即时通信——原理、技术和应用

王海涛　付　鹰
解放军理工大学通信工程学院　南京　210007

摘　要　首先，介绍即时通信的概念、特点和技术原理，较为全面地剖析了实现即时通信系统涉及的关键技术，包括即时通信传输协议、相关安全技术和音/视频编解码技术等；其次，简要概述了即时通信系统在我校的应用情况；最后，说明当前即时通信工具存在的问题及其发展趋势。

关键词　即时通信；音视频编码；对等通信；数据加密；防火墙

计算机网络的飞速发展，极大改变了人们的通信方式，网络交流已成为现代社会人际交往的一种时尚、便捷的方式，使得近年来迅速崛起的即时通信(Instant Messaging，IM)工具受到人们的普遍欢迎[1]。即时通信是基于计算机网络的一种新兴应用，它最基本的特征就是信息的即时传递和用户的交互性，并可将音、视频通信、文件传输及网络聊天等业务集成为一体，为人们开辟了一种新型的沟通途径。简单地讲，即时通信是一种使人们能在网络上方便快捷识别在线用户并与他们实时交换信息的技术，并逐渐成为继电子邮件之后最受欢迎的在线通讯和交流方式。与传统通信方式相比，即时通信具备快捷、廉价、隐秘性高的特点，在网络中可以跨年龄、身份、行业、地域的限制，达到人与人、人与信息之间的零距离交流。从这点上讲，网络即时通讯的出现改变了人们的沟通方式和交友文化，大大拓展了个人生活交流的空间。即时通信的出现和计算机网络和音/视频编码等技术的发展密不可分，本文将较为全面地介绍即时通信涉及的关键技术，并说明其存在的问题和发展趋势。

1　即时通信的技术原理

即时通信是一种基于Internet 的通信技术，涉及到IP/TCP/UDP/Sockets、P2P、C/S、多媒体音视频编解码/传送、Web Service等多种技术手段。无论即时通信系统的功能如何复杂，它们大都基于相同的技术原理，主要包括客户/服务器(C/S)通信模式和对等通信(P2P)模式[2]。

C/S结构以数据库服务为核心将连接在网络中的多个计算机形成一个有机的整体，客户机(Client)和服务器(Server)分别完成不同的功能。但在客户/服务器结构中，多个客户机并行操作，存在更新丢失和多用户控制问题。因此，在设计时要充分考虑信息处理的复杂程度来选择合适的结构。实际应用中，可以采用三层C/S结构，三层C/S结构与中间件模型非常相似，由基于工作站的客户层、基于服务器的中间层和基于主机的数据层组成。在三层结构中，客户不产生数据库查询命令，它访问服务器上的中间层，由中间层产生数据库查询命令。三层C/S结构便于工作部署，客户层主要处理交互界面，中间层表达事务逻辑，数据层负责管理数据源和可选的源数据转换。

P2P模式是非中心结构的对等通信模式，每一个客户(Peer)都是平等的参与者，承担服务使用者和服务提供者两个角色。客户之间进行直接通信，可充分利用网络带宽，减少网络的拥塞状况，使资源的利用率大大提高。同时由于没有中央节点的集中控制，系统的伸缩性较强，也能避免单点故障，提高系统的容错性能。但由

于P2P网络的分散性、自治性、动态性等特点，造成了某些情况下客户的访问结果是不可预见的。例如，一个请求可能得不到任何应答消息的反馈。

当前使用的IM系统大都组合使用了C/S和P2P模式。在登录IM进行身份认证阶段是工作在C/S方式，随后如果客户端之间可以直接通信则使用P2P方式工作，否则以C/S方式通过IM服务器通信。举例来说，在图1中，用户A希望和用户B通信，必须先与IM服务器建立连接，从IM服务器获取到用户B的IP地址和端口号，然后A向B发送通信信息。B收到A发送的信息后，可以按照A的IP和端口直接与其建立TCP连接，与A进行通信。此后的通信过程中，A与B之间的通信则不再依赖IM服务器，而采用一种对等通信(P2P)方式。由此可见，即使通信系统结合了C/S模式与P2P模式，也就是首先客户端与服务器之间采用C/S模式进行通信，包括注册、登录、获取通信成员列表等，随后，客户端之间可以采用P2P通信模式交互信息。

