在远程不可信服务器上使用混沌进行隐私保护生物认证

参考文献《Privacy preserving biometric authentication using Chaos on remote untrusted server》

隐私保护生物识别方案根据使用的方案分为四类：

- 模糊抽取器：基于模糊提取器的方案易于实现，然而为了提高系统准确性，辅助数据往往过大，而且辅助数据攻击中的操纵也是有可能的。
- 生物哈希：基于生物哈希的的系统易于实现，工作速度更快，生物哈希的一个主要缺点是它会遭受计算表攻击，如查找表攻击
- 安全多方计算：提供了加密域中的计算便利，但是限制是计算开销较大。
- 同态计算：随机数的创建需要计算，降低系统的速度，在实时业务应用程序中，高复杂同态计算速度非常慢。

概要

混沌函数通常用于产生置换和扩散所需的密钥流。根据系统的安全性，一维和多维的混沌映射是有用的，对于更多的数维，生成的结果包含复杂的混沌序值，然而由于其复杂性，实施这些多维解决方案的成本很高。而对于认证数据表ADT中包含的用户信息和特征码保护是最为重要的，不仅需要提高访问成本，而且由于其稳定性，保证几乎不可能替换生物特征码。

任何生物特征都可以使用对称加密进行加密解密；在提出的密码系统中使用混沌映射用于产生的密钥流，由种子值计算得出，种子来自于输入的生物特征图像；其次是在加密过程中进行了像素级的混淆和扩散，而且每一轮都会修改混淆密钥流，重复执行m和n次；最后使用Paillier加密系统进行加密，使用同态方式执行操作，并分析了身份验证数据表攻击，验证其安全级别。

接下来首先简要介绍Paillier算法和混沌映射，接下来阐述系统的主要组成及工作流程，最后是对于安全性的实验和讨论。

Paillier

对于了解密码学的同学来说，无论是同态加密(下称HE)还是本次要说的Paillier算法，都并不罕

见，但这里为了引出接下来的系统，我还是对这些方面稍作介绍。

根据支持的计算类型和支持程度，HE可以分为下三种类型:

- 半同态加密(PHE)：只支持加法或乘法中的一种运算。其中只支持加法运算的又叫做加法同态加密(AHE)。
- 部分同态加密(SWHE):可同时支持加法和乘法运算，但是支持的计算次数有限。
- 全同态加密(FHE):支持任意次数的加法和乘法运算。

由于FHE在计算有限次乘法后需要较复杂的去除噪声的操作，一般使用通用MPC协议通信开销较大。在复杂的计算场景中，单独使用某种通用方法通常得不到一个可用的落地方案。PHE诞生之后很快在隐私保护领域得到了大量使用，其高效且支持无限次加法或乘法的特点使其成为隐私计算的重要基本组件。在首次被提出后，学术界出现了多个支持PHE的方案，如RSA、GM、Elgama和Paillier等。

我们假定两个大素数U和V，并定义对于公钥的加密函数Encrypt()和Decrypt(),对于整数域ZN中的两个参数k1,k2，Paillier算法定义了如下同态性质：

$$R = U \times V.$$

$$Decrypt(Encrypt(k_1 * k_2)) = Decrypt(Encrypt(k_1)^{k_2} mod\ R^2)$$

$$Decrypt(Encrypt(k_1 + k_2)) = Decrypt(Encrypt(k_1) * Encrypt(k_2) mod\ R^2)$$

> 方案安全性可以归约到判定性合数剩余假设（DCRA），即给定一个合数n和整数z，判定z是否在n^2下是否是n次剩余是困难的。这个假设经过了几十年的充分研究，到目前为止还没有多项式时间的算法可以攻破，所以Paillier加密方案的安全性被认为相当可靠的。

混沌系统

由于图像数据本身具有大数据量，高冗余和相邻数据有强相关性的特点，使用AES等传统加密算法并不适合用于图像加密。
混沌系统具有遍历性，非周期性，对初始条件和控制参数具有高灵敏度和类随机行为等优良的内在特性。基于混沌映射的典型加密过程可以分为两个部分：置换和融合。在图像加密这一研究领域，之前的主流方式是在像素级排列的图像加密，像素级加密本质上就是对于图片以非线性形式进行重组排列，难以抵御选择明文和选择密文攻击。

