# 手机端F2FS镜像文件数据取证方法

翟雨佳，李涛，胡爱群
〔东南大学网络空间安全学院，江苏南京 211189；网络通信与安全紫金山实验室，江苏南京 211111；网络空间国际治理研究基地（东南大学），江苏南京 211189〕

**摘 要：** F2FS文件系统是一种基于NAND存储设备设计的新型开源Flash文件系统，具有良好的性能和Flash友好的特性，得到越来越多手机厂商的关注。在手机数据取证中，人们对F2FS文件系统镜像文件的数据解析相关研究还很少，支持F2FS文件系统近期删除文件的数据解析相关研究也还不充分。文章根据F2FS文件系统特有的读写访问机制和手机图形化界面的垃圾回收机制，提出一种F2FS文件系统镜像文件的通用数据解析取证框架。实验结果表明，该方法可以准确解析F2FS文件系统的目录结构，实现近期删除文件的数据取证，有效支撑F2FS文件系统的取证研究。

**关键词：** F2FS文件系统；镜像文件数据取证；删除文件数据取证
**中图分类号：** TP309.3 **文献标识码：** B

# Data forensics method of mobile terminal F2FS image file

Zhai Yujia[1,2,3], Li Tao[1,2,3], Hu Aiqun[1,2,3]
〔*School of Cyber Science and Technology, Southeast University, Jiangsu Nanjing 211189; Purple Mountain Laboratories for Network and Communication Security, Jiangsu Nanjing 211111; Research Base of International Cyberspace Governance （Southeast University）, Jiangsu Nanjing 211189*〕

**Abstract:** The Flash Friendly File System (F2FS) is a new type of open source Flash file system based on NAND storage device design, with good performance and Flash-friendly features. In the mobile phone data forensics, there are few related researches on data analysis of F2FS mirror files, and related research on the analysis of F2FS's recent deleted data is also insufficient. According to the unique read-write access principle of F2FS and the garbage collection mechanism of the graphical interface, this paper builds a general forensics framework for data analysis of F2FS mirror files. The experimental results show that this F2FS mirroring forensics method can effectively analyze the directory structure and files data. The method has high accuracy and can be effective support the forensic research of F2FS.

**Key words:** F2FS; data forensics of image files; data forensics of deleted files

## 1 引言

F2FS文件系统（Flash Friendly File System）是一种新型开源Flash文件系统[1]，和传统的Linux文件系统相比，F2FS文件系统更好发挥了闪存存储设备的性能优势。华为率先将F2FS文件系统加入到自己的产品中，随后大量厂商也采用该文件系统来取代Ext4文件系统[2]。到目前为止，有多家厂商的移动设备已经使用该文件系统，包括三星Galaxy Note 10、华为Mate20、Moto Z等。有

些厂商的设备虽仍然使用Ext4文件系统，但支持从Ext4到F2FS转换。随着F2FS文件系统在手机端的广泛应用，构建手机端F2FS文件系统通用取证框架具有重大意义。

一般情况下，手机文件系统的取证主要是对手机文件系统的镜像文件进行数据解析。美亚柏科公司把F2FS文件系统镜像文件视为一个磁盘，通过模拟Linux挂载磁盘，构建文件系统组织架构信息，实现解析整个F2FS文件系统目录结构[3]。该方法需要根据位图信息遍历整个文件系统，虽然有效解决了可移植性问题，但是解析大容量文件系统时，解析效率明显下降。梁效宁等人对于已删除文件提供了一种边恢复边解析删除文件元数据的方法，分别遍历F2FS文件系统的段摘要区（SSA）和主区（Main Area），比较文件合并恢复的同时完成对删除文件的解析[4]。梁效宁等人所提出的方法只恢复并解析删除文件，并不涉及解析文件系统中现存文件，而且分别遍历SSA和Main Area区域也需要花费更长的时间，效率较低。综上，目前对F2FS文件系统镜像文件的数据解析相关研究还不充分。

本文根据F2FS文件系统特有的读写访问机制以及垃圾回收机制，提出一种对F2FS文件系统镜像文件数据解析的通用取证方法。实验结果表明，本文提出的F2FS文件系统镜像取证方法可以有效的实现对F2FS文件系统的目录结构解析，完成对近期删除文件的数据解析，对电子取证具有实际应用意义[5,6]。

