

上海电力学院

本科毕业设计（论文）



题 目： 企业邮件的安全分发与检测

院 系： 计算机科学与技术学院

专业年级： 大四

学生姓名： 陈志成 学号： 20091699

指导教师： 刘大明

2013 年 6 月 30 日

企业邮件的安全分发与检测

摘要

企业邮箱是当今国内外互联网应用中使用最为广泛的应用之一。作为现代电子商务和社会联系的一个重要的沟通工具,企业邮箱承载着支撑整个商业社会进行运作的信息与数据。随着中小型企业如雨后春笋般的生长和企业内网的不断发展,企业邮箱及其携带的信息具备了一个快速上升的趋势。无论是对企业邮箱及其附带信息进行有效管理,还是对企业邮箱是否能具备一定的安全性来防止信息泄露,都提出了更高层次的需求。

课题通过对电子邮件到企业邮箱发展的深入调查和多方对比,探寻在企业邮箱的发展过程中,伴随着的安全的隐患及设计操作的不合理之处。之后再对国内现有的多个邮件管理系统进行安全、功能和交互分析,并进行邮件管理系统的优势和劣势的对比。随后,结合分析的结果提出对邮件管理系统各个设计方式和初步功能的构想。从核心算法思想、系统结构后台和前端交换展示优化等各个模块的设计进行了深入的解析,探讨了如何将安全及异常检测的算法融合到邮件系统中几个具备代表性的设计中的策略,说明基于数据异常分析的思想设计能为处理邮件带来巨大的优势。

课题研究的企业邮箱系统强化安全与异常检测的策略在于引入邮件重要度的加权算法,并结合使用基于图论的邮件社交网路分析方法,继而推导邮件管理系统的交互设计和功能,给出了一种全新基于数据分析和社交网络思想的新形态的邮件系统,并在此之上进行了多种安全加密算法的搭载和测试,以及系统后台设计稳定性、扩展性的优化设计,保证了企业邮箱系统具备一定的实用性。系统各个模块都已经通过了测试,测试结果均稳定正常,产生的数据结果能满足大量邮件的异常分析处理,完全符合设计的要求。

关键词 企业邮箱 邮件权重分类 邮件社交分析 加密算法 信息安全

THE SECURITY DISTRIBUTION AND MONITORING OF ENTERPRISE EMAIL SYSTEM

Abstract

Enterprise mailbox is the one of the most widely used application of the Internet application in domestic and overseas. As an important communication tool of e-business and social connection, Enterprise mailbox has the most important data and information of business society. With the fast growth of China's small and medium-sized enterprises and development of enterprise intranet, enterprise mailbox and information have turn into a rapidly rising trend. So, it put forward a higher level of demand not only to the effective management of enterprise mailbox attachment information, but also to have a certain security to prevent information leakage.

Firstly, this paper investigates history of development from Email System to Enterprise E-mail System, probes into the potential security risk and irrational points existing in design and operation. Furthermore, this paper analyses the security, functionality and interaction of E-Mail Management Systems of existence in the native , and compares the advantages and disadvantages of these Email Management System. On the basis of theoretical analysis, this paper brings the conception of design and primary function in E-Mail Management System. Each functional module such as core algorithm, structures has been thorough analyzed and researched.

The primary design of Enterprise E-mail System in this paper is introduced Weighted Algorithm, which in combination with email social network analyze method based on Graph Theory. Then the paper derives the interaction design and function of E-Mail Management System. Finally, give a brand new Mail System based on data analyze and social network. In addition, the Mail

System also configurate a set of encryption algorithms to provide security and a optimize for back-end design to provide stability. Each module designed in this paper have already passed the test, the test results were stable and normal, the result is reasonable, completely meet the design requirements.

Keywords enterprise mailbox email classify email social analyze encryption information security

目 录

1	引言.....	1
1.1	大数据下的邮箱管理背景.....	1
1.2	邮箱伴随企业的急速发展与扩张.....	2
1.3	企业邮箱安全检测的巨大隐患和损耗.....	4
1.4	多种邮箱系统的安全检测对比分析.....	5
1.5	研究内容与目的.....	8
2	系统平台搭建关键技术.....	10
2.1	J2EE 的 SSI 框架设计.....	10
2.2	JavaMail 邮箱技术原理.....	15
2.2.1	Message 邮件文本.....	16
2.2.2	Transport 邮件推送.....	18
2.2.3	Store 存储机制.....	19
2.3	前端设计框架+数据可视化展示.....	19
2.3.1	Bootstrap 前端设计框架.....	20
2.3.2	jQuery Ajax 实现 SPA 应用.....	23
2.3.3	Nvd3 数据可视化建模.....	24
3	企业邮箱异常检测算法研究.....	26
3.1	邮件重要度检测算法研究.....	26
3.2	安全加密算法研究.....	27
3.2.1	对称加密技术.....	28
3.2.2	非对称加密技术.....	28
3.3	基于图论和社会网络方法的邮件团体分析算法研究.....	29
4	安全与异常检测功能验证.....	35
4.1	发件安全检测模块设计.....	35
4.1.1	企业邮件重要度安全标记.....	35
4.1.2	安全锁定邮件设计.....	36
4.1.3	安全加密选择.....	36
4.2	收件安全检测模块设计.....	38
4.2.1	收件箱安全检测设计.....	39
4.2.2	邮件阅读安全锁.....	39
4.3	重要度排序异常检测模块设计.....	40
4.3.1	智能邮件检测.....	41
4.3.2	其他邮件检测.....	42
4.4	数据可视化异常检测模块设计.....	43
4.4.1	邮件联系人热度异常检测.....	43
4.4.2	邮件时间段分布异常检测.....	44
4.4.3	邮件联系人异常检测.....	45
4.4.4	邮件社会关系网络异常检测.....	46
5	结论.....	48
	谢辞.....	50

参考文献.....	51
-----------	----

1 引言

1.1 大数据下的邮箱管理背景

当今，高速的互联网极大的缩短了居住在不同国家，不同种族人们之间的距离，即使两者彼此位于天涯海角也能像室友一样畅所欲言。当然，互联网也是最容易使用和开销最小的一种通信手段。互联网的通讯工具包括处理基于音频、视频和文字的工具，例如邮箱，微博，IM等，这些工具大量的被网络用户或商务人士所使用，几秒钟内就可以将各种各样的信息传递给其他人。其中，邮件作为这样一个使用人群最为广泛且最受商务人士关注的通信工具，同样是历史最为悠久的电子通信方式，也迎来自己高速发展的一个阶段。

由于全球网络在企业内网和互联网两个方面持续性的进行增长和强化，电子邮件用户，特别是企业邮箱用户，也十分期待使用新的策略来管理收件箱。原始电子邮箱单一的管理方式，操作复杂的用户界面和十面埋伏的安全隐患已经饱受用户所诟病。在当今这个科技化的世界里，人们需要的是安全的，可靠的收件箱，如果还能拥有用户友好的界面、简单直观的数据分析和良好的邮件管理方式，就能更加受到用户的拥护。此外，即使随着通讯工具的不断增多，电子邮件用户的群体不但没有被蚕食，反而在以一个巨大的速度在持续不断的增加。根据由某咨询公司的市场调研，至2009年初世界各地已经大约拥有了2亿电子邮件用户。到现在，电子邮件不仅仅用于通信方面，高科技的企业邮箱更是象征着一个公司的实例，不得不说，电子邮件作为一个最为成功的电子通信工具典范，将会在市场上长期保存一个霸主的地位。

据统计，平均而言，一个普通邮箱用户每天将会收到40~50封电子邮件。而且更为突出的是，企业邮箱用户每天几乎将收到数百个电子邮件。这样看来，用户将会花费显著的时间量进行电子邮件的处理。这上百封邮件中，商业电子邮件是至关重要的，因为企业邮箱的用户不乏一些公司的决策人物，这些邮件往往涉及到产品需求或客户利益。但在过去的许多年中，管理收件箱仍然是一个非常繁琐的工作。垃圾邮件过滤的确是一个非常有效的方法，并已经成功能分辨出80%的垃圾邮件,为用户阅读和筛选邮件减轻了巨大的负担,但优先级的的问题还是没有

得到质的解决,用户还是可能花费巨额的时间用于滚动和筛选来找到需要的邮件。现在某些电子邮件系统让用户通过设置一个字段来区分重要消息邮件的优先级。但是,仍然有一个问题,很多用户会遗忘掉这个字段,对他们的电子邮件进行优先级划分将会失效。本文就是希望提出一种新的解决方案,在保证邮件平台系统安全、易用的基础上,能够提供一种持续的动态重要度分析算法进行邮件的排序,并结合数据分析视图和社交关系网状结构图,来达到一种对企业邮箱的持续安全分发和监测。

1.2 邮箱伴随企业的急速发展与扩张

电子邮箱产品也可以算是最基础最古老的互联网企业服务之一,随着全球信息化的发展,邮箱作为一种快捷有效的通讯方式其应用将更加广泛,而且伴随着他在商务领域得天独厚的优势,转型成为专业化的企业邮箱也成为大势所趋。电子邮箱在进入中国以来30多年从免费到专业企业邮箱的发展过程中,可以说主要经历了四个飞速成长的阶段:

(1)个人免费邮箱为主的诞生期

自1987年我国第一封电子邮件“跨越长城、通向世界”成功发送后,从此中国人便开启了由使用第一封电子邮件而开始使用互联网的历史。由于当时中国互联网刚刚起步,网民人数少,所以当时的电子邮箱大都指个人免费邮箱。直到1998年,电子邮箱的更高级产品“企业邮箱”概念开始传入中国,这是电子邮箱在中国的诞生阶段。

(2)电子邮箱高速发展的成长期

而从1999年-2002年期间,中国互联网进入普及和应用的高速成长期,政策对互联网发展也重视起来,据悉截止到2002年12月31日,中国共有上网计算机约2083万台,上网用户数约5910万,CN下注册的域名17.9万个,WWW站点约37.1万个,国际出口带宽9380M,中国网民绝对数量的上升使电子邮箱作为简单的通讯工具而成为互联网的基础服务。这一时期,随着中国互联网的发展,“企业邮箱”概念才真正被中国网民所接受,这个阶段属于中国电子邮箱的成长阶段。

(3)企业邮箱崭露头角的发展期

随着互联网及中国经济的蓬勃发展，到2003年以后，互联网已经进入千家万户，电子邮箱也进入了它的快速发展阶段。尤其在2005年、2006年电子邮箱自身的功能以及用户数量、使用频率等都有较大幅度的完善和提高。随着2007年邮箱技术及功能的逐步全面，电子邮箱已逐渐普及，企业邮箱也逐渐被企业所认可，一些企业通过自架邮局系统或者购买企业邮箱来进行业务往来。

(4)企业邮箱逐渐盛行的成熟期

据艾瑞咨询预测，在2007年—2008年期间，随着互联网的成熟稳定，再加上网民行为的更加理性，电子邮箱市场的发展将逐渐平稳，但是随着各类新技术的不断推出，电子邮箱将与其他技术不断融合，企业邮箱作为互联网普遍看好的电子邮箱产品进入激烈的角逐期，它以其不断创新的功能和日臻完善的服务相结合的企业邮箱，区别于免费电子邮箱而备受广大企业的关注。同时，企业不断的增长也是助推企业邮箱的一个重要因素，2010年，具有电子商务应用能力的中小企业数量超过4500万家，拥有企业邮箱的企业超过300万家。预计未来几年，金融危机对中小企业的影响进一步减退，随着经济的逐步复苏，中小企业数量开始稳步增长，预计增长率保持在8%左右，2014年将有望超过6000万家。

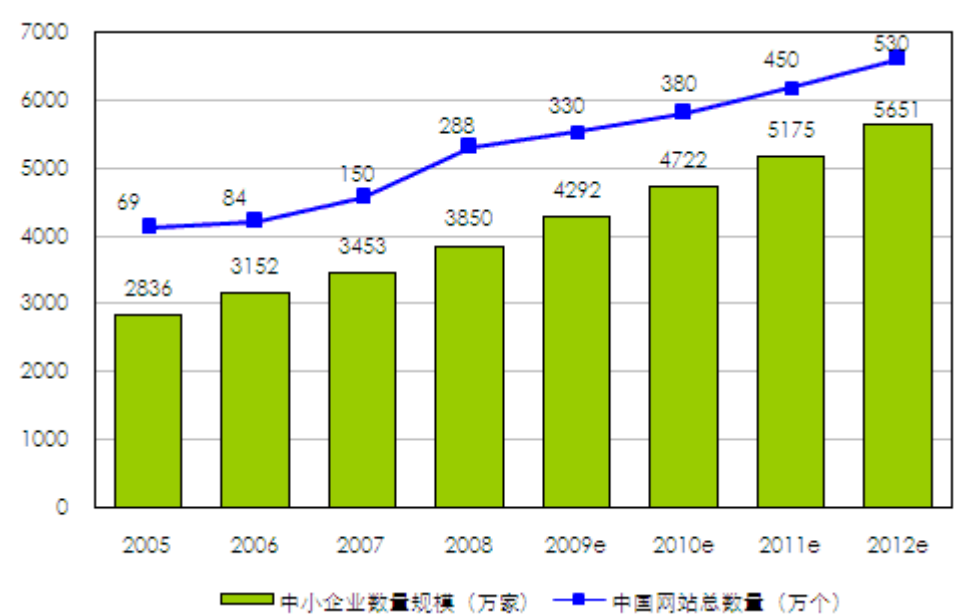


图 1-1 2005-2012 年中国中小企业和网站总体数量规模

企业邮箱作为电子邮箱市场一支新生的力量，毫无疑问将会在未来的数十年迅速发展壮大，并成为未来社会中必不可少的一个承重轮和助推器。

1.3 企业邮箱安全检测的巨大隐患和损耗

电子邮件在企业内部和外部沟通中扮演着十分重要的角色，用户对电子邮件作为生产力工具是越来越依赖，它的安全问题也越来越让公司管理层不容忽视。

首先，邮件具有很大的泄密隐患。据调查，大约有74%的邮件安全事件是由于内部人员将机密信息在没有适当加密的情况下发送到了公司的外部，公司职员也许不会通过邮件散播恶意软件，但是他们可能因疏忽而泄露公司机密，而且许多人对于邮件安全持有错误的观念。根据一项最新研究，大约有近94%的公司承认，他们在防止公司机密通过邮件外泄方面显得力不从心，只有6%的受访者表示有信心防止内部员工通过邮件泄露公司机密。研究还表明，有近32%的公司不清楚公司机密是否已被外泄，因而也就不能采取有效措施规避这一风险；62%的公司表示尽管不能阻止泄密事件的发生，但是可以查出泄密的邮件来源。另外，如果花费人力物力去追查的话，额外的成本也将是一个大问题，大部分企业无法承受这种高额投资。


