

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
Государственное образовательное учреждение высшего профессионального об-
разования
«Санкт-Петербургский государственный электротехнический университет
“ЛЭТИ” имени В.И. Ульянова (Ленина)» (СПбГЭТУ)

УДК 519.6+681.51

А.Ю. Дорогов

МАТЕМАТИЧЕСКИЕ ОСНОВЫ МЕТОДОВ ОПТИМАЛЬНОЙ ЧАСТИЧНОЙ
БАЛАНСИРОВКИ ЗНАКОВЫХ ГРАФОВ

Деп. в ВИНТИ 23.07.2007, №760-В2007, 33с

Санкт-Петербург
2007

1. Введение

Сигнальный граф, в котором все ребра имеют вес $+1$ или -1 называют знаковым. Граф подобного вида обычно используется как модель поведения исследуемого объекта или процесса. Вершинам графа сопоставляются переменные состояния объекта, а знак ребра представляет собой качественную оценку причинной связи между смежными вершинами. Если знак положительный, то возрастание значения данной переменной приводит к возрастанию значения смежной, и наоборот уменьшение значения данной переменной влечет за собой уменьшение значения смежной переменной. Все связи графа считаются знаково-симметричными в обоих направлениях для внутренних переменных и однонаправленными для внешних воздействий.

Одной из первой и до сих пор наиболее привлекательной областью приложения знаковых графов является социология где знаковые графы используются для моделирования поведения людей в малых группах (Cartwright и Harary (1956) [1], а также Harary (1954) [2]) Каждому индивидууму группы сопоставляется вершина знакового графа, а знак ребра определяет дружественное или недружественное отношении индивидуума с ближайшим соседом. Было показано, что баланс интересов индивидуумов в группе соответствует знаковой четности всех циклов графа. Цикл называется четным, если он имеет четное число ребер с отрицательным знаком (напомним, что число ноль также считается четным). В этом ключе задача поиска оптимального поведения в группе может быть сформулирована как достижение знакового баланса в графе.

В системах поддержки решений обычно используется так называемая когнитивная карта, которая выражает известные субъекту основные закономерности наблюдаемой ситуации в виде ориентированного знакового графа. Вершины графа – это факторы (признаки, характеристики ситуации), а дуги – причинно-следственные связи между факторами [3,11]. Задача поиска оптимального решения может быть сформулирована как задача достижения знакового баланса в когнитивной карте. В настоящее время теория знакового баланса используется в экономических моделях [4], мультиагентных системах [5], системах адаптивного поведения, телекоммуникационных системах [6] и многих других прикладных областях. Перспективным направлением является приложения методов знакового баланса для предварительного анализа данных в нейротехнологии, распознавании образов и кластеризации данных, извлечения и представления знаний.

Характеризация знаковых графов впервые была дана структурной теоремой Harary (1954). Позже Maybee [7], Harary [8], Kabell [9], и Hansen разработали алгоритмы оценки возможности балансировки знаковых графов. Не все системы являются полностью балансируемыми, важным направлением теории балансировки является оценка степени баланса (Taylor 1970 и Roberts 1976 [10,11]). Вопросы частичной балансировки знаковых графов для социальных систем исследовались в работах Doreian, P., Mrvar, A., 1996 [12], Doreian, P. [13].

Процедура согласования знакового графа связана с переменной знаков его ребер. Возможны три основных подхода к решению проблемы достижения знакового баланса:

1. Нахождение минимального набора ребер графа, изменение знаков которых обеспечит согласование.
2. Определение ограниченного множества наборов ребер, каждый из которых обеспечивает согласование графа.
3. Применение переборных процедур для перечисления всех возможных вариантов согласования.

В первом случае речь идет, как правило, о построении единственного решения, проблема смысловой мотивации такого решения не ставится. Второй подход может дать множественные решения. Проблема состоит в том, чтобы среди них отобрать только те варианты, которые адекватны реальному поведению моделируемой системы. Третий путь приводит к NP-полным задачам и упирается в ограничение размерности.

Не смотря на множество работ, математическое решение задачи балансировки знаковых графов больших размеров до сих пор не получено. Принципиально новый подход к решению данной проблемы представлен в работах Б.Ф. Фомина и Т.Л. Качановой (1978) [14] которые использовали модель знакового графа для реконструктивного анализа сложных систем [15,16]. Основная идея новой парадигмы заключается в оптимальном использовании внутренних симметрий знакового графа для балансировки его ограниченных областей (локальностей). Предложенный индуктивный подход позволяет максимально наращивать сбалансированные локальности, исключая прямой перебор вариантов. Отличительной особенностью является определение доминирующих вариантов согласования и наличие технических индикаторов полноты данных. Системное знание раскрываются через анализ особых структур обладающих различными видами симметрий и исследование их взаимодействий. Общее решение было получено на качественном уровне. В настоящей работе представлены основные положения новой парадигмы балансировки знаковых графов и сделана попытка построить ее математическое обоснование.

2. Треугольник противоречий

Нарушенный баланса в знаковом графе отражает наличие противоречия в данных, на основе которых этот граф был построен. Минимальный знаковый граф, позво-

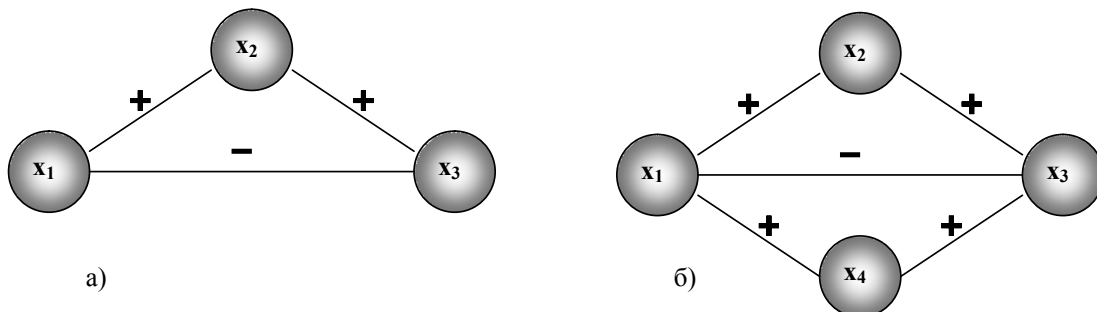


Рис. 1. а) Треугольник противоречий. б) Декомпозиция согласованного контура в два несогласованных треугольника

ляющий выразить противоречие, представляет собой замкнутый контур, состоящий из

трех вершин (рис. 1а)). В данном графе при передаче по контуру $x_1 - x_2 - x_3 - x_1$ положительное приращение переменной x_1 трансформируется в отрицательное приращение этой же переменной, т.е. контурный коэффициент передачи имеет значение (-1), что физически трактуется как противоречие. В этом случае контур называют несбалансированным или нечетным.

Особая роль фигуры треугольника заключается в том, что он может служить базой индукции в задаче балансировки сложных знаковых графов. Нетрудно показать, что многовершинный контур, образованный сторонами смежных треугольников, будет сбалансирован, если сбалансированы все образующие его треугольники. Обратное, вообще говоря, не верно, альтернативный пример показан на рис. 1б).

3. Матричная интерпретация

Каждое ребро знакового графа отражает взаимную зависимость двух системных переменных $x_i \leftrightarrow x_j$. Эта связь является двусторонней для внутренних переменных и односторонней для внешних воздействий. Рис. 2 демонстрирует знаковый граф, в котором явным образом выражена направленность и симметричность связей.

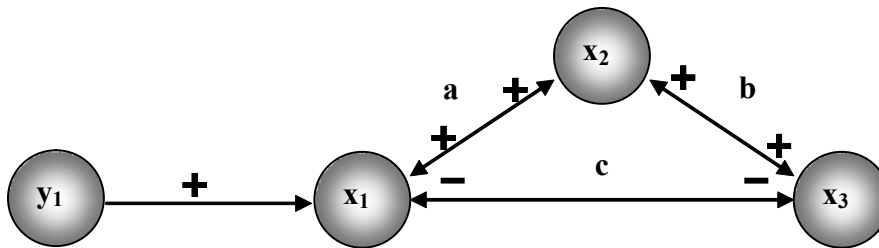


Рис. 2. Полная модель знакового треугольника

Вершины x_1, x_2, x_3 соответствуют внутренним переменным, а вершина y_1 - внешнему воздействию. Эквивалентным описанием знакового графа является матрица связей (матрица смежности). Для графа показанного на рис. 2 эта матрица имеет вид:

$$C = \begin{array}{c|ccc} & x_1 & x_2 & x_3 \\ \hline x_1 & 0 & 1 & -1 \\ x_2 & 1 & 0 & 1 \\ x_3 & -1 & 1 & 0 \end{array}$$

Элементы матрицы связей равны коэффициентам передачи между системными переменными. Матрица симметрична по отношению к главной диагонали, поскольку симметричны все внутренние связи. По графу можно построить также систему алгебраических уравнений, определяющих реакцию системы на внешние воздействия:

$$\begin{cases} x_1 - x_2 + x_3 = y_1 \\ -x_1 + x_2 - x_3 = 0 \\ x_1 - x_2 + x_3 = 0. \end{cases} \quad (1)$$

Матрицу A , составленную из коэффициентов алгебраической системы, назовем системной матрицей. Для нашего примера она имеет вид:

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}.$$

Системная матрица связана с матрицей связей соотношением: $A = I - C$, где I - единичная матрица. Системные переменные и внешние воздействия представим в виде векторов:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ 0 \\ 0 \end{bmatrix}.$$

Тогда реакция системы на внешние воздействия может быть выражена в виде матричного уравнения:

$$AX = Y. \quad (2)$$

Уравнение имеет единственное решение относительно вектора X , если определитель Δ_A системной матрицы A отличен от нуля. Нетрудно проверить, что для треугольника противоречий значение определителя равно нулю. При нулевом определителе существует множество решений, которые образуют плоскость некоторой размерности в пространстве состояний системы. Это свидетельствует о неопределенности поведения системы и ее реакции на внешние воздействия. Говорят, что алгебраическая система плохо обусловлена [17], если определитель ее системной матрицы близок или равен нулю. Таким образом, наличие противоречий в треугольнике эквивалентно плохой обусловленности соответствующей алгебраической системы. В целом это положение справедливо и для более сложных знаковых графов с противоречиями, хотя определитель их системной матрицы может быть отличен от нуля. Абсолютное значение определителя служит мерой обусловленности системы. В плохо обусловленной системе малые изменения внешних воздействий или элементов системной матрицы могут привести к значительным изменениям внутренних переменных. Поэтому при оптимальной балансировке графа, необходимо выбрать вариант с наилучшей обусловленностью системной матрицы.

Достигнуть баланса в треугольнике можно изменив знак связей. Четыре возможных варианта согласования треугольника показаны на рис. 3. Прямоугольниками выделены знаки изменяемых связей. Все согласованные треугольники имеют значение определителя $\Delta_A = -4$, поэтому нельзя отдать предпочтение, какому либо из вариантов по степени обусловленности. Однако в последнем варианте число перемен знаков больше чем в первых трех. Поэтому этот вариант можно считать менее предпочтительным, поскольку он приносит больше изменений в исходный граф связей.

Основываясь на данных рассуждениях задачу оптимальной балансировки знакового графа можно сформулировать следующим образом: необходимо найти минимальное подмножество ребер, изменение знаков которых приводит к наилучшей обусловленности системной матрицы.

Пример треугольника показывает, что существует конечное множество эквивалентных вариантов согласования знаковых графов. Число вариантов балансировки растет экспоненциально при увеличении количества вершин графа. Выполнить полный перебор и оценку всех вариантов для сколько-нибудь большой системы не представляется возможным. Необходимы иные принципы построения вариантов согласования, детали новой парадигмы будут представлены в последующих разделах.