# 2 F2FS文件系统概述

## 2.1 F2FS文件系统布局结构

F2FS文件系统中最小的数据存储单元为块（Block），每个块的大小默认为4K[7]。逻辑上，F2FS文件系统分为元数据区域（Meta Area）和主数据区域（Main Area）两部分。

元数据区域包括超级块（SB）、检查点（CP）、节点地址表（NAT）等五个部分。元数据区域一般存储在系统空间最开始的多个连续块中，主要存储F2FS文件系统基本参数配置、段的使用分配信息、数据块地址等基础信息。

主数据区域分为节点段区域和数据段区域两部分。主数据区域主要用来存储用户文件信息，其中与文件属性相关的信息一般存放于节点段区域，文件的数据信息存放于数据段区域。F2FS文件系统通过Node-Date的索引关系管理文件，节点段区域和节点地址表之间存在映射关系[8]。具体布局结构如图1所示。

## 2.2 文件元数据信息块

F2FS文件系统采用多级索引机制保存大文件数据，采用内联机制保存小文件数据，文件元数据信息块是实现这些机制的重要结构。

文件元数据信息块又称为inode_block，每个文件元数据信息块都对应一个普通文件或者目录文件，普通文件和目录文件的元数据信息块的区别在于普通文件在元数据信息块内偏移量为0x168处存储文件物理内存地址，而目录文件在偏移量为0x168处存储该目录下子文件的ino节点号和对应子文件的文件名称等信息。一般普通文件元数据信息块的数据结构如表1所示，目录文件元数据信息块的数据结构如表2所示。

表1 普通文件元数据信息数据结构

| 块内偏移 | 字节长度 | 内容描述 |
|---|---|---|
| 0x10 | 8 | 文件大小 |
| 0x20 | 24 | 时间戳 |
| 0x5C | 255 | 文件名称 |
| 0x168 | 3692 | 923组索引地址 |
| 0x0FE8 | 20 | 文件标识号nid |
| 0x0FEC | 20 | 文件节点号ino |

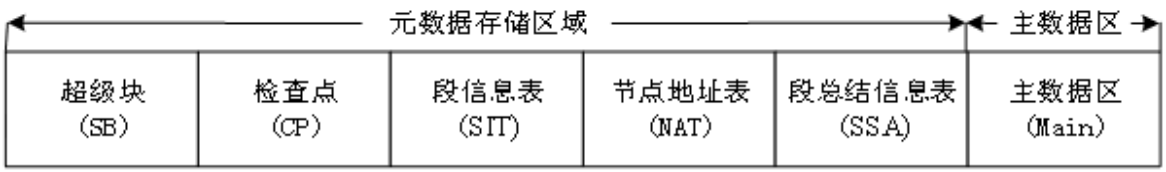


图1 F2FS文件系统布局结构

表2 目录文件元数据信息块数据结构

| 块内偏移 | 字节长度 | 内容描述 |
|---|---|---|
| 0x20 | 24 | 时间戳 |
| 0x58 | 4 | 目录名的字节长度 |
| 0x5C | 255 | 目录名称 |
| 0x168 | 3692 | 目录子文件信息 |
| 0x0FE8 | 20 | 目录文件标识号nid |
| 0x0FEC | 20 | 目录文件节点号ino |

目录文件元数据信息块中偏移量为0x168处目录子文件数据信息区存储了182个子文件目录项，用来描述目录下子文件的信息，如子文件名的哈希值及子文件节点号等。

节点地址表中的nat_entry表项存储了主数据区中各文件对应的文件节点号ino和节点数据块的地址信息。结合节点信息表和子文件目录项，即可构建当前目录文件的目录树。

## 2.3 近期删除数据解析原理

用户手动删除文件后并不会直接清除F2FS文件系统中的文件内存数据块，只是改变该文件相关标志位，清除nat_entry表项和文件元数据信息块的对应关系，只有当新的数据写入到该删除文件的内存数据块后，原有的信息才会因被覆盖而无法完整还原[9,10]。考虑到用户手机端删除文件操作等同于用户在图形界面删除文件，根据在图形界面删除文件并不直接删除文件而是将文件先放置在.Trash目录下的特点，可以采用寻找并解析.Trash目录文件，并结合有效节点地址表的方法，实现对用户近期删除文件的数据解析。

