文件运动轨迹追踪技术的研究及实现①

王文宇 (中国科学院 研究生院 北京 100049)

摘 要： 详细阐述了文件运动轨迹追踪技术的实现方法，并提供基于该技术的文件运动轨迹追踪系统的解决方案。利用 Windows 文件系统过滤驱动对文件透明加密，并追踪文件运动轨迹。追踪文件离开工作安全域后，无法使用；在工作安全域内，只允许满足安全策略的进程操作追踪文件，其他进程则拒绝访问。对追踪文件的任何操作，文件运动轨迹追踪系统均记录日志并上传至服务器，既保证实时监控文件流向，杜绝危险性的操作，又便于以后对操作信息进行统计分析和审计。

关键词： 文件系统过滤驱动；透明加密；运动轨迹；追踪；审计

Design and Implementation of Tracing File Trajectory

WANG Wen-Yu

(The Chinese Academy of Science, Beijing 100049, China)

Abstract: In this article, the implementation of tracing file trajectory is described in details. The solution of the tracing system based on tracing file trajectory is also provided. Taking advantages of transparent encrypt, which is based on the windows file system filter driver, the file trajectory is tracing as follows: the file is unable to be accessed once the tracing file is taken out of the working security domain. Only when the file in the security domain and the process is secure at the same time, it can be accessed by the handler. The tracing system will record all information about the operations on the file, besides log of the submission to the server. According to the processes above, the file trajectory is tracing on time, and hence can put an end to the dangerous operations. It is convenient to analyze and audit the results.

Keywords: file system filter driver; transparent encrypt; trajectory; tracing; audit

合理有效的管理涉密文件，是信息安全领域的重中之重。传统管理纸介质涉密文件，通过严格的登记、审批等保密制度，并确定涉密责任人，进行文件的收文、发送、借阅、修改及销毁。

随着计算机应用的普及和互联网日新月异的发展，电子文件在信息领域占据的比重日益增大。电子文件与纸介质文件相同，也涉及保密、安全、管理、归档、使用等问题。同时，电子文件具有复制和传递时的高保真性、非直读性、对硬件和软件的依赖性、信息与载体相对分离性、易更改性以及信息共享的便利性等特征，这些特征无形中增加了电子文件管理和维护的复杂性，也向管理者提出严峻的挑战。如何对涉密电子文件进行现代化、标准化、规范化的管理，并确保其在整个生命周期中的安全分发和使用，是时代赋予电子文件管理的历史使命。因此，如何防止电子文件泄密，有效监控文件的传播、修改、销毁，是信息安全领域亟待解决的问题。

管理涉密电子文件，需从以下几个方面保证电子文件的安全：阻断网络黑客的攻击，以防止涉密信息被恶意窃取、篡改和破坏；完善内部控制策略，以防止用户有意或无意的传播和泄密；若发生威胁文件的操作，则跟踪、监控文件的流转，保证文件在生命周期内的安全。近几年，终端涉密文件的保护已得到广泛深入的研究[1-3]。

目前，保护主机终端电子文件安全的方法主要包括以下几种：

① 收稿时间:2009-07-31;收到修改稿时间:2009-09-16

(1) 加密文件

通过对文件加密，可确保电子文件内容的秘密性。电子文件的加密方法有多种，如全盘数据加密，特定进程透明加解密等。全盘数据加密，因存在系统盘数据未加密等漏洞，其推广度和普及面受到一定程度的制约和限制。透明加解密[4]是目前电子文件保护的新热点。

(2) 控制文件的传播途径

该方式通过控制文件的泄密途径，如网络途径、打印途径、复制途径和外设接口等，达到保护文件的目的。随着公司企业信息化建设的推进，办公自动化已迅速发展。电子文件这一浩大的信息流，具有网络间共享和交流的基本要求，若控制网络途径，电子文件便失去其共享的重要作用。而禁用外设端口，则无法控制和管理文件在网络间的传播。

(3) 格式转换

该方式将涉密电子文件转化成特定的格式。电子文件完成格式转换后，访问权限已受到控制，如将文件设置为“只读”状态。在此状态下，用户只能读取信息，无法对其做任何修改。此方式可有效防止用户非法更改电子文件内容，保证电子文件的原始性和真实性。

以上方法在一定的程度上达到保护文件的目的。目前，运用上述方法实现防泄密的产品，仍缺乏一套完善的防范体系，无法在文件整个生命周期内实现对文件的实时追踪和审计。文件运动轨迹追踪则填补这一空白。通过透明加解密和电子审批，结合系统化的管理，对文件在生命周期内进行全程的运动轨迹追踪。

