

本科生毕业设计

|  |
| --- |
| 基于QT的网络数据包分析软件 |

|  |  |
| --- | --- |
| 院 系 | **计算机科学与技术** |
| 专业班级 | **计科201408班** |
| 姓 名 | **陈延良** |
| 学 号 | **U201414782** |
| 指导教师 | **周正勇** |
| 2018年06月10日 | |

**学位论文原创性声明**

本人郑重声明：所呈交的论文是本人在导师的指导下独立进行研究所取得的研究成果。除了文中特别加以标注引用的内容外，本论文不包括任何其他个人或集体已经发表或撰写的成果作品。本人完全意识到本声明的法律后果由本人承担。

作者签名： 年 月 日

**学位论文版权使用授权书**

本学位论文作者完全了解学校有关保障、使用学位论文的规定，同意学校保留并向有关学位论文管理部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权省级优秀学士论文评选机构将本学位论文的全部或部分内容编入有关数据进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于 1、保密囗，在 年解密后适用本授权书

2、不保密囗 。

（请在以上相应方框内打“√”）

作者签名： 年 月 日

导师签名： 年 月 日

摘 要

当今随着信息化发展，网络安全问题日渐突出，网络攻击日益泛滥，因此网络信息安全显得尤为重要。网络主机必须有足够的安全措施，否则网络的价值就会贬值。网络数据包分析，已不是一个新鲜课题，通过对实时网络流量的采样分析，从而能够在一定程度上帮助网络管理人员及时地发现和排除网络异常，提高网络可靠性的监控和保障能力。本设计为开发一款基于Qt的网络数据包分析软件，能够读取pcap文件，将其中的网络数据包内容显示出来，并进行统计分析，以表格的形式展示给用户。本次设计借助了libpcap函数库，在Qt的框架下进行编程。设计最终实现了对TCP、UDP、ICMP、ARP、HTTP、DNS、DHCP等协议的支持。

网络数据包分析软件在以后日益复杂的网络中还将持续发挥着其不可替代的作用，也将随着网络的发展不断完善，能够更加精确地分析并定义数据包，让网络管理员能够轻易地区别哪些为正常数据包，哪些为异常数据包，从而能够更加准确的地调控网络。

**关键词**：网络数据包分析；网络协议解析；pcap文件格式；Qt界面开发；libpcap函数库

Abstract

Nowadays, with the development of information technology, network security is becoming more and more prominent, and network attacks are becoming more and more widespread. Therefore, network information security is of great importance. The network host must have adequate security measures, otherwise the value of the network will depreciate. Analysis of network packets, is not a fresh topic, by analyzing the sampling of real-time network traffic, to some extent, so that they can help network manager timely discovery and eliminate abnormal, improve the reliability of the network monitoring and support capability. This graduation design is a network packet analysis software based on Qt framework. Design can read pcap file, and display the content of network packets. The data packets are then statistically analyzed and presented to users in the form of charts. This design is designed by using the libpcap function library, which is programed under the framework of Qt. This design finally realizes the support for TCP, UDP, ICMP, ARP, HTTP, DNS, DHCP and other protocol.

Network packet analysis software in the future in the increasingly complex network will continue to play its irreplaceable role, will also continuously improve with the development of the network, can more accurately analyze and define packet, let the network administrators can easily distinguish which is normal packets, which is abnormal packets, which can more accurately the regulation network.

**Keywords:** Network packet analysis, Network protocol, Pcap file format, Qt interface development, libpcap library function

目 录

[摘 要 I](#_Toc513803305)

[Abstract II](#_Toc513803306)

[1 绪 论 1](#_Toc513803307)

[1.1 课题研究的背景和意义 1](#_Toc513803308)

[1.2 国内外研究现状 2](#_Toc513803309)

[1.3 课题要求 3](#_Toc513803310)

[1.4 研究方案 3](#_Toc513803311)

[1.5 论文结构 4](#_Toc513803312)

[2 网络数据包介绍 5](#_Toc513803313)

[2.1 pcap文件类型 5](#_Toc513803314)

[2.2 网络数据包 7](#_Toc513803315)

[2.3 本章小结 21](#_Toc513803316)

[3 系统设计与实现 22](#_Toc513803317)

[3.1 系统结构 22](#_Toc513803318)

[3.2 数据结构 26](#_Toc513803319)

[3.3 主要函数 33](#_Toc513803320)

[3.4 实现细节 34](#_Toc513803321)

[3.5 本章小结 36](#_Toc513803322)

[4 性能测试与分析 37](#_Toc513803323)

[4.1 测试环境 37](#_Toc513803324)

[4.2 运行测试 37](#_Toc513803325)

[4.3 本章小结 42](#_Toc513803326)

[5 总结与展望 45](#_Toc513803327)

[致 谢 46](#_Toc513803328)

[参考文献 47](#_Toc513803329)

[附录：英文缩写 49](#_Toc513803330)

# 绪 论

## 课题研究的背景和意义

随着计算机网络的发展，全球信息化已成为社会发展的必然趋势。但是由于计算机网络有其自身的特点，其连接形式的多样性，和网络的开放性、连通性等，导致一些网络容易受到黑客、罪犯分子和一些病毒的攻击。因此，网络信息的安全性和保密性一直是一个至关重要的问题。网络信息的安全和保密对于传输敏感数据的计算机网络尤其重要，比如军事自动化指挥网络或银行。因此，网络必须有足够的安全措施，否则网络将是一个无用的，甚至危及国家安全的网络。在计算机网络世界中，存在着许多潜在的威胁，因此网络安全措施应该能够应对各种威胁。

计算机网络所面临的威胁大体可分为两种：一是对网络中信息的威胁；二是对网络中设备的威胁。影响计算机网络的因素有很多，有些因素可能是有意的，又有可能是无意的；有些可能是人为的，也可能是非人为的；有些可能是外来黑客对网络系统的资源非法占有，归结起来，针对网络安全的威胁主要有三种：

1. 人为的无意失误：如操作员安全配置不当所造成的安全漏洞，用户安全意识不强，用户口令选择的不慎，用户将自己的账号随意转借他人或与别人共享等都会对网络安全带来威胁。
2. 人为的恶意攻击：这是计算机网络所面临的最大威胁，敌手的攻击和计算机犯罪就属于这一类。此类攻击又可分为以下两种：一种是主动攻击，它以各种方式有选择地破坏信息的有效性和完整性；另一类是被动攻击，它是在不影响网络正常工作的情况下，进行截获、窃取、破译以获得重要机密信息。这两种攻击均可对计算机网络造成极大的危害，并导致机密数据的泄漏。
3. 网络软件的漏洞和“后门”：网络软件不可能是百分之百的无缺陷和无漏洞的，然而这些漏洞和缺陷恰恰是黑客进行攻击的首选目标，曾经出现过黑客攻入网络内部的事件，这些事件的大部分就是因为安全措施不完善所招致的苦果。另外软件的“后门”都是软件公司的设计编程人员为了自便而设置的，一般不为外人所知，但一旦“后门”洞开，其造成的后果不堪设想。

网络数据包协议分析程序是一种用来收集网络数据中有用数据的程序，这些数据可以是用户的帐户密码或一些商业机密数据，它是一种利用计算机的网络接口来截获预定用于其他计算机的数据包的工具。为了及早发现和制止网络上的各种攻击，我们需要对网络上的数据进行分析，找出问题并加以防范。网络数据包协议分析程序的存在对于系统管理员是至关重要的，系统管理员通过网络数据包协议分析程序可以被诊断为大量的不可见的模糊问题，这些问题涉及两个或两个以上的计算机之间的通信异常，有些甚至涉及各种协议，借助网络数据包协议分析系统管理员可以方便的确定有多少流量属于哪个网络协议，占主要的通信的协议是哪一个，大多数流量的目的地是哪一个主机，报文信息发送占用时间多少，或一系列消息传输间隔时间是多少。这些都会为管理员判断网络问题，管理区域网路提供一些非常有价值的信息。

## 国内外研究现状

目前国内外已经有相当多的成熟且功能强大的网络数据包协议分析软件。比较著名的网络数据包分析软件有：wireshark、TcpDump、EtherPeek等。

Wireshark： Wireshark是一个开源的网络分析系统，目前是Linux和Windows平台上最好用的开源网络协议分析器。Wireshark最初是由Gerald Combs开发的，然后由一个松散的Wireshark团队维护。它目前提供的强大的协议分析完全可以与商业网络分析系统相媲美。自1998年发布第一个0.2版本以来，大量志愿者加入了Wireshark的新协议解析器，Wireshark现在已经支持了500多个协议。很难想象有这么多人开发适合系统的代码；而且很容易在系统中添加一个新的协议解析器，一个不知道系统结构的新手可以根据接口的集合开发自己的协议。所有这些都是由于Wireshark的良好设计结构。事实上，由于网络上各种各样的协议，各种各样的新协议层出不穷，一个好的协议分析器必须具有良好的可伸缩性和结构。为了满足网络开发的需要，不断增加新的协议解析器。

TcpDump： 顾名思义，TcpDump可以完全拦截网络中传输的数据包头，提供对网络层、协议、主机、网络或端口过滤器的支持的分析，并提供逻辑语句，例如and、or、not等逻辑语句，帮助你过滤掉一些无用的信息。用简单的词来定义TcpDump是：dump the trafiic on a network，根据用户的定义网络数据包截获数据包分析工具，TcpDump以其强大的功能、灵活的拦截策略，成为每一个高级系统管理员分析网络，排查问题的一个必备基本工具之一。

EtherPeek： EtherPeek开始作为一个数据包检测的网络包分析工具软件，经过多年的发展已经成为一个真正的网络管理工具并具有一些新的特性，如网络监视和分析。由美国联邦调查局用于追踪逃犯，贩毒，电脑黑客，和一些人被怀疑为外国间谍的人。是一种直观、强大的以太网网络和协议分析器。支持Macintosh和Windows平台。EtherPeek简化了在多个平台上查找和修复复杂网络任务。EtherPeek使用了非常容易使用的行业标准，提供了解码、筛选和诊断网络的能力。EtherPeek以其友好的图形界面而闻名，它提供了非常详细和多样化的网络使用信息、网络节点会话和数据包内容。使用EtherPeek在有问题的LAN中执行自定义诊断测试，监视网络通信和事件，跟踪非法网络活动，测试和调试网络硬件和软件。

## 课题要求

使用libpcap库+Qt编程设计一个能够解析pcap文件类型的网络数据包分析软件，能够显示出数据包内容并对其进行分析统计。具体要求如下：

1. 能够读取pcapa文件，并能将其中数据包解析出来。
2. 能够支持几种常见的协议类型，包括网络层、运输层和应用层。
3. 具有数据包内容查看功能，能够将文件中所包含的数据包展示出来。
4. 具有统计分析功能，能够将文件中不同类型的数据包分类，并解析。
5. 具有良好的用户界面，便于用户直观的使用网络数据包分析软件。

## 研究方案

在本次课题的研究中，网络数据包分析软件实现对pcap文件格式类型的网络数据包解析和统计功能，完成一些常用协议的识别和分析：实现TCP/IP协议簇几个基本协议的分析——ARP、TCP、UDP、ICMP以及一些应用层的常用协议如DHCP、DNS、HTTP等。使用libpcap函数库来进行pcap类型文件的读取操作，然后使用Qt的框架进行编程。首先实现一个软件初始界面，提供对待解析文件的选取功能，然后打开文件进入分析界面，首先分析界面中对整个pcap文件数据包进行一个概览，然后可以对感兴趣的单个数据包进行详细解析，可以显示出数据包二进制的数据，同时包括链路层、网络层、运输层、应用层的各层控制信息，以及数据包的应用层详细信息显示。另外统计功能通过对协议、IP地址和端口号的筛选功能来实现。

## 论文结构

本文共分为五个章节：

第一章，首先介绍了一下网络数据包分析的研究背景，研究现状，介绍了当下比较主流的几个网络分析软件，然后根据需求设计了研究方案，以及要最终达到的预期目标。

第二章，根据课题需要，我们首先介绍了一下网络数据包的分类，以及不同类型数据包的结构，以及协议的一些详细实现细节。

第三章，在整个系统的构建介绍完成之后，开始介绍一些系统操作的设计与实现，比如如何对协议进行判断，如何对协议中不同长度的数据包进行解析，如何对文件中的不同类型数据包进行统计等。

第四章，我们将会进行实验测试结果的展示，主要包括软件的完整性和正确性，对预期支持协议类型的测试，对pcap数据包的实际解析测试，以及用户界面的良好性测试。

第五章，我们将对整个项目进行总结，总结其已经实现的所有功能和达到的性能以及展望项目未来的发展前景和研究方向。

# 网络数据包介绍

在这一章本文将介绍各种不同网络数据包的结构以及保存网络数据包的pcap文件类型。

## pcap文件类型

pcap文件格式是一种常用的网络数据包存储格式，或者说是一种数据流格式，包括wireshark、tcpdump在内的很多软件可以直接把网络数据流转换为这种格式。下面对pcap格式进行详细解析。

Pcap格式结构如下：

文件头 24字节

数据包头 + 数据包（数据包头为16字节，后面紧跟数据报）

数据包头 + 数据包

… …

下面对各个部分进行详细介绍：

* Global Header（pcap文件头 共24byte）

作为pcap文件头，在一个pcap文件中，只有一个global header，它包含了文件的读取规则、版本号、时间戳精度、最大存储长度限制等内容；

* + Magic（4byte）：作为文件开始的标记，并用来规定文件自己的字节读取顺序。0xa1b2c3d4用来表明按原本的顺序读取，0xd4c3b2a1表示后面的数据字节都要交换顺序读取。考虑到计算机内存的存储结构，一般该字段为0xd4c3b2a1，即所有字节都需要交换顺序读取。
  + Major（2byte）：当前文件主要的版本号，通常为0x0200
  + Minor（2byte）：当前文件次要的版本号，通常为0x0400
  + ThisZone（4byte）：当地的标准时间，如果是GMT则该字段全是0，通常都直接写0000 0000
  + SigFigs（4byte）：时间戳的精度，通常设置为全零即可
  + SnapLen（4byte）：最大的存储长度，如果想抓整个包，设置为0xffff0000，但通常来说0xff7f0000就行了
  + LinkType（4byte）：链路类型，常用类型有以下几种：

0 BSD loopback devices,except for later OpenBSD

1 Ethernet,and Linux loopback devices

6 802.5 Token Ring

7 ARCnet

8 SLIP

9 PPP

10 FDDI

100 LLC/SNAP-encapsulated ATM

101 “raw IP”,with no link

102 BSD/OS SLIP

103 BSD/OS PPP

104 Cisco HDLC

105 802.11

108 later OpenBSD loopback devices(with the AF\_value in network byte order)

