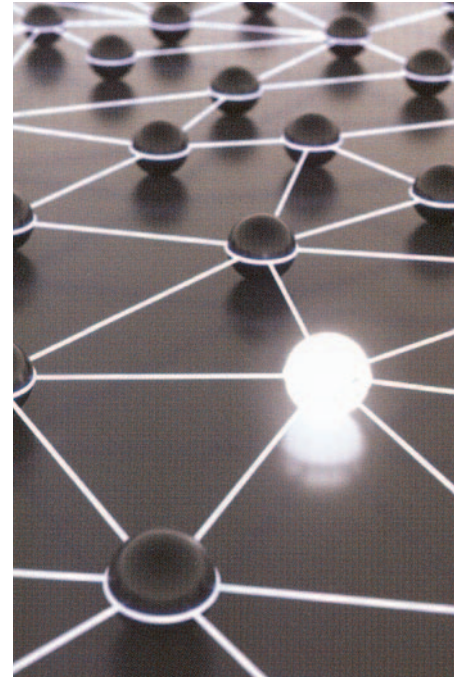


Ольга Жернакова

# Системы рекомендаций и поиска контента

Количество телеканалов, которые может принимать абонент платного ТВ, исчисляется сотнями. Библиотеки видео по запросу содержат тысячи фильмов и передач. Если телевизор подключен к интернету, дополнительно появляется доступ к неисчислимому количеству видеороликов и каналам web-ТВ. Однако, что парадоксально, абоненты по-прежнему жалуются, что им нечего смотреть. Чтобы помочь им найти то, что понравится, операторы задумываются о введении системы поиска и рекомендаций. Задача оператора при этом не только улучшить впечатление абонентов от услуги и уменьшить их отток, но и увеличить продажи видео по запросу и премиальных телеканалов.



На рынке существует множество систем поиска и рекомендаций. Некоторые из них — совсем новые, некоторые существуют уже лет десять. Но все они решают более-менее одинаковые задачи. Сначала нужно понять, что нравится зрителю, потом — найти соответствующее видео, потом его предложить — причем так, чтобы пользовательский интерфейс никого не напугал, а управление осуществлялось несколькими кнопками пульта дистанционного управления. Ну, и завершающий этап — оценить эффективность и посчитать доходы.

## Сбор информации о видео

Теоретически в каждом видео — как в отдельном файле библиотеки видео по запросу, так и в потоке цифрового телеканала, хранится некая дополнительная информация о видео — так называемые метаданные. Что-то добавляется автоматически — например, камера отмечает время съемки, часть заполняют монтажеры и редакторы телеканала. Для чтения и редактирования метаданных существуют специальные программы. Помимо встроенных в транспортный контейнер метаданных информация о видео может храниться отдельно — например, при размещении

файла на видеохостинге обычно создается текстовое описание и ключевые слова, а при передаче телеканала оператору редакторы отдельно составляют файл с программой передач с их краткими описаниями для включения в EPG.

Однако единообразия в подходах нет, ошибок в этой информации тоже хватает, и в результате использовать такой уже имеющийся у оператора бесплатный источник информации как метаданные и EPG для системы рекомендаций оказывается недостаточным.

Компании, реализующей систему поиска и рекомендаций, приходится искать надежный источник информации и заполнять описания самостоятельно. Иногда почти всю работу делает сам оператор. Например, в Рикоре (Активное ТВ) для заполнения описаний передач работает специальная команда модераторов.

Человека можно попытаться заменить алгоритмами искусственного интеллекта, который будет также изучать доступные источники информации. В израильской компании Jinni по имеющимся в открытом доступе текстовым описаниям умеют отбирать список из 40-60 ключевых слов, которые характеризуют настроение видео, его сюжет и стиль.

В промежуточном варианте оператор может получать такие автоматические данные и вносить какие-то правки и добавления — например, в решении TiVo оператор может выделить специального редактора, который отбирает самые интересные передачи определенного направления.

Но откуда, собственно, взять все эти текстовые описания и градации по жанрам?

