

*Игорь Николаевич ГЛУХИХ —
проректор по информационным технологиям,
доктор технических наук, профессор*

*Антон Павлович МООР —
аспирант кафедры информационных систем*

УДК 004.9:345.5

ПРИМЕНЕНИЕ НЕЧЕТКОЙ ЛОГИКИ В МОДЕЛИРОВАНИИ ИНФОРМАЦИОННОЙ СИСТЕМЫ ОБНАРУЖЕНИЯ МОШЕННИЧЕСТВА С ИСПОЛЬЗОВАНИЕМ ПЛАСТИКОВЫХ КАРТ

АННОТАЦИЯ. В статье описана математическая модель информационной системы обнаружения мошенничества с использованием пластиковых карт на основе методов нечеткой логики. Рассматриваются различные аспекты построения экспертной системы, реализующей функции обнаружения мошенничества с использованием пластиковых карт.

The article is devoted to fuzzy mathematical model of card fraud detection information system. It considers various aspects of fraud management information system development.

В современных условиях интенсивного развития финансового рынка, с увеличением количества и суммы банковских операций, совершаемых с использованием пластиковых карт, а также с проникновением электронных средств платежа во многие сферы жизни, потери от мошенничества с пластиковыми картами составляют значительную долю затрат финансовых институтов, работающих с электронными средствами платежа.

Преступления, связанные с пластиковыми картами, являются весьма привлекательными в силу разнообразных причин. В обществе сложился образ преступника, совершающего мошенничество с пластиковыми картами — хорошо одетый образованный молодой человек, зарабатывающий деньги своим умом. Деньги перетекают в карман преступника без насилия и конфронтации [1].

Мошенничество с использованием пластиковой карты — это манипуляции с платежной системой, в результате которых денежные средства держателя карты или другого полноправного участника платежной системы незаконным образом переходят во владение другого лица или группы лиц. Необходимо отметить, что денежные средства могут перейти в другие руки косвенным способом: украденная карта может послужить мошенникам средством оплаты при покупке товара.

Мошенничество может быть совершено лицами, так или иначе причастными, к платежной системе. Наиболее часто встречающимся видом мошенничества является кража или подделка пластиковой карты посторонним лицом, но, к сожалению, известны случаи, когда к мошенничеству причастны сотрудники торгово-сервисного предприятия или банка, работающего с пластиковыми картами [3].

Даже если участнику платежной системы удалось избежать финансовых потерь в результате мошенничества, факт удавшегося мошенничества оказывает резко негативное влияние на престиж банка, как среди его клиентов, так и среди других участников платежной системы.

Таким образом, методы построения информационных систем обнаружения мошенничества являются весьма актуальной темой научных исследований.

Задача обнаружения мошенничества с использованием пластиковых карт может быть в общем виде сформулирована следующим образом.

Имеется набор финансовых операций с пластиковыми картами в платежной системе. Из множества всех операций необходимо выделить мошеннические, то есть неправомерные операции, которые были совершены с умыслом незаконного получения денежных средств.

Операции с пластиковыми картами имеют множество характеристик, таких как, например, дата, время, сумма, место операции. Задачу выявления мошеннических операций можно осуществлять с помощью анализа этой информации. В силу того, что вся информация создается и передается в электронной среде, а также по причине большого количества операций, решение задачи должно воплотиться в виде информационной системы, которая предоставляет возможность давать ответ на основной вопрос задачи.

Очевидно, что ответ на вопрос задачи не может быть получен с полной достоверностью. Естественной мерой эффективности системы, решающей задачу обнаружения мошенничества, будет экономическая и статистическая эффективность.

Математическая модель информационной системы обнаружения мошенничества с использованием пластиковых карт на основе нечеткой логики может быть описана следующим образом.

- множество всех операций T ;
- признаки операции — набор векторов $H(t)$, $t \in T$;
- характеристическая векторная функция — $Z(H(t))$, определяющая вычисляемые параметры операции;
- попарно непересекающиеся четкие множества A_r (определенно мошеннические операции), B_r (определенно аутентичные операции), C_r (неопределенные операции), $A_r \cup B_r \cup C_r = T$. Множества A_r и B_r могут быть пустыми;
- набор знаний экспертов, выраженных в лингвистической форме.

На основе этих данных необходимо определить степень принадлежности элемента t из множества T нормальному [2] нечеткому подмножеству A мошеннических транзакций, то есть найти значение $F_A(t)$, где F_A — характеристическая функция принадлежности [5], принимающая значения в диапазоне $[0;1]$.