113 special Linux “cooker” capture

114 LocalTalk

* Packet Header（数据包头 共16byte）

Packet Header一般都有很多，每个packet header后面会紧跟着一串数据内容Packet data，packet header定义了packet data数据的长度、时间戳等信息。

* + Timestamp：被捕获时间戳高位，单位是seconds
  + Timestamp：被捕获时间戳地位，单位是microseconds
  + Caplen：当前数据区的长度，即当前抓取到的数据帧长度，不包括packet header头部本身的长度，单位是字节，由此可以得到下一个数据帧的位置
  + Len：离线数据长度：网络中实际数据帧的长度，一般不大于caplen，多数情况下和caplen数值相等。
* Pcaket Data（共Caplen Byte）

包头之后就是数据包的数据了，数据长度的字节数就是Caplen个，在这之后便是一个新的Packet Header，其内容定义一个新的Packet Data属性，后面又接一个新的Packet Data，如此循环。

也就是说pcap文件中没有规定捕获packet数据包之间的间隔字符，下一组的数据包起始是由上一个包的packet header中caplen决定的。以下是对pcap文件的详细结构图（以一个TCP类型的普通报文为例）：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | | 2 | 3 | 4 | 5 | 6 | | | 7 | | 8 | 9 | A | | B | C | D | | E | F |
| **Pcap文件头（24字节）** | | | | | | | | | | | | | | | | | | | | | |
| **Pcap文件头** | | | | | | | | | | | **pcap数据包头（16字节）** | | | | | | | | | | |
| **Pcap数据包头** | | | | | | | | | | | **MAC以太网帧头（14字节）** | | | | | | | | | | |
| **MAC头** | | | | | | | | | **IP数据报头（一般20字节）** | | | | | | | | | | | | |
| **IP头** | | | | | | | | | | | | | | | **TCP头** | | | | | | |
| **TCP报文段头（一般20字节）** | | | | | | | | | | | | | | | | | | | **数据域** | | |
| **数据域** | | | | | | | | | | | | | | | | | | | | | |
| **Pcap数据包头（16字节** | | | | | | | | | | | | | | | | | | | | | |
| **MAC以太网帧头（14字节）** | | | | | | | | | | | | | | | | | | | **IP头** | | |
| **IP数据报头（一般20字节）** | | | | | | | | | | | | | | | | | | | | | |
| **IP头** | | **TCP报文段头（一般20字节）** | | | | | | | | | | | | | | | | | | | |
| **TCP头** | | | | | | | | **数据域** | | | | | | | | | | | | | |
| **数据域** | | | | | | | | | | | | | | | | | | | | | |
| **……** | | | | | | | | | | | | | | | | | | | | | |

图2-1 pcap文件格式

## 网络数据包

### 以太网帧

在链路类型中，通常为以太网类型，即LinkType = 1;

下面介绍一些以太网链路层的帧头，即MAC头：

* 目的MAC地址（共6byte）

目的MAC地址，表明数据包在链路层当前起始的源地址。以太网中MAC地址长度为6字节，每一个物理设备都有全球唯一的MAC地址。

* 源MAC地址（共6byte）

源MAC地址，表明数据包下一个链路层传输的目标地址。

* 上层协议类型（16bit）
  + IPv4协议类型，代码为0x0800
  + ARP协议类型，代码为0x0806
  + IPv6协议类型，代码为0x86DD
  + PPPoE协议类型，代码为0x8864
  + 802.1Q tag协议类型，代码为0x8100

### IP数据报

IP（Internet Protocol 因特网互联协议），是TCP／IP协议族中最为核心的协议。它提供不可靠，无连接的服务，它为运输层提供服务，而需要差错控制的功能也只能由其上层实现。在以太网中，以太网帧中承载着IP数据报在链路层中传送。而所有的TCP，UDP，ICMP，IGMP等上层数据都被封装在IP数据报中传送。

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | | 3 | 4 | 5 | 6 | 7 | 8 | 9 | a | b | c | d | e | f |
| MAC帧头（14字节） | | | | | | | | | | | | | | |  | |
| IP报头（20字节） | | | | | | | | | | | | | | | | |
|  | | | IP数据报 | | | | | | | | | | | | | |
| …… | | | | | | | | | | | | | | | | |

图2-2 IP数据报

IP数据报头一般为20字节，如下：

* 版本信息（4 bit）；

用来表明IP协议当前使用的版本号，作为最常使用的IPv4，其值为0100。

* 头长度（4 bit）；

是头部占32比特的数字，包括可选项。普通IP数据报，没有任何选项的，该字段的值为5，即160比特=20字节。此字段最大值为60字节。

* 服务类型（8 bit）；

其中前3比特为优先权字段（Precedence，现已被忽略）。第8比特保留未用。第4至7比特分别代表延迟，吞吐量，可靠性和花费。他们取值为1时分别代表要求最小延时，最大吞吐量，最高可靠性和最小费用。这4比特的服务类型中只能置其中1比特为1。可以全为0。

* 数据包长度（16 bit）；

指明整个数据报的长度（以字节为单位）。最大长度为65535字节。

* 数据包标识（16 bit）；

用来唯一标识主机发送的每一份数据报。通常每发一份报文，他的值会加1。

* 分片使用，标志（3 bit）+片偏移（13 bit）；

前3比特标志一份数据是否要求分段。后13比特指明该段偏移距原始数据报开始的位置。

* 存活时间（8 bit）；

用来设置数据报最多可以经过的路由器数。由发送数据的源主机设置，通常为32，64，128等，每经过一个路由器，其值减1，直到0时该数据报会被丢弃。

* 协议类型（8 bit）；

指明IP层所封装的上层协议类型，如ICMP（1），IGMP（2），TCP（6），UDP（17）等。

* 首部校验和（16 bit）；

内容是根据IP头计算得到的校验和码。计算方法是，对头部中两个16比特进行二进制反码求和。（和ICMP，IGMP，TCP，UDP不同，IP不对头部后的数据进行校验）。

* 源IP地址（32 bit）；

指明数据包在网络层中源IP地址，作为IPv4通常为4字节。

* 目的IP地址（32 bit）；

指明数据包在网络层中的目的IP地址，作为IPv4通常为4字节。

* 选项（如果有的话，32 bit）；

用来定义一些任选项：如记录路径，时间戳等，这些选项很少被使用，同时并不是所有主机和路由器都支持这些选项。可选字段的长度必须是32比特的整数倍。如果不足必须填充0以达到此长度要求。

### ARP报文

ARP（Address Resolution Protocol 地址解析协议），是根据IP地址获取物理地址的一个TCP／IP协议。在网络中ARP请求报文通常是以广播的形式发送到局域网中，请求一个IP地址对应的物理地址，其数据包格式如下

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | | 3 | | 4 | | 5 | 6 | 7 | 8 | 9 | a | | b | c | | d | e | f |
| MAC帧头（14字节） | | | | | | | | | | | | | | | | | | | HType | |
| PType | | | Hl | | Pl | | Op | | 发送端MAC地址 | | | | | | | | 发送端IP地址 | | | |
| 目的MAC地址 | | | | | | | | | 目的IP地址 | | | | |

图2-3 ARP报文格式

前14字节为以太网首部，对于以太网首部来说，如果是请求报文的话，以太网目的地址为ff:ff:ff:ff:ff:ff；即广播报，帧类型的话ARP对应的值为0x0806。报文部分的各字段含义如下：

* 硬件类型（HType）

硬件地址不只以太网一种，如果是以太网类型，此时值为1

* 协议类型（PType）

表示要映射的协议地址的类型，要对IPv4地址进行映射，此时值为0x0800

* 硬件地址长度（Hl）

以太网类型地址长度对应6字节

* 协议地址长度（Pl）

IPv4地址类型对应长度4字节

* 操作类型（OP）

操作类型字段，值为1，表示进行ARP请求；值为2，表示进行ARP应答；值为3，表示进行RARP请求；值为4，表示进行RARP应答。

* 发送端以太网地址

发送ARP请求或应答的硬件地址，这里是以太网地址，和MAC头的源地址相同。

* 发送端IP地址

发送ARP请求或应答的IP地址

* 目的以太网地址

目的以太网地址，请求时通常为广播，即为00:00:00:00:00:00。

* 目的IP地址

目的IP地址，请求中为请求MAC的IP地址，应答时为请求的源IP地址

### TCP报文段

TCP（Transmission Control Protocol 传输控制协议），是一种面向连接的、可靠的、基于字节流的传输层通讯协议。由于IP仅提供不可靠的包交换，而通常应用层需要一种可靠的像管道一样的连接，因此TCP的作用就是如此，位于IP层和应用层之间提供了可靠的数据传输。报文段头一般为20字节，结构如下：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | | 3 | 4 | 5 | 6 | 7 | 8 | 9 | a | b | c | d | e | f |
| MAC帧头（14字节） | | | | | | | | | | | | | | |  | |
| IP数据报头（20字节） | | | | | | | | | | | | | | | | |
|  | | | TCP报文段头 | | | | | | | | | | | | | |
| …… | | | | | | | | | | | | | | | | |

图2-4 TCP报文格式

* 源端口号（16 bit）

源端口号，在一个操作系统中，通常有多个进程，源端口号标识了发送主机的进程。

* 目的端口号（16 bit）

目的端口号标识了接收方主机的进程。

* 序列号（32 bit）

发送序号（Sequence Number）标识从TCP源端向TCP目的端发送的数据字节流，表示在这个报文段中的第一个数据字节的顺序号。序号时32bit的无符号数。

* 确认号（32 bit）

确认序号（Acknowledgment Number）是包含发送确认的一端所期望收到的下一个顺序号，通常是上次已成功收到的数据字节顺序号加1。只有ACK标志为1时确认字段有效。

* TCP头长度（4 bit）

TCP头部的长度，用来表明TCP首部中32比特字的数目数，通常为5。

* 保留部分（6 bit）

保留字段，目前没有使用，都为0。

* 标志位（6 bit）

共有6个标志位比特，他们中可以多个同时置1。

URG 紧急指针有效

ACK 确认序号有效

PSH 指示接收方应该尽快将此报文段交给应用层而不用等缓冲区满

RST 一般表示断开一个连接

SYN 同步序号用来发起一个连接

FIN 发送端完成发送任务，即断开连接

* 窗口大小（16 bit）

窗口大小，表示最多能接受的字节数。

* 校验和（16 bit）

检验和字段检验的范围包括首部和数据这两部分，强制性字段，一定由发送端计算和存储，并由接受端验证。

* 紧急指针（16 bit）

在URG标志位1时生效，紧急指针是一个正偏移量，和序号字段中的值相加表示紧急数据最后一个字节的序号，TCP的紧急方式是发送端向另一端发送紧急数据的一种方式。

* 选项（如果有）

选项的长度可变，TCP最初值规定一种选项，即最大报文段长度MSS。（Maximum Segment Size）

### UDP报文段

UDP（User Datagram Protocol 用户数据报协议），相对于TCP协议 ，UDP是无连接的，提供不可靠的信息传送服务，是一种尽力而为的服务。报文段头为8字节，如下：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | | 3 | 4 | 5 | 6 | 7 | 8 | 9 | a | b | c | d | e | f |
| MAC帧头（14字节） | | | | | | | | | | | | | | |  | |
| IP数据报头（20字节） | | | | | | | | | | | | | | | | |
|  | | | UDP报文段头 | | | | | | | | | | | | | |
| …… | | | | | | | | | | | | | | | | |

图2-5 UDP报文格式

* 源端口号（16 bit）

源端口号标识了发送主机的进程。

* 目的端口号（16 bit）

目的端口号标识了接收方主机的进程。

* 数据包长度（16 bit）

标识数据包的数据长度，单位为字节

* 校验和（16 bit）

校验和，包含UDP头和数据两部分

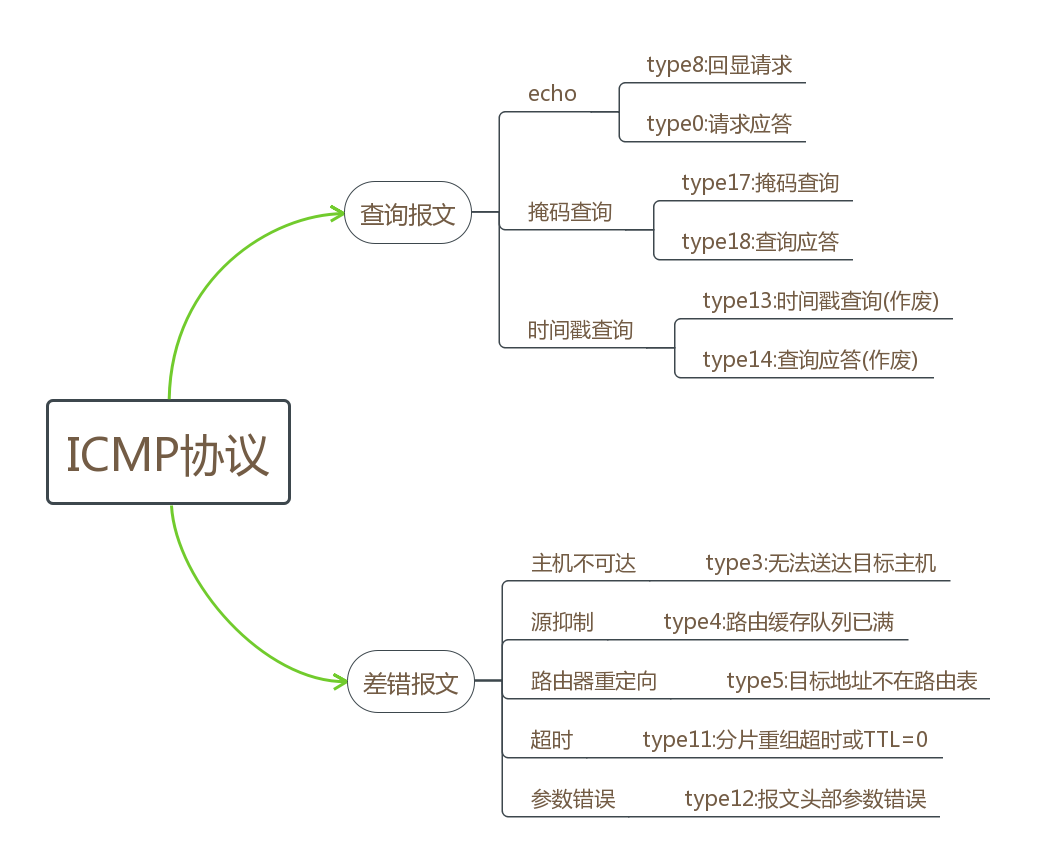
### ICMP报文

ICMP（Internet Control Message Protocol 网络控制信息协议），协议是IP协议的补充，用于IP层的差错报告，拥塞控制，路径控制以及路由器或主机信息的获取。ICMP与IP协议位于同一层（IP层），而ICMP报文是封装在IP数据包的数据部分进行传输的，不同于TCP和UDP，他并不传输用户数据，但对于用户数据的传递起者至关重要的作用。如下位ICMP的报文结构