文件运动轨迹追踪是保护文件安全，实现灾情评估的基础，是信息安全领域新的迫切需求。对文件在整个生命周期中的变化以及运行轨迹进行跟踪和审计，以监控不符合公司安全策略的非法行为，从而为公司安全策略、规则制度的完善提供依据，实现电子文件现代化、标准化、规范化的管理，以健全的信息安全体系保证涉密信息的安全性和共享性，进一步推进办公自动化的进程。

1　文件轨迹追踪系统的功能

电子文件运动轨迹追踪系统，通过实时记录电子文件的操作，追踪电子文件的运动轨迹，达到监控电子文件的运动状态、审计非法人员(或程序)窃取用户权限对文件实施越权操作、给管理员提供有力证据以追究泄密责任人的目的。

电子文件运动轨迹追踪系统运行在被监控机上，实时记录进程对受控文件的访问。对标识追踪标志位的文件进行读、写、删除、关闭、打印、复制、粘贴、另存为、剪切、重命名等操作时，追踪系统记录文件的 GUID、操作用户、Agent ID、操作时间、文件完整路径、文件调用进程等相关信息，并将相关信息保存、上传至服务器，后台则根据文件操作的日志信息，对文件运动轨迹进行还原，有效监控文件的操作和流向，实现对文件整个生命周期内的监控。

文件运动轨迹追踪系统应具备以下功能项：

(1) 设置文件运动轨迹的追踪标识

文件运动轨迹追踪系统，简称为追踪系统，只针对特定的文件进行跟踪。通过追踪系统，设定涉密电子文件的追踪标志，该过程称之为设置追踪标识。

设置追踪标识后的文件，称之为追踪文件。追踪系统根据当前文件是否具有追踪标识，决定是否进行轨迹追踪，以获取精确、实用的信息，实现涉密电子文件的高效率管理。

(2) 监控文件操作

追踪文件运动轨迹时，追踪系统监控文件的多种操作，并对已监控的操作进行日志记录。文件的主要操作如下：

① 创建：文件创建时，追踪系统根据策略决定是否进行运动轨迹追踪。若需追踪，则自动设置追踪标识，并记录、上传日志。记录的内容包括文件的 GUID、操作用户、Agent ID、操作时间、文件名、文件完整路径、文件调用进程等相关信息。

② 读：任何进程访问追踪文件时，均记录访问信息。访问包括查看文件属性、打开文件等读文件操作。已授权的进程具有读权限，而未授权的进程则拒绝访问。若以 windows PE 系统或者第二操作系统登录被监控机，此时追踪系统未启动，通过以下两种方式防止非法读操作：①直接拒绝对追踪文件任何形式的访问；②通过木马技术在文件中加入相关的控制信息，一旦发现非法读操作，即自动删除文件。若已联网，则向服务器发送日志。

③ 写：通过监控保存动作，追踪文件写操作。授权进程保存追踪文件时，追踪系统接收文件内容已更

新的消息，即记录写操作日志并传至服务器。以文件打开——关闭为一个完整周期，同一周期内的写操作，追踪系统上传一次监控日志，以保证服务器收集到的信息规范、有效，便于以后查询和审计。

④ 另存为：追踪系统监控文件另存为的操作，记录日志并传至服务器。文件另存后，其属性与源文件一致，包括追踪标识、文件 GUID 等信息。对另存后的文件的所有操作，均可通过追踪系统进行再跟踪。

⑤ 打印：打印方式包括本地打印、网络打印等。以任何方式打印追踪文件，追踪系统均记录日志并上传。

⑥ 复制：复制追踪文件时，追踪系统记录文件源文件名、源文件路径、源文件的 GUID 以及目标文件名、目标文件路径、用户名、应用进程名、复制时间、操作结果、Agent ID。目标文件属性与源文件一致，如追踪标识，文件 GUID 等。对目标文件的操作，记录日志并上传至服务器，该轨迹信息归并至源文件的 GUID 中。