本文所指的解析近期删除文件主要包括获取删除文件的元数据信息块、原存储路径信息等文件基本信息。解析.Trash目录文件结合节点信息表的近期删除文件数据取证方法主要分为三个步骤。

（1）确定.Trash目录文件节点号ino以及文件元数据信息块地址：.Trash目录文件属于F2FS文件系统根目录文件的子文件，可以根据2.2节原理解析根目录文件的元数据信息块，确定.Trash目录文件节点信息号ino，获取该ino所对应的文件元数据信息块地址。

（2）解析.Trash目录文件元数据信息块：通常情况下，.Trash目录文件下的info普通子文件存储被删除文件的原存储路径；.Trash目录文件下的files子目录文件存储被删除文件的节点号ino和文件名称等基本信息。

（3）获取被删除文件的元数据信息块和原存储路径：结合节点地址表，解析删除文件的元数据信息块，得到删除文件的删除时间、存储地址等信息，实现删除文件取证。

## 2.4 F2FS文件系统镜像文件

镜像文件是将特定的一系列文件按照一定格式制作成的单一文件。文件系统镜像文件不仅包含了文件系统的数据文件，还包含了文件系统的基本属性信息和分区信息，文件系统镜像文件包含了文件系统所有信息[11]。

本文模拟手机文件系统，制作了多组F2FS文件系统镜像文件用作实验研究和结果分析，主要分为三类镜像文件：

（1）F2FS文件系统刚被创建，仅包括几个单一文件的初态F2FS镜像文件；

（2）模拟多次重复挂载、创建或修改已有文件、移动文件位置等用户操作，创建已使用一段时间后的F2FS镜像文件；

（3）存在各种不同类型或不同大小文件的大型F2FS镜像文件。

针对以上不同类型的镜像文件，运用WinHex分别获取超级块只读区域，分析校验F2FS镜像文件的健全性和有效性。超级块区校验的具体过程为：

（1）魔数标志位“0x1020F5F2”检查；

（2）扇区大小等默认参数检查；

（3）根索引节点号及各元数据索引检查。

若超级块和备份超级块校验均不通过，则表明文件系统发生损毁或镜像文件操作有误。

# 3 F2FS镜像文件目录项取证分析

为实现对镜像文件目录项的取证，首先应获取当前F2FS的有效节点地址表，然后根据有效节点地址表的nat_entry表项获取解析文件和文件夹

的元数据信息块，逻辑关系完成目录树的构建，本章将主要从基于根目录文件时间戳获取有效节点地址表法、通用F2FS镜像文件目录树取证方法这两个方面展开[3]。

## 3.1 基于根目录文件时间戳获取有效节点地址表法

为防止宕机对元数据造成不可恢复的损害，一般F2FS文件系统中存有两个及以上节点地址表，但只有一个有效节点地址表存储最新的节点信息。由于F2FS文件系统的jounal缓存机制的影响，存在有效nat_entry表项未更新到节点地址表的现象[12]。根据节点地址表的写入原理，通过检查点段定位到最新有效的节点地址表是镜像文件目录项取证方法中主要的技术难点。

针对上述现象，本文提出一种基于根目录文件时间戳获取有效节点地址表方法，该方法的实现流程图如图2所示。基于根目录文件时间戳获取有效节点地表址方法的步骤为:

（1）从超级块中确定各个节点地址表和版本号较高的有效检查点段的起始地址;

（2）在有效检查点段内查找根目录节点号ino（0x03），记录元数据信息块的地址，标记为根目录元数据信息块Ⅰ;

（3）在各个节点地址表中查找根目录节点号ino（0x03），分别记录各元数据信息块的地址，依次标记为根目录元数据块Ⅱ、Ⅲ等;

（4）比较根目录元数据信息块Ⅰ、Ⅱ、Ⅲ等数据块中时间戳，获取最新时间戳根目录元数据块所在的节点地址表和检查点段的缓存nat_entry表项;

（5）若节点地址表段的根目录时间戳最新，则节点地址表即为当前有效的节点地址表。若检查点段的根目录时间戳最新，需手动将检查点段的缓存nat_entry表项更新到节点地址表，得到当前有效的节点地址表;