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | | 3 | 4 | 5 | 6 | 7 | 8 | 9 | a | | b | c | d | e | f |
| MAC帧头（14字节） | | | | | | | | | | | | | | | |  | |
| IP数据报头（20字节） | | | | | | | | | | | | | | | | | |
|  | | | ICMP固定报文头（8字节） | | | | | | | | | ICMP数据部分 | | | | | |
| …… | | | | | | | | | | | | | | | | | |

图2-6 ICMP报文格式

ICMP分很多类，总的来看分为两类1）通知错误原因的差错报文 2）请求应答的查询报文，其大致分类如下：

图2-7 ICMP报文类型

常见类型：

* 回显请求（0）：ping工具通过发送ICMP回显消息检查特定节点的IPv4连接已排查网络问题。
* 回显应答（8）：节点发送回显答复消息响应ICMP回显消息。
* 重定向（5）：路由器发送“重定向”消息，告诉发送主机到目标IPv4地址更好的路由。
* 源抑制（4）：路由器发送“源结束”消息，告诉发送主机以较低的频率发送数据报，以免因为路由器拥塞而导致数据报被丢弃。
* 超时（11）：当超过IP生存期时和分段IP数据报没有在期限内重新组合都会发送这种消息。
* 无法达到目标（3）：路由器和目标主机通知发送主机数据无法传送。

由于不同报文类型通常会有不同的格式，以下主要介绍常见类型中两种通用格式的ICMP报文格式。ICMP报文格式及类型，前8字节固定，通用格式。

|  |  |  |
| --- | --- | --- |
| 类型 | 代码 | 校验和 |
| （这4字节取决于ICMP报文类型） | | |
| ICMP数据部分 | | |

图2-8 ICMP通用格式

当发送一份差错报文时，报文始终包含IP的首部和产生ICMP差错报文的IP数据报的前8字节，这样接受ICMP差错报文的模块就会吧它与某个特定协议和用户进程联系起来（分别是IP头中协议字段和IP数据报前8字节中的端口号来判断）。对于类型目的不可达（3），超时（11），参数错误报文（12），源抑制（4），重定向（5）：

|  |  |  |
| --- | --- | --- |
| 类型 | 代码 | 校验和 |
| 未用（全为0） | | |
| IP数据报头（包括选项）+IP数据报原始数据字段前8字节 | | |

图2-8 ICMP差错类型格式

对于类型Echo(8、10)、信息请求或信息响应：

|  |  |  |
| --- | --- | --- |
| 类型 | 代码 | 校验和 |
| 标识符 | | 序列号 |
| ICMP数据部分 | | |

图2-9 ICMP查询类型格式

### DNS报文

DNS（DomainNameSystem 域名系统），是用于命名组织到域层次结构中的计算机和网络服务，其报文一般包含在UDP报文段中。

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | | 3 | 4 | 5 | 6 | 7 | 8 | 9 | a | | b | c | d | e | f |
| MAC帧头（14字节） | | | | | | | | | | | | | | | |  | |
| IP数据报头（20字节） | | | | | | | | | | | | | | | | | |
|  | | | UDP报文段头（8字节） | | | | | | | | | DNS报文 | | | | | |
| …… | | | | | | | | | | | | | | | | | |

图2-10 DNS报文格式

DNS报文格式及类型：

|  |  |
| --- | --- |
| Transaction ID（会话标识） | Flags(标识) |
| Questions(问题数) | Answer RRs(回答 资源记录数) |
| Authority RRs(授权 资源记录数) | Additional RRs(附加资源记录数) |
| Name(查询名，长度不固定) | |
| Type(查询类型) | Class（查询类） |
| Name(查询名，长度不固定) | |
| Type(查询类型) | Class（查询类） |
| Time to live(生存时间) | |
| Data lenth(资源数据长度) | Data |
| Data（资源数据，长度不固定） | |
| Name(查询名，长度不固定) | |
| Type(查询类型) | Class（查询类） |
| Time to live(生存时间) | |
| Data lenth(资源数据长度) | Data |
| Data（资源数据，长度不固定） | |
| Name(查询名，长度不固定) | |
| Type(查询类型) | Class（查询类） |
| Time to live(生存时间) | |
| Data lenth(资源数据长度) | Data |
| Data（资源数据，长度不固定） | |

图2-11 DNS详细格式

* 头部(蓝色区域)
  + 回话标识（2字节）：DNS报文的ID标识，对于请求报文和其对应的应答报文，这个字段是相同的，通过它可以区分DNS应答报文是哪个请求的响应。
  + 标志（2字节）：

QR：查询/响应标志，0为查询，1为响应

Opcode：0表示标准查询，1表示反向查询，2表服务器状态请求。

AA：表示授权回答

TC：表示可截断的

RD：表示期望递归

RA：表示可用递归

Rcode：表返回码，0表没差错，3表名字差错，2表服务器错误。

* + 数量字段

Questions、Answer RRs、Authority RRs、Additional RRs各自表示后面的四个区域的数目。Questions表示查询问题区域节的数量，Answer RRs表示回答区域的数量，Authority RRs表示授权区域的数量，Additional RRs表示附加区域的数量。

* 查询问题区域（绿色区域）
  + 查询名：长度不固定，且不使用填充字节，一般该字段表示的就是需要查询的域名，一般格式如下：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | h | u | b | 4 | h | u | s | t | 3 | e | d | u | 2 | c | n | 0 |

数字表明后面的字节数，hub前为数字3,hust前为4,edu前为3，cn前为2，最后结尾必须是0,所以域名为hub.hust.edu.cn。

* + 查询类型：

表2-1 DNS查询类型

|  |  |  |
| --- | --- | --- |
| 类型 | 助记符 | 说明 |
| 1 | A | 由域名获得IPv4地址 |
| 2 | NS | 查询域名服务器 |
| 5 | CNAME | 查询规范名称 |
| 6 | SOA | 开始授权 |
| 11 | WKS | 熟知服务 |
| 12 | PTR | 把IP地址转换为域名 |
| 13 | HINFO | 主机信息 |
| 15 | MX | 邮件交换 |
| 28 | AAAA | 由域名获得IPv6地址 |
| 252 | AXFR | 传送整个区的请求 |
| 255 | ANY | 对所有记录的请求 |

* + 查询类

通常为1，表明是internet数据

* 回答区域（紫色区域）
  + 域名

格式与Queries区域查询名字字段一样，有一点不同的是，当域名重复出现时，该字段使用2字节的便宜指针表示，具体格式时，前面两个最高位为11.用于识别指针，其余14位表示从DNS报文开始出计数，指出字段在报文中相应偏移字节数。典型的一个列子就是c00c(11000000 00001100)12正好时头部长度，正好指向Queries区域的查询名字字段。

* + 查询类型

同上，见表格

* + 查询类

对于internet信息，总是IN

* + 生存时间

以秒为单位，表示的是资源记录的生命周期，一般用于当地地址解析程序取出资源记录后决定保存及使用缓存数据的时间，它同时也可表示该资源的稳定程度，极为稳定的信息会被分配一个很大的指，如84600，这是一天的秒数。

* + 资源数据

该字段也是一个变长字段，表示按照查询段的要求返回的相关资源记录，例如如果查询报文想要的回应是一个IP地址，这里就对应一个IP地址。

* 授权区域（黄色区域）

同上回答区域

* 附加区域（橙色区域）

同上回答区域

### DHCP报文

DHCP（Dynamic Host Configuration Protocol 动态主机配置协议），是一个局域网的网络协议，主要负责给内部网络或网络服务供应商自动分配IP地址。通常是通过UDP协议工作，其对应端口号分别，服务器为67，用户为68。

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | | 3 | 4 | 5 | 6 | 7 | 8 | 9 | a | | b | c | d | e | f |
| MAC帧头（14字节） | | | | | | | | | | | | | | | |  | |
| IP数据报头（20字节） | | | | | | | | | | | | | | | | | |
|  | | | UDP报文段头（8字节） | | | | | | | | | DHCP报文 | | | | | |
| …… | | | | | | | | | | | | | | | | | |

图2-12 DHCP报文格式

DHCP报文详细结构如下：

|  |  |  |  |
| --- | --- | --- | --- |
| 报文类型（1B） | 硬件类型（1B） | 硬件地址长（1） | 跳数（1B） |
| 事务ID（4B） | | | |
| 秒数（2B） | | 标志位（2B） | |
| 客户端IP地址（4B） | | | |
| 你的IP地址（4B） | | | |
| 服务器IP地址（4B） | | | |
| 网关IP地址（4B） | | | |
| 客户机硬件地址（16B） | | | |
| 服务器主机名（64B） | | | |
| 启动文件名（128B） | | | |
| 选项（变长） | | | |

图2-13 DHCP报文详细格式

* 报文类型

报文类型码，只能取值1和2 ，请求包request为1，应答包reply为2

* 硬件类型

硬件类型代码，1代表千兆以太网

* 硬件地址长度

单位为字节，表明硬件地址长度，以太网对应值为6

* 跳数

DHCP跨网段使用时由reply agent使用，DHCP client发出的包此字段清零。

* 事务ID

值为DHCP client发送数据包时填充的一个随机数，用来在client和server之间表示一组request和reply。

* 秒数

由客户端填充，表示客户端开始请求分配IP或请求更新IP租期所经历的时间。

* 标志

标志字段，目前只使用左边最高位的bit，其余bit必须由client发包时清零，用来表示自己想从server或是reply agent那里得到的回复是unicase类型还是braodcast类型。

* 客户端IP地址

只有在client拥有IP，能够响应ARP请求时，并且client端处于BOUND，RENEW，或者REBINDING状态时才会被client填充

* 你的IP地址

DHCP server分配给client的IP地址。

* 客户端IP地址

DHCP流程的下一个服务器地址，由DHCP server在DHCPOFFER和DHCPACK保证返回。

* 网关IP地址

DHCP的中继IP地址，当跨网段使用DHCP时，该字段由一个relay agent填充，并且每个agent都将hops加一，通常hops值有一个限制，超过此限制，该包将会被丢弃。

* 客户端硬件地址

占十六个字节

* 服务器名称

占64字节，可选的DHCP server的Hostname,是一个以null结尾的字符串。

* 文件名

启动文件名，128字节是一个以null结尾的字符串，当主机从网络启动时会使用。

* 选项

可选参数列表，长度不固定。

### HTTP报文

HTTP（HyperText Transfer Protocol 超文本传输协议），是一种使用极为广泛的网络协议。协议设计之初就是为了提供一种发布和接受HTML页面的方法。报文通常通过TCP报文段来进行传输，其对应的端口号为80。

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | | 3 | 4 | 5 | | 6 | 7 | 8 | 9 | a | b | c | d | e | f |
| MAC帧头（14字节） | | | | | | | | | | | | | | | |  | |
| IP数据报头（20字节） | | | | | | | | | | | | | | | | | |
|  | | | TCP报文段头（一般20字节） | | | | | | | | | | | | | | |
|  | | | | | | | HTTP报文 | | | | | | | | | | |
| …… | | | | | | | | | | | | | | | | | |

图2-14 HTTP报文格式

## 本章小结

这一章主要是对网络中常见数据包类型进行详细介绍，这些协议类型也是本次网络数据包分析系统设计的基础。协议主要包括TCP/IP协议簇几个基本协议的分析——ARP、TCP、UDP、ICMP以及一些应用层的常用协议如DHCP、DNS、HTTP等。本章对各个协议有详细的协议实现细节，而在实现过程中需要对各个协议的数据结构，数据包长度以及高层协议的类型等都要有比较深的理解。

# 系统设计与实现

## 系统结构

整个网络数据包分析工具采用模块化的设计思想，原因是许多程序太长或太复杂，很难写在单一单元中。如果把代码分为较小的功能单元，这样将大大简化编程过程。模块化程序一般比单独程序容易编写、调试和修改。所以本系统采用模块化的设计思想。该网络数据包分析工具主要分三个模块：界面显示模块，数据包解析存储模块，数据包分析统计模块。接下来我们将对这三个模块的关系和各自的功能进行详细介绍。

### 界面显示模块

界面模块主要负责将数据显示给用户，并提供数据输入功能。在整个程序中，界面模块与其他两个模块密切合作。原则是尽量合理的显示数据包的有用信息。

该模块中负责以下几个功能：

* + 初始界面显示

在运行程序后的初始界面，主要负责引导用户进行待解析文件的选择，以及对

* + 总揽表显示

总揽表主要是将待解析文件的所有数据包的大致信息展示给用户，方便用户对整个文件中的数据包有一个大致的了解，以便查看。其中每个数据包的信息主要有捕获时间戳、包大小、包类型、发送方地址、接收方地址、报文的初略信息概括等。

* + 数据包二进制内容显示

当用户选取一个数据包时，数据包的二进制数据将会展示给用户，所以该模块主要是负责将二进制数据展示到一个输出表中，并给予其对应的ASCII码的数值。

* + 数据包解析数据显示

当用户选取一个数据包时，数据包的链路层、网络层、运输层和应用层的数据包头将被详细解析出来，并通过输出框展示给用户。对于一些协议，如HTTP，可以直接显示出其应用层HTTP报文，另一些协议，如DHCP，将会显示出其网络控制的类型和控制信息。

* + 文件中数据包统计

文件中的数据包类型通常有很多种，每一种的数量不一，通过对其进行不同类型的数量统计，最终将统计结果通过输出框显示给用户。

* + 筛选信息的控制

筛选信息控制主要负责三个筛选条件的控制，协议类型、IP地址和端口号，不同的协议支持筛选的类型不一，例如对于ICMP报文就不存在端口号。

最后通过Qt中的窗口机制，设计一个合理的窗口调用规则，例如，在初始界面中可以选取文件进入该文件的详细解析界面，然后关闭解析界面后还能正常返回到初始界面，供用户继续解析其他pcap文件，而不是直接结束程序。

