Eurocrypt'18 - Dupont - Fuzzy Authenticated key Exchange

概述：本文在UC框架下提出了fuzzy PAKE用于解决带噪声的低熵口令的密钥协商问题，它对于口令无熵要求，且只要两个pass-string足够接近即可完成密钥协商。通过Yao's Garbled Circuits和low Hamming distance分别构造了fPAKE，其中后者基于robust secret sharing。本文进行了协议的设计并对安全性和性能进行了比较与分析。

优点：①首次提出fuzzy PAKE来解决带噪声的低熵口令的密钥协商问题
②文中给出了两种构造方法，详细描述了协议的细节，安全性分析和性能比较
③对于研究历史的介绍较好。

问题：①该方案的实践性未知，且作者也未分析应用场景(研究动机较弱)。
-低熵口令若存在字符不一致(噪声)，按照PAKE应无法完成协议，文中的方案无疑让PAKE的敌手的优势增大
②第1个方案假设口令为n个字符，一方面，应分析双方持有口令不等长的情况，另一方面，执行PAKE后产生n个值，泄漏了口令的长度。
③在第2个方案 fPAKE^M 中，方案在A猜测较接近时会泄漏每个字符是否匹配已的信息，比起单纯回答正确与否，A拿到了更多信息，本文没有对此画好敌手的优势
④本文没有与其它类似方案的安全性和性能对比。

基于公共秘密的密钥协商可能有两种complications：
1) pass-string可能来自于非均匀且低熵的分布 → PAKE，生成高熵密钥用于后续通信
2) 双方持有的pass-string可能存在noise，不相等 → information-reconcilication
原有的带有noise的pass-string为高熵，未考虑低熵的场景。
在noise的情况下的构造取决于特定的noise model，最多研究的为binary Hamming distance
→双方实体持有的n bit的字符串中有δ位不同，其余位相同。

研究目标：基于低熵的带有noise的pass-string的密钥协商协议，满足：
①抵抗离线字典攻击 ②可以处理多种noise类型且具有较高的容错能力
③可通过UC框架进行组合。
▲只要敌手无法猜测出足够接近的pass-string即可保证协议是安全的，对pass-string的熵与error的数目的关系无要求
▲Universal Composability (UC)框架的优点
①确保协议在任何环境下运行都安全 ②对pass-string分布的熵无要求或约束。
▲本文在有恶意敌手和无认证通道的情况下构建fuzzy PAKE，
①使用Yao's garbled circuits和oblivious transfer
②使用Hamming distance：双方持有的n个字符的pass-string中不相同的字符较少，双方对n个字符执行PAKE协议，最终得到n个值，其中每一位是否匹配已是未知的

安全模型
▲使用 $d(pw, pw')$ 代表 $pw, pw' \in F_p^n$ 之间的距离，δ为阈值，则 $d(pw, pw') \le \delta$ 时认为 pw, pw' 足够相似。
对于双方协议来说，P_0 和 P_1 允许敌手猜测口令(若足够接近可设置sk)，若 P_i corrupt，则敌手也可以设置它的sk。但是，若 P_{1-i} corrupt，但敌手未猜测到 P_i 的pass-string时，敌手不能设置 P_i 的sk。
▲TestPwd 接口，包含3个泄漏函数：
L_c：对pass-string的猜测足够接近，$d \le \delta$
L_n：对pass-string的猜测较为接近，允许敌手获取部分信息但无法完成协议 $\gamma \ge d > \delta$
L_f：对pass-string的猜测相差较远，猜测失败。
→当收到S的query (TestPwd, sid, P_i, pw_i') 时，若存在1个fresh的record (P_i, pw_i)，则设置 $d \leftarrow d(pw_i, pw_i')$：①$d \le \delta$，record ← compromised，给S响应 $L_c(pw_i, pw_i')$
②$\delta < d \le \gamma$，record ← compromised，给S响应 $L_n(pw_i, pw_i')$ ③$\gamma < d$，record ← interrupted，响应 $L_f(pw, pw')$

根据几种泄漏函数给出4种不同的构造方式：

①不给敌手提供他可信息，$\mathcal{F}_{\text{PAKE}}^{N}$，$L_C^N(pw_i, pw_{i'}) = L_{Im}^N(pw_i, pw_{i'}) = L_f^N(pw_i, pw_{i'}) = \perp$

