В.С. Левченков, Л.Г. Левченкова

ИСПОЛЬЗОВАНИЕ МЕТОДА МУЛЬТИОТНОШЕНИЙ ДЛЯ ПРОГНОЗИРОВАНИЯ СТРУКТУРЫ СВЯЗЕЙ В СОЦИАЛЬНЫХ СЕТЯХ [1]

Введение

Под социальной сетью обычно понимают структуру связей, возникающих в коллективе, участники которого обладают определенными общими интересами. Примером социальной сети является совокупность ученых, работающих в некоторой области науки и связанных друг с другом совместными научными работами. Другой пример: социальная сеть большой программистской компании, группы сотрудников которой выполняют совместные проекты.

С формальной точки зрения социальная сеть – это граф с множеством вершин, фиксирующих совокупность индивидуумов в сети, и дугами, показывающими взаимодействие между людьми и порождающими структуры с разнообразными свойствами (см. обзорную работу [1]). Очень интересной проблемой, связанной с описанием временной эволюции таких сетей является прогнозирование новых связей, появляющихся в сети с течением времени [2]. Она формализуется следующим образом. Пусть мы знаем в какой-то момент t (или точнее, промежуток времени I) состояние социальной сети, т.е. вершины графа сети и все дуги. Можно ли только на основе информации, описывающей это состояния сделать оценку того, какие новые дуги появятся в этом графе к другому моменту времени t' (или промежутку времени I'). В работе [2] эта проблема рассматривается для конкретного случая социальной сети ученых, работающих в ряде областей физики и публикующих статьи по определенной тематике. Для оценки различных теоретических моделей авторы используют базы данных по научным статьям в пяти областях физики за период 1994-1999гг. Исходные данные в каждой базе разбиваются на две группы: в первой содержатся работы за промежуток времени $I = [1994, 1996]$, а во второй – за промежуток $I' = [1997, 1999]$. Социальная сеть на каждом интервале I и I' описывается графами $G = (A, E)$ и $G' = (A, E')$, соответственно. Множество вершин A

[1]Работа выполнена при поддержке Гранта Президиума РАН по проекту ИКС.

отвечает только тем авторам, которые в течение I и I' имели не менее трех публикаций. Этим приемом снимается проблема, связанная с различием состава авторов в промежутках I и I'. Множество дуг E (E') в графах G (G') будем строить так: неориентированная дуга ab между вершинами $a, b \in A$ графа G (G') проводится тогда, когда авторы a и b в течение промежутка времени I (I') участвуют хотя бы в одной совместной работе. Петли в графах G и G' отсутствуют. Далее, по этим графам находится множество $E_0 = E' \cap \bar{E}$ новых дуг, появившихся в G' по сравнению с графом G. Здесь

$$\bar{E} = L(A) \setminus E, \tag{1}$$

а $L(A) = \{ab : \ a \neq b \& \ a, b \in A\}$. Таким образом, $\bar{E}$ – это множество дуг дополнения графа G.

Пусть число новых дуг равно n, $|E_0| = n$. Задача предсказания новых связей, согласно [2], состоит в нахождении по структуре графа G такого упорядочения множества $\bar{E}$, чтобы E_{max} – совокупность n лучших элементов по этому упорядочению – содержала максимально возможное число элементов из множества E_0. Численной оценкой эффективности прогноза служит при этом мощность m множества $E_{max} \cap E_0$, $m = |E_{max} \cap E_0|$. В работе [2] на примере пяти баз данных было проанализировано действие более десятка различных методов упорядочения (простых и весьма замысловатых) и высказано аргументированное суждение, что знание структуры графа G позволяет сделать определенный вывод о появлении новых связей между заданными вершинами. Отметим, что наиболее успешные упорядочения приводили к значениям m, в 30-50 раз превосходящим среднее число совпадений в случае равновероятного упорядочения множества $\bar{E}$.

В настоящей работе мы предлагаем новый метод упорядочения, основанный на самосогласованном подходе [3]-[5], развитом в теории выбора. Этот метод позволяет учесть то обстоятельство, что интервал I' имеет конечную длину (в эксперименте его длина составляет три года), а появление дуг не носит одномоментный характер, но распределено по времени, что приводит к влиянию вновь появившихся дуг на характер прогнозируемого упорядочения.

