

基于 Agent 的 CSCW 多媒体交互环境的模型和实现

Modeling and Implementation of the Multimedia Interactive Environment Based on Agent in CSCW

张鹏程 李人厚 秦 明 田 锋 顾新华

(西安交通大学系统工程研究所 西安710049)

Abstract Based on the characteristics of the multimedia interaction in CSCW, this paper presents a model of the multimedia interactive environment utilizing Agent technology. The architecture model includes a multimedia interactive management server, one or more Agent-based MCUs and multimedia interactive clients. Then analysis of the model is given. Finally, an application example of the model is shown.

Keywords Agent, CSCW, Multimedia

1. 引言

1984年 Iren Greif 和 Paul Cashman 提出了计算机支持协同工作(Computer Supported Cooperative Work, 即 CSCW)的概念,这是一个利用计算机技术、网络与通信技术、多媒体技术、分布处理技术以及人机接口技术,将时间上分离、空间上分布而工作上又相互依赖的多个协作成员以及活动有机地组织起来,以共同完成某一项任务的新型分布式计算机系统^[1,2]。特别是近年来,随着计算机技术和网络技术的飞速发展,CSCW 理论和技术也得到了极大的进步,在电子商务、远程医疗、远程教育、远程协作编辑、远程协同设计等领域获得广泛的应用^[3]。

CSCW 中,以音频和视频为主要内容的多媒体技术向用户提供了 WYSIWIS(What You See Is What I See)的交互环境,是当前 CSCW 中一个重要的研究热点。目前对 CSCW 中多媒体交互技术的研究主要集中在视音频压缩技术、多媒体通信模型、传输质量控制技术以及系统结构模型等方面。CSCW 中,参加协作的人员可能处于同一位置,也可能分布在不同国家的不同地方,因而分布性是多媒体交互的一个基本特性。其次,多媒体交互环境是 CSCW 中实时交互的主要手段,它向协作用户提供及时的视音频交流,因而实时性也是多媒体交互环境的一个重要特点。还有,随着信息技术日新月异的发展,系统结构的开放性已经成了应用软件结构发展的方向,为了适应这一趋势,CSCW 中多媒体交互环境还应该具备结构的开放性以及易集成性。

CSCW 中,多媒体交互环境应该具有如下功能:

- 支持地域上分布的多个小组的实时全双工视音频交互和交互成员管理功能。
- 对不同节点的视音频数据源提供多点数据合成和路由转发功能。
- 灵活的多媒体交互策略控制管理功能。

本文根据 CSCW 系统中多媒体交互环境分布性、开放性、实时性的特点和功能目标,提出了基于 Agent 的多媒体交互环境结构模型(MIEBA),并对提出的系统结构模型进行了分析。最后,给出了基于 MIEBA 模型的一个实现例子。

2. 基于 Agent 的多媒体交互模型

CSCW 中,多媒体交互环境向用户提供的面对面的交互环境,也可归结为一个人人交互的空间。根据前文提出的多媒体交互环境的功能目标,基于 Agent 的多媒体交互环境可以用 BNF 范式表示为如下形式:

$\langle \text{MIEBA} \rangle ::= \langle \text{MIMServer} \rangle (\langle \text{BAMCU} \rangle)^+ (\langle \text{BAClient} \rangle)^+$

其中: $\langle \text{MIMServer} \rangle$ 为多媒体交互管理服务器 Agent,它负责管理多媒体客户端 Agent 以及多点控制单元 Agent 的注册服务、多媒体转发路径建立以及多媒体交互策略的控制; $\langle \text{BAMCU} \rangle$ 为多媒体多点控制单元 Agent,它负责对多媒体数据流的合成和不同网段间多媒体数据流的多点数据传送。 $(\langle \text{BAMCU} \rangle)^+$ 表示多媒体交互环境中包括一个或多个多点控制单元 Agent; $\langle \text{BAClient} \rangle$ 为多媒体交互客户端 Agent,负责本地视音频数据的采集和各节点视音频数据的表现。 $(\langle \text{BAClient} \rangle)^+$ 表示多媒体交互环境中包括一个以上的多媒体交互客户端。

图1给出的是多媒体交互环境的结构模型。

图中,多媒体交互环境各组成部分间的交互信息主要包括视音频多媒体数据流和交互环境中 Agent 间的控制消息。

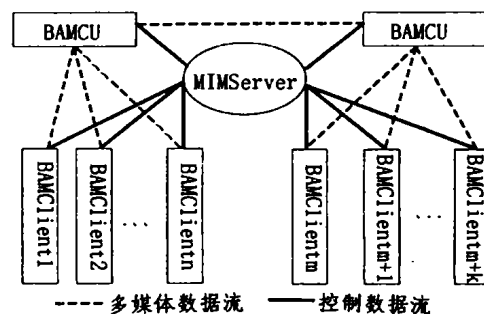


图1 CSCW 中多媒体交互环境结构

对于多媒体数据流,由于在全双工多媒体交互情况下,网络间的数据流量比较大,因而对各多媒体交互客户 Agent 发送的数据流,采用分布式多 MCU 策略来对音频数据进行合

