
МАТЕМАТИЧЕСКИЙ АНАЛИЗ ЭКОНОМИЧЕСКИХ МОДЕЛЕЙ

ПРОГНОЗИРОВАНИЕ ЧИСЛЕННОСТИ КЛИЕНТСКОЙ БАЗЫ КОМПАНИИ НА ОСНОВЕ МАРКОВСКИХ ЦЕПЕЙ

© 2016 г. А.В. Андреева, Т.К. Богданова

(Москва)

Построена информационно-логическая комплексная модель управления клиентской базой компании для расчета показателя долгосрочной стоимости клиента. В отличие от предыдущих исследований в модели учтены особенности покупательского поведения и социально-демографические характеристики групп клиентов, а перемещение клиентов внутри клиентской базы представлено в виде марковской цепи. Разработанная комплексная динамическая модель управления клиентской базой компании позволяет прогнозировать численность кластеров и клиентской базы в целом на любом заданном интервале времени, проводить анализ динамики изменения численности клиентских кластеров и выделить наиболее важные этапы формирования и развития клиентской базы компании. На каждом этапе выделенные группы клиентов оцениваются с точки зрения будущей прибыли для компании и потенциала перекрестных продаж продуктов/услуг. Полученные оценки позволяют рассчитать допустимый бюджет на маркетинговые мероприятия при сохранении заданного уровня прибыли и составить наиболее эффективный план проведения адресно-направленных маркетинговых мероприятий.

Ключевые слова: управление клиентской базой компании; розничная торговля; группы клиентов; лояльность клиента; покупательское поведение; марковские цепи; динамическая модель; долгосрочная стоимость клиентской базы.

Классификация JEL: C15, C32, C53.

1. ВВЕДЕНИЕ

Проблема управления клиентской базой компании, повышение лояльности клиентов на протяжении последних двух десятилетий остается актуальной. Последний мировой кризис показал, что компании с высоким уровнем лояльности клиентов в условиях нестабильного финансового положения смогли удержать лидирующие позиции на рынке. Большинство исследователей отмечают, что качественное управление клиентской базой является одной из важнейших составляющих эффективного управления компанией (Мейер, 2004; Пепперс, Роджерс, 2006, Сьюэл, Браун, 2007).

Переход компаний к клиентоориентированному ведению бизнеса (Черкашин, 2004) позволяет компании увеличить прибыль и повысить эффективность, а также доходность за счет роста выручки от существующей клиентской базы и оптимизации операционных затрат. В работе Ф. Райчхелд (Reichheld, 1996) отмечает два ключевых фактора повышения прибыли компании от повышения лояльности клиентов: влияние уровня лояльности на динамику численности потребителей и рост прибыли на одного клиента. Между уровнем лояльности и коэффициентом выбытия из клиентской базы существует статистически показанная обратная зависимость. Рост прибыли от клиента с течением времени объясняется следующими причинами: ростом с течением времени числа перекрестных продаж и потребностей клиентов. Коэффициент лояльности влияет и на другие параметры, а это прежде всего снижение издержек на привлечение клиента, получение базовой прибыли, снижение операционных издержек, увеличение числа рекомендаций, возможность повышения ценовой премии.

В рамках клиентоориентированной стратегии управления компанией можно выделить *два ключевых подхода* к оценке эффективности управления клиентской базой компании – на основе

показателя лояльности клиентов и оценки долгосрочной стоимости клиента (customer lifetime value, CLV).

В настоящее время первое направление наиболее широко разработано как в западной, (Ламбен, 1996; Reichheld, 1996; Aaker, 1991; Hofmeyr, Rice, 2000), так и в российской литературе (Цысарь, 2002; Добровидова, 2003; Карасев, 2008). В качестве основных показателей рассматриваются, в частности, коэффициент повторных покупок, доля кошелька клиента и уровень удовлетворенности.

Следует отметить, что модели управления клиентской базой на основе показателя лояльности достаточно просты и интуитивно понятны, но имеют скорее экспертный, чем формальный характер. Исключение составляет модель повторных покупок, позволяющая проводить анализ текущего состояния клиентской базы на основе данных прошедших периодов. Эти модели не позволяют прогнозировать состояния клиентской базы на будущих интервалах времени и не могут быть использованы для разработки долгосрочной стратегии предприятия. Кроме того, в них не учитываются затраты компании на привлечение и обслуживание клиентов, что ограничивает возможности компании в выборе оптимального распределения финансовых ресурсов на управление клиентами.

Поэтому компании, которые при разработке стратегии управления клиентами основываются только на повышении уровня лояльности и удовлетворенности потребителей, могут оказаться заложниками своей лояльности: с одной стороны, максимальный уровень удовлетворенности клиентов и высокие оценки качества обслуживания, с другой – низкие показатели прибыльности и рентабельности вложенных средств.

Преодолеть ограничения моделей управления клиентской базой компании, основанных на показателе лояльности клиентов, позволяют подходы на основе показателя долгосрочной стоимости клиента CLV. Долгосрочная стоимость клиента представляет собой совокупность чистого дохода, ожидаемого от клиента в будущем. Подход на основе показателя CLV позволяет соотнести доходы, полученные от клиентов в течение всего периода их взаимодействия с компанией, с расходами, которые компания несет на привлечение и обслуживание клиентов с учетом временной стоимости денег. Показатель CLV позволяет предоставить объективную оценку эффективности управления клиентской базой компании, численно измеримую и сравнимую между собой во времени и в совокупности с другими компаниями, как того же сегмента экономики, так и различных отраслей.

Этот подход рассматривается в зарубежной литературе (Berger, Nasr, 1998; Reichheld, 1996; Fader, Hardie, 2009; Malthous, Blattberg, 2005; Dipak, Siddhartha, 2002) и в ряде работ российских авторов (Полежаев, 2006б; Третьяк, Слов, 2012; Крюкова, Кузьмин, 2009).

Показатель CLV в качестве критерия управления клиентской базой фактически сводится к максимизации прибыли на ограниченном интервале времени, следовательно, критерий управления клиентской базой в этом случае – максимизация прибыли, которую приносят клиенты за заданный период, или максимизация показателя CLV. К ограничениям, которые затрудняют использование данного подхода в организациях, относится сложный и многошаговый алгоритм расчета, требующий в том числе консолидации всех расходов компании в расчете на одного клиента, а также отсутствие однозначной интерпретации полученных значений.

2. ПОДХОДЫ К УПРАВЛЕНИЮ КЛИЕНТСКОЙ БАЗОЙ КОМПАНИИ НА ОСНОВЕ ПОКАЗАТЕЛЯ CLV

Можно выделить три основных направления решения задачи управления клиентской базой компании на основании показателя CLV.

1. Построение регрессионных моделей (Berry, Linoff, 2004; Malthous, Blattberg, 2005). К сожалению, несмотря на гибкость и универсальность регрессионных моделей, они имеют ряд недостатков, что и отмечается в обзорах (Dipak, Siddhartha, 2002; Fader, Hardie, 2009): необходимость адаптации и пересчета основных коэффициентов для каждой конкретной организации; высокая чувствительность к исходным данным, что повышает вероятность получения ошибоч-

ных результатов при незначительных отклонениях в данных; резкое снижение точности прогнозирования с увеличением числа периодов прогнозирования.

