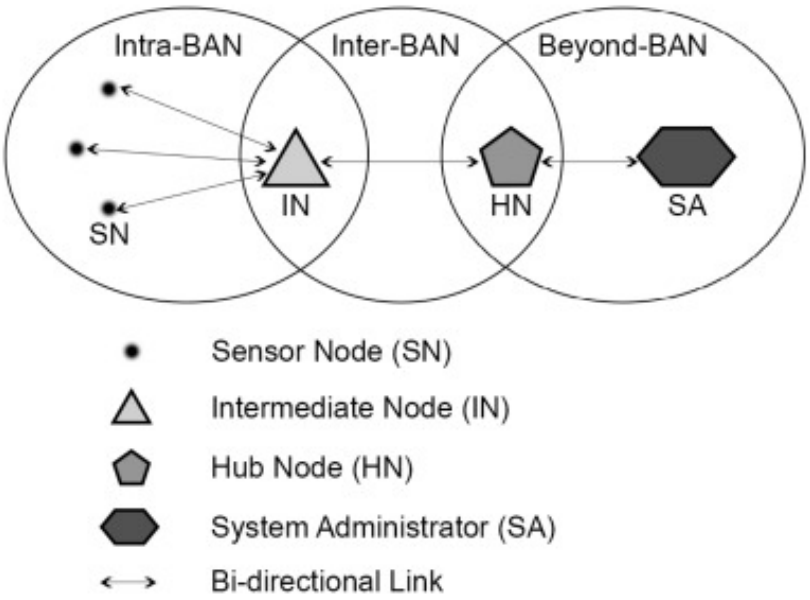

图 5: 系统模型

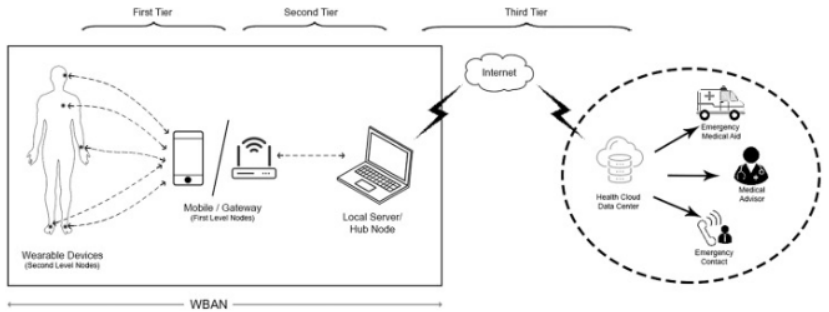

图 6: 访问控制示例

问题与相关工作

Farash 等人提出了一种**用户认证和密钥建立方案**。当用户请求在自身和网络内的传感器设备之间建立通信通道时，系统与该特定设备进行通信，并指示该设备执行身份验证过程。然后，在传感器设备和用户之间通过可信网络设备开始身份验证过程。此外，该方案**使用简单的对称密码学**，使整个系统轻量级。遗憾的是，**该方案无法抵抗某些加密攻击**。尽管许多研究人员在文献中针对不同领域提出了不同的隐私保护、身份验证和访问控制技术，但**设计一种低资源消耗和高效率的技术仍然具有很高的挑战性**。