1. Метод мультиотношений для предсказания связей в социальной сети

Пусть нам известна структура графа $G = (A, E)$, и наша задача состоит в том, чтобы упорядочить множество $\bar{E}$ из (1) в соответствии с возможностью появления его элементов в качестве дуг в графе $G' = (A, E')$ (естественно, при построении прогноза мы не знаем дуг графа G'). Подойдем к этой задаче с точки зрения группового выбора. Будем считать, что каждый элемент $a \in A$ ($|A| = \Delta$) служит источником парной оценки элементов $\bar{E}$, порождая бинарное отношение r_a на $\bar{E}$, строящееся только на основе структуры множества дуг E. Отношение r_a будем описывать его характеристической формой

$$r_a(cd, gh) = \begin{cases} 1, & \text{если справедливо } cd\, r_a\, gh \\ 0 & \text{в противном случае.} \end{cases} \tag{2}$$

Наличие отношения $cd\ r_a\ gh$ содержательно означает, что "возможность образования дуги cd в графе G не меньше возможности образования дуги gh". Формально бинарное отношение r_a имеет следующий вид

$$r_a(cd, gh) = \begin{cases} 0, & \text{если } ag \in E \& ah \in E \& (ac \in \bar{E} \vee ad \in \bar{E}) \\ 1 & \text{во всех остальных случаях.} \end{cases} \tag{3}$$

Его выбор основан на следующих соображениях. Автор $a \in A$, оценивая возможность появления совместных работ с участием $\{c, d\}$ или $\{g, h\}$ руководствуется имеющимся у него опытом написания совместных работ с каждым из рассматриваемых авторов. Если с g и h у него уже есть совместные работы, а с автором c или автором d таких работ нет, то он считает, что возможность возникновения совместной работы у $\{c, d\}$ меньше, чем у $\{g, h\}$. Во всех остальных случаях предполагается, что такая возможность у пары $\{c, d\}$ не меньше, чем у пары $\{g, h\}$.

Построенную согласно (3) систему бинарных отношений на множестве $\bar{E}$ мы объединим в профиль отношений $(r_a)_{a \in A}$ и рассмотрим задачу упорядочения множества $\bar{E}$ с точки зрения теории группового выбора.

Применим для построения упорядочения специальный метод (будем называть его в контексте данной работы *методом мультиотношений*), позволяющий по профилю отношений вида (3) найти скалярный критерий, определяемый с точностью до умножения на постоянный для всех элементов $\bar{E}$ множитель (см. [3]-[5]). Это обстоятельство важно по той причине, что такой критерий, нормированный на 1, находится однозначно и может быть проинтерпретирован как вероятность возникновения новой дуги в графе G.

Прежде чем написать систему уравнений, которой удовлетворяет искомый критерий, преобразуем отношения профиля. Потребность в этом видна из следующих соображений. Значения отношения r_a на элементах (дугах) $l, l' \in \bar{E}$ содержательно можно интерпретировать посредством следующих высказываний:

(a) $(r_a(l,l') = 1) \,\&\, (r_a(l',l) = 0)$ – "возможность появления дуги l больше, чем возможность появления дуги l'" (обозначим эту ситуацию как $l > l'$);

(b) $(r_a(l,l') = r_a(l',l) = 1)$ – "возможности появления дуг l и l' одинаковы" ($l \sim l'$);

(c) $(r_a(l',l) = 1) \,\&\, (r_a(l,l') = 0)$ – "возможность появления дуги l меньше, чем возможность появления дуги l'" ($l < l'$).