2. Построение Pareto/NBD моделей (Schmittlein, Morrison, Colombo, 1987). Эти модели имеют ряд ограничений: получение ошибочных результатов, если интервал между транзакциями клиента превышает два года; трудности использования моделей в компаниях с большим числом клиентов, поскольку расчет CLV ведется для каждого клиента; более сложный расчет параметров в сравнении с регрессионными моделями, что препятствует их широкому использованию для решения задач оперативного управления. Модификация Pareto/NBD модели, предложенная в работе (Fader, Hardie, Lee, 2005), позволила частично решить проблему сложности оценки параметров модели. Однако чувствительность к исходным данным и недостаточная гибкость модели существенно ограничивают область применения и частоту использования данной модели в реальном секторе экономики.

3. Построение моделей марковских цепей (Markov chain models, MCM) (Pfeifer, Carraway, 2000; Полежаев, 2006б; Третьяк, Слое, 2012; Крюкова, Кузьмин, 2009).

В отличие от вероятностных моделей модели на основе марковских цепей (MC-модели) отличаются относительной простотой и требуют меньших затрат при расчете своих показателей. При этом уровень точности прогнозирования MC-моделей сопоставим с Pareto/NBD моделями. Кроме того, MC-модели отличаются высокой гибкостью и легко адаптируются к специфике работы предприятия.

Анализ существующих моделей управления клиентской базой компании на основе показателя долгосрочной стоимости клиентской базы показывает, что, несмотря на разнообразие представленных моделей, всем им присущи следующие ограничения:

1) концентрация на управлении поведением отдельного клиента, что требует значительных затрат времени в случае использования моделей в компаниях с многомиллионной клиентской базой;

2) предположение, что покупательское поведение зависит только от размера маркетинговых расходов; таким образом, профиль клиента и влияние способа коммуникаций с клиентом не учитываются;

3) для сегментации клиентской базы, как правило, учитываются только число и давность покупок клиента и не учитываются его время взаимодействия с компанией, категории покупаемых товаров и сумма покупки;

4) общая сумма затрат учитывается только на маркетинговые мероприятия. Такие параметры, как тип проводимых маркетинговых мероприятий, способы коммуникации с клиентом и характер рекламного предложения, не рассматриваются, не анализируется характер их влияния на каждую группу клиентов.

Схема существующих направлений исследования задачи управления клиентской базой компании и подходов к ее решению приведены на рис. 1.

Подходы, рассматривающие клиентскую базу как одну большую группу, не позволяют учесть того факта, что в большинстве случаев у клиентов – различные модели покупательского поведения. Вместе с тем, как показано в работе (Андреева, 2011), клиентская база компании не является однородной и представляет собой непересекающиеся между собой кластеры. Каждый кластер характеризуется рядом показателей: доходами, получаемыми от клиентов; расходами на привлечение и удержание клиентов; временем взаимодействия клиента с компанией; давностью покупок; категорией покупаемых товаров.

Поведение отдельного клиента будет не постоянным во времени, т.е. существует некий жизненный цикл в отношении с клиентами, с помощью которого можно описать изменение характера покупательского поведения во времени. На поведение клиента влияют внешние факторы: маркетинговые инициативы, рост популярности определенной категории товаров, изменение образа жизни и т.д., которые приводят к изменению интенсивности покупок, т.е. переходу клиента из одной группы в другую.

В качестве методов совершенствования моделей управления клиентской базой компании в работах (Андреева, 2010, 2011) предлагается перейти от моделей управления поведением



Рис. 1. Схема существующих направлений исследований задачи управления клиентской базой компании

отдельного клиента к модели управления группами клиентов, обладающих сходными характеристиками покупательского поведения. В работе (Андреева, 2012) предложен подход к построению комплексной модели управления клиентской базой компании, позволяющий учитывать особенности покупательского поведения и социально-демографические отличия групп клиентов.

3. ИНФОРМАЦИОННО-ЛОГИЧЕСКАЯ МОДЕЛЬ УПРАВЛЕНИЯ КЛИЕНТСКОЙ БАЗОЙ КОМПАНИИ

Основные этапы разработки комплексной модели управления клиентской базой компании с учетом подхода, предложенного в (Андреева, 2012), приведены на рис. 2.

Ключевым моментом успешного решения задачи управления клиентской базой является ее сегментация. Критерии разделения клиентов на группы зависят от специфики деятельности компании и доступной для анализа информации о клиентах, но в качестве базовых характеристик необходимо учитывать покупательское поведение: частоту покупок, размер среднего чека, период неактивности клиента (так называемое “время сна”) и общее время взаимодействия клиента с компанией. Для расчета показателя долгосрочной стоимости клиента необходимо использовать три ключевых параметра: размер дохода от группы клиентов; расходы по группе клиентов; численность групп клиентов в каждый момент времени. Третьим важным этапом разработки комплексной модели управления клиентской базой является анализ проводимых в компании маркетинговых мероприятий и оценка характера влияния этих мероприятий на группы клиентов. Также при анализе динамики покупательского поведения необходимо учитывать возможное влияние факторов внешнего воздействия – таких как рыночные тенденции или появление новых технологий, а также естественные изменения покупательского поведения клиентов, связанные с изменением уровня жизни или социального статуса групп клиентов. На основе проведенного анализа строятся регрессионные модели, отражающие характер покупательского поведения