## Дополнительные текстовые источники

Лучшим источником информации о фильмах и телепередачах считаются специализированные журналы с целым штатом редакторов. В Англии это, например, IPC Media — издающий аж пять разных журналов с программкой телепередач, каждый для своей аудитории. Именно данные IPC Media используются в системе поиска и рекомендаций в сети Fetch TV IPVision. В Германии есть издательство Axel Springel, тоже печатающее журналы-программки и предоставляющее базы с метаданными для разработчиков поисковых систем; с этим издательством работает компания Aprico, предоставляющая свое решение для системы поиска и рекомендаций. Решение ROVI использует электронную программу передач от другой известной европейской компании, Informedia из Люксембурга. Когда-то на заре

Полная версия статьи выложена на сайте [telemultimedia.ru](http://telemultimedia.ru)

создания Теле-Спутника мы вели с этой компанией переговоры о получении программы передач западных спутниковых каналов.

Для видео по запросу и популярных сериалов в качестве источника информации хорошо подойдут базы данных фильмов в интернете. Например, известнейшая база на сайте [imdb.com](http://imdb.com) содержит почти 2 миллиона описаний фильмов и телепередач, и информация динамически обновляется. Но не стоит думать, что пользоваться ими можно бесплатно. База доступна операторам кабельного ТВ и прочим желающим за плату от 15 тысяч долларов в год.

Есть более экзотические решения. Например, можно распознать текст аудиодорожки или взять уже имеющиеся субтитры. Для того чтобы справиться с полученным потоком данных, потребуются специальные алгоритмы, использующие информацию о семантике текста (такая работа ведется в BBC).

Таким образом, у оператора есть несколько возможных источников информации, и очень часто они используют сразу несколько.

После того как оператор собрал описание передач и фильмов — как в библиотеке видео по запросу, так и для линейных телеканалов, абоненту можно дать возможность фильтровать передачи по жанру, любимым актерам, ключевым словам, длительности показа, времени трансляции и т.п. Но как все-таки достигнуть того, чтобы абоненту почти ничего не нужно было делать, а система магическим образом догадывалась бы, что ему нравится?

### Сбор информации об абоненте

Современные телеприставки и телевизоры лишь немногим уступают компьютерам и могут собирать информацию о каждом действии (нажатии кнопки пульта) абонента. Если есть обратный канал, оператор может анализировать эту информацию, но, вообще говоря, это не обязательно — приложение может работать и локально. Например, в решении Актив ТВ информация о предпо-

чтениях абонента хранится на приставке. Если у оператора есть обратный канал, он может собирать информацию у себя. В целом ряде иностранных проектов платного ТВ операторы собирают статистическую информацию о просмотрах конкретных передач — и в рамках линейного вещания, и в варианте видео по запросу. На основе этой информации можно сделать так называемый статистический collaborative filtering. В самом простом варианте отбирается самое популярное видео, которое и предлагается абонентам в виде отдельного списка, отсортированного по жанрам и категориям. В более сложном варианте строятся графы — Байесовские сети, описывающие статистическую связь между разными элементами контента. Они дают возможность предсказать, какова вероятность, что пользователь, посмотревший данную передачу, посмотрит и какую-то другую. Эти графы можно передавать от оператора к оператору, тем самым снижая время, необходимое для запуска услуги. Особенно хорошо они работают в случае сериалов или повторов передач. Но для совсем новых программ схема не работает — их нужно добавлять в диаграмму на основе каких-то других данных и эвристических алгоритмов. В качестве основного элемента системы поиска и рекомендаций подобные диаграммы на базе паттернов просмотра используют такие производители систем поиска и рекомендаций, как TV Genius (ныне — подразделение компании Red Bee) и Gravity.

Плюсом подобного подхода является то, что можно использовать уже построенные другим оператором графы и не покупать дополнительные метаданные помимо того, что есть в EPG, так как абоненты фильтруют передачи сами. Персональная информация об абоненте тоже может не собираться — достаточно общей статистики просмотров всеми абонентами сети. Наиболее известный пример действия passive collaborative filtering в интернете — это окно «пользователи, которые смотрят этот фильм, смотрят также вот это» на [amazon.com](http://amazon.com) или [ozon.ru](http://ozon.ru).

У collaborative filtering есть и противники. Например, специалисты ThinkAnalytics считают, что этот подход не позволяет добавить новую передачу кроме как вручную силами оператора — а кредо этой компании состоит в том, что оператор не должен напрягаться. Вместо этого компания предлагает predictive modelling — предсказание выбора пользователя на основании статистических исследований и моделей искусственного интеллекта. К примеру, оказывается, что зрители, которым нравится american wrestling, любят классическую музыку. Фильтры такие результаты не дадут.