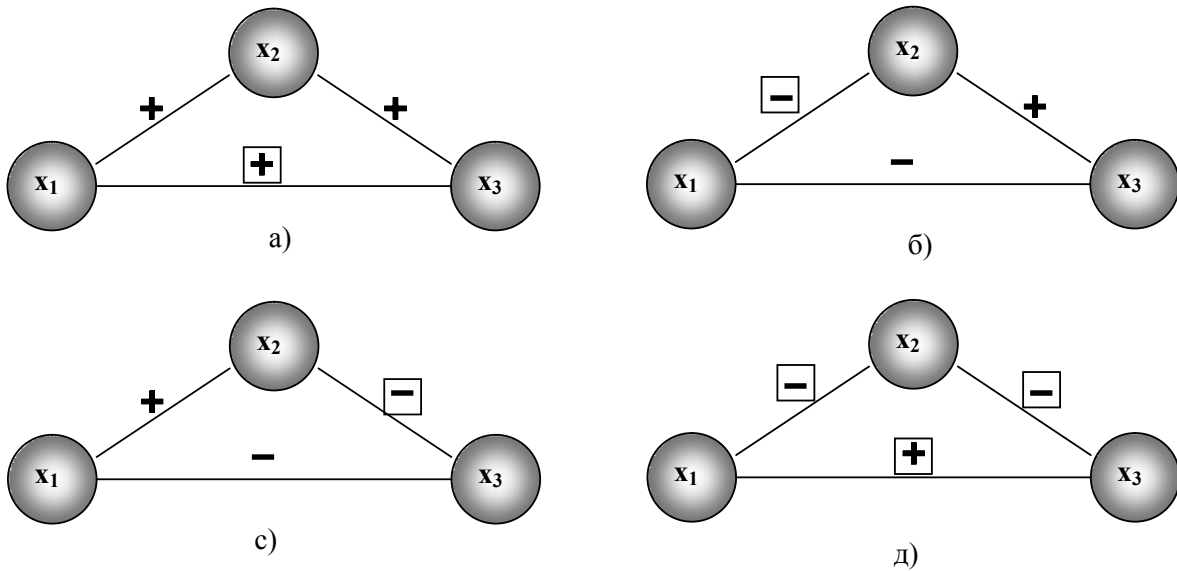


Рис. 3. Варианты согласования треугольника противоречий. Прямоугольниками помечены изменяемые связи

Построение системной матрицы и вычисление определителя при высокой размерности задача достаточно трудоемкая и требует большого объема вычислений. Кроме того, использование численных методов исключает возможность построения аналитических выражений для оценки степени влияния отдельных контуров графа на значение определителя системной матрицы. В следующем разделе будут рассмотрены средства, упрощающие вычисления значения определителя и главное позволяющие получить аналитическую форму выражения баланса графа, через баланс треугольников.

3. Топологические формулы

Вычисление определителя системной матрицы непосредственно по графу связей можно эффективно выполнить, используя топологические методы расчета [18]. Знаковый граф принципиально ничем не отличается от обобщенного сигнального графа используемого для расчета электрических цепей. Представление знакового графа, показанное на рис. 2 совпадает с формой обобщенного сигнального графа, за исключением весов вершин. Весами вершин обобщенного сигнального графа являются соответствующие элементы главной диагонали системной матрицы A . Для знакового графа эти значения равны 1 для всех вершин и, поэтому их можно условно не отображать в поле графа. Топологическая формула для вычисления определителя обобщенного сигнального графа имеет вид:

$$\Delta = \delta + \sum L^{(j)}\delta^{(j)} + \sum L^{(j)}L^{(k)}\delta^{(j,k)} + \dots + \sum L^{(j)}L^{(k)} \dots L^{(s)},$$

где $\delta = \prod a_{ii}$ - произведение весов всех вершин графа, $\delta^{(j)}$ - произведение весов всех вершин графа, кроме вершин, входящих в j -й контур с весом $L^{(j)}$, $\delta^{(j,k)}$ - произведение весов всех вершин графа, кроме вершин, входящих в j -й и k -й не касающиеся контуры с весом $L^{(j)}$ и $L^{(k)}$ соответственно и т. д. Последний член формулы включает в себя максимальное количество не касающихся контуров, которые очевидно покрывают все вершины графа, поэтому значение $\delta^{(j,k,\dots,s)}$ равно единице. Следует обратить внимание на то обстоятельство, что вес контура $L^{(j)}$ равен коэффициенту контурной передачи, взятому с обратным знаком, это позволяет исключить чередование знаков «плюс» и «минус» в данной формуле.

Для упрощения расчета часто используют формулы разложения сигнальных графов по ребру или по вершине. Разложение графа заключается в том, что ребро или вершина удаляется из графа, определитель рассчитывается для упрощенного графа, а затем по топологическим формулам реализуется переход к значению определителя исходного графа. Приведем только топологическую формулу разложения по вершине, которая оказывается наиболее применимой для знаковых графов:

$$\Delta = a_{ii}\Delta_i + \sum L_{ii}^{(s)}\Delta^{(s)},$$

здесь Δ_i - определитель графа G_i образуемого из исходного графа G удалением вершины x_i , $L_{ii}^{(s)}$ - вес s -го контура проходящего через вершину x_i , $\Delta^{(s)}$ - определитель части графа, не касающийся s -го контура. Поскольку для знакового графа все вершины имеют единичный вес, топологические формулы упрощаются и примут вид:

$$\Delta = 1 + \sum L^{(j)} + \sum L^{(j)}L^{(k)} + \dots + \sum L^{(j)}L^{(k)} \dots L^{(s)}, \quad (3)$$

$$\Delta = \Delta_i + \sum L_{ii}^{(s)}\Delta^{(s)}. \quad (4)$$

Реакцию переменной состояния x_k на внешнее воздействие y_l можно определить с помощью топологической формулы:

$$W_{kl} = \frac{x_k}{y_l} = \frac{\sum_s P_{kl}^{(s)}\Delta^{(s)}}{\Delta}, \quad (5)$$

где Δ - определитель системной матрицы, $P_{kl}^{(s)}$ - вес s -го пути в вершину x_k из вершины внешнего воздействия y_l , $\Delta^{(s)}$ - определитель части графа не касающийся s -го пути. Вес пути вычисляется как произведение весов всех дуг, принадлежащих данному пути.

Обобщенный сигнальный граф треугольника, показанный на рис. 2 включает в себя пять контуров: два вложенных друг в друга контура из трех вершин с противоположным направлением обхода и три контура из двух смежных вершин треугольника. Коэффициент контурной передачи для треугольника равен $\alpha = abc$. Если треугольник не имеет противоречий, то $\alpha = 1$, в противном случае $\alpha = -1$. Все двухвершинные кон-

туры имеют коэффициент передачи равный +1. Граф не содержит не касающихся контуров. Используя топологическую формулу (3) немедленно получим:

$$\Delta = 1 - 2\alpha - 3 = -2(1 + \alpha).$$

Отсюда в частности следует, что для треугольника противоречий $\Delta = 0$. Рассмотрим некоторые типичные примеры топологического расчета, которые понадобятся далее для анализа структур.

- 1) Диполь – это знаковый граф из двух вершин (см. рис. 4а)).

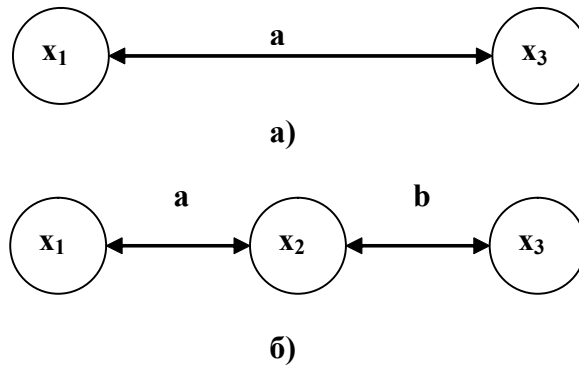


Рис. 4. Диполь – а) и ленточный граф – б)

Применяя топологическую формулу (3) для знакового диполя будем иметь: $\Delta = 1 - a^2 = 0$. Таким образом, определитель системной матрицы диполя равен нулю в не зависимости от знака связи.

- 2) Ленточный граф (см. рис. 4б)). Разложим граф по вершине x_1 , в результате получим $\Delta = 1 \cdot 0 - a^2 = -1$; Нетрудно видеть, что такое же значение определителя будет иметь ленточный граф с произвольным количеством вершин.
- 3) Замкнутый ленточный граф (см. рис. 5). Разложим граф по вершине x_4 . Новый граф будет разомкнутым ленточным графом, состоящих из трех вершин, его определитель (как было показано выше) равен -1. Определитель полного графа может быть представлен выражением:

$$\Delta = -1 - a^2 - c^2 - 2\lambda = -3 - 2\lambda,$$

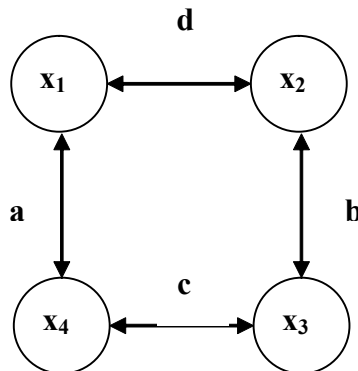


Рис. 5. Замкнутый ленточный граф

где $\lambda = adbc$ - контурная передача вдоль замкнутой ленты. Если контур согласован, то $\Delta = -5$. Значение определителя не зависит от числа вершин в замкнутой ленте.

4) Несвязанные графы. Несвязанным графом [18] называется топологическая структура, содержащая m несвязанных между собой дугами подграфов G_1, G_2, \dots, G_m , каждый из которых может иметь произвольную структуру. Несвязанному графу соответствует квазидиогональная блочная системная матрица. Как известно [19] определитель квазидиогональной блочной матрицы с квадратными блоками равен произведению определителей матриц блоков, т.е.

$$\Delta = \prod_{i=1}^m \Delta_i.$$

Используя результаты рассмотренных примеров, определим реакции системных переменных знакового треугольника (см. рис. 2) на внешнее воздействие y_1 . По топологической формуле (5) получим:

$$W_{11} = \frac{x_1}{y_1} = 0, \quad W_{21} = \frac{x_2}{y_1} = \frac{a}{\Delta}, \quad W_{31} = \frac{x_3}{y_1} = \frac{c}{\Delta}.$$

4. Разделяющие структуры

Определяющей позицией новой парадигмы балансировки знаковых графов [14] является концепция последовательной индукции от минимальной цикловой фигуры – треугольника противоречий. Совокупность всех треугольников противоречий служит базой индукции при поиске оптимальных вариантов согласования. Рассматриваемые отдельно от графа связей, все эти треугольники совершенно равноправны. В формальном исследовании допустимость смены знака на любом ребре не обсуждается и считается всегда возможной. Каждый треугольник противоречий обладает симметрией по-

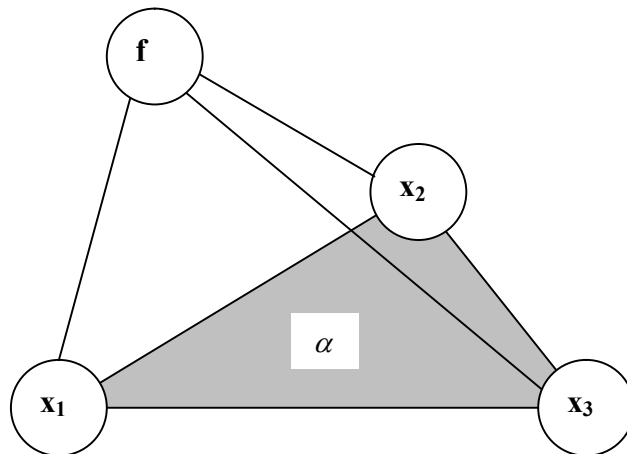


Рис. 6. Многовершинное окружение треугольника противоречий

воротам, что не дает возможность точно установить несогласованные пары связей и все варианты согласования являются допустимыми. Была высказана идея, что нарушить симметрию поворота треугольника можно за счет информации имеющейся в его бли-

жайшем окружении. Окружение для треугольника противоречий должно быть таким,

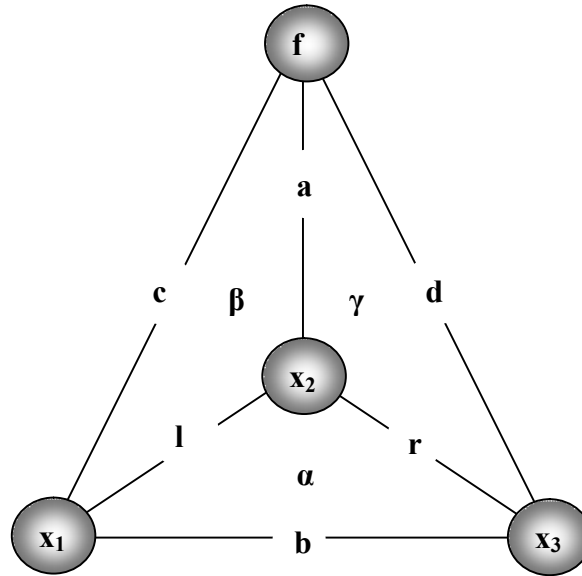


Рис. 7. Многовершинная структура. Знаки ребер обозначены латинскими символами, контурные передачи – греческими символами

чтобы подграф каждой вершины окружения включал в себя треугольник как единое целое. Это условие приводит к многовершинной структуре – тетраэдру (рис. 6) в которой сам треугольник α служит основанием тетраэдра, а вершина окружения f связана симметричными связями со всеми вершинами треугольника. Каждый тетраэдр многовершинной структуры имеет три боковых грани и к ним также применим критерий знакового баланса. Отправной точкой для последующих рассмотрений является следующая теорема.