Нахождение явного вида функции F_A является затруднительным в силу большого количества параметров транзакции t и неопределенности ее значений для каждого элемента из множества T . Для того чтобы вычислить значение функции F_A для любой заданной $t \in T$, воспользуемся аппаратом нечеткого вывода.

Введем лингвистические переменные x_i , то есть характеристики транзакции или ситуации, выраженные экспертами в лингвистической, словесной форме, сопоставим им нечеткие переменные A_1, A_2, \dots, A_n с функциями принадлежности $F_{A_1}, F_{A_2}, \dots, F_{A_n}$. Высказывание « x_i есть A_i » назовем частным критерием мошенничества.

Введем также специальную лингвистическую переменную P «степень принадлежности транзакции множеству мошеннических», которая принимает значения нечетких переменных Q_i .

Сформулируем систему предикатных правил вида «Если x_i является A_i , то P является Q_i » (где i — номер правила) и, используя один из алгоритмов нечеткого вывода, получим степень принадлежности элемента t множеству A .

В качестве простейшего примера рассмотрим систему, состоящую из двух правил:

«Если количество предыдущих транзакций по карте за сутки очень велико, то степень принадлежности транзакции мошенническим = 0.9»

«Если сумма транзакции незначительна, то степень принадлежности транзакции мошенническим = 0.3»

В этих правилах:

x_1 = «количество предыдущих транзакций по карте за сутки»;

A_1 = «очень велико»;

Q_1 = «=0.9»;

x_2 = «сумма транзакции»;

A_2 = «незначительна»;

Q_2 = «=0.3».

Легко заметить, что переменные Q_1 и Q_2 определены на четких подмножествах. Но, принимая во внимание, что четкие множества являются частным случаем нечетких множеств, их использование здесь допустимо. Кроме того, такой вид правил позволяет нам использовать для этого примера алгоритм Сугэно 0-го порядка (известный также как упрощенный алгоритм нечеткого вывода) [5].

Этот алгоритм состоит в следующем: после вычисления точных значений функций $\alpha_i = F_{A_i}(x_i)$ для заданных входных значений x_i значение переменной P вычисляется по формуле

$$P^* = \frac{\sum_i \alpha_i * Q_i}{\sum_i \alpha_i}.$$

То есть, например, для заданной ситуации, «снятие 100 рублей в банкомате, пятое за сутки» этот алгоритм может работать так:

1. По заданным F_{A_i} будет посчитано, что $\alpha_1=0.4$ (5 транзакций — это не очень много), а $\alpha_2=0.6$ (100 рублей — сумма чуть выше средней);

2. $P^* = \frac{0.4 * 0.9 + 0.6 * 0.3}{0.4 + 0.6} = 0.54$, что можно интерпретировать, как «транзакция может быть мошеннической, а может быть аутентичной, оснований для особой тревоги нет».

В качестве функции принадлежности F_{A_i} можем выбрать, например, кусочно-линейную функцию

$$F_{A_i} = \begin{cases} 0, & \text{при } x \leq 0 \\ \frac{x}{\lambda}, & \text{при } 0 < x < \lambda \\ 1, & \text{при } x \geq \lambda \end{cases},$$

где x — количество транзакций за сутки, λ — параметр, задаваемый экспертом, как число транзакций, определяемое как «безусловно большое». График этой функции изображен на рис. 1.