### 数据包解析存储模块

数据包分离模块主要负责读取pcap文件后将其中的数据包解析分离出来，并保存在对应的数据结构中。

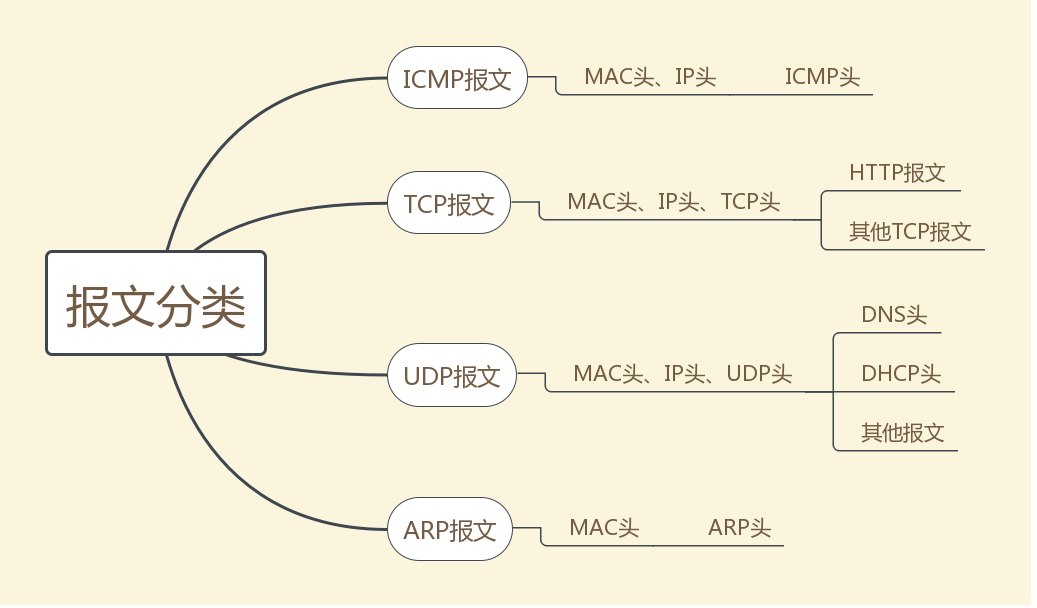


图3-1 报文分类

该模块主要是程序一开始对文件读取成功后进行调用该模块。主要是根据数据包的协议包头进行数据信息的解析，并作出适当的解释，最后根据类型存储到相应数据结构中，以便其他模块读取。

### 数据包分析统计模块

数据包分析统计模块主要负责对每一个数据包的解析以及对整个pcap文件中的数据包进行统计。本模块就是对捕获的数据包按照数据链路层（MAC）、网络层（IP，ARP/RARP）、传输层（TCP，UDP，ICMP）和应用层（HTTP，DNS，DHCP等）的层次结构自底向上进行解析，并将解析结果显示输出。

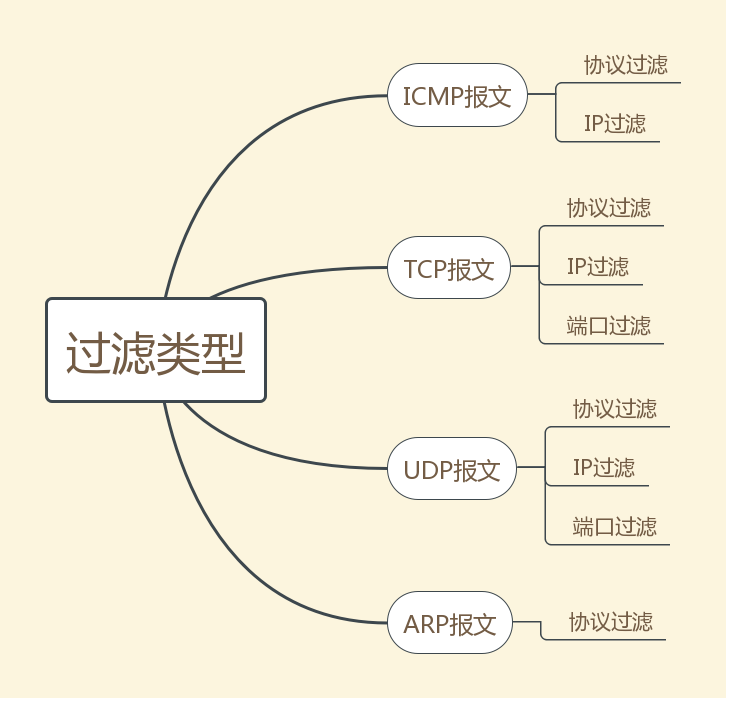


图3-2 过滤类型

在进行协议解析时要注意各个协议报头多字节整形字段从网络字节序到主机字节序的转换。对于多字节的数据（如short、int、long等），不同CPU在其内部表示一个数时会使用不同字节顺序方案，将低序字节存储在起始地址的称为小段存储，否则称为大端存储。网络中有不同类型的计算机，在网络数据的传输中要统一数字的表示方式，即统一采用大端存储，所以数据包头中表示长度或类型等的多字节数据字段应按需转换为本地主机的表示形式。可以用ntohs()或ntohl()将双字节或四字节整型由网络字节序转换为主机字节序。

数据包分析统计，主要任务是分析数据包，首先的将协议之间的层次结构划分清楚。

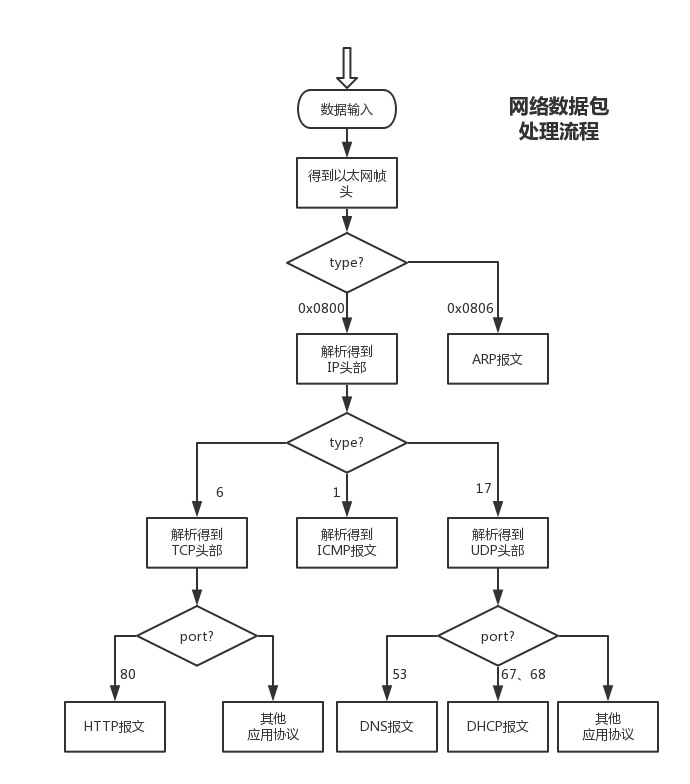


图3-3 网络数据包处理

## 数据结构

本设计中非常重要的一部分便是数据结构了，每一个数据包的解析过程离不开数据结构的支撑。整个程序中主要包括一些头部数据结构，如PCAP文件头、PCAP数据包头、MAC头，IP头，TCP头，UDP头，DHCP头，DNS头，ICMP头，ARP头等。下面逐一介绍各个头部的结构体。

* 首先是一个pcap文件中只存在一个的PCAP文件头结构体，共24字节。

typedef struct pcapFileHeader{

uchar8\_t magic[4]; ——读取顺序(4字节)

uint16\_t version\_major; ——主要版本号(2字节)

uint16\_t version\_minor; ——次要版本号(2字节)

int32\_t thiszone; ——时区修正(4字节)

uint32\_t sigfigs; ——精准时间戳(4字节)

uint32\_t snaplen; ——抓包最大长度(4字节)

uint32\_t linktype; ——链路类型(4字节)

}pcapFileHeader\_t;

* 每个数据包都包含的pcap数据包头，共16字节，主要包含数据包的一些基本信息

typedef struct ts{

uint32\_t ts\_sec; ——秒数(4字节)

uint32\_t ts\_usec; ——毫秒数(4字节)

}ts\_t;

typedef struct pcapPktHeader{

ts\_t ts; ——文件捕获时间（8字节）

uint32\_t caplen; ——数据包长度(4字节)

uint32\_t len; ——文件数据包长度(4字节)

}pcapPktHeader\_t;

* 每个以太网帧中都包含的MAC头

typedef struct EtherHeader{

uchar8\_t DstMAC[6]; ——目的MAC地址(6字节)

uchar8\_t SrcMAC[6]; ——源MAC地址(6字节)

uint16\_t FrameType; ——帧类型(2字节)

}EtherHeader\_t;

typedef struct EtherTail{

uint32\_t Checksum; ——数据帧尾校验和(4字节)

}EtherTail\_t;

* IP数据包中都包含的IP头，本文主要介绍IPv4数据包，共20字节

typedef struct IPHeader{

uchar8\_t Ver\_Hlen; ——版本+报头长度(1字节)

uchar8\_t TOS; ——服务类型(1字节)

uint16\_t TotalLen; ——总长度(2字节)

uint16\_t ID; ——标识(2字节)

uint16\_t Flag\_Segment; ——标志(3bit)+片偏移(13bit)

uchar8\_t TTL; ——生存周期(1字节)

uchar8\_t Protocol; ——协议类型(1字节)

uint16\_t Checksum; ——头部校验和(2字节)

uint32\_t SrcIP; ——源IP地址(4字节)

uint32\_t DstIP; ——目的IP地址(4字节)

}IPHeader\_t;

* TCP报文段包含有TCP头，报文头实际长度由前20字节中的报头长度HeaderLen字段决定，其一般为20字节，有时还包括长度不定TCP可选部分。

typedef struct TCPHeader{

uint16\_t SrcPort; ——源端口(2字节)

uint16\_t DstPort; ——目的端口(2字节)

uint32\_t SeqNum; ——序号(4字节)

uint32\_t AckNum; ——确认号(4字节)

uchar8\_t HeaderLen; ——报头的长度(4bit)+保留(4bit)

uchar8\_t Flags; ——保留(2bit)+标识控制信息(6bit)

uint16\_t Window; ——窗口大小(2字节)

uint16\_t Checksum; ——校验和(2字节)

uint16\_t UrgentPoint; ——紧急指针(2字节)

}TCPHeader\_t;

typedef struct TCPOptions{

char8\_t m\_ckind; ——类型（1字节）

char8\_t m\_cLength; ——长度（1字节）

char8\_t m\_cContext[]; ——内容（不定）

}TCPOption\_t;

* UDP报文段头(8字节）

typedef struct UDPHeader{

uint16\_t SrcPort; ——源端口号(2字节)

uint16\_t DstPort; ——目的端口号(2字节)

uint16\_t SegLen; ——报文段长度(2字节)

uint16\_t Checksum; ——校验和(2字节)

}UDPHeader\_t;

* ICMP固定头部部分，为4字节。每一个ICMP类型的报文都包括的字段。其标识了后面紧跟数据的类型和代码。

typedef struct ICMPHeader{

uchar8\_t ICMPType; ——类型(1字节)

uchar8\_t ICMPCode; ——代码(1字节)

uint16\_t Checksum; ——校验和(2字节)

}ICMPHeader\_t;

* ICMP报文类型0，应答包

typedef struct ICMPType0{

uint16\_t ICMPID; ——ID标识（2字节）

uint16\_t ICMPSeq; ——序列号（2字节）

uchar8\_t ICMPData[]; ——数据内容（不定）

}ICMPType0\_t;

* ICMP报文类型3，目的地不可达

typedef struct ICMPType3{

uint16\_t ICMPPmvoid; ——空指针（2字节）

uint16\_t ICMPNextmtu; ——下一个MTU（2字节）

IPHeader\_t IPH; ——IP头（20字节）

uchar8\_t IPData[8]; ——IP数据前8字节（8字节）

uchar8\_t AddData[]; ——额外数据（1字节）

}ICMPType3\_t;

* ICMP报文类型8，请求包

typedef struct ICMPType8{

uint16\_t ICMPID; ——ID标识（2字节）

uint16\_t ICMPSeq; ——序列号（2字节）

uchar8\_t ICMPData[]; ——数据内容（不定）

}ICMPType8\_t;

* ICMP报文类型11，超时

typedef struct ICMPType11{

uint32\_t ICMPVoid; ——空指针（4字节）

IPHeader\_t IPH; ——IP头（20字节）

uchar8\_t IPData[8]; ——IP数据前8字节（8字节）

uchar8\_t AddData[]; ——内容（不定）

}ICMPType11\_t;

* ARP头

typedef struct ARPHeader{

uint16\_t ARPHrd; ——硬件类型（2字节）

uint16\_t ARPPro; ——协议类型（2字节）

uchar8\_t ARPHLn; ——硬件地址长度（1字节）

uchar8\_t ARPPln; ——协议地址长度（1字节）

uint16\_t ARPOP; ——操作码OP（2字节）

}ARPHeader\_t;

* DNS报文头部(长度固定部分)

typedef struct DNSHeader{

uint16\_t ID; ——标识（2字节）

uint16\_t Flags; ——标志（2字节）

uint16\_t Qst\_num; ——问题数（2字节）

uint16\_t Rsc\_num; ——资源记录数（2字节）

uint16\_t Aut\_num; ——授权资料记录数（2字节）

uint16\_t Adt\_num; ——额外资源记录数（2字节）

uchar8\_t data[]; ——数据内容（不定）

}DNSHeader\_t;

* DHCP报文头部(长度固定部分)

typedef struct DHCPHeader{

uchar8\_t OP; ——报文类型（1字节）

uchar8\_t HrdType; ——硬件类型代码（1字节）

uchar8\_t HrdLen; ——硬件地址长度（1字节）

uchar8\_t Hops; ——跳数（1字节）

uint32\_t XID; ——事务ID 随机数（4字节）

uint16\_t Sec; ——秒数（2字节）

uint16\_t Flags; ——标志（2字节）

uint32\_t Ciaddr; ——客户端IP地址（4字节）

uint32\_t Yiaddr; ——你的IP地址（4字节）

uint32\_t Siaddr; ——服务器IP地址（4字节）

uint32\_t Giaddr; ——网关IP地址（4字节）

uchar8\_t Chaddr[16]; ——客户端硬件地址（16字节）

uchar8\_t Sname[64]; ——服务器名（64字节）

uchar8\_t File[128]; ——引导文件名（128字节）

}DHCPHeader\_t;

接下来介绍程序中存储整个pcap文件中数据的几个链表。链表结构主要时有4个独立的链表。（ICMP、TCP、UDP、ARP）

* ICMP链表，主要保存ICMP报文

typedef struct ICMP\_List{

struct ICMP\_List \*next; ——下一个ICMP结构体指针

int Seq\_num; ——序号（识别作用）

pcapPktHeader\_t pkthdr; ——pcap数据包头结构体

EtherHeader\_t etherhdr; ——以太网头结构体

IPHeader\_t iphdr; ——IP报头结构体

ICMPHeader\_t icmphdr; ——ICMP头部固定部分

uchar8\_t data[]; ——ICMP数据

}ICMP\_List\_t;