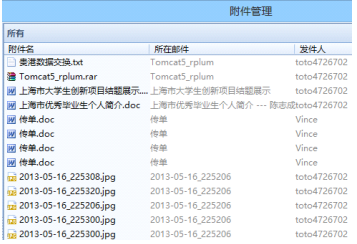
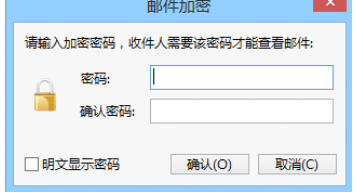
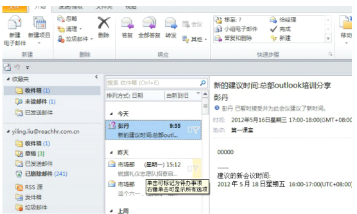
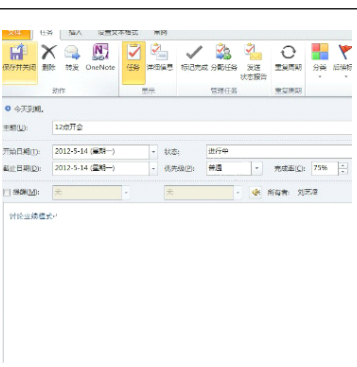
其次，垃圾邮件严重浪费生产力。其实垃圾邮件的定义很宽泛，它的意思大体是指未受请求而主动发送的邮件，它和你家邮筒无端端收到的垃圾邮件异曲同工。但是这种垃圾邮件绝对不仅仅是一堆不受欢迎而又花样百出的混合物，它会阻碍电子邮件账户(以及网路和服务器)的通信。同时，被塞满垃圾邮件的电子邮件不仅耗费了公司资金，而且让员工很伤神。研究发现美国公司因为垃圾邮件每年丧失的生产力的价值达到了710亿美元，研究公司是通过调查850家企业电子邮件用户进行调查，这些企业在调查中表示如果没有合适的保护措施，垃圾邮件占公司所有邮件数量的20%。同时，他们还要花费工作总时间的5%来处理这些垃圾邮件，从而妨碍员工生产力和企业成本竞争优势。虽然垃圾邮件看起来不需要花很多时间去处理，例如普通用户每天要花大约10分钟时间去阅读或者删除这些垃圾邮件，但长年累月来看这些浪费的生产力将是惊人的。


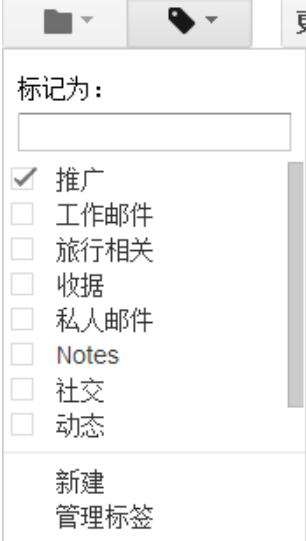
最后，邮件病毒破坏力惊人。可以说在企业网络中，最为惧怕的就是邮件病

毒了，由于他并没有明确的指向和来源，所以往往没有具体的防范措施。邮件病毒其实和普通的电脑病毒一样，只不过由于它们的传播途径主要是通过电子邮件，所以才被称为邮件病毒。邮件病毒除了具备普通病毒可传播性、可执行性、破坏性、可触发性特征之外，还有一定的特殊性。例如，一是感染速度、快扩散面广：由于企业邮箱的邮件不仅仅在单个企业内部传播，这直接致使邮件病毒的扩散不仅快，而且扩散范围很大，不但能迅速传染局域网内所有计算机，还能通过将病毒在一瞬间传播到千里之外。二是清除病毒困难：单机上的计算机病毒有时可通过删除带毒文件，格式化硬盘等措施将病毒彻底清除，这对于企业极其重要的内部数据可以说是毁灭性的打击。而企业中的计算机一旦感染了病毒，清除病毒变得非常困难，刚刚完成清除工作的计算机就有可能被网络中另一台带毒工作站所感染，使得邮件病毒变得非常困难了。三是破坏性大：网络中的计算机感染了邮件病毒之后，将直接影响网络的工作，轻则降低速度，影响工作效率，重则使网络及计算机崩溃，资料丢失。

1.4 多种邮箱系统的安全检测对比分析

目前国使用人数比较多的邮件管理系统包括两种，一种是使用本地邮件客户端的邮件系统，较为出名的包括Foxmail，Outlook等。另一种是在浏览器打开的WEB电子邮箱，现在使用较多又具有良好的邮件管理功能的包括Gmail，Hotmail，163邮箱，Yahoo邮箱和QQ邮箱等。在这里主要对这些主流的邮箱系统的基本功能、管理功能和数据处理功能进行对比分析，为后文邮件系统的设计提供借鉴的作用。

邮箱	功能	图例
Foxmail	邮件联系人查询： 可以查询某联系人发送的所有邮件	
	附件统一管理： 可以直接访问所有邮件中不同格式的数据	
	邮件附带密码： 可以在发送邮件的同时，为邮件设置一份打开密码	
Outlook	日程管理功能： 可以向联系人直接发送一段日程表，提示对方会议时间，地点，主题	
	新建任务功能： 能够设置定时的任务发送给对方，让对方及时反馈完成的程度，类似于 Todo	

邮箱	功能	图例
Gmail	<p>邮件环聊功能： 可以围绕一个邮件主题自由进行发言，并进行视频通话</p>	<p>环聊让对话更生动有趣</p> <ul style="list-style-type: none"> • 发送照片和表情符。 • 最多与 10 位朋友免费视频通话。 • 在电脑、Android 设备和 Apple 设备上均可连线聊天。  <p>下一步</p>
	<p>邮件标签功能： 可以对邮件的主题进行分类，方便找到同类型的邮件</p>	

邮箱	功能	图例									
163	操作消息提醒: 可以方便查看到每一个操作产生的效果										
	异常手机绑定: 任何进行邮箱配置操作或功能异常都会用短信发送提示给用户	 <table><tr><th colspan="2">*****497 开通的手机服务</th></tr><tr><td>手机扫码解绑</td><td>*****497@163.com 已激活</td></tr><tr><td>邮箱异常提醒通知</td><td>已开通</td></tr><tr><td>账号异常通知</td><td>已开通</td></tr></table>	*****497 开通的手机服务		手机扫码解绑	*****497@163.com 已激活	邮箱异常提醒通知	已开通	账号异常通知	已开通	
	*****497 开通的手机服务										
手机扫码解绑	*****497@163.com 已激活										
邮箱异常提醒通知	已开通										
账号异常通知	已开通										
邮箱健康指数: 可以对邮箱整体的安全性和稳定性进行评分, 并给予用户建议	 <p>体检得分: 84</p> <p>您的邮箱存在一些风险, 请根据提示进行相关修复, 共扫描14项, 其中 3 项有问题</p> <p>重新体检</p> <p>全相关</p> <table><tr><td>密码强度</td><td>密码强度为中, 建议定期修改。</td></tr><tr><td>密码保护</td><td>已设置1项密码保护, 忘记密码时</td></tr><tr><td>绑定手机号码</td><td>已绑定手机号码。</td></tr><tr><td>登录情况</td><td>登录未发现异常。</td></tr><tr><td>是否存在被盗隐患</td><td>您的帐号未发现被盗隐患。</td></tr></table>	密码强度	密码强度为中, 建议定期修改。	密码保护	已设置1项密码保护, 忘记密码时	绑定手机号码	已绑定手机号码。	登录情况	登录未发现异常。	是否存在被盗隐患	您的帐号未发现被盗隐患。
密码强度	密码强度为中, 建议定期修改。										
密码保护	已设置1项密码保护, 忘记密码时										
绑定手机号码	已绑定手机号码。										
登录情况	登录未发现异常。										
是否存在被盗隐患	您的帐号未发现被盗隐患。										

1.5 研究内容与目的

现行企业邮箱基本都具备处理垃圾邮件的引擎及杀毒引擎, 这些能够有效的查杀一些垃圾和病毒邮件, 但还是有很多的缺陷。现在邮箱系统普遍缺少对企业内部邮件的自动化分析, 不能够整体体现出企业内部的邮件信息交流状况, 更不具备团队分析后的分发功能和交流密度的全面分析, 无法为企业内部的信息交流趋势提供分析和参考, 而且对于一些邮件欺诈和保密也无法做出及时正确的分析和保障。

针对这些情况, 本文目的在于设计一个企业团体邮件安全分发的异常检测系统, 具有良好的自动化分析反馈, 能对大批量的邮件进行分类检测, 能对邮箱用户群体的联络社交网进行检测, 并对数据安全也做出了一定强化。

研究内容主要包括以下几个方面:

- (1)实现企业大批量邮件自动扫描, 根据重要度计算智能的隔离垃圾邮件, 病毒文件等等危及企业安全的不良邮件。
- (2)实现企业邮件对于邮件团体的内省, 监测企业内部邮件的分布和动向, 并解析邮件社交网络的产生。
- (3)通过邮件传递分析企业内部的联系密度, 分析企业成员的凝聚力。
- (4)对重要文件以及邮件网关, 采用成型加密算法, 进行快速高效加密。
- (5)以前沿的UED前端技术进行设计, 保障系统对用户友好, 实现SPA邮箱应用。

其中拟解决的关键问题包括如何实现以数据结构为基础的智能分析算法, 算法能不能正确分析并反馈真实有效的数据, 如何实现大量数据处理下的算法持续有效性及记忆功能, 如何实现对文件资源的有效加密方法等问题。

设计的目的也主要包含两个方面, 首先, 使用快速发展的前沿技术和思想帮助企业对邮件这个重要的信息源产生的异常进行检测和反馈, 并提供邮件的安全加密加锁措施, 以保证企业管理者对企业内部的有效分析和管理, 避免不可预制的危害企业发展的事件产生。其次, 为企业安全管理这个原本重视度较低却又异常重要的课题提供了良好的实践资料和模板, 并倡导了对信息安全的重视。

2 系统平台搭建关键技术

2.1 J2EE 的 SSI 框架设计

J2EE是一种利用Java平台来简化企业解决方案的开发、部署和管理相关的复杂问题的体系结构。而Java本身就是一种非常流行的语言，因为Java作为一门完全开源的编程语言，完全颠覆了原来语言注重精英和封闭式的微软模式，尊重了开源社区的力量。Java本身包含三个板块，分别是J2SE, J2EE, J2ME。事实证明，除了J2ME后来被Android取代，无论在哪个方面Java都取得了空前的突破和成功。J2EE技术的基础就是核心Java平台或Java 2平台的标准版，J2EE不仅巩固了标准版中的许多优点，例如"编写一次、随处运行"的特性、方便存取数据库的JDBC API、CORBA技术以及能够在Internet应用中保护数据的安全模式等等，同时还提供了对 EJB (Enterprise JavaBeans)、Java Servlets API、JSP (Java Server Pages) 以及XML技术的全面支持。其最终目的就是成为一个能够使企业开发者大幅缩短投放市场时间的体系结构。

本文中的企业邮箱系统后台使用J2EE进行开发，并与最为成熟的Struts、Spring和iBatis框架相结合，可以最大发挥出J2EE可伸缩性、灵活性、易维护性的特点。这套框架的系统从职责上分为四层：视图层、控制器层、业务逻辑层和数据持久层，以帮助开发人员能迅速搭建搭建松耦合、复用性高、扩展性强且维护方便的Web应用。其中使用Struts作为系统的整体基础架构，负责MVC的分离，在Struts框架的模型部分，控制业务跳转，利用iBatis框架对持久层提供支持，Spring做管理，管理Struts和iBatis。

采用上述开发模型，不仅实现了视图、控制器与模型的彻底分离，而且还实现了业务逻辑层与持久层的分离。这样无论前端如何变化，模型层只需很少的改动，并且数据库的变化也不会对前端有所影响，大大提高了系统的可复用性。而且由于不同层之间耦合度小，有利于团队成员并行工作，大大提高了开发效率。而且这里其实上就是应用了一个软件设计时非常经典的一个设计模式——MVC设计模式，他能有效的打造一个体系非常清晰，架构非常有弹性的应用，下面文章会具体介绍MVC的定义。

2.1.1 MVC 设计模式

MVC英文即Model-View-Controller，即把一个应用的输入、处理、输出流程按照Model、View、Controller的方式进行分离，这样一个应用被分成三个层——模型层、视图层、控制层。这样做虽然要花费额外的工作，但产品的结构清晰，设计弹性力度很高，产品的应用通过模型可以得到更好地体现。

首先，最重要的是应该有多个视图对应一个模型的能力。在目前用户需求的快速变化下，可能有多种方式访问应用的要求。例如，订单模型可能有本系统的订单，也有网上订单，或者其他系统的订单，但对于订单的处理都是一样，也就是说订单的处理是一致的。按MVC设计模式，一个订单模型以及多个视图即可解决问题。这样减少了代码的复制，即减少了代码的维护量，一旦模型发生改变，也易于维护。其次，由于模型返回的数据不带任何显示格式，因而这些模型也可直接应用于接口的使用。再次，由于一个应用被分离为三层，因此有时改变其中的一层就能满足应用的改变。一个应用的业务流程或者业务规则的改变只需改动MVC的模型层。最后，它还有利于软件工程化管理。由于不同的层各司其职，每一层不同的应用具有某些相同的特征，有利于通过工程化、工具化产生管理程序代码。下面就详细的对MVC每个层中的定义进行简要的介绍：

(1)视图层(View)

代表用户交互界面，对于Web应用来说，可以概括为HTML界面，J2EE系统中有可能为JSP。随着应用的复杂性和规模性，界面的处理也变得具有挑战性。一个应用可能有很多不同的视图，MVC设计模式对于视图的处理仅限于视图上数据的采集和处理，以及用户的请求，而不包括在视图上的业务流程的处理。