Заметим, что поскольку в каждом из этих случаев дается сравнительная оценка пары дуг, то фактически возникает не одна, а две ситуации. Действительно, в случае (a) имеет место не только $l > l'$, но и $l' < l$; в случае (b) – $l \sim l'$ и $l' \sim l$; в случае (c) – $l < l'$ и $l' > l$. При этом значение отношения r_a на паре элементов l и l' определяет некоторую численную оценку соответствующей ситуации. Если использовать терминологию количественной оценки спортивных состязаний, то в случае (b) элементы l и l' получают по одному очку: $r_a(l,l') = 1$; $r_a(l',l) = 1$; в случае (a) элемент l получает одно очко, а элемент l' ни одного: $r_a(l,l') = 1$; $r_a(l',l) = 0$; и аналогично для случая (c). Однако, при такой интерпретации сразу же бросается в глаза расхождение численных оценок ситуаций в (a), (c) и (b): в первых двух случаях элементы вместе получают одно очко, а в случае (c) – два очка. Если мы хотим приписать элементам из $\bar{E}$ некоторые "справедливые" оценки, то такое расхождение необходимо устранить. В теории мультиотношений [3]-[5] это достигается на основе использования операции z-нормализации, когда вместо значения характеристической функции $r_a(l,l')$ рассматривается число

$$\tilde{r}_a(l,l') = \frac{z r_a(l,l')}{r_a(l,l') + r_a(l',l)}, \tag{4}$$

где z выбирается таким образом, чтобы величины $\tilde{r}_a(l,l')$ при всех $l, l' \in \bar{E}$ были целыми неотрицательными числами (такая функция $\tilde{r}_a$ в [3]-[5] называется *мультиотношением*). В частном случае r_a вида (3) z достаточно выбрать равным 2. Тогда при любых $l, l' \in \bar{E}$ выполняется соотношение

$$\tilde{r}_a(l,l') + \tilde{r}_a(l',l) = 2. \tag{5}$$

Аналог соотношения (5) легко обнаружить в системах оценки результатов спортивных турниров. Например, результаты игр футбольных команд (победа – 2 очка, ничья – 1 очко) непосредственно удовлетворяет условию (5). В шахматных турнирах, чтобы обнаружить соотношение (5), достаточно результат (победа – 1 очко, ничья – 1/2 очка) домножить на 2.

Нормализованные отношения используются при групповой оценке однородным образом путем перехода к групповому мультиотношению $m(l, l')$ на $\bar{E}$

$$m(l, l') = \sum_{a \in A} \tilde{r}_a(l, l'). \tag{6}$$

Систему чисел $m(l, l')$ можно представить несколько иным способом, более удобным для проведения вычислений и допускающим интересное обобщение. Введем в рассмотрение функцию "ближайшего соседства" дуги l и элемента a, $f_a(l)$, определенную на множестве $\bar{E}$,

$$\forall cd \in \bar{E} \quad f_a(cd) = \begin{cases} 1, & \text{если } ac \in \bar{E} \& ad \in \bar{E} \\ 0 & \text{в противном случае.} \end{cases} \tag{7}$$

Припишем каждой дуге $l \in \bar{E}$ неотрицательное число

$$\rho(l) = \sum_{a \in A} f_a(l). \tag{8}$$

Значение этой функции на дуге $cd \in \bar{E}$ равно суммарному числу общих "ближайших соседей" элементов c и d, то есть числу тех вершин в графе G, которые соединены дугой как с вершиной c, так и с вершиной d. Заметим, что поскольку в графе G нет петель, то справедливо неравенство

$$\forall l \in \bar{E} \quad \rho(l) \leq \Delta - 2. \tag{9}$$

Утверждение 1.

$$m(l, l') = \Delta + \rho(l) - \rho(l'). \tag{10}$$

Доказательство. Согласно (3) и (4) при $z = 2$ значения $\tilde{r}_a(l, l')$ можно записать в виде

$$\tilde{r}_a(l, l') = \begin{cases} 2, & \text{если } f_a(l) > f_a(l') \\ 1, & \text{если } f_a(l) = f_a(l') \\ 0, & \text{если } f_a(l) < f_a(l'). \end{cases}$$

Поскольку $f_a(l)$ принимает только два значения (0 или 1), то отсюда вытекает, что

$$\tilde{r}_a(l, l') = 1 + f_a(l) - f_a(l'),$$

откуда согласно (8) мы и получим соотношение (10). *Q.E.D.*

Сопоставим теперь мультиотношению (10) целочисленную неотрицательную матрицу $M = (m_{ll'})_{l,l' \in \bar{E}}$ с элементами

$$m_{ll'} = \begin{cases} m(l', l), & \text{если } l' \neq l \\ 0, & \text{если } \ l' = l. \end{cases} \tag{11}$$

Утверждение 2. Матрица M неразложима.

