

ГОЛОСОВОЕ И ВИЗУАЛЬНОЕ ОБЩЕНИЕ В ОНЛАЙНЕ

О.В. МАТЮШЁНОК¹, В.С. ОСКЕРКО²

¹ - студентка 1 курса, факультета МЭО, группы УВЭД-5, Белорусского государственного экономического университета

² - научный руководитель, доцент кафедры информационных технологий, Белорусского государственного экономического университета, Минск, 220672, Партизанский пр., 26, тел. (8017) 249-19-81, e-mail: oskerko@bseu.minsk.by

Аннотация: Общение является одним из важнейших компонентов нашей повседневной жизни. Поэтому тема общения всегда будет оставаться актуальной. В данной статье описаны основные способы общения в сети. Это голосовые чаты, Интернет-вещание, видеоконференции, IP-телефония и VoIP. В статье также дан обзор программных продуктов, обеспечивающих общение в онлайне.

Ключевые слова: Общение в онлайне, e-mail, Интернет-вещание, IP-телефония, голосовые и видеоконференции.

1. ВВЕДЕНИЕ

Людям необходимо общение. И если личная встреча затруднена, обходится слишком дорого из-за большого расстояния или неудобна еще по каким-либо причинам, *на помощь* чаще всего *приходят телефон, e-mail, Instant Messenger.* Сеть предлагает удобные и недорогие способы общения для сотен миллионов людей во всем мире.

Без электронной почты сегодня не обходится практически ни один пользователь РС. Текстовый обмен решает множество проблем, ведь обычно требуется всего несколько секунд, чтобы сообщение облетело половину земного шара и нашло своего адресата.

После e-mail быстрое развитие получили технологии Instant Messaging (IM) – «сверхбыстрая почта». С появлением IM у пользователей появилась возможность ведения интерактивного диалога в реальном времени с контролем статуса (presence management) собеседника, т.е. возможностью быть уверенным, что пользователь получил сообщение на основе информации о его онлайн-статусе. Таких возможностей e-mail не дает.

Но продолжала существовать фундаментальная сложность текстового общения в реальном времени. Выразить мысль в виде текста при дефиците времени не всегда удобно — надо записывать все, что хочешь сказать. *Оптимальным способом общения* в реальном времени, конечно, *является голос*, а не текст. Кроме того, текстовым IM не достает эмоциональности. Недаром сразу после появле-

ния электронного текстового обмена люди стали использовать смайлики и прочие символы выражения настроения. Тем не менее текстом невозможно описать то, что мы передаем в обычном разговоре интонацией или мимикой и жестами.

Неудивительно, что популярность сетевых программных продуктов для коммуникации растет невиданными темпами. Технологии быстро сменяют друг друга, и если еще год назад мало кто верил, что голосовое общение по сети быстро станет реальностью, то это мнение кардинально изменилось с возникновением Yahoo! Messenger и Windows Messenger, которые предлагают естественное общение с помощью голоса, а не только обмен текстовыми сообщениями. *Голосовые чаты, Интернет-вещание, IP-телефония и VoIP (Voice over IP — технология передачи речи по пакетным сетям)* быстро внедряются в нашу компьютерную жизнь.

2. ГОЛОСОВЫЕ И ВИДЕОКОНФЕРЕНЦИИ

2.1. Принципы передачи аудио- и видеoinформации через сеть

Следующим шагом онлайн-общения стало появление голосовых и видеоконференций. С помощью недорогой цифровой камеры и микрофона пользователь может проводить вещание в Интернете. Однако для более или менее качественного видеосеанса нужен не только компьютер с подключенной видеокамерой, но и высокоскоростной канал

в Интернете. Сегодня большинство новых компьютеров производители стали комплектовать аппаратной поддержкой коммуникаций, звуковыми и видеокартами, микрофоном, колонками и даже Web-камерами. Платы расширения обеспечивают ускоренное кодирование/декодирование видео- и аудиосигналов. Такой прогресс в техническом оснащении стандартных бытовых PC вызвал появление *голосовых конференций (чатов)*, в которых пользователи могли общаться между собой, не набирая текст, а говоря через сеть как по телефону. Одновременно с этим появились услуги по проведению *видеоконференций*, примерами которых могут служить такие продукты, как CU-SeeMe, Microsoft NetMeeting и др. В обоих случаях принцип передачи аудио- и видеoinформации — схожий. Она компрессируется, пакетируется и стандартными протоколами передается через каналы Интернет-адресатам.