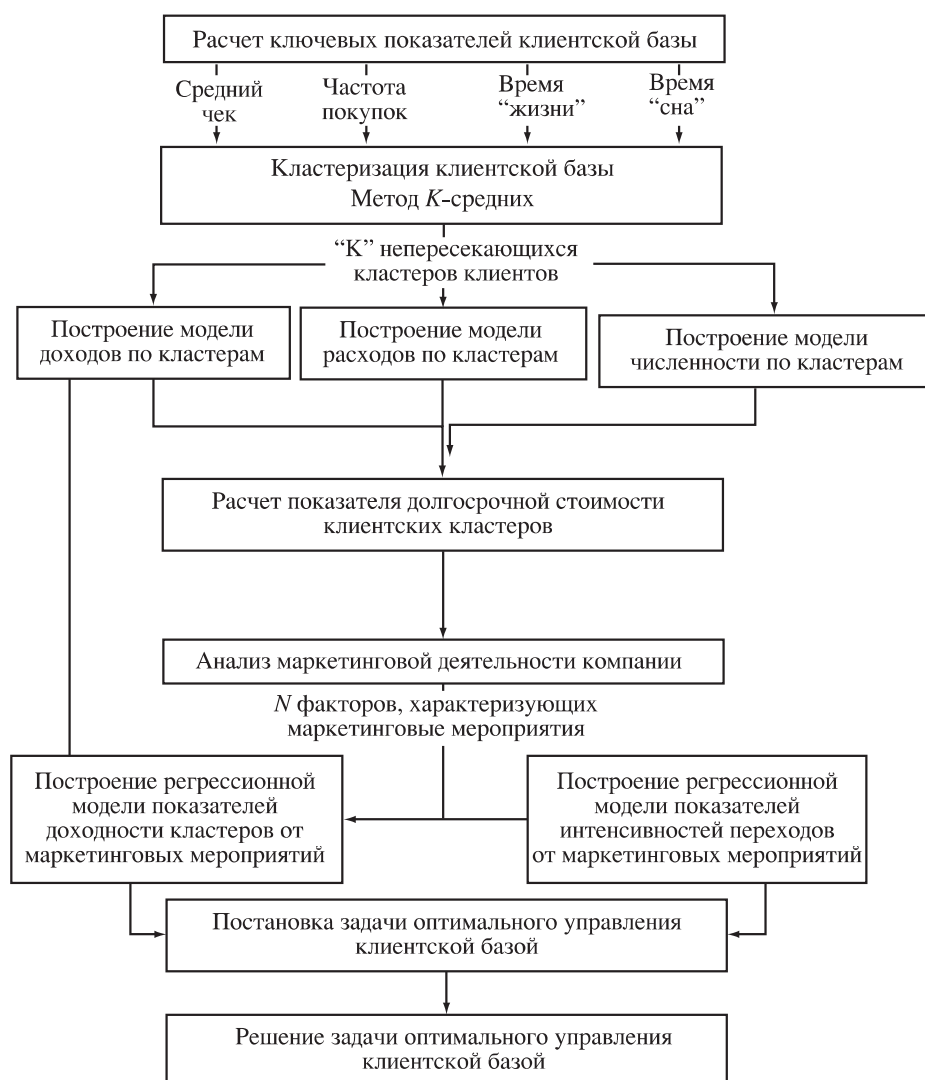


Рис. 2. Информационно-логическая модель управления клиентской базой компании

групп клиентов и уровень влияния маркетинговых кампаний и внешних факторов на динамику его изменения.

По итогам разработки указанных выше моделей осуществляется постановка задачи оптимального управления клиентской базой компании, т.е. формулируется критерий оптимального управления, существующие ограничения и требования к виду получаемого оптимального решения.

В рамках задачи управления клиентской базой компании критерием оптимального управления является достижение клиентской базой максимального значения долгосрочной стоимости на горизонте планирования, а результатом решения – набор параметров и требований к проводимым маркетинговым мероприятиям, в результате которых долгосрочная стоимость клиента достигает оптимального значения. В качестве ограничений в компании выступает конечный бюджет на проведение маркетинговых мероприятий и технические возможности сотрудников их организовать.

4. МОДЕЛИРОВАНИЕ ЧИСЛЕННОСТИ КЛИЕНТСКОЙ БАЗЫ КОМПАНИИ

Кластерный подход к прогнозированию численности клиентской базы, предложенный в (Андреева, 2011), базируется на следующих предпосылках:

- клиентская база компании может быть разбита на n непересекающихся кластеров;
- каждый кластер имеет свои характеристики по доходам клиентов, времени взаимодействия клиента с компанией, социально-демографическим характеристикам клиентов, категориям покупаемых товаров;
- клиентская база компании динамически обновляется за счет постоянного притока новых клиентов и выбытия клиентов, зарегистрированных в базе ранее;
- каждый новый клиент в зависимости от его характеристик автоматически попадает в один из выделенных на этапе кластеризации кластеров;
- в течение жизненного цикла происходит изменение покупательского поведения клиента, что обуславливает его перемещение из одного кластера в другой или прекращение взаимодействия клиента с компанией.

Тогда динамика формирования кластеров может быть описана стохастическим процессом, где перемещение клиентов внутри клиентской базы компании представляется в виде марковской цепи, состоящей из $K + 1$ состояний. Состояния с 1 по K характеризуют один из выделенных клиентских кластеров со своими покупательскими характеристиками: частота покупок, средний чек, число месяцев от первой покупки и последней покупки, а состояние $K + 1$ характеризует внешний мир. Переходы от состояния $K + 1$ в другие состояния марковской цепи возможны в любых направлениях.

Перемещение клиентов внутри выделенных клиентских кластеров рассматривается для случая, когда влияние прошлой истории взаимодействия клиента и компании на будущее поведение клиента невыявлено. Иными словами, процесс описывается как цепь Маркова первого порядка, предлагается подход к прогнозированию численности групп клиентов, в основе которого лежит адаптация модели движения кадров О.В. Староверова (Староверов, 1997). Этот подход позволяет учесть различия в покупательском поведении клиентов и отследить изменение качественных характеристик клиентской базы, что повышает точность расчета показателя CLV.

В этом случае модель прогнозирования численности клиентской базы по кластерам, как показано в (Андреева, 2012, с. 63), имеет вид:

$$N^t = N^{t-1} - RN^{t-1} + P^T RN^{t-1} + (c_2 - c_1), \quad (1)$$

где N^t – численность кластеров в момент времени t (вектор); N^{t-1} – численность кластеров в момент времени $t-1$ (вектор); $R = \{\delta_{ij} r_{ij}\}$ – матрица интенсивностей выходов клиентов из кластера i , где δ_{ij} – символ Кронекера; $P = \{p_{ij}\}$ – матрица вероятности нахождения индивида из кластера i в кластере j ; t – интервал времени, за который производится расчет численности кластеров; c_1 и c_2 – вектора выбытия и прибытия клиентов соответственно.

Показатели R и P можно оценить с помощью матрицы привлечения и потери клиентов. Впервые матрицу привлечения и потери потребителей использовал Ф. Райчхелд для анализа покупательского поведения клиентов (Reichheld, 1996). Этот подход был использован для оценки параметров модели прогнозирования численности клиентской базы компании. В матрице привлечения и потери клиентов (табл. 1) на пересечении строк и столбцов указано число клиентов x_{ij} , перешедших из кластера i в кластер j . Интенсивность перехода a_{ij} клиентов из кластера i в кластер j определяется как отношение числа клиентов, перешедших из кластера i в кластер j к численности кластера на момент выхода клиентов, и рассчитывается по формуле

$$a_{ij} = x_{ij} / \left(\sum_{l=1}^K x_{il} + c_{ij} \right). \quad (2)$$

Тогда интенсивность выхода клиентов из кластера i (r_i) представляет собой сумму всех переходов из кластера i в другие кластеры, т.е.

$$r_i = \sum_{j=1}^K a_{ij} - a_{ii}. \quad (3)$$

Таблица 1. Матрица привлечения и потери клиентов

	Кластер 1	...	Кластер K	Кластер $K + 1$ (внешний мир)
Кластер 1	x_{11}	...	x_{1k}	c_{11}
...	...	...	...	...
Кластер K	x_{k1}	...	x_{kk}	c_{1k}
Кластер $K + 1$ (внешний мир)	c_{21}	...	c_{2k}	

Вероятность нахождения клиента из кластера i в кластере j рассчитывается как

$$p_{ij} = a_{ij}/r_i \quad (4)$$

Таким образом, используя матрицу привлечения и потери клиентов, можно оценить значения показателей R и P в модели прогнозирования численности клиентской базы.

