



РАСШИРЕННАЯ CRM ТЕХНОЛОГИЯ

КАК ЕСТЕСТВЕННЫЙ ИСТОЧНИК ОБЪЕКТИВНЫХ МАРКЕТИНГОВЫХ ДАННЫХ

А.В. Тимофеев (ООО "Центр речевых технологий")

Рассмотрена методика сбора и фиксации численных параметров динамики рыночных тенденций, встроенная в методологию CRM и имеющая целью естественного включения менеджеров по продажам в процесс сбора маркетинговой информации. Для автоматической обработки и оценивания информативности полученных оценок, использован оригинальный алгоритм, основанный на применении α -stable моделей. Описан пример реализации предложенной методологии на платформе MS CRM 3.0.

Описание проблемы

Постоянный мониторинг рыночных тенденций чрезвычайно важен на конкурентном рынке в качестве базового фактора в обеспечении повышения рыночной эффективности компании. Без мониторинга и последующего анализа данных, полученных в результате этого процесса, а также дальнейшего использования этих данных для улучшения бизнес-процессов компании, само существование бизнеса рано или поздно будет неминуемо поставлено под вопрос. Практически все успешные компании с той или иной степенью эффективности реализуют мониторинговые процессы, используя для этих целей различные источники информации и данные специализированных служб. Среди источников оперативных данных о рыночных тенденциях менеджеры по продажам играют далеко не последнюю роль. Именно они непосредственно контактируют с клиентами и имеют значительные возможности по сбору релевантной информации из первых рук как той, что может быть выражена в численных данных, так и той, что носит качественный характер. Однако вопрос стимулирования продавцов к аккуратному и системному вводу данных, полученных от клиента, в большинстве компаний является далеко не столь простым и очевидным делом, как это может показаться на первый взгляд. Не последнюю роль в решении этой проблемы играет наличие или отсутствие удобных и достаточно наглядных способов фиксации данных с обязательной возможностью их дальнейшего автоматизированного анализа в информационной системе, которой пользуются продавцы компании.

Целью работы является описание методологии построения принципов фиксации маркетинговых данных в информационной системе продавцами, а также формальных способов анализа этих данных как с целью улучшения бизнес-процессов компании, так и с целью изменения фокуса маркетинговой активности.

Основной результат

В настоящее время основным информационным инструментом, которым пользуются продавцы в своей ежедневной бизнес-деятельности, являются CRM сис-

темы. О сути CRM технологий написано и сказано достаточно много, чтобы не повторять эти тезисы в очередной раз. Поэтому рассмотрим дополнительные возможности CRM технологий в части сбора сегментированной информации о рыночных тенденциях, а также формальных принципах анализа этой информации в предположении, что результаты этого анализа в дальнейшем будут использованы для коррекции параметров бизнес-процессов компании. Технологию такого рода будем называть расширенной CRM технологией.

Введем некоторые обозначения и понятия. Обозначим: m_s – продавец; c_d – клиент, $c_d \in C$, C – множество клиентов; $z(m, c_d, \theta)$ – сделка (проект) типа θ , который ведут продавец $m_1, m_2, \dots$, с клиентом c_d , здесь $m = \{m_1, m_2, \dots\}$; $\theta = \theta(z(m, c_d, \theta))$ – тип сделки (проекта); $\theta = \theta(z(m, c_d, \theta)) \in \Theta$, Θ – множество допустимых типов сделок; $s_j(\theta, z(m, c_d, \theta))$ – номер этапа сделки $z(m, c_d, \theta)$, $j \in J$, J – множество допустимых индексов этапов. Совокупная последовательность этапов сделки типа θ совместно с группой обязательных артефактов каждого этапа и правил их активации образуют шаблон (модель) $\Lambda(\theta)$ эскалации сделок типа θ . Под обязательными артефактами этапа понимается минимальная совокупность бизнес-активностей, которые необходимо реализовать на конкретном этапе в определенной последовательности с целью полной реализации данного этапа.