### Персонализация — как догадаться, кто перед экраном?

Телевизором обычно пользуется несколько человек, и рекомендовать одно и то же бабушке и внуку будет неразумно. Есть несколько подходов, позволяющих более точно определить, кто в данный момент находится перед экраном. В случае использования описанного выше collaborative filtering учитывается, какая передача просматривалась только что.

Кроме этого, практически все компании используют разные системы выдачи, зависящие от времени дня и дня недели. Некое представление о некоей обычной среднестатистической семье дает возможность оператору считать, что днем ТВ смотрят дети и пенсионеры, вечером и в выходные — вся семья, ночью — молодежь.

Практически все системы предлагают возможность создания профилей — как при совместном пользовании ПК. Но, как показывает практика, абоненты этим почти не пользуются. Заинтересовать абонентов разные производители пытаются по-разному. Aprico, например, предлагает создание персональных каналов. Каждый член семьи отбирает то, что ему интересно в собственный телеканал — то есть занимается программированием и сам заинтересован создавать свой отдельный профиль. По данным компании, 70 процентов зрителей этой возможностью пользуются.

Индивидуальные профили отдельных пользователей помогают создать и мобильные телефоны или компьютеры, на которых можно отмечать телепрограммы, требующие записи. Например, оператор Freeview в Англии предлагает зрителям единую программу передач, которая доступна не только с ТВ, но и с компьютера и смартфона. Компьютер и телефон используются для того, чтобы пометить какую-то передачу — чтобы потом получить напоминание о ней с возможностью отдать команду на запись. Поскольку компьютер и, тем более, мобильный телефон — персональные устройства, информация об абоненте получается более точной.



Так выглядит схема системы рекомендаций по мнению Aprico

Абонента можно и вовлечь в сбор информации. Например, можно предлагать отмечать понравившиеся ему передачи — ставить звездочки-рейтинги, «лайки» в фейсбуке и т.п. В решении TiVo, которое только что внедрил английский оператор Virgin Media, абонент может ставить передаче «плюсы» или «минусы» — от одного до трех, то есть получается оценка из 7 баллов. Недавно в ряде EPG в Западной Европе появилась возможность отмечать через фэйсбук просмотренные передачи. Большинство посетителей конференции на IBC2011 отмечали, что это лишнее действие, поэтому они его не делают, даже если смотрят телевизор с планшетником на коленях, но нашлись и те, кто считает эту функцию удобной — например, при просмотре длинного сериала с повторами или таймшифтом удобно отметить, какие именно серии уже просмотрены.

**Полная персонализация — интеграция с социальными сетями**

Пользователи социальных сетей обычно добровольно предоставляют о себе самую полную информацию — не только пол и возраст, но еще и предпочтения и круг друзей, к мнению которых они готовы прислушаться. Поэтому производители систем поиска и рекомендаций стремятся интегрировать свои решения с социальными сетями. На

самом первом этапе телеканалы пользуются социальными сетями для привлечения посетителей на свои сайты, создания сообщества и повышения узнаваемости бренда. Российский телеканал Дождь, по словам Полины Козловской на конференции в сентябре в Санкт-Петербурге, имеет от 10 до 20 процентов заходов на сайт телеканала из социальных сетей. Следующий шаг в интеграции телевидения и соцсетей — интеграция контента и вывод на телеэкран записей твиттера или Facebook. В ряде телепередач это уже используется - например, так можно делать при просмотре телеканала Eurora Plus. Естественным продолжением будет введение эффективной системы рекомендаций оператора платного ТВ для своих абонентов-участников социальных сетей, которых становится все больше. Тенденции в этом направлении заметны каждому — телевизор все больше смотрится не совместно всей семьей, а по отдельности — на разных экранах, а чувство общности создают соцсети. Операторы платного ТВ должны заранее подготовиться и продумать соответствующие системы рекомендаций для этой аудитории.

potube.tv ориентируется именно на пользователей компьютеров и планшетников и занимается разработкой систем рекомендаций на основе социальных сетей. Остальные разработчики решений менее

радикальны, но тоже начинают с ними работу. На IBC 2011 TV Genius показала новую разработку — рекомендации на базе Facebook. Исследования показали, что в США у каждого пользователя в среднем около 150 друзей и каждый отметил около трех телепередач, поэтому пользователь Facebook, подключающийся к приложению TV Genius, сразу получает рекомендации от своего круга. Другая новинка компании — оценка количества твитов с упоминанием передачи. Интересно, что большое количество твитов, как показала практика, характерно как для очень интересных, так и для самых «отстойных» передач. «Серединка на половинку» не вызывает у телезрителей желания высказаться.