Перемещение клиентов из группы в группу происходит при изменении ключевых показателей их покупательского поведения: частоты покупок, размера среднего чека, категории покупаемых товаров (Андреева, 2012). Динамическая модель прогнозирования численности клиентской базы с учетом интенсивности перехода a_{ij} клиентов из группы i в группу j будет иметь вид:

$$N_i^t = (1 + a_{ii}^t)N_i^{t-1} + \sum_{j=1}^K a_{ji}^t N_j^{t-1} - \sum_{j=1}^K a_{ji}^t N_i^{t-1} + (c_{2i}^t - c_{1i}^t), \quad (5)$$

где N_i^t – численность кластера i в момент времени t (вектор); N_i^{t-1} – численность кластера i в момент времени $t-1$ (вектор); a_{ij}^t – интенсивность перехода из кластера i в кластер j в момент времени t ; t – интервал времени, за который производится расчет численности кластеров; c_1^t и c_2^t – вектора выбытия и прибытия клиентов в момент времени t соответственно.

Таким образом, на основе статистических данных о перемещениях клиентов внутри клиентской базы, а также на основе информации о числе новых и ушедших клиентов, предлагаемая модель позволяет определить интенсивности переходов клиентов между группами и оценить вероятность нахождения клиента после перехода в одной из групп. Это, в свою очередь, позволяет спрогнозировать численность групп клиентов на долгосрочном интервале времени.

5. ИНФОРМАЦИОННАЯ БАЗА ИССЛЕДОВАНИЯ

Для апробации предложенной динамической модели прогнозирования численности клиентской базы были использованы данные сегмента салонов сотовой связи одной из крупнейших торгово-розничных компаний России. Сеть магазинов, расположенных на всей территории Российской Федерации, в общей сложности имеет более 2000 торговых точек. Основные категории продаваемых товаров: аппараты сотовой связи, смартфоны, фото- и видеотехника, ноутбуки и планшеты, прочие электронные товары (электронные книги, плееры, GPS-навигаторы и т.д.). Также компания предоставляет ряд финансовых услуг: прием платежей от клиентов (сотовая связь, Интернет, ЖКХ), денежные переводы, погашение кредитов, страхование. Для анализа использовались данные о покупках клиентов компании в Казани в период с ноября 2007 г. по март 2012 г. Общее число уникальных клиентов – 220 292 записи.

В рамках данного исследования была использована информация о покупках только тех клиентов, которые оформили карту лояльности компании и в момент оплаты предъявили ее на кассе. Оформление карты лояльности происходит при совершении первой покупки клиентом и заполнении специальной анкеты (фамилия и имя клиента, дата рождения, пол, возраст, контактная информация). При предъявлении карты лояльности в момент совершения покупки клиенту начисляются бонусные баллы в размере 1% суммы покупки. Бонусные баллы в дальнейшем могут быть использованы в качестве скидки при последующих покупках.

Для анализа было использовано семь следующих переменных: 1) уникальный идентификатор клиента (*Client_ID*); 2) пол (*Sex*, принимает значения: 1 – мужчина, 0 – женщина); 3) возраст (*Age*); 4) семейное положение (*Family_Status*, принимает значения: 1 – женат/замужем, 2 – холост/не замужем, 0 – не указано); 5) дата и время покупки (*Purchase_Date*); 6) сумма покупки в рублях РФ (*Purchase_Sum*); 7) категория купленного товара (*Purchase_Category*, принимает значения: 1 – сотовые телефоны, смартфоны; 2 – фото- и видеотехника; 3 – ноутбуки, нетбуки, планшеты; 4 – аксессуары; 5 – платежи; 6 – финансовые продукты (денежные переводы, погашения кредитов, страховые продукты); 7 – непрофильные товары; 8 – товары, бывшие в употреблении; 9 – тарифные планы, услуги операторов сотовой связи).

Эти наборы параметров характеризуют каждого уникального клиента компании, т.е. база представляет собой панельные данные. Для проведения анализа покупательского поведения в динамике необходимо осуществить переход к временным рядам, т.е. к показателям, характеризующим покупательское поведение клиента в каждый из заданных интервалов планирования, – месяц.

Для решения данной задачи нами были рассчитаны следующие агрегирующие переменные: 1) дата первой покупки клиента – *First_Purchase_Date*; 2) дата последней покупки клиента – *Last_Purchase_Date*; 3) число покупок клиента: за весь период – *Purch_Num_All*, ежемесячно – *Purch_Num_i*, ежемесячно накопленным итогом – *Purch_Cumul_Num_i*; 4) сумма покупок клиента: за весь период – *Purch_Sum_All*, ежемесячно – *Purch_Sum_i*, ежемесячно накопленным итогом – *Purch_Cumul_Sum_i*; 5) частота покупок клиента: за весь период – *Freq_All*, ежемесячно накопленным итогом – *Freq_i*.

Частота покупок в момент времени t рассчитывается как отношение числа покупок клиентов, совершенных в момент времени t , к числу месяцев, прошедших от даты первой покупки до момента времени t :

$$Freq_t = \sum_{i=1}^t Purch_Num_i / t. \quad (6)$$

Замечание. Частота покупок в последнем месяце совпадает с частотой покупок при расчете за весь период времени.

Дополнительно были рассчитаны следующие четыре производные от основных переменных показатели покупательского поведения.

1. Средний чек клиента: за весь период *Avg_All*; ежемесячно накопленным итогом *Avg_i*.

Средний чек клиента в момент времени t в соответствии с (7) рассчитывается как отношение накопленной суммы покупок клиента на момент времени t к числу покупок клиента на момент времени t :

$$Avg_t = \sum_{i=1}^t Purch_Sum_i / \sum_{i=1}^t Purch_Num_i. \quad (7)$$

2. Число месяцев, прошедшее от первой до последней покупки.

3. Число месяцев, прошедшее от первой покупки до выбытия клиента или до марта 2012 г.¹, если клиент активен.

4. Число месяцев, прошедшее от последней покупки до выбытия клиента или до марта 2012 г., если клиент активен.

Моделирование проводилось с использованием следующих программных инструментов анализа данных: IBM SPSS Statistics, версия 19; Eviews 7 Standart Edition.

В процессе проведения валидации исходных данных был выявлен ряд нарушений в заполнении анкеты и некоторые аномальные значения и выбросы: ошибки в заполнении даты рождения; мошеннические действия: выявление аномально большого числа покупок; технические ошибки: выявление клиентов с аномально маленькой или большой суммой покупок и среднего чека;

¹ Последний период, вошедший в используемый для расчетов набор данных.

Таблица 2. Область допустимых значений переменных, руб.

Наименование переменной	Минимальное значение	Максимальное значение
Возраст клиента	18	90
Сумма покупки	100	2 000 000
Средний чек	10	2 000 000
Совокупное число покупок	1	800

отрицательные значения; незаполненные обязательные поля анкеты (возраст, пол), клиенты, не оформившие ни одной покупки. В результате анализа из первичной базы данных было удалено 38 365 записей. Для построения модели использовалась выборка из 181 927 записей.