②提供给敌手猜测的结果，$\mathcal{F}_{\text{PAKE}}$，$L_C(pw_i, pw_{i'}) =$ "correct guess"，$L_{Im}() = L_f() =$ "wrong guess"

③泄漏不匹配之的谓词的下标：$\mathcal{F}_{\text{PAKE}}^{M}$，$L_f^M() =$ "wrong guess"

$L_C^M() = (\exists\ i \text{ s.t. } pw_i[1] = pw_{i'}[1], \text{"correct guess"})$，$L_{Im}^M() = (\{\cdots\}, \text{"wrong guess"})$

④当猜测足够接近时，泄漏真实的 pass-string：$L_C^P(pw_i, pw_{i'}) = L_{Im}^P(pw_i, pw_{i'}) = pw_i$，

$L_f^P(pw_i, pw_{i'}) =$ "wrong guess"　$\mathcal{F}_{\text{PAKE}}^{P}$

使用 Garbled Circuit 构造 $\mathcal{F}_{\text{PAKE}}$ $(\gamma = \delta, \mathcal{F}_{\text{PAKE}}^{P})$

▲ 两个优点：① 比起其它方案更灵活，可以用 circuit 计算的距离均可使用

② $\delta = \gamma$，可在实现功能的情况下保证安全.

▲ 基于 Oblivious Transfer (OT) 和 Yao's Garbled Circuits (YGC)

OT：sender 将两个 secret 中的1个发送给 receiver，receiver 可选择想要的 secret，安全性保证 sender 无法获知 choice bit，receiver 不知道另一个 secret 的信息 (认证通道)

YGC：双方输入敏感数据 (pass-string) 计算输出，输出不会暴露输入的信息，一方 garble 它们要评估的函数，另一方以 garbled 的形式评估.

→ 可抵抗 malicious evaluator，但不可抵抗 malicious garbler (可 mis-garble 函数)

通常使用 cut-and-choose 转换来保证安全，但开销过大

▲ 一种转换方式

①定义 Randomized Fuzzy Equality Functionality，$\mathcal{F}_{\text{RFE}}^{P}$，敌手只能以真实的输入猜测 (即通过 corrupt 参与者) 而不可通过中间人攻击

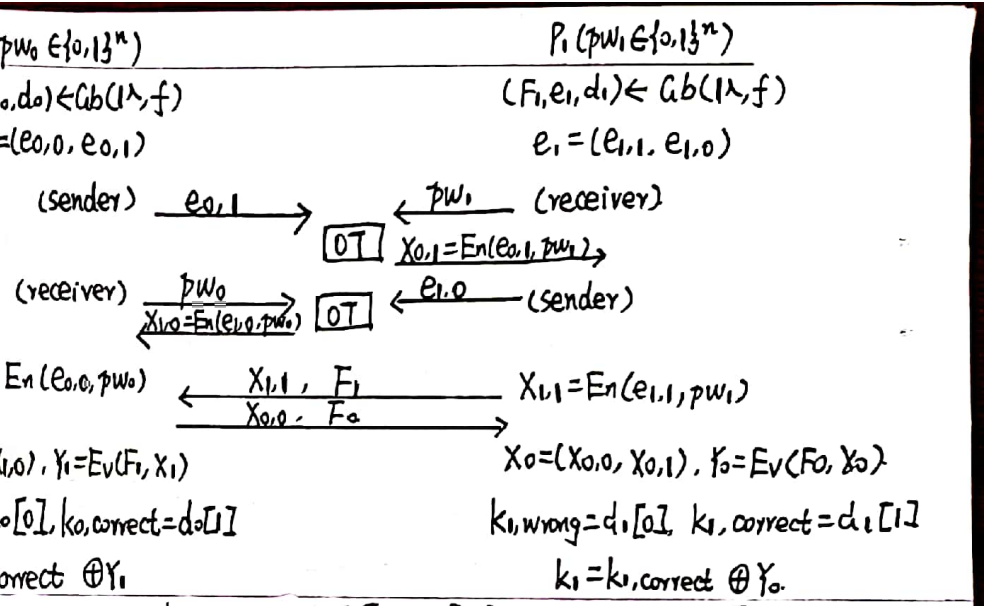
②引入协议 Π_{RFE}，使用 YGC 安全实现 $\mathcal{F}_{\text{RFE}}$，通过两个实体分别扮演 evaluator 和 garbler 来实现安全性，在认证通道实现