● Jabber：Jabber是一种开放的、基于XML的协议，用于即时通信消息的传输与表示。因特网上成千上万台服务器都使用基于Jabber协议的软件。Jabber系统中的一个关键理念是网关，支持用户使用其它协议访问网络，如AIM和ICQ、MSN Messenger 和Windows Messenger、SMS或E-mail。

● 可扩展通讯和表示协议(XMPP)：XMPP基于Jabber协议，用于流式传输实时通信、表示和请求-响应服务等的XML元素，是用于即时通信的一个开放且常用的协议。XMPP基于XML，具有语法清晰、易于实现等特点，因其专门面向即时通信，具有即时通信特需的一些功能特性，如好友列表、群组功能等。XMPP常常用于客户机/服务器架构当中，客户机需要利用XMPP协议通过TCP连接来访问服务器，而服务器也是通过TCP连接进行相互连接。

● 即时通信对话初始协议和表示扩展协议(SIMPLE)：SIMPLE协议基于SIP(session initial protocol)，为SIP协议指定了一整套的架构和扩展方面的规范。SIP是一种网际电话协议，可用于支持IM消息表示。SIP能够传送多种方式的信号，如INVITE信号和BYE信号分别用于启动和结束会话。SIMPLE协议在此基础上还增加了用于发送单一分页的即时通信内容的MESSAGE信号和用于传输显示信息的NOTIFY信号。

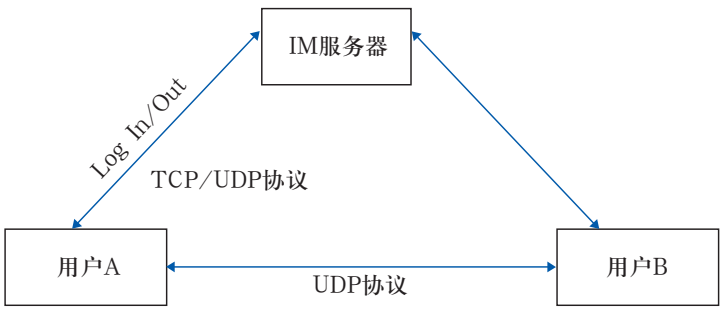


图1　IM技术原理示意图

2　即时通信的传输协议

现阶段的即时通信系统大多数是非标准系统，其通信协议与接口由厂商定义。为了解决即时通信的标准问题，IETF成立了专门的工作小组，研究和开发IM相关的协议。目前已提出了多个IM技术标准。其中比较有影响的是[3]：

● 即时通信通用结构协议(CPIM)：CPIM定义了通用协议和消息的格式，即时通信和显示服务都是通过CPIM来达到IM系统中的协作。

3　即时通信的安全技术

大多数的即时通信系统在设计时都考虑了系统的稳定性、兼容性和可扩展性，但在安全方面考虑不够，而

确保用户通信安全是一个不容忽视的问题[4]。

3.1 即时通信系统安全模型

即时通信系统的工作模式有两种：一种是传统的C/S模式，在这种模式下，用户首先与服务器建立连接，然后由服务器为用户提供信息发送通道，用户的交流信息都要通过服务器中转；另一种方式是在通信阶段IM客户端直接建立连接，构成P2P方式的IM。在即时通信系统中，P2P模型下的节点并非完全对等。在P2P通信中至少一方具有公网IP地址，可作为服务器监听并接受连接请求，以让另一方发起并建立TCP或UDP通信连接。实际的IM系统都是混合应用了C/S和P2P。在登录IM进行身份认证阶段是工作在C/S方式，随后如果客户端之间可以直接连接则使用P2P方式工作，否则以C/S方式通过IM服务器通信。典型的即时通信系统的安全模型如图2所示。此模型基于RFC2778定义的即时通信实体模型，涵盖了网络通常的连接形式和攻击行为模式，包括IM系统的三类参与者：IM服务器、IM客户端和攻击者。图中虚线表示点到点安全，点划线表示端到端安全。