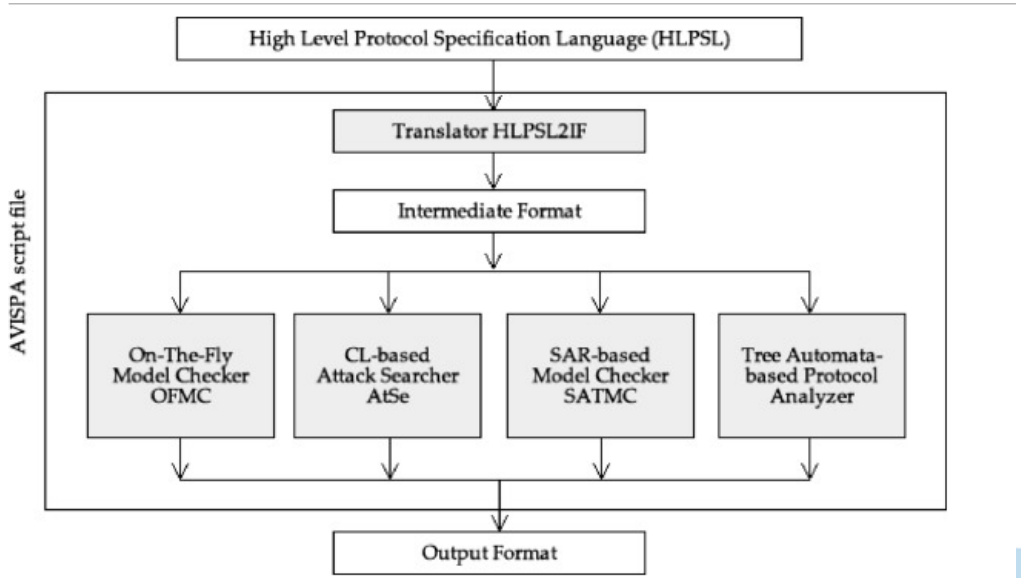

图 7: AVISPA 工具的安全验证流程图

系统模型

系统模型包括 4 个主要实体：

- **物联网设备 (IoT Devices, IoTD)**：与患者身体相关的资源受限设备，用于收集患者的实时医疗数据，并通过网关设备发送给用户。
- **网关 (Gateway)**：**不受资源限制的设备**，充当物联网设备与用户之间的中介。
- **中央管理员 (Central Administrator, CA)**：负责初始化系统并对系统中的每个实体进行注册和身份验证。
- **用户 (Users)**：可能是医生或其他医疗工作者，通过网关设备访问物联网设备收集的患者数据。

设计目标

该方案有 4 个设计目标：

- **双向认证**：在医疗网络中，设备和用户之间需要相互认证，并同意会话密钥以建立安全通信。
- **消息完整性**：确保医疗数据不会被未授权的用户篡改，从而保障数据的完整性。
- **身份匿名性**：保护用户和设备的身份信息不被攻击者获取和跟踪，防止仿冒攻击和中间人攻击。
- **轻量级设计**：方案必须适合资源受限的物联网设备，使用如 XOR 和哈希等轻量级操作，以延长设备的使用寿命。

威胁模型

论文采用了 Dolev-Yao 安全模型，该模型假设攻击者可以完全控制公共通信信道，并具备以下攻击能力：

- 攻击者可以在不安全信道上拦截、修改、删除或重放消息。
- 攻击者可以通过物理攻击获取 IoT 设备存储的敏感数据。
- 攻击者能够进行功率分析攻击来窃取存储在设备内存中的数据。