**Абсолютный контроль — знать о пользователе все**

Существуют еще более радикальные решения, нежели интеграция с социальными сетями, где пользователь все-таки сам рассказывает о себе то, что считает нужным. Например, японская вещательная корпорация NHK рассматривает возможности распознавания голоса и выражения лица пользователя перед экраном, а к тому же изучает окружающие его предметы — «контекст» пользователя в данный конкретный момент. В зависимости от этого контекста абоненту можно предлагать нужную инфор-

**Все для профессионального приема телевидения и радио со спутников**

**Профессиональные спутниковые конвертеры**

- SMW PLL конвертеры со стабильностью от ±5 до ±250 KHz
- SMW PLL Ku-band конвертеры Twin и Quattro
- Norsat DRO C-band конвертеры от ±100 до ±500 KHz
- Norsat PLL C-band конвертеры от ±2 до ±25 KHz

**Облучатели и волноводные узлы**

- C-band, Ku-band и совмещенные C/Ku-band
- С диэлектрическим (Ku) и штыревым (C) деполаризатором
- С разделителем поляризации и высоким уровнем развязки

**Усилители спутниковой ПЧ**

- С различным усилением, стандартным и высоким IP3

**Спутниковые конвертеры:** SMW PLL, Quattro PLL Twin PLL, QPLL, N8515, N3220, SMW ILA, OMT-30, XM-611, C-band Of, CR/CL/Ku Of

ooo САТПРО, [www.satpro.ru](http://www.satpro.ru)  
**SATPRO**  
 г. Екатеринбург, ул. Красноармейская, 76  
 +7 (343) 379-0807, 379-0857, 221-0012  
 e-mail: mail@satpro.ru, ICQ: 220225630

Logos: SMW SWEDISH MICROWAVE AB, Satpro, Norsat International Inc.

мацию. К счастью, это пока только планы. В системе рекомендаций HybridCast, которую разработала NHK для программ VOD в рамках гибридного решения для вещательного телевидения и интернета, используются обычные социальные сети и рейтинги, которые составляют абоненты.

### Паранойя

Пользователи интернета давно привыкли, что многие сайты анализируют их поведение на странице, поисковые запросы, тексты и видео, которые они смотрят. Аналогично производители предполагают, что и сбор информации о пользователе для системы поиска и рекомендаций вызовет негативную реакцию только у незначительного количества телезрителей. Ясно, что ее нужно предусмотреть и давать пользователю возможность отключать систему рекомендаций, или хранить персональную информацию о пользователе на устройстве пользователя, или использовать только общую статистическую информацию без ее привязки к конкретным абонентам.

### Как это совместить — как люди выбирают, что им смотреть?

И вот тут начинают работу социологи и психологи. Насколько человек склонен следовать рекомендациям? Каким он доверяет больше?

Результаты исследований поведения пользователей, проведенные в Западной Европе компанией TV Genius, показывают, что при поиске видео в интернете люди действительно следуют советам друзей или случайным рекомендациям. А вот при поиске и просмотре телепередач на телевизоре все происходит не так. Абоненты внимательно слушают советы родных и друзей о том, что им стоит посмотреть, но в 95 процентах случаев этим рекомендациям не следуют! Так что рекомендации друзей компания советует ставить на последнее место.

В остальном производители решений делятся на две группы. Одни считают, что основой для систем рекомендаций нужно брать описания передач, составленные специалистами — редакторами и критиками из специализированных журналов и сайтов. Другие — что в первую очередь нужно использовать collaborative filtering. Если есть обратная связь, то по мере накопления информации об абоненте система collaborative filtering начинает работать все лучше — например, компания Gravity сначала предлагает абоненту передачи, которые выбирает большинство, с учетом его местонахождения и времени суток, а затем «группа по интересам» определяется все более точно.