Для упрощения изложения, будем считать, что $z(m, c_d, \theta)$ обозначает сделку типа θ на всех этапах ее развития, начиная от этапа первичного контакта до этапа завершения сделки. Фиксация множеств Θ и J , а также шаблоны эскалации продаж $\Lambda = \{\Lambda(\theta) | \theta \in \Theta\}$ являются естественным результатом внедрения CRM системы. Далее, пусть $P = \{p_1, p_2, \dots, p_m\}$ – множество характеристических параметров (ХП), о сути использования которых будет более подробно сказано далее. Каждый параметр имеет тип, который обозначается символом $H(p_i)$ и принадлежит множеству $\{C, K\}$, C – числовой, K – качественный. То есть: $\forall p_i \in P: H(p_i) \in \{C, K\}$. Если параметр числовой, то в качестве его значений принимается некая скалярная величина из множества рациональных чисел, имеющая понятийно интерпретируемый смысл. Если параметр качест-

венный, то в качестве его значений выбираются числовые строки типа "хороший", "плохой", "малопригодный" и др.

Естественно, что множества значений качественных параметров должны быть оцифрованы с использованием техники размытых множеств Лотфи-Заде [1, 2, 3]. Подход, основанный на квантификации по технике Лотфи-Заде, дает очевидные преимущества при автоматизированном анализе данных. Стоит отметить, что в ряде случаев может оказаться достаточным задание отношений частного порядка на множествах значений качественных признаков, которые могут быть использованы при реализации процедур предпочтения. При использовании аппарата размытых множеств значения степени принадлежности соответствующему множеству должны быть нормированы к единице. Заметим, что в случае использования для квантификации значений качественных параметров техники размытых множеств, все ХП как "числовые", так и "качественные" будут иметь исключительно численные и положительные значения.

Обозначим символом $\chi(p_i) = \{p_{1i}, p_{2i}, p_{3i}, \dots\}$ — множество допустимых значений параметра $p_i \in P$. Каждый этап сделки характеризуется соответствующим подмножеством ХП: $P(s_j(\theta, \cdot)) \subseteq P$. Эти подмножества априорно фиксированы для каждого этапа воронки продаж по конкретному типу сделок θ . Состав подмножеств $P(s_j(\theta, \cdot))$, $j \in J$ может изменяться после апостериорного анализа реализации множества воронок продаж. Суть этих изменений определяется дополнительными значениям параметров или новыми параметрами, которые создают продавцы в процессе продаж.

Предполагается, что в процессе реализации предпродажной активности продавец, описывая сделку на конкретном этапе ее развития, будет обязан назначить каждому из априорно заданных ХП конкретное значение. В результате работы продавцов над сделками определенного типа в течение периода T , длительность которого зависит от типа сделки, будет собран материал, характеризующий в терминах соответствующих подмножеств $P(s_j(\theta, \cdot)) \subseteq P$ поэтапную реализацию соответствующей воронки продаж. Итак, реализуется следующая совокупность подмножеств: $P(T, \theta) = P(s_j(\theta, z(m_z, c_d, \theta))), j \in J, z(m_z, c_d, \theta) \in Z(T, \theta)$. Здесь $Z(T, \theta)$ — множество сделок типа θ , реализованных в периоде T , m_z — множество продавцов, работающих над конкретной сделкой $z(m_z, c_d, \theta) \in Z(T, \theta)$ типа θ , при этом $p \in P(T, \theta)$, $p \in \chi(p)$. Если ХП $p \in P(T, \theta)$ используется для характеристики j -го этапа сделки $z(m_z, c_d, \theta)$, будем обозначать реализацию этого параметра следующим образом: $p(z|j)$. Для j -й стадии воронки продаж типа θ в результате ввода значений соответствующих ХП менеджерами по продажам будет фиксировано множество $R_j(p, T, \theta) = \{p(z|j) | p(z|j) \in \chi(p), p \in P(T, \theta), z \in Z(T, \theta)\}$, состоящее из реализаций ХП $p \in P(T, \theta)$. Распределение значений параметра p носит случайный характер, который определяется конкретными обстоятельствами каждой конкретной

сделки, рыночными тенденциями, особенностями клиента и прочими обстоятельствами.