张鹏程 博士研究生,主要研究网络与多媒体应用、CSCW、工程数据库理论、计算机辅助设计技术。李人厚 教授,博士生导师,主要研究领域为 CSCW、计算机辅助设计技术、网络安全、智能控制理论和方法。

成后再发送到会议中的其它节点的客户 Agent。为了保证数据传输的实时性,多媒体数据的传输采用实时传输协议(RTP)。

为了保证多媒体交互环境能够和其它基于 Agent 的协作工具进行集成,对于 Agent 间的信息交流,必须遵循共同的通讯语言协议。因为在协作中,Agent 不但要具有和人交互的能力,还要具有和其它 Agent 交互的社会协作能力。而且,采用共同的通讯语言协议和公共的内容表示方式也是实现不同用户不同时期开发的 Agent 实体间交互的重要保证。基于“speech acts”理论的 KQML 语言是目前基于 Agent 的系统中广泛采用的 Agent 通信语言,KQML 语言是一种描述性语言,其中定义了 Agent 间传递的消息格式和消息处理协议,能清楚地表示 Agent 间的交互行为^[4]。因而多媒体交互环境中,各 Agent 间的交互采用 KQML 语言。

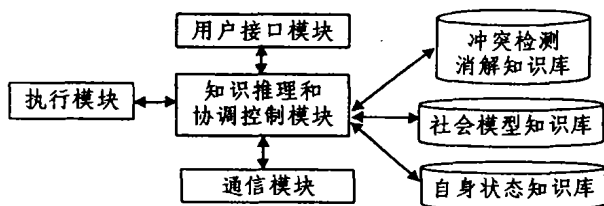


图2 Agent 结构模型

图2是我们给出的多媒体交互环境中的 Agent 结构模型。它由冲突检测消解知识库、自身状态信息知识库、社会模型知识库、知识推理和协调控制模块、执行模块、用户接口模块、通信模块等几部分组成。其中:

- 用户接口模块是 Agent 和人交互的主要通道,用户的意图和操作经过用户接口模块传送到 Agent 的知识推理和协调控制模块进行进一步的处理;Agent 对接收到的人和它 Agent 的消息处理后,经用户接口模块反馈给用户。

- 通信模块是 Agent 间传递消息的通道,Agent 通过通信模块进行协商交互,从而共同完成单独无法完成的功能目标。

- 知识推理和协调控制模块是 Agent 的神经中枢部分。该模块接收到人和其它 Agent 的消息后,根据冲突检测消除知识库中的知识,进行基于规则和事例知识推理的冲突检测和消解;冲突检测通过后,该模块利用社会模型和自身状态知识库中的知识对消息进行进一步的处理,并把处理的结果反馈给用户接口模块、其它 Agent 或传送到执行模块执行某一指定的功能。

- 执行模块是 Agent 具体功能的实现部分,该模块在知识推理和协调控制模块的控制下执行,并把执行结果反馈给知识推理和协调控制模块作进一步的处理。

- 冲突检测消解知识库中包括基于规则和事例的冲突检测知识、冲突消解知识。它是 Agent 进行冲突检测和消解的主要知识依据。

- 社会模型知识库中包括和该 Agent 已经建立熟人关系的 Agent 的名字、能力等社会信息,它是 Agent 间进行协同工作的前提和保证。

- 自身模型状态知识库是对 Agent 的领域问题求解功能、目标的说明以及自身状态信息的描述。它是 Agent 进行知识推理和协调控制的主要依据。

MIEBA 环境中,MIMServer 提供有注册管理服务,它负责管理所有节点 Agent 的注册。当各 BAMCU Agent 启动后,首先向 MIMServer 发送注册消息,使得 MIMServer 中拥

有关于该 BAMCU Agent 的社会知识和状态知识。同样多媒体交互客户 BAMClient Agent 启动后,向 MIMServer 发送注册消息。当 BAMClient 加入一个多媒体交互会议后,准备发送多媒体数据时,BAMClient 将向 MIMServer 发送 KQML 消息,查询环境中可以提供多媒体数据合成和多点数据传输的 BAMCU 的信息情况。MIMServer 在收到该查询信息后,根据自身知识库中的多媒体交互会议分组情况、BAMCU 的网络分布和负载情况,选择出相对最优的 BAMCU Agent,并以 KQML 消息告知多媒体交互客户 BAMClient Agent。多媒体交互过程中,当客户 BAMClient 使用的 BAMCU Agent 出现意外事件(例如 BAMCU 所处的机器死机,或所处的网络出现故障)时,同一多媒体交互会议中的节点 Agent 通过通信和协商,将能发现问题所在,并向 MIMServer 重新发送请求,查询获得别的可用 BAMCU,使得多媒体交互能够不因局部的节点问题或网络问题而瘫痪。