На входе адресата информация обрабатывается в соответствии с протоколом, декомпрессируется и преобразуется в стандартный звуковой и/или видеосигнал. Казалось бы, Web-камера дала практически все необходимое для полноценного общения и миг всеобщего ликования близок. Но трудности подстерегают сразу — как только пользователь подключает камеру к своему компьютеру. Для полноценной видеоконференции необходим канал с высокой пропускной способностью, иначе вы получите видеоокошко в 1/16 экрана и смену кадров со скоростью 2-3 в секунду. Это скорее напоминает фотографии, чем живое видео. Легко подсчитать, что стандартная камера, подключенная к USB-порту компьютера и способная выдавать до 30 кадров в секунду, формата 160x120, RGB, в секунду пытается передать около 7 Мбайт информации. В случае недостаточной пропускной способности канала стандартные видеоконференции прореживают видеосигнал, предоставляя максимальный приоритет звуку. В результате получается аналог замедленного и одновременно плохо смонтированного (с разрывами) кино. На обычной же телефонной линии (dial-up, 28 Кбит/с) даже один голос плохо проходит, в сети большие задержки, прерывания, шумы, эхо, а о видео вообще говорить не приходится. Кроме того, здесь есть еще целое множество подводных камней. Например, регламент общения звуковых чатов, особенно в

случае, когда пользователей более двух и требуется соблюдать определенный этикет, чтобы каждый смог услышать всех.

2.2. Конфиденциальность в сети

Фундаментальное ограничение, однако, состоит в том, что *пользователи стремятся контролировать, что передается в сеть*, и часто не хотят передавать всю информацию о себе и окружающем мире. Например, вам звонят по делу домой рано утром, а вы только что вышли из ванной. Или шеф вызывает вас на видеосвязь в офисе, а у вас на рабочем месте беспорядок, да еще в гостях — знакомые. В этих и многих других подобных случаях полный видеосеанс вам будет явно не на пользу. Ведь хотелось бы, чтобы шеф всегда видел вас в строгом деловом костюме, а приятель не шутил над вашим дежурным галстуком, как бы вы ни выглядели на самом деле.

Можно ли себе представить, что *современные средства цифровой обработки аудио- и видеoinформации позволят убрать лишние детали?* Можно ли, «разобрать на части» голос и изображение, затем убрать или заменить все то, что считается конфиденциальным или неудобным, а далее «собрать» и визуализировать/озвучить оставшееся? Это возможно уже сегодня!

Техническая возможность обнаруживать и выделять лицо, находящееся в поле зрения камеры (на любом двигающемся фоне), и возможность выделять голос говорящего в условиях фоновых шумов (музыки, разговоров и пр.) может служить подспорьем в области онлайн-коммуникаций.

Что же нам сможет дать выделенное на изображении лицо? Во-первых, можно определить пространственную ориентацию головы, включая углы наклона и поворота, сдвиг, приближение-удаление. Во-вторых, можно получить характеристики мимики. В-третьих, можно заменить фон, который по каким-либо причинам не устраивает говорящего. В результате получается описание пространственного положения лица, взаимного расположения черт лица и мимических признаков в виде некоторого вектора состояния, занимающего объем памяти не более чем 200 байт.

Таким образом, мы в сотни раз можем снизить объем информации, которую необ-

ходимо передать по коммуникационным сетям, и создаем принципиально новый способ управления. Кроме того, в руках самого пользователя (а не производителя программного обеспечения) будет находиться тот самый «механизм управления конфиденциальностью», о котором говорилось выше.