⑦ 重命名：重命名追踪文件，追踪系统记录源文件名、目标文件名、GUID、操作结果等信息传至服务器。

⑧ 删除：删除操作是文件生命周期内的终点。删除追踪文件时，追踪系统记录文件名、文件路径、文件的 GUID、时间、操作结果、Agent ID 等信息传至服务器。

(3) 监控文件位置移动

监控文件操作是追踪文件运动轨迹的核心，而文件的移动行为是追踪系统的重要组成部分。

文件位置改变，追踪系统详细记录移动轨迹，包括移动的方式，如以移动存储设备方式转移，或以邮件方式发送至网络或其他终端，并将详细日志上传至服务器。

(4) 监控系统剪贴板

文件之间的剪贴，严重威胁涉密文件的安全。对文件剪贴板的监控是解决剪贴板泄密的关键方法。剪贴板内容在文件间流转时，追踪系统进行如下监控：

追踪文件间的剪贴　追踪文件间剪贴时，追踪系统记录源文件名、源文件路径、源文件的 GUID、目标文件名、目标文件路径、目标文件 GUID、用户名、时间、操作结果、事件描述、Agent ID 等信息并传至服务器。

非追踪文件与追踪文件间的剪贴　非追踪文件内容剪贴至追踪文件，追踪系统记录源文件名、源文件路径、源文件的 GUID、目标文件名、目标文件路径、目标文件 GUID、用户名、时间、操作结果、事件描述、Agent ID 等信息并传至服务器。追踪文件内容剪贴至非追踪文件，追踪系统禁止该操作。

非追踪文件间的剪贴，追踪系统不记录日志。

(5) 控制进程

在追踪系统中，进程分为涉控进程和非涉控进程。涉控进程指允许打开追踪文件的进程。同时，追踪系统禁止非涉控进程访问追踪文件。

(6) 系统审计

通过文件名 GUID、事件类型、时间范围、Agent ID、IP 地址等单一条件或组合条件进行审计。相同事件可归并，以防大量非重要的相同事件充斥审计结果，影响关键信息的查询和提取，从而有效跟踪文件整个生命周期中的运动轨迹。

通过不同条件的查询，审计结果如图 1 所示。横轴表示时间，原点表示文件初始信息，即文件设置追踪标识的时间；节点表示对文件进行的具体行为或操作，连线表示文件运动轨迹。以事件类型为基础，不同颜色代表各种文件操作。当鼠标定位至某节点，则显示该事件的详细信息，如文件 GUID、操作时间、Agent ID 等。

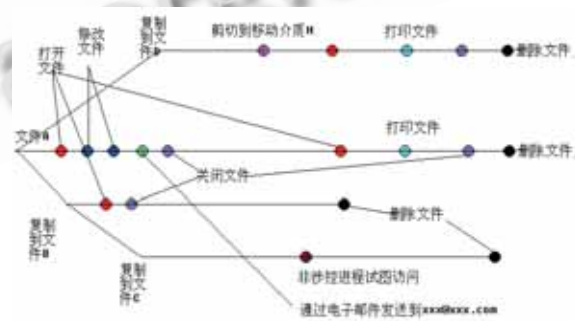

图 1　文件运动轨迹追踪审计结果

2　系统的组成

文件运动轨迹追踪系统通过客户端(Agent)、服务器、控制台三部分分工合作，实现上述功能。

(1) 客户端(Agent)

安装在被监控机，与服务器进行通信。客户端读

取监控策略信息和涉控进程，设置文件跟踪标识、记录文件操作日志、上传日志到服务器等。

(2) 服务器

服务器保存监控策略和涉控进程列表、监控文件操作信息、接收客户端上传的日志以及其他辅助功能。

(3) 控制台

控制台编辑监控策略、下发监控策略、编辑涉控进程列表、日志审计、跟踪文件移动轨迹等。

3 关键技术

3.1 透明加解密技术

文件运动轨迹追踪系统追踪的文件，均经过 windows 文件系统过滤驱动[5]进行强制加密，即透明加解密技术[4]。透明加解密技术对文件进行强制加密后，保证在工作安全域内不影响操作者对文件的访问，在工作安全域外任何人均无法获取文件内容，从而防止用户主动或者被动的泄密。透明加解密工作原理如图 2 所示。

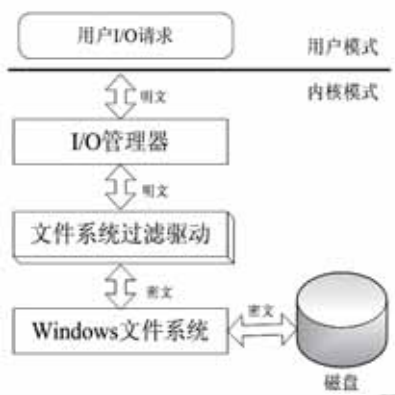

图 2　透明加解密技术工作原理

当用户保存文件时，应用层将用户 I/O 请求发送至 I/O 管理器，I/O 请求到达文件系统过滤驱动后，过滤驱动对数据进行加密，windows 文件系统接收加密后的数据并传至存储设备，从而保证文件在磁盘上密文存储。当用户读取文件时，文件系统过滤驱动解密数据，用户获得明文。文件加解密过程在后台自动运行，根据服务器下发的安全策略，由文件系统过滤驱动识别文件是否可明文访问。