方案工作流程

- 本节介绍了所提出方案的构建过程，包含四个算法，分别为系统初始化、用户注册、设备注册以及登录与认证。
- 图8展示了该方案的工作流程。

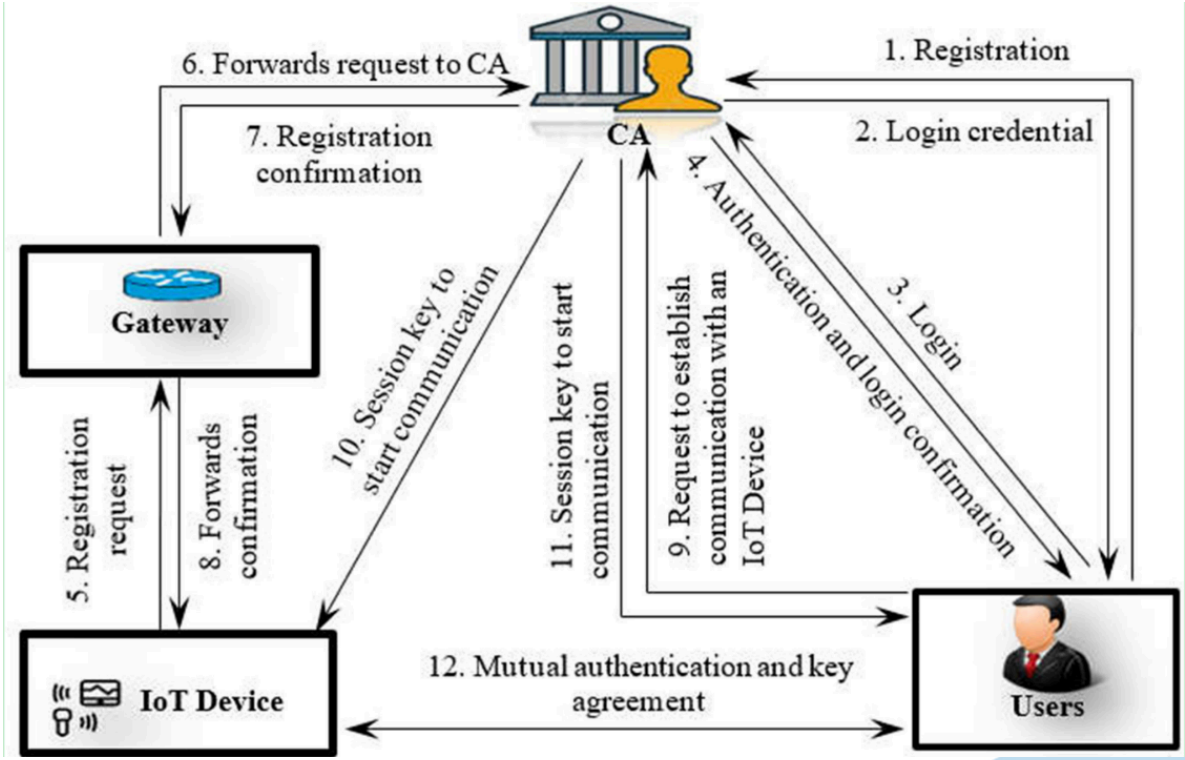

图 8: 拟议方案的工作流程

系统初始化

- 中央管理员首先通过选择用户、物联网设备 (IoTD) 和网关的密钥值 RU、RD 和 RG 来初始化系统。
- CA 通过安全通道将密钥发送给用户、IoTD 和网关。这些密钥会分别安全地存储在各自的内存中。
- 图9展示了系统初始化过程。