Похожая ситуация наблюдается и со звуком. Различия чисто технологические — это распознавание фонем для синхронизации движения губ, подавление акустического эха, искажение (преобразование) голосового тембра или подмена одного тембра голоса на другой (*voice disguise* — технология, позволяющая изменить голос до неузнаваемости). Например, можно говорить своим обычным голосом на входе, а на выходе будет звучать голос Мадонны или Буша-младшего.

Иными словами, возможности современных технологий анализа видеосигнала и голоса настолько широки, что в основном ограничиваются лишь качеством и мощностью используемых аппаратных средств. Более того, ряд компаний уже предлагает такие возможности на обычных PC — например, *российская SeeStorm* и *американская Eye-matic*, которые ведут разработки систем управления анимацией виртуальных персонажей на основе анализа аудио- и видеопотоков в реальном масштабе времени. Пока разрабатываемые технологии весьма оригинальны, и поэтому фирмы тщательно оберегают рецепты своего успеха, демонстрируя лишь анимационные рекламные ролики.

Именно задачу «*достоверного присутствия*» с использованием *трехмерных моделей* и ставят перед собой разработчики новейших коммуникационных систем. Уже сегодня коммуникационные программные продукты, ориентированные на массового пользователя и использующие трехмерные персонажи, работают по принципу управления голосом. Система анализирует голос человека, распознает звуки, которые он произносит, и на принимающем конце проигрывает этот голос наряду с трехмерным изображением лица человека, мимика и движения губ которого соответствуют произносимым словам в реальном времени. Результатом является вполне реалистичный эффект присутствия собеседника.

Желание людей использовать трехмерные персонажи для общения вполне понятно — люди получают *новое средство визуальной*

коммуникации, которое не требует высокоскоростных каналов, сохраняет эмоции и оберегает личную жизнь (устраняет лишние детали), передавая только лицо человека. Ряд компаний уже предлагает свои услуги по созданию таких трехмерных персонажей по фотографии для широкого рынка, например австралийская BioVirtual (программный продукт SDMeNow), английская Digimask и российская SeeStorm (программный продукт AvatarMe). Результаты работы продуктов SDMeNow и AvatarMe достаточно схожи и позволяют непрофессиональному пользователю в короткое время создать трехмерный персонаж по фотографии.

3. IP-ТЕЛЕФОНИЯ

3.1. Принципы работы IP-телефонии

Вернёмся ещё раз к голосовому общению в сети. Чтобы передать голос через Интернет, присоединяем к компьютеру микрофон, оцифровываем звук, нарезаем его небольшими кусками, оформляем в IP-пакеты и отправляем их по Интернету. На другом конце содержимое пакетов склеиваем и воспроизводим через колонки или наушники, подключенные к звуковой карте. То же самое делаем с ответными репликами. Что ж, в теории все просто, но на практике возникает масса непредвиденных осложнений.

Первое из них не является, правда, особенно неожиданным. Давайте подсчитаем объем передаваемых данных. При частоте оцифровки 8 кГц и разрядности 8 бит получаем 64 Кбит/с. Именно этот подсчет определил пропускную способность каналов, заложенную в стандарт ISDN. Но сегодня такой подход считается крайне расточительным.

Прорыв наметился с появлением *алгоритмов, позволяющих эффективно сжимать поток аудиоданных*. По аналогии с модемами (модулятор-демодулятор) такие программы стали называть *кодеками* (кодер-декодер). Широко распространенные алгоритмы сжатия (тот же MP3, к примеру) позволяют уплотнять аудиоинформацию в 10 и более раз.

Конечно, звук при этом искажается, однако подбором параметров кодека можно добиться вполне удовлетворительного качества. Объем же передаваемых данных заметно уменьшается.