（6）若出现根目录时间戳相同的情况，再比较根目录下的子文件的时间戳信息。

## 3.2 F2FS镜像目录树取证方法

对F2FS文件系统镜像文件进行目录取证最重要的步骤是确定各目录项间的链接关系，构建最新的目录树结构。根据上述分析，可以得到一种通用的F2FS文件系统镜像文件目录树取证方法，该方法取证流程图如图3所示。

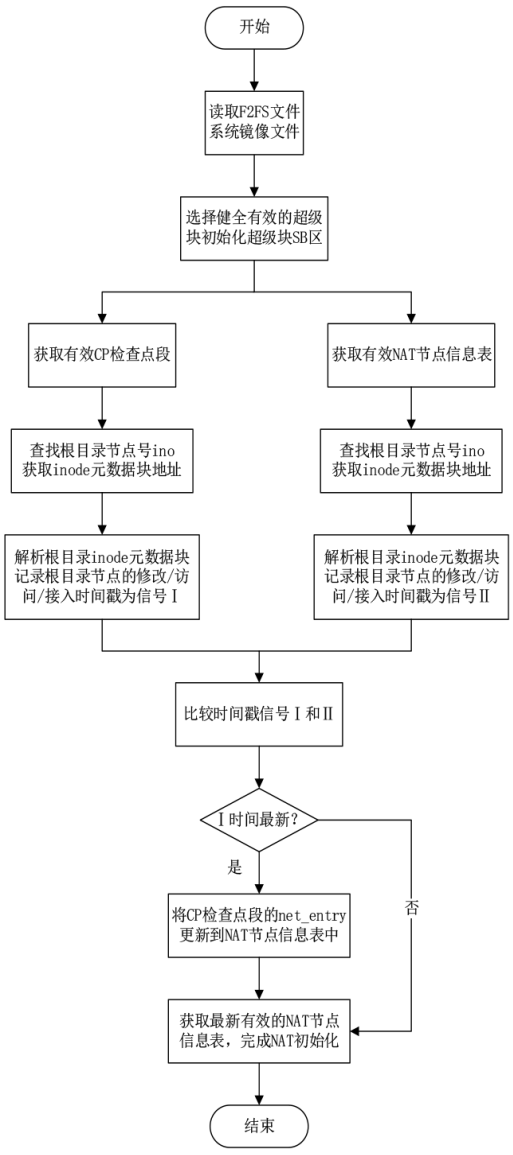

图3 FCL办理流程

通用的F2FS文件系统镜像文件目录树取证方法的具体步骤流程为:

（1）载入F2FS镜像文件，读取镜像文件超级块，确定检查点段、各节点地址表地址;

（2）运用根目录文件时间戳方法得到有效节点地址表并读取nat_entry表项;

（3）以根目录作为镜像文件目录树起点，在有效节点地址表查找根目录文件节点号ino，解析根目录的元数据信息;

（4）在有效节点地址表查找根目录各子文件

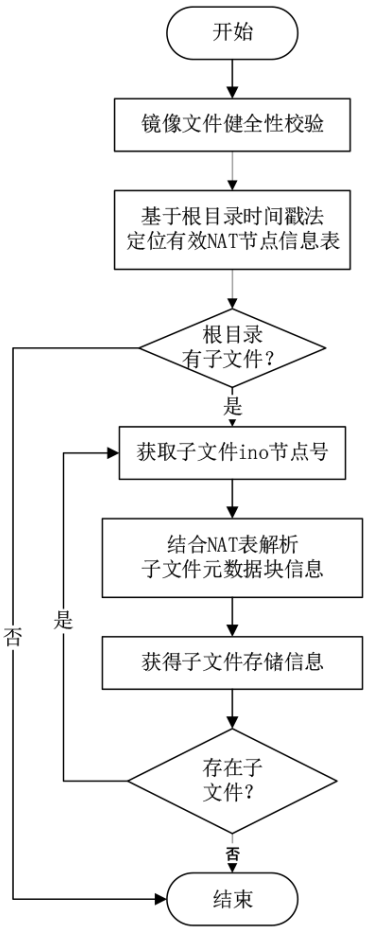


图3 通用F2FS镜像文件目录树取证流程图

节点号ino，解析各子文件的元数据信息；

（5）不断重复对每个子文件进行解析，构建完整F2FS文件系统镜像文件目录树。

# 4 F2FS镜像文件的删除文件取证分析

删除文件的数据取证主要是指获取用户近期删除文件的元数据信息和文件内容[4]。在F2FS镜像文件目录结构解析的基础上，进一步提出一种F2FS镜像文件删除文件取证方法，该方法具体流程如图4所示。