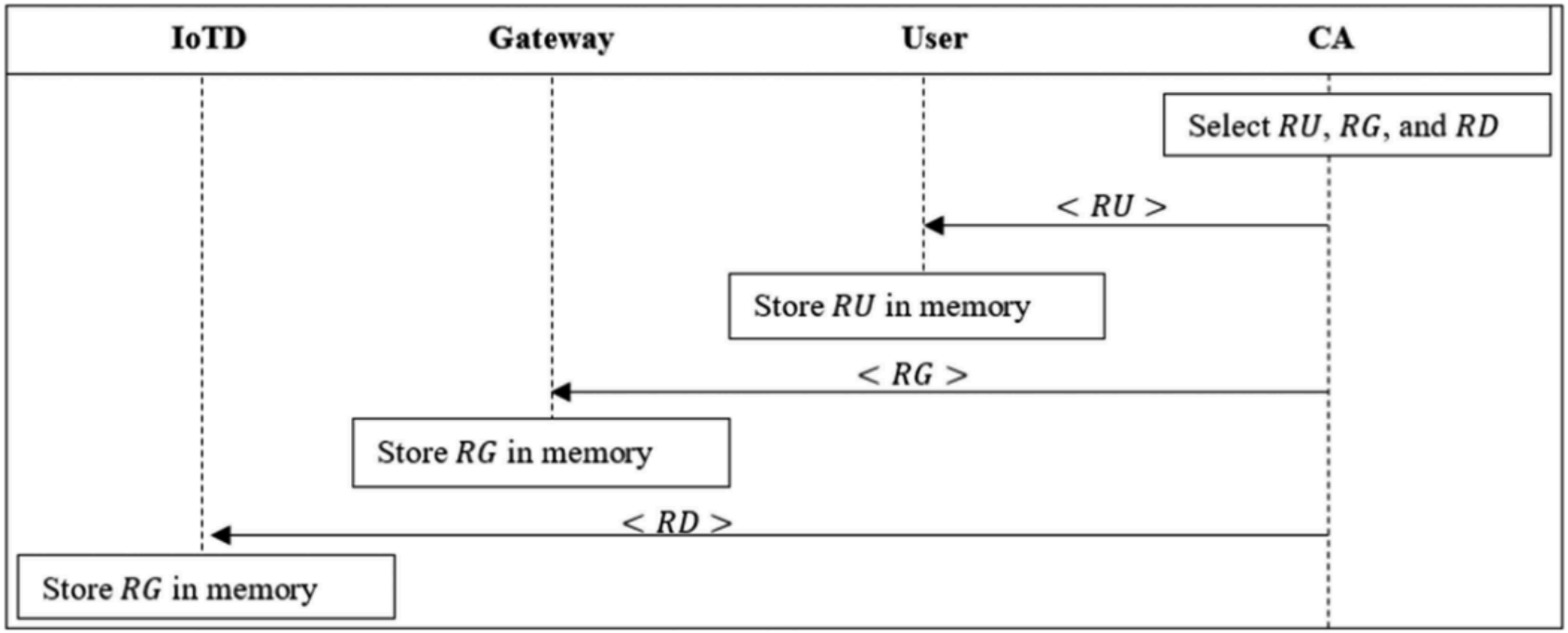

图 9: 系统初始化过程

用户注册过程

- 用户需要将其设备与中央管理员 (CA) 注册以实时获取患者的健康数据。
- 整个注册过程如图10所示：

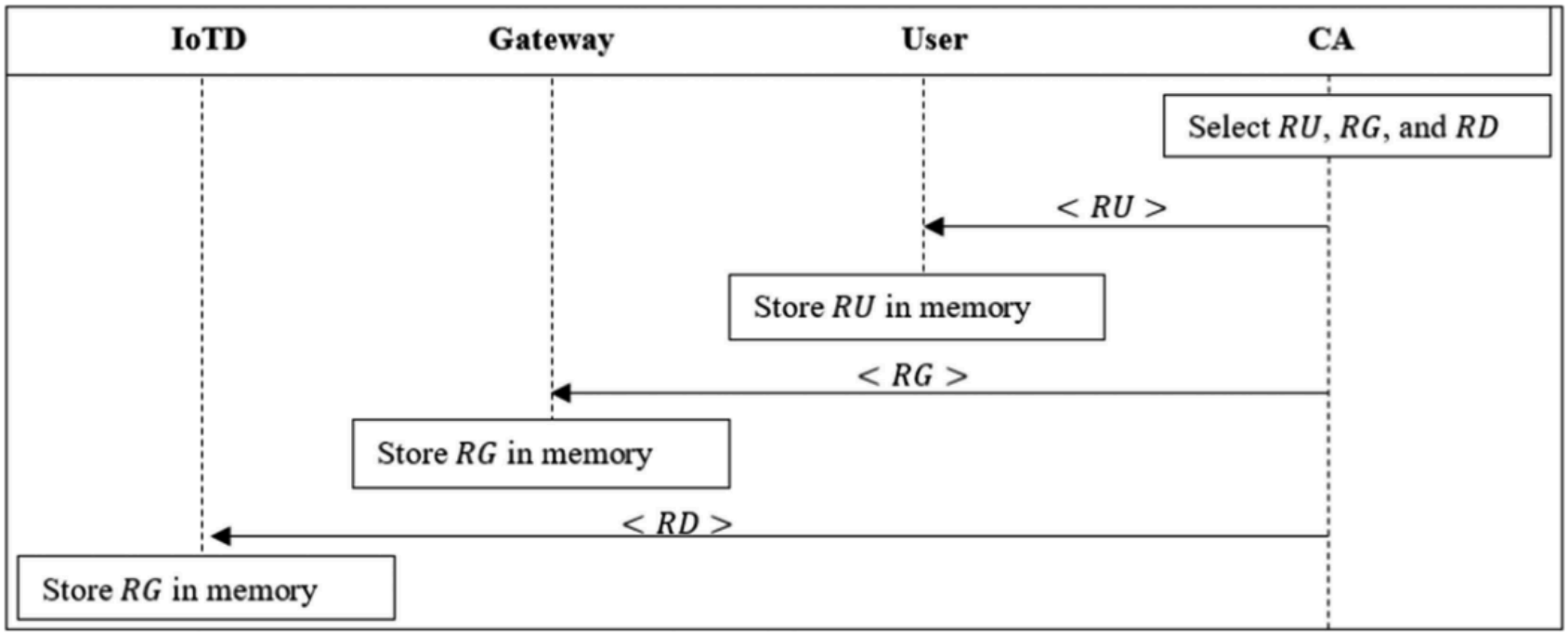

图 10: 用户注册过程

设备注册过程