Современные кодеки, разработанные специально для задач IP-телефонии, используют гораздо более изощренные алгоритмы, основанные на изучении особенностей человеческого слуха и голоса, фонетики различных языков. С их помощью можно добиться приемлемого качества передачи голоса на скорости всего 2,4 Кбит/с. То есть по обычному модемному каналу 33,6 Кбит/с можно передавать сразу 14 (!) телефонных разговоров. Стоимость интернет-трафика оказывается просто несопоставимо меньше стоимости телефонных разговоров, которые с его помощью можно передать. На этом-то различии и основывается *главная бизнес-идея IP-телефонии*. Именно эта несопоставимость позволяла мечтать о грядущих сверхприбылях в данном секторе высоких технологий. Однако от технической идеи до работающего бизнеса - долгий путь.

Прежде всего требовалось навести порядок в использовании кодеков. Ведь как бы хороши ни были патентованные алгоритмы, пока они не получают широкого распространения, пользы от них будет немного - абоненты с разными кодеками будут просто не в состоянии услышать друг друга. Разными фирмами было разработано множество несопоставимых друг с другом кодеков. Одни стремились обеспечить максимальное качество звука, другие - увеличить степень сжатия, третьи - повысить устойчивость к сбоям. Относительный порядок стал устанавливаться только в 1996-1998 годах с принятием стандарта H.323, определяющего правила передачи мультимедийной информации по пакетным сетям.

3.2. Проблемы IP-телефонии

Теперь самое время поговорить о пакетной передаче данных. *Протокол IP*, лежащий в основе Интернета, как известно, *не гарантирует доставку IP-пакетов* от отправителя к получателю. При возникновении ошибки или задержки пакет-неудачник просто уничтожается без уведомления как отправителя, так и получателя. Кроме того, пакеты порой доставляются адресату разными маршрутами и в результате поступают к нему не в том порядке, в каком отправлялись. При передаче файлов за доставкой пакетов и их порядком следит вышестоящий протокол TCP. Однако использовать его для передачи

голоса невозможно, поскольку, обнаружив нехватку пакета, TCP просто приостановит передачу данных и начнет выяснять отношения с отправителем. В разговоре возникнет перерыв, который для абонентов гораздо хуже безвозвратной потери 0,1 с звучания.

Таким образом, *для нормальной передачи голоса через пакетную сеть необходим свой транспортный протокол* (например, RTP/RTCP - Real-time Transport [Control] Protocol или RTSP - Real-Time Streaming Protocol), сохраняющий работоспособность даже в условиях частичной потери передаваемой информации.

Если исходный звуковой поток будет просто «нарезан» на кусочки, скажем, по 0,1 с, то *потери пакетов* будут приводить к коротким «выпадениям» *в разговоре*. Перетерпеть их, конечно, можно, но все же они очень неприятны. Чтобы обойти эту проблему, информацию о каждом таком коротком интервале «размазывают» по нескольким соседним пакетам. В случае утраты одного или даже нескольких пакетов разрывов в передаваемой речи не возникнет, просто несколько снизится качество звукопередачи. Причем нередко это сказывается лишь на тембре голоса - появляется характерная «металлическость», - но не мешает понимать произносимые слова.

Еще одна проблема связана со *скоростью передачи пакетных данных*. В традиционной телефонии скорость передачи сигнала от одного абонента к другому ограничена лишь скоростью света, и при длине экватора 40 тыс. км задержка сигнала редко превосходит 0,1 с. Если же пакеты время от времени задерживаются в пути, кодеки будут увеличивать задержку до 300-500 мс, а то и больше, чтобы отстающие пакеты успевали добраться по назначению к моменту их очереди на воспроизведение. Задержка свыше 100 мс уже заметна для говорящих, а если она достигает полусекунды, то ощущается дискомфорт.

Причина задержек обычно кроется в перегрузке каналов связи или обслуживающих их маршрутизаторов. В Интернете нагрузка может меняться в широких пределах как из-за естественных ритмов жизни планеты - суточных, недельных, сезонных, - так и в результате непредвиденных событий - технических аварий, хакерских атак, ажиотажных нагрузок.

