安全性強化的密文可搜尋公開金鑰加密系統

Security Enhanced Public Key Encryption with Keyword Search

丁培毅 吳宗杉 溫玎蕙

國立台灣海洋大學 國立台灣海洋大學 國立台灣海洋大學

Email:pyting@mail.ntou.edu.tw Email:y456@mail.ntou.edu.tw Email:M96570025@mail.ntou.edu.tw

摘 要 — 本文主要改進 Boneh 等人在 2004 年提出的 "可進行關鍵字搜尋的公開金鑰加密系統", 改進的系統裡伺服器必須擁有密鑰才能進行搜尋, 如此可以去除原本系統需要透過安全管道傳送關鍵字暗門的假設, 我們也重新檢視 Baek 在 2008 年所提出的安全性定義, 我們指出其中的弱點, 強化其定義使之具有抵抗 "選擇測試密文 - 關鍵字暗門"攻擊的能力,指出 Baek 的系統在針對此嚴格的安全性定義證明時遇見的困難, 然後我們提出一個改進的系統,並且在隨機智者模型下證明此系統滿足前述強化之安全性定義。

關 鍵 詞 — 關鍵字搜尋, 公開金鑰加密系統, 選擇測試密文-關鍵字暗門攻擊, 隨機智者模型

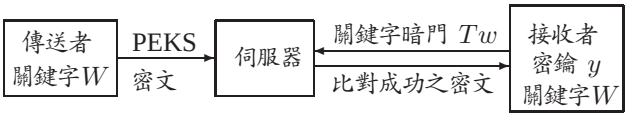
***Abstract*—We propose a security enhanced version of Boneh's "Public Key Encryption with Keyword Search" system. The server in the new system is equipped with a key pair for performing the search operations. This new system eliminates completely the "secure channel" assumption for the keyword trapdoor. We reexamine the security definition by Baek, point out the weakness of it, and strengthen it such that it is secure against chosen test ciphertext - keyword trapdoor attacks. We discuss the problems met by Baek's scheme and propose a modified system with full security proofs according to the enhanced security definition in the random oracle model.**

***Index Terms:*—Keyword Search, Public Key Encryption, Chosen Test Ciphertext - Keyword Trapdoor Attack, Random Oracle Model**

一、簡介

可搜尋的加密系統主要分爲三大類, 第一類是私密資訊擷取 (Private Information Retrieval, PIR) 系統[4], [6], [9] 其中一方提供資料庫給另一方查詢, 查詢的一方除了得到符合條件的資料之外, 無法得到其他的資訊, 另一方面查詢者也希望資料擁有者不會直接看到他的查詢條件, 此類系統通常爲多方安全運算 (Secure Multiparty Computation) 的應用[7]; 第二類是資料儲存伺服器提供承租者儲存、查詢承租者個人擁有的密文資料[10], [5], 此類系統主要運用對稱式的密碼系統加密, 伺服器在搜尋過程中無法得到資料密文或是查詢關鍵字密文的內容; 第三類是可進行關鍵字搜尋的公開金鑰加密系統[3], [1], 這類系統可以運用在安全的電子郵件伺服器上, 基本上有傳送者、接收者和伺服器三個參與者, 任何一個傳送者可以用接收者的公鑰加密訊息傳遞至郵件伺服器, 接收者可以製作關鍵字的暗門傳遞給伺服器以搜尋具有指定關鍵字的電子郵件。

本文探討的系統爲前述第三類的應用, Boneh[3] 等人在 2004 年提出第一套基於雙線性配對 (Bilinear pairing)、支援關鍵字搜尋的公開金鑰加密系統 (Public-key Encryption system with Keyword Search, PEKS), 如圖一所示該系統基於雙線性配對中



系統參數:

大質數 p, 群 G_1, G_2, $|G_1| = |G_2| = p$, g 爲 G_1 的生成元素, 雙線性配對 $e : G_1 \times G_1 \rightarrow G_2$, 雜湊函式 $H_1 : \{0,1\}^* \rightarrow G_1^*, H_2 : G_2 \rightarrow \{0,1\}^k, k = \lceil \log_2 p \rceil$

接收者: 挑選 $y \in_R Z_p^*$, 計算 $Y = g^y$, 公鑰 $pk_r = Y$, 密鑰 $sk_r = y$

密文 $PEKS(pk_r, W)$:

傳送者任選 $r \in_R Z_p^*$,

$S = (U_1, U_2) = (g^r, H_2(e(H_1(W), Y^r)))$

關鍵字暗門 $Trapdoor(sk_r, W)$:

接收者計算 $Tw = H_1(W)^y$

關鍵字比對 $Test(Tw, S)$:$S = (U_1, U_2)$,

伺服器計算並比對 $H_2(e(Tw, U_1)) \stackrel{?}{=} U_2$

圖一. Boneh[3] 之 PEKS 系統

三個人可以運用三方 Diffie-Hellman 金鑰交換協定[8]共享一個秘密數值的特性來設計, 亦即當三個人的公鑰分別爲 g^x, g^y, g^z, 則三人都可以自行計算出共享的秘密 $e(g^y, g^z)^x = e(g^x, g^z)^y = e(g^x, g^y)^z = e(g, g)^{xyz}$, 系統運作時傳送者挑選暫時性、密文相依的密鑰 r 及公鑰 g^r, 接收者的密鑰爲 y, 公鑰爲 g^y, 另外針對密文中的關鍵字 W, 運用雜湊函式計算 $H_1(W)$ 作爲一個輔助的公鑰, $\text{dlog}_g H_1(W)$ 爲對應的密鑰, 沒有任何人知道其數值, 此時傳送者與接收者共享一個秘密 $e(H_1(W), g^y)^r = e(H_1(W)^y, g^r)$, 前者在加密時由傳送者計算, 後者則分爲關鍵字暗門 (trapdoor) $H_1(W)^y$ 由接收者計算, $e(\cdot, g^r)$ 則由保存所有密文的伺服器計算, 如此伺服器不知道搜尋時的關鍵字 W, 接收者也不會得到不含指定關鍵字的文件。

上述系統中由於任何人只要攔截到密文 $(g^r, H_2(e(H_1(W), g^y)^r))$ 以及關鍵字暗門 $H_1(W)^y$, 不需要任何私密資訊即可測試密文中是否含有關鍵字 W, 因此需要假設運用安全的管道來傳送密文或是關鍵字暗門。Baek[1] 在 2008 年針對此問題提出一個解決的方案, 然而因爲在安全性定義裡使用相當弱的攻擊者, 其安全性仍然有待改進。本論文首先探討 Baek 提出的安全性定義; 修改其定義成爲在選擇關鍵字 (chosen keyword)、選擇測試密文-關鍵字暗門 (chosen test ciphertext-trapdoor) 攻擊下之密文不可分辨 (message-indistinguishability) 安全性以抵抗攻擊力較強的敵人; 指出 Baek 的系統在嘗試證明此強化之安全性時遇見的困難; 接著提出一個改進的機制, 並且在隨機智者 (random oracle) 模型[2]下證明此改進的機制符合強化的安全性定義, 如此我們可以設計一個接收者與伺服器各自具有密鑰的系統, 任何第三者縱然攔截密文以及關鍵字暗門仍然無法自行比對測試, 進而提升系統運作的安全性。

由於無法列舉所有可能的攻擊方法, 近二十年來密碼系統的開發相當倚賴理論的證明, 許多沒有安全性證明的系統陸續都遭到破解, 於是很多系統的設計架構中自然地使用到一些安全性證明中需要的元件, 有時甚至不是很容易解釋該元件直覺上的設計理念, 然而因爲任何有心或是無心的疏失都可能被攻擊者運用來破解或是干擾系統正常的運作, 在設計系統時我們還是應該要採取防禦性的安全概念- 也就是由證明中逐步建構可安全運作的系統機制。

本論文第二節中將簡述 Baek 的安全性定義以及其所提出的 SCF-PEKS 系統, 第三節提出強化的安全性定義, 闡述其意涵並且指出 Baek 的機制證明此強化的安全性定義時的問題, 第四節針對強化的安全性提出改良的 SCF-PEKS-1 系統, 第五節證明 SCF-PEKS-1 系統滿足強化之安全性定義, 第六節爲結論。

二、Baek 的 SCF-PEKS 系統與安全性定義

圖二中摘要描述 Baek 的 “不需安全管道 (Secure Channel Free, SCF) 的 PEKS” 系統, 符號上依循 Boneh 的 PEKS 系統以方便比較及閱讀, 因此與 Baek 的文章略有出入。

系統中最主要替伺服器設計了公鑰 pk_s 與密鑰 sk_s, 並且由傳送者與伺服器共享一個秘密 $e(Q, X)^r =$

系統參數:

大質數 p, 群 G_1, G_2, $|G_1| = |G_2| = p$, g 爲 G_1 的生成元素, 雙線性配對 $e : G_1 \times G_1 \rightarrow G_2$, 雜湊函式 $H_1 : \{0,1\}^* \rightarrow G_1^*, H_2 : G_2 \rightarrow \{0,1\}^k, k = \lceil \log_2 p \rceil$

伺服器: 挑選 $x \in_R Z_p^*$, 計算 $X = g^x$, 挑選 $Q \in_R G_1^*$, 公鑰 $pk_s = (Q, X)$, 密鑰 $sk_s = x$