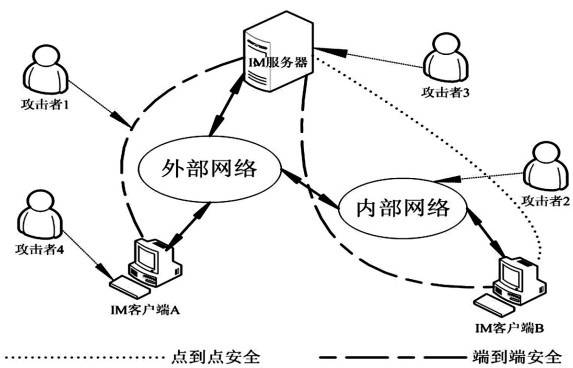


图2 IM的安全模型

3.2 数据加密技术

由于即时通信系统很可能受到安全攻击，必须对重要的数据进行加密。数据加密过程就是通过加密系统把原始的数字信息(明文)，通过加密算法变换成与明文完全不同的数字信息(密文)的过程[5]。

数据加密技术主要分为数据传输加密和数据存储加密。在即时通信系统中，数据存储加密主要是指后台的服务器端对用户信息(帐号、密码)、用户关系列表等重要数据以及前台客户端的个人资料、通信记录的加密；数据传输加密技术主要是对传输中的数据流进行加密，常用的有链路加密、节点加密和端到端加密三种方式。

链路加密是仅在数据链路层加密传输数据，不考虑信源和信宿，用于保护通信节点间的数据。信息在每台节点机内都要被解密和再加密，依次进行，直至到达目的地。节点加密方法是在节点采用一个与节点机相连的密码装置，密文在该装置中被解密并被重新加密，明文不通过节点，避免了链路加密节点易受攻击的缺点。端到端加密是为数据从一端到另一端提供的加密方式。数据在发送端被加密，在接收端解密。在端到端加密中，除报头外的报文均以密文的形式贯穿于全部传输过程。因此，在中间节点不需要有密码设备，同链路加密相比，可减少密码设备的数量。另一方面，在链路加密时，报文和报头两者均须加密。而在端到端加密时，为将报文传送到目的地，必须检查路由选择信息，因此，只能加密报文，而不能对报头加密。这样就容易被某些通信分析工具发觉，而从中获取某些敏感信息。

3.3 防火墙技术

在对抗各种网络攻击的技术中，防火墙技术是有效而常用的防御技术之一。防火墙实质上就是在网络的不同层次上所设置的电子屏障。防火墙的核心是安全策略，这是一种对什么内容能够通过防火墙边界的技术规范。管理员可以根据以下两个原则为防火墙设置访问控制策略[6]：除了明确的合法流量，禁止其它所有流量；除了明确的非法流量，允许其它所有流量。防火墙较好地解决了内部网的安全问题，并且由于它易用易配易扩展，因而成为网络安全防卫的首选工具。

4 音视频编码技术

在即时通信系统的实现过程中，必须考虑充分利用网络带宽。音视频数据对带宽的要求比较高，只有选择合适的音频和视频编码算法并对之进行优化，才能保证软件的实用性。

4.1 H.323协议

H.323是一套在分组网上提供实时音频、视频和数据通信的标准，是ITU-T制订的在各种网络上提供多媒体通信的系列协议H.32x的一部分[7]。H.323协议被普遍认为是目前在分组网上支持语音、图像和数据业务最成熟的协议。