* TCP链表，主要保存TCP报文段，包括HTTP报文在内的其他TCP应用报文，全部保存在该链表中

typedef struct TCP\_List{

struct TCP\_List \*next; ——下一个TCP结构体指针

int Seq\_num; ——序号（识别作用）

pcapPktHeader\_t pkthdr; ——pcap数据包头结构体

EtherHeader\_t etherhdr; ——以太网帧头结构体

IPHeader\_t iphdr; ——IP报头结构体

TCPHeader\_t tcphdr; ——TCP头结构体

uchar8\_t data[]; ——应用层数据

}TCP\_List\_t;

* UDP链表，主要保存UDP报文段，包括DNS、DHCP报文在内的其他UDP应用层报文，全部保存在该链表中。

typedef struct UDP\_List{

struct UDP\_List \*next; ——下一个UDP结构体指针

int Seq\_num; ——序号（识别作用）

pcapPktHeader\_t pkthdr; ——pcap数据包头结构体

EtherHeader\_t etherhdr; ——以太网头结构体

IPHeader\_t iphdr; ——IP报头结构体

UDPHeader\_t udphdr; ——UDP头结构体

uchar8\_t data[]; ——应用层数据

}UDP\_List\_t;

* ARP链表，主要保存ARP 报文

typedef struct ARP\_List{

struct ARP\_List \*next; ——下一个ARP结构体指针

int Seq\_num; ——序号（识别作用）

pcapPktHeader\_t pkthdr; ——pcap数据包头结构体

EtherHeader\_t etherhdr; ——以太网头结构体

ARPHeader\_t arphdr; ——ARP头结构体

uchar8\_t data[]; ——ARP数据

}ARP\_List\_t;

* 总表（包含所有表的表头指针）

typedef struct All\_list\_hdr{

int icmp\_num; ——ICMP报文数量

int tcp\_num; ——TCP报文数量

int udp\_num; ——UDP报文数量

int arp\_num; ——ARP报文数量

ICMP\_List\_t \*icmp\_listhdr;——保存ICMP链表头指针

TCP\_List\_t \*tcp\_listhdr； ——保存TCP链表头指针

UDP\_List\_t \*udp\_listhdr; ——保存UDP链表头指针

ARP\_List\_t \*arp\_listhdr; ——保存ARP链表头指针

}All\_list\_hdr\_t;

以下为程序中数据包的存储结构图，本次设计中设计支持的数据包均存入一下四个链表当中，分别为ICMP、TCP、UDP、ARP链表。其头部设置为一个结构体成员，这样就可以在程序中随时遍历查找需要的数据包。同时，为了方便在点击事件时能有效的节省遍历链表的开销，结构体中额外设置了每个数据包的编号，该标号并不是数据包中本身具有的标号，而是在解析过程中对数据包依次解析后人为添加，这样在遍历过程中不必对每一个数据包进行比对，而是通过编号即可进行查找匹配。

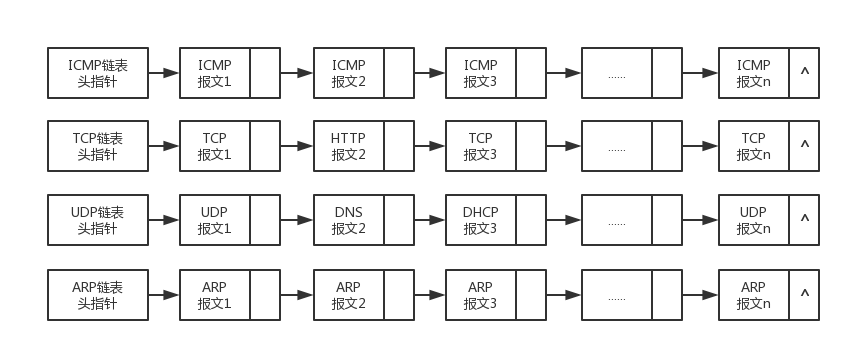


图3-4 数据存储结构

## 主要函数

本节主要介绍主要函数。首先是pcap文件读取函数——pcap\_loop();

其中数据包处理函数为一个回调函数，每次循环读取数据包后都会调用该函数。函数的处理流程如下：

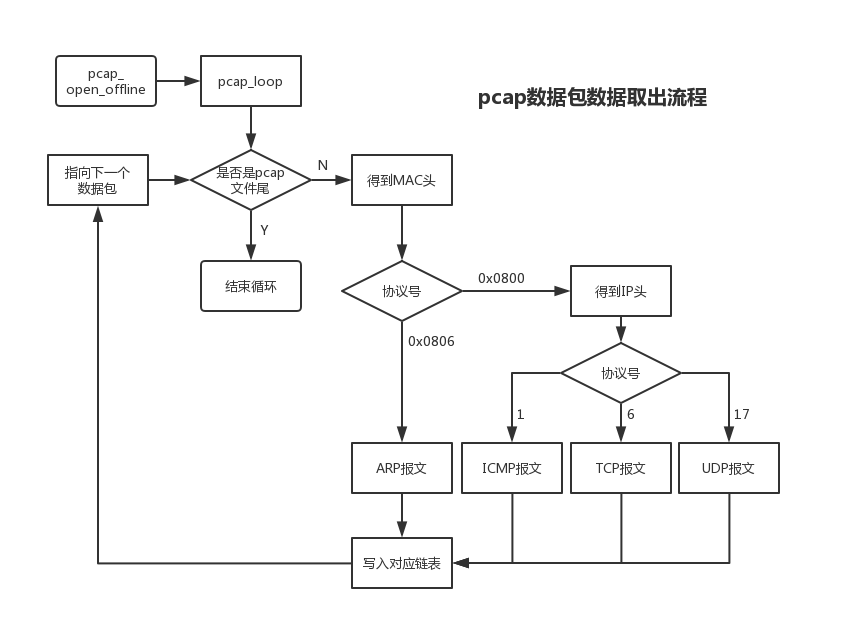


图3-5 pcap文件取出流程

再介绍Qt中的主要槽函数

首先是最主要的总揽表点击事件对应的槽函数on\_overviewTable\_clicked()

函数主要负责用户在点击总揽表中感兴趣数据包时显示其对应数据包二进制数据以及对数据包解析出的各层（包括链路层、网络层、运输层、应用层）的头部信息。

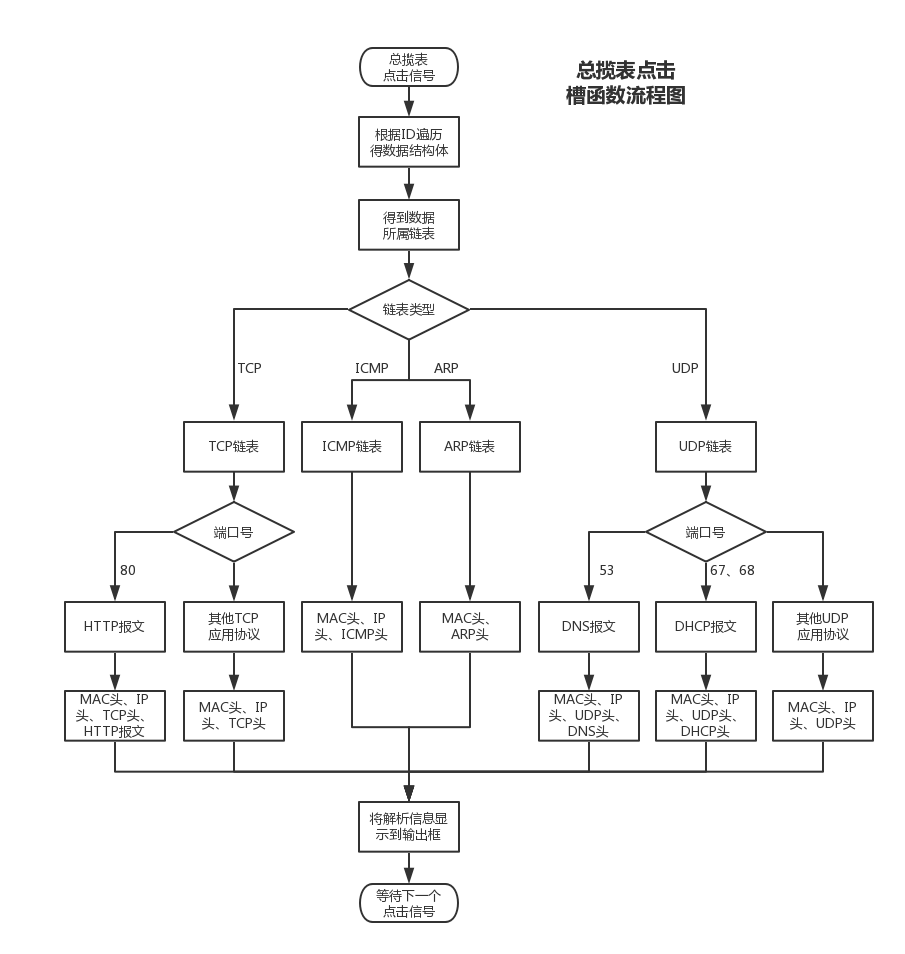


图3-6 点击事件槽函数流程

## 实现细节

实现细节，在程序中有许多具体的实现细节需要注意。比如网络字节序和主机字节序列的区别，TCP头部的长度一般情况下为20字节，但有时会有选项字段，所以导致最终不只20字节，如果没有进行判断的话会导致后面的数据解析错误。对与过滤类型，有的协议具备端口号，例如TCP、UDP等，但有些不具备端口号，例如ICMP、ARP等。最后对整个程序的流程图进行一些总结如下：

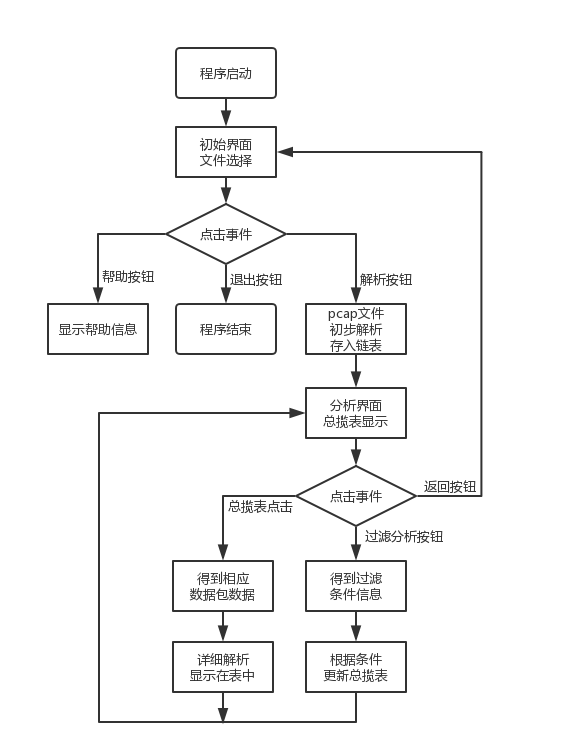


图3-7 程序运行流程

## 本章小结

本章对网络数据包分析软件各个组成部分的设计与实现进行了详细的论述。首先介绍了界面模块，数据包分离模块，数据包分析统计模块三个模块的具体设计与实现。分析了在运行过程中可能遇到的种种情况并且综合各种情况选择最有利的设计方案。然后。本章对网络数据包分析软件各个组成部分的设计与实现进行了详细的论述。首先介绍了系统的总体结构，每个模块之间的合作与分工。然后介绍了整个程序用到的数据结构，包括数组、结构体和链表等。还介绍了程序的一些主要函数实现，最后介绍了一些具体实现细节与特点，如何合理运用Qt程序设计方法来实现程序预期要达到的效果。

# 性能测试与分析

## 测试环境

测试硬件环境如下：

系统：MacOS 10.13.3

处理器：2.3 GHz Intel Core i5

内存：8GB 2133 MHz LPDDR3

软件环境如下：

Qt：Qt 5.10.0（Clang 7.0（Apple）,64 bit）

QtCreator：Qt Creator 4.5.0

WireShark：Wireshark Version 2.4.3

## 运行测试

首先我们通过已有的抓包软件（如wireshark）来捕获网络接口上的数据包，由于本数据包分析软件暂不支持网络上所有协议，所以本通过过滤规则来仅抓取支持类型的数据包（包括ICMP、TCP、HTTP、UDP、DHCP、DNS 、ARP等）。

然后本次测试依次对各个协议的支持效果进行检验，并通过过滤特性来找出想要的数据包 。首先在将pcap数据包文件放在应用程序所在文件夹内，如图

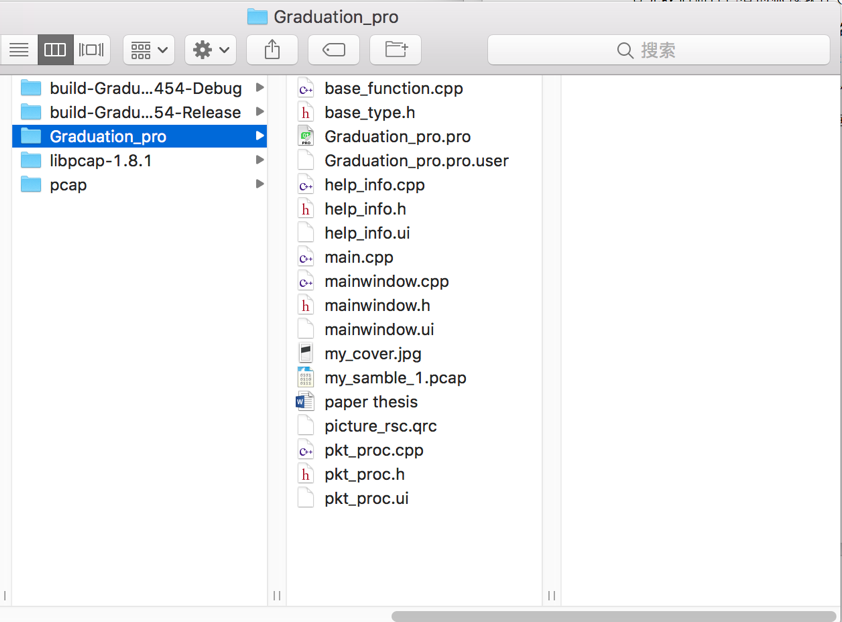


图4-1 文件放入程序所在目录

然后打开程序。在文件选择框中选择要解析的文件。程序的初始界面如下，当新加入文件时点击刷新即可识别到pcap文件。同时方便新用户的使用，程序添加了帮助按钮，可以通过点击查看程序的使用方法。如下图