Теорема «О разделении граней тетраэдра». Если в основании тетраэдра размещен несбалансированный треугольник, то только две из четырех граней тетраэдра могут быть согласованы. Невозможными являются случаи, когда согласованы три грани или только одна какая-то грань.

Доказательство. Рассмотрим тетраэдр в плоском отображении (рис. 7). По условию грань $x_1 - x_2 - x_3$ есть треугольник противоречий, поэтому $\alpha = lrb = -1$. Предположим, что треугольник $x_1 - f - x_3$ согласован, тогда его контурная передача определяется равенством $cdb = 1$. Выразим это равенство, через передачи контуров, используя следующее тождество:

$$cdb = (cal)(adr)(blr).$$

Переменные a, l, r являются знаковыми и добавлены в выражение парами, поэтому они не меняют произведение. В результате получим:

$$cdb = \beta\gamma\alpha = 1.$$

Поскольку $\alpha = -1$, то последнее равенство будет выполняться, когда β и γ будут иметь разные знаки. Таким образом, при данном предположении только две из четырех граней тетраэдра могут быть согласованы.

Предположим теперь, что $cdb = \beta\gamma\alpha = -1$ и $\alpha = -1$. Равенство будет выполняться, только если β и γ имеет одинаковые знаки. Когда эти знаки положительны, то две грани в тетраэдре согласованы, когда знаки отрицательны, то в тетраэдре нет согласованных граней. Случаи, когда согласованы три или одна грань тетраэдра (с треугольником противоречий в основании) существовать не могут. *Теорема доказана.*

Для полноты картины рассматриваются все возможные тетраэдры имеющими в основании заданный треугольник противоречий. С каждой вершиной окружения связаны три грани, по доказанной теореме две из этих граней согласованы. Это может быть любая пара из трех возможных граней. Каждая из граней вершиной окружения ассоциируется с одной из сторон треугольника в основании тетраэдра. По ассоциативному признаку все вершины окружения могут быть разделены на три категории. Каждой категории соответствует ассоциированная пара сторон треугольника основания, с которой вершина окружения образует согласованные треугольники.

Если одна из трех категорий является пустым множеством, то многовершинная

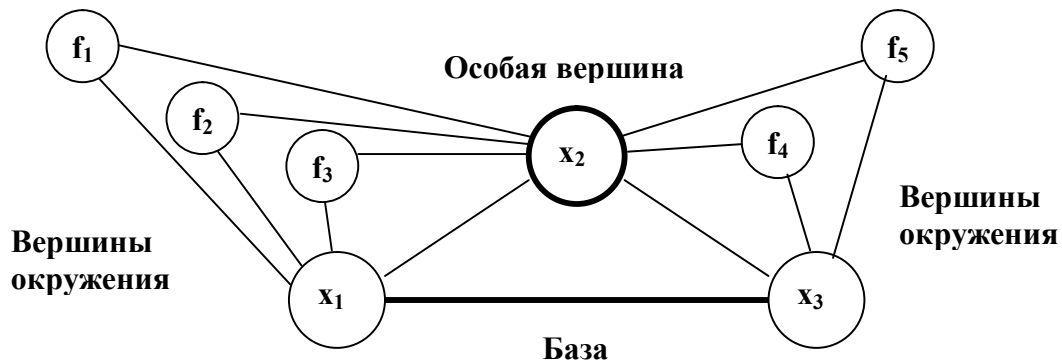


Рис. 8. Морфология многовершинной структуры (для упрощения рисунка часть ребер не показана)

структура приобретает осевую симметрию. При такой симметрии элементы треугольника становятся различимыми, в нем выделяются особое ребро – база и противоположное этому ребру вершина – особая вершина (рис. 8). База основания играет особую роль и отличается по своей значимости от боковых ребер; все треугольники, образованные вершинами окружения и замкнутые на базу не противоречивы, напротив все противоречивые треугольники замкнуты на боковые ребра. В силу этого изменение знаков боковых ребер при согласовании будут мотивированы принципом минимальных изменений, тогда как перемена знака базы никак не обоснована. Таким образом, в многовершинной структуре остаются только два обоснованных варианта согласования, порожаемых изменением знака какого либо одного бокового ребра (рис. 9).

Если из трех категорий вершин окружения две являются пустыми, то многовершинная структура ассиметрична. Для нее существует единственный вариант согласования, связанный с изменением только одной стороны треугольника основания, а именно той с которой вершины окружения образуют противоречивый треугольник. Выбор базы из двух оставшихся сторон не определен, но это не препятствует использованию ассиметричных структур в процедурах балансировки наравне со структурами, обладаю-

щими осевой симметрией. Более того, эти структуры наилучшим образом отвечают принципу минимальных изменений.

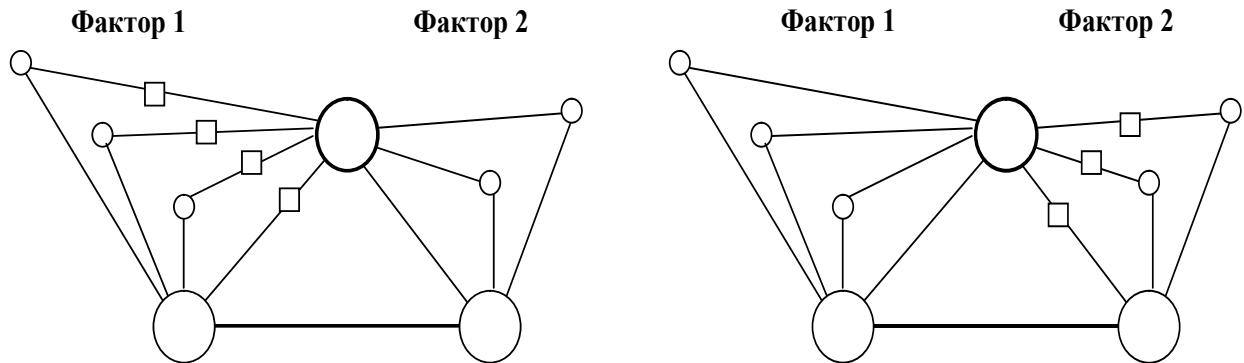


Рис. 9. Варианты согласования многовершинной структуры. Прямоугольниками помечены связи одновременно изменяемые при согласовании

Многовершинные структуры обоих типов состоящие из N тетраэдров окружения удовлетворяют следующим условиям:

1. Количество согласованных треугольников в структуре равно $2N$, тогда как число не согласованных треугольников в ней, включая основание равно $(N + 1)$.
2. Только одно ребро основания (база) определяет с вершинами окружения N согласованных треугольников.
3. Сумма согласованных треугольников, образованных вершинами окружения и двумя другими ребрами основания структуры равна N .

Три перечисленных условия полностью соответствуют теореме «о разделении граней тетраэдра» в предположении, что в каждом тетраэдре многовершинной структуры существуют сбалансированные треугольники. Многовершинные структуры, удовлетворяющие данным трем условиям, были названы *разделяющими* и обозначаются далее идентификатором «SHARE». Разделяющие структуры с осевой симметрией называются также *синглетами* [16].

Если все три вышеупомянутые категории вершин окружения пустыми не являются, то для данной многовершинной структуры осевая симметрия не может быть выделена. Для подобных структур не существует мотивированных вариантов согласования треугольника противоречий; изменение знака любой стороны треугольника основания всегда будет порождать новые не согласованные грани вершин окружения. Данный тип структуры назовем *предсинглетом*, при удалении части показателей в системе данных, эти структуры могут перейти в разряд разделяющих структур.

Осевая симметрия SHARE разделяет вершины окружения на два непересекающихся подмножества – фактора. Каждый из факторов образует согласованные треугольники с одним из боковых ребер основания, а треугольники противоречий с другим боковым ребром. Боковые ребра отличаются вершинами баз, поэтому и факторы также сопоставляются вершинам баз (рис. 10). Более того, вершины баз включаются в факторы, это позволяет рассматривать симметричные и ассиметричные разделяющие

структуры в одном ключе. Допустимое согласование достигается изменением знаков связей вершин одного из факторов с особой вершиной.

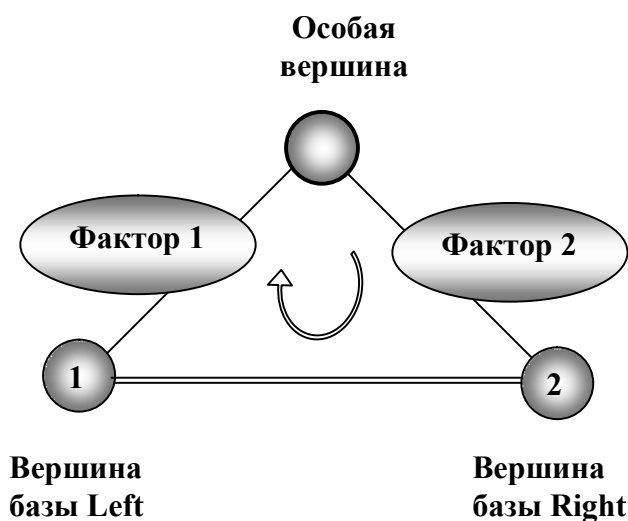


Рис. 10. Ориентированная разделяющая структура. Вершины баз условно выделены из факторов.

Для различения факторов и соответствующих базовых вершин удобно ввести ориентацию синглета, задавая для него направление обхода треугольника. Тогда фактором 1 (или Left-фактором) будем считать фактор, который расположен слева от осо-

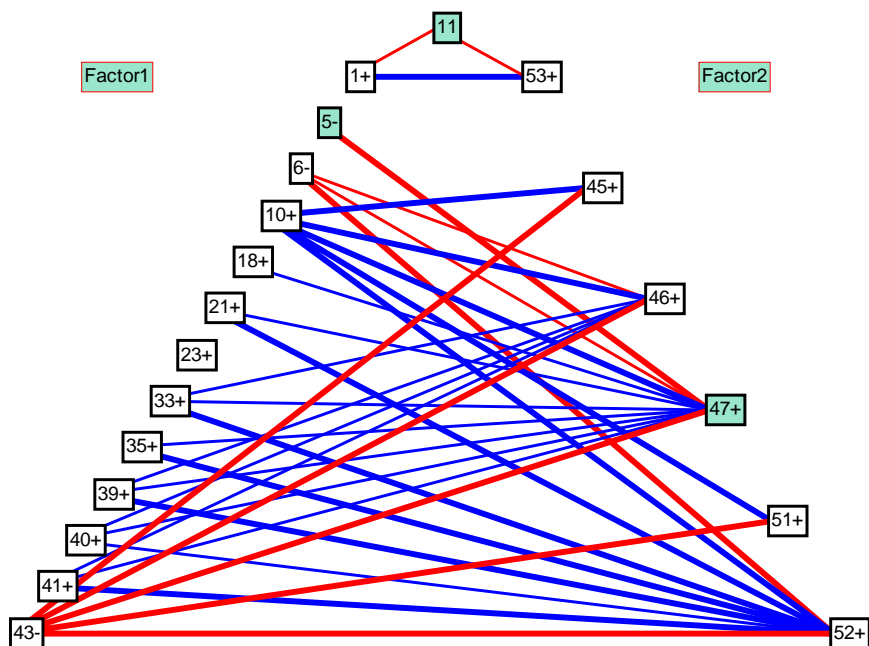


Рис. 11. Портрет разделяющей структуры SHARE в реальных данных. Положительные связи отмечены красным цветом (светлый тон), отрицательные – синим (темный тон). Утолщенными линиями показаны базы. Треугольник противоречий фрагментарно выделен. Особые вершины имеют темную окраску. Связи вершин факторов с особой вершиной не показаны, но их знаки отображаются после номера факторной вершины

бой вершины по направлению обхода треугольника, а фактором 2 (или Right – фактором) будем считать фактор, расположенный справа от особой вершины по направлению обхода. Вершины окружения фактора Left образуют согласованные треугольники с левым ребром треугольника противоречий, а вершины окружения фактора Right с правым ребром. Разделяющая структура, в которой задана ориентация треугольника противоречий, называется поляризованной или ориентированной. Эта же терминология используется и для факторов. Частичный знаковый граф, являющийся носителем разделяющей структуры в дальнейшем называется *SHARE-локальностью*, а треугольник противоречий – *ядром локальности*. В общем случае локальность может быть носителем для нескольких SHARE.

На рис. 11 приведен пример SHARE-структуры в реальной системе данных. (здесь и далее для примеров используется база данных, описанная в работе [20], знаковый граф связей был построен по матрице парных корреляций атрибутов базы данных).