H.323制定了无QoS保证的分组网络上的多媒体通信系统标准，为LAN、WAN、Internet、因特网上的数据通信的应用提供了技术基础和保障。从整体上来说，H.323是一个框架性建设，它涉及到终端设备、音频、视频和数据传输、通信控制、网络接口方面的内容，还包括了组成多点会议的多点控制单元(MCU)、多点控制器(MC)、多点处理器(MP)、网关以及关守等设备。H.323规定了不同的音频、视频或数据终端协同工作所需的操作模式。它将是下一代因特网电话、音频会议终端和视频会议技术的主要标准。它所包含的主要功能单元及其标准如下：

视频编码(H.263/H.261)：完成对视频码流的冗余压缩编码。

音频编码(G.723.1等)：完成语音信号的编码，并在接收端可选择加入缓冲延迟以保证语音的连续性。所采用的标准为ITU-T的G.723.1，它提供5.3kbit/s和6.3kbit/s两种码率，采用线性预测综合分析编码方法，分别使用代数码本激励线性预测和多脉冲最大似然量化，从而各自获得编码复杂度和质量的优化。

控制单元(H.245)：提供端到端信令，以保证H.323终端的正常通信。所采用的协议为H.245(多媒体通信控制协议)，它定义了请求、应答、信令和指示四种信息，通过各种终端间进行通信能力协商，打开/关闭逻辑信道，发送命令或指示等操作，完成对通信的控制。

H.225层：将视频、音频、控制等数据格式化并发送，同时从网络接收数据。另外，还负责处理一些诸如逻辑分帧、加序列号、错误检测等功能。

4.2 音频编码技术

音频编码技术主要分为三类：

第一类是波形编码，力图使重建语音的波形保持原始语音的波形形状，如PCM和ADPCM(G.711、G.721、G.722、G.723、G.726、G.727)。

第二类是参数编码，通过提取、编码语音的特征参数，保持重建语音的可懂度，如LPC-10e等(G.723.1)。

第三类是混合编码，结合了上述两种方法的优点，能重构高质量的语音，如矢量和激励线性预测和码激励线性预测(CELP)等，如G.728、G.729、GIPS。

在即时系统中，用户可以通过使用自己的声音输入设备(声卡)进行语音的直接交流，同时可以根据网络带宽情况进行自适应调节，确保最佳的表现效果。

在即时通信系统的实现过程中，经常选择基于H.323的G.723.1音频编码算法，是因为它具有以下特点：能提供极低的码率和较高的重建语音质量；具有中等复杂程度的编码算法；具有较佳抗噪声性能；编码延迟较长是G.723.1音频编码算法的不足，在实际应用中应该考虑语音与其它信号的同步问题。

4.3 视频编码技术

H.263基于H.323，为视频会议和视频电话应用程序提供压缩。H.263视频编码标准是专为中高质量运动图像压缩所设计的低码率图像压缩标准[8]。与H.261的p×64K的传输码率相比，H.263的码率更低，单位码率可以小于64K，且支持的原始图像格式更多。

H.263将编码过程分为帧内编码和帧间编码两个部分。在帧内用改进的DCT变换并量化，在帧间采用1/2象素运动矢量预测补偿技术，使运动补偿更加精确。H.263的编码速度快，其设计编码延时不超过150ms；码率低，在512K乃至384K带宽下仍可得到相当满意的图像效果，十分适用于需要双向编码并传输的场合(如可视电话)和网络条件不是很好的场合(如远程监控)。

运动估计在H263压缩编码方案中占到总计算量的60%～80%，因此改进的重点在于寻找一种简便、快捷、高效的运动估计算法。目前H263采用基于块匹配的运动估计法。在各种实现算法中，全搜索得到的运动估计最有利于视频压缩，但巨大的时间开销成为其致命

的弱点。因此有必要对原有的钻石搜索算法作一定的优化，尤其针对钻石搜索算法容易产生重复计算的特点进行相应优化，同时结合先进的MMX技术进行进一步的优化，这样大大提高了相应的运算速度和压缩率[9]。