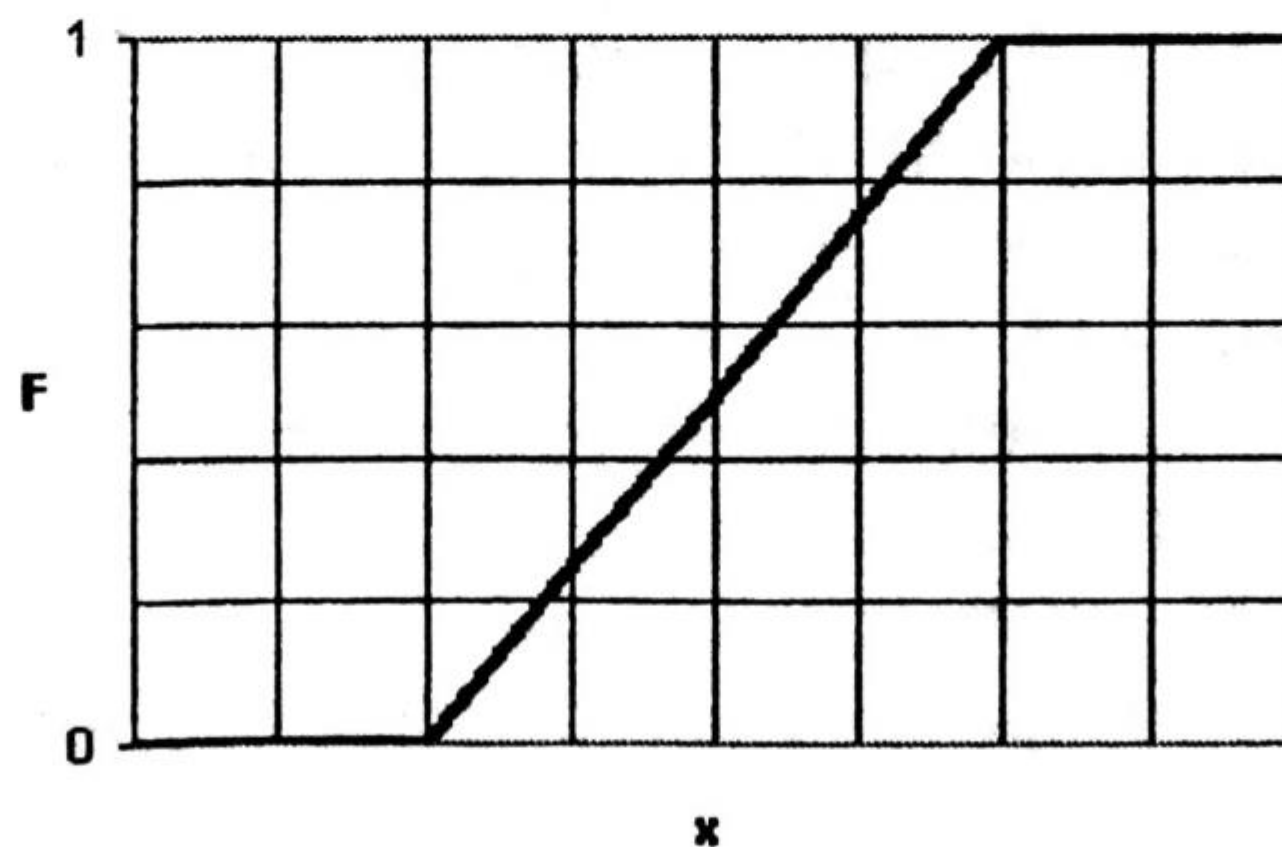


Рис. 1. Кусочно-линейная функция принадлежности

На практике, впрочем, применение кусочно-линейных функций принадлежности ограничено грубостью их аппроксимации и слабой гибкостью. Поэтому

может оказаться целесообразным использование более сложных функций, таких как, например, показательные или Гауссовы [5] функции, которые, несмотря на большую, по сравнению с кусочно-линейными функциями, сложность вычисления, предполагают более точное отображение знаний экспертов.

Следует заметить, что выбор параметра λ в случае кусочно-линейной функции из примера или параметров других функций принадлежности является фундаментом функционирования системы. Очевидно, что самостоятельное определение параметров окажется затруднительным для пользователей. В качестве способов определения типов и параметров функций принадлежности можно предложить следующие: построение дерева диалога с пользователем, проверка результатов установки параметров на обучающей выборке транзакций, спектральный анализ распределения значений функции.

От того, будет ли выбран в качестве алгоритма решения задачи один из стандартных алгоритмов, реализована их композиция или разработан качественно новый алгоритм, зависит эффективность работы информационной системы. Опыт разработки прототипа ИС обнаружения мошенничества показал, что преобразование и синтез алгоритмов нечеткого вывода в алгоритм, адаптированный к особенностям задачи, позволяет оптимизировать скорость и эффективность ее решения.

Итак, имея некоторый набор лингвистических переменных и аппарат для работы с ними, эксперт может составлять систему предикатных правил, а ИС на их основе — находить значение функции принадлежности F_A для любой транзакции t , то есть решать поставленную задачу.