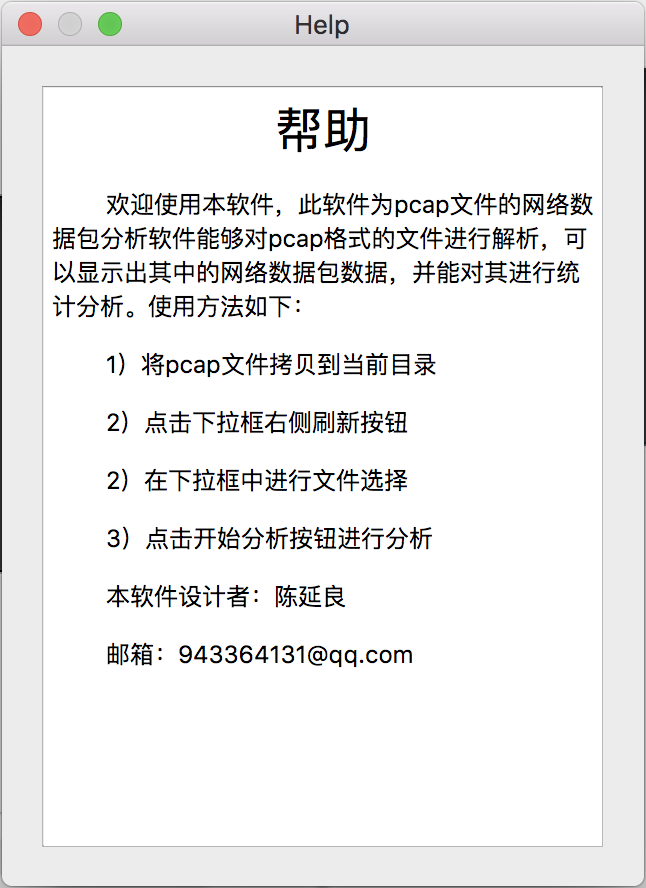


图4-2 程序初始界面

程序的使用分为4步打开：

* 的将pcap文件拷贝到程序当前目录
* 点击下拉框，若无待选择文件可点击右侧刷新按钮
* 从下拉框中选定待解析文件
* 点击开始解析按钮开始解析，进入下一界面

分析界面主要由4部分组成，从左至右从上至下共有四个输出显示区域，区域1为总揽表，显示了整个pcap文件中所有的数据包，包括数据包的一些基本信息，例如捕获时间、源地址、目的地址、协议类型、长度、应用层初略信息等；区域2为程序详细解析的输出框，对于不同网络数据包类型来说，有着不同的头部结构，此输出框详细的列出数据包各个层次的头部信息；区域3为二进制数据包内容，用十六进制来简化表示，以及其对应的ASCII码；区域4为整个pcap数据包统计的基本信息，包括pcap文件头的信息。

打开文件my\_samble\_1.pcap，该文件的总揽界面如下：

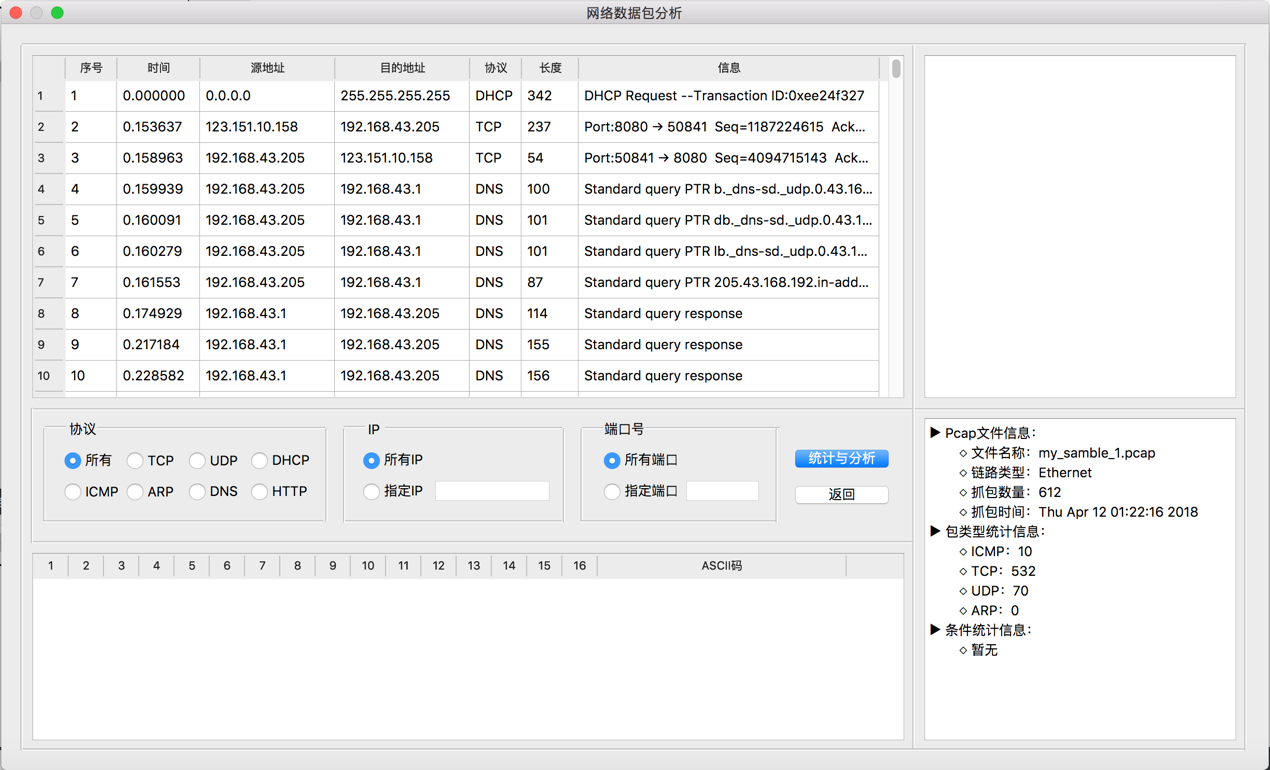


图4-3 解析初始界面

在总揽表中选取一个ICMP数据包，可以在下放的数据框中看到包内数据的二进制信息，同时在右侧可以看到ICMP数据包的各层包头的详细解析信息，包括MAC头、IP头、ICMP头，然后还有ICMP的报文详细信息解析，例如改ICMP是一个类型代码为3的端口不可达的报文。

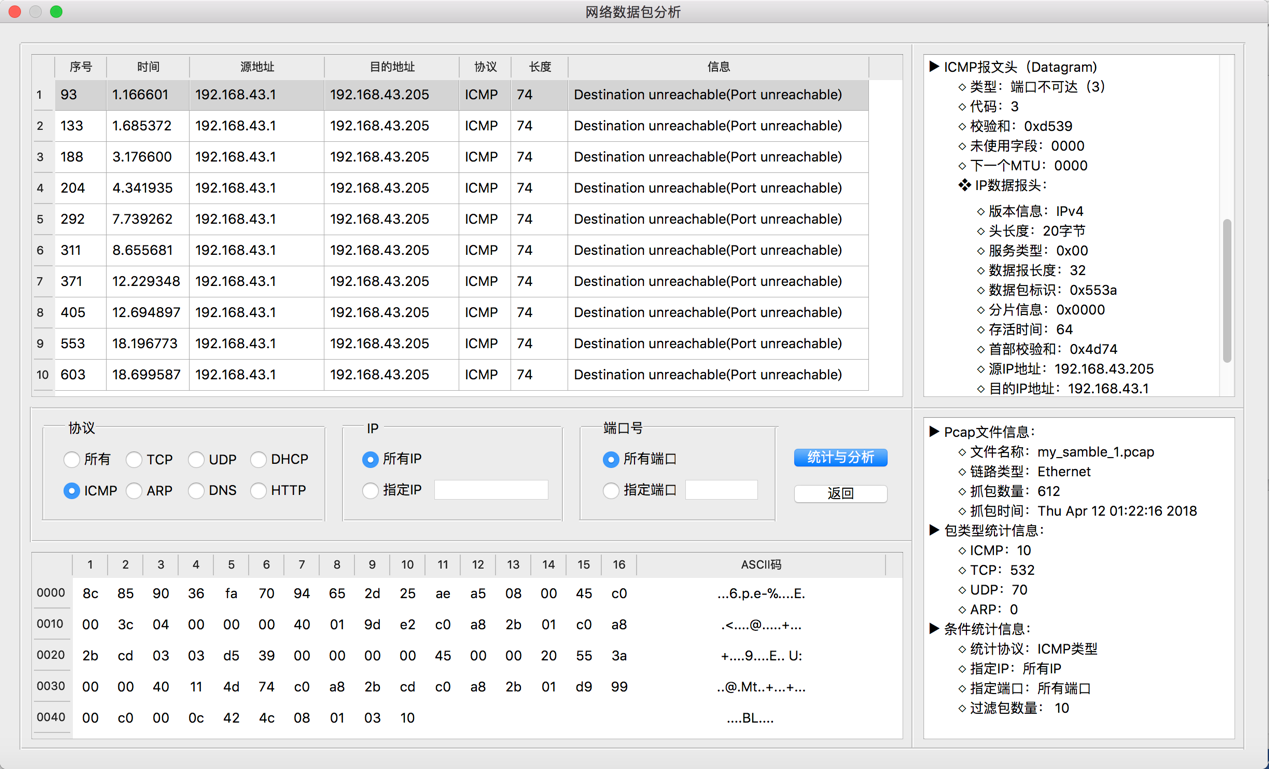


图4-4 ICMP报文解析

在总揽表中选取一个TCP数据包，在信息概览栏中可以看到，该TCP报文由源地址123.151.10.158发送到192.168.43.205的网络主机中，长度为237，端口号是由8080到50841。

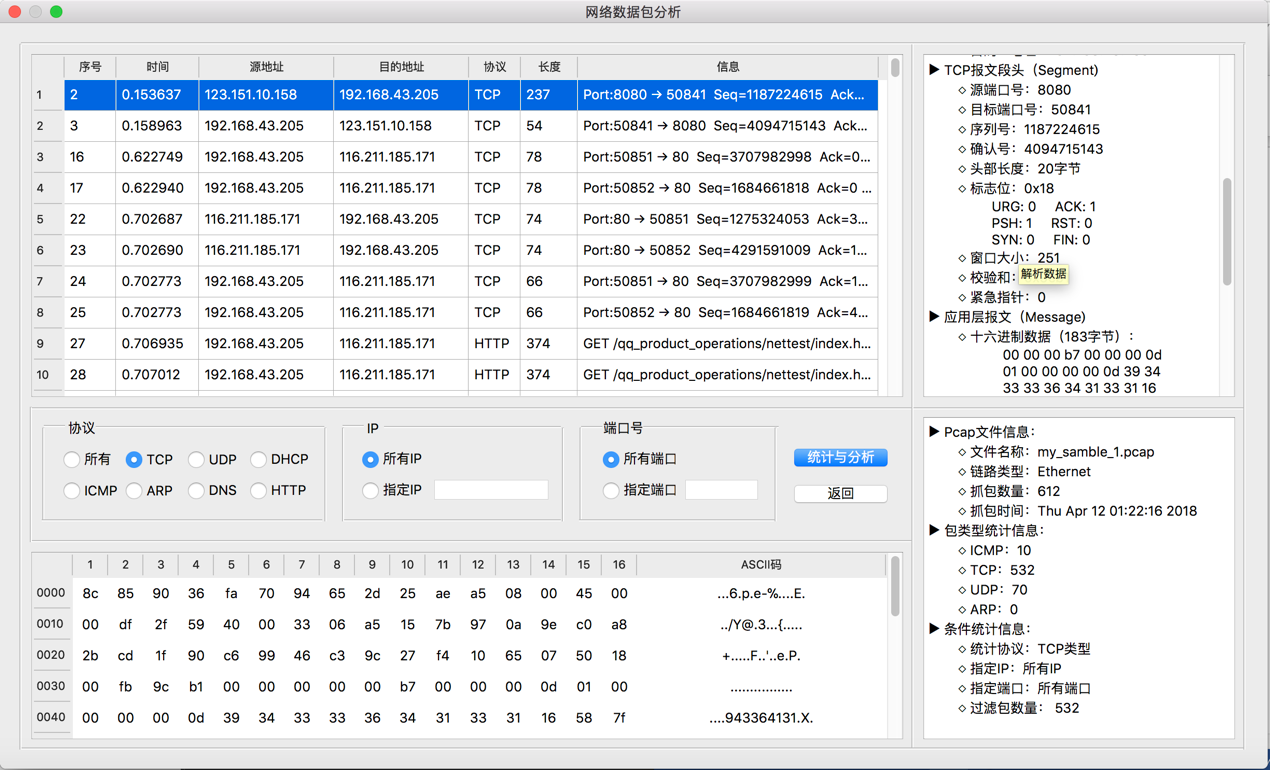


图4-5 普通TCP报文解析

在总揽表中选取一个HTTP数据包，可以在一旁的解析数据框中看到该HTTP报文的源地址和目的地址，TCP包头信息以及HTTP报文的内容。

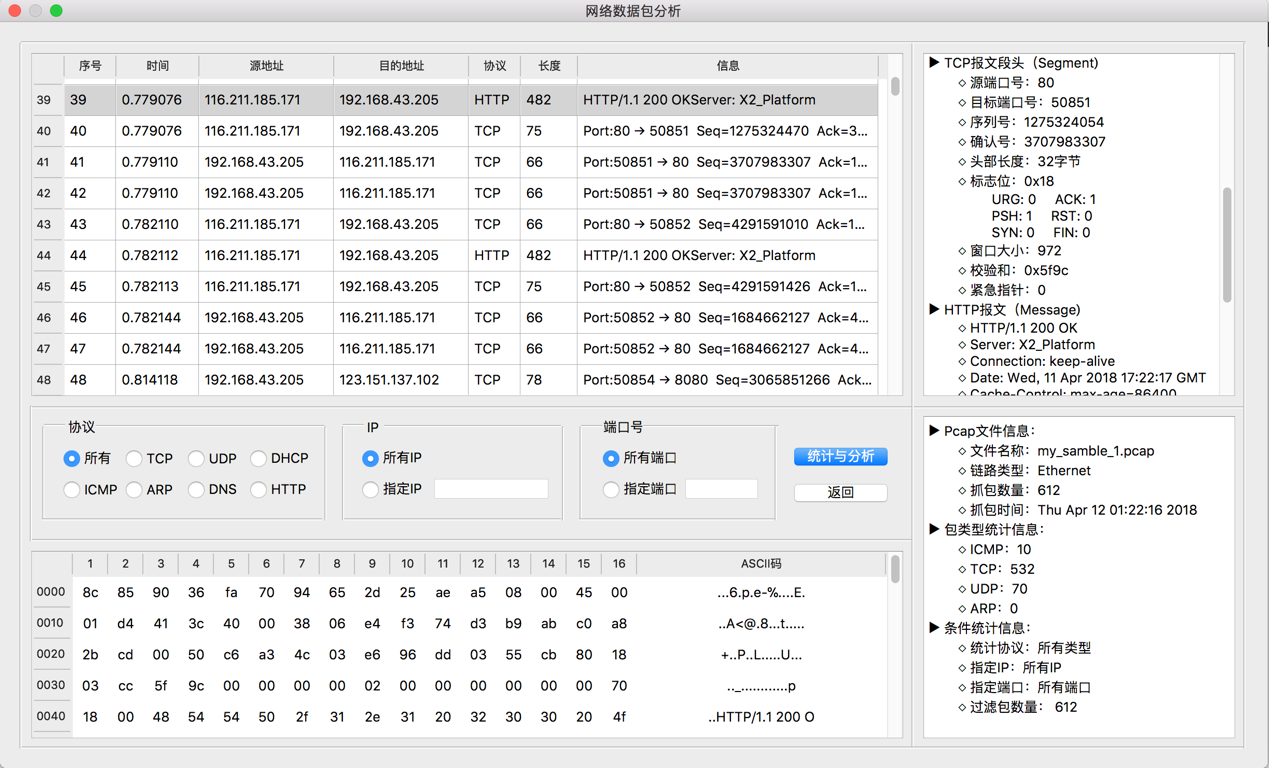


图4-6 HTTP报文解析

在总揽表中选取一个UDP数据包，该UDP由源地址为192.168.43.205主机中的端口55705发往目的地址为192.168.43.1的目的主机的192号端口，其内容为4字节。