О возможном вероятностном законе распределения статистики $R_j(p, T, \theta)$ следует поговорить отдельно и более подробно. Бизнес-процессы компании и рыночный сегмент, в котором компания использует эти бизнес-процессы, в совокупности образуют информационную систему открытого типа (ИСОТ), состояние которой косвенно описывает множество ХП этой системы. Не все элементы этого множества представляют одинаковую ценность для анализа и принятия решений, но так как априорно трудно фиксировать наиболее значимые ХП, окончательная фиксация подмножества значимых ХП реализуется в процессе обработки результатов наблюдения реализации ХП. Распределение значений конкретного ХП системы может характеризоваться большей или меньшей степенью энтропии. Если в ИСОТ имеет место некая доминирующая тенденция, она имплицитно повышает концентрацию вероятностного распределения значений тех ХП, которые наиболее адекватно отражают параметры динамики ИСОТ. В этом случае энтропия ХП, а следовательно, и энтропия ИСОТ, будут снижаться. Если же доминирующей тенденции нет, энтропия ХП, а следовательно, и энтропия ИСОТ, будут повышаться. Если конкретный ХП слабо отражает динамику ИСОТ, энтропия его распределения не будет обусловлена динамикой ИСОТ, поэтому этот ХП желательно исключать из дальнейшего анализа. Связь динамики ХП с динамикой ИСОТ фиксируется методом обычного корреляционного анализа, по результатам которого следует оставить в множестве P только те ХП, которые характеризуют динамику ИСОТ с требуемой чувствительностью, определяемой величиной коэффициента корреляции. Проблема, однако, состоит в том, что для полноценной реализации корреляционного анализа нужны выборки репрезентативного объема, получить которые при условии постоянной коррекции совокупности бизнес-процессов компании — нереально. Поэтому на практике приходится обрабатывать все множество P , принимая решение об аналитической значимости конкретных ХП лишь на основании загрубленных критериев, некоторые из которых будут рассмотрены далее. Если энтропия распределения конкретного ХП системы велика, то принять управленческое решение на основе анализа конкретных реализаций этого параметра будет затруднительно, так как все реализовавшиеся значения будут в определенном смысле равноправны, и как следствие, — неинформативны.

Как правило, практический интерес представляют именно распределения ХП с низкой энтропией, которые характеризуют распределения, возникшие в результате влияния какого-либо доминирующего фактора или тенденции. Именно ХП этого типа обладают наибольшей аналитической значимостью, а следовательно, именно они должны быть использованы при коррекции фокуса маркетинговой активности.

Пусть $f(p|j, T, \theta)$ — плотность распределения параметра p . В этом случае, энтропия характеристического параметра p будет определяться простой зависимостью: $H(p|j, T, \theta) = -E(\text{Log}(f(p|j, T, \theta)))$. ХП p_i будем считать аналитически более значимым (содержательным) в сравнении с ХП p_k , если $H(p_i|j, T, \theta) < H(p_k|j, T, \theta)$. Очевидно, что аналитически более значимые параметры обладают более высокой степенью кумулятивности (концентрированности) плотности их вероятностного распределения. В качестве аппроксимации закона распределения статистики $R(p, T, \theta)$ разумно использовать класс так называемых α -stable вероятностных распределений. Несмотря на определенные сложности технического плана, обусловленные тем, что у подавляющей части законов этого класса отсутствуют моменты первого и второго порядков, использование этой аппроксимации оказывается практически целесообразным и очень эффективным. Дело в том, что класс α -stable кроме универсального нормального закона включает группу законов краевого типа, так называемых power laws, которые являются достаточно адекватными моделями распределений с низкой энтропией. А ведь законы с низкой энтропией чаще всего встречаются в ИСОТ, вследствие воздействия на нее определенных доминирующих тенденций, возникающих, например, на конкурентном рынке в условиях значительного дисбаланса ресурсов рыночных игроков.

