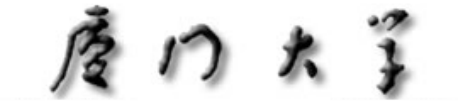
学校编码：10384　　分类号_____密级_____
学　　号：22320051302486　　UDC_____



硕　士　学　位　论　文

电子邮件调查分析系统的设计与实现

The Design and Implementation of Email Investigation and Analysis System

曾 春 溪

指导教师姓名：　吴 顺 祥 教授

专 业 名 称：　系　统　工　程

论文提交日期：　2008 年 4 月

论文答辩时间：　2008 年 5 月

学位授予日期：　2008 年　　月

答辩委员会主席：_____

评　　阅　　人：_____

2008 年 5 月

厦门大学学位论文原创性声明

兹呈交的学位论文，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的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文产生的权利和责任。

声明人（签名）：

年　　月　　日

厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

1、保密（　），在　　年解密后适用本授权书。

2、不保密（　）

（请在以上相应括号内打“√”）

作者签名：　　日期：　　年　　月　　日

导师签名：　　日期：　　年　　月　　日

摘　要

随着信息网络的高速发展，电子邮件作为一种快捷便利的通信手段，已经深入普及到人们的日常工作与生活中，其间蕴含着丰富的个人信息，是进行计算机调查取证的重要途径，能为案件侦破提供有力的线索。电子邮件客户端为邮箱用户提供了友好的管理界面，受到网民的普遍青睐，使用率高，其保存的邮件数据文件也是调查取证的重要对象，挖掘分析其中的有用线索是计算机取证的重要手段。

本文研究的电子邮件调查分析系统主要是针对 Microsoft Outlook 和 Outlook Express 这两种常用的电子邮件客户端软件，通过解析客户端下保存的邮件数据文件，即从 Pst、Dbx 两种格式的邮件存储复合文档中提取相关邮件的收发帐户邮件地址、主题、发送时间、邮件内容和附件等信息，然后根据收发件人邮件地址进行统计归类，运用网络分析、可视化等技术绘制邮件时间图和网络图，进而为分析收、发件人之间隐藏的内部关系网提供有力依据。

本文首先分析了电子邮件调查分析的重要意义，介绍了国内外电子邮件调查分析的现状；接着，详细阐述了本系统开发的关键技术：基于 COM 技术的 Dbx 邮件文件解析、基于 OLE 自动化技术的 Pst 邮件文件解析，以及在绘制邮件时间关系图和网络关系图过程中用到的数据库访问技术、图形绘制基本理论和基于遗传算法的图自动布局算法等；然后，介绍了系统的总体目标、运行环境、功能需求，以及详细功能设计和具体实现的基本情况等；最后，对本系统的研究开发情况进行了总结，并就开发中的不足之处提出了下步的努力方向。

本文研究的电子邮件调查分析系统经过反复测试和完善，目前性能良好、运行稳定，达到了系统开发的目标，它的深入研究和使用将对我国计算机调查取证方式的发展起到积极的作用。

关键词： 电子邮件；文件解析；COM；OLE 自动化；ADO；遗传算法

ABSTRACT

Along with the rapid development of information network, E-mail, as an important and quick means of communication, has been deeply into people's daily work and life. The rich personal information contained inside is significant for investigation and evidence collection, providing important clues to handle cases. E-mail clients offer a friendly user interface, generally favored by netizens. The mail data files preserved become an important target of computer forensics, from which we can mine out and analyze the useable clues.

The Email Investigation & Analysis System mentioned in this paper is mainly for Microsoft Outlook and Outlook Express, which are two commonly used e-mail clients. Through parsing of data files kept by mail-client, that is, extract the address, subject, date, mail-content and attachment of related mails from the Pst and Dbx format compound documents, then make statistic and classification according to sender-address or receiver-address. Using social networking analysis, visualization, and other technique to draw mail time-line layout and network layout, we can discover the hidden internal relation network between senders and receivers, and get powerful evidence for cases investigation.

This paper described the significance of e-mail survey firstly. Then, introduced in detail the key technologies of the system development: Dbx mail file parsing based on COM technology; Pst mail file parsing based on OLE Automation technology; database access technology used in the system; basic theory of graph drawing and Automating layout algorithm based on GA (Genetic Algorithm). The overall goal of the system, running environment, functional requirements, as well as designing details and other features are also described then. Finally, summarize the system development and on the lack propose the next step to make efforts.

The system discussed in the paper has gone through repeated testing and perfecting. At the moment it is in good performance, operation stability and achieves the objectives of the system development. So the system will have some active action for the development of computer forensics in our country, and has some value to spread.

Keywords: E-mail; File Parsing; COM; OLE Automation; ADO; Genetic Algorithm

目　录

Contents