В табл. 2 приведена область допустимых значений переменных.

Анализ демографических характеристик клиентской базы показал, что среди клиентов компании имеется небольшое преобладание мужчин; из числа указавших свое семейное положение – небольшое преобладание “холостых” над теми, кто обозначил свой статус, как “женат/замужем”; и значительная группа клиентов, которые не отметили своего семейного положения. Преобладающая возрастная категория как среди мужчин, так и среди женщин, – молодые люди в возрасте от 25 до 40 лет. Потребительское поведение клиента характеризуется его покупательской активностью – совокупным числом покупок; совокупной суммой покупок; размером среднего чека; частотой покупок; временем “активной жизни” и “неактивной жизни” (“сна”) клиента. Большая часть клиентской базы имеет довольно пассивное покупательское поведение: небольшое число покупок из бюджетной категории товаров (средний чек от 1–2 тыс. руб.), частота покупок низкая: в среднем – менее одной покупки в год.

Было выявлено, что число клиентов со сроком жизни менее 10 месяцев значительно ниже числа клиентов, привлеченных на этапе формирования клиентской базы (срок жизни от 50 до 53 месяцев). То есть можно сделать вывод, что происходит постепенное насыщение клиентской базы компании. Примерно 12 500 клиентов активны и совершили покупку в последнем месяце (около 7% рассматриваемой клиентской базы). Распределение клиентов по давности покупок довольно равномерное.

6. КЛАСТЕРИЗАЦИЯ КЛИЕНТСКОЙ БАЗЫ КОМПАНИИ В РАМКАХ ПРОВЕДЕННОГО ИССЛЕДОВАНИЯ

Кластеризация клиентской базы была проведена методом *k*-средних с использованием программного инструментария IBM SPSS Statistics. Для кластеризации клиентской базы использовались нормированные значения следующих наборов переменных: частота покупок клиента (за весь период); средний чек клиента (за весь период); число месяцев с первой покупки до выбытия клиента или до марта 2012 г., если клиент активен². В дальнейшем данный показатель был обозначен как время “жизни” клиента; количество месяцев с последней покупки до выбытия клиента или до марта 2012 г., если клиент активный, как время “сна” клиента.

В результате анализа результатов кластеризации для дальнейшего исследования было выбрано пять кластеров из шести кластерных решений (табл. 3).

На основе проведенного нами анализа представлено сводное описание получившихся клиентских кластеров.

Кластер 1. “Рядовые покупатели”: клиенты, которые приобретают в компании профильные товары и услуги – сотовые телефоны недорогих (бюджетных) моделей, аксессуаров к ним, а также услуги приема платежей в адрес различных поставщиков услуг. Около половины клиентов вместе с телефоном приобретают тарифный план. Средний чек составляет 1462,48 руб., невысокая частота покупок – 0,45 покупки в месяц (или, примерно, 1 раз в 2 месяца).

² Март 2012 г. – дата последнего периода в наборе для проведения анализа данных.

Таблица 3. Кластеризация клиентской базы (кластеры 1–5)

Наименование переменной	Кластер 1	Кластер 2	Кластер 3	Кластер 4	Кластер 5
Средний чек	1462,48	755,04	14747,44	921,27	1918,12
Частота покупок	0,45	2,50	0,48	7,48	0,15
Время “жизни”	30,4	27,7	23,2	24,9	41,7
Время “сна”	8,8	3,0	13,8	1,5	33,7
Численность кластера	59 518	11 198	2382	1384	36 686

Кластер 2. “Плательщики”: клиенты, имеющие высокую частоту пользования услугами по приему платежей в адрес различных поставщиков услуг. Эти клиенты также приобретают в компании телефоны и аксессуары к ним. Около 60% также дополнительно приобретают тарифные планы сотовых операторов. При этом средний чек ниже, чем у рядовых покупателей, и составляет 755,04 руб., но большая частота покупок – в среднем 2,5 покупки в месяц.

Кластер 3. “Средний класс”: клиенты, приобретающие дорогие модели сотовых телефонов, а также активно пользующиеся финансовыми услугами компании: денежные переводы, погашение кредитов, страхование. Средний чек у данной группы значительно выше, чем у других, и составляет 14 747,44 руб., частота покупок – 0,48 покупки в месяц. Клиенты этого кластера относительно пассивны – период времени от последней покупки составляет в среднем 13,8 месяцев.

Кластер 4. “Приверженцы”: клиенты, которые активно приобретают товары/услуги в данной компании, осуществляют платежи в пользу различных поставщиков услуг. Доля покупок в категории фото и видео, ноутбуки, непрофильные товары – самая высокая среди всех рассматриваемых кластеров. Средний чек у данной группы составляет 921,27 руб., самая большая частота покупок – 7,48 покупки в месяц.

Кластер 5. “Спящие”: клиенты, первоначально обращающиеся в компанию преимущественно за финансовыми услугами – денежный перевод, погашение кредита, страховые услуги. В дальнейшем приобретают базовые категории товаров: сотовые телефоны, аксессуары, тарифные планы. Самый длительный период неактивности – 33,7 месяцев. Средний чек – 1918,12 руб., частота покупок самая низкая – 0,15.

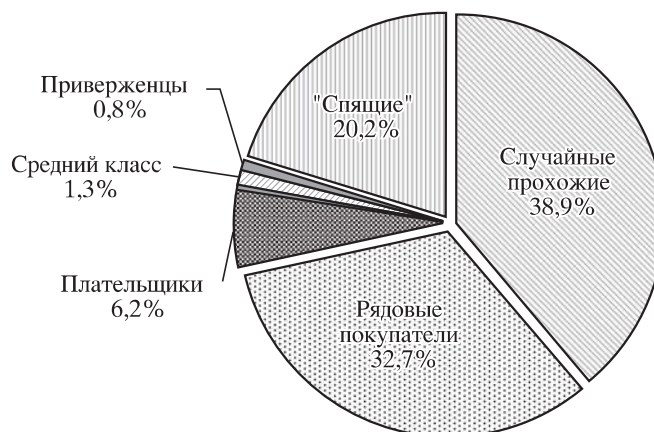
Кластер 6. “Случайные прохожие”: клиенты, совершившие за рассматриваемый период только 1 покупку. Как правило, это покупка профильного товара компании – сотовый телефон. Размер среднего чека – 2901,83 руб.

Как видно из анализа численности кластеров клиентской базы, 59,1% клиентов характеризует весьма низкая покупательная активность (рис. 3).

Анализ динамики изменения численности полученных кластеров позволяет выделить следующие основные этапы формирования и развития клиентской базы компании.

Этап 1 (рост). На данном этапе происходит интенсивный рост числа клиентов компании во всех кластерах. Основной рост происходит за счет притока новых клиентов, за исключением пятого кластера, который характеризует “спящих” клиентов.