- 每个物联网设备 (IoTD) 在向用户传输患者的健康数据之前，需要与中央管理员 (CA) 进行注册。
- 设备注册过程如图11所示：

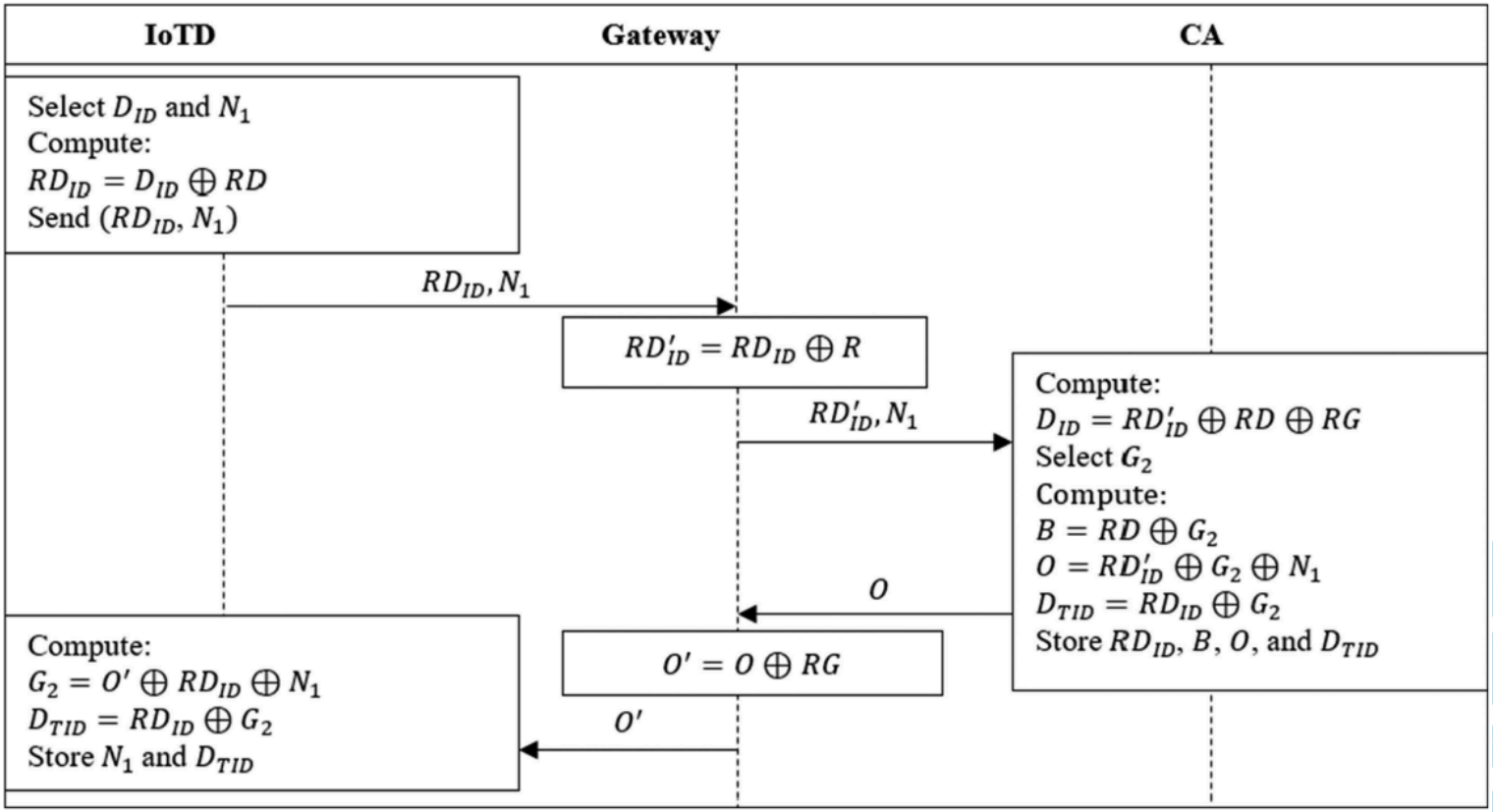

图 11: 设备注册过程

登录与认证过程 (1/2)

- 该部分展示了用户和物联网设备 (IoTD) 之间的登录和相互认证过程，用户需要通过 CA 和网关进行身份认证。
- 登录和认证过程如图12所示：

登录与认证过程 (2/2)