Реализация описанной математической модели в информационной системе предполагает наличие двух хранилищ данных, выполняющих разные функции:

- база данных, в которой хранятся транзакции со своими атрибутами, а также другая формализованная информация. С точки зрения отображения математической модели база данных будет представлять множество операций T , набор признаков операций — H и характеристическую функцию $Z(H)$;

- база знаний, являющаяся проекцией знаний экспертов на информационную систему, содержащая в себе нечеткие правила вывода решения, лингвистические и нечеткие переменные, определения термов, семантики и синтаксиса.

Пусть имеется структура базы данных, содержащая таблицу с транзакциями и их атрибутами, а также другие вспомогательные таблицы. И пусть построена база знаний, которая содержит определения лингвистических и нечетких переменных, а также нечеткие правила вывода решения, являющиеся отображением на ИС математической модели обнаружения мошенничества на основе нечеткой логики. Система получает данные о транзакциях и знания экспертов, на выходе системы для каждой новой транзакции должен быть ответ на вопрос «является ли транзакция мошеннической».

Система должна, во-первых, обеспечивать ввод и вывод информации, во-вторых, распределять и направлять потоки данных между ее составными частями, и, в-третьих, реализовывать способы вывода решения, заложенные в базе знаний, на транзакциях из базы данных.

Рассмотрим потоки информации в ИС, принимая во внимание не только основную, но и дополнительные, ранее поставленные нами задачи.

На рис. 2 можно увидеть потоки данных и основные этапы их обработки, хранения и вывода.