Этап 2 (развитие). На данном этапе рост численности клиентской базы в це-

**Рис. 3.** Кластеризация клиентской базы компании

лом замедляется, и число клиентов по кластерам стабилизируется. Возрастает доля изменения численности клиентов в кластерах за счет межкластерных переходов, но при этом доля новых клиентов в общем притоке клиентов в кластер остается преобладающей.

Этап 3 (насыщение). На данном этапе приток новых клиентов практически прекращается, увеличивается доля неактивных клиентов. Численность клиентов в кластерах начинает постепенно убывать. С точки зрения управления клиентской базой компании на данном этапе значительно возрастает значение эффективных маркетинговых мероприятий, чтобы сократить выбытие клиентов и повысить их покупательскую активность.

7. АНАЛИЗ ДИНАМИКИ ИЗМЕНЕНИЯ КЛИЕНТСКИХ КЛАСТЕРОВ

Изменение численности кластеров вызывается рядом факторов. Для анализа перемещения клиентов между кластерами используется марковская цепь, состоящая из шести состояний: внешнего мира и пяти состояний, каждое из которых представляет собой кластер клиентов со следующими показателями: частота покупок; средний чек; время “жизни” клиента; время “сна” клиента. Каждый показатель определяется интервалом допустимых значений (минимальное и максимальное значение).

Каждый новый клиент в соответствии с величиной суммы совершенной покупки попадает в один из пяти сформированных ранее кластеров. При этом для клиентов, совершивших одну покупку, не предусматривается отдельного состояния. Если в течение заданного количества месяцев клиент не совершил повторной покупки, фиксируется его автоматическое выбытие из данного кластера. Клиент находится в состоянии i в момент времени t , если на момент времени t показатели его покупательского поведения удовлетворяют допустимому интервалу значений для данного состояния.

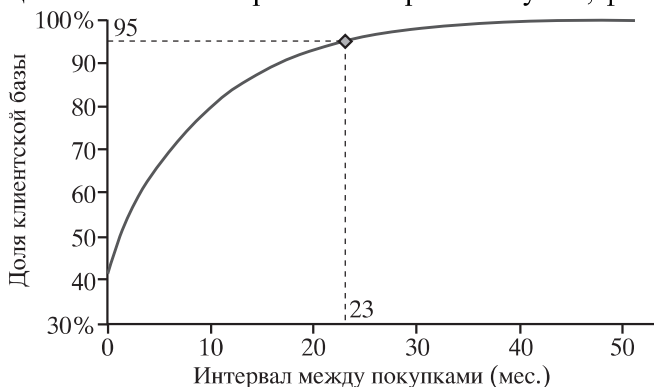


Рис. 4. Доля клиентов, совершивших повторную покупку

В рамках данного исследования в качестве периода времени, после которого считается, что клиент разорвал свои взаимоотношения с компанией, выбирается максимальный интервал времени между последовательными покупками, в течение которого 95% клиентов, попавших в выборку, неоднократно совершали покупки (рис. 4).

Как видно из рис. 4, максимальный временной интервал, в течение которого 95% клиентов совершают повторную покупку, составляет 23 месяца (или в среднем через 2 года). Это значение (23 месяца) было выбрано в качестве критерия отнесения клиента к активному или неактивному классу покупателей.

Полученные значения граничных интервалов для отнесения клиента к определенному состоянию марковской цепи сведены в таблицу ниже (табл. 4).

Замечание. Анализ на интервале длительностью 53 месяца (с ноября 2007 г. по март 2012 г.) в общем случае требует при расчетах среднего чека учитывать инфляцию цен на потребительские товары. Но на рынке товаров электронной техники наблюдается высокий темп технологического устаревания моделей товаров и услуг, что вызывает снижение цен на предыдущие модели, сопоставимое с размером инфляции, а иногда и значительно ниже. Поэтому в рамках данного исследования инфляционные изменения при расчете среднего чека клиента не учитывались. Тем не менее при адаптации модели для компаний, работающих в других секторах экономики, необходимо корректировать стоимостные показатели на величину инфляции.

В соответствии с выявленными граничными условиями было определено нахождение клиентов в выделенных состояниях для каждого момента времени (месяц) (см. табл. 4) и составлены матрицы переходов клиентов между кластерами.

Таблица 4. Границы интервалов ключевых показателей по кластерам

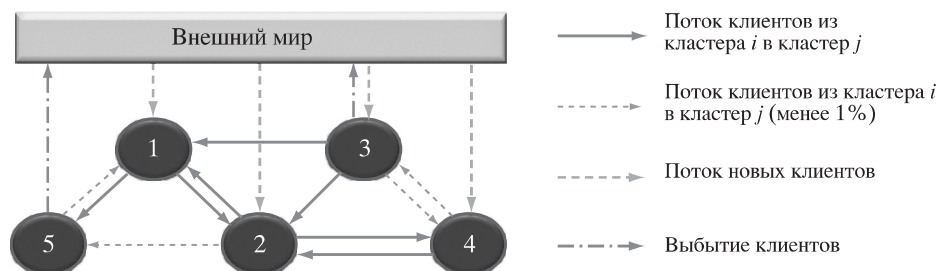
Кластер	Средний чек		Частота покупок		Период “активной жизни” клиента		Период “сна” клиента	
	Нижняя граница	Верхняя граница	Нижняя граница	Верхняя граница	Нижняя граница	Верхняя граница	Нижняя граница	Верхняя граница
1	10	8000	0	1,5	1	53	0	23
2	10	12 540	1,5	5	1	53	0	23
3	8001	122 000	0	5	1	53	0	23
4	10	33 100	5,01	42	1	53	0	23
5	10	8 000	0	1,5	20	53	19	23

По формуле (2) была рассчитана интенсивность переходов a_{ij} клиентов из кластера i в кластер j , по формуле (5) – численность клиентов по каждому кластеру – N_i^t . В результате было получено 20 временных рядов, содержащих 52 наблюдения, и проведен детальный анализ полученных данных.

Анализ полученных результатов показал, что основные переходы из кластера 1 происходят в кластеры 2 и 5. При этом интенсивность перехода в кластер 2 вначале резко падает, а с июля 2008 г. стабилизируется на уровне 0,6–1%. Переходы в кластер 5 начинают происходить с мая 2009 г. при достижении критичной границы по интервалу неактивности клиентов для этого кластера (19 месяцев). Переходы в кластеры 3 и 4 незначительные (менее 1%). Наибольшая интенсивность переходов из кластера 2 наблюдается в кластеры 1 и 4. При этом переход в кластер 1 характеризует снижение интенсивности в период наблюдения, а интенсивность перехода в кластер 4 со временем стабилизируется на уровне 0,5–1,5%. Переходы в кластеры 3 и 5 показывают незначительную интенсивность (менее 1% численности группы). Согласно полученным результатам основной поток перемещений клиентов из кластера 3 происходит в кластеры 1 и 2, при этом интенсивность переходов со временем снижается. Интенсивность переходов в кластер 4 – менее 1%, а переходы в кластер 5 отсутствуют. Основной поток клиентов из кластера 4 перемещается в кластер 2, интенсивность перехода в кластер 3 незначительная, в кластеры 1 и 5 клиенты из кластера 4 не переходят. Особенность кластера 5 состоит в том, что клиенты, попавшие в кластер 5, в дальнейшем прекращают взаимодействие с компанией, за исключением небольшой части клиентов – около 1–0,5% переходящих в кластер 1.