5 即时通信系统在我校的建设和应用

随着校园网基础设施的建设及网络应用的普及，对信息交流的即时性和多元性要求越来越迫切。如何能够借助网络工具拓展工作空间，提高工作效率，是我校信息化建设迫切需要解决的问题之一。为此，我校研制开发了校园网即时通信系。校园网即时通信系统由客户端和服务器端两部分组成，主要包括以下几个模块：个人信息管理、用户管理、信息操作、文件操作、语音功能、视频功能。

服务器端的功能主要包括：

1）服务器IP地址的获取：服务器端软件自动从运行的计算机获取本机地址作为服务器IP地址，客户端登录时必须使用服务器的IP地址。

2）用户管理：在服务器端，管理员具有高权限，可以进行如下操作(打开/关闭服务器，允许/禁止用户注册，允许/禁止用户登录)。

3）信息操作：包括信息的广播和对用户私聊信息的发送。

客户端的功能涉及以下几方面：

1）信息操作：对即时通信信息的发送与接收。

2）文件操作：文件的安全发送与接收。

3）语音功能：调用语音模块，实现语音数据的传输。

4）视频功能：调用视频模块，实现视频数据的传输。

5）用户管理：包括用户的查找，添加好友，删除好友。

6）消息管理：对服务器端发送至个人用户的广播和私聊信息进行存储与管理，对个人的通信记录进行管理(导出或者完全删除通信记录)。

7）个人信息管理：包括用户注册、密码更改、个人资料，个人状态(在线、离开、隐身、离线)。

校园网即时通信系统目前已在我校多媒体教学、远程教学、学术交流和科研管理等方面得到了广泛的应用，取得了很好的成效。利用该工具能够加强信息传送、文件传送、音频对话、视频对话、消息管理和信息记录回放等功能，提高了广大员工的办公效率、加强了教员和学员的交流沟通、随着应用不断铺开，解决了目前信息沟通手段的信息失真、不对称和时效性不强等问题，信息获得变得更加快捷准确，降低了通信成本，拓展了工作空间。

6 即时通信产品的现状和发展分析

目前即时通信产品市场竞争异常激烈，以腾讯QQ、微软MSN、网易POPO等为代表的众多即时通信产品拥有大批使用者。但是，这些产品功能上大同小异，并且缺乏互操作性，从而限制了即时通信市场的进一步发展，归纳而言存在以下几个问题。

6.1 产品同质化严重

众多即时通信工具功能相似。国内最主要的即时通信工具QQ是模仿借鉴ICQ起家的， 经过不断创新，QQ超越了ICQ并占据市场一般以上的份额。QQ的成功也为后来者提供了借鉴模式，模仿是最节省力气的办法，于是不同的通信软件推出了相似的面孔。

6.2 用户定位不清

QQ最先的定位是大众娱乐，为用户提供交流的平台。但随着用户群越来越大，用户需求越来越复杂。为满足不同用户需要，QQ不得不增加了更多的功能。很多人都感到QQ有些功能对自己并不需要，而特别需要的功能又得不到很好的满足。MSN最初定位是高端商务人士和高学历、高收入人群的一款相对专业、用于工作用途的即时通信软件，但为了抢夺更大的市场份额，MSN也增加了很多娱乐功能，使得它同其他软件的区分越来越模糊。国内其他通信软件也陷入了同QQ和MSN相似的怪圈。

6.3 各通信工具互不沟通

虚拟的互联网世界需要沟通，作为主要的沟通工

具之一，即时通信软件大有取代电子邮件成为互联网上第一大应用软件的发展趋势。即时通信为沟通而生，因差异而存在，但不同软件之间缺乏互通和互操作性却阻碍着它的进一步发展。因为互不沟通，有的用户只好同时安装多个即时通信软件，既占用了空间，又消耗了资源。越来越多的用户希望目前的即时通信软件之间能够实现兼容，做到互联互通，满足最大的用户需要。

6.4 即时通信工具成为电脑病毒传播新途径

据一份国外调查数据显示，病毒可以通过即时通信工具在30秒内感染多达50万台电脑，一种传播速度极快的IM病毒将可以利用一种软件漏洞突破电脑的安全防御系统，然后执行一些非授权命令。