3.2 设置文件访问安全属性

追踪文件由加密的文件内容和文件访问安全属性组成。文件访问安全属性置于加密内容的尾部，包含文件跟踪标识和文件全局唯一标识符 GUID 以及若干保留位。

文件访问安全属性结构体定义如下：

```
Struct FileAccessAttribute {
    UNIT SetAttribute; //  设置默认值，标识文件已设置安全属性
    UINT GUID;        //文件全局唯一标识符
    UINT RESERVE；//保留位
}
```

通过 HOOK API 控制文件操作函数，从而监控各种文件操作。解析文件跟踪标识时，追踪系统根据安全策略识别涉控进程操作的文件是否设置文件追踪标识。文件追踪标识的设置按以下步骤进行：

① 判断是否存在文件追踪标志位

② 判断是否需要设置文件追踪标志位

③ 系统进程通过与文件系统过滤驱动交互，生成文件访问安全属性并设置文件跟踪标识，同时将文件唯一标识符 GUID 和文件路径记录到日志中并上传至服务器。

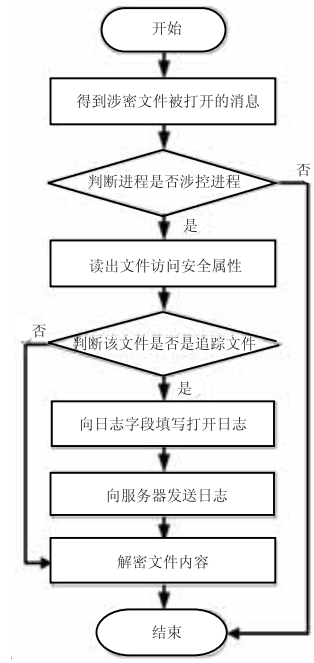

图 3　文件运动轨迹追踪系统追踪文件的流程示意图

客户端用户操作文件时，追踪系统追踪文件的流程如图 3 所示。文件系统过滤驱动判断当前进程是否为涉控进程。若当前进程为非涉控进程，则拒绝访问

该文件。若当前进程为涉控进程，文件系统过滤驱动读取文件访问安全属性，解析文件是否含追踪标识。若无追踪标识，截除文件访问安全属性并解密文件内容，用户可自由操作该文件，且无须向服务器发送日志；若解析到追踪标识，可确定该文件为追踪文件。追踪系统监控用户对文件的打开操作并将日志上传至服务器。文件系统过滤驱动截除文件访问安全属性，解密文件内容，此时用户已获得文件信息。

3.1 系统消息监控的实现

系统消息监控模块主要实现剪贴板消息的监控和HOOK API 子模块的启动。通过创建特定线程实现系统消息的监控。

该线程的主要工作流程包括：RegisterClass(注册窗体类，设置窗体消息回调函数)——〉Create-Window(创建窗体)——〉GetMessage(循环获取窗体消息，利用 TranslateMessage，DispatchMessage 处理获得的消息)。

发送至窗体 hwnd(CreateWindow 返回的窗体句柄)的消息由窗体消息回调函数处理，在该回调函数中完成剪贴板消息监控及 HOOK API 子模块的启动。各消息处理机制如下：

WM_CREATE 消息：①通过函数 SetClipboard-Viewer(hwnd)设置该窗体为剪贴板消息监控窗体。②加载挂钩 API 函数的动态链接库文件，通过函数 GetProcAddress 获取该动态链接库文件中建立全局钩子的函数 InstallHook(),在该函数中建立全局 WH_CBT 钩子，启动 HOOK API 模块。③通过 EnumWindows 函数刷新已经打开的进程窗口，从而使全局钩子对当前所有正在运行的进程有效。