(2)控制层(Controller):

可以理解为从用户接收请求，将模型与视图匹配在一起，共同完成用户的请求。划分控制层的作用也很明显，它清楚地告诉你，它就是一个分发器，选择什么样的模型，选择什么样的视图，可以完成什么样的用户请求。控制层并不做任何的数据处理，而是全权交给模型层进行处理。

(3)模型层(Model):

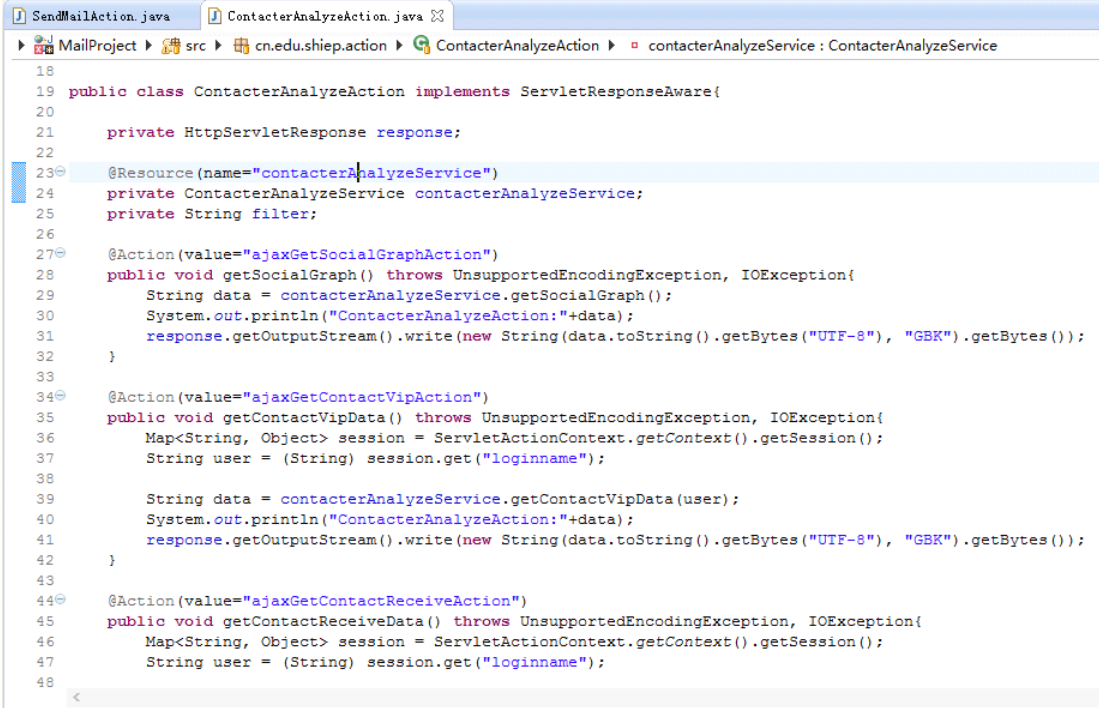
也可以称为业务逻辑层，就是业务流程/状态的处理以及业务规则的制定。业务流程的处理过程对其它层来说是黑箱操作，模型接受控制层请求的数据，并返

回最终的处理结果。业务模型的设计可以说是MVC最主要的核心。

2.1.2 控制层设计

在企业邮箱系统的控制层的设计之中，我们使用Struts2来承担整个系统的页面业务请求处理调度，例如发件、收件的行为，均是通过两个叫SendMailAction和ReceiveMailAction的Action类经行处理，图2-1就是企业邮箱系统中大量使用Action的示例。

实际上在Struts2中，承担MVC中控制层角色的就是一个名称为 Action的类。Action是一个通用的控制组件。这个控制组件提供了处理所有发送到 Struts2的HTTP请求的入口点。另外控制组件也会将从页面带过来的参数自动反射到Action的属性字段中,并进行校验。Action类实现网站所有的跳转逻辑，它可以访问视图层或调用业务逻辑层。最后Action类从业务逻辑层得到结果后就会把控制权传给接下来要显示JSP文件，生成视图。所有这些控制逻辑利用struts.xml文件来配置。



```
18
19 public class ContacterAnalyzeAction implements ServletResponseAware{
20
21     private HttpServletResponse response;
22
23     @Resource(name="contacterAnalyzeService")
24     private ContacterAnalyzeService contacterAnalyzeService;
25     private String filter;
26
27     @Action(value="ajaxGetSocialGraphAction")
28     public void getSocialGraph() throws UnsupportedEncodingException, IOException{
29         String data = contacterAnalyzeService.getSocialGraph();
30         System.out.println("ContacterAnalyzeAction:"+data);
31         response.getOutputStream().write(new String(data.toString().getBytes("UTF-8"), "GBK").getBytes());
32     }
33
34     @Action(value="ajaxGetContactVipAction")
35     public void getContactVipData() throws UnsupportedEncodingException, IOException{
36         Map<String, Object> session = ServletActionContext.getContext().getSession();
37         String user = (String) session.get("loginname");
38
39         String data = contacterAnalyzeService.getContactVipData(user);
40         System.out.println("ContacterAnalyzeAction:"+data);
41         response.getOutputStream().write(new String(data.toString().getBytes("UTF-8"), "GBK").getBytes());
42     }
43
44     @Action(value="ajaxGetContactReceiveAction")
45     public void getContactReceiveData() throws UnsupportedEncodingException, IOException{
46         Map<String, Object> session = ServletActionContext.getContext().getSession();
47         String user = (String) session.get("loginname");
48     }
```

图 2-1 企业邮箱使用 Struts2 来控制跳转

2.1.3 业务逻辑层设计

在企业邮箱系统的业务逻辑层设计之中，我们使用 Spring 对所有的业务类进行自动的装配。例如在 ServiceBean 中，我们会将 DAO 层的数据链接处理类进行自动注入，以方便 ServiceBean 能够快速方便地对数据层进行操作，例如，在 ContacterAnalyzeService 类中，我们就为其自动进行了 IContacterDao 和 IMailDao 的装配，这是因为这个 Service 的业务就是实现对联系人的数据进行分析以及取出所有邮件进行数据统计的功能。不止在 ServiceBean 中，我们在 Controller 中也会注入 ServiceBean，例如 ContacterAnalyzeService 类就会被注入到 ContacterAnalyzeAction 中，确保控制器能够调用到业务层。图 2-2 就是企业邮箱系统中通过 Spring 自动为 ServiceBean 进行装配的示例。

Spring 是一个开源框架，它是为了解决企业应用开发的复杂性而创建的。Spring 使用基本的 JavaBean 来完成以前只可能由 EJB 完成的事情。然而，Spring 的用途不仅限于服务器端的开发。从简单性、可测试性和松耦合的角度而言，任何 Java 应用都可以从 Spring 中受益。使用 Spring 的原因非常简单，他有两个可以说是当今软件设计中的两个哲学思想：控制反转 (IOC) 和面向切面 (AOP)。控制反转是指，一个对象依赖的其它对象会通过被动的方式传递进来，而不是这个对象自己创建或者查找依赖对象。这正好与 JNDI 相反的——不是对象从容器中查找依赖，而是容器在对象初始化时不等对象请求就主动将依赖传递给它。面向切面则允许通过分离应用的业务逻辑与系统级服务（例如审计（auditing）和事务（transaction）管理）进行内聚性的开发。应用对象只实现它们应该做的，就是完成业务逻辑。它们并不负责其它的系统级关注点，例如说日志或事务支持等等。

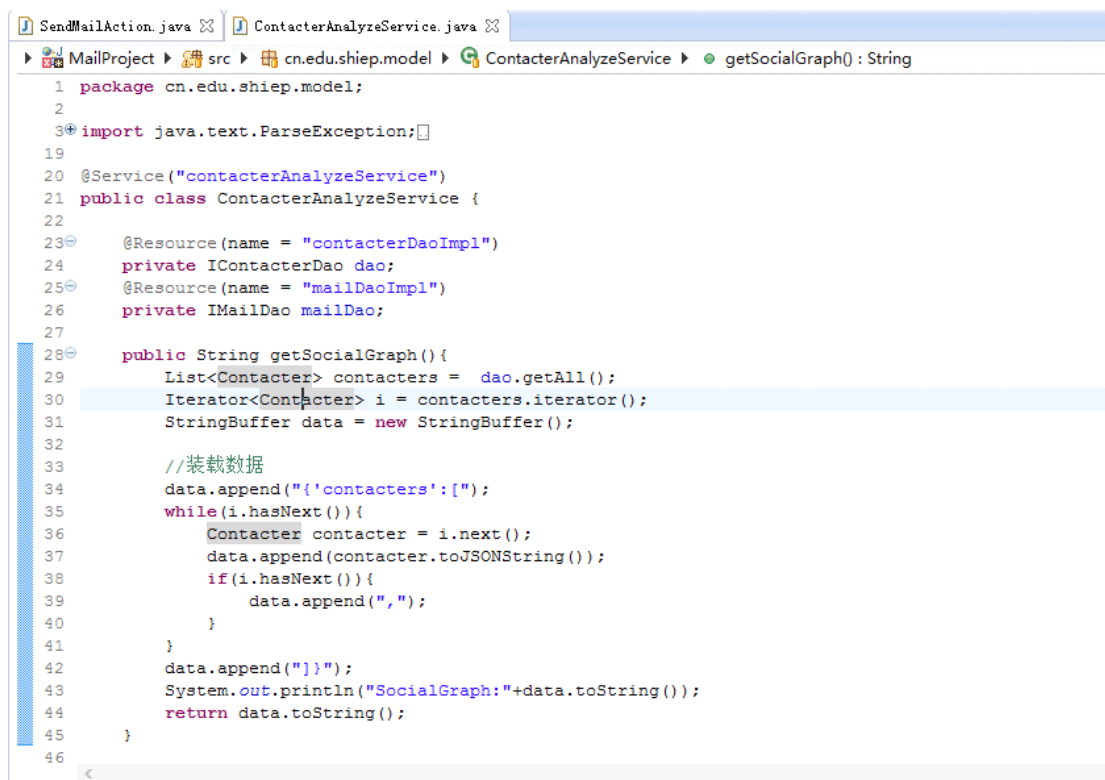


图 2-2 通过 Spring 来进行自动装配

2.1.4 数据持久层设计

在企业邮箱系统的数据持久层设计之中，我们使用iBatis来处理所有SQL到类的映射，也就是进行ORM处理。对于每个EntityBean系统中都会对应生成一个同名的Map映射XML，这个XML中就涵盖了所有关于这个EntityBean的数据操作。例如在Mail这个EntityBean，我们会在其对应的mailsql.xml文件中描述其所有相关的数据库操作SQL映射，包括Mail的增加、更新、统计等等操作的SQL语句。图2-3就是企业邮箱系统中iBatis配置数据库ORM映射文件的示例。

iBATIS被广泛认为是当今可用的最简单的持久化框架之一。简单性的理念根植于iBATIS，并在其所有开发目标中居于首位。这种简单性的取得是因为iBATIS直接构建于JDBC和SQL之上，实际上iBATIS在类和数据库表之间建立了一个额外的间接层，也就是SQL的映射层，我们直接通过类调用SQL的Map映射与数据库进行链接处理，并将把SQL语句的参数与结果（也即输入和输出）映射为类。iBATIS对于Java开发者来说非常简单，因为它除了不用编写那么多代码外与JDBC的工作机制非常相像。这样简单的框架带给我们的好处是非常明显的，那

就是生产效率的大幅度提高。据统计，iBATIS减少了持久层大约62%的代码量。所以能减少如此之多的代码量，究其原因还是开发者再不需要编写繁琐的JDBC代码了，转而关注到SQL语句的高质量实现。

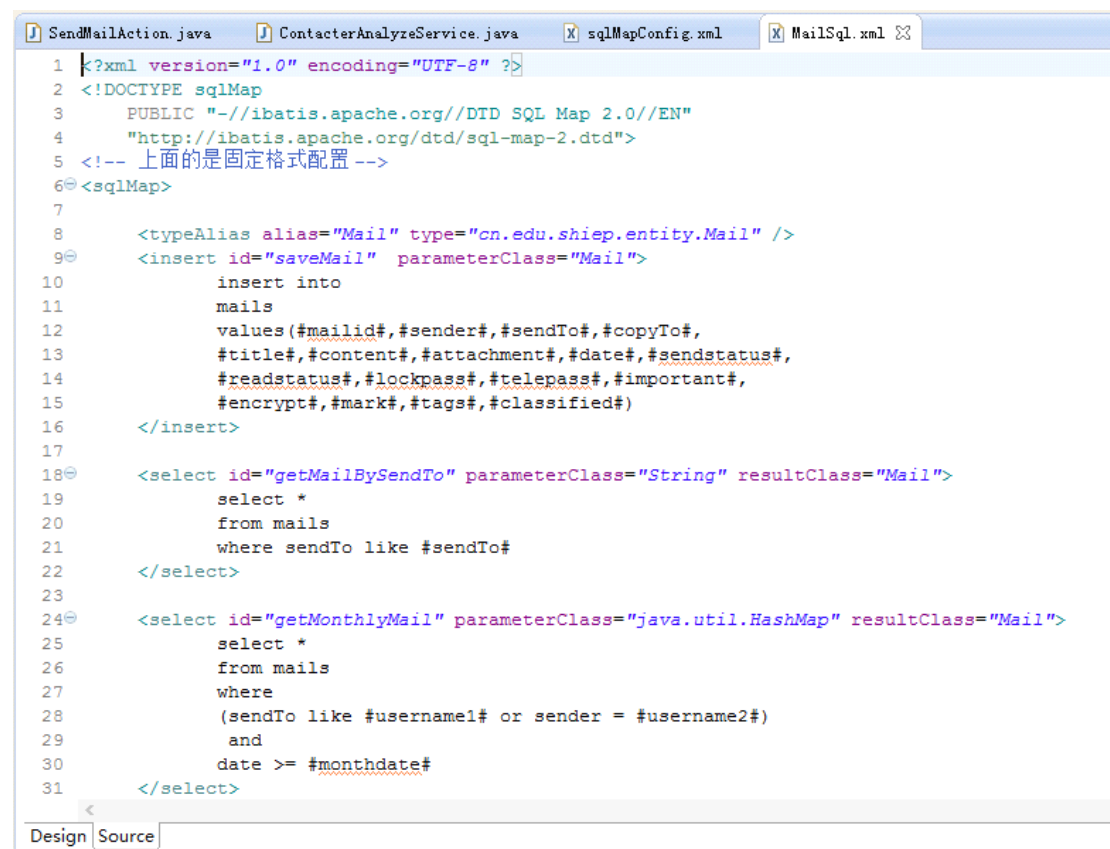


图 2-3 通过 iBatis 来进行处理 SQL 到类的映射

2.2 JavaMail 邮箱技术原理

本系统在基于J2EE的平台上，选取了JavaMail组件模块作为实现邮件发送的核心模块。JavaMail，顾名思义，就是提供给开发者处理电子邮件相关的编程接口。它是Sun发布的用来处理email的API，从Sun被Oracle收购以后就一直由Oracle在维护。它可以方便地执行一些常用的邮件传输。JavaMail API是Sun公司为方便Java开发人员在应用程序中实现邮件发送和接收功能而提供的一套标准开发包，它支持一些常用的邮件协议，如：SMTP、POP3、IMAP。开发人员使用JavaMail API编写邮件处理软件时，无须考虑邮件协议底层的实现细节，只要调用JavaMail开发包中相应的API类就可以了，以下将着重说明JavaMail各个核心模块的作用