Доказательство следует из того факта, что в силу (9) все недиагональные элементы матрицы M положительны. Такая матрица, очевидно, неразложима.

Поскольку матрица M неразложима, то метод мультиотношений [3]-[5] строит упорядочение R на множестве $\bar{E}$ следующим образом. Матрица M преобразуется в так называемую матрицу переходов $T = (t_{ll'})_{l,l' \in \bar{E}}$, элементы которой полностью определяются мультиотношением (10),

$$t_{ll'} = \begin{cases} m(l', l), & \text{если } l' \neq l \\ \sum\limits_{l'' \in \bar{E}} m(l, l''), & \text{если } \ l' = l. \end{cases} \tag{12}$$

Искомое упорядочение R на $\bar{E}$ является слабым порядком (т.е. рефлексивным, связным и транзитивным бинарным отношением) и определяется скалярным критерием $\sigma = (\sigma(l))_{l \in \bar{E}}$

$$\forall l, l' \in \bar{E}, \quad lRl' \Leftrightarrow \sigma(l) \geq \sigma(l'). \tag{13}$$

Этот скалярный критерий в общем случае находится из следующей системы уравнений

$$\sum_{l \in \bar{E}} \sigma(l) = 1, \qquad \sigma(l) = u(l)v(l), \qquad Tu = \lambda_0 u, \qquad vT = \lambda_0 v, \tag{14}$$

где λ_0 – максимальное собственное значение неразложимой матрицы T.

Утверждение 3. Для мультиотношения вида (10) скалярный критерий σ удовлетворяет упрощенной системе уравнений

$$\sum_{l \in \bar{E}} \sigma(l) = 1, \qquad \sigma T = 2\Delta(|\bar{E}| - 1)\sigma, \tag{15}$$

где $|\bar{E}|$ – число элементов в множестве $\bar{E}$.

Доказательство. Матрица T, строящаяся согласно (12) по мультиотношению (10), имеет постоянную строчную сумму. Действительно,

$$\sum_{l' \in \bar{E}} t_{ll'} = \sum_{l' \in \bar{E} \setminus \{l\}} (m(l, l') + m(l', l)) = 2\Delta(|\bar{E}| - 1). \tag{16}$$

Согласно теореме Фробениуса-Перрона [6], отсюда следует, что $\lambda_0 = 2\Delta(|\bar{E}| - 1)$, а вектор u в (14) можно положить равным единичному, т.е. считать $u = (1, ..., 1)$. В результате, критерий σ оказывается пропорциональным вектору v и система (14) переходит в (15). *Q.E.D.*

Утверждение 4. Упорядочения, порождаемые критерием $\sigma = (\sigma(l))_{l \in \bar{E}}$ и функцией $\rho(l)$ из (8), совпадают, т.е.

$$\forall l, l', \quad \begin{array}{l} \rho(l) \geq \rho(l') \Leftrightarrow \sigma(l) \geq \sigma(l') \\ \rho(l) = \rho(l') \Leftrightarrow \sigma(l) = \sigma(l'). \end{array}$$

Доказательство.

1. Пусть l и l' таковы, что $\rho(l) \geq \rho(l')$. Сравним между собой значения критерия σ на этих элементах. Для этого выделим из системы (15) следующие два уравнения

$$\begin{array}{l} \sum\limits_{l'' \neq l, l'} t_{l''l}\sigma(l'') + t_{l'l}\sigma(l') = \sum\limits_{l'' \neq l} t_{ll''}\sigma(l), \\ \sum\limits_{l'' \neq l, l'} t_{l''l'}\sigma(l'') + t_{ll'}\sigma(l) = \sum\limits_{l'' \neq l'} t_{l'l''}\sigma(l'). \end{array} \tag{17}$$

При получении уравнений (17) мы использовали соотношение (16), связывающее максимальное собственное значение матрицы T с ее строчной суммой.

Предположение $\rho(l) \geq \rho(l')$ в силу (10), (12) приводит к следующим условиям на элементы матрицы T:

(a) $t_{l'l} \geq t_{ll'}$,

(b) $\forall l'' \in \bar{E} \setminus \{l, l'\} \quad t_{l''l} \geq t_{l''l'}$.