F2FS文件系统镜像文件删除文件取证方法的具体操作步骤为：

（1）读取镜像文件超级块，通过根目录文件时间戳的方法获取有效的nat_entry表项；

（2）结合有效节点地址表获取根目录的.Trash子文件元数据信息；

（3）解析.Trash文件元数据信息块及info和files子文件的元数据信息块，获取被删除文件的

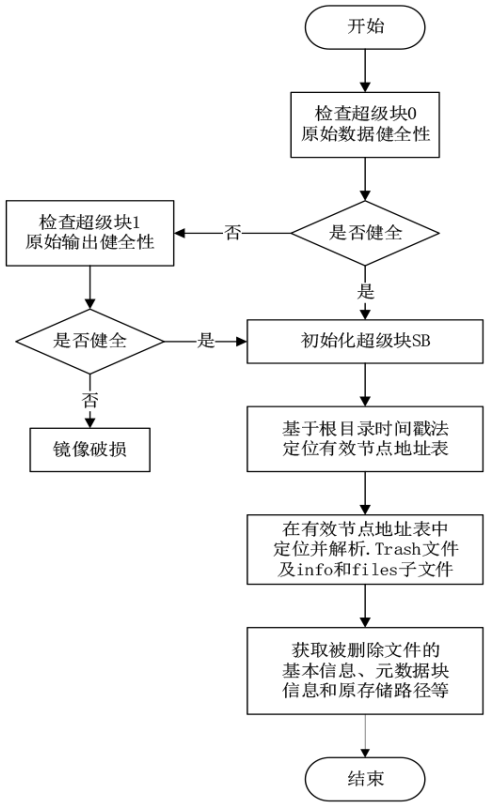


图4 通用的F2FS镜像文件删除文件取证方法

元数据信息块和原存储路径。

# 5 系统实现与结果分析

## 5.1 镜像文件目录树取证结果

运用F2FS镜像文件目录树取证方法，分别构建不同类型的F2FS镜像文件目录树结构，获取各个文件的基本信息，包括文件名称、大小、物理存储地址等。

初态F2FS镜像文件目录解析结果如图5所示，结果表明该初态F2FS文件系统有一个文件夹和两个文件。每个文件的基本信息在取证结果报告一栏有详细说明。

使用一段时间后的F2FS镜像文件目录解析结果如图6所示，结果表明使用过后的F2FS包括myfolder、test、picture等多个文件夹、5个文本文件和13个图片文件。每个文件的基本信息在取证结果报告一栏均有详细说明。

大型F2FS文件系统镜像文件目录解析结果如图7所示，结果表明大型F2FS包括300个文件夹，文件夹下存放多个文本文件。文件的基本信息在

图5 初态F2FS镜像目录解析结果

图6 使用一段时间后F2FS镜像文件目录解析结果

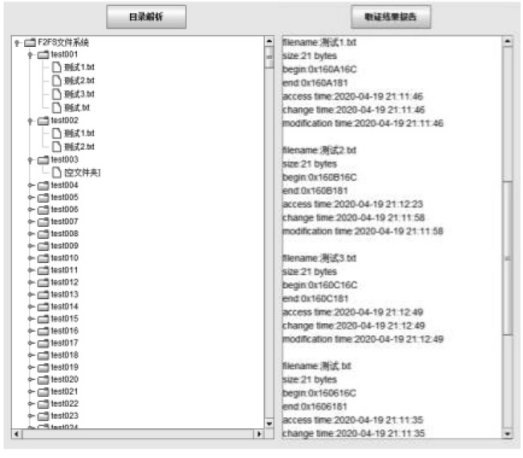

图7 多文件大型F2FS镜像文件目录解析结果

取证结果报告一栏有详细说明。

不同Linux系统之间文本编译器的差别，文本文档的名称上可能有乱码现象。上述结果表明本文提出的F2FS镜像文件目录树取证方法可以实现对不同类型F2FS镜像文件构建目录树，获取各文件信息，验证了该方法可行性。