基于上述问题，各主流即时通信开发和运营公司纷纷采取了相应的措施，并取得了一些进展。腾讯推出的QQ2005中加入了全新的功能“群”， 它的功能有点类似校友录上的留言版，利用该功能可以同时给多位离线朋友发信息。这样，即时通信软件也将由纯粹的P2P(Peer To Peer)变成G2G(Group To Group)。MSN采用邮件地址作为用户名，使用电子邮件地址进行注册，借此建立MSN与Hotmail之间的连接，同时还可给用户提供BLOG空间。而网易POPO新增加的一个功能就是紧密结合手机，给手机用户和泡泡用户一样的待遇，手机用户同样出现在泡泡的通信面板上。

沟通是即时通信软件的核心，如何为用户提供更优秀、更方便的沟通服务是用户最关心的功能和需求，也是即时通信软件发展的必然趋势。如果各通信软件能够互联互通，将节省大量的软件开发和维护成本，并且极大地扩大用户资源，提高广告效应，达到事半功倍的效果。虽然各通信软件融合是大势所趋，但在实施过程中还是会遇到很大的阻力。各公司出于商业考虑，很难与竞争对手互联互通。尤其是处于领先地位的QQ和MSN等。所以，差异与融合将在未来的一段时间里互相博弈，各通信软件也将使出浑身解数留住用户，而提供给用户的将会是更加精彩的服务。

7 结束语

本文对即时通信涉及到的关键技术进行了剖析，并说明了当前即时通信工具存在的问题及其今后的发展趋势。计算机网络的飞速发展，极大改变了人们的通信方式，近年来迅速崛起的即时通信工具受到了人们的普遍欢迎。即时通信系统作为电子邮件和网络电话的有益补充，丰富了当前计算机网络用户的通信手段，为政府机关、科研院所和个人用户提供了一种有效的信息交流方式，提高了工作效率，拓展了工作空间。可以预见，即时通信的发展和应用有着光明的前景。

参考文献

[1] 庞怡,许洪光,姜媛.即时通讯工具现状及发展趋势分析,科技情报开发经济,2006,16

[2] 段翰聪,卢显良,宋杰.面向连接的P2P即时通信中继策略研究,计算机科学,2005,6

[3] 李远杰,刘渭峰,张玉清.主流即时通软件通信协议分析,计算机应用技术

[4] 代印唐,张世永.即时通信安全研究,研究与设计,2005,4

[5] 张世永.网络安全原理与应用,科学出版社,2003

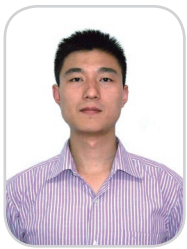
[6] 程玮玮,王清贤.防火墙技术原理及其安全脆弱性分析,计算机应用,2003,10

[7] 姚志恒,田栋,沈兰荪.基于H_263的实时视频编码技术研究,电路与系统学报,2002,3

[8] 徐丽琨,黄登山.基于H.263视频压缩的新钻石搜索算法,2005,9(5)

[9] 落红卫,吉宏宇.H.323安全研究与标准制定,现代电信科技,2006,8

作者简历



王海涛

博士，现为解放军理工大学通信工程学院副教授。研究方向为宽带网络、无线自组网和服务质量保障。

付　鹰

硕士，现为解放军理工大学通信工程学院讲师。研究方向为网络信息管理和应用。

Instant Messaging—Theory, Technique And Application

Wang Haitao
Fu Ying

Institute of Communication Engineering, PLAUST, Nanjing 210007, China

Abstract　Firstly the development status of IM are introduced and the key techniques are analyzed in detail, including transmission protocols, security schemes and audio/video coding techniques. Secondly, construction and application of IM in our school is introduced. In the end, problems in current IM tools and theirs future development trends are expounded.

Keywords　Instant Messaging; Audio/Video Coding; P2P; Data Encryption; Firewall