和功能，图2-4为JavaMail内部核心模块的一个关系图。

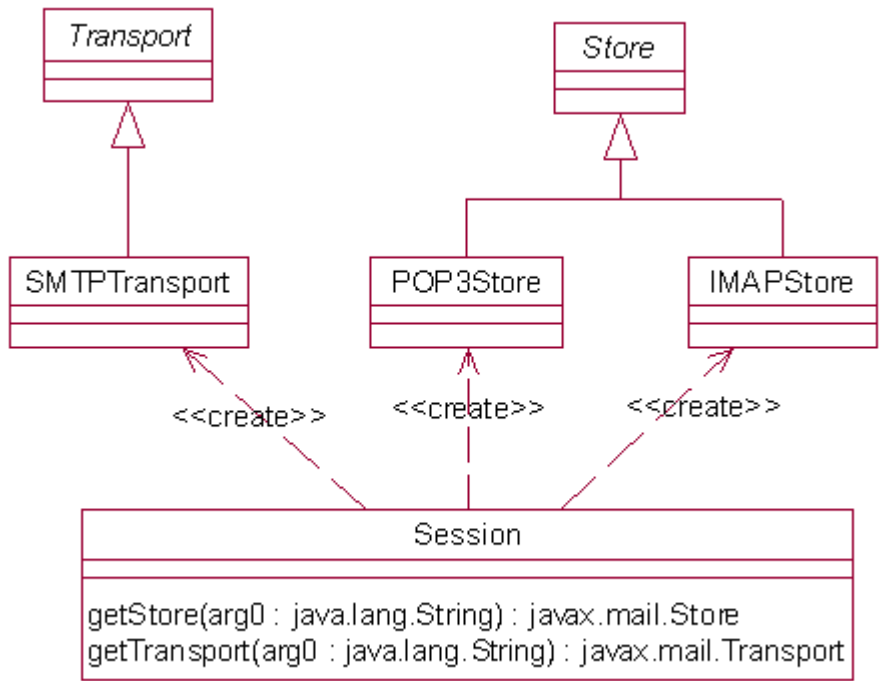


图 2-4 JavaMail 核心模块关系图

2.2.1 Message 邮件文本

邮件（Message）对象用于封装客户端要发送给收件人的全部数据信息。与地址对象类似，JavaMail API中定义的Message对象采用抽象类定义形式，通常在程序设计中利用 javax.mail.internet.Message 对象来封装数据信息。javax.mail.Message是创建和解析邮件的核心类，他的实例代表一封电子邮件。客户端程序发送邮件时，首先使用创建邮件JavaMailAPI创建出封装了邮件数据的Message对象，然后把这个对象传送给发送API发送。客户端接收邮件时，邮件接收API把收到的邮件数据封装在Message类的实例对象中，客户端程序再使用邮件分析API从这个对象中介析出接收到的邮件数据，图2-5为邮件Message组成结构的说明。

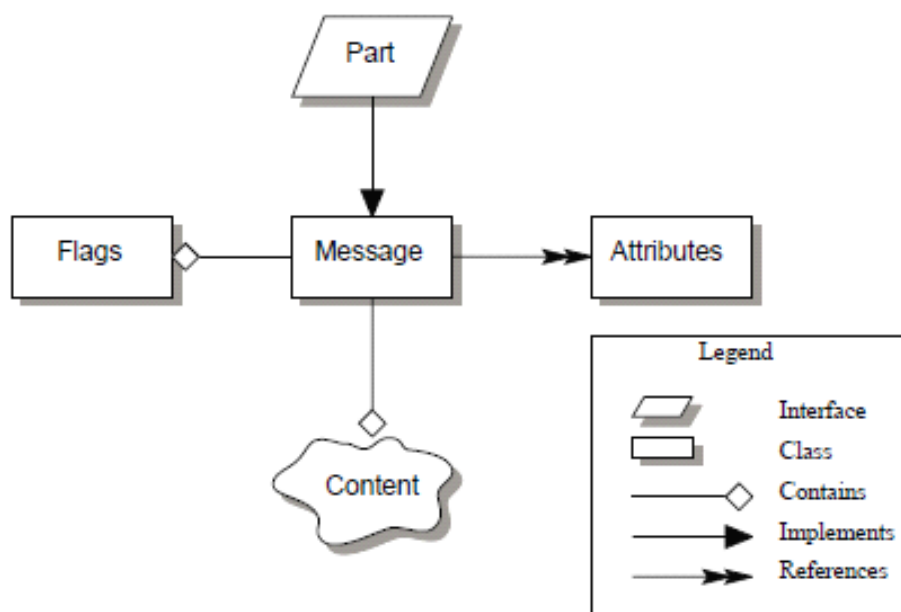


图 2-5 JavaMail 中 Message 的结构组成

邮件对象的构造方法中需要会话对象实例作为参数，如下面创建邮件对象实例的代码：

```
Message message = new MimeMessage(mailSession);
```

创建邮件对象实例后，可以调用该对象的setContent方法在邮件对象中封装特定类型的数据对象，例如文本类型或者超文本类型等，见下面的代码：

```
message.setText("This is a E-mail");
```

在上面setText方法的参数内容就是这封Email的邮件内容。另外，每个邮件对象均需要定义一个邮件标题（Subject），因此，可以调用邮件对象的setSubject方法为邮件对象创建标题：

```
message.setSubject(subject);
```

在邮件对象中还定义了setFrom方法和addRecipient方法，用于在邮件中标明该邮件对象的发送人和接收人地址。下面的代码将在2.2节中创建的地址对象添加到邮件对象实例中：

```
message.setFrom(new InternetAddress(from));
message.addRecipient(Message.RecipientType.TO, new
InternetAddress(to));
```

从上面的代码可以看出：邮件对象中定义的接收对象类型 TO、CC和BCC分别对应接收人、抄送人和密送人地址。需要注意的是：在创建地址对象实例以及在邮件对象中封装地址信息时，JavaMail API中没有定义检查邮件地址有效性的方法。因此，我们需要编写相应的邮件地址有效性检查代码，例如：

```
result    =    loginService.validateUser(username,    password,new
HashMap<String, Object>());
```

2.2.2 Transport 邮件推送

邮件发送（Transport）对象定义了客户端应用程序中采用特定的传输协议将邮件对象发送到邮件服务器的基本过程。javax.mail.Transport类是发送邮件的核心API，他的实例对象代表实现了某个邮件发送协议的邮件发送对象，例如SMTP，客户端程序创建好Message对象后，只需要使用邮件发送API得到Transport对象，然后把Message对象传递给Transport对象，并调用它的发送方法，就可以把邮件发送给指定的SMTP服务器。

创建邮件发送对象实例的方法很多，通常采用调用会话对象中定义的相应方法获取邮件发送对象实例的方式，这些方法的定义如下：

```
Transport transport = mailSession.getTransport("smtp");
```

从上述方法定义可以看出：在创建邮件发送对象时可以采用不同的策略，例如可以指定邮件传输目标的地址对象或者指定采用的传输协议类型等，包括SMTP、IMAP协议等等。通常一般在发送邮件中我们采用SMTP协议，SMTP是一种简单邮件传输协议(Simple Mail Transfer Protocol, SMTP)，它定义了发送邮件的机制。在JavaMail环境中，基于JavaMail的程序将和因特网服务供应商ISP(internet Service Provider 's)SMTP服务器通信。SMTP服务器会中转消息给接收方SMTP服务器以便最终让用户经由POP或者IMAP获得。

另外，在程序设计中可以不创建传输对象实例而是调用Transport对象中定义的静态方法send将邮件发送到邮件服务器，但是首先还是要与SMTP服务进行连接校验，以确定用户的身份，如下面的代码：

```
transport.connect(SMTPHost, user,password);
transport.sendMessage(message, message.getAllRecipients());
```


最后，不要忘记关闭Transport对象，这是一个非常容易忽略的地方，而且会引发严重的问题。如果你忘记关闭，邮件服务器会一直保持与Transport对象进行联系，等待Transport的下一步操作，浪费了十分珍贵的连接资源，关闭代码示例如下：

```
transport.close();
```

JavaMail API缺省在Transport对象中利用SMTP协议来发送邮件对象。读者在程序设计中可以采用IMAP等功能更强的协议类型。但需要注意的是针对的邮件服务器必须支持客户端应用程序采用的协议类型。

2.2.3 Store 存储机制

邮件接收对象（Store）用于客户端建立与邮件服务器的连接并在邮件服务器中对客户端指定的邮件账户和密码进行认证，最终将属于该邮件账户的电子邮件传输到客户端。javax.mail.Store类是接收邮件的核心，他的实例对象代表实现了某个邮件接收协议的邮件接收对象，例如POP3。客户端程序接收到邮件时，只需要有事邮件接收API得到Store对象的接收方法，就可以从制定的POP3服务器获得邮件数据，并把这些数据封装带表示邮件的Message的对象中。

在基于POP3协议的邮件服务器中，针对每个注册的邮件账户均创建唯一的名称为"INBOX"的文件夹对象（Folder），用于存储邮件服务器接收的电子邮件，而基于IMAP协议的邮件服务器可以创建多个文件夹对象。

在邮件接收程序中，通常利用会话对象实例的getStore方法并在方法参数中指定接收邮件的协议类型来创建邮件接收对象实例，例如：

```
store = session.getStore("pop3");
```

基于POP3协议的邮件接收程序可以调用邮件接收对象的getFolder方法并在方法参数中指定字符串"INBOX"来创建文件夹对象实例，如下面的代码：

```
folder = store.getDefaultFolder().getFolder("INBOX");
```

2.3 前端设计框架+数据可视化展示

近年来随着Web应用交互复杂度的提升，前端开发也迎来了一个高速发展的

时期，已经迎来了Web2.0的大时代。在互联网的演化进程中，Web 1.0时代，网站的主要内容都是静态的，用户使用网站的行为也以浏览为主，已经不能再满足现在我们对用户交互体验如此高的要求。2005年以后，互联网进入Web 2.0时代，各种类似桌面软件的Web应用大量涌现，网站的前端由此发生了翻天覆地的变化。网页不再只是承载单一的文字和图片，各种富媒体让网页的内容更加生动，网页上软件化的交互形式为用户提供了更好的使用体验，这些都是基于前端技术实现的。

随着Web 2.0概念的普及和W3C组织的推广，网站重构的影响力正以惊人的速度增长。XHTML+CSS布局、DHTML和Ajax像一阵旋风，铺天盖地席卷而来，包括新浪、搜狐、网易、腾讯、淘宝等在内的各种规模的IT企业都对自己的网站进行了重构。除了一些老牌框架纷纷推出改动较大的升级之外，还涌现出一批新生代的开源库和框架，包括Bootstap，FlatUI，AngularJS等等，推动着Web应用开发理念向越来越强调前端架构的方向发展。

本文设计的企业邮箱系统中主要使用了Bootstrap担任前端界面的主要设计框架，并大量使用到了jQuery的Ajax方法实现服务器的异步交互，着力于实现一个SPA的应用。另外，因为企业邮箱中使用到了多种算法进行数据处理，为了使得用户能够更加直观的认识到的数据分析结果带来的好处，系统大量使用到了数据可视化的技术进行数据处理，并进行了分类的展示。现在具体介绍一下本系统使用到前端框架的一些特性：

2.3.1 Bootstrap 前端设计框架

Bootstrap 是Twitter编写的一个前端工具箱，目的是为了帮助设计师和 Web 前端开发人员快速有效地创建一个结构简单、性能优良、页面精致的 Web 应用。它使用了最新的浏览器技术，可以提供精致的网页排版方式以及表单、按钮、表格、网格栅格化、导航等诸多元素。Bootstrap 的内置样式继承了简洁亮丽的设计风格，任何开发团队都能使用它提供的 HTML 模板、CSS 样式和 jQuery 组件来布署或者重建一个外观漂亮的页面应用。

BootStrap的核心模块从大的方面可以分为布局框架、页面排版、基本组件、

jQuery 插件以及变量编译的 Less 几个部分。下面就来简单了解一下Bootstrap中几个主要用到模块的功能和效果。

(1)页面布局

布局在每个项目中都必不可少，Bootstrap 在 960 栅格布局的基础上扩展了一套优秀的 Grids 布局，而在“响应式布局”中有更强大的功能，能让网格布局适应各种设备。此外，改变模板中的类名，就能实现不同的布局风格。比如常见的“固定布局”，只需要在 HTML 中添加“container”类名；而要实现“流体布局”，只需要在 HTML 中添加“container-fluid”类名。 Bootstrap 还为开发者设计了“Responsive”，令布局框架更为出色。开发者可以在此基础上进行任何样式的覆盖，从而实现理想中的响应式设计。图2-6即为Bootstrap中独特的栅格设计。



