学校编码：10384　　分类号_____密级_____
学　　号：23120101152971　　UDC_____

厦门大学

硕　士　学　位　论　文

基于口令的密钥导出算法安全性分析

On the Security Analysis of Password-Based Key Derivation Function

周　君

指导教师姓名：李晓潮　副教授
专　业　名　称：电子与通信工程
论文提交日期：2013 年 5 月
论文答辩时间：2013 年 6 月
学位授予日期：2013 年　　月

答辩委员会主席：_____
评　　阅　　人：_____

2013 年　月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（　　）课题（组）的研究成果，获得（　　）课题（组）经费或实验室的资助，在（　　）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年　　月　　日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（　　）1.经厦门大学保密委员会审查核定的保密学位论文，于　　年　　月　　日解密，解密后适用上述授权。

（　　）2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年　　月　　日

摘 要

文件存储的安全性主要是通过身份认证和数据加密实现，在身份认证和数据加密密钥的产生中都需要由基于口令的密钥导出算法生成的密钥，生成密钥的随机性决定了文件安全机制整体的安全性，所以基于口令的密钥导出算法是文件认证机制中的关键。基于口令的密钥导出算法分为 PBKDF1 和 PBKDF2。目前，新的文件认证机制中往往推荐使用密钥导出函数 PBKDF2 来取代 PBKDF1，从而获得更高的安全性。为此，本文以 OpenOffice 的安全性为例，证明了其中密钥导出算法 PBKDF2 的理论安全性，并详细分析了 OpenOffice 的实际安全性。主要工作成果如下：

（1） 在 CCS 安全定义下，利用随机预言机模型与 Game-Playing 技术界定了密钥导出函数 PBKDF2 与随机函数不可区分优势上界 $Adv^{prf}_{PBKDF2}(t) < \frac{\lfloor t/c \rfloor}{|PW|} + \frac{t^2}{2^n}$，其中 t 代表攻击者询问次数，c 代表迭代次数，|PW|代表口令空间大小，n 是导出密钥长度；在 CPM 安全定义下，本文证明了 PBKDF2 存在的安全缺陷，证明了迭代次数 c 在 CPM 安全定义下并没有提高 PBKDF2 的安全性，其与随机函数不可区分优势上界 $Adv^{prf}_{PBKDF2}(t) < \frac{t}{|PW|} + \frac{t^2}{2^n}$；最后提出几种改进方案，并论证各个方案的安全性；根据对 PBKDF2 的理论安全性证明结果，给出 OpenOffice 的口令认证机制的理论安全论断，当 t<<c|PW|时，OpenOffice 的口令认证机制是安全的；

（2） 在目前主流 Nvidia 显卡上利用 CUDA 技术，对 OpenOffice 进行并行口令穷举攻击，结合实验结果分析了 OpenOffice 加密文件在并行口令穷举攻击下的实际安全性。利用以上的两个成果，在使用 OpenOffice 加密文档时，尽量选择由数字、字母、特殊符号共同组成的字符集，并且字符长度应不小于 7 位来增大口令空间，从而获得较高的安全性。

关键词：密钥导出算法；随机预言机模型；并行计算

ABSTRACT

The file security is mainly based on identity authentication and data encryption. The password-based key derivation function (PBKDF) is used to generate derived key during authentication and encryption, the randomness of a derived key is mainly influence the whole file security mechanism, so password-based key derivation function is the key point in file authentication mechanism. Password-based key derivation function is consist of PBKDF1 and PBKDF2. Nowadays, more and more applications prefer to use PBKDF2 instead of PBKDF1, which could produce a derived key more than 160 bits long and be more secure. Thus, in the paper we take OpenOffice security as an example to prove the theoretical security of PBKDF2, and analysis the practical security of OpenOffice. The main results are as follows:

(1) In CCS security mode, we prove the theoretical security of PBKDF2 in the random oracle by using Game-Playing technology to quantify the upper bound of Adversary's Advantage between PBKDF2 and ideal random function satisfies $Adv^{prf}_{PBKDF2}(t) < \frac{\lfloor t/c \rfloor}{|PW|} + \frac{t^2}{2^n}$, where t is the number of query, c denotes iteration count, $|PW|$ is password space, n is the derived key length; In CPM security mode, we find a security flaw in PBKDF2, which proves the iteration count c doesn't improve the security of PBKDF2. Thus, the security upper bounder of Adversary's Advantage is $Adv^{prf}_{PBKDF2}(t) < \frac{t}{|PW|} + \frac{t^2}{2^n}$; Finally, we propose some cases to improve the security of PBKDF2 and give the security discussion. According to the PBKDF2 security analysis, we give a security conclusion: when t<<c|PW|, OpenOffice password authentication is secure.