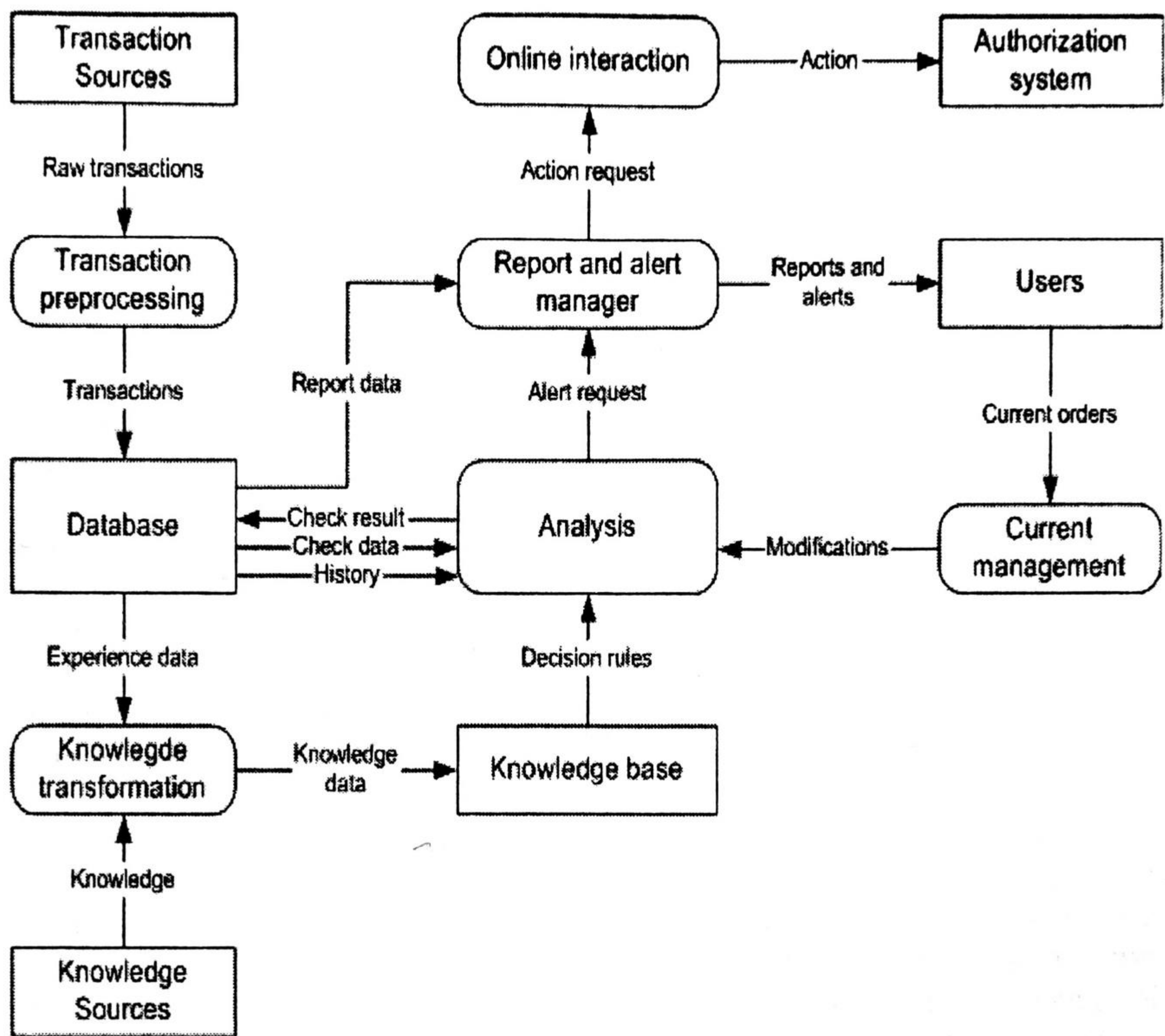


Рис. 2. Диаграмма потоков данных системы

Хранилища данных на рисунке изображены в виде двух горизонтальных линий и вертикальной линии слева от них, интерфейсы — в виде прямоугольников, а процессы, обрабатывающие данные, — в виде прямоугольников со скругленными углами.

Данные в систему попадают из трех источников:

- Transaction sources — источники транзакций, платежная или авторизационная система;
- Knowledge sources — источники знаний, эксперты;
- Users — пользователи системы, риск-менеджеры.

Полученные данные и знания после обработки попадают в одно из двух хранилищ: Knowledge base (база знаний) или Database (база данных), или, как управляющие импульсы (Current Orders), отправляются на обработку, минуя хранилища данных.

Центральным звеном системы, решающим основную задачу, является модуль анализа (Analysis), в который стекаются данные из всех источников. После получения ответа на вопрос «является ли транзакция мошеннической», данные о нем могут отправляться пользователям, в базу данных и в авторизационную систему. База данных осуществляет хранение результатов проверки для последующего использования, информация пользователям обеспечивает обратную связь и позволяет риск-менеджерам предпринять какие-либо действия в отношении мошенников и пострадавших от мошенничества. Информация авторизационной системе поступает в виде запросов на те или иные действия, предотвращающие дальнейшее использование карты в преступных целях.

Другой важный элемент системы — модуль трансформации знаний (Knowledge transformation) получает на входе знания, вводимые экспертами в

систему и данные об уже проверенных транзакциях для накопления опыта и корректировки знаний. Эксперты имеют возможность вводить новые знания, а также изменять и исправлять уже существующие знания на основании собственных суждений или результатов работы системы, получаемых ими от риск-менеджеров (в случае совпадения эксперта и риск-менеджера — на основе собственного опыта).

Сущность модуля отчетов и оповещений (report and alert manager) достаточно проста: функции этого модуля — создавать отчеты для просмотра риск-менеджерами и своевременно предупреждать о мошенничествах заданный круг лиц и систем (риск-менеджеры, клиенты, авторизационная система). Однако при всей простоте модуля способы и критерии оповещения и методы могут быть так разнообразны, что его разработка потребует значительных трудовых затрат. Вместе с тем без этого модуля работа системы будет замыкаться внутри себя. Как можно заметить, все исходящие из системы потоки данных имеют источником именно этот модуль, поэтому его важность для системы нельзя недооценивать.

Построенная система должна отвечать следующим требованиям:

1. Быстрые алгоритмы вывода решения. Объем поступающих на обработку транзакций может быть большим и решение нужно успевать получать в среднем быстрее поступления следующей транзакции на проверку;

2. Надежность и отказоустойчивость. Отсутствие ошибок в решении — одно из важнейших требований к системе, каждая ошибка которой может повлечь финансовые потери. Система оперирует знаниями, ее алгоритмы не могут быть полностью понятны пользователям, поэтому любая вычислительная или алгоритмическая ошибка может породить недоверие пользователей к системе;

3. Гибкость и мобильность. Приемы и способы мошенников развиваются со временем, меняются также платежные системы, пластиковые карты и другие внешние условия. Система должна быть способна реагировать на такие изменения без внесения исправлений в ее исходный код;

4. Прозрачность и понятность. В любой момент работы системы должно быть возможным дать ответ на вопросы «как?» и «почему?», касающиеся алгоритма вывода решения и результата работы. Поэтому алгоритмы, которые используются в системе и могут наблюдаться явным образом, должны быть просты. К тому же простота — одно из требований эффективности, и, следовательно, скорости.