Согласованная многовершинная структура, полученная изменением связей Left-фактора с особой вершиной называется *Left-эталонной моделью*, а изменением связей Right-фактора называется *Right-эталонной моделью*. Каждая из эталонных моделей определяет одно из двух возможных поведений системы в границах SHARE. Применимость Left или Right моделей может быть потенциально ограничена, если одна или несколько изменяемых связей фактора являются базами других SHARE. Декларирован-

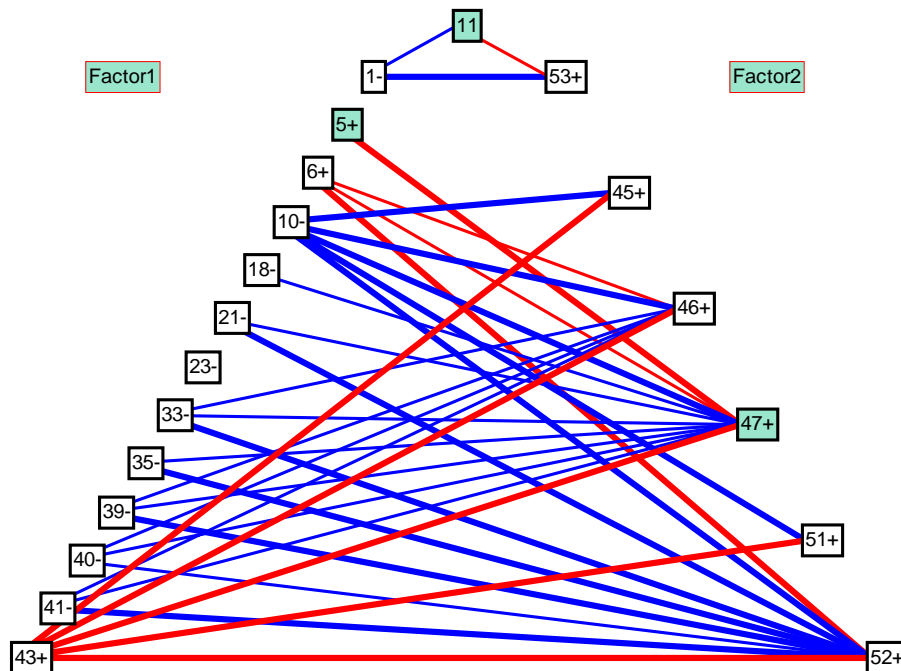


Рис. 12. Эталонная модель SHARE-Left

ный выше принцип минимизации изменений запрещает изменять знаки баз, это ведет к мотивированному ограничению в применимости модели, когда база попадает в число изменяемых связей. Для характеристики ограничений эталонных моделей введем для SHARE вектор маркировки $P = (p_l \ p_r)$, состоящий из двух бинарных элементов. Пер-

вый элемент – p_l , соответствует эталонной модели Left и равен 1, если эта модель не имеет ограничений и нулю в противном случае. Второй элемент – p_r , подобным образом описывает эталонную модель Right. Вектор маркировки не препятствует использованию эталонной модели при независимом применении SHARE, но содержит информацию о потенциальных ограничениях которые могут иметь место при взаимодействии ее с другими локальностями. На рис. 12 показан пример эталонной модели Left. Граф модели полностью сбалансирован.

5. Концепция открытой локальности

Построение разделяющих структур было связано с обособлением вершин окружения и их связей с вершинами ядра. Обратное погружение выделенной структуры в исходный граф, приводит к появлению дополнительных связей, которые могут быть не

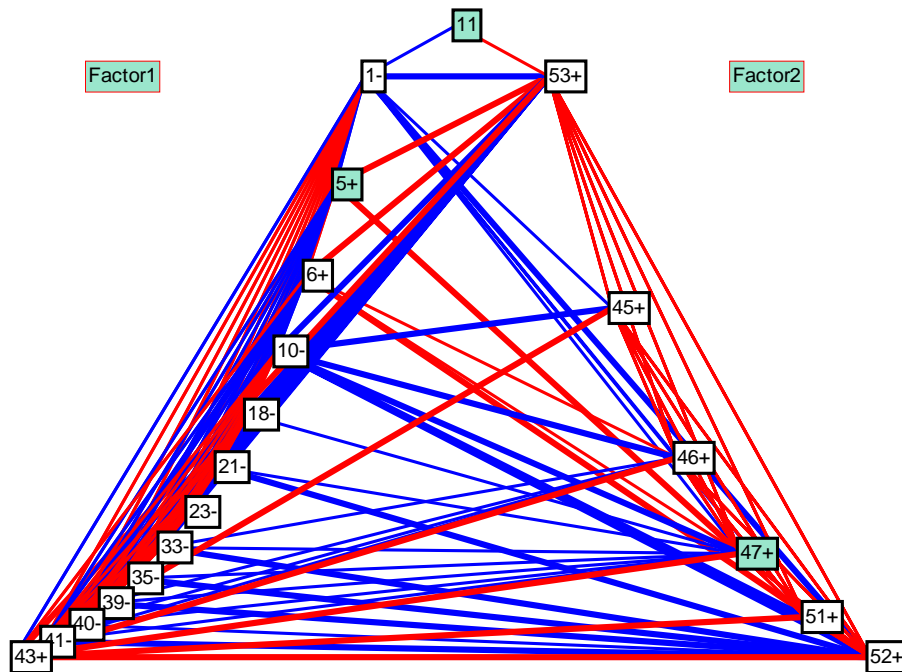


Рис. 13. Открытая локальность SHARE -UNION

охвачены процедурой балансировки по эталонным моделям. Подграф, который наследует структурную модель и дополненный связями исходного графа определяет модель открытой локальности (UNION). На рис. 13 приведен пример открытой локальности SHARE-UNION. Для следующего шага итерации отбираются SHARE, которые сохраняют согласованность в открытой локальности, по крайней мере, для одной эталонной модели.

Для открытой локальности может быть применена дополнительная балансировка несогласованных треугольников, которые не входят ни в одну из SHARE. Возможные варианты согласования определяются прямым перебором альтернатив и оценкой степени обусловленности каждого из вариантов для поиска наилучшего. Поскольку в основном баланс в UNION достигается через одну из эталонных моделей, число несба-

лансированных связей относительно не велико и использование переборных процедур вполне допустимо.

6. Интегрированная SHARE

Локальности SHARE можно рассматривать как «точки кристаллизации», которые позволяют запустить механизм индукции в процедуре согласования знакового графа. Одним из вариантов расширения локальности состоит в выделении на множестве SHARE однородных классов по особой вершине. В каждый класс включаются разделяющие структуры, которые имеют общую особую вершину (рис. 14) и согласуются в собственной UNION. Структура, соответствующая классу SHARE с общей особой

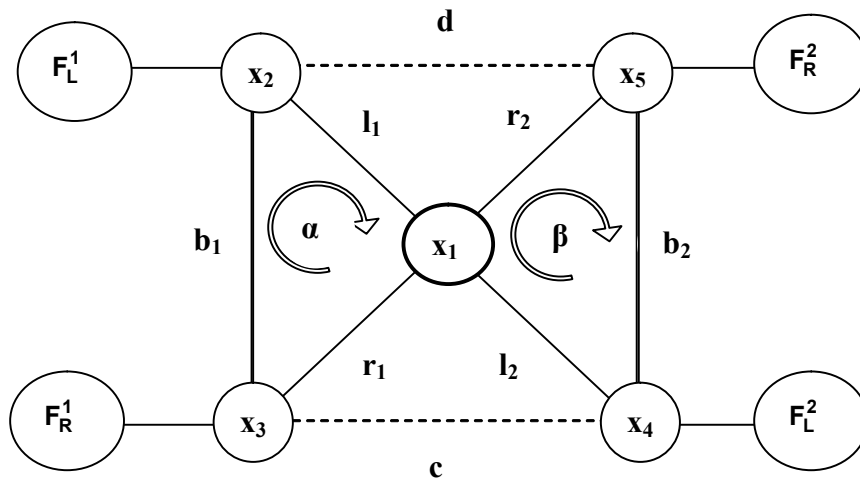


Рис. 14. Интегрированная структура, объединяющая две SHARE с общей особой вершиной. Символы b, r, l, d, c обозначают знаки связей, символы α, β – контурные передачи, символы F – факторы SHARE. Пунктирными линиями показаны возможные дополнительные связи

вершиной называется интегрированной SHARE (ISHARE). Частичный граф, объединяющий все треугольники противоречий в ISHARE, называется её ядром.

Разделяющие структуры SHARE имеют двухфакторную организацию. Целесообразно поддерживать ориентированную двухфакторность и в интегрированной структуре, это в полной мере отвечает декларированному принципу минимальности изменений. При построении ISHARE возникает проблема корректного объединения факторов SHARE в два фактора ISHARE. Задача ориентации имеет однозначное решение только в том случае, если граф ядра является связным и решается в результате анализа пересечений факторов SHARE и связей ядра ISHARE.

Например, для структуры показанной на рис. 14 в предположении, что для факторов SHARE выполнены условия $F_L^1 \cap F_R^2 = \emptyset$, $F_R^1 \cap F_L^2 = \emptyset$. Факторы ISHARE формируются следующим образом $F_L = F_L^1 \cup F_L^2$, $F_R = F_R^1 \cup F_R^2$ (см. Рис. 15). Процедура поляризации ISHARE начинается с поляризации баз. Определяющим является требование, что вновь образуемые полярные базовые множества не должны пересе-

каться между собой. Если после поляризации по множеству баз для некоторых SHARE

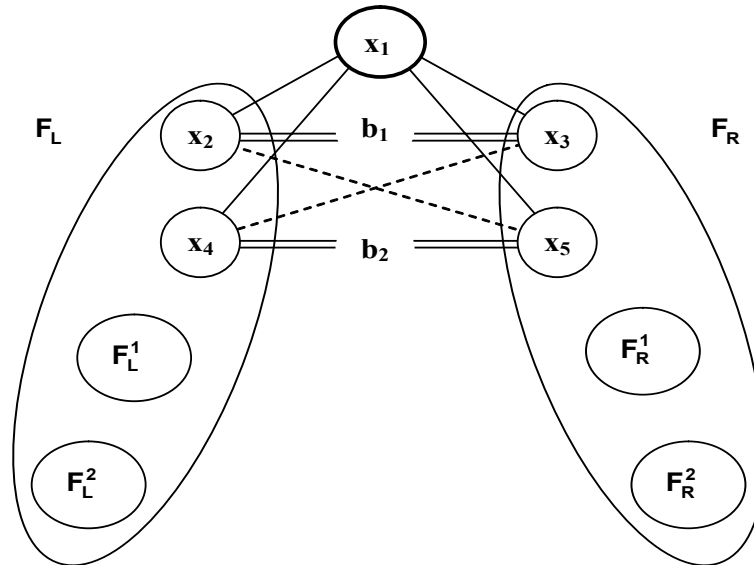


Рис. 15. Поляризация факторов ISHARE

осталось неопределенность, то поляризация выполняется по факторным вершинам, не принадлежащим множеству баз.

Назначение имен (ориентация) для факторов ISHARE (Left или Right) достаточно условно (как впрочем, и для факторов SHARE) однако эта операция необходима как средство корректного построения эталонных моделей. Структура ISHARE формируется последовательным накоплением локальностей с общей особой вершиной, поэтому

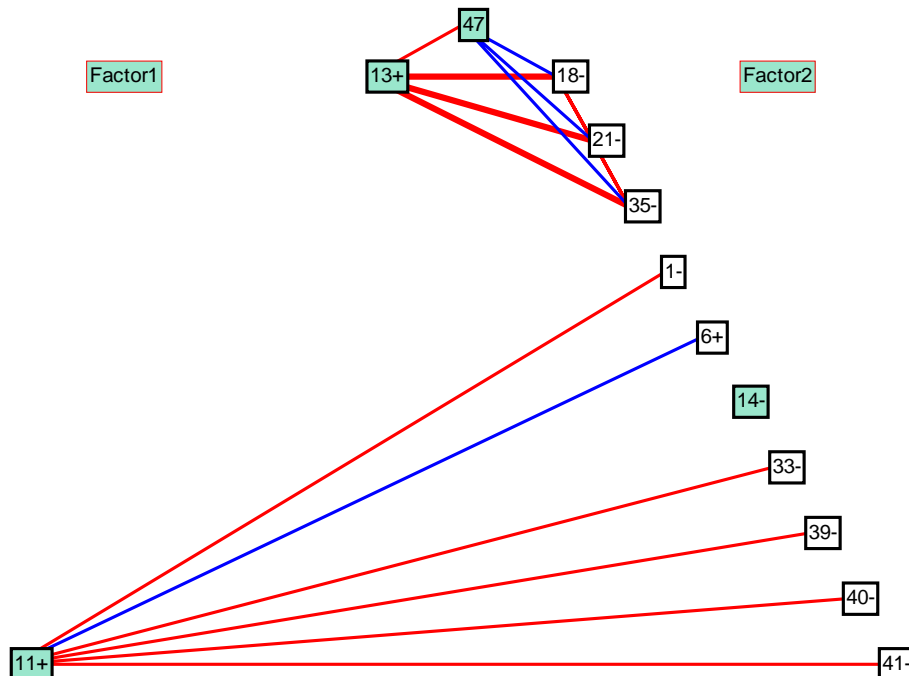


Рис. 16. Интегрированная структура ISHARE, образованная объединением трех структур SHARE с общей особой вершиной.