3. MIEBA 中的多媒体数据传送机制实现

MIEBA 中,客户 BAMClient Agent 间进行多媒体交互时,可以通过和 MIMServer 的交互动态寻找多媒体数据转发路径。因而模型中多媒体数据传送机制的算法是该模型的一个重要组成部分。

BAMClient Agent 间的多媒体数据流是采用 UDP 协议发送的 RTP 包。两个 BAMClient Agent 间存在如下三种关系:

- 1)两个 BAMClient Agent 属于同一网段,它们之间的多媒体数据可以通过 UDP 组播直接实现。

- 2)两个 BAMClient Agent 分属不同的网段,并且根据网络状况和会议分组情况,两个节点又采用了同一个 BAMCU Agent,这时两者间的多媒体数据传输必须通过 BAMCU 的直接转发才能完成。

- 3)两个 BAMClient Agent 分属不同的网段,并且在选择多点控制单元时选择了不同的 BAMCU,这时两个客户 Agent 间的通信必须通过两个 BAMCU 的间接转发才能完成。

因此,任意两个 BAMClient Agent 间的多媒体数据流,最多经过两个 BAMCU Agent 节点的转发即可完成。

多媒体交互环境中,任意两个客户节点间要实现多媒体数据传输,必须首先确定多媒体传输的方式和传输经过的 BAMCU Agent 路径。这些工作是通过 MIMServer Agent 和 BAMClient 的协商交互完成的。由于多媒体数据流的传输采用的是 UDP 方式,因此在确定了本节点的 BAMCU Agent 后,传送路径上的各 Agent 还必须根据会议组成情况,动态分配多媒体数据流的 UDP 发送、接收端口以及负责多媒体发送、接收线程的控制。MIEB 中多媒体转发的 UDP 端口分配和转发过程描述如下:

一个 BAMClient 选择了 BAMCU,如果同一个多媒体会议中,已经有其它的 BAMClient Agent 建立了通过该 BAMCU 的多媒体数据传输,则该建立该节点多媒体数据的传输和接收转发过程如下:

- a)数据发送 如果 BAMClient 和 BAMCU 在同一物理网段,则 BAMClient 在本网段内组播多媒体数据;如果 BAMClient 和 BAMCU 不在同一物理网段,则 BAMClient 向 BAMCU 定向放送数据。当本会议组中,其它节点采用的 BAMCU 的个数和本节点对应的 BAMCU 上本组远程节点的个数大于0时,则本节点对应的 BAMCU 开始接收数据。而

对接收到的数据的进一步转发处理过程如下:

- 如果本会议小组内还有其它 BAMClient 使用了同一 BAMCU,则 BAMCU 向这些节点转发数据。

- 如果本会议小组内有其它 BAMCU 存在,则为本 BAMCU 分配向其它 BAMCU 转发数据所需要的转发端口并开始转发数据。其它, BAMCU 对接收到的数据根据相应的 BAMClient 的性质不同分别启动本地组播和远程定向发送数据进程。

b)数据接收 如果同一多媒体会议小组,除去自己外还有别的节点使用该 BAMCU 来转发多媒体数据,则所需做的只是把 BAMCU 接收到的其它 BAMClient 节点的多媒体数据转发过来即可。

如果同一多媒体会议小组中,只有新加入节点自己采用该 BAMCU 来转发多媒体数据,那么还必须把其它所有节点的多媒体数据首先转发到新加入节点对应的 BAMCU,然后再转发到新加入节点 BAMClient。

4. 应用例子

采用本文提出的基于 Agent 的 CSCW 多媒体交互环境模型,我们实现了一个支持多媒体交互的协作设计系统——CoopDesigner。CoopDesigner 系统提供了一个包括多媒体交互环境、电子白板、混合机制应用共享等协作工具的基本协作环境,用户可以在这个环境中根据自身需求灵活裁减和配置

系统组成。CoopDesigner 系统向用户提供了基于 IP 网络的全双工视听交互功能,并且通过 Agent 间的协商交互,系统可以动态确定多媒体数据的转发路径,实现跨物理网段的多媒体数据的多点传送。