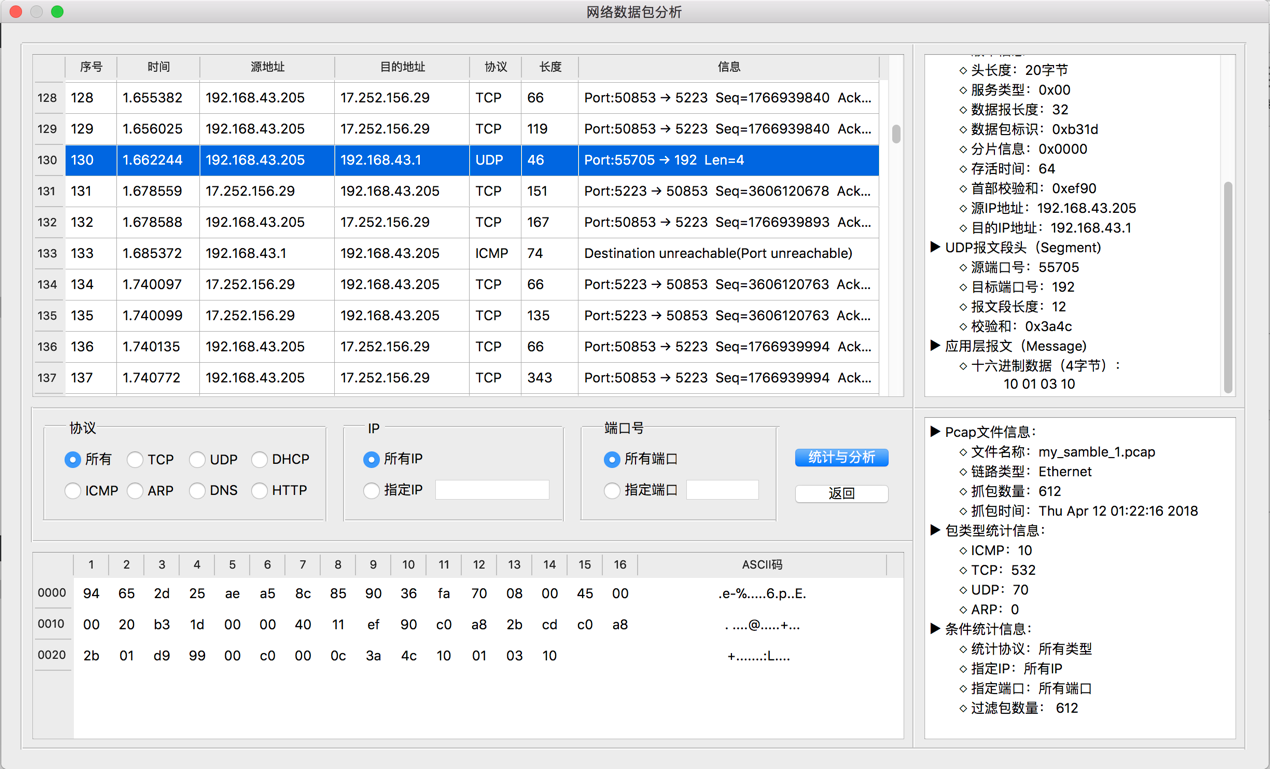


图4-7 普通UDP报文解析

在总揽表中选取一个DNS数据包，可以在一旁的解析框中看到，该DNS报文为一个查询报文。查询地址类型为A，域名为3gimg.qq.com。

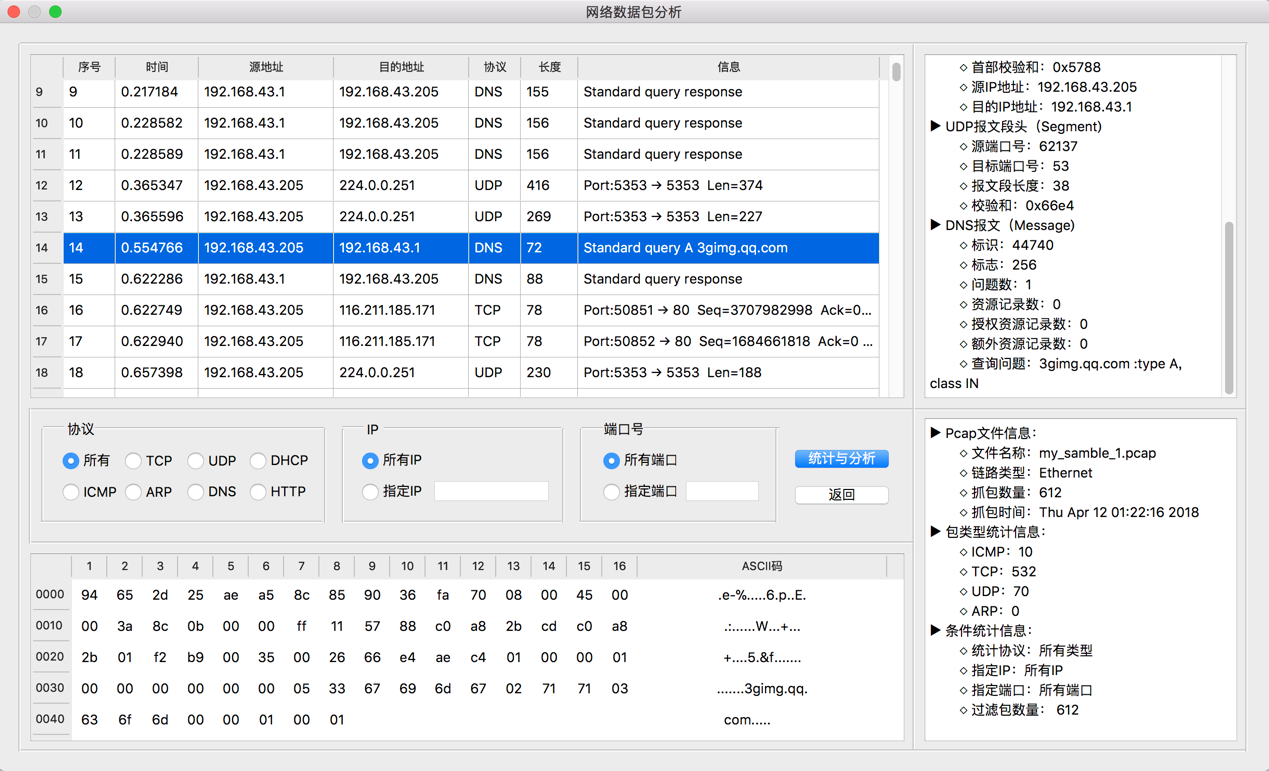


图4-8 DNS报文解析

在总揽表中选取一个DHCP数据包，可以看到该DHCP报文为请求报文，是一个目的地址为255.255.255.255的广播报文，其请求者的硬件地址为8c:85:90:36:fa:70。

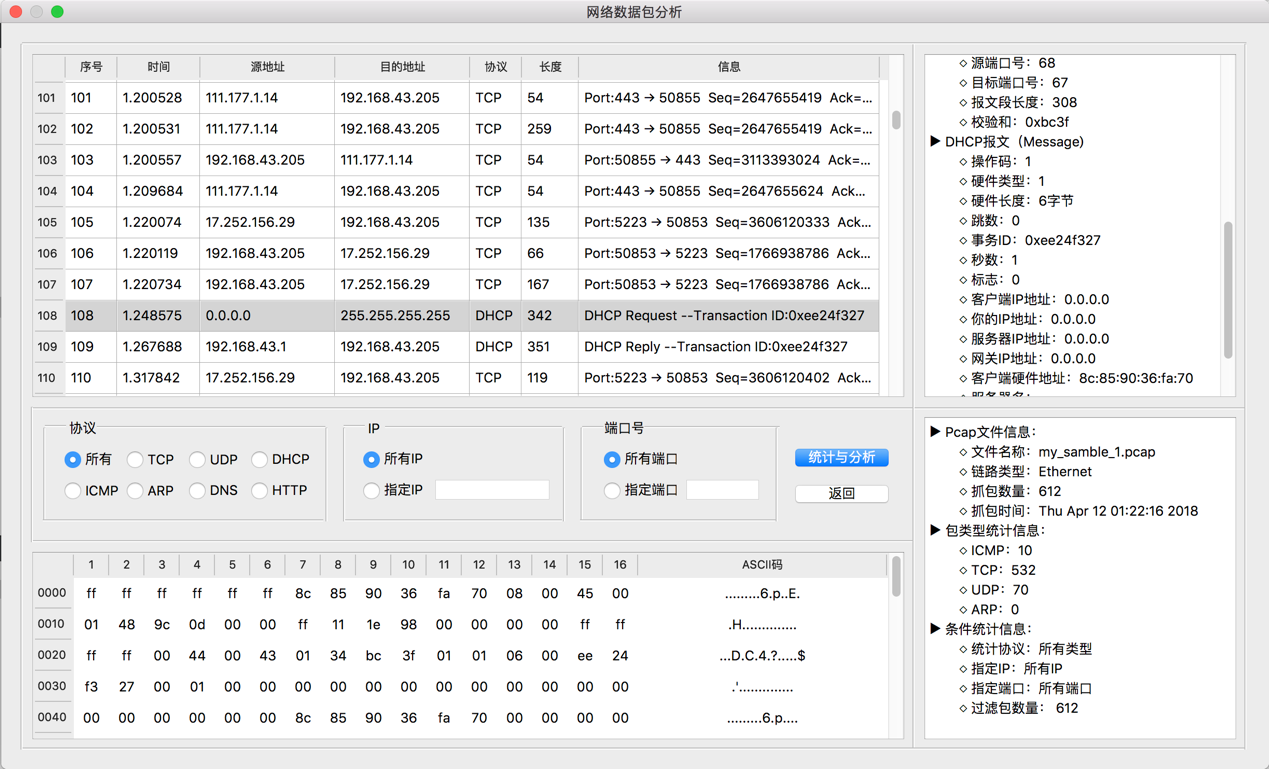


图4-9 DHCP报文解析

在总揽表中选取一个ARP数据包，该ARP报文为请求报文，是一个局域网的广播报文，请求IP地址为192.168.31.131主机的MAC地址。



图4-10 ARP报文解析

在协议选择框中选择一种协议类型，例如在测试时选择了ICMP协议类型，然后指定IP为192.168.43.1的主机相关的ICMP报文，可以在总揽表中看到所有的ICMP报文，同样右侧在点击相应报文后提供了解析后的包头数据。