В общем случае компании-разработчики стараются внедрить максимальное количество разнообразных фильтров. В решении Compass компании ORCA Interactive, на-

пример, их семь. Самое популярное видео, похожее видео, продвигаемое оператором, выбор критиков, рекомендации из соцсетей, и т.д. Фильтров так много потому, что Orca предлагает операторам самим настроить систему под себя и выбрать то, что кажется правильным.

В Вее подход обратный. Операторам не предлагается проводить исследовательскую работу, количество фильтров сокращено до четырех (контекст, информация из соцсетей, collaborative filtering и семантический фильтр), причем зритель видит результаты сразу трех. В результате настройкой занимается не оператор, а зритель. Вее также отмечает важность способа представления результатов, и нужно выдавать не текстовую информацию, а трейлеры.

Все опрошенные компании-разработчики в своих системах поиска и рекомендаций используют все доступные пользователю источники видео — линейное вещание, VOD, отложенный просмотр (catch UP TV), видео, записанное на цифровой видеомagneфон. При этом результаты выдачи зависят от устройства, с которого приходит запрос. Например, в решении Вее при поиске с мобильного в выдаче будут более короткие фильмы.

Специалисты компании Gravity отмечают, что вкусы и предпочтения отдельного телезрителя и всего общества со временем меняются, поэтому система должна не просто накапливать данные, но и отслеживать эти изменения.

Параметров поиска и выдачи настолько много, что системы становятся действительно сложными. Чаще всего операторам предлагаются системы с максимально возможным числом параметров, на базе которых они смогут понять, как эффективно работать с системой непосредственно на их абонентской базе.

Но на то, чтобы адаптировать решение под своих абонентов и добиться его эффективной работы, по мнению TV Genius, у оператора уходит около двух лет.

### Кому это выгодно, и кто станет это внедрять?

Как уже говорилось выше, оператор платного ТВ может быть заинтересован в введении системы рекомендаций, чтобы просто повысить лояльность абонента, потому что абонент, который считает, что смотреть нечего, склонен отказаться от услуг. Аналогичные причины — желание максимально соответствовать персональным потребностям абонента — подвигают оператора предлагать абоненту возможность составлять фаворитные списки и вводить поканальную подписку.

Таким образом, внедрение системы рекомендаций может теоретически снизить отток абонентов.

Есть еще и другой момент. Обычно производители решений дают возможность оператору влиять на выдачу результатов поиска. Абоненту в первых рядах будут предложены релевантные, но платные материалы — видео по запросу, передачи, идущие на каналах премиального пакета. Как поясняет английский оператор Fetch TV, сам абонент заходит в раздел VOD, только если смотреть ему совсем уж нечего. Поэтому оператору выгодно предлагать фильмы из библиотеки VOD среди прочих программ.

Библиотека видео по запросу, в которой работает система рекомендаций, начинает более эффективно предлагать зрителям редко просматриваемые фильмы — так называемые long tail — и поэтому становится менее затратной для оператора, чем когда всем зрителям предлагаются популярные блокбастеры. Оператор также сможет размещать в результатах поиска рекламу в соответствии с интересами абонента.

Однако нужно помнить, что система рекомендаций не бесплатна, оператор платит и за алгоритмы, и за дополнительную информацию о контенте. Пока системы рекомендаций только развиваются, и ни один оператор не дает точных данных об экономическом эффекте от их внедрения.

Пионерами в использовании систем поиска и рекомендаций стали операторы IPTV самого разного масштаба. Например, решение Orca Compass начинают внедрять грузинский оператор Silknet и французский France Telecom. Спутниковое ТВ тоже использует рекомендации — вспомним российский Актив ТВ, где не требуется даже обратного канала. Если же он есть, возможностей еще больше. Решение Think Analytics, например, использует английский Sky с его 11 миллионами пользователей (и обратным каналом через наземную сеть).

Введение системы рекомендаций оказывается экономически выгодным для производителей телевизионных приставок, работающих с видео в интернете. Хороший пример эффективности — цифровой видеомagneфон TiVo, записывающий как операторский контент, так и видео из интернета. TiVo пользуется успехом именно потому, что позволяет абонентам искать и записывать телепрограммы, и занимается этими разработками с самого начала своего существования — около десяти лет. В 2011 году TiVo запустила решение для Virgin, английского оператора кабельного ТВ, которое, по словам оператора, оказалось очень успешным и позволило заметно увеличить продажи новых дорогих приставок. У Orca Interactive тоже есть решение для приставок, которое реализовано совместно с ее родительской компанией Viaccess и внедрено у шведского оператора Boxer.