这里的混沌并非指的是无序，而是当一个简单确定的系统不仅可以产生简单确定的行为，还可以产生貌似随机的不可确定行为。对于混沌的定义是指确定的宏观非线性系统在一定条件下所呈现的不确定的或不可预测的随机现象；是确定性与不确定性，规则性与非规则性；有序性与无序性融为一体的现象。这也和香农提出的密码学扩散混淆理论有一定相似之处。我们对于混沌定义以下几个特征:

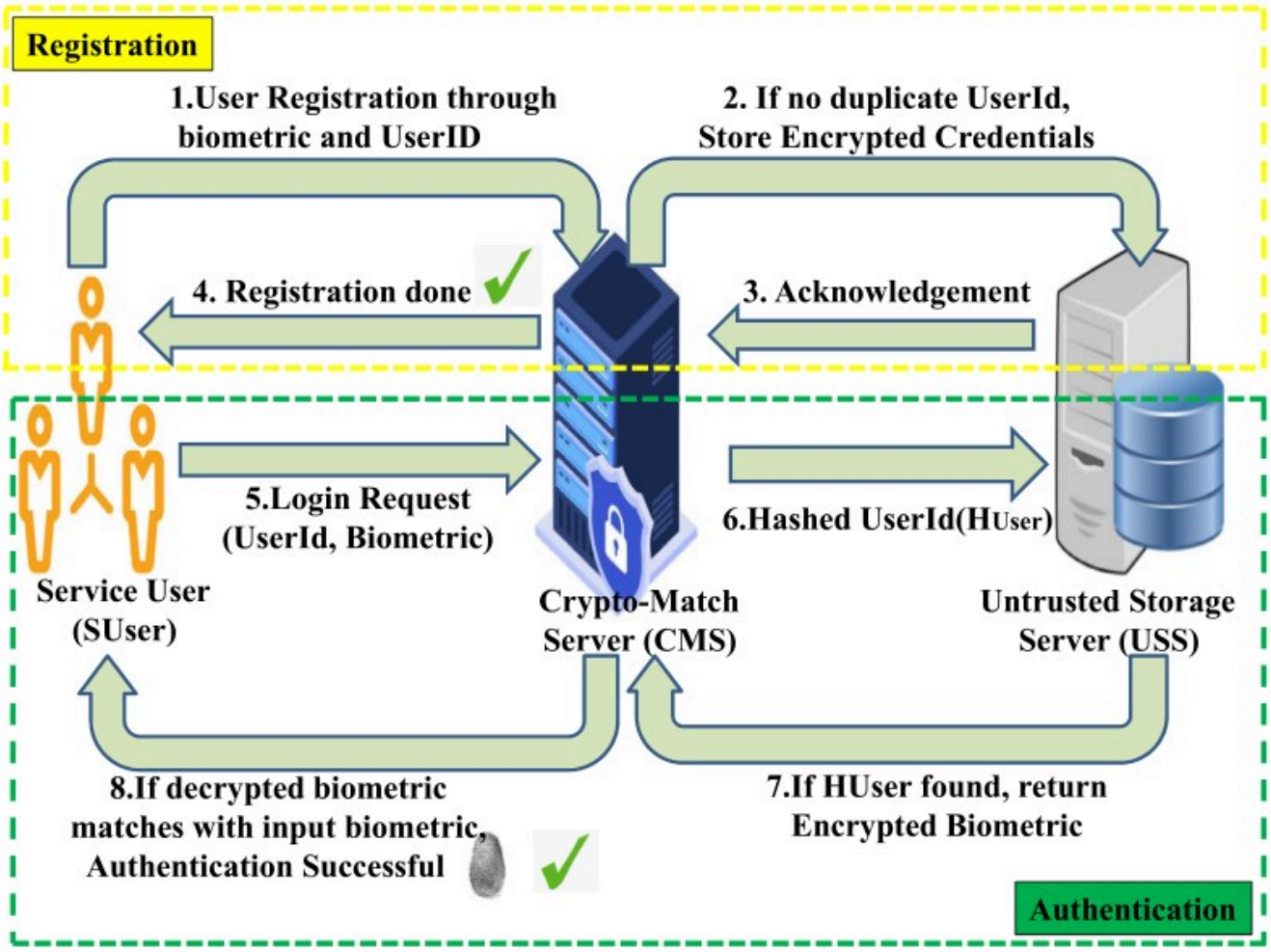
1. 首先是对于初始条件的灵敏依赖性，通俗来说就是蝴蝶效应，“失之毫厘差之千里”，局部或足够微小的扰动都会对于系统最后的结果产生较大的改变。
2. 长期不可预测性，混沌的非线性动力学特性决定了混沌是不可预测的。
3. 分形性，分形性指混沌的运动轨线在相空间中的行为特征，表示混沌运动状态具有多叶，多层结构，且叶层越分越细，表现为无限层次的自相似结构。
4. 有界性，混沌运动轨线始终局限于一个确定区域，混沌吸引子是混沌有界性的最好体现。
5. 遍历性，混沌运动在其混沌吸引域内是各态历经的，在有限时间内混沌轨道不重复地经历吸引子内每一个状态点的邻域。