接收者: 挑選 $y \in_R Z_p^*$, 計算 $Y = g^y$, 公鑰 $pk_r = Y$, 密鑰 $sk_r = y$

密文 $SCF{-}PEKS(pk_s, pk_r, W)$:

傳送者挑選 $r \in_R Z_p^*$, $S = (U_1, U_2) = (g^r, H_2(e(Q, X^r) \cdot e(H_1(W), Y^r)))$

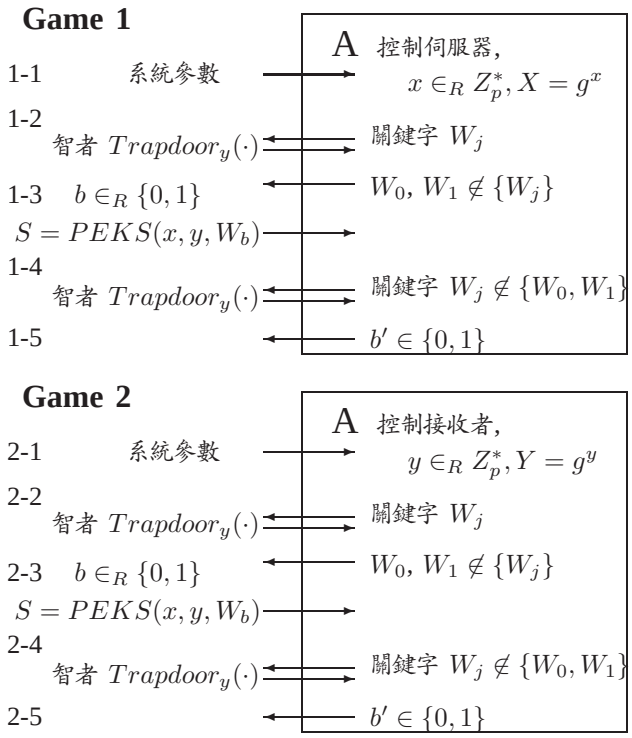
關鍵字暗門 $Trapdoor(sk_r, W)$:

接收者計算 $Tw = H_1(W)^y$

關鍵字比對 $Test(sk_s, Tw, S)$:$S = (U_1, U_2)$,

伺服器計算並比對 $H_2(e(Q^x \cdot Tw, U_1)) \stackrel{?}{=} U_2$

圖二. Baek 的 SCF-PEKS[1] 系統



圖三. Baek 的安全模型

$e(Q, g^r)^x$, 伺服器在比對時需要運用其密鑰才能計算出 $e(Q^x, g^r)$, 乘上 $e(Tw, U_1)$, 再計算 $H_2(\cdot)$ 雜湊函式值即可與 U_2 比對, 如此的設計直覺上的確可以去除安全傳送管道的假設。其安全模型如圖三所示, 主要分爲兩個部份, Game 1 基本上就是 Boneh 的 PEKS 的安全模型, 是一個攻擊者完全控制了伺服器, 擁有可以回答任意關鍵字暗門的智者 (oracle) 的密文分辨模型; Game 2 則是當攻擊者完全控制了接收者, 同樣擁有可以回答任意關鍵字暗門智者的密文分辨模型。Baek 也證明圖二的系統符合圖三的安全性定義, 如果有一個有效率的攻擊者可以破解 Game 1 或是 Game 2, 運用這個攻擊者就可以設計一個有效率的演算法來破解 BDH(Bilinear Diffie-Hellman) 問題。

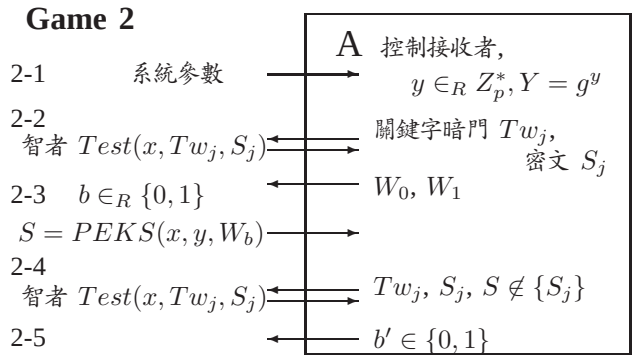
所謂 BDH 問題即爲給定 $< g, g^\alpha, g^\beta, g^\gamma >$ 求取 $e(g,g)^{\alpha\beta\gamma}$ 的問題, 本文中所有系統之安全性都基於 BDH 假設, 亦即假設任意機率式多項式時間的演算法都無法解出 BDH 問題。