практически удобно присвоить ISHARE ориентацию первой SHARE в алгоритме накопления.

При объединении в ISHARE факторы образующих локальностей изменяют свою ориентацию, адаптируясь к заданной ориентации факторов ISHARE. Вместе с факторами SHARE ориентацию изменяют и соответствующие вектора маркировки. Правило формирования вектора маркировки для ISHARE определяется выражением:

$$P = \bigcap_i P_i$$

где операция \bigcap обозначает поразрядное логическое произведение предварительно поляризованных векторов маркировки объединяемых SHARE. На рис. 16 приведен пример структуры ISHARE в реальных данных.

Оценим теперь степень обусловленности структуры ISHARE. Рассмотрим знаковый граф представленный на рис. 14, и вычислим определитель его ядра. Связи, показанные пунктиром в первичном анализе, будем полагать отсутствующими. Разложим граф по вершине x_1 и воспользуемся топологической формулой (4) для вычисления определителя. После разложения по вершине граф будет состоять из двух диполей. В п.3 показано, что определитель диполя равен нулю при любых знаках связи. Отсюда следует, что определитель системной матрицы ядра равен нулю, в независимости от того согласованы или не согласованы треугольники SHARE. Этот результат показывает, что эффект от объединения разделяющих структур в ISHARE может быть нулевым, или даже приводить к отрицательным результатам, если вершины ядра не связаны между собой.

Пусть в графе присутствует дополнительное ребро d . Вновь выполним разложение по вершине x_1 , после разложения получим ленточный граф, его определитель равен -1 (точнее d^2 , см. п.3). В целом определитель матрицы ядра можно записать в виде следующего выражения:

$$\Delta = -1 - 2\alpha \cdot 0 - 2\beta \cdot 0 - 2l_1 dr_2 - 2b_1 dr_2 r_1 - 2l_1 db_2 l_2.$$

После элементарных преобразований получим:

$$\Delta = -1 - 2\gamma - 2\alpha\gamma - 2\beta\gamma,$$

где $\gamma = l_1 dr_2$. Определитель будет максимален по модулю и равен (-7) , когда все треугольники в ядре ISHARE согласованы. Согласование в двухфакторной структуре достигается изменением связей вершин одного из факторов с особой вершиной. Если после изменения факторных связей необходимо согласование в треугольнике γ , то оно достигается изменением знака ребра d .

Пусть теперь в графе присутствуют два дополнительных ребра d и c . Вычислим определитель, разложив граф по вершине x_1 . После разложения получим замкнутый ленточный граф, его определитель равен $-3 - \lambda$ (см п.3), где λ - коэффициент контурной передачи. После несложных преобразований определитель полного графа ядра можно записать в виде:

$$\Delta = -3 - 2\lambda - 2\alpha\gamma - 2\beta\eta - 2\alpha\gamma\chi - 2\beta\eta\alpha,$$

где $\eta = l_2 cr_1$. Определитель будет равен (-13), если все треугольники в ядре ISHARE согласованы. Согласование достигается изменением одного из факторов, и изменением знаков ребер c и d , если это необходимо.

Представленные количественные результаты свидетельствуют о том, что увеличение связей в ядре улучшает степень обусловленности локальности ISHARE. Для дальнейшего использования отбираются структуры ISHARE, которые согласуются в открытой локальности. Вектор маркировки определяет потенциальные ограничения эталонных моделей конкретной ISHARE.

7. Двухфакторные модели взаимодействий

Локальности SHARE взаимно пересекаются, порождая различные виды взаимодействий. Среди множества возможных вариантов можно выбрать такие, которые приводят к двухфакторным структурам. На основе этой концепции можно построить новую линию индуктивного расширения балансируемых локальностей. Исследование взаимодействий осуществляется поэтапно – от структур состоящих из двух SHARE к структурам с большим числом взаимодействующих локальностей.

Дуплеты. Дуплеты это локальности образованные двумя SHARE, которые пересекаются по общему ребру. Из трех возможных вариантов пересечений: база-база, база – боковое ребро, боковое ребро – боковое ребро, к двухфакторным структурам приводит только последний вариант (рис. 17). Пунктиром на рисунках показано дополнительное ребро, которое может существовать в дуплетах.

Наличие общего бокового ребра ведет к необходимости выполнить согласование для каждой SHARE либо по общему ребру, либо по противоположным боковым реб-

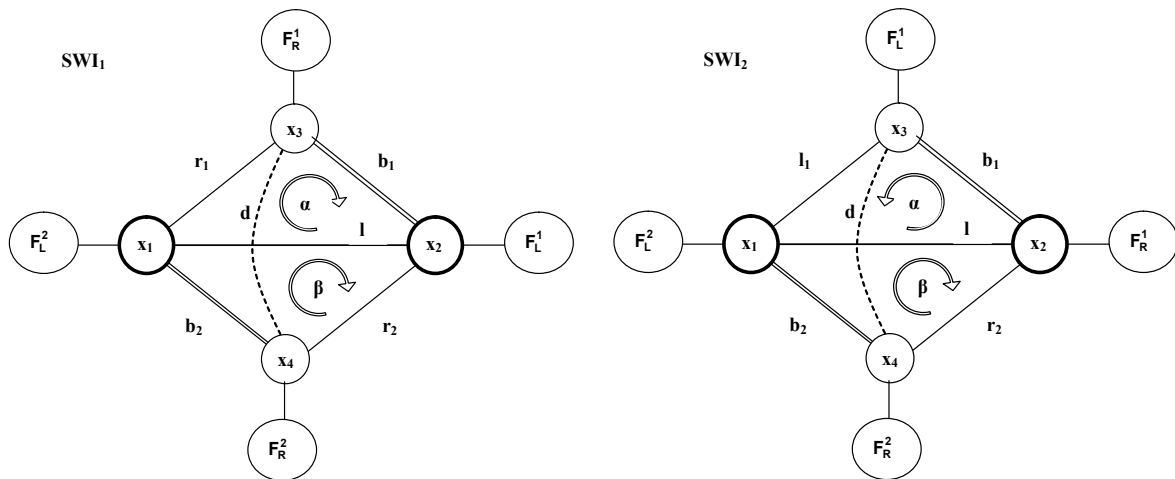


Рис. 17. Двухфакторная модель SWI – «Переключение»

рам, других вариантов быть не может. Модели SWI₁ и SWI₂ («Переключение») отличаются только ориентацией треугольников SHARE. Объединение факторов реализуется по правилам:

$$F_L = F_L^1 \cup F_L^2, \quad F_R = F_R^1 \cup F_R^2 \quad \text{для модели SWI}_1;$$

$$F_L = F_R^1 \cup F_L^2, \quad F_R = F_L^1 \cup F_R^2 \quad \text{для модели SWI}_2;$$

Поляризация факторов зависит от типа модели. Правило формирования вектора реализуемости для дуплетов имеет вид:

$$P = P_1 \cap P_2.$$

Маркировочные вектора P_1, P_2 предварительно согласуются с измененной поляризацией факторов SHARE. Факторы дуплета SWI различаются отношением к общему ребру, эту асимметрию можно использовать для мотивированной ориентации дуплетов. Назовем фактор дуплета левым (Left), если он объединяет факторы образующих SHARE действующие на общее ребро и правым (Right), если он объединяет SHARE-факторы различных ребер. Выше приведенные правила соответствуют данному соглашению. На рис. 18 показана модель SWI₂ в реальной системе данных.

Вычислим степень обусловленности ядра и рассмотрим влияние дополнительного ребра d . Обе модели не отличаются друг от друга по степени обусловленности, для определенности для анализа выберем модель SWI₁. Вначале рассмотрим модель без дополнительного ребра d . Разложим граф по вершине x_1 . После разложения получим ленточный граф из трех вершин, его определитель равен (-1) (см. п.3). Определитель для полного графа ядра представляется выражением:

$$\Delta = -1 - r_1^2 \cdot 0 - b_2^2 \cdot 0 - l^2 - 2\alpha - 2\beta - 2r_1 b_1 r_2 b_2 = -2 - 2\alpha - 2\beta - 2\alpha\beta.$$

При согласованных треугольниках получим $\Delta = -8$. При наличии дополнительного ребра d , после разложения по вершине x_1 получим граф в виде треугольника со-

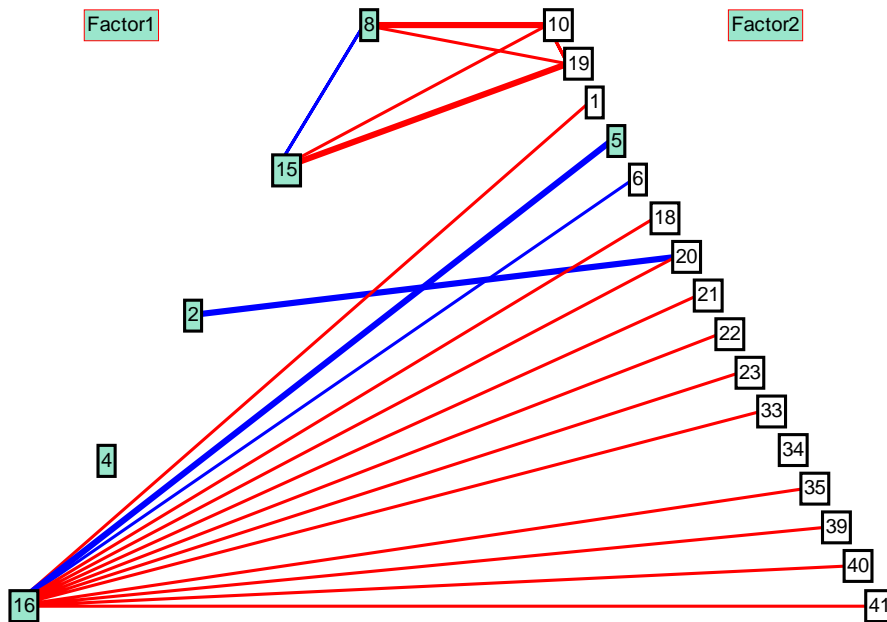


Рис. 18. Дуплет SWI в реальных данных.

стоящего из вершин x_2, x_3, x_4 с контурной передачей $\gamma = db_1 r_2$ и определителем $-2(1 + \gamma)$. В целом для ядра дуплета определитель представляется выражением:

$$\begin{aligned} \Delta &= -2(1 + \gamma) - 2\alpha - 2\beta - 2\eta - 2\alpha\beta - 2r_1 dr_2 l - 2b_2 db_1 l = \\ &= -2(1 + \gamma) - 2\alpha - 2\beta - 2\eta - 2\alpha\beta - 2\eta\beta - 2\eta\alpha. \end{aligned}$$

где $\eta = r_1 db_2$. При согласованных треугольниках получим $\Delta = -16$. Согласование достигается изменением одного из факторов, и изменением знака ребра d (если это необходимо). Таким образом, наличие дополнительного ребра улучшает степень обусловленности локальности дуплета. Для использования отбираются дуплеты, которые согласуются в открытых локальностях. Ограничения по эталонным моделям устанавливаются вектором маркировки. В реальных данных наблюдается сильная корреляция между вектором маркировки и степенью согласованности эталонных моделей. Для показанного примера (рис. 18) вектор маркировки $P = [0 \ 1]$ указывает, что реализуемой без ограничений является эталонная модель Right. Баланс графа для обеих моделей оценивался по числу несогласованных треугольников. Для модели Left это значение оказалось равным 2, а для модели Right – 31 при общем числе треугольников – 910.