下面首先引出一个概念：李雅普诺夫特征指数，是用于识别混沌运动的一个特征。常用于判定一个系统的混沌性，通过图像可以直观地看出某个系统或者映射是否是混沌系统或映射。
判定依据是当其大于0时，说明系统会进入混沌态，对应的映射也具备被称为混沌映射的资格。小于零时说明会趋于稳定且此时对系统的初始状态不敏感，等于零时说明系统稳定。

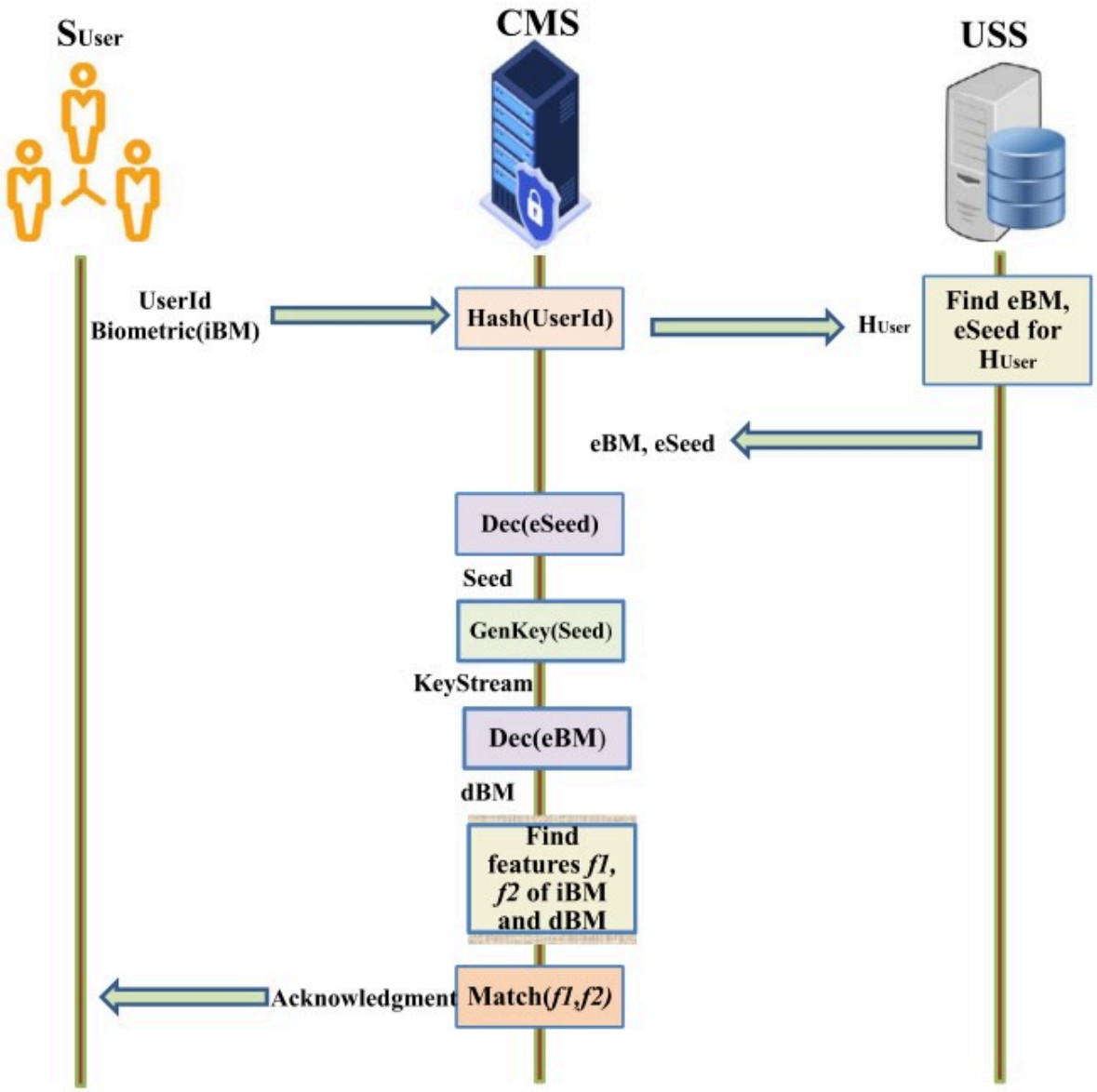
系统和工作流程

该方案包含三个实体，服务用户，加密匹配服务器和不受信任的存储服务器，服务用户指的是经过系统认证的用户，加密匹配服务器在建议的系统中，被视为可信实体，必须严格遵守协议，该服务器负责提供两种服务：用户注册和身份验证，除此之外，它还维护Paillier密码系统的密钥，用于种子的加密解密计算。

初始化工作在加密匹配服务器上完成，服务器生成用于加密和解密种子值的密钥吗，并维护保存这些数据。
注册阶段开始时，客户提供输入用户ID和用户生物特征并提交到加密匹配服务器，加密匹配服务器使用SHA-256计算ID的哈希值，并提交给远程不受信服务器，远程服务器检查ADT的输入ID，如果已注册，则返回冲突，如果ID的哈希值为新，则让加密匹配服务器执行生物特征数据的加密，并计算生物特征图像的种子值。种子用于生成密钥流并加密生物特征图像并由加密匹配服务器将这些数据上传至远程服务器。接收完成后，并将注册结果反馈给用户。