图 2-6 Bootstrap 中栅格系统的设计

(2)基本组件

基本组件是 Bootstrap 2 的精华之一，里面都是开发者平时需要的交互组件。比如“网站导航”、“Tabs”、“工具条”、“面包屑”、“分页栏”、“提示标签”、“产品展示”、“提示信息块”和“进度条”等。这些组件都配有“jQuery”插件，运用它们可以大幅度提高用户的交互体验，使产品不再那么呆板无吸引力。要想使用 Bootstrap 基本组件，必须满足三点：第一，最基本的 HTML 结构要对号；第二，需要 Bootstrap 中的“jQuery”插件提供相应功能；第三，在项目中对应的“Tabs”元素上启用“Tabs”功能。至于如何使用，在什么地方使用最好，还需要根据 Bootstrap 所提供的文档，以及各插件的参数，具体问题具体分析。因为只有充分了解，才能灵活运用。图2-7则为Bootstrap基本组件中非常实用的对话框组

件的展示。



图 2-7 Bootstrap 中模态对话框的设计

(3)ICON功能

Bootstrap的ICON是一个继承自jQuery十分具有特色的功能，很多情况下当我们觉得单纯用文字表示苍白无力，就可以使用ICON小图标来让用户增加一点理解程度。Bootstrap中为大家准备了上百种常用的“Icon”图标应用，可以说是能很好的涵盖到各个方面的设计之中。

说起 ICON，到底什么才是ICON？ICON其实是字体文件，用符号和字形的轮廓(像箭头、文件夹、放大镜等) 代替标准的文字数字式字符。但ICON是专门为用户界面设计的，就像其它网站字体一样，使用CSS@font-face在web浏览器里展示。他们可以很好的拥有跨浏览器支持，并实现动态缩放来适应各种分辨率的屏幕，同时还可以呈现不同颜色、阴影或渐变效果。所以在新晋的网页设计中，ICON可以说占到了一个非常重要的地位。图2-8为Bootstrap控件结合ICON使用的效果展示。

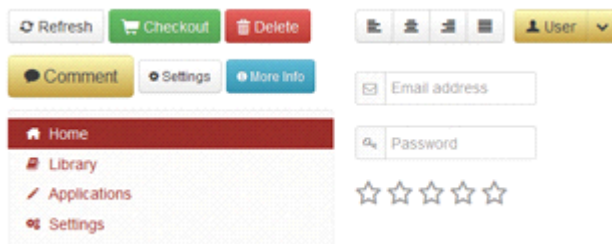


图 2-8 Bootstrap 中小图标的设计

Bootstrap 是一套前端开发利器。它可以帮助我们加速项目开发，让我们身处在一个完备的系统中，拥有一致的设计和实现方法。不需要在外观上花费过多时间，使开发者能将精力集中于更重要的功能。

2.3.2 jQuery Ajax 实现 SPA 应用

Ajax即为异步JavaScript和XML，可以说在某种程度上就是Web2.0的象征。通过 AJAX，JavaScript可以使用JavaScript的XMLHttpRequest对象来直接与服务器进行通信。如果使用AJAX的异步模式，浏览器就不必等用户请求操作，也不必更新整个窗口就可以显示新获取的数据。只要来回传送采用XML格式的数据，在浏览器里面运行的JavaScript代码就可以与服务器进行联系。JavaScript代码还可以把样式表加到检索到的数据上，然后在现有网页的某个部分加以显示，有效的提高用户的体验程度。

虽然Ajax技术并不复杂，但是实现方式还是会因为每个开发人员而有所差异。jQuery提供一系列非常优秀的Ajax函数来帮助我们统一这种差异，这些函数真正做到了高度的封装，重载全面，并且让调用Ajax更加简单。jQuery提供了几个用于发送Ajax请求的函数。其中最核心也是最复杂的是jQuery.ajax()，所有的其他Ajax函数都是它的一个简化调用。当我们想要完全控制Ajax时可以使用此结果，否则还是使用简化方法如get, post, load等更加方便。具体调用方式在这里就不再陈述了，可以前往查看jQuery的官方API文档。

例如本文设计的企业邮箱系统中就大量使用到了通过jQuery实现的Ajax技术，在收取邮件中，我们可以无刷新的获取到所有邮件，这样的好处不仅仅是提升了用户的交互体验，并且用户还可以通过点击功能按钮来即时刷新并获取到新

的邮件，并具有通知的推送功能，这是以往传统邮箱必须反复全盘刷新邮件页面所不能做到的。同时，本文设计的企业邮箱系统意在于设计一个SPA应用，即Simple Page Application，这是一个全新的设计理念，也就是整个邮箱应用完全不进行页面的跳转，所有的数据均是使用Ajax技术与后台进行交互和处理，保持数据的高度统一性。这样设计的好处在于，用户不用再担心数据统一的问题而反复刷新页面，异步数据的传输能迅速提高用户的使用效率并节省页面跳转带来的资源浪费，当然也极大的提高了用户体验。

2.3.3 Nvd3 数据可视化建模

现在互联网上的数据可视化框架已经呈雨后春笋般的势头出现，强大的图表类库非常多，但是在企业邮箱系统的数据可视化设计中，我们只使用到了NVD3.js，首先这个类库的强大是一定的，并且能够优雅的进行展示，另外主要还是这个类库基于知名的d3.js类库。可以帮助你创建可重用的d3.js图表和图表组件而不降低d3.js的强大功能。

NVD3.js是一种奇妙的基于javascript的数据可视化库，允许使用者将数据绑定到DOM，然后再转换成为图表。NVD3不是一个庞大的框架，像想象那样提供许多功能。相反，这个类库非常新，拥有的组件和图表不是很多，但是可以帮助你轻松的管理和自定义图表，而不必自己去构建所有的开发细节，同时提供良好适应性。公开全部功能，如CSS3、HTML5和SVG网页标准，很少的开销，NVD3速度非常快。支持大数据集和动态交互和动画效果。支持的图形包括：线状图，柱状图，气泡图，区域图及其组合图等。支持所有主流的浏览器。图2-9即为Nvd3生成的一些图表示例。



图2-9 一些Nvd3生成图表的示例

3 企业邮箱异常检测算法研究

3.1 邮件重要度检测算法研究

在本文中，所提出的一个重要的方法就是邮件重要度的算法。邮箱系统使用了动态的权重分配技术，通过观察用户与用户之间的联系密度，以及用户在对收发邮件时所做出的行为来动态分配给电子邮件以重要度。对于企业用户来说，这是非常具有实际意义的，往往在企业环境中，邮箱中大量充斥着会议通知，日常活动，商务信息，任务计划，工资表单等，企业用户往往困扰于在大量邮件中找到真正需要，重要的邮件。这些邮件的重要度越高，它们所在邮箱的顺序加权也就更高。这意味着高重要度的邮件总会放置在用户看到邮件列表的前列，通过这样的方式来提高用户处理大量邮件的效率，并且通过重要度来自动将那些垃圾邮件过滤到底部，企业用户就能轻松分辨这些垃圾邮件。

企业电子邮件的重要度通过计算两种类型重要度的混合重要度而得到。重要度1是邮件发送者的一个重要度，因为电子邮件的重要度非常受其发件人影响，例如一个垃圾邮件或宣传邮件的发送者可能总是同一人，始终发送垃圾的邮件会严重降低他的重要度，导致即使其发送的某封垃圾邮件经过误操作不小心在重要度2上体现比较高，在重要度1上还是能将其分辨成为重要度低的邮件。重要度2是电子邮件上执行的操作的权重的总和，这些执行操作包括，当邮件被执行到这些操作后，会相应地对这些行为进行重要度记分，加分的行为包括设置邮件为重要、将其加锁，多次阅读邮件等等，这些行为都证明这封邮件的价值相当高。而减分的行为则包括将其划分为垃圾邮件，从未读过邮件等等，这些行为将大大的降低其重要度2的分数。

通过反复研究和测试重要度对邮件排名的影响，最终总结出来重要度1，重要度2，总重要度相应的计算公式。重要度1的计算公式为：

$$(1) ID1 = 0.5 + \frac{send - rec}{send + rec}$$

在这个公式中ID1就是重要度1(Important Degree)简称，send指用户针对某个联系人发送邮件的总数，rec则指用户针对某个联系人收取邮件的总数。可以看

出这个式子的计算分成两个部分，前者是一个固定的常数，这是用于控制重要度1在总重要度计算中所占的权重，所以可以根据平衡来进行配置；后者则是一个针对联系人发件和收件数量的一个计算比值，根据这个比值我们可以判别联系人的一个权重，发件数量越多，证明这个联系人是联系的重要对象，他的重要度就会呈上升趋势，反之收件越多，则会使重要度呈下降趋势。

重要度2的计算公式为：

$$(2) ID2 = 1 + (\sum Act)$$

在这个公式中ID2即为重要度2， $\sum Act$ 指邮件操作权重的总和。同样，这个计算公式也被分成两个部分，前者同样是一个固定的常数，这是用于控制重要度2在总重要度计算中所占的权重；后者是用户在发送和接收中对这个邮件的操作，这些加权包括设置邮件为重要、将其加锁，多次阅读邮件，其划分为垃圾邮件等等，这个数字越高，证明这封邮件对于用户来说价值是非常高的，反之则又成为垃圾邮件的可能。

总重要度的计算公式为：

$$(3) Total = ID1 * ID2$$

通过重要度1(ID1)和重要度2(ID2)的一个乘积计算就可以得到总重要度。可以看得出来，总重要度是邮件联系人重要度和邮件操作重要度的一个平衡值，所以有效的提高的公式的稳定程度。通过总重要度的评估，我们便可以为每个邮件赋予一个重要度的加权值，这个加权值将会在用户查看企业邮箱邮件的时候得到最大的体现。

3.2 安全加密算法研究

通过数据加密技术，可以极大的提高企业邮箱数据传输的安全性，保证邮件的数据不会轻易的外泄。一个数据加密系统包括加密算法、明文、密文以及密钥，密钥控制加密和解密过程，数据加密过程就是通过加密系统把原始的数字信息(明文)，按照加密算法变换成与明文完全不同得数字信息(密文)的过程。

数据加密算法也有很多种，按照发展进程来分，经历了古典密码、对称密钥密码和非对称密钥密码阶段，古典密码算法有替代加密、置换加密；对称加密算

法包括DES和AES；非对称加密算法包括RSA、背包密码、McEliece密码、Rabin、椭圆曲线、ElGamal D_H等。目前在数据通信中使用最普遍的算法有DES算法、RSA算法等。在企业邮箱系统中，我们实验性的将最常用的DES，AES算法和自行研究的ECC算法融合到平台之中，对平台中的邮件文本进行加密。保证从用户点击发送开始到接受者打开邮件为止全程的文本数据加密，有效的防止了第三方攻击者截取邮件并盗取信息的行为，为企业邮箱提供了一份更高的安全保障。以下本文就从对称加密和非对称加密两个方面来具体介绍以下各种加密方法的强度和优劣。

3.2.1 对称加密技术

(1)DES加密算法

DES是一种对二元数据进行加密的算法，数据分组长度为64位，密文分组长度也是64位，使用的密钥为64位，有效密钥长度为56位，有8位用于奇偶校验，解密时的过程和加密时相似，但密钥的顺序正好相反。

DES算法的弱点是不能提供足够的安全性，因为其密钥容量只有56位。由于这个原因，后来又提出了三重DES或3DES系统，使用3个不同的密钥对数据块进行(两次或)三次加密，该方法比进行普通加密的三次块。其强度大约和112比特的密钥强度相当。

(2)AES加密算法

AES是美国联邦政府采用的一种区块加密标准，它的英文名称就是高级加密标准(Advanced Encryption Standard)的缩写，并且设计的初衷就是用于取代DES加密算法。

AES依然采用对称分组密码体制，密钥长度的最少支持为128、192、256，分组长度128位这就比DES的56密钥要强的多，并且直到现在都没有有效的破解方法。

3.2.2 非对称加密技术

(1)RSA算法

RSA算法既能用于数据加密，也能用于数字签名，RSA的理论依据为：寻找两个大素数比较简单，而将它们的乘积分解开则异常困难。在RSA算法中，包含两个密钥，加密密钥PK，和解密密钥SK，加密密钥是公开的，能够进行不具备安全条件的远程传输。

RSA算法的优点是密钥空间大，缺点是加密速度慢，如果RSA和DES结合使用，则正好弥补RSA的缺点。即DES用于明文加密，RSA用于DES密钥的加密。由于DES加密速度快，适合加密较长的报文；而RSA可解决DES密钥分配的问题。

(2)ECC算法

ECC又名椭圆曲线算法，椭圆曲线密码体制来源于对椭圆曲线的研究，指的是由韦尔斯特拉斯（Weierstrass）方程： $y^2+axy+a_3y=x^3+a_2x^2+a_4x+a_6$ 所确定的平面曲线，其安全可靠性的保证也来自于有限域 F_p 上的离散对数问题是公认的难题，比有限域上的离散对数问题还要难上许多，所以其安全性也更高，是未来非对称密码发展的重要方向。

但ECC现在还处在高速发展的情况下，目前密码学组织对椭圆曲线的掌握还并不深入，也没有一个成熟，稳定的实用模型。毕竟椭圆曲线本身就处于解析几何这个数学中就较难的概念之中，现在又赋予其离散对数的难题理论，研究起来还是非常有难度和深度的，但是这种情况持续时间越长，更会让人们相信这种密码是非常难以破解的。

3.3 基于图论和社会网络方法的邮件团体分析算法研究

图论是数学的一个分支。它以图为研究对象。图论中的图是由若干给定的点及连接两点的线所构成的图形，这种图形通常用来描述某些事物之间的某种特定关系，用点代表事物，用连接两点的线表示相应两个事物间具有这种关系，整体结构图如图3-1所示。

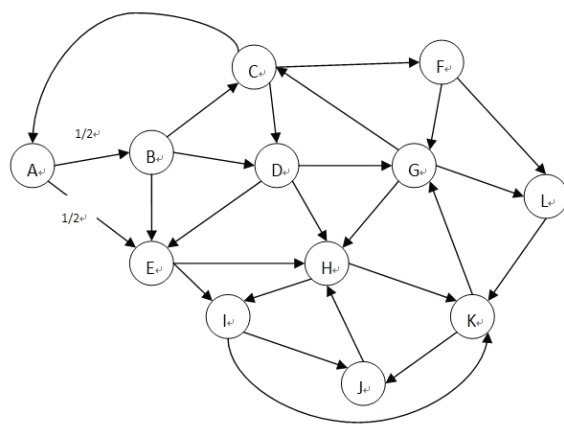


图 3-1 有向图示例

在本文中的企业邮箱系统，正是使用图论作为邮件交互关系网分析的基础理论。整个企业邮箱平台的收发邮件数据可以通过一个数据矩阵来表示，而如果将这个矩阵通过图论的知识来体现，就可以一张直观可视化的有向图来表示。