Проведенный анализ показал, что интенсивность притока новых клиентов во все кластеры клиентов с течением времени уменьшается. Особенно резкое снижение притока новых клиентов наблюдалось из кластеров с 1 по 4-й в июле–августе 2008 г. (возможно, данный спад связан с произошедшим в этот период мировым финансовым кризисом). Все это подтверждает выдвинутую гипотезу о насыщении клиентской базы компании. Преимущественное выбытие клиентов наблюдается из кластеров 3 и 5. Поток выбытия клиентов из остальных кластеров незначительный и составляет менее 1% численности группы.

На основании проведенного анализа были скорректированы потоки внутри марковской цепи с учетом полученных значений интенсивностей перехода (рис. 5).

**Рис. 5.** Скорректированная схема потоков между кластерами

Таким образом, в результате проведенного нами анализа было выявлено, что перемещение клиентов между кластерами клиентской базы характеризуется семью значимыми потоками между состояниями марковской цепи. Основной поток выбытия клиентов происходит из кластера 5, фактически представляющего собой поглощающее состояние, характеризующее класс “спящих” клиентов компании, из которого практически отсутствуют переходы в другие состояния. Небольшой поток выбытия клиентов наблюдается из кластера 3, или кластера клиентов, приобретающих дорогие товары и пользующихся финансовыми услугами компании. Приток клиентов возможен во все состояния, за исключением кластера 5.

8. ЗАКЛЮЧЕНИЕ

В ходе проведенного нами исследования были получены следующие результаты.

1. Построена информационно-логическая комплексная модель управления клиентской базой компании на основе показателя долгосрочной стоимости клиента, учитывающей характеристики покупательского поведения и социально-демографические отличия групп клиентов.

2. Выявлены группы клиентов, имеющие значимые различия по основным характеристикам покупательского поведения групп клиентов, характеризующегося частотой покупок, средним размером чека, общим временем взаимодействия клиента с компанией, давностью последней покупки.

3. Выявлен характер внешнего воздействия рынка и показано, что:

– изменение численности кластеров клиентов в значительной степени происходит за счет притока новых клиентов, а не межкластерных переходов;

– с течением времени на фоне снижения интенсивности притока новых клиентов в каждый кластер возрастает доля межкластерных переходов. Поэтому одним из возможных путей повышения прибыльности клиентской базы является стимулирование перехода клиентов в наиболее прибыльные кластеры.

4. Выявлены межкластерные перемещения клиентов, обусловленные изменением собственных потребностей клиентов.

5. На основе динамики изменения численности клиентских кластеров выделены основные этапы формирования и развития клиентской базы компании: рост, развитие и насыщение. Показано, что с точки зрения управления клиентской базой компании на этапе насыщения для сокращения выбытия клиентов и повышения их покупательской активности существенно возрастает значение адресно-направленных маркетинговых мероприятий.

Таким образом, проведенное нами исследование подтверждает, что подход к прогнозированию численности клиентской базы компании на основе расчета CLV клиентских кластеров позволяет выявить кластеры наиболее ценных для компании клиентов, обеспечивающих наибольшую долю доходов. Сделав акцент в предстоящей маркетинговой кампании на эти кластеры, можно повысить эффективность управления клиентской базой компании. В дальнейшем необходимо разработать инструментарий, который позволит менеджерам принимать более обоснованные решения, касающиеся маркетинговых мероприятий конкретной направленности.

СПИСОК ЛИТЕРАТУРЫ

- Андреева А.В.** (2010). Модель управления клиентской базой – новый шаг в развитии CRM? // *Директор информационной службы (CIO.RU)*. № 3. М.: Открытые системы. С. 26–28.
- Андреева А.В.** (2012). Оптимальное управление клиентской базой на основе показателя долгосрочной стоимости клиента // *Бизнес-информатика*. № 4(22). С. 61–68.
- Андреева А.В.** (2011). Разработка модели прогнозирования численности клиентской базы компании // *Аудит и финансовый анализ*. № 6. С. 104–109.
- Добровидова М.А.** (2003). Эффективные технологии повышения лояльности потребителей // *Маркетинг и маркетинговые исследования*. № 3(45). С. 48–53.

- Карасев А.П.** (2008). Разработка факторной модели лояльности для рынка услуг сотовой связи // *Маркетинг и маркетинговые исследования*. № 2(74). С. 98–111.
- Крюкова А.А., Кузьмин Е.В.** (2009). Разработка концепции комплексного управления клиентами // *Вестник СГЭУ: ежемесячный журнал*. № 7. С. 61–64.
- Ламбен Ж.-Ж.** (1996). Стратегический маркетинг. Европейская перспектива. СПб.: Наука.
- Мейер М.В.** (2004). Оценка эффективности бизнеса. М.: Вершина.
- Пепперс Д., Роджерс М.** (2006). Управление отношениями с клиентами. М.: Манн, Иванов и Фербер.
- Полежаев И.Е.** (2006б). Метод сегментации клиентских баз данных на основе жизненного цикла клиента. [Электронный ресурс] // *Исследовано в России*. С. 1875–1902. Режим доступа: <http://zhurnal.apec.relarn.ru/articles/2006/200.pdf>, свободный. Загл. с экрана. Яз. рус. (дата обращения: июль 2015 г.).
- Полежаев И.Е.** (2006а). Марковская модель для прогнозирования состояния клиентской базы данных. [Электронный ресурс] // *Исследовано в России*. С. 1903–1907. Режим доступа: <http://zhurnal.apec.relarn.ru/articles/2006/201.pdf>, свободный. Загл. с экрана. Яз. рус. (дата обращения: июль 2015 г.).
- Староверов О.В.** (1997). Азы математической демографии. М.: Наука. С. 56–59.
- Сьюэлл К., Браун П.** (2007). Клиенты на всю жизнь. М.: Манн, Иванов и Фербер.
- Третьяк О.А., Слоев И.А.** (2012). Оценка маркетинговой деятельности по состоянию клиентского потока // *Российский журнал менеджмента*. Т. 10. № 1. С. 29–50.
- Цысарь А.В.** (2002). Лояльность покупателей: основные определения, методы измерения, способы управления // *Маркетинг и маркетинговые исследования*. № 5(41). С. 55–61.
- Черкашин П.А.** (2004). Готовы ли Вы к войне за клиента? Стратегия управления взаимоотношениями с клиентами. М.: ИНТУИТ.РУ.
- Aaker D.A.** (1991). *Managing Brand Equity*. N.Y.: The Free Press.
- Berger P.D., Nasr N.L.** (1998). Customer Lifetime Value: Marketing Models and Applications // *Journal of Interactive Marketing*. Vol. 12. No. 1. Winter. P. 17–30.
- Berry M.J.A., Linoff G.S.** (2004). *Data Mining Techniques*. Indianapolis: Wiley Publishing, Inc.
- Dipak J., Siddhartha S.** (2002). Customer Lifetime Value Research in Marketing: A Review and Future Directions // *Journal of Interactive Marketing*. Vol. 16. No. 2. P. 34–45.
- Fader P.S., Hardie B.G.S., Lee K.L.** (2005). Counting Your Customers the Easy Way: An Alternative to the Pareto/NBD Model // *Marketing Science*. Vol. 24(2). P. 275–284.
- Fader P.S., Hardie B.G.S.** (2009). Probability Models for Customer-Base Analysis // *Journal of Interactive Marketing*. Vol. 23. P. 61–69.
- Hofmeyr J., Rice B.** (2000). *Commitment-Led Marketing*. Chichester: John Wiley and Sons.
- Malthous E.C., Blattberg R.C.** (2005). Can We Predict Customer Lifetime Value? // *Journal of Interactive Marketing*. Vol. 19. P. 2–16.
- Pfeifer P.E., Carraway R.L.** (2000). Modeling Customer Relationships as Markov Chains // *Journal of Interactive Marketing*. Vol. 14. Spring. P. 43–55.
- Reichheld F.** (1996). *The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value*. Boston: Harvard Business School Press.
- Schmittlein D.C., Morrison D.G., Colombo R.** (1987). Counting Your Customers: Who Are They and What Will They Do Next? // *Management Science*. Vol. 33. January. P. 1–24.