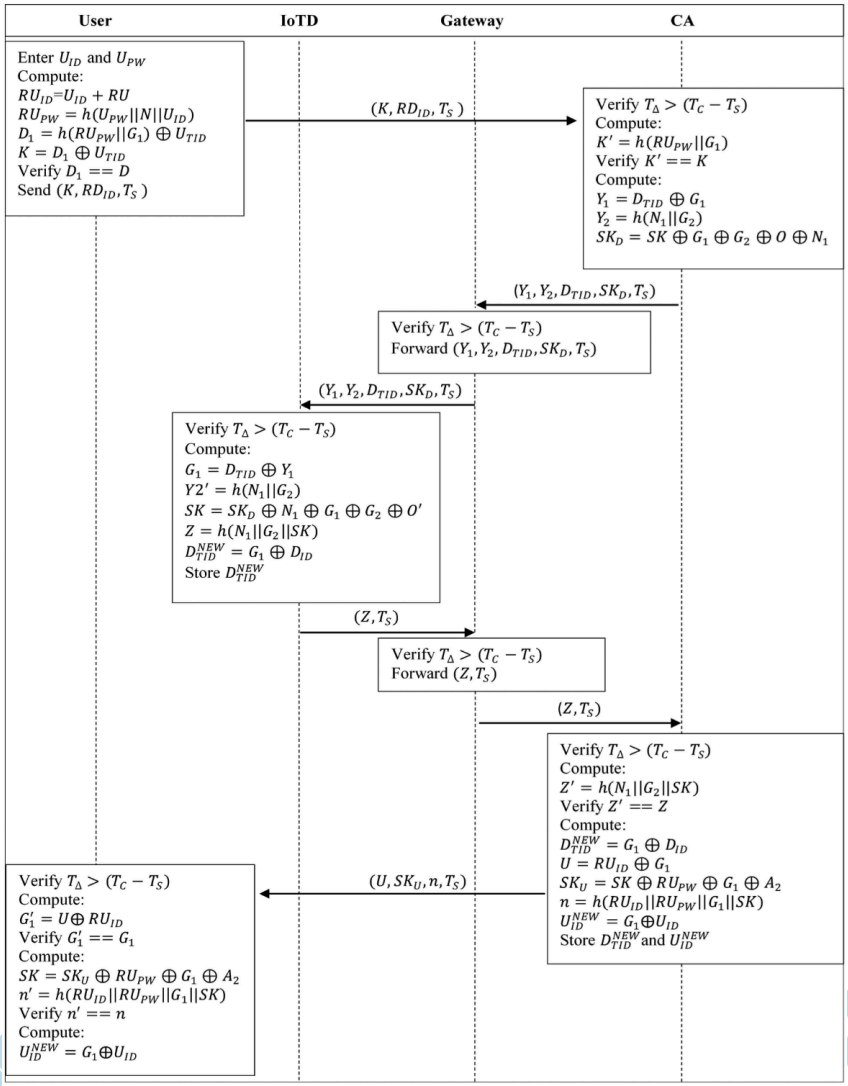

图 12: 登录与认证过程

使用 ROR 模型的形式化分析 (1/4)

- 本节介绍通过使用 ROR 模型对提出方案的安全性进行分析。ROR 模型是一种用来验证会话密钥安全性的概率模型。
- 假设攻击者 **A** 完全控制通信信道，能够捕获、修改、篡改、重定向、删除或重放通过网络发送的所有消息。
- 攻击者 **A** 可以执行以下查询：
  - **Exe(ξ)**：执行窃听攻击，获取有效方之间交换的消息。
  - **Snd(ξ, $\mathcal{M}$)**：攻击者可以发送消息 $\mathcal{M}$ 给 ξ，并接收 ξ 的回复。
  - **Cpt(ξ)**：攻击者可以执行捕获攻击，获取存储在物联网设备(IoTD)、网关和用户设备中的所有参数。

使用 ROR 模型的形式化分析 (2/4)

- 继续介绍攻击者 **A** 的查询：
  - **Hsh**(S)：攻击者通过此查询获取输入字符串 S 的固定长度哈希值。
  - **Tst**(ξ)：在此查询中，攻击者随机选择一个会话密钥 (SK)，如果 $c = 1$，攻击者获取真实的 SK；如果 $c = 0$，获取随机数。
- **A** 赢得游戏的事件称为 **Suc**，并且 **A** 破坏方案的优势定义为：

$$Adv_{\mathcal{A}}^{\xi} = |2Prob(Suc) - 1|$$

使用 ROR 模型的形式化分析 (3/4)

定理

对于一个多项式时间 t 内运行的攻击者 $\mathcal{A}$，其破坏方案的优势满足：

$$Adv_{\mathcal{A}}^{\xi} \leq 2\max\left\{D' \cdot q_{snd}, \frac{q_{snd}}{2^f}\right\} + \frac{q_{snd}}{2^{f-2}} + \frac{3q_{hsh}^2}{2^{f-1}}.$$

其中，D' 和 b 是两个常数，f 是用户输入信息的长度。

- 证明使用六个游戏序列来展示。每个游戏 G_i 都计算了攻击者 **A** 在不同查询规则下的成功概率。

使用 ROR 模型的形式化分析 (4/4)

- 从 G_0 到 G_6，展示了攻击者 **A** 的各种攻击行为对破坏方案的成功概率影响，结果是每个游戏的成功概率不断递减。

$$\text{Prob}\left[\text{Suc}_{\mathcal{A}}^{G_6}\right] - \text{Prob}\left[\text{Suc}_{\mathcal{A}}^{G_5}\right] \leq \frac{q_{hsh}^2}{2^{f+1}}.$$

最终得出：

$$\text{Adv}_{\mathcal{A}}^{\xi} \leq 2\max\left\{D' \cdot q_{snd}^{b}, \frac{q_{snd}}{2^f}\right\} + \frac{q_{snd}}{2^{f-2}} + \frac{3q_{hsh}^2}{2^{f-1}}.$$

使用 AVISPA 工具的形式化分析 (1/2)