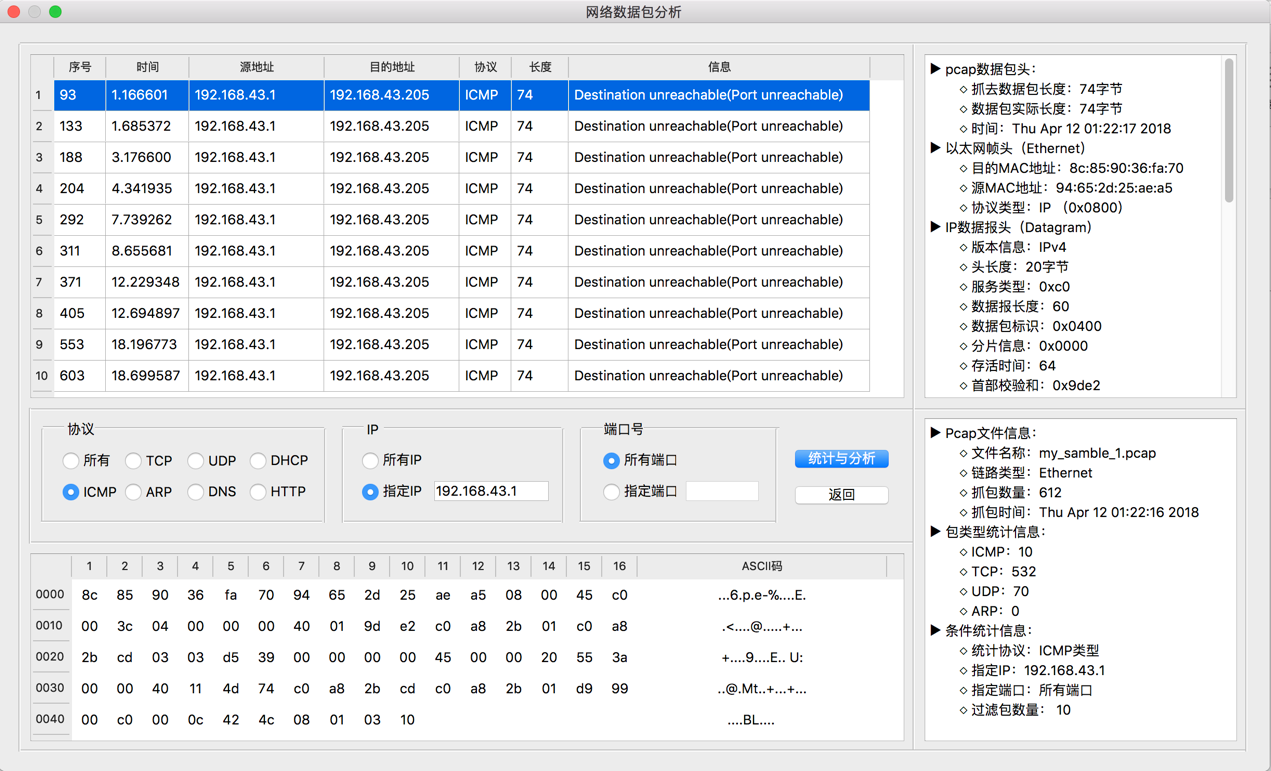


图4-11 过滤测试一

在端口输入框中输入一个端口号，然后对一个特定的协议如TCP进行过滤测试，在这里选择了端口号8080，可以看到表中都是与端口8080相关的数据。

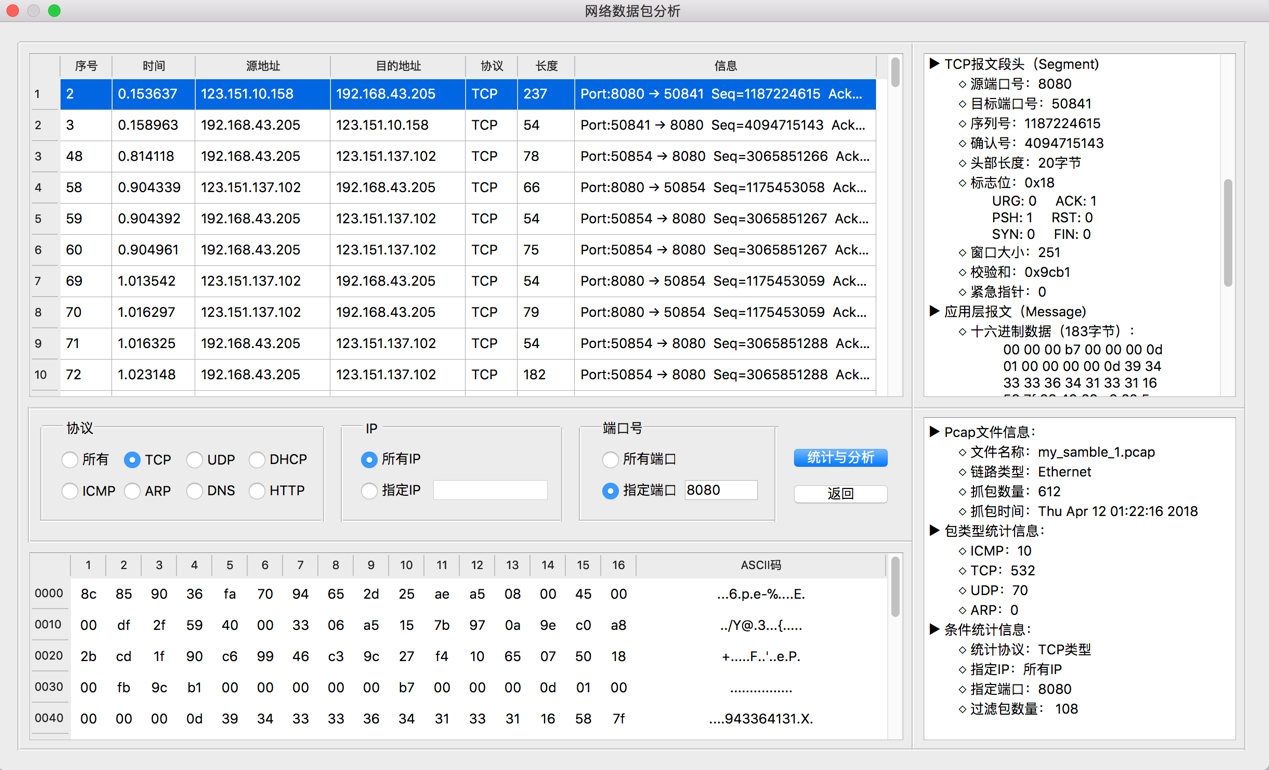


图4-13 端口过滤测试

## 本章小结

本章通过对不同类型的数据包进行解析测试。展示了系统对多个协议的支持情况。目前系统主要支持一下协议：ICMP、TCP、UDP、ARP等，同时在应用层协议有HTTP、DNS、DHCP等。系统通过表格的形式将数据包显示出来，然后通过点击表格的数据包条目，在文本框中进一步显示详细包信息，同时在下方将当前数据包的二进制数据显示给用户。

# 总结与展望

在这几个月的设计与实现中，网络数据包分析软件基本实现了预期的要求：

1. 能够读取pcap文件类型的数据包文件，能够正常解析出文件中信息。
2. 实现对多个协议类型的解析，目前支持网络层支持对ARP、IP等协议解析，运输层支持对ICMP、TCP、UDP等协议解析，应用层支持对HTTP、DNS、DHCP等协议解析。
3. 能够将支持对pcap文件中所有数据包进行分类统计，并对统计分析结果呈现给用户。
4. 能够通过协议、IP地址、端口号对数据包进行过滤，得到特定的数据包。
5. 用户界面相对友好，用户能够方便地对数据包进行概览，选择特定的数据包类型进行分析查看，得到有用的数据。

当然，由于时间有限，接下来讨论一些本课题还有一些目前存在的问题和一些日后可进一步完善的地方。

1. 支持协议不完全，对于一个普通网络适配器正常情况下所捕捉流量尚不能完全解析出来。
2. 尚不具备对异常流量的分类以及分析功能。
3. 只能对协议、端口号、IP地址进行过滤，不具备对MAC地址等其他筛选条件的过滤功能。
4. 尚未实现以图的形式显示统计结果的功能。

本次设计的网络数据包分析软件在对HTTP包进行解析时可能会对错误地处理。这个功能我们也打算在以后的研究中实现。

致 谢

整个毕业设计从题目的选择到最后实现所有的功能用了大概四个月左右的时间，在这段时间里，我得到了很多的帮助，也从中学到了很多。

首先要感谢的是我的导师周正勇老师，在整个设计过程中，周正勇老师给予我无私的帮助，为我提供了种种专业知识上的指导，和一些富于创造性的建议，在我遇到问题，苦于无法解决时，老师总是能及时给予我细心而专业的回答，在后期答辩之前，周老师也仔细浏览了我的开题报告，文献翻译等内容并提出了许多珍贵的修改意见，周老师一丝不苟的作风，严谨求实的态度使我深受感动。正是在周老师的帮助下，我得以顺利的完成此次毕业设计。在此向周正勇老师表示深深的感谢和崇高的敬意！

其次，在临近毕业之际，我还要借此机会感谢这四年中给予我诸多教诲和帮助的各位老师们，是你们的教诲让我对计算机领域有了大致的了解并为以后进行更深入的研究打下了夯实的基础。感谢他们四年来的辛勤栽培，不积跬步何以致千里，各位任课老师认真负责，在他们悉心的帮助和支持下，我能够很好的掌握和运用专业知识，并在设计中得以体现，顺利完成毕业论文。

此外，我还要感谢各位同学和我的各位室友，在毕业时即这段时间里，你们给予了我很多启发，提出了许多宝贵的意见。当然，除了在学习上的互帮互助，还有在生活上的互相照顾。对于你们的帮助和支持，在此我表示深深的感谢！

最后，感谢母校，感谢学院。学院为我们提供了丰富的学习资源与优秀的学习环境，让我们在一个良好的氛围中度过了这充实的大学四年。

参考文献

1. Stanford L. Levin,Stephen Schmidt. IPv4 to IPv6: Challenges, solutions, and lessons[J]. Telecommunications Policy,2014,38(11).
2. Jingyi He,S.-H.Gary Chan. TCP and UDP performance for Internet over optical packet-switched networks[J]. Computer Networks,2004,45(4).
3. Bruce J. Nikkel. An introduction to investigating IPv6 networks[J]. Digital Investigation,2007,4(2).
4. Eric Gamess,Rina Surós. An upper bound model for TCP and UDP throughput in IPv4 and IPv6[J]. Journal of Network and Computer Applications,2007,31(4).
5. Li Wenwei,Zhang Dafang,Yang Jinmin,Xie Gaogang. On evaluating the differences of TCP and ICMP in network measurement[J]. Computer Communications,2006,30(2).
6. S. Saurabh,A.S. Sairam. ICMP based IP traceback with negligible overhead for highly distributed reflector attack using bloom filters[J]. Computer Communications,2014,42
7. Lakshmanan Senthilkumar,Venkadachalam Sankaranarayanan. Integrated Proactive Admission Control Technique For both UDP And TCP Traffic Flows[J]. Journal of Networks,2007,2(1).
8. Prof.N.Penchalaiah,Dr.R.Seshadri. Performance of TCP/IP/ UDP adaptive header compression algorithm for wireless network[J]. International Journal on Computer Science and Engineering,2010,2(5).
9. Chrome G. HTTP persistent connection[J]. 2015.
10. Biju Issac. Secure ARP and Secure DHCP Protocols to Mitigate Security Attacks[J]. International Journal of Network Security,2009,8(2).
11. Woo-Yong Choi. Reliable Multicast MAC Protocol for IEEE 802.11 Wireless LANs with Extended Service Range[J]. Frequenz,2011,65(11-12).
12. Nora G. Uberti Manassero,Ivana L. Viola,Elina Welchen,Daniel H. Gonzalez. TCP transcription factors: architectures of plant form[J]. BioMolecular Concepts,2013,4(2).
13. J. Domańska,A. Domański,T. Czachórski,J. Klamka. Fluid flow approximation of time-limited TCP/UDP/XCP streams[J]. Bulletin of the Polish Academy of Sciences Technical Sciences,2014,62(2).
14. Ghazi Al Sukkar,Ramzi Saifan,Sufian Khwaldeh,Mahmoud Maqableh,Iyad Jafar. Address Resolution Protocol (ARP): Spoofing Attack and Proposed Defense[J]. Communications and Network,2016,08(03).
15. Craig Langston,Esther Hiu-Kwan Yung,Edwin Hon-Wan Chan. The application of ARP modelling to adaptive reuse projects in Hong Kong[J]. Habitat International,2013,40.
16. Xie Chao Guo. Design of ARP Intrusion Detection System[J]. Applied Mechanics and Materials,2014,3066(539)..
17. 武焱,袁焊忠,席振元.TCP/IP协议分析器的设计与实现[J].中小企业管理与科技(下旬刊),2016(01):251-252.
18. 史靓菲.网络监控之数据包分析工具的开发与设计[J].科技资讯,2006(22):96.
19. 张静. 网络协议分析技术的研究与应用[D].中南大学,2012.
20. 许应康,张阿莉.基于PCAP格式网络数据包分析软件设计[J].现代电子技术,2013,36(10):49-51.
21. 唐小婕.一种网络数据包分析工具的设计与实现[J].科技信息,2013(26):298-299.
22. 高凯,赵登攀.PCAP文件格式网络数据包分析软件设计与实现[J].软件导刊,2013,12(12):150-151.
23. 马畑名. 网络通讯数据包截获与解析研究[D].陕西科技大学,2015.
24. 艾翔.网络数据包分析系统的设计与实现[J].商品与质量,2010(S7):135-136.
25. 李振学. 基于协议分析的网络管理技术研究与实践[D].东北石油大学,2014.

附录：英文缩写

|  |  |  |
| --- | --- | --- |
| 英文缩写 | 英文全称 | 中文全称 |
| IP | Internet Protocol | 因特网互联协议 |
| MAC | Media Access Control | 媒体访问控制（硬件地址） |
| ICMP | Internet Control Message Protocol | 因特网控制报文协议 |
| TCP | Transport Control Protocol | 传输控制协议 |
| UDP | User Datagram Protocol | 用户数据报协议 |
| ARP | Address Resolution Protocol | 地址解析协议 |
| HTTP | HyperText Transfer Protocol | 超文本传输协议 |
| DNS | Domain Name System | 域名系统 |
| DHCP | Dynamic Host Configuration Protocol | 动态主机配置协议 |

**成 绩 评 定**

**指导教师评定意见**

一、对毕业设计（论文）的学术评语（应具体、确切、实事求是）

|  |
| --- |
| 陈延良同学学习态度评价…工作时间进度评价…论文内容与价值…论文结构评价…论文规范性评价…作者学术水平评价…，论文（基本）达到本科毕业设计论文的要求，同意参加本科毕业答辩。  导师评语建议学生先写个草稿，由导师审阅论文时一并修订成型，最终评语可以直接打印，导师签字即可，）  （答辩前导师需要给出具体成绩，下表评分中合计部分更新域可自动运算求和） |

二、对毕业设计评分

(1)理工医科评分表

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 评分项目  (分值) | 调研论证  (10分) | 外文翻译  (5分) | 设计(论文)撰写质量  (10分) | 学习态度  (10分) | 基本理论和基本技能  (50分) | 创 新  (15分) | 合 计  (100分) |
| 得分 | **8** | **4** | **8** | **8** | **40** | **12** | **80** |

(2)文科评分表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 评分项目  (分值) | 文献阅读与文献综述  (10分) | 外文翻译  (10分) | 论文撰写质量  (10分) | 学习态度  (10分) | 学术水平、论证能力和创新  (60分) | 合 计  (100分) |
| 得分 |  |  |  |  |  |  |

指导教师签字：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 年 月 日

**答辩小组评定意见**

一、评语（根据学生答辩情况及其设计（论文）质量综合评价）

|  |
| --- |
|  |

二、评分

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 评分项目  (分值) | 答 辩 情 况 | | 论 文 质 量 | | 合 计  (100分) |
| 答辩情况  (15分) | 回答问题情况  (25分) | 规范要求与文字表达  (20分) | 学术水平  (40分) |
| 得分 |  |  |  |  |  |

答辩小组长签字：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 年 月 日

**毕业答辩及成绩评定说明**

1. 毕业答辩
2. 答辩前，答辩小组应详细审阅每个学生的毕业设计（论文），为答辩做好准备。
3. 严肃认真组织答辩，开好答辩会。
4. 指导教师应参加所指导学生的答辩会，但评定其成绩时宜回避。
5. 答辩中要做好记录以供成绩评定时参考。
6. 成绩评定
7. 答辩前每个学生都要将自己的毕业设计（论文）在指定时间内交给指导教师，由指导教师审阅，写出评语并预评分。
8. 答辩工作结束后，答辩小组应举行专门会议进行讨论，在参考指导教师预评结果的基础上，结合学生毕业设计（论文）质量和学生答辩情况，综合评定每个学生的成绩。
9. 院（系）对专业答辩小组提出的优秀和不及格的毕业设计（论文），要组织院（系）级答辩，最终确定成绩，并向学生公布。
10. 各专业学生的最后成绩应符合正态分布规律。
11. 请用蓝、黑钢笔手写或五号宋体字编辑，签名须手写，A4纸双面打印。

**毕业设计（论文）成绩评定**

|  |
| --- |
| 班号：**计科201408班** 学生姓名：**陈延良**  综合成绩：\_\_\_\_\_\_**86**\_\_\_\_\_\_\_分（折合等级\_\_\_\_\_**良**\_\_\_\_\_\_\_）  答辩小组长（签名）：\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 年 月 日 |