Если оператор IP-телефонии станет полагаться на использование для связи обычного интернет-трафика, обеспечиваемое им качество обслуживания будет напрямую зависеть от неподконтрольной ему погоды в Большом Интернете. Решить эту проблему можно, передавая телефонные разговоры между узлами пакетной телефонной сети по соединениям. Кроме того, такие прямые каналы уменьшают общую задержку доставки пакетов за счет сокращения числа транзитных узлов, производящих маршрутизацию.

Итак, применение кодеков с высокой степенью сжатия, специальных протоколов и выделенных каналов связи обеспечивает достаточно качественную передачу голоса по пакетным сетям. Однако если абонентским устройством будет оставаться компьютер, вряд ли новую технологию ждет блестящее будущее.

И дело даже не в малой по сравнению с телефонами распространенности компьютеров. Просто по своим функциональным характеристикам даже самый мультимедийный компьютер пока остается менее удобным для голосового общения, чем обычный телефонный аппарат. Компьютер может подавать звуковые сигналы, воспроизводить музыку или аудиодорожку к принимаемому видеосигналу. Все эти аудиопотоки привычно слышать из колонок. В то же время мало кто чувствует себя комфортно, пользуясь громкой связью при телефонных разговорах. Постоянное же ношение наушников мешает обычному общению, будь то в офисе или дома. Ничего удобнее телефонной трубки пока все-таки не придумано. А компьютер, приспособленный к использованию в качестве телефона, по удобству пока находится на уровне старинных аппаратов, состоявших из отдельных телефона и микрофона, держать которые приходилось в обеих руках.

Пока не решены эргономические проблемы и пока IP-телефония не добилась доминирующего положения в мире, *интеграция с сетями традиционной телефонии* является для IP-телефонии вопросом жизни и смерти. Интеграция эта должна обеспечивать не только преобразование цифрового сигнала в аналоговый и обратно, но и возможность передачи вызова с компьютера на телефонный аппарат, с телефонного аппарата на компьютер и, что самое важное, с одного телефонного аппарата на другой посредством пакетных

сетей передачи данных.

3.3. Деньги и IP-телефония

IP-телефония может приносить большие прибыли, т.к. у неё есть много преимуществ. Во-первых, это значительно *более высокая ёмкость системы*. Проложить оптоволоконный кабель ничуть не сложнее, чем обычный телефонный, а возможностей у него на несколько тысяч одновременно разговаривающих абонентов. Во-вторых (и это главное), *инфраструктуру, необходимую для IP-телефонии, можно с успехом использовать и для любых других цифровых коммуникаций*. И наоборот, естественно. Таким образом, кабели только для IP-телефонии никто не тянет. Эта технология неплохо себя чувствует в крупных корпоративных IP-сетях и там, где развиты Интернет-коммуникации.

Итак, компания, имеющая доступ к высокоскоростным каналам цифровой связи, может прийти на рынок IP-телефонии практически «на всем готовом». Все, что ей остается, - это закупить оборудование, наладить его и ждать клиентов. А первыми клиентами будут как раз компании, которым тоже хочется поучаствовать в дележе IP-пирога, но которые не обладают десятками оптоволоконных каналов. Типичным примером их деятельности является перепродажа в регионах услуг первичного провайдера технологий IP-телефонии. Цены на эти услуги, как правило, устанавливаются существенно ниже тех, что существуют в телефонии традиционной. «Существенно» - это в два-три раза. Срок мучительной смерти традиционной телефонии от невостребованности отодвигают лишь отмеченные выше недостатки новой технологии, да еще тот факт, что IP-телефония - дело пока новое и рядовому пользователю непривычное. Хотя, говорить о каком-либо бизнесе в области IP-телефонии можно лишь в случае соединения «телефон-телефон». Проекты с использованием соединений «телефон-компьютер», «компьютер-телефон» и «компьютер-компьютер» существуют и даже иногда работают, но лишь в виде эксперимента. Лет через пять все это благополучно канет в Лету - использовать дорогостоящие линии связи только для телефонных переговоров расточительно уже сегодня.