这张邮件网络有向图是由“点”和“边”组成，其中“点”代表在这个有向图中的联系人；“边”则代表连接的两“点”之间存在联系。这种关系可以是单向的，也可以是双向的。在企业邮箱系统中，边的方向具体就是指代“发送邮件”或者“回复邮件”的行为，例如A指向B代表A发送邮件给B，我们也可以认为这样一种行为指向代表A将B视为重要的联络人，这点在社交关系学中也有证明，如果在研究朋友社交关系的有向图时，边的方向可以代表“喜欢”的倾向性。图论中有大量概念和方法来观测邮箱图的结构模型，并用于分析有向图中节点之间的结构内涵。其中主要概念包括度数，密度，距离，中心度等，方法则包括有向图深度遍历和有向图广度遍历两种，现在我们就通过图2-2来具体介绍一下每个概念和方法的定义：

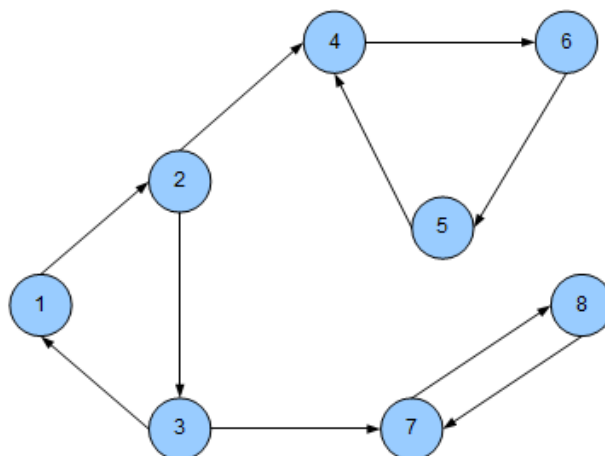


图 3-2 有向图概念解说图

(1)度数(Degree)

在图论中，由一条线连着的两个点是“相邻的”。与某个点“相连”意味这两点是一个邻居的关系。于是，我们就用“度数”来表示某点邻居的总数量。如图2-2中的有向图，点1的度数就是2，应为他与点2和点3相邻。并且在有向图中还可以划分得更细一点，度数又可以根据边的方向分为“入度”(in-degree)和“出度”(out-degree)。“入度”是指直接指向该点邻居的总数，例如图2-2中点7的入度就是2。“出度”则相反，是指该点指向的邻居的总数，例如图2-2中点2的出度就是2。在企业邮箱系统中，我们大量使用出入度来体现和监测邮件网络关系图中的核心人物和边缘人物，用以决策图中的关键联系人。

(2)密度(Density)

图论中有向图的密度指图中实际拥有的连线总数与可能最多拥有的连线总数之比。密度是有向图中的一个重要的测量变量。密度越大的有向图具有越高的紧密性，反之紧密性也越低。在企业邮箱系统中，我们通过密度来检测邮件网络关系图中的聚合程度，密度越大说明企业内部的员工得到了充分的交流和沟通，说明企业的营造的环境氛围是和谐向上的；如果密度越低则说明企业内部员工要不各自为政，要不自己行事孤僻，说明企业内部环境出现了问题，此时我们就可以利用邮件网络关系图来找到这些员工，并解决他们的问题。

(3)距离(Distance)

两点之间最短的连接线路的长度叫做两点的距离，如图2-2的有向图中，点1和点6的距离就是3。特别注意的是，距离的测量必须考虑边的方向性，到达一点必须其中间所有的边都是指向目的地一点的，这样才能认为是连通的线路。在企业邮箱系统中，我们可以通过距离来定位最远的交际范围，从而界定他工作的空间，还可以用于确定企业间各部门的联系情况。

(4)中心度(Center-Degree)

在有向图中，“中心度”主要包括“中间中心度”和“接近中心度”。“中间中心度”是指某点处于中间人位置的高度程度，表示一个点有多大程度是其他点的“中介者”，有多大程度掌握信息的能力。“接近中心度”表示某点与其他点的紧密程度，表示某点与其他点的紧密程度，表示一个点有多大程度被其他点掌握，当某点“接近中心度”越高，表明他与其他行动者非常亲密。在企业邮箱系统中，我们可以用这个方法找到突出的企业员工，并进行人才的筛选和提拔。

(5)深度优先遍历(Depth First Search)

深度优先搜索法是树的先根遍历的推广，它的基本思想是：从图G的某个顶点 v_0 出发，访问 v_0 ，然后选择一个与 v_0 相邻且没被访问过的顶点 v_i 访问，再从 v_i 出发选择一个与 v_i 相邻且未被访问的顶点 v_j 进行访问，依次继续。如果当前被访问过的顶点的所有邻接顶点都已被访问，则退回到已被访问的顶点序列中最后一个拥有未被访问的相邻顶点的顶点 w ，从 w 出发按同样的方法向前遍历，直到图中所有顶点都被访问。其递归算法如下：

DFS(G)

```
1  for each vertex  $u \in G.V$ 
2       $u.color = WHITE$ 
3       $u.\pi = NIL$ 
4   $time = 0$ 
5  for each vertex  $u \in G.V$ 
6      if  $u.color == WHITE$ 
7          DFS-VISIT( $G, u$ )
```

DFS-VISIT(G, u)

```
1   $time = time + 1$                 // white vertex  $u$  has just been discovered
2   $u.d = time$ 
3   $u.color = GRAY$ 
4  for each  $v \in G.Adj[u]$           // explore edge  $(u, v)$ 
5      if  $v.color == WHITE$ 
6           $v.\pi = u$ 
7          DFS-VISIT( $G, v$ )
8   $u.color = BLACK$                 // blacken  $u$ ; it is finished
9   $time = time + 1$ 
10  $u.f = time$ 
```

(6)广度优先遍历(Breadth First Search)

图的广度优先搜索是树的按层次遍历的推广，它的基本思想是：首先访问初始点 v_i ，并将其标记为已访问过，接着访问 v_i 的所有未被访问过的邻接点 $v_{i1}, v_{i2}, \dots, v_{it}$ ，并均标记为已访问过，然后再按照 $v_{i1}, v_{i2}, \dots, v_{it}$ 的次序，访问每一个顶点的所有未被访问过的邻接点，并均标记为已访问过，依次类推，直到图中所有和初始点 v_i 有路径相通的顶点都被访问过为止。其非递归算法如下：

```

BFS( $G, s$ )
1  for each vertex  $u \in G.V - \{s\}$ 
2       $u.color = \text{WHITE}$ 
3       $u.d = \infty$ 
4       $u.\pi = \text{NIL}$ 
5   $s.color = \text{GRAY}$ 
6   $s.d = 0$ 
7   $s.\pi = \text{NIL}$ 
8   $Q = \emptyset$ 
9  ENQUEUE( $Q, s$ )
10 while  $Q \neq \emptyset$ 
11      $u = \text{DEQUEUE}(Q)$ 
12     for each  $v \in G.Adj[u]$ 
13         if  $v.color == \text{WHITE}$ 
14              $v.color = \text{GRAY}$ 
15              $v.d = u.d + 1$ 
16              $v.\pi = u$ 
17             ENQUEUE( $Q, v$ )
18      $u.color = \text{BLACK}$ 

```

在本文设计的邮件系统中，广泛应用到了以上图论的各个概念和方法的知识，并具体的进行了设计和融合，通过遍历方法的搜索和图论概念的分析，最后得到的数据能够证明这些方法在企业邮件系统中的联系网络分析中是行之有效的。

4 安全与异常检测功能验证

4.1 发件安全检测模块设计

发送邮件模块是企业邮箱系统一个重要的组成要素，在本文设计的企业邮箱系统中，使用了Bootstrap，Flat UI和CKEditor做为前端的设计框架，后台接收前端的输入参数，并使用JavaMail中的Message组装邮件文本，通过Transport组件发送到邮件服务器，图4-1为发件页面的展示。下面就对发件模块中一些设计重点进行描述和分析。

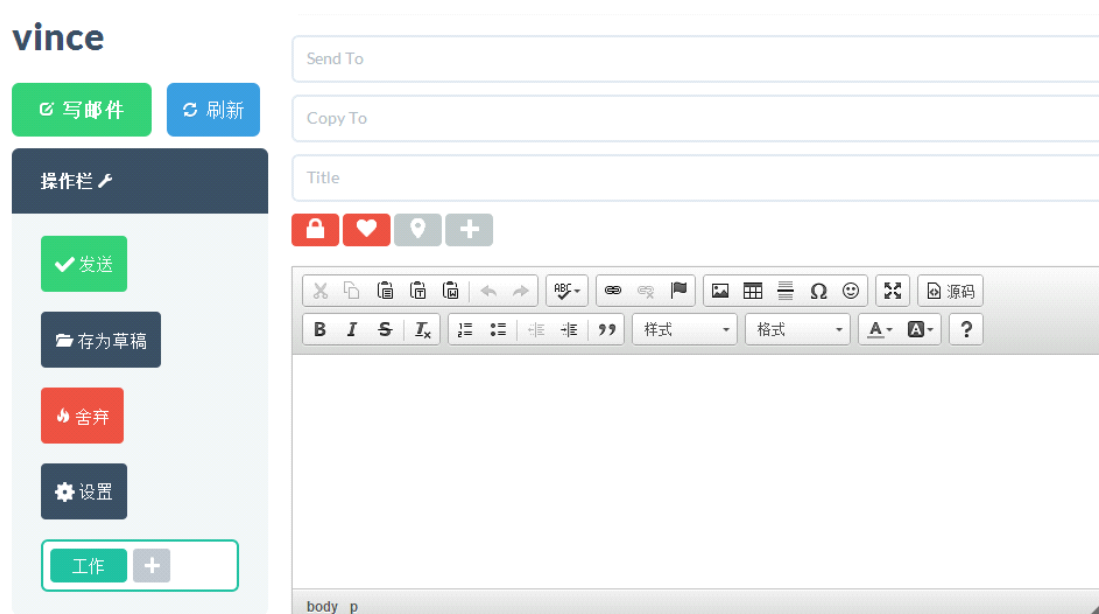


图 4-1 邮箱系统中发件模块的设计页面

4.1.1 企业邮件重要度安全标记

邮件重要度是与邮件重要度算法关系十分紧密的一个功能，通过点击图5-1中红色的“心”型按钮，我们可以打开或关闭这封邮件对于发往的联系人是否重要的标记。

这个功能主要是使用在当邮件包含有重要事件或任务时，提示对方需要认真阅读此封邮件，同时会提高此封邮件的重要度。后续增强设计还可以附加请求回执等等。图4-2为重要度标记打开时的标志。



图 4-2 重要度标记已打开

4.1.2 安全锁定邮件设计

邮件锁定是一个十分具有特色的功能，我们通过点击图5-1中红色的“锁”型按钮可以打开一个模态对话框，这个对话框包含两个输入的内容：加锁密码和选择是否用短信发送给收件人，如果选择发送，会提示输入对方的联系号码。

邮件锁定的设计思路是取自我们在打开RAR文件时会提示要输入密码，这其实提升了很多的安全性。同样的，我们把这个功能配置到邮箱时，它也能取得很好的效果，避免无关人员打开这封邮件，假设这封邮件涉及到工资，家庭事务等私密信息时，它在用户离开座位时能提供更高一层的保护。同时，如果填好对方的手机号码，内部会通过调用飞信模块自动将密码发送到对方手机上，并附上说明，对方就可以很清楚的知道了。图4-3即为提示输入密码和手机号码的输入框。

图 4-3 提示输入密码和手机号码

4.1.3 安全加密选择

安全加密为企业邮件系统更加提升了一层安全性的保障。我们通过点击图5-1中左侧工具栏中的“设置”按钮即可以打开一个模态对话框。在模块对话框中可以自由选择选择邮件文本加密时使用的算法，默认是使用DES进行加密。

安全加密能有效的阻止第三方从网络链路中截取到原始的信息，通过加密以后，对方没有密钥所以很难破解加密后的信息，知道收件人打开邮件时系统才会自动进行解密。有效的保证了邮件信息的安全。图4-4为在设置选择加密的方式。

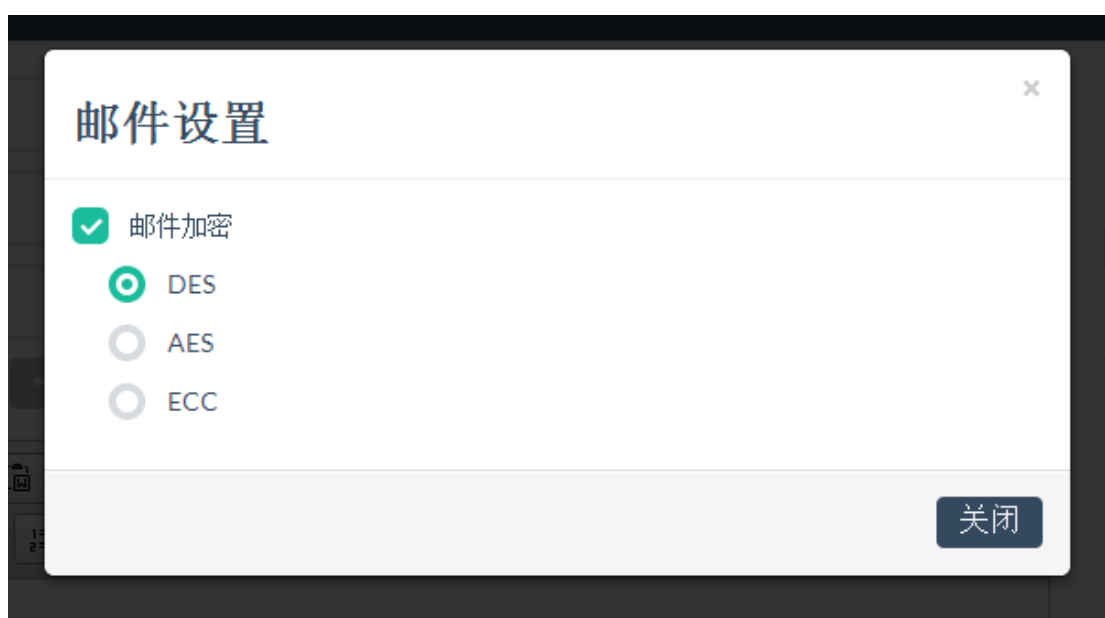


图 4-4 提示选择邮件文本加密的方式

4.1.4 富文本邮件编辑

富文本编辑器是现在非常流行的一种互联网文本强化组件，图4-1中右下角的编辑器即为富文本编辑器。我们可以首先在编辑器内写上邮件的文本，在通过选择编辑器上方的功能按钮来提供额外的效果，例如字体的颜色、大小和修饰标记。也可以通过功能按钮来插入表情和图片，使得邮件的内容更加多样化。