结论 多媒体交互是 CSCW 中的一种重要的交互协作方式。但现有 CSCW 的多媒体交互环境中对多媒体数据的跨网段传输采用的基本都是静态的 MCU 分布结构,使得当局部 MCU 节点出现故障后,不能及时建立新的多媒体数据的传送通道。而我们提出的基于 Agent 的多媒体交互模型中,多媒体数据转发路径是通过 Agent 间的协商交互动态建立的,当局部多点控制单元出现问题后,系统可以迅速建立新的多媒体转发路径,因而基于该模型的系统具有较强的可靠性和适应性。

参 考 文 献

- 1 Grudin J. Computer-supported cooperative work: History and focus. IEEE Computer, 1994, (5): 19~26
- 2 郑庆华. CSCW 的理论、技术与实现: [博士论文]. 西安: 西安交通大学, 1997
- 3 倪强, 朱光喜. 计算机支持下的协同工作的研究现状综述. 计算机工程与应用, 2000, 36(4)
- 4 Finin T, Fritzson R, et al. KQML as an agent communication language. In: the Proc. of the third intl. conf. on information and knowledge management, ACM press, Nov. 1994

(上接第56页)

时多媒体业务成为网络进一步的研究重点。当前存在的问题主要是: 如果不考虑网络的实际容量而大量发送非拥塞控制的 UDP 业务, 很容易导致网络过载和高丢失率, 严重时会引起“拥塞崩溃”现象; 非拥塞控制业务还会大量侵占一丢包就减少发送速率的 TCP 业务的带宽, 有时会导致 TCP 协议根本不能建立连接, 更谈不上传送数据了。因此, Sally Floyd 等提出 TCP 友好(TCP-friendly)的概念, 并由此产生了一系列 TCP 友好控制机制。本文从实现方式的角度, 将这些 TCP 友好控制机制总结成几类, 并分析了它们的优势和不足。通过文中分析, 我们可以看到基于 AIMD 和基于建模的控制机制各有利弊, 但如果将两者结合, 可以更好地把握流量的控制力度, 既实现了 TCP 友好的流量控制, 又满足了实时媒体流的性能要求。从另一方面来说, 端到端的机制与逐跳的机制也是这样。端到端的机制对业务端到端的流量进行控制, 提供 TCP 友好的流量特性, 并保障实时业务的质量, 但它对网络拥塞的反应较慢, 控制不及时; 逐跳的控制机制对网络内部的拥塞状况了如指掌, 采取控制时快速、准确, 但对业务流的端到端性能不能很好地把握, 也不能完全实现 TCP 友好的控制。因此, 最佳的方案是将这几种方式结合起来, 提出全局 TCP 友好的控制机制, 这样, 不仅可以更好地实现网络拥塞控制, 对于实时媒体流而言, 还能提供 TCP 友好的流量特性, 确保网络响应流与非响应流间的公平性, 保障网络的总体性能。

参 考 文 献

- 1 Floyd S, Fall K. Promoting the use of end-to-end congestion control in the Internet. IEEE/ACM Trans. On Networking, 1999, 7

- (4)
- 2 Floyd S. Congestion control principle. IETF draft (draft-floyd-cong-02.txt), March 2000
- 3 Bansal D, Balakrishnan H. TCP-friendly congestion control real-time streaming application. [MIT technical report]. May 2000
- 4 Floyd S, Handley M, Padhye J. A comparison of equation-based and AIMD congestion control. May 2000. URL <http://www.aciri.org/tfrc>
- 5 Rejaie R, Handley M, Estrin D. RAP: an end-to-end rate-based congestion control mechanism for realtime streams in the Internet. In: Proc. of INFOCOMM'99, volume 3, March 1999
- 6 Sisalem D, Schulzrinne H. The loss-delay based adjustment algorithm: a TCP-friendly adaptation scheme. In: Proc. of NOSSDAV'98, Cambridge, England, July 1998
- 7 Padhye J, et al. Modeling TCP throughput: a simple model and its empirical validation. In: Proc. of SIGCOMM'98, Aug. 1998
- 8 Floyd S, Handley M, Padhye J. Equation-based congestion control for unicast application. In: Proc. of ACM SIGCOMM 2000, Aug. 2000
- 9 Padhye J, et al. A model based TCP-friendly rate control protocol. In: Proc. of NOSSDAV'99, June 1999
- 10 Sisalem D, Emanuel F, Schulzrinne H. The direct adjustment algorithm: a TCP-friendly adaptation scheme. August 1997. URL: <http://www.focus.gmd.de/usr/sislem>
- 11 Mishra P, Kanakia H. A hop by hop rate-based congestion control scheme. In: Proc. on Communications Architecture & Protocols. Baltimore, MD USA. Aug. 1992
- 12 Sisalem D, Schulzrinne H. ALS: an ABR-like service for the Internet. In: Proc. of ISCC'2000, France, July 2000