當一個同時符合 Game 1 安全性與 Game 2 安全性的系統運作時, 伺服器和接收者必須合作才能夠進行密文與關鍵字暗門的比對, 缺一不可。

三、強化的安全模型

仔細檢討 Baek 的安全性定義, 可以發現模型中所探討的敵人較弱, 在 Game 2 中, 可回答任意關鍵字暗門的智者 $Trapdoor_y(\cdot)$ 基本上是多餘的, 因爲攻擊者 A 完全控制了接收者, 也知道接收者的密鑰 y, 因此攻擊者可以隨時計算任意關鍵字的暗門, 根本不需要詢問該智者, 攻擊者無法藉由詢問該智者而加強其攻擊的能力; 由另一方面來看, 由於攻擊者並不知道伺服器的密鑰 x, 因此攻擊者沒有辦法自行比對密文與關鍵字暗門, 亦即計算 $Test(x, Tw, S)$, 如果能夠如圖四提供可計算 $Test(x, \cdot, \cdot)$ 的智者給攻擊者, 則可以加強攻擊者的攻擊能力, 如此定義的安全模型包含了原來 Baek 的安全性定義, 所得到的安全性較高。

在嘗試證明 Baek 的 SCF-PEKS 具有圖四中 Game 2 的安全性時, 必須在不知道 x 的情況下模擬計算出 $Test(x, Tw_j, S_j)$, 亦即需要計算 $e(Q^x, U_1) \cdot$



圖四. 加強之安全模型

$e(Tw_j, U_1)$, 其中 $S_j = (U_1, U_2)$, Q 可以由挑戰者挑選 δ, 令 $Q = g^\delta$, 如此可以計算 $e(Q^x, U_1) = e(X^\delta, U_1)$, 由於我們假設 H_2 爲隨機智者, 因此如果 Tw_j 即爲 S_j 中的關鍵字 W_j 所對應的暗門, 則 $e(Q^x, U_1) \cdot e(Tw_j, U_1)$ 一定詢問過 H_2 智者, 可以在 H_2-list 中尋找 $t_j = e(Q^x, U_1) \cdot e(Tw_j, U_1)$ 所對應的 V_j, 如果與 U_2 相同則 $Test(x, \cdot, \cdot)$ 智者回答 1 否則回答 0。然而在運用 Game 2 的攻擊者攻擊 BDH 問題 $< g, g^\alpha, g^\beta, g^\gamma >$ 時 (參考圖七及第五節的說明), 由於 $e(Q, X^\gamma) = e(g^\delta, g^{\alpha\gamma})$, 並無法很容易地得到 $e(g,g)^{\alpha\beta\gamma}$ 之目標值, 因此在下一節中我們會引入第三個雜湊函式, 以完成 Game 2 的證明。

四、針對強化安全模型改良之系統

如上節所述, 我們不知道 Baek 的 SCF-PEKS 系統是否符合強化的安全性定義, 如圖五所示, 我們針對上述証明遇見的問題, 提出適當的修正, 使用三個雜湊函式 H_1, H_2, 及 H_3, 如此可以在隨機智者模型下證明此改良系統滿足圖三中 Game 1 以及圖四中 Game 2 的安全定義。

此系統之正確性驗證如下: $H_3(e(H_1(U_1)^x \cdot Tw, U_1)) = H_3(e(H_1(g^r), g^{rx}) \cdot e(H_2(W)^y, g^r)) = H_3(e(H_1(g^r), X^r) \cdot e(H_2(W), Y^r)) = U_2$

SCF-PEKS-1 和 SCF-PEKS 兩個系統之間最主要的差別在於前者的密文中使用 $e(H_1(g^r), X)^r$ 而不是 $e(Q, X)^r$, 概念上這一項的主要目的仍然是由傳送者和伺服器交換一個祕密值, 傳送者可以

系統參數:
大質數 p, 群 G_1, G_2, $|G_1| = |G_2| = p$, g 爲 G_1 的生成元素, 雙線性配對 $e : G_1 \times G_1 \to G_2$, 雜湊函式 $H_1 : G_1 \to G_1^*$, $H_2 : \{0,1\}^* \to G_1^*$, $H_3 : G_2 \to \{0,1\}^k, k = \lceil \log_2 p \rceil$

伺服器: 挑選 $x \in_R Z_p^*$, 計算 $X = g^x$, 公鑰 $pk_s = X$, 密鑰 $sk_s = x$

接收者: 挑選 $y \in_R Z_p^*$, 計算 $Y = g^y$, 公鑰 $pk_r = Y$, 密鑰 $sk_r = y$

密文 $SCF{-}PEKS{-}1(pk_s, pk_r, W)$:
傳送者任選 $r \in_R Z_p^*$, $S = (U_1, U_2) = (g^r, H_3(e(H_1(g^r), X^r) \cdot e(H_2(W), Y^r)))$

關鍵字暗門 $Trapdoor(sk_r, W)$: 接收者計算 $Tw = H_2(W)^y$

關鍵字比對 $Test(sk_s, Tw, S)$:$S = (U_1, U_2)$,
伺服器比對 $H_3(e(H_1(U_1)^x \cdot Tw, U_1)) \stackrel{?}{=} U_2$