在协议过程中，若存在消息未到达或格式不正确，则实体输出随机的 key，

下图的协议中，$e_{0,1}$ 表示 P_0 编码 P_1 的 pass-string 获取的编码信息

pw_0 由 P_0 在本地编码，pw_1 则是通过 OT 编码

▲ output label：在 output-projective garbling 方案中的 garbled 输出，使用 $k_{i,\text{correct}}$ 代表 1，$k_{i,\text{wrong}}$ 代表 0.



③使用 split RFE 构建 $\mathcal{F}_{\text{PAKE}}$ (将需要认证通道的协议转换成不需要的)

SRFE 只能提供会话认证，但不可提供实体认证 (通过签名和验签的方式)

▲ 使用 Hamming Distance 的 Circuit $f_\cdot$ (代表两个等长串中不同字符的位置的个数)

$d(pw, pw') := |\{j \mid pw[j] \neq pw'[j], j \in [n]\}|$

①f 对相应的二进制串执行 XOR，构建标识着相等或不等的 bit 列表.

②f 将这些 bit 填充到 threshold gate 处，若至少 $n-\delta$ 个为 0 则返回 1，否则返回 0

→ 易于 garble，需要 n 条密文，而 $\mathcal{F}_{\text{PAKE}}$ 需 2 个 garbled circuits，故需 $2n$ 条密文交换

使用 Hamming Distance 来构造　$\mathcal{F}_{\text{PAKE}}^{M}$

▲ Robust Secret Sharing (RSS)

对于 vector $C \in \mathbb{F}_q^n$，集合 $A \subseteq [n]$，C_A 为映射 $\mathbb{F}_q^n \to \mathbb{F}_q^{|A|}$，即 sub-vector $(c_i)_{i \in A}$.

$C_A : (C_i)_{i \in \bar{A}}$，其中 $\bar{A} := [n] \setminus A$

→ (n, t, γ)-RSS 有两个算法：Share：$\mathbb{F}_q \to \mathbb{F}_q^n$，Reconstruct：$\mathbb{F}_q^{\gamma} \to \mathbb{F}_q$.

①t-privacy：$\forall s, s' \in \mathbb{F}_q$，$A \subseteq [n]$，若 $|A| < t$，则 $C_A (C \leftarrow\$ \text{share}(s))$ 和 $C'_A (C' \leftarrow\$ \text{Share}(s'))$ 同等分布

②γ-robustness：$\forall\ s \in \mathbb{F}_q$，$A \subseteq [n]$ 有 $|A| \geq \gamma$，则 Share(s) 的输出 C 和 $\tilde{C}$ 有 $C_A = \tilde{C}_A$，Reconstruct 后 s 即只要有 γ 个 share，就可以完成重构

▲ Linear Codes：一个长度为 n，rank 为 k 的 linear code 为维度为 k 的向量空间 $\mathbb{F}_q^n$ 的子空间 C，C 的最小距离 d 为任意两个 code words 的最小距离

$(n,k,d)_q$-code：大小的q，长度为n，rank为k，最小距离为d. 可检测出至多d-1个错误，并纠正至多$\lfloor (d-1)/2 \rfloor$个错误

→ Singleton bound：对于任何 linear code，有 $k+d \le n+1$，则 maximum distance separable (MDS) code 满足：$k+d=n+1$，$d=n-k+1$ 表示为：$(n,k)_q$-MDS code.