Известно [4], что переменная X подчинена закону из класса α -stable, если характеристическая функция этого распределения имеет вид:

$$\begin{aligned} \log \varphi(t) &= \\ \{j\mu t - \sigma^\alpha |t|^\alpha [1 - \beta \text{sign}(t) \tan(\alpha\pi/2)]\}, &\text{ если } \alpha \neq 1 \\ \log \varphi(t) &= \\ \{j\mu t - \sigma^\alpha |t|^\alpha [1 + \beta \text{sign}(t) \tan(\alpha\pi/2)]\}, &\text{ если } \alpha = 1 \end{aligned} \quad (1)$$

Здесь $\mu \in R^1$, $\sigma > 0$, $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$. Параметры $\Omega = (\alpha, \beta, \gamma, \mu)$ полностью определяют α -stable распределение и имеют следующий смысл: α — показатель стабильности, который характеризует импульсивность распределения. При малых значениях этого параметра соответствующие распределения имеют более тяжелые хвосты и характеризуются большей степенью импульсивности. По мере роста этого параметра к величине 2, хвосты распределения все более облегчаются, а вид распределения становится все менее импульсивным, постепенно приближаясь к гауссиану. Далее, β есть показатель асимметричности (skewness parameter), при $\beta = 0$ распределение становится симметричным (такие распределения образуют класс *symmetric α -stable (SaS)*). Параметр $\sigma > 0$ — параметр масштаба (scale parameter), который характеризует степень рассеяния (дисперсии) распределения и в случае гауссиана совпадает со средним квадратическим отклонением. μ — параметр сдвига определяет положение распределения и в случае гауссиана совпадает с математическим ожиданием. Для классов α -stable, при $\alpha = 1$ имеем дело с распределением Коши, а при $\alpha = 2$ — с обычным нормальным распределением, причем оба эти распределения принад-

лежат классу *SaS*. Напомним, что в классе законов распределения с фиксированными дисперсией и математическим ожиданием максимальная энтропия достигается именно на гауссовом распределении вероятностей ($\alpha = 2$) [8]. При значении дисперсии, стремящейся к бесконечности, максимальная энтропия достигается на распределениях, близких к равномерному распределению. При $\alpha < 2$ законы распределения рассматриваемого класса иногда называют Парето-подобными.

Итак, необходимо по выборке $R(p, T, \theta)$ определить оценку $\Omega'(R(p, T, \theta))$ параметров $\Omega = (\alpha, \beta, \sigma, \mu)$ плотности модельного распределения $f(p|\Omega, j, T, \theta)$ из класса α -stable и на основании полученной аппроксимации $f(p|\Omega', j, T, \theta)$ построить доверительное (критическое) множество значений ХП. Эти множества будут являться основой для принятия решения о наличии определенной тенденции в динамике ИСОТ, что является естественной базой для выстраивания адекватной маркетинговой коррекции. Информативность (информационная ценность) полученных выводов можно характеризовать параметрами $\Omega' = (\alpha', \beta', \sigma', \mu')$: в частности, чем ближе α' будет к 1, и чем меньше при этом по величине будет параметр σ' , тем более информативное множество ХП было получено. Для оценивания вектора параметров θ разумно использовать метод сведения к линейной регрессии [5]. Рассмотрим оценку:

$$\varphi'_n(t) = \sum_{j=1}^n e^{itx_j}, \quad x_j \in R(p, T, \theta).$$

Используя (1), получим:

$$\log(-\log |\varphi(t)|^2) = \log(2\sigma^2) + \alpha \log |t| \quad (2)$$

Для $\alpha \neq 1$ имеем:

$$\begin{aligned} \text{Re } \varphi(t) &= \exp(-|\sigma t|^\alpha) \cos[\mu t + |\sigma t|^\alpha \beta \text{sign}(t) \tan(\pi\alpha/2)], \\ \text{Im } \varphi(t) &= \exp(-|\sigma t|^\alpha) \cos[\mu t + |\sigma t|^\alpha \beta \text{sign}(t) \tan(\pi\alpha/2)]. \end{aligned}$$

Откуда получаем:

$$\text{Arctan}(\text{Im } \varphi(t) / \text{Re } \varphi(t)) = \mu t + \beta \sigma^\alpha \tan(\pi\alpha/2) \text{sign}(t) |t|^\alpha. \quad (3)$$

Уравнение (2) зависит только от параметров α, σ , что дает возможность оценить их из следующей регрессии

$$y(t_k) = m + \alpha \omega(t_k) + \varepsilon(t_k), \quad k = 1, 2, \dots, K.$$

Здесь $y(t_k) = \log(-\log |\varphi'(t_k)|)$, $\omega(t_k) = \log |t_k|$, $\varepsilon(t_k)$ — ошибки наблюдений, $m = \log(2\sigma^2)$. Следуя [5], определим $K \in [9, 134]$, $t_k = \pi k / 25$. Выбор конкретного значения K определяется объемом наблюдаемой выборки и значением параметра α . Чем больше значение параметра α , тем больше значение величины K . В [5] предложено для диапазона значений α [1.5, 2] рассматривать $K = 10$, для диапазона [0.6, 1.5] значение $K = 20$, для диапазона [0.4, 0.6] значение $K = 60$, а при $\alpha < 0.4$ использовать $K = 120$. После получения оценок α', σ' , параметры β, μ оцениваются с использованием соотношения (3) с применением оценок α', σ' . В частности, используя (3), получим следующее уравнение регрессии:

$$z(\mu, \beta | s) = \mu u(s) + \beta \sigma^\alpha \tan(\pi\alpha/2) \operatorname{sign}(u(s)) |u(s)|^\alpha + \eta(s), s = 1, 2, \dots, S, \quad (4)$$

вместо параметров α, σ используются их оценки α', σ' , $\{\eta(s) | s = 1, 2, \dots, S\}$ – последовательность ненаблюдаемых случайных помех. Следуя Koutrouvelis [5], определяем $u(l) = \pi l/50$, $L \in [9, 50]$.

Получив, таким образом, оценку вектора параметров θ , можно оценить плотность распределения численно [6, 7]. Программы численного оценивания плотности α -stable доступны по ссылке: <http://academic2.american.edu/~jpnolan/stable/stable.html>. Далее, используя построенные оценки плотности, легко получить доверительные интервалы с требуемым коэффициентом доверия для ХП. При этом оценка параметра α' может быть использована в качестве грубой оценки аналитической значимости ХП. Чем меньше соответствующее значение α или чем ближе значение этого параметра к 0,5 (распределение Леви), тем более информативной можно считать оценку конкретного ХП, и тем большее внимание ему следует оказывать в процессе коррекции маркетинговой активности. Однако более корректным способом оценивания аналитической значимости ХП следует признать метод непосредственного оценивания величины выборочной энтропии $H(p | \Omega', j, T, \theta)$, которое основано на использовании численных оценок плотности распределения $f(p | \Omega', j, T, \theta)$ (класс α -stable). В этом случае, чем меньше значение выборочной энтропии $H(p | \Omega', j, T, \theta)$, тем более информативен соответствующий ХП.