圖五. 安全性強化之 SCF-PEKS-1 系統

用 $e(H_1(g^r), X)^r$ 來計算此值, 而伺服器可以用 $e(H_1(g^r), g^r)^x$ 來計算此值, 運用這個共享的祕密值實現只有伺服器可以比對密文與關鍵字暗門的需求, 因此去除了需要透過安全管道傳送關鍵字暗門的假設, 下一節的證明中顯示使用 $H_1(g^r)$ 取代 Q 除了可以在大部分情況下正確模擬 $Test(x, \cdot, \cdot)$ 智者, 同時透過隨機智者 H_1 的模擬也可以順利地將破解 Game 2 的問題轉化爲破解 BDH 問題, 從而建立此系統的 Game 2 安全性。

五、選擇關鍵字攻擊與選擇測試密文-關鍵字暗門攻擊下密文不可分辨之安全性

定理 1. 在 BDH 問題爲計算上無法破解的假設下, 圖五中 SCF-PEKS-1 系統爲隨機智者模型下選擇關鍵字攻擊與選擇測試密文 - 關鍵字暗門攻擊下 密文不可分辨的。

證明.

Game 1:

圖六爲破解 BDH 的演算法 B 之配置, 其中假設

攻擊者 A 完全控制伺服器 (擁有其密鑰 x), 且有不可忽略的優勢 ϵ 分辨 SCF-PEKS-1 密文, B 在取得一個 BDH 問題 $< g, g^\alpha, g^\beta, g^\gamma >$ 後, 在圖三 Game 1 步驟 1-1 中將接收者的公鑰 Y 設爲 g^α, 此時接收者的密鑰即爲未知的 α, B 需要以下列的方法模擬 $H_1(U_{1j})$, $H_2(W'_j)$, $H_3(t_j)$ 三個隨機智者, 其中 q_T 爲詢問 $Trapdoor_\alpha$ 智者的次數上限:

$\underline{\underline{H_1(U_{1j})}}$: 於 Z_p^* 中隨機挑選 a_j, 於 $\{0,1\}$ 中挑選 c_j, 滿足 $\Pr\{c_j = 0\} = \frac{1}{1+q_T}$, 如果 $c_j = 0$ 則令 $h_j = g^\beta \cdot g^{a_j}$, 如果 $c_j = 1$ 則令 $h_j = g^{a_j}$, 記錄 (U_{1j}, a_j, c_j, h_j) 資料於 H_1-list 中。

$\underline{\underline{H_2(W'_j)}}$: 於 Z_p^* 中隨機挑選 a'_j, 於 $\{0,1\}$ 中挑選 c'_j, 滿足 $\Pr\{c'_j = 0\} = \frac{1}{1+q_T}$, 如果 $c'_j = 0$ 則令 $h'_j = g^\beta \cdot g^{a'_j}$, 如果 $c'_j = 1$ 則令 $h'_j = g^{a'_j}$, 記錄 (W'_j, a'_j, c'_j, h'_j) 資料於 H_2-list 中。

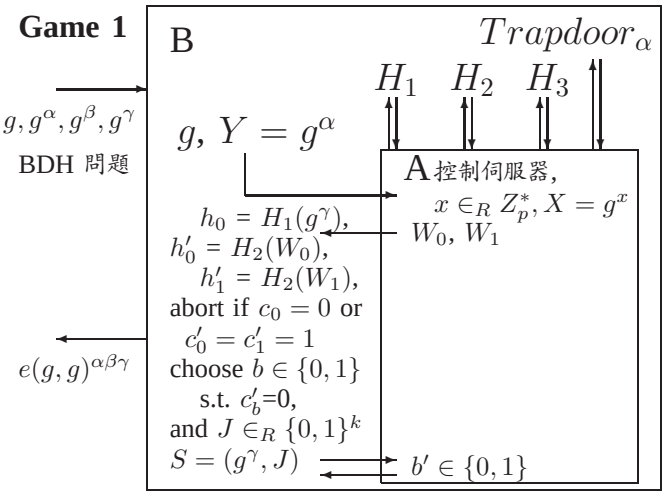
$\underline{\underline{H_3(t_j)}}$: 於 $\{0,1\}^k$ 中隨機挑選 V_j, 記錄 (t_j, V_j) 於 H_3-list 中。

由於 B 不知道 α, 智者 $Trapdoor_\alpha(W) = H_2(W)^\alpha$ 的模擬方法如下:

$\underline{\underline{Trapdoor_\alpha(W)}}$: 於 H_2-list 中尋找 $(W'_j = W, a'_j, c'_j, h'_j)$ 之資料, 如 c'_j 爲 0 則中斷模擬, 如 c'_j 爲 1 則 $H_2(W)^\alpha = \left(g^{a'_j}\right)^\alpha = (g^\alpha)^{a'_j} = Y^{a'_j}$,

攻擊者 A 完成步驟 1-2 後, 在步驟 1-3 中輸出沒有詢問過暗門的任意兩個關鍵字 W_0 與 W_1, 演算法 B 詢問 H_1 及 H_2 智者, 得到 $h_0 = H_1(g^\gamma)$, H_1-list 中對應的資料爲 (a_0, c_0), $h'_0 = H_2(W_0), h'_1 = H_2(W_1)$, H_2-list 中對應的資料爲 (a'_0, c'_0) 與 (a'_1, c'_1), 如果 $c_0 = 0$ 或是 $c'_0 = c'_1 = 1$ 則停止模擬, 否則挑選 $b \in \{0,1\}$, 使得 $c'_b = 0$, 隨機由長度 k 的字串中挑選 J, 令 SCF-PEKS-1 密文 S 爲 (g^γ, J), 傳送給攻擊者 A, A 進行步驟 1-4 之詢問, A 最後輸出 b', 亦即 A 認爲與 S 對應之關鍵字爲 $W_{b'}$。

在上述 A 的攻擊過程中由於 S 的第一部份是 g^γ, 爲了使 b' 等於眞正的 b, 隨機智者模型要求 A 至少需要詢問 H_3 智者下列兩者之一: $e(H_1(g^\gamma), X)^\gamma \cdot e(H_2(W_0), Y)^\gamma$ 或是 $e(H_1(g^\gamma), X)^\gamma \cdot e(H_2(W_1), Y)^\gamma$, 因爲 B 不知道 γ 的數值, 因此 B 無法由 H_3-list 中直接找到該筆詢問記錄, 只能盲目地在 H_3-list 中挑選一個 $(t, \cdot)$, 如果剛好挑到, B 再假設他挑到的 t 對應到 $W_{\hat{b}}$, $\hat{b} \in_R \{0,1\}$, 亦即 $t = e(H_1(g^\gamma), X)^\gamma \cdot e(H_2(W_{\hat{b}}), Y)^\gamma$, 此時由 H_1-list 中尋找 $(U_{1j} = g^\gamma, a_j, c_j, h_j)$, 由 H_2-list 中尋找 $(W'_i = W_{\hat{b}}, a'_i, c'_i, h'_i)$, 則 $t = e(g^{a_j}, X)^\gamma \cdot e(g^{\beta+a'_i}, g^\alpha)^\gamma = e(g^\gamma, X)^{a_j} \cdot e(g,g)^{\alpha(\beta+a'_i)\gamma}$, 可推得 $e(g,g)^{\alpha\beta\gamma} = t/e(g^\alpha, g^\gamma)^{a'_i}/e(g^\gamma, X)^{a_j}$。B 成功地破解 BDH 問題的優勢如下: $Adv_{PEKS,B}^{Game1,CKA} \approx \epsilon \cdot \frac{1}{q_{H_3}} \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{1+q_T}\right)^{q_T} \cdot \frac{q_T}{1+q_T} \cdot \left(1 - \left(1 - \frac{1}{1+q_T}\right)^2\right) \approx \epsilon \cdot \frac{1}{q_{H_3}} \cdot \frac{1}{2} \cdot \frac{1}{e} \cdot \frac{1+q_T}{q_T} \cdot \frac{q_T}{1+q_T} \cdot \frac{q_T^2+2q_T}{q_T^2+2q_T+1} \approx \frac{\epsilon}{2 \cdot e \cdot q_{H_3}}$, 其中 q_{H_3} 爲隨機智者 H_3 的詢問次數上限。



圖六. Game 1 中破解 BDH 的演算法 B 之配置

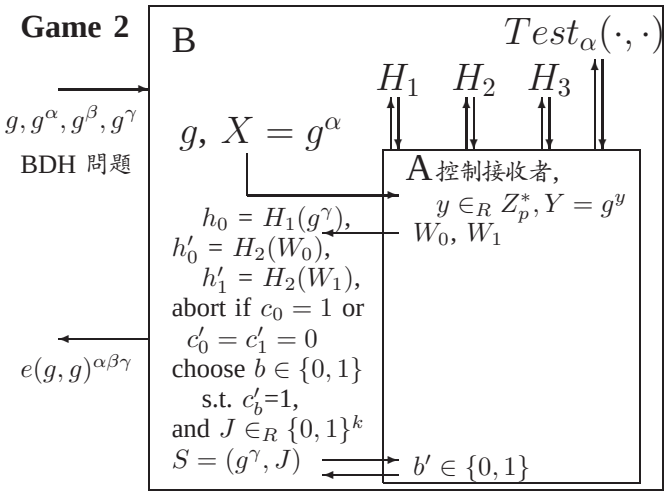
Game 2:

圖七爲破解 BDH 的演算法 B 之配置, 假設其中攻擊者 A 完全控制接收者, 且有不可忽略的機率 ϵ 分辨 SCF-PEKS-1 密文, B 在取得一個 BDH 問題 $< g, g^\alpha, g^\beta, g^\gamma >$ 後, 在圖四 Game 2 步驟 2-1 中將伺服器的公鑰 X 設爲 g^α, 此時伺服器的密鑰即爲未知的 α, B 需要以 Game 1 的證明中模擬 $H_1(U_{1j})$, $H_2(W'_j)$, 和 $H_3(t_j)$ 的方法來模擬 H_1, H_2, H_3 三個隨機智者,

其中 q_T 改爲 A 詢問 $Test_\alpha(\cdot,\cdot)$ 智者的次數上限。由於 B 不知道 α, 智者 $Test_\alpha(Tw,S)$ 的模擬方法如下:

$Test_\alpha(Tw,S)$: 令 $S=(U_1,U_2)$, 由於 Tw 必須運用 H_2 來產生, 因此在 H_2-list 中尋找 (W_j',a_j',c_j',h_j') 滿足 $e(h_j',Y)=e(Tw,g)$, 由於產生 S 時必須同時詢問 H_1 與 H_2 智者, 因此在 H_1-list 中尋找 (U_{1i},a_i,c_i,h_i) 滿足 $U_{1i}=U_1$, 如果 c_i = 1, 則比對 $H_3(e(H_1(U_{1i})^\alpha,U_1)\cdot e(Tw,U_1))=H_3(e(g^{a_i\alpha},U_1)\cdot e(Tw,U_1))=H_3(e(X^{a_i},U_1)\cdot e(Tw,U_1))$ 與 U_2 即爲 $Test_\alpha(Tw,S)$, 如果 c_i = 0 則停止模擬。

攻擊者 A 完成步驟 2-2 後, 在步驟 2-3 中輸出任意兩個關鍵字 W_0 與 W_1, 演算法 B 詢問 H_1 及 H_2 智者, 得到 $h_0=H_1(g^\gamma)$, H_1-list 中對應的資料爲 (a_0,c_0), $h_0'=H_2(W_0),h_1'=H_2(W_1)$, H_2-list 中對應的 (a_0',c_0') 與 (a_1',c_1'), 如果 $c_0=1$ 或是 $c_0'=c_1'=0$ 則停止模擬, 否則挑選 $b\in\{0,1\}$, 使得 $c_b'=1$, 隨機由長度 k 的字串中挑選 J, 令 SCF-PEKS-1 密文 S 爲 (g^γ,J) 並傳送給 A, A 進行步驟 2-4 之詢問, A 最後輸出 b', 亦即 A 認爲與 S 對應之關鍵字爲 $W_{b'}$。在上述 A 的攻擊過程中由於 S 的第一部份是 g^γ, 爲了使 b' 等於眞正的 b, 隨機智者模型要求 A 至少需要詢問 H_3 智者下列兩者之一: $e(H_1(g^\gamma),X)^\gamma\cdot e(H_2(W_0),Y)^\gamma$ 或是 $e(H_1(g^\gamma),X)^\gamma\cdot e(H_2(W_1),Y)^\gamma$, 因爲 B 不知道 γ 的數值, 因此 B 無法由 H_3-list 中直接找到該筆詢問記錄, 只能盲目地在 H_3-list 中挑選一個 $(t,\cdot)$, 如果剛好挑到, B 再假設他挑到的 t 對應到 $W_{\hat b}$, $\hat b\in_R\{0,1\}$, 亦即 $t=e(H_1(g^\gamma),X)^\gamma\cdot e(H_2(W_{\hat b}),Y)^\gamma$, 此時由 H_1-list 中尋找 $(U_{1j}=g^\gamma,a_j,c_j,h_j)$, 由 H_2-list 中尋找 $(W_i'=W_{\hat b},a_i',c_i',h_i')$, 則 $t=e(g^{\beta+a_j},g^\alpha)^\gamma\cdot e(g^{a_i'},Y)^\gamma=e(g,g)^{\alpha(\beta+a_j)\gamma}\cdot e(g^\gamma,Y)^{a_i'}$, 可推得 $e(g,g)^{\alpha\beta\gamma}=t/e(g^\alpha,g^\gamma)^{a_j}/e(g^\gamma,Y)^{a_i'}$。B 成功地破解 BDH 問題的優勢如下: $Adv^{Game2,CTCTA}_{SCF-PEKS-1,B}\approx\epsilon\cdot\frac{1}{q_{H_3}}\cdot\frac{1}{2}\cdot\left(1-\frac{1}{1+q_T}\right)^{q_T}\cdot\frac{1}{1+q_T}\cdot\left(1-\left(\frac{1}{1+q_T}\right)^2\right)\approx\epsilon\cdot\frac{1}{q_{H_3}}\cdot\frac{1}{2}\cdot\frac{1}{e}\cdot\frac{1+q_T}{q_T}\cdot\frac{1}{1+q_T}\cdot\frac{q_T^2+2q_T}{q_T^2+2q_T+1}\approx\frac{\epsilon}{2\cdot e\cdot q_{H_3}\cdot q_T}$, 其中 q_{H_3} 爲隨機智者 H_3 的詢問次數上限。 ∎