Триплеты. Триплеты это локальности образованные объединением трех SHARE, которые пересекаются по общим ребрам. Двухфакторная структура триплета возникает, когда SHARE пересекаются по боковым ребрам. Двухфакторный триплет типа INT (INTERACTION) можно рассматривать как объединение двухфакторного ду-

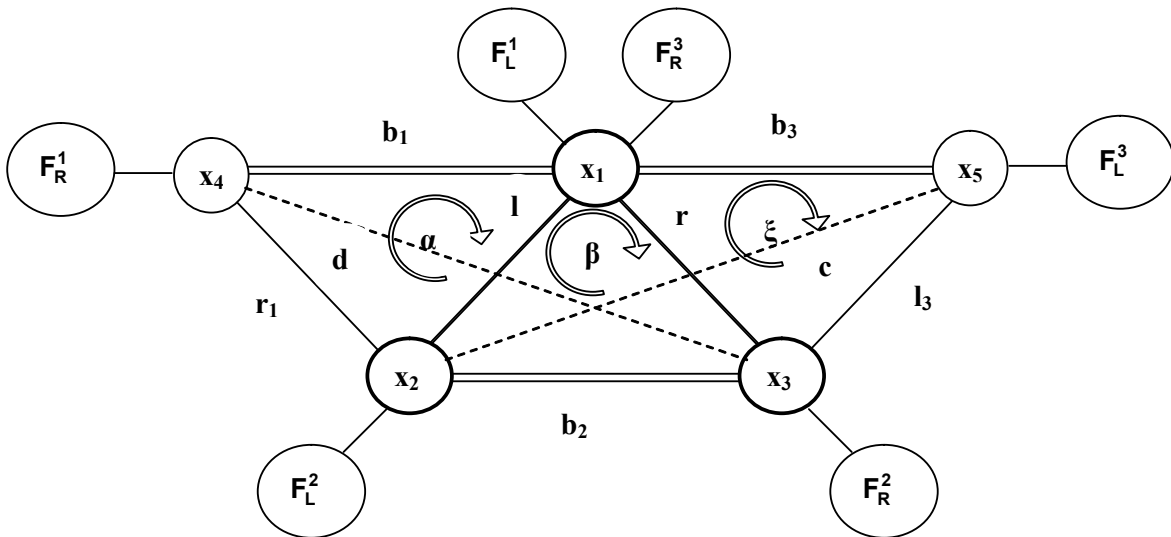


Рис. 19. Триплет, образован тремя структурами SHARE

плета SWI и одной SHARE (Рис. 19). Структура балансируется либо за счет изменения знаков ребер l, l_3 либо за счет изменения знаков ребер r, r_2 . Фактор дуплета образуется объединением факторов общего ребра и альтернативного фактора удаленного треугольника. Объединение факторов общего ребра выполняется с учетом их ориентации в каждом треугольнике: если общее ребро является левым в обоих смежных треугольниках, то объединяются факторы Left, если левым в одном и правым в другом, то объединяются соответствующие факторы Left и Right. Для структуры приведенной на рис. 19 факторы объединяются по следующему правилу:

$$F_L = F_L^1 \cup F_L^2 \cup F_L^3,$$

$$F_R = F_R^1 \cup F_R^2 \cup F_R^3.$$

В этих выражениях принято соглашение, что ориентация триплета, наследуется от ориентации дуплета.

Следует отметить, что приведенные формулы объединяют подмножества вершин, при этом для триплета может возникнуть ситуация когда сформированные множества F_L и F_R имеют непустое пересечение, что на первый взгляд противоречит двухфакторной концепции. Однако фактически объединяются не вершины факторов, а ребра, связывающие их с особыми вершинами, поскольку знаки именно этих ребер изменяются при балансировке графа. Такое понимание факторов позволяет избежать неинтерпретируемых пересечений. Для структур SHARE и ISHARE, где существует только одна особая вершина, подобных проблем не возникает.

Вектор маркировки для триплета INT определяются как пересечение векторов маркировки образующих SHARE предварительно поляризованных, по тем же правилам, что и факторы.

$$P = \bigcap_i P_i.$$

Оценим степень обусловленности структуры INT. Рассмотрим вначале вариант, когда дополнительные ребра d и c отсутствуют. Разложим граф по вершине x_1 . После удаления вершины образуется незамкнутый линейный граф, который имеет определитель равный (-1) (см. п.3). Определитель полного графа ядра следует из топологической формулы (4):

$$\Delta = -1 + b_1^2 + b_3^2 - 2\beta - 2r_1 b_1 r b_2 - 2b_3 l_3 b_2 l - 2b_2 b_3 l_3 b_2 r_1 =$$

$$= 1 - 2\beta - 2\alpha\beta - 2\xi\beta - 2\alpha\beta\xi.$$

Если все треугольники согласованы, то определитель равен $\Delta = -7$.

Рассмотрим теперь вариант с дополнительными связями. Для вычисления определителя разложим граф по вершине x_3 . После удаления вершины получим дуплет, его определитель (см. предыдущий пункт) равен:

$$\Delta^0 = -2(1 + \gamma) - 2\alpha - 2\beta - 2\eta - 2\alpha\beta - 2\eta\beta - 2\eta\alpha.$$

Определитель полного триплета получим, следуя формуле (4):

$$\Delta = \Delta^0 - 2\xi - 2c r_1 b_1 b_3 - 2c l b_3 - 2c b_2 l_3 - 2b_3 l_3 b_2 r_1 b_1 =$$

$$= \Delta^0 - 2\xi - 2\alpha\mu - 2\mu - 2\chi - 2\alpha\beta\xi,$$

где $\mu = c l b_3$, $\chi = c b_2 l_3$. При полном согласовании значение определителя будет равно $\Delta = -26$. Также как и выше рассмотренных структурах, дополнительные ребра улучшают степень обусловленности ядра. Согласование триплета достигается изменением одного из факторов и изменением знака ребер d и c (если это необходимо). Для использования отбираются триплеты, которые согласуются в их собственных открытых локальностях. Ограничения по использованию проверяются по вектору маркировки. В рассматриваемой базе данных триплеты не обнаружены.

Двухфакторные структуры высшего порядка. Структуры высшего порядка образуются как комбинации рассмотренных SHARE, дуплетов и триплетов. Двухфакторность сохраняется, когда новые структуры образуются за счет пересечения боковых ребер. Факторы объединяются и поляризуются согласно общим правилам, которые были сформулированы выше при рассмотрении триплетов. На множестве всех разде-

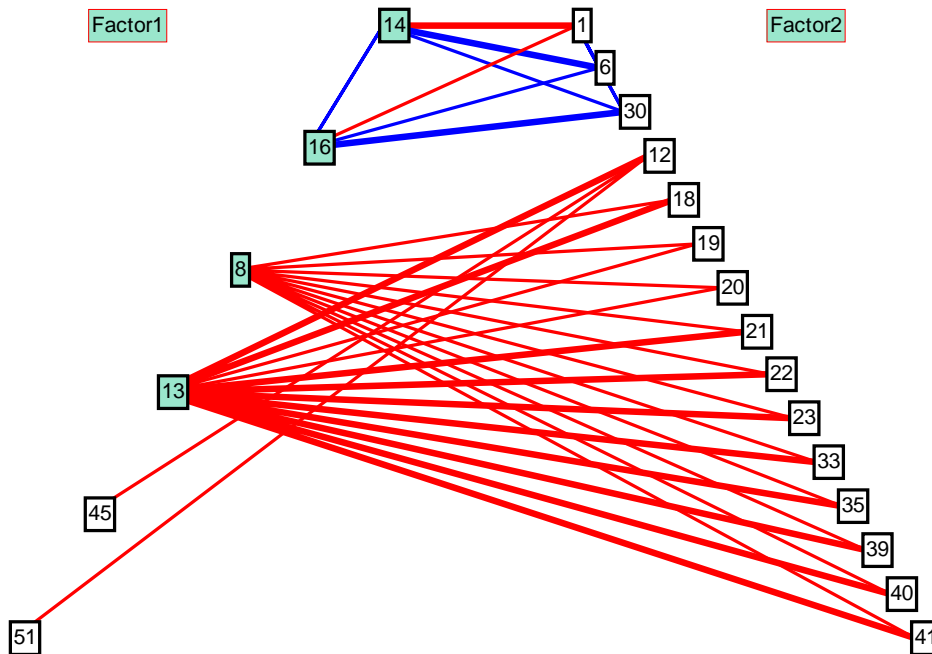


Рис. 20. Модель i3Edge. Объединяет модели SWI (вершины 30-6-14-16) и SHARE (вершины 1-14-16) которые пересекаются по общему ребру 14-16

ляющих структур можно выделить двухфакторные классы, которые имеют общее боковое ребро и две различные особые вершины, назовем такие структуры i2Edge.

Объединяя структур i2Edge с SHARE по общему ребру, получим модели i3Edge. На рис. 20 приведен пример структуры i3Edge в реальных данных. Подобную операцию системного развития можно итерационно продолжить, отбрасывая на каждом шаге структуры не балансируемые в открытой локальности.

8. Многофакторные модели взаимодействий

Знаковый баланс следует понимать как необходимое условие реальности состояний системы с согласованными значениями всех ее показателей. Каждому варианту согласования соответствует класс возможных состояний системы, которые определяют какой-то один из возможных типов поведения системы. Двухфакторные модели мотивируют два эталонных типа поведения и позволяют получить наиболее простую аппроксимацию неоднородностей сложной системы. Однако многообразие систем далеко не исчерпывается двухфакторными моделями. Концепция взаимодействия разделяющих структур легко развивается и на многофакторные представления, позволяю-

шие выразить неоднородность поведения в нескольких эталонных вариантах. Рассмотрим многофакторные модели на примере многофакторных дуплетов.

Дуплет «Поглощение (ABS)». Дуплет образуется пересечением двух разделяющих структур SHARE по общему ребру, которое является боковым ребром в одной структуре и базой в другой. Четыре возможных варианта дуплета ABS представлены на Рис. 21. Варианты отличаются ориентацией треугольников и размещением общего

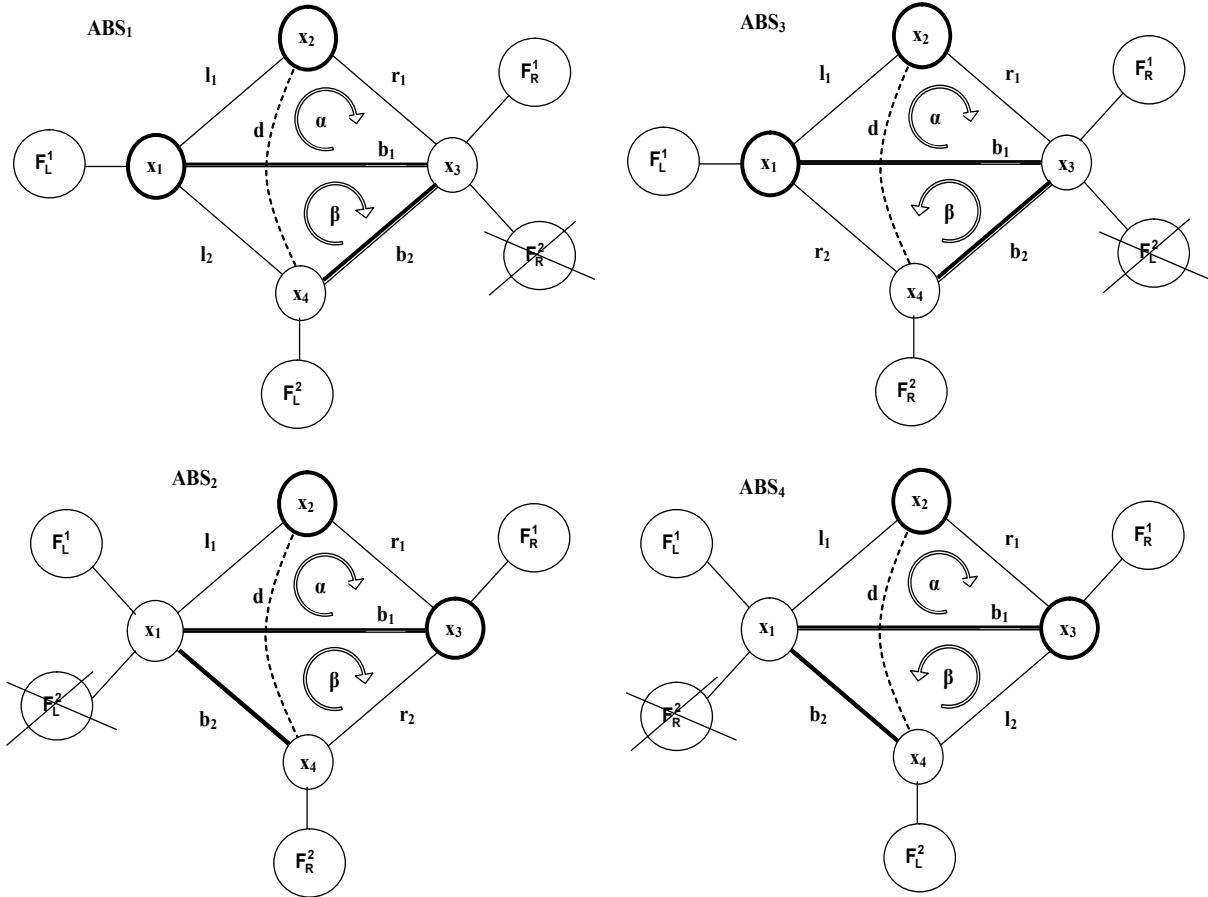


Рис. 21. Модели трехфакторного взаимодействия ABS

ребра. Для каждой модели на рисунке перечеркнут фактор, изменение связей с которым недопустимо в процедуре согласования (по принятому соглашению о константности баз). Маркерные матрицы структур представлены в таблице 1.