Системы рекомендаций и поиска часто предлагаются и в виде онлайн-версии — как приложение к программе передач операторо-

ра, которую тот выкладывает на сайте, или даже как отдельный продукт для поиска интернет-видео — см, например, watchmi.tv от Aprico, bee.tv, jinni.com и tivo.com — но для конечного пользователя в России это достаточно бесполезно, так как сервисы ориентированы на жителей определенных стран, и программы передач туда заложены соответствующие.

### Проблемы

Итак, системы поиска и рекомендаций начали свое развитие. Как же участники рынка видят стоящие перед ними проблемы? Как мне пояснили в компании ORCA, это:

1. Необходимость создания убедительной бизнес-модели. Важно доказать операторам, что с помощью системы поиска и рекомендаций можно заработать денег, а не просто в какой-то мере улучшить предоставляемые услуги. В компании Gravity считают, что проблема решается легко, если стоимость решения зависит от финансовых результатов его внедрения. Доход Gravity зависит от увеличения продаж/монетизации на стороне клиента относительно показателей, зафиксированных до начала эксплуатации рекомендательного решения. Компания специализируется на рекомендациях, связанных с рекламой продуктов и услуг, и поэтому преимущества от внедрения ее решения поддаются подсчету. В общем случае это, конечно, не так.

2. Разработка единых стандартов для метаданных. Решением этой задачи занимаются исследовательские организации типа BBC, NBC и NoTube, и не только с точки зрения улучшения систем рекомендаций, но и в рамках проектов гибридного ТВ, использующего интернет-каналы в дополнение к вещанию.

3. Абоненту должно быть просто и удобно пользоваться системой. (Нужно помнить, что телевизор — это не компьютер с его мышкой и клавиатурой!)

4. ThinkAnalytics добавляет, что пользователь не очень склонен вводить информацию о себе и пока не убедится в удобности системы рекомендаций, дает минимальную. Соответственно, система рекомендаций должна работать и в этом случае — то есть необходимо поддерживать несколько стратегий.

5. Проблемы, которые добавляют специалисты компании TV Genius, — это недостаточность ресурсов устройств типа телевизоров и приставок. Производители стремятся экономить и поэтому не используют самые современные процессоры. В результате памяти и мощности процессора недостаточно для быстрого и качественного поиска на самом устройстве. Именно поэтому тема поиска и рекомендаций появляется только теперь, когда эти устройства подключаются к интернету и есть возможность

организовать процесс на стороне сервера оператора или даже на серверах производителя решения — предоставить услугу в формате SaaS (как готовую услугу, полностью реализуемую на стороне поставщика решения).

6. Еще одна проблема — трудности разработки поисковых алгоритмов. Видео ищется по-другому — нужны, например, дополнительные алгоритмы для отбора самых свежих версий телесериала.

7. Прозрачность поиска. По мнению TV Genius, для пользователей должно быть понятно, почему им рекомендуется та или другая передача, а если результат не соответствует ожиданиям, должна быть возможность скорректировать результаты — убрать передачу, которая не нравится так, чтобы больше ее не было в результатах выдачи.

8. Со стороны операторов при выборе системы нужно, конечно, учитывать и то, как она будет взаимодействовать с имеющейся инфраструктурой, сколько места и ресурсов потребует работа приложения в STB, а также сколько она будет стоить.

Не смотря на изобилие проблем, количество операторов, использующих системы рекомендаций, растет во всем мире. Несомненно, и в России ими будут пользоваться не только интернет-сервисы. ■

**Лардотелеком**

**Оборудование для DVB-C вещания**

**TELEVIEW & PBI**

**DVB-C ресиверы**  
Conax, DVCrypt, Cerber Crypt

**LATEL**

**LC-808HD**  
VoD  
Гибридный (DVB-C+IP) ресивер стандарта HD  
встроенный медиаплеер

**LC-908**  
от 1200 РУБ  
DVB-C ресивер стандарта MPEG-2 SD

**Головная станция от 180 тыс. руб.**

Екатеринбург      Петропавловск

тел. (343) 3840-440, 261-35-89      тел. +7 (7152) 39-00-68  
www.latel.ru      www.latel.kz  
e-mail: lt@latel.ru      e-mail: lt@latel.kz

реклама