圖七. Game 2 中破解 BDH 的演算法 B 之配置

一個系統滿足 Game 1 的安全性表示: 如果攻擊者沒有取得接收者運用密鑰所計算的關鍵字暗門, 他就算知道伺服器的密鑰 x, 也沒有辦法比對密文與關鍵字, 進而搜尋整個資料庫裡面的資料; 一個系統滿足 Game 2 的安全性則表示: 如果攻擊者沒有伺服器的密鑰 x, 就算知道接收者的密鑰 y, 也沒有辦法直接比對密文與關鍵字, 進而搜尋整個資料庫。

六、結論

本論文主要加強 Baek 不需要安全管道假設、關鍵字可搜尋的公開金鑰加密系統的安全性定義, 使得安全模型中的攻擊者可以進行適應式的 "密文-關鍵字暗門 測試" 詢問, 指出 Baek 的 SCF-PEKS 系統在證明此強化的安全定義時的困難, 提出一個修改過的系統 SCF-PEKS-1, 並且証明在隨機智者模型下可以滿足強化的安全性定義。相較之下, 原本 Boneh 的 PEKS 就算運用 SSL 或是其他密碼機制來實作安全管道, 安全地傳送關鍵字暗門, 這樣子的系統仍然無法滿足本文中 Game 2 的安全性, 因爲攻擊者只要完全控制接收者, 他就可以自行比對密文與關鍵字。在應用層面上來說, 本文所提出的 SCF-PEKS-1 系統和 Boneh 有實作安全管道傳送關鍵字暗門的 PEKS 有一些差異, 前者的應用中接收者沒有伺服器幫忙的話, 無法自行搜尋, 後者則

可, 因此需要視實際應用而有所取捨。而 Baek 的 SCF-PEKS 和本文的 SCF-PEKS-1 則是屬於同樣類型的應用, 本文提出的 SCF-PEKS-1 系統具有較佳的安全性。

七、致謝

本論文相關研究承蒙中華電信研究所 (計畫編號:TL-98-1501) 及行政院國科會 (計畫編號:NSC 97-2221-E-019-014) 經費補助, 得以順利完成, 特此致謝。

八、參考文獻

[1] J. Baek, R. Safavi-Naini, and W. Susilo, "Public key encryption with keyword search revisited", in Proc. of 2008 International Conference on Computational Science and its Applications, *ICCSA 2008*, LNCS 5072, pp.1249–1259, 2008.

[2] M. Bellare and P. Rogaway, "Random oracles are practical: A paradigm for designing efficient protocols", in Proc. of the 1st ACM conference on Computer and Communications Security, *CCS 1993*, pp.62–73, 1993.

[3] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search", in Proc. of *Advances in Cryptology - Eurocrypt 2004*, LNCS 3027, pp.506–522, 2004.

[4] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval", in Proc. of the 36th annual symposium on Foundations of Computer Science, *FOCS 1995*, pp.41–50, 1995.

[5] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions", in Proc. of the 13th ACM conference on Computer and Communications Security, *CCS 2006*, pp.79–88, 2006.

[6] G. Di Crescenzo, T. Malkin, and R. Ostrovsky, "Single-database private information retrieval implies oblivious transfer", in Proc. of *Advances in Cryptology - Eurocrypt 2000*, LNCS 1807, pp.122–138, 2000.

[7] W. Du and M. J. Atallah, "Secure multi-party computation proglems and their applications: a review and open problems", in Proc. of New Security Paradigms Workshop, pp.11–20, 2001.

[8] A. Joux, "A one round protocol for tripartite Diffie-Hellman", in Proc. of the 2000 Algorithmic Number Theory Symposium, *ANTS 2000*, LNCS 1838, pp.385–394, 2000.

[9] Y. Lindell and B. Pinkas, "Privacy preserving data mining", in Proc. of *Advances in Cryptology - Crypto 2000*, LNCS 1880, pp.36–54, 2000.

[10] D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data", in Proc. of the 2000 IEEE symposium on Security and Privacy, *S&P 2000*, pp.44–55, 2000.