这里的富文本编辑器是使用当前设计比较优秀的CKEditor，CKEditor是一个专门使用在网页上属于开放源代码的所见即所得文字编辑器。它志于轻量化，不需要太复杂的安装步骤即可使用，并且还可以根据自己的需要来调整功能按钮的

多少，来让编辑器看上去更加的整洁。图4-5为CKEditor的用户使用界面。

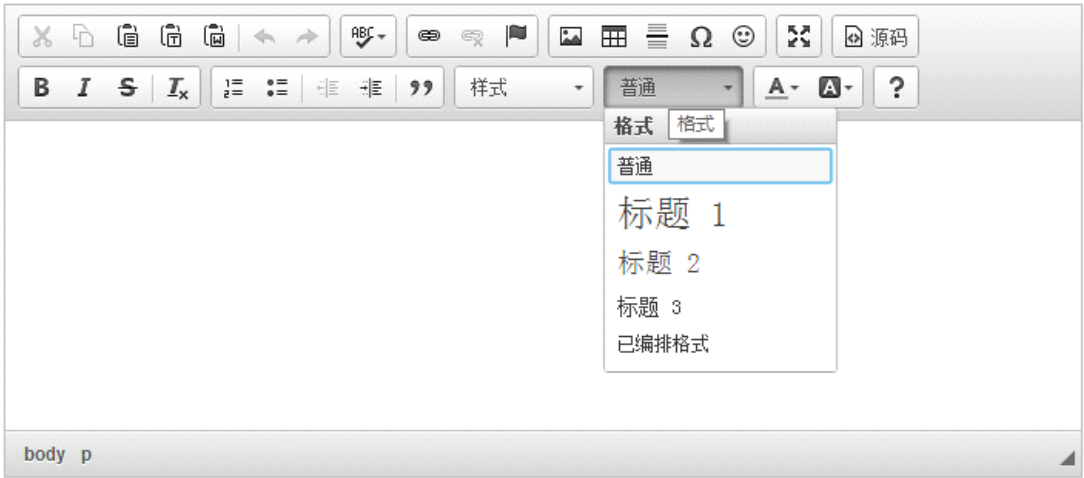


图 4-5 富文本编辑器界面

4.2 收件安全检测模块设计

收件邮件模块同样是企业邮箱系统重点设计的一块内容，其中包括收件和阅读邮件两大功能。收件模块使用大量Ajax的方式从后台传输邮件数据，并通过邮件重要度算法进行测试和计算，提供了大量相关的排序分类方式，使得用户能简单地对邮件进行管理和监测。另外收件预览界面还提供了数个图标来控制邮件的分类。图4-6为收件页面的展示。下面就对收件模块中一些设计的重点进行描述和分析。



图 4-6 邮箱系统中收件模块的设计页面

4.2.1 收件箱安全检测设计

收件箱包含几个重要的部分，如图4-6所示，左侧部分是收件箱的操作栏，提供多种基于邮件重要度的排序算法，这一部分的内容放在下一节进行描述。右侧部分就是收件箱的部分了。

可以观察到邮件的预览包含阅读状态，摘要这两个字段，其中阅读状态通过绿和黄两种颜色来区分已读和未读邮件，通过点击这个图标我们可以对邮件进行阅读；摘要可以对邮件的部分内容进行预览，这样做的好处是可以通过这些摘要信息来预先判断这封邮件是否需要阅读，而不用打开邮件。另外如果邮件处于加锁状态时，系统将会显示锁定信息，保证私密信息不会泄露。

另外邮件预览的右侧部分包括三个标记属性，其中标记和重要可以点击。“标记”是用户认为这封邮件可能在短期内再派上用场而设置的标识，“重要”则是指发件人的提示，用户可以手动取消掉这个重要标识。最后的加锁的点亮取决根据发件人是否进行了邮件加锁，这是不可更改的。图4-7展示了三种标记属性的使用状态。

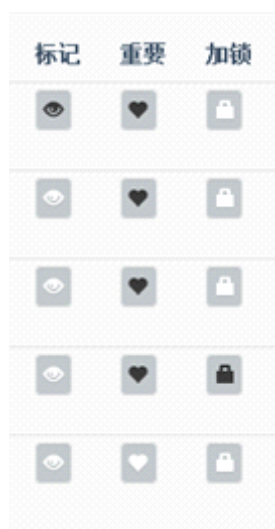


图 4-7 打开加锁邮件时出现的对话框

4.2.2 邮件阅读安全锁

邮件阅读设计相对而言较为简单，主要包含发件人，标题和邮件文本等信息，

此时邮件文本会以HTML来显示富文本编辑器所编辑的内容。图4-8为邮件阅读时的界面。

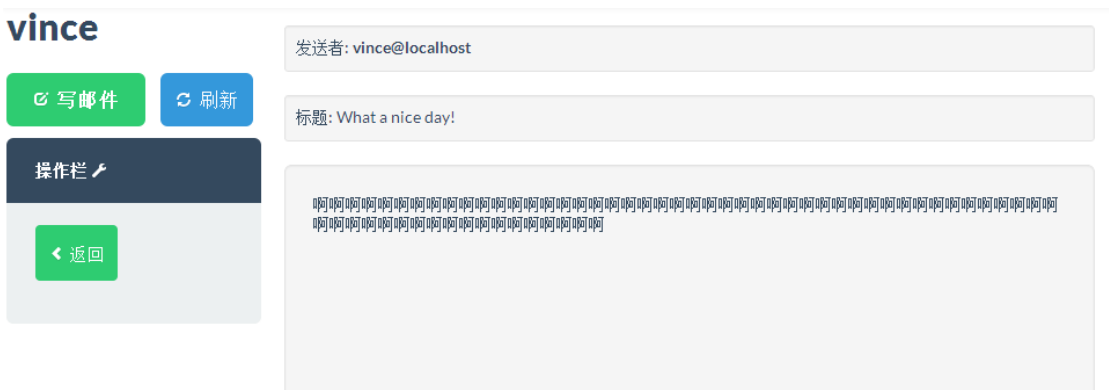


图 4-8 邮件阅读界面

注意的是如果进行了邮件加锁，在打开邮件的时候会提示输入加锁密码，输入正确才会继续打开。图4-9为提示输入加锁密码的对话框。



图 4-9 打开加锁邮件时出现的对话框

4.3 重要度排序异常检测模块设计

邮件重要度排序模块设计是企业邮箱系统设计的核心内容之一，其中包括只智能排序，未读排序，加锁排序等多种排序方式。其中智能排序就是通过取出邮件重要度算法计算后的结果，将数值的大小转化为排序的依据，有效的将垃圾邮

件和商务重要邮件分离开来，提高了用户使用邮箱的效率，降低了滚动过度带来的负面影响，并使得收件箱干净整洁。

4.3.1 智能邮件检测

智能邮件排序的依据就是邮件重要度公式在对邮件分析后产生的数值结果，这个值越高，说明邮件的重要度也越高，在排序是就会在体现在收件箱的顶部，相反垃圾邮件则不会出现在首要位置。重要度计算的公式即为：

$$ID = (0.5 + (Send - Rec) / (Send + Rec)) * (1 + E(Operation))$$

其中ID就是邮件的重要度，Send为与联系人发送邮件的数量，Rec为与联系人接收邮件的试炼，E(Operation)为对于邮件操作权重的综合(例如标记、重要、加锁等都会赋予不同的权重)。

图4-10和图4-11分别标示了邮件在智能排序前后的状态，我们通过仔细的对比观察，发现Emily同学发来的一封未读邮件的位置明显出现的上升，这是因为其被用户主动点开了“标记”，并保留了“重要”的状态，对于用户来说重要性不言而喻。另外由Emily发送的一封带锁邮件也提高了排名，这是因为带锁的邮件涉及到私密的信息，对于用户来说会相当重要。另外Emily的一封没有任何加权的邮件毫无疑问的降到了排名最低端，证明重要度算法是正确的。

已读	发件人	标题	摘要	时间	标记	重要	加锁
已读	emily	1234	1234...	2013-04-25 15:20:30			
已读	emily	What a nice day!	啊啊啊啊啊啊啊...	2013-04-26 15:21:35			
已读	emily	This is a secret mail!	内容被锁定	2013-04-26 16:03:21			
未读	emily	Today is Sunday!	Hello! ...	2013-05-05 13:09:53			
已读	alice	123456	Hello!<...	2013-05-07 00:56:46			

图 4-10 未使用智能排序时的收件箱

已读	发件人	标题	摘要	时间	标记	重要	加锁
未读	emily	Today is Sunday!	Hello! ...	2013-05-05 13:09:53			
已读	emily	This is a secret mail!	内容被锁定	2013-04-26 16:03:21			
已读	alice	123456	Hello!<...	2013-05-07 00:56:46			
已读	emily	What a nice day!	啊啊啊啊啊啊啊啊...	2013-04-26 15:21:35			
已读	emily	1234	1234...	2013-04-25 15:20:30			

图 4-11 使用智能排序后的收件箱

4.3.2 其他邮件检测

除了使用智能排序以外，还设计了多种方便用户进行分类的排序，例如未读标识排序，重要标记排序，加锁标识排序等等。后续增强设计时，还会考虑到使用邮件的标签进行排序。图4-12和图4-13分别展示了邮件经过重要标识排序和加锁标识排序后的结果。

已读	发件人	标题	摘要	时间	标记	重要	加锁
未读	emily	Today is Sunday!	Hello! ...	2013-05-05 13:09:53			
已读	emily	What a nice day!	啊啊啊啊啊啊啊啊...	2013-04-26 15:21:35			
已读	alice	123456	Hello!<...	2013-05-07 00:56:46			
已读	emily	This is a secret mail!	内容被锁定	2013-04-26 16:03:21			
已读	emily	1234	1234...	2013-04-25 15:20:30			

图 4-12 使用重要标识排序后的收件箱

已读	发件人	标题	摘要	时间	标记	重要	加锁
已读	emily	This is a secret mail!	内容被锁定	2013-04-26 16:03:21			
已读	emily	1234	1234...	2013-04-25 15:20:30			
已读	alice	123456	Hello!<...	2013-05-07 00:56:46			
已读	emily	What a nice day!	啊啊啊啊啊啊啊啊...	2013-04-26 15:21:35			
未读	emily	Today is Sunday!	Hello! ...	2013-05-05 13:09:53			

图 4-13 使用加锁标识排序后的收件箱

4.4 数据可视化异常检测模块设计

数据分析及可视化模块设计同样是企业邮箱系统设计的核心内容之一，可以发现，整个在本文设计的企业邮箱平台中，在页面每个合理的位置都放置了可以供用户参考的数据可视化图表。这些图表包括对三个模块：用户发送邮件的数据分析，用户和联系人交互的数据分析以及用户邮件社交网络关系。通过这些模块提供直观的数据显示，用户就能根据关系事实来进行发件的决策，我们要求的目的是无需去告诉用户哪些地方异常，而是让用户一眼就能明白并清楚哪些地方出现了异常。下面就分别对这些可视化模块进行解析：

4.4.1 邮件联系人热度异常检测

邮件联系人热度榜出现在收件页面的右侧，如图 4-14 所示。当用户登陆进入企业邮箱时，就可以看到自己最近联系人收发邮件的状况，数据视图中的信息包括最近联系最多的 8 个用户的收件、发件数。用户可以通过点击顶部的周，月，年来切换统计时间的长度，还可点击寄件或收件按钮来具体关注一方面的信息。

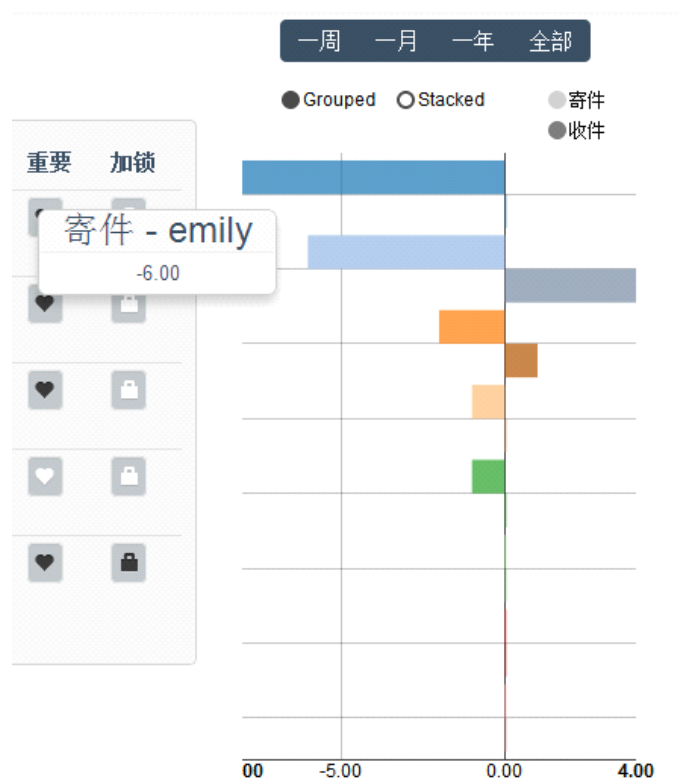


图 4-14 邮件联系人热度排行榜图示

4.4.2 邮件时间段分布异常检测

邮件时间段分布图需要进入系统数据栏目下，并点击“我的月报”，就可以查看邮件分布的数据分析图了。分布分析图包括两个部分，一部分是一个月中按星期来分布的邮件统计图，另一部分是一个月中按一天时间段来分布的邮件统计图。通过分布图我们可以有效的确定我们在星期几，哪个时间段是收到邮件的高峰期，需要我们特别去注意。同时也能对用户的一个工作状态产生一定的体现。图 4-15 为一个月中按一天时间段来分布的邮件统计图示例。