## 5.2 镜像文件删除文件取证结果

删除多个文件的F2FS镜像文件近期用户删除文件取证结果如图8所示，结果表明大型F2FS文件系统内存在3个用户近期删除文件，被删除文件的名称和对应的原存储路径信息、被删除时间、文件大小、文件存储地址等基本信息在取证结果报告一栏有详细说明。

图8 F2FS多文件删除取证结果

上述结果图表明本文提出的手机端F2FS镜像文件数据取证方法可以准确定位并获取近期被删除文件的元数据信息。

## 5.3 平均取证时间分析

如表3所示，本文提出的手机端通用F2FS镜像文件取证方法通过定位有效文件夹以及.Trash文件夹快速解析获取用户当前目录结构和近期删除的所有文件数据信息。本文方法的平均取证时间约等于定位各个文件夹的时间之和，时间数量级控制在毫秒级别，且本文方法的平均算法时间只和当前系统中文件的数量有关，不受文件系统自身大小的影响，因此本文方法解析镜像文件的取证效率更高。尤其是在系统容量较大、初态文件

表3 不同方法的平均取证时间

| 镜像类型 | 本文取证时间 | 遍历法取证时间 |
|---|---|---|
| 初态F2FS镜像文件 | 110ms | 580ms |
| 使用一段时间F2FS镜像文件 | 130ms | 580ms |
| 大型F2FS镜像文件 | 340ms | 580ms |

系统以及已经反复使用过一段时间的手机端F2FS文件系统中，该方法同目前广泛使用的磁盘遍历搜索法相比更高效。

## 6 结束语

本文根据F2FS特有的读写访问机制和手机图形化界面的垃圾回收机制提出一种F2FS镜像文件数据解析的通用取证方法。通过多组不同类型镜像数据取证实验表明，本文提出的F2FS镜像取证方法可以有效解析F2FS镜像文件目录结构，准确快速实现删除文件的取证操作，对F2FS取证具有重要的意义。

**基金项目：**

1.国家自然科学基金“基于量化可信模型的信息系统智能安全机制研究”（项目编号：616011131004911）；

2.信息通信与安全前沿科学研究中心资助；

3.至善青年学者支持计划资助。

**参考文献**

[1] C. Lee, D. Sim, J-Y. Hwang, S. Cho, F2FS: A New File System for Flash Storage [C]. Proceedings of the 13th Usenix conference on File And Storage Technologies (FAST' 15), 2015: 273-286.

[2] 张辉极.基于Flash存储的智能手机文件系统解析[J].电信科学,2010(S2):39-44.

[3] 潘泽汇;陈明辉;张辉极,等.一种通用F2FS文件系统解析方法、终端设备及存储介质:中国，201711338307.2[P].2018-05-29.

[4] 梁效宁,许超明,赵飞,等.一种恢复F2FS文件系统中被删除文件的方法：中国，201711485304.1[P].2018-06-12.

[5] 刘卫华.大数据环境下的电子取证研究[J].科技创新与应用,2018,000(035):75-76.

[6] 黄少荣,陈丹.大数据环境下的电子取证研究[J].网络空间安全,2019,10(07):79-82.

[7] Liang Y , Fu C , Du Y , et al. An empirical study of F2FS on mobile devices[C]. 2017 IEEE 23rd International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA). IEEE, 2017: 1-9.

[8] 邓傲松.闪存友好型文件系统性能优化技术的设计[D].重庆:重庆大学,2018.

[9] 杨泽明,许榕生,刘宝旭.文件删除的恢复与反恢复[J].信息网络安全,2002,4:38-41.

[10] 刘春枚.基于计算机取证的Linux文件系统解析与设计[J].中国测试,2013, 39(S2):108-111.

[11] 李继伟.计算机取证磁盘镜像研究与虚拟仿真实现[D].重庆:重庆邮电大学,2016.

[12] 艾绍新.Windows数据恢复技术在电子取证中的应用研究[D].大庆:东北石油大学,2013.

**作者简介：**

翟雨佳（1996-），女，汉族，山西晋中人，东南大学，硕士；主要研究方向和关注领域：网络内生安全。

李涛（1984-），男，汉族，江苏镇江人，东南大学，博士，东南大学网络空间安全学院，副教授；主要研究方向和关注领域：可信计算、移动终端安全。

胡爱群（1964-），男，汉族，江苏如皋人，东南大学，博士，东南大学，教授；主要研究方向和关注领域：无线通信安全。