用户验证阶段如图所示，用户首先在客户端上传自己的认证生物图像和用户ID到加密匹配服务器，加密匹配服务器计算用户ID并上传到远程不受信服务器进行匹配，匹配成功时向加密匹配服务器传递该用户的注册加密图像和加密种子值进行解密，解密的过程具体为首先使用种子值生成解密用的密钥流，再解密图像数据，通过生物图像特征生成特征代码F1,F2,并将其参与匹配。



安全性分析

第一个指标是直方图分析测试，通过对于像素分布生成的直方图，保证原始图像和加密图像中不应存在统计相似性，如果生成的加密图像中的像素是均匀的，那么加密方法就是安全的，这样能够使攻击者无法泄漏有关原始图像和加密图像的任何信息。

第二个指标是相关性分析，根据定义，图像的内置属性是相邻或连接的像素高度相关的，从加密系统的目的出发，需要打破所有相邻像素之间的整体相关性，从具体实现来看，如果垂直、水平和对角线方向的相关性大约为0，则密码系统的安全水平被认为是良好的。

第三个指标是差分攻击分析，首先简要讲述差分攻击的定义，在差分攻击中，选择特定的普通图

像，通过单个比特进行修改接着通过密码系统进行加密从而研究差异性，为了抵抗这种差分攻击，每次加密中密码系统应当给出不同的加密图像。

第四个指标是信息熵分析，信息熵用于检查源信息，衡量随机性，如果分布均匀，则熵最大为8，在信息熵接近8的情况下，很难预测加密图像。