一个月中按时间段分布的邮件

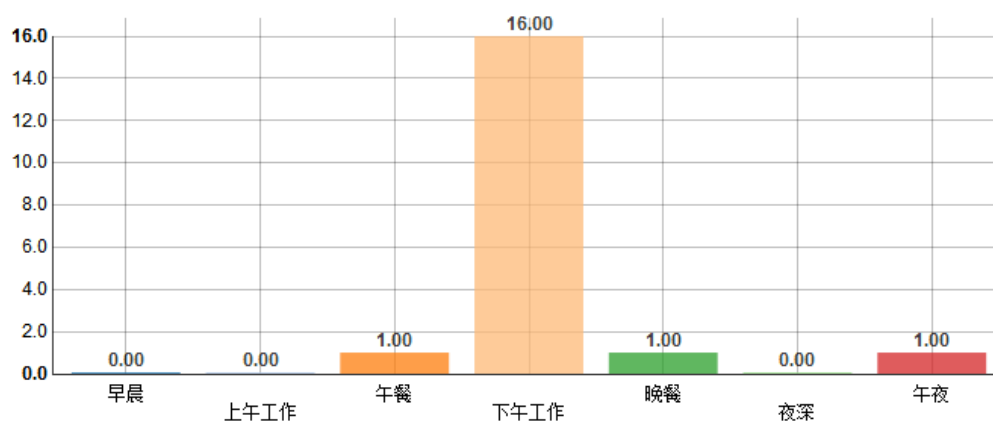


图 4-15 邮件时间段分布图示

4.4.3 邮件联系人异常检测

邮件联系人分析图需要进入系统数据栏目下，并点击“联系总汇”，就可以查看邮件联系人的数据分析图了。联系人的分析图也包括两个部分，一部分是一个月中按邮件收发比例进行排序的联系人分析图，另一部分是按照联系人 VIP 分数排名的联系人分析图。通过分布图我们可以有效的确定与那些人保持了密切的邮件联系，那些人我们可能忽略没有重视(可能这个联系人本身是我们需要重视的)，我们就可以同过这个图对自己最近的联系行为进行一个审视，并决策下一步是应该继续和某联系人保持密切联系或者是要开始重视一个原本被忽略的联系人。图 4-16 为按联系人 VIP 分数排名的联系人分析图。

联系人按VIP分数排名

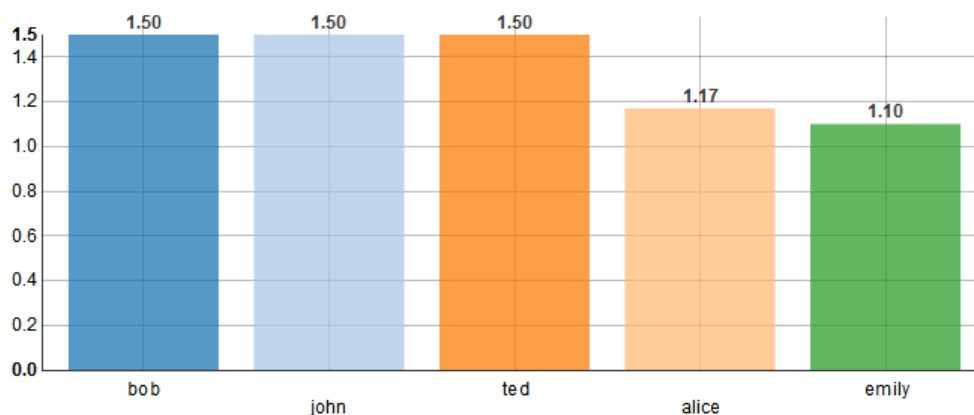


图 4-16 邮件联系人 VIP 分数图示

4.4.4 邮件社会关系网络异常检测

邮件社会关系网络图是一个非常具有代表意义的图表，也是本系统中一个最符合核心思想的一张图。他的特点在与能够对整个企业邮件生态系统中的人进行一个全面的审视。哪些人和哪些人有联系，哪些人形成了一个密集交流的圈子，哪些人在圈子中成为核心人物，通过这个邮件关系网络，都可以一目了然的展现在用户面前，分辨垃圾邮件发送者还是重要邮件发送者可以是件很轻松的事情。当然，邮件网络图只会延伸到一个限定的层数，并形成一个人联系集合。用户还可以自由的对这张图表进行操作，比如删去一条连接线，添加一个联系人等等。

我的圈子

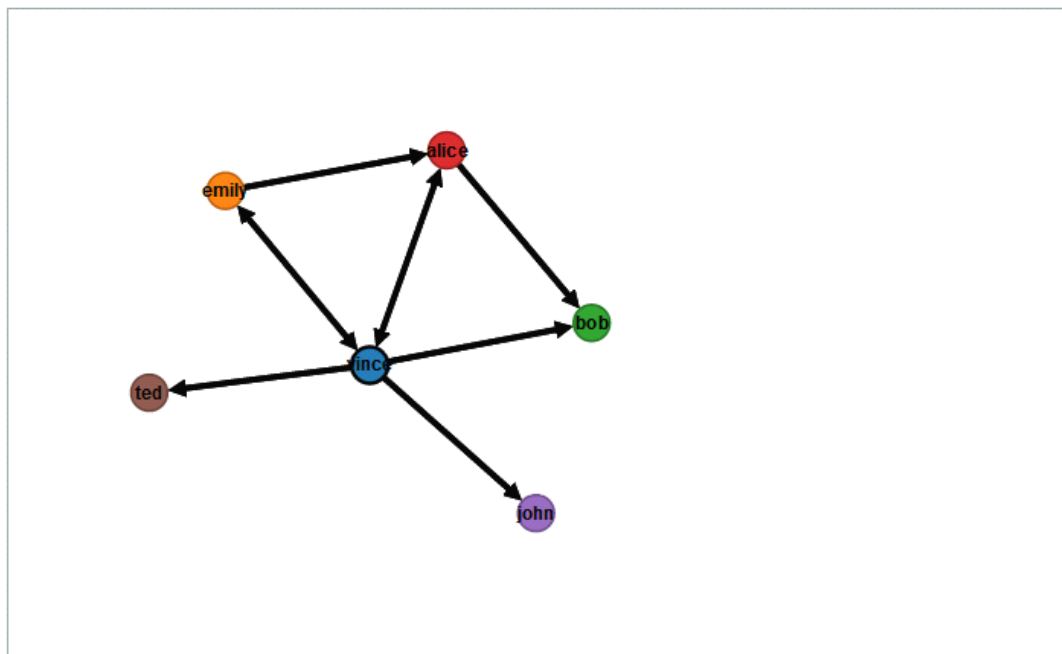


图 4-17 邮件社会网络关系图示

5 结论

本文是以自身用J2EE平台搭建的企业邮箱系统软件作为基础，主要使用基于权重的邮件重要度算法和基于图论的邮件社交网络算法进行邮件的数据分析，主要针对邮件和联系人这两个集合进行异常分析与检测。数据分析的方式是从搭建的企业邮箱系统中抽取日常邮件的收发数据进行统计分析，再通过搭载邮件权重算法进行整体的反复测试和计算，提取出邮件的一个重要度算法模型，然后搭载到邮件上进行分类和检测。另一方面，系统也将这些邮件数据对应的联系人进行一个二维关系的转换，形成一张具体映射到每个联系人的邻接矩阵，构建数据图表，再转化到图论中的有向图，最后构造一张关系网络图，直观的体现网络中的变化和扩展。

本文对基于权重的邮件重要度算法和基于图论的邮件社交网络算法的探讨过程，主要分为几个步骤进行：

(1)对现有各种邮件系统的研究和应用的设计进行对比分析，得到现有系统各方面设计的优点。

(2)对邮件系统安全性的问题和对于数据直观分析的漠视进行了总结。

(3)通过设计邮箱权重算法，反复在搭建的J2EE企业邮件平台进行数据测试，得到并验证重要度算法公式。

(4)对计算机和数学领域中的图论背景进行了分析和理解，探讨了图论在社交网络上的应用，并分析社交网络的重要性。

(5)基于图论对邮件网络进行分析，通过对在J2EE企业邮件平台上的数据进行测试和设计，形成数据可视化的图表，并映射成为一张总的交互关系表。

(6)通过与现代技术的结合，研究和分析如何将重要度算法和邮件网络分析结果良好的映射到交互设计中，为用户提供安全性和可操作性的高度结合，并给出了实现。

完成这些分析步骤后，我们最后通过系统的测试得到了一个可用的系统，在每个副本的研究探讨中本文主要得到出以下结论：

(1)对现有各种邮件系统的研究和应用的设计进行对比分析：

通过对现有邮件客户端或者浏览器端的基本功能和设计管理功能的分析，得到了很多现有邮件系统很多优秀的设计思路和交互体验提升方式，但可以得出现有的邮件系统只关注邮件层面，缺少深层次的利用邮件的“关系属性”来进一步地为用户提供以关系事实为基础的邮件管理方式的结论。

(2)总结对邮件系统安全性的问题和数据分析的漠视

当前多种邮件系统确实已经开始在很多方面上注重了安全的问题，但是配置都有些生涩难懂，用户不能的对自己的邮箱的安全性有一个直观的认识，或是对自己的行为数据进行观察和分析，这样很可能用户根本就不会考虑去使用。

(3)设计邮箱权重算法，提取重要度算法公式

邮件的价值可以根据各种属性来进行体现，通过权重算法提取出的算法公式，我们可以将邮件的价值进行量化。乐观的行为将会加分，悲观的行为将会减分，再通过公式权重的调整，可以达到一种平衡的邮件价值体现。根据这个价值我们就可以对邮件进行分析，进行检测，进行分类，那些是垃圾邮件，那些是重要的商务邮件，一目了然。

(4)基于图论对邮件网络进行分析，形成数据可视化的图表

通过对邮件网络的关系分析，系统可以发掘多种模型并建模，例如在一天的邮件收发周期内，其邮件沟通网络存在的相对活跃时间段在于下班时间前两小时和相对沉寂时间段在于中午，帮助决策在哪个时间段需要密切关注邮箱或者在何时发送邮件给对方比较适宜。这样一个数据的关系分析能够帮助用户很好的决策他的邮箱行为，并分辨异常。

(5)将重要度算法和邮件网络分析结果良好的映射到交互设计中

将基于重要度算法和邮件网络分析的结果数据通过与现代发达前端技术中的可视化框架相结合，所形成的企业邮箱管理系统能够解决电子邮件中部分设计社会关系的安全问题和可操作性问题，具有可视化和辅助目标用户决策的特点，满足需要大量处理邮件或联系人关系的用户对邮件管理系统的期望和需求。

谢辞

时间不知不觉从弹指间溜过，眼看四年的大学生活已经接近尾声，回首过去四年的大学时光，酸、甜、苦、辣，大学的点点滴滴都让我难以忘怀。大学生活将是我一生的回忆，也将是我一生的财富，在这里，我学会了怎么立足，怎么与人和平相处，更重要的是让我学会了在社会上立足的本领。学校的呵护、老师的教导、同学的帮助，让我内心充满无限的感动和感激。

论文结束的时候，我心里异常激动。首先应当感激的是我的导师，在我学习、工作和生活方面给予了悉心指导和无微不至的关怀，特别是在论文选题、课题研究及论文写作方面，倾注了很多心血。然后我要感谢我的朋友们，和你们度过了我一生中最值得纪念的时光，正因为有你们的支持和帮助，才得以顺利地完成毕业设计的工作。

最后感谢所有关心和支持我的人，祝愿他们健康、快乐，我也祝愿自己能如愿在自己热爱的计算机领域中，做出自己的努力。

参考文献

- [1]M. T. Pervez, M. Shoaib, S. Shoaib, K. Karim and S. Majid AN APPROACH TO FILTER VIP EMAILS USING DYNAMIC WEIGHT ASSIGNMENT[J]. 2011
- [2]祁巍, 李一军, 李英来, 谢伟全. 智能化企业电子邮件分类管理模型研究[J]. 中国软件学 2007
- [3]马会忠. 企业邮件系统的安全性设计[J]. 计算机与网络, 2003
- [4]严辉容, 胡小青, 西庆坤. 在大中型企业总构建安全可靠的 Internet 电子邮件系统 [J]. 四川工程职业技术学院学报 2008
- [5]Peng, Z. and J.D. Dai Multiple-Criteria Linear Programming for VIP E-Mail Behavior Analysis[J]. The Sixteenth National Conference on Artificial Intelligence 2007
- [7]M. R. And W. Zhou An Innovative Analyzer for Email Classification Based on Grey List Analysis[J]. IFIP International Conference on Network and Parallel Computing' 07 2007
- [8]郭绍忠, 段丹, 刘晓楠, 甄涛. 邮件挖掘技术在社会网络分析中的研究与应用[J]. 计算机工程与设计, 2008
- [9]孙卫琴. 精通 Struts: 基于 MVC 的 Java Web 设计与开发[M]. 北京: 电子工业出版社, 2004.
- [10]吴炜煜. 面向对象分析设计与编程[M]. 北京: 清华大学出版社, 2007.
- [11]朱少民. 软件测试方法和技术 [M]. 北京: 清华大学出版社, 2005.
- [12]刘中兵. 开发者突击: Java Web 主流框架整合开发: J2EE+Struts+Hibernate[M]. 北京: 电子工业出版社, 2008.
- [13]孙鑫. Struts2 深入详解[M]. 北京: 电子工业出版社, 2008.
- [14]李刚. Struts2 权威指南—基于 WebWork 核心的 MVC 开发[M]. 北京: 电子工业出版社, 2007.
- [15]洪丹丹. 基于 J2EE 轻量级框架的安全 Web 架构研究与应用[D]. 成都: 电子科技大学, 2009.
- [16]蒋军强. MySQL 数据库中乱码成因及解决办法探究[J]. 科技信息 (科学教研), 2007

- [17] 耿伟, 刘振海, 孙磊. Struts2 框架技术的研究与分析[J], 电脑知识与技术, 2008
- [18] John R. Cameron. JSP and JSD: The Jackson approach to software development [M]. IEEE Computer Society Press, 1989.
- [19] Bruce Eckel. Thinking in Java[M]. Beijing: China Machine Press, 2004.