Матрицы образуются вертикальной конкатенацией маркерных матриц образующих структур $P_1 = [l_1 \ r_1]$ и $P_2 = [l_2 \ r_2]$, с последующей коррекцией запрещенного поля. Предварительная поляризация факторов образующих SHARE выполняется так, что-

Табл. 1. Варианты маркерных матриц модели ABS

ABS ₁	ABS ₂	ABS ₃	ABS ₄
$P = \begin{bmatrix} l_1 & r_1 \\ l_2 & 0 \end{bmatrix}$	$P = \begin{bmatrix} l_1 & r_1 \\ 0 & r_2 \end{bmatrix}$	$P = \begin{bmatrix} l_1 & r_1 \\ r_2 & 0 \end{bmatrix}$	$P = \begin{bmatrix} l_1 & r_1 \\ 0 & l_2 \end{bmatrix}$

бы в одном столбце сформировать максимальный фактор, полностью балансирующий структуру дуплета. Ассиметрия дуплета позволяет ввести мотивированную ориентацию факторов, например, назвав максимальный фактор левым, а не полный фактор – правым.

Дуплет «Подобие (SIM)» и структуры iBase. Дуплет образуются пересечением двух разделяющих структур SHARE по общей базе (рис. 22). Модель описывает четы-

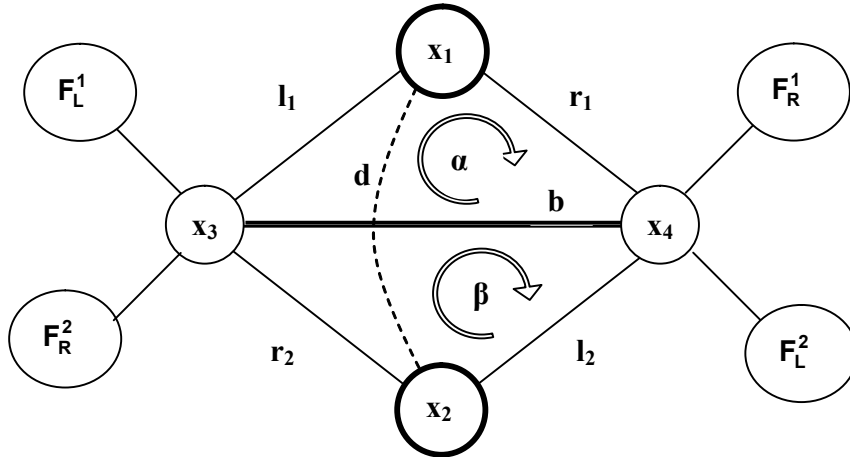


Рис. 22. Дуплет SIM

рехфакторное взаимодействие. Согласование дуплета может быть достигнуто независимым изменением пары факторов взятых по одному от каждой из образующих струк-

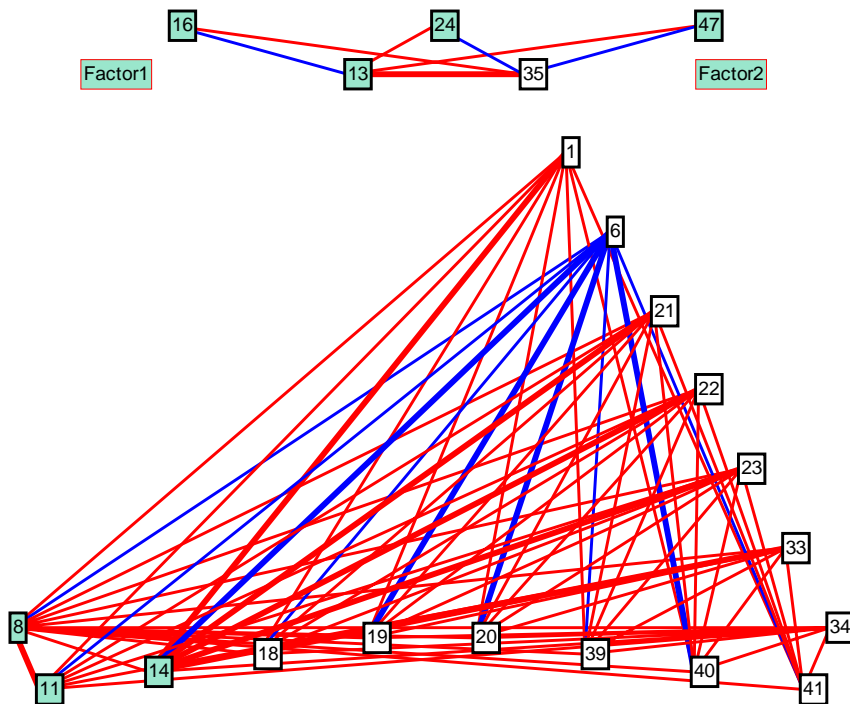


Рис. 23. Модель iBase в реальных данных. Ребро 13-35 – общая база

тур. Поляризация факторов выполняется так, чтобы при объединении SHARE совпали одноименные вершины баз. Маркерная матрица для модели формируется конкатенаци-

ей маркерных матриц образующих структур, с учетом предварительной поляризации. Для модели показанной на **рис. 22** маркерная матрица будет иметь вид:

$$P = \begin{bmatrix} l_1 & r_1 \\ r_2 & l_2 \end{bmatrix}$$

Ориентация дуплета SIM не мотивирована и может быть произвольной, фактически она назначается также как и в модели ISHARE.

На множестве разделяющих структур SHARE можно выделять однородные классы по общности базового ребра. Назовем такие структуры iBase. Структуры являются многофакторными, все образующие SHARE могут балансироваться независимо. Пример структуры iBase в реальных данных приведен на Рис. 23. Факторы образующих SHARE поляризуются также как и в модели SIM. В данной структуре вершины факторов могут пересекаться, в нижней части рисунка показаны вершины общие для обоих факторов iBase.

9. Сценарий оптимальной балансировки

Конечной целью процедуры частичной балансировки знакового графа является нахождение всех балансируемых локальностей максимального размера. Индуктивная процедура балансировки начинается с поиска треугольников противоречий, которые определяют базу для последующих шагов индукции. Если треугольники противоречий отсутствуют, то с большой вероятностью можно утверждать, что система данных сбалансирована. Для полной уверенности необходимо проверить наличие циклов из четырех и более вершин, которые не декомпозируются в треугольники.

На втором шаге находятся все разделяющие структуры SHARE. Если при наличии треугольников противоречий структуры SHARE не образуются, то это говорит о неполноте данных, и необходимости расширения их объема.

Следующий этап связан с выявлением системных взаимодействий. На этом шаге находят все двухфакторные структуры ISHARE, дуплеты, триплеты, iEdge высшего порядка и многофакторные структуры iBase, ABS и их комбинации. На каждом шаге выделяется подмножество балансируемых структур, которые являются базой следующего шага системного развития. Количество шагов индукции ограничено, поскольку с увеличением степени сложности структурных моделей происходит уменьшение числа локальностей. Сценарий завершается анализом состава обнаруженных локальностей, исследованием их на тождественность и пересечения.

Результатом реализации сценария является структурный портрет данных с набором балансируемых локальностей и выделенных ролей отдельных вершин. Каждая обнаруженная локальность несет информацию о вариантах возможного поведения системы в условиях части (компонента поведения) и является основой для выделения областей интереса и принятия компетентных решений. В своей совокупности множество балансируемых локальностей образуют базу для построения предметных знаний и разработки прикладных предметных заключений и рекомендаций.

10. Структурный портрет реальной системы данных

Для исследования методов частичной балансировки была использована база данных, полученная для задачи семантической классификации изображений [20]. Система данных представляет собой таблицу, состоящую из 56 количественных показателей и 214 объектов наблюдений. В качестве меры связи показателей использовался коэффициент парной корреляции. Связи считались значимыми, если вероятность гипотезы ошибочного значения коэффициента корреляции не превышала уровня 0.05. Построенный знаковый граф покрывает 55 вершин и имеет плотность 0.406. Граф состоит из 648 ребер, из которых 401 являются положительными и 247 отрицательными. В графе обнаружено 3281 треугольников, из которых 393 являются треугольниками противоречий. Максимальная степень вершины 38, минимальная 2. Треугольников противоречий образуют 108 разделяющих структур SHARE.

Структурный анализ. В таблице 2 представлены обнаруженные структуры различного типа. Для каждого типа указана мощность множества структур, число несущих локальностей, размер вершинного покрытия, минимальное и максимальное число вершин в локальности (объем локальностей), число особых вершин, число баз и число

Табл. 2. Структурный портрет системы данных

Тип структурной модели	Число моделей	Число локальностей	Вершинное покрытие	Объем локальностей	Число особых вершин	Число баз	Число базовых вершин
SHARE	108	52	40	4-21	13	98	39
Балансируемые SHARE	103	47	40	4-20	13	94	38
ISHARE	13	13	40	4-26	13	88	39
Балансируемые ISHARE	9	9	36	4-20	9	19	24
DOUBLET	53	17	37	6-24	12	47	30
Балансируемые DOUBLET	32	12	35	8-24	11	29	26
iBase	7	6	32	9-24	7	7	10
Балансируемые iBase	7	6	32	9-24	7	7	10
i2Edge	38	8	33	6-21	9	41	38
Балансируемые i2Edge	19	3	28	8-19	6	20	22

вершин образующих множество баз. Отдельными строками выделены подмножества структур, балансируемых в открытых локальностях. Локальности могут быть носителями одновременно для нескольких структур, распределения разделяющих структур SHARE по несущим локальностям показан на рис. 24.

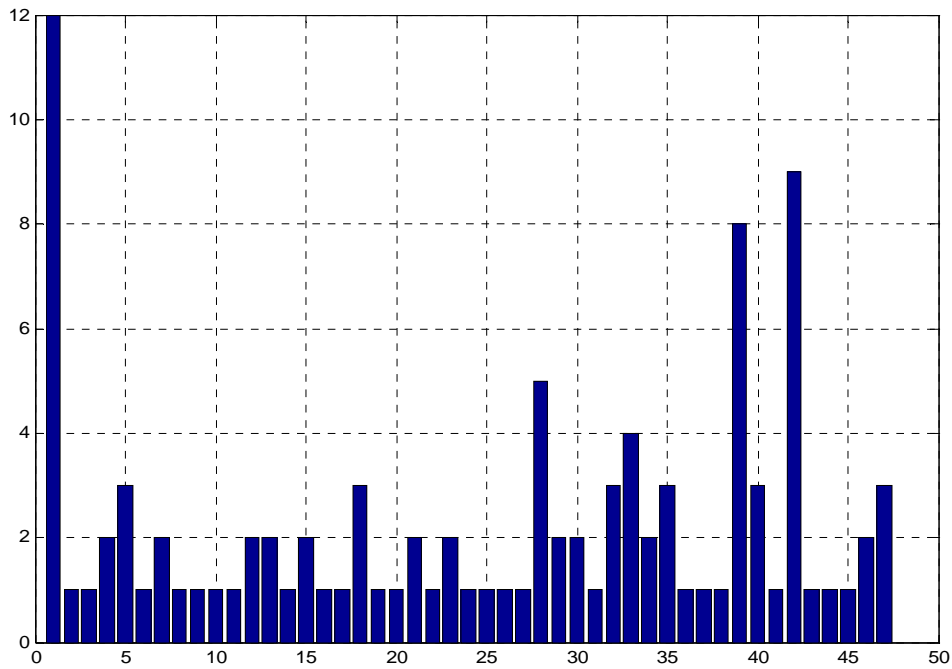


Рис. 24. Представительность разделяющих структур SHARE в локальностях

Балансируемые SHARE были использованы для выделения дуплетов SWI, SIM, ABS. Обнаружено 53 дуплета, из них 32 балансируемых. На рис. 25 показано распределение вершин балансируемых дуплетов на множестве показателей. Ряд дуплетов имеют общую локальность. Диаграмма пресечений локальностей для дуплетов показана на рис. 26.