Очень показателен пример Молдавии, где всего за год в одном только Кишиневе число

пользователей IP-телефонии увеличилось на 60 тысяч человек. Популярность новой технологии в этой стране достигла такого уровня, что ею, к сожалению, всерьез заинтересовалось государство. К сожалению - потому что, заметив надкушенную часть телефонного пирога, министерство транспорта и связи Молдавии просто взяло и запретило негосударственным компаниям оказывать услуги международной IP-телефонии. Впоследствии это постановление было отменено, но ценой драматической закулисной борьбы и выплаты «дотаций» государственному монополисту в области телекоммуникаций - «Молдтелекому».

3.4. Конструктор для IP-телефонии

Простейший абонентский комплект для работы с IP-телефонией - это компьютер, имеющий звуковую плату и колонки и подключенный к Сети. Понятно, что полноценной телефонией такой контакт назвать нельзя, ведь оба абонента обязательно должны иметь компьютер, подключенный к Сети, а это существенно сужает сферу общения.

Второй вариант (компьютер-телефон) уже требует соответствующей поддержки со стороны оператора связи: если вызывающий абонент по-прежнему подключен к Интернету, то его собеседник отвечает с помощью обычного телефонного аппарата.

Клиентским решением профессионального уровня (соединение телефон-телефон) с набором телекоммуникационных услуг, таких как определение номера, принятие двух звонков на одну телефонную линию одновременно, перевод звонка, организация телефонных конференций и т. д., может считаться телефонный аппарат стандарта H.323 с интерфейсом Ethernet (RJ-45) либо телефонная PC-карта, которая делает компьютер, подключенный к IP-сети, совместимым с обычным телефонным аппаратом. Существуют модели, позволяющие подключать к одному устройству до 6 телефонных линий. Такие аппараты целесообразно использовать при создании телефонной сети в организациях, имеющих развитую сетевую инфраструктуру и непосредственный выход к магистральным каналам операторов телекоммуникаций.

4. ЗАКЛЮЧЕНИЕ

По данным экспертов, ежегодный прирост мирового рынка коммуникационных услуг должен составить не менее 30%. Провайдеры Интернет-видеоконференций, такие как WebEx Communications, Net2000 Communications, предоставляющие услуги через стандартные Web-браузеры, постоянно заявляют об увеличении пользователей и росте доходов. Голосовые чаты HearMe, PalTalk, ChatVoice пользуются растущей популярностью. А ведь это утилитарные способы общения, ограниченные как пропускной способностью каналов, так и слишком реалистичным отображением общения, сдерживающими свободу собеседников.

Один из первых программных продуктов на мировом рынке, ориентированный на массового пользователя и доступный уже сегодня (причем бесплатно) для визуального и голосового общения по стандартным телефонным каналам (dial-up, 28 Кбит/с), — продукт российской компании SeeStorm (www.see-storm.com).

И среди IP-телефонов встречаются достаточно неординарные разработки. В частности, компанией Symbol предложена модель *портативного аппарата NetVision Data*, который соединяется с телефонным шлюзом (от Cisco или Nortel) посредством протокола беспроводного доступа IEEE 802.11b и имеет в числе прочих достоинств 40-разрядные механизмы шифрования трафика.

Программные продукты для голосовых и видеокommunikаций в сети, использующие сложнейшие алгоритмы распознавания голоса и видео, уже покинули лаборатории и очень скоро войдут в обиход миллионов людей во всем мире.

ЛИТЕРАТУРА

1. Прохоров А., Букина В. Общение в онлайне. Сегодня и завтра.// Компьютер Пресс, №1, 2002, с.45-49.
2. Сергеев А., IP-телефония.// Мир Internet, №11, 2001, с.18-23.
3. Поваляев Е., Разговор с компьютером.// Мир Internet, №6, 2001, с.172-177.
4. Темерев А., День X, или немного о правильной почте.// Мир Internet, №11, 2001, с.64-65.