WM_DRAWCLIPBOARD 消息：当用户进行系统剪贴板操作时接收到该消息。

3.3 剪贴板监控的实现

Windows 粘贴板是一种简单且开销较小的 IPC(InterProcess Communication，进程间通讯)机制。Windows 系统支持剪贴板 IPC 的基本机制，该机制实现的方式是由系统预留一块全局共享内存，暂存各进程间交换的数据。提供数据的进程创建一个全局内存块，并将要传送的数据移到或复制到该内存块，接受数据的进程(也可以是提供数据的进程本身)获取此内存块的句柄，并完成对该内存块数据的读取。为了实现上述功能，Windows 提供了存放于 USER32.dll 中的一组 API 函数、消息和预定义数据格式等，并通过对这些函数、消息的使用来管理在进程间进行的剪贴板数据交换。

系统调用粘贴板的简化步骤如下：首先通过调用 OpenClipboard 函数打开剪贴板，GetClipboard-Data 函数获取剪贴板的内容；若设置剪贴板内容，则先调用 EmptyClipboard 函数清空剪贴板，然后调用 SetClipboardData 设置内容(在获取和设置粘贴板内容的函数的参数中均用相应的数据格式)。文件运动轨迹追踪系统对剪贴板的监控，主要通过监控表 1 中的 API 函数的方式实现。监控过程如下：

表 1 剪贴板的 API 函数

API函数	函数功能描述
OpenClipboard	打开剪贴板
EmptyClipboard	清空剪贴板
SetClipboardData	设置剪贴板数据
GetClipboard	获取剪贴板数据
CloseClipboard	关闭剪贴板
IsClipboardFormatAvailable	检查剪贴板中某一数据格式是否存在
CountClipboardFormats	获取目前剪贴板所存在的数据格式数目

① 在挂钩 API 函数前，利用函数 CreateWindow 建立用于监控剪贴板信息的全局窗口(hClip)。根据发送到窗口的消息类型，窗口回调函数进行相应处理。WM_COMMAND 消息中，处理剪贴板保护状态的切换，该消息所附带的参数 wParam 中则包含启动或取消剪贴板保护信息(wParam 的低十六位)。

②替代 SetClipboardData 的 mySetClipboard-Data 伪函数的处理机制

首先判断源文件是否设置追踪标识，然后调用系统 API，根据是否设置追踪标识决定是否记录日志并上传至服务器。日志包括文件名、文件路径、文件的 GUID、复制的内容、内容格式、用户名、用户组、用户权限、应用进程名、应用进程 ID、应用进程权限、时间、函数类型、函数返回结果、事件描述、IP、agent ID。

③ 替 代 GetClipboardData 的 myGetClip-boardData 伪函数的处理机制

首先判断目标文件是否设置追踪标识，然后调用系统 API，并根据是否设置跟踪标志位来决定是否记录日志并上传至服务器。

3.4 打印监控功能实现

文件打印监控模块，建立独立线程实时监控受控终端的文件打印作业。监控线程启动后，利用时间轮

询的方式实施监控功能。一旦发现追踪文件打印作业，即记录日志并上传至服务器。日志内容包括打印的文件名、文件路径、文件的 GUID、用户名、用户组、用户权限、时间、事件描述、IP、Agent ID。

4 结论

本文针对电子文件易泄密和难管理的突出问题，详细阐述了通过文件运动轨迹追踪技术跟踪和监控涉密电子文件的解决方案。该方案以透明加解密技术为基础，通过文件跟踪标识确立跟踪目标、记录详细跟踪信息并上传日志，以便于审计和管理。

通过文件运动轨迹追踪系统，监控、跟踪、记录用户对文件的所有操作，实现最高的系统安全。分析上传的日志，可从庞大的数据信息中抽取有用信息，对用户操作进行分类整理。通过对用户当前操作的跟踪，可实时发现非法或者危险的操作，在泄密事件发生前发出警报，及时阻止泄密。一旦泄密事件发生，通过审计可回溯历史活动，发现泄密渠道，在第一时间取得有力证据。

对文件运动轨迹进行追踪，能够有效防止敏感信息的泄密，为实现电子文件现代化、标准化、规范化的管理提供必要的保障，对信息安全的发展具有重大的现实意义。

参 考 文 献

1 许访.沈昌祥. Windows2000／XP 安全文件系统的设计与实现.计算机工程与应用，2004,40(15):l07－109．

2 黄革新．Windows 加密文件系统核心技术分析.电脑与信息技术, 2005,13(4):1－4.

3 于江,苏锦海,张永福.基于 Windows 2000 的存储加密系统设计与实现.计算机应用，2006,26(5):1084－1086．

4 陈尚义.透明文件加解密技术及其应用.信息安全与通信保密, 2007,11:75－77.

5 Nagar R, Mason T. Windows NT File System Internals.