Реализация методологии на платформе MS CRM 3.0

Технологические аспекты реализации расширенной CRM технологии на платформе MS CRM 3.0 достаточно очевидны. В качестве носителя данных о значениях ХП используется специальный объект "Маркетинговые параметры", созданный средствами дизайнера MS CRM 3.0. Атрибуты этого объекта, с одной стороны, содержат искомое значение ХП, определенное для конкретного этапа j процесса продажи $z(m, c_d, \theta)$, а с другой – четко связаны с шаблоном эскалации продаж $\Lambda(\theta) \in \Lambda$, который в свою очередь соответствует конкретному типу сделки θ . Управление эскалацией продаж в MS CRM реализуется с использованием правил, задаваемых с помощью инструмента MS Workflow Manager и инициирующихся при создании события создания объекта "Возможная сделка". Объект "Маркетинговые параметры" включается в шаблоны эскалации продаж стандартными средствами MS WorkFlow Manager. При этом правила эскалации процесса продаж настраиваются так, чтобы обеспечить полное заполнение всех обязательных к заполнению полей этого объекта.

В результате, за период T формируется совокупность статистик $R_j(p, T, \theta)$, $p \in P$, $\theta \in \Theta$, $j \in J$, на осно-

вании которых строится множество интервальных оценок характеристических параметров. Для построения этого множества используется специальная утилита, реализующая алгоритмы построения оценок плотностей α -stable распределений с дальнейшим интервальным оцениванием ХП, описанные в предшествующем разделе статьи. Утилита реализована в виде специальной сборки и получает доступ к совокупности экземпляров объекта "Маркетинговые параметры" посредством MS CRM SDK 3.0. Результатом этой обработки является набор интервальных оценок для характеристических параметров маркетинговой активности, построенных для фиксированного значения коэффициента доверия. Аналитическая важность (информационная полезность) каждого ХП характеризуется как значением параметра α , которое вычисляется одновременно с границами доверительного интервала, так и прямой оценкой энтропии распределения ХП.

Заключение

Предложенная методология оценивания числовых характеристических параметров динамики рыночных тенденций позволяет естественным образом включить менеджеров по продажам в процесс маркетинговых исследований. Это разрешает массу проблем, связанных с информационными нестыковками, часто существующими между отделом продаж и маркетинговым отделом. Предлагаемый алгоритм автоматического оценивания аналитической значимости параметров основан на использовании α -stable моделей и является новационным, позволяя объективно оценивать степень информационной полезности каждого из используемых характеристических параметров. Предложенная методология естественным образом встраивается в обычную CRM методологию, что доказывается относительной легкостью ее реализации на платформе MS CRM 3.0.

Список литературы

1. Zadeh L. Toward a Generalized Theory of Uncertainty. GTU, Information Sciences, Elsevier. Vol. 172. 2005.
2. Zadeh L. From Search Engines to Question-Answering Systems-The Role of Fuzzy Logic, Progress in Informatics, № 1, 1-3, 2005.
3. Мацневский С.В. Нечеткие множества: Учебное пособие. – Калининград: Изд. КГУ. 2004.
4. Zolotarev V.M. One-Dimensional Stable Distributions, American Mathematical Society, 1980.
5. Koutrouvelis I.A. Regression-Type Estimation of the Parameters of Stable Laws, Journal of the American Statistical Association 75, 1980.
6. Mittnik S., Doganoglu T. and Chenyao D. Computing the Probability Density Function of the Stable Pareto Distribution, Mathematical and Computer Modelling 29, 1999.
7. Borak Sz., Härdle W. and Weron R. Stable Distributions, in P. Cizek, W. Härdle, R. Weron (eds.) Statistical Tools for Finance and Insurance, 1980, Springer
8. Пугачев В.С. Введение в теорию вероятностей. М.: Наука, 1968.

Тимофеев Андрей Владимирович – д-р техн. наук, начальник ИТ-департамента ООО "Центр речевых технологий".
E-mail: timofeev.andrey@gmail.com