- 本节展示了使用 AVISPA 工具对提出的方案进行形式化验证。AVISPA 工具是一种自动化的形式化验证工具，用于分析协议的安全性。
- AVISPA 使用高级协议规范语言 (HLPSL) 描述协议，并通过转换器 (HLPSL2IF) 将其转换为中间格式 (IF)，然后使用工具的四个后端进行分析：
  - 基于安全协议分析的树自动机 (TA4SP)
  - 基于约束逻辑的攻击搜索 (CL-AtSe)
  - 基于 SAT 的模型检查器 (SATMC)
  - 基于在线模型检查的 OFMC
- 结果显示协议的安全或不安全状态。

使用 AVISPA 工具的形式化分析 (2/2)

- 提出的方案的 HLPSL 代码使用 AVISPA 工具的 Security Animator (SPAN) 模拟工具进行分析。分析结果如图13所示。
- 该分析显示，方案在 CL-AtSe 和 OFMC 后端上都是安全的，能够抵御已知攻击，且会话密钥的机密性得到了验证。

File

```
SUMMARY
 SAFE

DETAILS
 BOUNDED_NUMBER_OF_SESSIONS
 TYPED_MODEL

PROTOCOL
 /home/span/span/testsuite/results/Authentication.if

GOAL
 As Specified

BACKEND
 CL-AtSe

STATISTICS

 Analysed  : 29503 states
 Reachable  : 16691 states
 Translation: 0.00 seconds
 Computation: 0.09 seconds
```

File

```
% OFMC
% Version of 2006/02/13
SUMMARY
 SAFE
DETAILS
 BOUNDED_NUMBER_OF_SESSIONS
PROTOCOL
 /home/span/span/testsuite/results/Authentication.if
GOAL
 as_specified
BACKEND
 OFMC
COMMENTS
STATISTICS
 parseTime: 0.00s
 searchTime: 3.02s
 visitedNodes: 4600 nodes
 depth: 16 plies
```

图 13: AVISPA 工具生成的输出结果

非正式分析 (1/2)

- 提出的方案能够抵御多种攻击，如窃听攻击、重放攻击、特权内部攻击以及被盗凭证攻击。该方案还保持了匿名性和不可追踪性。
- **窃听攻击**：在窃听攻击中，攻击者 **A** 可以截获、删除或修改传输的数据。即使攻击者获取了所有传输的消息，也无法获得会话密钥，因为 **A** 不知道 *G1* 和 *G2*，以及 CA 在注册过程中选择的密钥值 *RD*、*RU* 和 *RG*。
- **重放攻击**：在重放攻击中，攻击者通过重放之前会话中的消息来获得网络访问权限。提出的方案通过使用时间戳来防止这种攻击，确保消息在有效的传输延迟范围内被处理。

非正式分析 (2/2)

- **特权内部攻击**：在这种攻击中，内部攻击者可能获取到用户的身份和密码。然而，由于 *A2* 使用密钥值 *RU* 加密，且攻击者不知道 *RU* 和 *G1*，因此攻击者无法计算出会话密钥。
- **被盗验证攻击**：在这种攻击中，攻击者通过获取某些节点的验证凭证尝试计算会话密钥。但由于 **A** 无法获取 *RU*，无法成功计算出用户的身份和会话密钥，从而抵御被盗验证攻击。
- **匿名性与不可追踪性**：提出的方案确保用户的真实身份 *UID* 不会被攻击者获取，且用户的活动不会被追踪。即使攻击者获取了 *RUID* 和 *RUPW*，也无法通过这些信息追踪到用户。

安全特性比较

- 表2显示了提出的方案与其他方案在安全特性方面的比较，包括抵抗特权内部攻击、离线密码猜测、MITM 攻击、以及 IoTD 伪装等。

方案	特权内部攻击	离线密码猜测	前向保密性	MITM 攻击	IoTD 伪装	IoTD 捕获	被盗验证攻击
[30]	Y	N	N	N	Y	Y	Y
[32]	N	Y	N	Y	N	N	N
[33]	N	Y	N	Y	N	N	N
[34]	Y	Y	Y	Y	Y	N	N
提出方案	Y	Y	Y	Y	Y	Y	Y

表 2: 不同方案的安全特性比较

存储需求比较

- 提出的方案使用 SHA-256 哈希操作，其密钥长度为 256 位。
- 表3显示了不同方案在各个实体的存储需求比较。

Scheme	User	IoTD	IN	CA	Total
[30]	1088b	-	992b	160b	2240b
[32]	576b	576b	640b	1792b	2944b
[33]	-	640b	640b	480b	1780b
[34]	-	1088b	320b	320b	1728b
Proposed scheme	736b	320b	160b	1216b	2432b

表 3: Proposed Scheme's Storage Cost in Bits (b)

计算成本比较 (1/2)