Поступила в редакцию
10.08.2015 г.

REFERENCES (with English translation or transliteration)

- Aaker D.A.** (1991). *Managing Brand Equity*. N.Y.: The Free Press.
- Andreeva A.V.** (2010). A Model of Client Based Management – a New Step in CRM? *IT Director (CIO.RU)* 3, 26–28 (in Russian).
- Andreeva A.V.** (2011). Developing a Model Predicting the Size of Customer Base. *Audit and Financial Analysis* 6, 104–109 (in Russian).

- Andreeva A.V.** (2012). Optimal Control of a Company's Customer Base Using the Customer Lifetime Value Parameter. *Business Informatics* 4(22), 61–68 (in Russian).
- Berger P.D., Nasr N.L.** (1998). Customer Lifetime Value: Marketing Models and Applications. *Journal of Interactive Marketing* 12, 1, Winter, 17–30.
- Berry M.J.A., Linoff G.S.** (2004). Data Mining Techniques. Indianapolis: Wiley Publishing, Inc.
- Cherkashin P. A.** (2004). Are You Ready to War for the Client? Strategy of Customer Relationship Management. Moscow: Intuit.ru (in Russian).
- Dipak J., Siddhartha S.** (2002). Customer Lifetime Value Research In Marketing: A Review And Future Directions. *Journal of Interactive Marketing* 16, 2, 34–45.
- Dobrovidova M.A.** (2003). Effective Techniques to Increase Customer Loyalty. *Marketing and Market Research* 3(45), 48–53 (in Russian).
- Fader P.S., Hardie B.G.S.** (2009). Probability Models for Customer-Base Analysis. *Journal of Interactive Marketing* 23, 61–69.
- Fader P.S., Hardie B.G.S., Lee K.L.** (2005). Counting Your Customers the Easy Way: An Alternative to the Pareto/NBD Model. *Marketing Science* 24(2), 275–284.
- Hofmeyr J., Rice B.** (2000). Commitment-Led Marketing. Chichester: John Wiley and Sons.
- Karasev A.P.** (2008). Development of the Loyalty Factor Model for the Telecom Market. *Marketing and Market Research* 2(74), 98–111 (in Russian).
- Krukova A.A., Kuzmin E.V.** (2009). Developing the Concept of CRM Complex. *Vestnik of Samara State University of Economics* 7, 61–64 (in Russian).
- Lamben Zh.–Zh.** (1996). Marketing Strategy: A New European Approach. SPb.: Nauka (in Russian).
- Malthous E.C., Blattberg R.C.** (2005). Can We Predict Customer Lifetime Value? *Journal of Interactive Marketing* 19, 2–16.
- Meyer M.W.** (2004). Rethinking Performance Measurement: Beyond the Balanced Scorecard. Moscow: Vershina (in Russian).
- Peppers D., Rogers M.** (2006). Managing Customer Relationships. A Strategic Framework. Moscow: Mann, Ivanov, and Ferber (in Russian).
- Pfeifer P.E., Carraway R.L.** (2000). Modeling Customer Relationships as Markov Chains. *Journal of Interactive Marketing* 14(Spring), 43–55.
- Polezhaev I.E.** (2006a). Customer Database Segmentation Based on Customer Lifecycle Techniques. *Investigated in Russia*, 1875–1902. Available at: <http://zhurnal.ape.relarn.ru/articles/2006/200.pdf> (accessed: July 2015, in Russian).
- Polezhaev I.E.** (2006b). Markov Model to Predict the Client Database State. *Investigated in Russia*, 1903–1907. Available at: <http://zhurnal.ape.relarn.ru/articles/2006/201.pdf> (accessed: July 2015, in Russian).
- Reichheld F.** (1996). The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value. Boston: Harvard Business School Press.
- Schmittlein D.C., Morrison D.G., Colombo R.** (1987). Counting Your Customers: Who Are They and What Will They Do Next? *Management Science* 33(January), 1–24.
- Sewell C., Braun P.** (2007). Customers for Life. How to Turn That One-Time Buyer Into a Lifetime Customer. Moscow: Mann, Ivanov, and Ferber (in Russian).
- Staroverov O.V.** (1997). Foundations of Mathematical Demography. Moscow: Nauka, 56–59 (in Russian).
- Tretyak O.A., Sloev I.A.** (2012). Client's Flow Based Evaluation of Marketing Activities. *Russian Management Journal* 10, 1, 29–50 (in Russian).
- Tsisar A. V.** (2002). Customer Loyalty: Basic Definitions, Measurement and Control Methods. *Marketing and Market Research* 5(41), 55–61 (in Russian).

Predicting the Size of the Company-Customer Base Based on the Markov Chains

A.V. Andreeva, T.K. Bogdanova

The article represents the information and logical complex model to manage the company's customer base in order to calculate the index of long-term value of the client. In contrast to the previous issues this model takes into account the peculiarities of consumer behavior and socio-demographic characteristics of the client groups, and the movement of customers within the customer base is represented as a Markov chain. Developed complex dynamic model to manage the company's customer base makes possible to forecast the cluster frequency and the client base at all on any period, to analyze the population-change-dynamics of client cluster and to highlight the most important stages in the formation and development of the customer base. On each stage highlighted groups of clients are estimated according to future company's profit and potential for cross selling of products/services. The estimates allow to calculate the budget for marketing activities on retention of given profit level and to compose the most effective plan of address-directed marketing activities.

Keywords: management of customer base of the company; retail trade; group of clients; customer loyalty; buying behavior; Markov chain; dynamic model; long-term value of the customer base.

Classification JEL: C15, C32, C53.