(2) We use CUDA technology to make parallel password exhaustion attack on OpenOffice by mainstream Nvidia graphics card. With the theoretical and experimental results, we recommend that a secure password should be choosen from a mixing up character set of numbers, letters and special symbols, and the length of a password should longer than 7.

Key Words: Password-based Key Derivation Function; Random Oracle Model; Parallel Computing

目 录

CONTENTS

第1章. 绪论

1.1. 研究背景和意义

当今，我们正处于一个信息大爆炸的时代，随着网络和信息技术的高速发展，越来越多的信息依赖于个人 PC 和移动终端的存储，同时出现了各种存储技术：如分布式文件系统、云存储、高速磁盘阵列等，但这些存储技术都离不开最基本的存储介质——计算机硬盘以及其他便携式大容量存储设备，这些存储介质作为信息的主要载体，在给信息存储带了巨大便利的同时也带来了许多安全隐患。不法攻击者可以通过多种攻击手段，如窃听、截取或者盗取等方法获取敏感信息。这些信息一旦泄露，可能会给个人、企业和国家带来巨大的损失。因此，如何最大限度地保证存储设备上数据的安全性，正受到越来越多的企业与信息安全研究机构的关注。

在文件存储系统中，为了保护敏感信息的安全性，主要通过数据加密和身份认证两种技术手段实现[1]。

1.1.1. 数据加密技术

数据加密技术是将敏感信息（或称明文，Plain text），经过加密密钥（Encryption key）与加密函数转换，变成无意义的密文（Cipher text）进行存储，当合法用户需要使用数据时再将此密文经过解密函数和解密密钥（Decryption key）还原为明文形式，从而有效的减小了敏感信息直接泄露的危险性。数据加密的实现一般分为软件实现和硬件实现两大类，硬件实现一般通过特定的设备实现特定功能，而在文件存储系统中，较多采用软件实现数据加密，比如目前主流的 PDF[2]、Microsoft Office[3]、7-zip[4]、OpenOffice[5]等应用软件都自带了文件加密功能，像 PGP WDE[6]、BestCrypt[7]、Truecrypt[8]等专业存储加密软件中采用了 on-the-fly 的加密技术[9]，即当用户存取存储设备上的数据时，加密系统自动、透明的对数据进行实时加密或解密处理。通过软件实现的数据加密可分为对称加密和非对称加密两种方式，对称加密算法的特点是运算相对非对称加密解密简单来说、计算量小、

加密速度快，主要应用于需要加密大量数据的情况，而非对称加密算法安全性是基于复杂数学难题，运算复杂、速度慢，主要应用于对通信安全性极高的系统。由于两种加密方式的各自特点，且对称加密算法比非对称加密算法有着更快的运算速度，因此在对速度要求更高的存储加密系统中通常采用对称加密算法来对敏感数据内容进行加解密处理。对称加密算法又可分为流加密和分组加密。在 Microsoft Office 2003 及 PDF 的早期版本中使用 RC4 流加密算法，在 7-zip、PGP WDE、PGP SDA、OpenOffice 与 Truecrypt 中则采用 AES[10]、Twofish[11]、CAST5[12] 等分组加密算法。与流加密算法相比，分组加密算法更广泛的应用于存储加密系统的数据加密。

无论是对称加密还是非对称加密，它们对所需的密钥具备一定要求，理想的密钥是具备足够的信息熵在密钥空间呈随机均匀分布，这样的密钥往往需要通过密钥导出算法将源密钥材料（包括用户口令，密钥的概率分布信息，源密钥泄露的信息等）进行特定的算法处理，生成一个或者多个具有足够信息熵且随机均匀分布的加密密钥。因此，数据加密技术中密钥导出算法的导出密钥随机性成为存储加密系统中数据安全的核心问题。

1.1.2. 身份认证技术

身份认证是文件加密系统中数据机密性的另一个重要的认证方式，是证实用户的真实身份与其所声称的身份是否相符，以防止非法用户通过身份欺诈访问系统资源的过程。虽然身份认证方法各异，但基本可以分为以下三类[13]：

（1） 用户知道的信息（something the user knows），如口令，这是最传统也是最广泛的身份认证方法，一般用“口令”,“用户名+口令”,“用户名+口令+验证码”的形式；

（2） 用户所拥有的东西（something the user has），如动态口令卡、令牌等物理介质[14]，这些介质中通常存储了多组一次性密码，安全认证证书或者包含能根据用户参数生成动态密码的系统；

（3） 用户的身体特征（something the user is），如脸部识别、虹膜等物理信息[15][16]，它是以用户生物特征作为网上识别身份的要素，以唯一识别用户身份。

这三种身份认证方法有各自的优缺点[17][18][19]。相比于动态口令卡和生物