→ 通过构造 Share(s)，Reconstruct(w)，将 MDS code 转换成 (n,t,γ)-RSS，其中 $t=k-1$，$\gamma=\lceil (n+k)/2 \rceil$

▲ Implicit-Only PAKE (iPAKE)

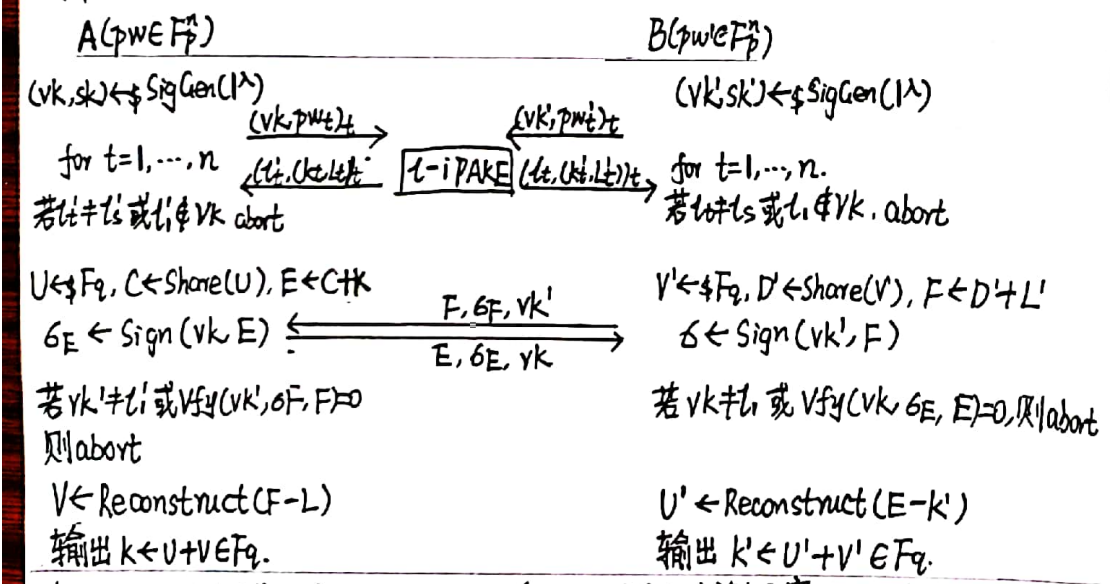
$\mathcal{F}_{iPAKE}$ 与 $\mathcal{F}_{pwKE}$ 的区别

① TestPwd silently 更新记录的内部状态，不会提供给敌手S.

② Newkey query 更改为敌手若 corrupt P_i 但没有猜对 P_{i-1}，则它不可为 P_{i-1} 设置sk.

→ 进一步扩展到 labeled implicit-only PAKE. (t-iPAKE). labels 为公共认证字符串，用于及时检测到协议流程被篡改.

▲构造 (RSS和 t-iPAKE → fPAKE)



其中，Share: $\mathbb{F}_q \to \mathbb{F}_q^n$. (SigGen → vk×sk, Sign, Vfy) 为签名方案，给定输入，从 label space vk和 key space $\mathbb{F}_q^2$ 中重复执行 t-iPAKE

▲为了避免BDK中的问题，未使用pass-string作为 one-time pad，而是通过 t-iPAKE 得到高熵的sk作为 one-time pad

▲首先对 pass-string 的每一个字符执行 t-iPAKE 得到高熵的会话密钥，通过对 nonce 执行 RSS，并使用高熵的sk作为 one-time pad 发送给另一方，RSS 的鲁棒性可保证部分不匹配的字符不会影响构造出 nonce 生成最终的 key

$\mathcal{F}_{PAKE}$RSS 的安全性

若 (Share: $\mathbb{F}_q \to \mathbb{F}_q^n$; Reconstruct: $\mathbb{F}_q^n \to \mathbb{F}_q$) 为 (n,t,γ) RSS 且 (SigGen, Sign, Vfy) 为一个EUF-CMA(Existential Unforgeability under Chosen Message Attack)安全的 one-time 签名方案，则 $\mathcal{F}_{PAKE}$RSS 可以在 $\mathcal{F}_{t\text{-}iPAKE}$-hybrid 模型中安全实现 $\mathcal{F}_{fPAKE}^{M}$.

- 被动攻击：通过RSS保证鲁棒性，再通过 one-time pad 保证 sk 的随机性
- 中间人攻击：通过 labeled 版本的 iPAKE 和 one-time 签名方案来防止
- 主动攻击：通过证明当不使用诚实实体的 pass-string 时也可以像真实执行协议一样输出.