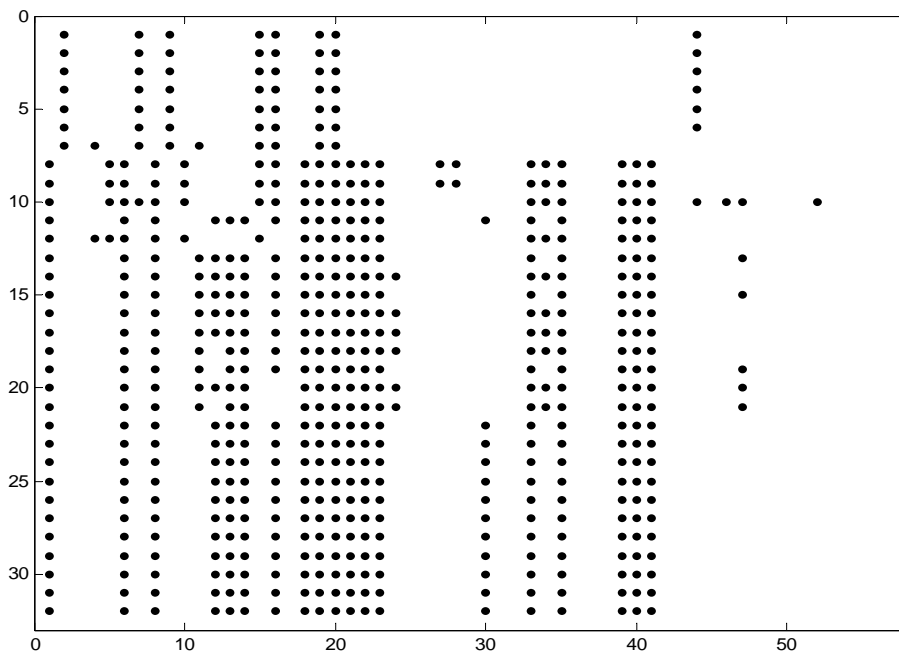


Рис. 25. Распределение вершин балансируемых дуплетов на множестве показателей

На множестве SHARE обнаружено 7 структур iBase все структуры данного типа

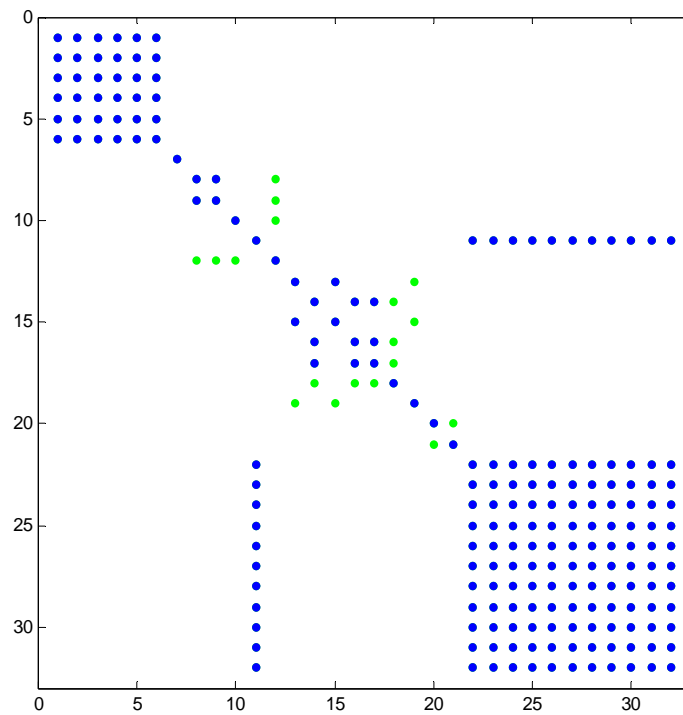


Рис. 26. Совпадение локальностей для дуплетов (темные точки), светлыми точками выделены совпадения локальностей по ядрам являются балансируемыми. Локальности i2Edge состоят из дуплетов SWI.

Оценка ролей вершин. После выделения разделяющих структур вершины графа связей становятся различимыми по своей роли в выявлении присущих системе неоднородностей. Классификация множества вершин для системы данных показана на

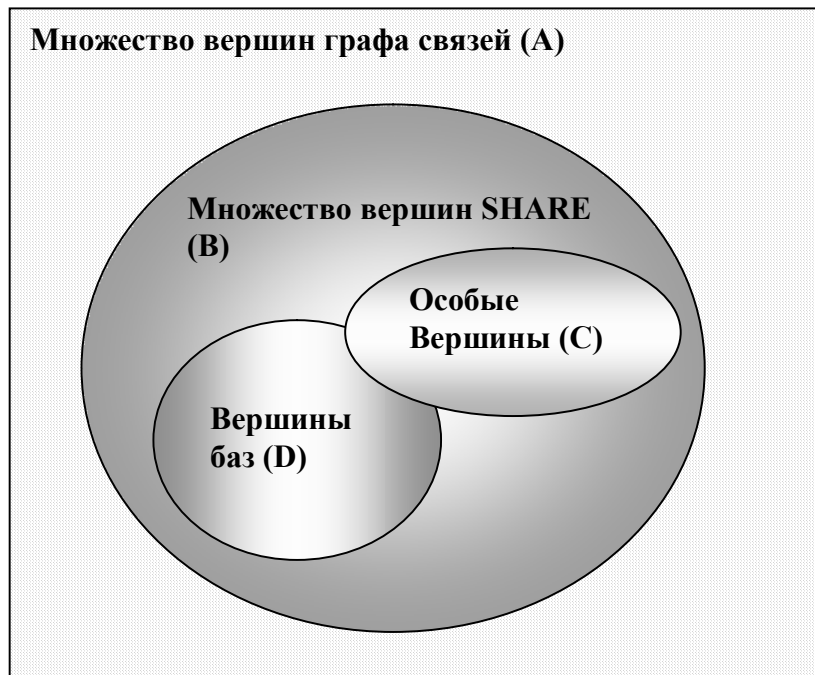


Рис. 27. Классификация вершин

рис. 27. Для оценки роли вершин сделаны несколько экспериментов. Результаты экспериментов представлены в таблице 3.

В первом эксперименте граф связей включал 40 вершин, которые составляют полное покрытие локальностями и полное ядерное покрытие. Этим вершинам оказалось достаточно, чтобы проявились практически все характерные неоднородности системы данных.

Во втором эксперименте сохранены только вершины баз (39). Изменение покрытия на одну вершину, привело к уменьшению числа SHARE и особых вершин, но число локальностей сохранилось. Таким образом, базы влияют на представительность локальностей и отражают природу неоднородностей.

В третьем эксперименте из базы данных были удалены особые вершины. В результате исчезли практически все локальности, но появились две дополнительные особые вершины, которые отсутствовали в исходной базе. Объясняется это тем, что в результате удаления некоторых вершин окружения два предсинглета перешли в разряд синглетов и образовали 20 структур SHARE. Вторая итерация по удалению особых вершин (эксперимент 4) полностью завершила процесс удаления локальностей. Редуцированная система данных сохранила противоречия, которые не могут быть разрешены в рамках данной парадигмы. Это указывает на факт недостаточной представительности системы данных для 40 оставшихся показателей.

В пятом эксперименте база данных содержала 13 показателей, которые являются особыми вершинами в исходной базе данных. В результате сохранено 13 различных локальностей с единичным наполнением, число особых вершин уменьшилось до 8. Эти

Табл. 3. Оценка роли вершин знакового графа реальной системы данных

Состав системы данных	Число вершин в графе связей	Общее число треугольников	Число треугольников противоречий	Число локальностей	Число SHARE	Размер множества вершин SHARE	Число особых вершин	Число вершин баз
0. (A)	55	3281	393	47	108	40	13	39
1.(B)	40	2535	292	47	108	40	13	39
2. (D)	39	2501	272	47	102	39	11	38
3. (A-C) Итер. 1.	42	1290	129	3	20	17	2	12
4.(A-C) Итер. 2.	40	1032	87	0	0	0	0	0
5. (C)	13	36	17	13	13	11	8	11
6. (A-B)	14	49	0	0	0	0	0	0

13 показателей выделяют кластер неоднородностей, обусловленных сложными внутрисистемными взаимодействиями.

В последнем эксперименте были удалены все вершины покрываемые множеством локальностей SHARE. В результате 14 оставшихся показателей образовали граф без противоречий, с низкой плотностью связей.

Оценочный анализ позволяет выделить наиболее информативный состав вершин, сделать выводы о полноте системы данных

11. Заключение

В данной работе показано, что проблема знакового баланса, может быть разрешена через индуктивное исследование структуры знакового графа. Основная конструктивная идея может быть выражена фразой «если нельзя достигнуть полного баланса, то можно попытаться выделить ограниченные области знакового графа, для которых возможности и варианты балансировки максимально обоснованы структурой графа». Рабочим инструментом исследования являются первичные системные единицы (SHARE), которые представляют собой минимальные артефакты, локально выражающие логические принципы симметрии и двухфакторного разделения. Доказанная теорема «О разделении граней тетраэдра» дает строгое математическое обоснование новой парадигме балансировки знаковых графов. Первичная системная единица наделена морфологией: в ней имеется особая вершина, вершины базы, база, боковые ребра и вершины непересекающихся факторов. Эти морфологические различия – дополнительная информация, позволяющая реализовать итерационный механизм структурного анализа знакового графа. Структуры SHARE локализуют неоднородности системы данных и через критерий знакового баланса выражают локальные механизмы их порождения. В работе показано, что критерий знакового баланса коррелируется с мерой обусловленности системной матрицы ядра структурных единиц. Для анализа степени обусловленности предложено использовать топологические формулы обобщенного сигнального графа. Оптимальность процедур балансировки в локальностях базируется на принципе минимальных изменений знакового графа и оценки степени обусловленности структурных единиц.

Концепция взаимодействий первичных единиц, определяет пути расширения балансируемых локальностей и одновременно выявляет возможные варианты балансировки. Обоснование выбора моделей взаимодействия подтверждается количественной оценкой степени обусловленности структурных образований. По существу совокупность эталонных моделей выражает метафорические аттракторы реальной системы, т.е. множества состояний которые являются наиболее типичными проявлениями поведения системы, а портретные образы системы данных образуют базу для построения предметных знания и разработки прикладных проектных решений.

Литература

1. Cartwright, Dorwin, and Harary, F. (1956). Structural balance: A generalization of Heider's theory. *Psychological Review* 63: 277-292.
2. Harary, F.: On the notion of balance of a signed graph. *Michigan Mathematical Journal*, 2, 143-146 (1953-54).
3. Abelson, R.P., Aronson, E., McGuire, T.W., Newcomb, T.M., Rosenberg, M.J., & Tannenbaum, O.H. (Eds.), 1968. *Theories of cognitive consistency*. Holt, Rinehart and Winston, New York.
4. Карбаев Д.С., Кораблин М.А., Смирнов С.В. Регрессионно-когнитивные графы в задачах макроэкономического прогнозирования / Вестник компьютерных и информационных технологий №03, 2007.
5. Notsu Akira, Ichihashi Hidetomo and Katsuhiro Honda. Agent Simulation Based on Perceptual Balance / *IJCSNS International Journal of Computer Science and Network Security*, VOL.6 No.12, December 2006.- pp.50-54.
6. Кораблин М.А., Мелик-Шахназаров А.В., Салмин А.А. Регрессионно-когнитивные графы в задачах анализа биллинговых систем // Информационные технологии, 2005, № 8, стр. 35-39.
7. Conditional colorability in graphs. *Graphs and Applications*. (F. Harary, J. Maybee, Eds.) Wiley, New York (1985) 127-136.
8. Harary, F., 1959. On the measurement of structural balance. *Behavioural Science* 4, 316-323.
9. Kabell J.A. A simple algorithm to detect balance in signed graphs. *Math. Social Sci.* 1 (1980) 131-136.
10. Roberts F. S. *Discrete Mathematical Models_ With Applications to Social, Biological and Environmental Problems Chapter _ Signed Graphs and the Theory of Structural Balance* p.64-81 (1976) Prentice Hall, Inc. Englewood Clis. New Yersey.
11. Робертс Ф.С. Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам. Пер. с англ. М.: Наука, 1986.
12. Doreian P., Mrvar A., 1996. A partitioning approach to structural balance. *Social Networks* 18, 149–168.
13. Doreian P., Hummon N. P. Some dynamics of social balance processes: bringing Heider back into balance theory, *Social Networks*, Vol. 25, 2003, pp. 17-49.
14. Качанова Т.Л., Фомин Б.Ф. Симметрии, взаимодействия в локальностях, компоненты поведения сложных систем / Издательско-полиграфический центр СПбГЭТУ, СПб, 1998.- 126с. (Препринт № 2).
15. Качанова Т.Л., Фомин Б.Ф. Метатехнология системных реконструкций: СПб: Изд-во СПбГЭТУ «ЛЭТИ», 2002. 336с.
16. Качанова Т.Л., Фомин Б.Ф. Технология системных реконструкций.- СПб.: Политехника, 2003.- 146с. – (Проблемы инновационного развития. Вып.2).

-
17. Бахвалов Н.С. Численные методы (анализ, алгебра, обыкновенные дифференциальные уравнения) М.: Наука, 1975.- 631с.
 18. Анисимов В.И. Топологический расчет электронных схем. Л., «Энергия», 1977. 240с.
 19. Гантмахер Ф.Р. Теория матриц.- 4-е изд. – М.: Наука, 1988.- 552с.
 20. Дорогов А.Ю., Курбанов Р.Г., Разин В.В Быстродействующий алгоритм семантической классификации JPEG-изображений / Электронный журнал «Нейроинформатика» 2006, том 1, № 2 с.124-144. – <http://www.ni.iomt.ru/Journal/N2/>.