- 表4展示了不同方案在用户注册、设备注册、登录和认证阶段的计算成本。
- 哈希操作的总时间也列在表中。

Phase	User	IoTD	CA	Total	Total Time (ms)
User registration	$2H_{CC}$	-	-	$2H_{CC}$	0.0008
Device registration	-	-	-	-	-
Login and authentication	$3H_{CC}$	$2H_{CC}$	$4H_{CC}$	$9H_{CC}$	0.0036
Total cost	$\mathbf{5H_{CC}}$	$\mathbf{2H_{CC}}$	$\mathbf{4H_{CC}}$	$\mathbf{11H_{CC}}$	**0.0044**

表 4: Proposed Scheme's Computation Cost in Different Phases

计算成本比较 (2/2)

- 表5展示了不同方案在计算成本上的比较，提出的方案表现出较优的计算效率。

Scheme	User	IoTD	CA	Total	Time (ms)
[30]	$10H_{CC}$	$19H_{CC}$	$7H_{CC}$	$36H_{CC}$	0.0144
[32]	$3H_{CC}$	$2H_{CC}$	$3H_{CC}$	$8H_{CC}$	0.0032
[33]	-	$3H_{CC}$	$10H_{CC}$	$13H_{CC}$	0.0052
[34]	-	$7H_{CC}$	$10H_{CC}$	$17H_{CC}$	0.0068
Proposed scheme	$2H_{CC}$	$3H_{CC}$	$4H_{CC}$	$9H_{CC}$	0.0036

表 5: Comparison of Computation Cost

通信成本比较

- 表6展示了不同方案在消息交换数量和总比特数上的通信成本比较。
- 提出的方案表现出较好的通信效率，特别是在消息交换数量与比特数的平衡方面。

Scheme	No. of message exchanged	Total no. of bits
[30]	4	3840
[32]	4	1024
[33]	6	5376
[34]	4	4096
Proposed scheme	6	1536

表 6: Comparison of Communication Cost

执行时间比较

- 图14展示了不同方案在用户注册、设备注册、以及认证阶段的执行时间比较。
- 提出的方案在这些阶段的执行时间相对较短，表明其具有较好的性能表现。

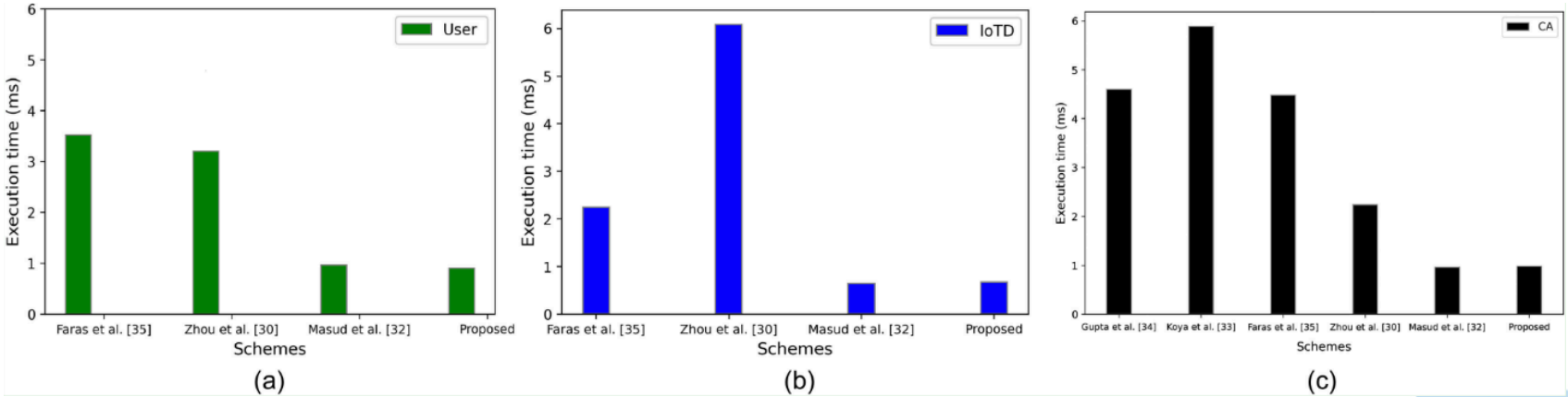

图 14: Comparisons of execution time (a) User registration, (b) Device registration, and (c) Authentication phase

结论与未来工作

- 本文提出了一种基于简单加密操作的轻量级用户与设备认证方案，能够保证用户匿名性，并抵抗多种密码学攻击。
- 安全性和性能分析验证了方案的有效性和效率，该方案在注册时间和认证时间上优于许多现有方案。
- 未来工作可通过改进方案的开销问题，并为物联网医疗应用引入多因素认证机制。

感谢大家的聆听