У ИС обнаружения мошенничества имеются две основные группы пользователей: эксперты, осуществляющие ввод и корректировку знаний в систему, и риск-менеджеры, которые обеспечивают ежедневную работу, осуществляют управление и получают результаты работы системы.

Информационная система обнаружения мошенничества, построенная с использованием описанной модели, отвечает требованиям, предъявляемым к таким системам. Гибкость, мобильность и простота использования обеспечиваются лингвистической, естественной формой задания знаний. Методы нечеткой логики показали свою эффективность на примере многих экспертных систем, а прозрачность вывода решений основана на интуитивном понимании человеком принципа работы правил вида «если — то».

Нечеткая логика представляется перспективным математическим методом решения задачи обнаружения мошенничества.

На основе описанной модели реализована информационная система обнаружения мошенничества, которая в настоящее время проходит этап накопления знаний и доработки базовых алгоритмов.

СПИСОК ЛИТЕРАТУРЫ

1. Гаврилов В. Махинации с кредитными карточками // <http://www.bankir.ru/analytics/cards/4/269>.
2. Круглов В. В., Дли М. И., Голунов Р. Ю. Нечеткая логика и искусственные нейронные сети: Учеб. пособие. М.: Изд-во физико-математической литературы, 2001. 224 с.
3. Лукацкий А. В. Системы обнаружения мошенничества // <http://www.infosec.ru/press/pub/p40.htm>.
4. Моор А. П. Нечеткая модель информационной системы обнаружения финансового мошенничества с использованием пластиковых карт // Интеллектуальные технологии в образовании, экономике и управлении: Сб. ст. II Межд. конф. Воронеж: Изд-во им. Е. А. Болховитинова, 2005. С. 404-406.
5. Усков А. А., Кузьмин А. В. Интеллектуальные технологии управления. Искусственные нейронные сети и нечеткая логика. М.: Горячая Линия-Телеком, 2004. 143 с.

Василий Александрович БАРИНОВ —
доцент кафедры математического моделирования,
кандидат физико-математических наук

Светлана Евгеньевна ХОЛОДОВА —
докторант факультета прикладной
математики - процессов управления СПбГУ,
кандидат физико-математических наук

УДК 532.591

ВОЛНЫ НА ТЕЧЕНИИ БАРОКЛИННОЙ ЖИДКОСТИ НАД ГОРИЗОНТАЛЬНЫМ И НЕРОВНЫМ ДНОМ

АННОТАЦИЯ. В статье рассматривается движение сжимаемой жидкости над горизонтальным и неровным дном. При решении задачи использован метод разделения переменных Фурье. Указана возможность применения метода Галеркина. Проведено сравнение с постановкой задачи о движении баротропной жидкости.

Motion of compressible liquid over horizontal and rough bottom is considered in article. The Fourier method of separation of variables was used. The possibility to use the Galerkin method is pointed. The comparison is conducted with statement of the problem about motion barotropic liquid.

Рассмотрим слой сжимаемой жидкости, ограниченный снизу твердым непроницаемым дном $y = -H(x, t)$, сверху — свободной поверхностью $y = \zeta(x, t)$. Здесь система осей xu прямоугольная, ось x совпадает с невозмущенной свободной поверхностью, ось y направлена вертикально вверх. Жидкость будем считать бароклиной идеальной тяжелой. Уравнения движения жидкости при использовании эйлеровых переменных имеют вид

$$\begin{aligned} \frac{\partial v_x}{\partial t} + v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y} &= -\frac{1}{\rho} \frac{\partial p}{\partial x}, \quad \frac{\partial v_y}{\partial t} + v_x \frac{\partial v_y}{\partial x} + v_y \frac{\partial v_y}{\partial y} = -\frac{1}{\rho} \frac{\partial p}{\partial y} - g, \\ \frac{\partial \rho}{\partial t} + \frac{\partial \rho v_x}{\partial x} + \frac{\partial \rho v_y}{\partial y} &= 0, \quad \frac{\partial}{\partial t} \frac{p}{\rho^\kappa} + v_x \frac{\partial}{\partial x} \frac{p}{\rho^\kappa} + v_y \frac{\partial}{\partial y} \frac{p}{\rho^\kappa} = 0, \end{aligned} \quad (1)$$

где $v_x = v_x(x, y, t)$, $v_y = v_y(x, y, t)$ — компоненты скорости, $\rho = \rho(x, y, t)$ — плотность, $p = p(x, y, t)$ — давление, g — величина ускорения силы тяжести.