Вычтем теперь из первого уравнения системы (17) второе уравнение. С учетом условия (b), эта процедура дает такое неравенство

$$\sigma(l) \left(\sum_{l'' \neq l} t_{ll''} + t_{ll'} \right) \geq \sigma(l') \left(\sum_{l'' \neq l'} t_{l'l''} + t_{l'l} \right). \tag{18}$$

Согласно (a), (b) и соотношению $m(l, l') + m(l', l) = 2\Delta$, вытекающему из (10), имеем: $t_{ll'} \leq t_{l'l}$ и $\sum\limits_{l'' \neq l} t_{ll''} \leq \sum\limits_{l'' \neq l'} t_{l'l''}$. Значит,

$$\left(\sum_{l'' \neq l} t_{ll''} + t_{ll'} \right) \leq \left(\sum_{l'' \neq l} t_{l'l''} + t_{l'l} \right).$$

Тогда из (18) вытекает $\sigma(l) \geq \sigma(l')$.

2. Пусть теперь $\rho(l) = \rho(l')$. Вычитая в (17) из первого уравнения второе и подставляя вместо элементов матрицы их выражения через значения функции $\rho(l)$, получим

$$0 = \sum_{l''\neq l,l'} \sigma(l)\,(\rho(l) - \rho(l')) =$$

$$= \sum_{l''\neq l,l'} (\Delta + \rho(l'') - \rho(l))(\sigma(l) - \sigma(l')) + 2\Delta(\sigma(l) - \sigma(l')),$$

откуда в силу положительности коэффициентов при $(\sigma(l) - \sigma(l'))$ в правой части, получим $\sigma(l) = \sigma(l')$.

3. Предположим, что $\sigma(l) = \sigma(l')$. Тогда проводя соответствующее вычитание в системе (17), придем к уравнению

$$\sum_{l''\neq l,l'} \sigma(l'')\,(\rho(l) - \rho(l')) + 2\sigma(l)\,(\rho(l) - \rho(l')) =$$
$$= \sum_{l''\neq l,l'} \sigma(l)\,(\rho(l') - \rho(l)) + 2\sigma(l)\,(\rho(l') - \rho(l))\,,$$

которое в силу положительности компонент $\sigma(l)$ собственного вектора, отвечающего максимальному собственному значению, дает $\rho(l) = \rho(l')$.

4. Пусть $\sigma(l) \geq \sigma(l')$. Предположим, что $\rho(l') > \rho(l)$. Тогда согласно п.1 доказательства, $\sigma(l') \geq \sigma(l)$, что с учетом посылки приводит к $\sigma(l') = \sigma(l)$. Значит, согласно п.3 доказательства, $\rho(l') = \rho(l)$, что противоречит предположению. Таким образом, $\sigma(l) \geq \sigma(l') \Rightarrow \rho(l) \geq \rho(l')$. *Q.E.D.*

Следствие. Упорядочение по методу мультиотношений совпадает с упорядочением по правилу "ближайших соседей" [2].

Как выяснено в работе [2], правило "ближайших соседей" оказывается весьма хорошим методом для прогнозирования. К сожалению, оно имеет весьма существенный недостаток: это правило ставит на последние места в упорядочении те дуги из $\bar{E}$, вершины которых вообще не имеют общих соседей. Однако, как следует из анализа баз данных, дуги такого типа довольно часто возникают на практике. Этот эффект можно учесть на основе естественного обобщения метода мультиотношений.

2. Влияние "виртуальных" дуг на эволюцию социальной сети

Как мы отмечали ранее, появление новых дуг в графе $G = (A, E)$ в процессе эволюции социальной сети рассматривается на протяжении некоторого периода времени I' и фиксируется в виде графа $G' = (A, E')$ (в работе [2] I' равняется трем годам). Предположим, что в течение первого года периода I' возникло несколько новых дуг (т.е. появился ряд совместных работ авторов, ранее не работавших вместе). Образование новых связей изменяет исходную информацию, представленную графом G, что, очевидно, должно сказаться на прогнозе, который должен формироваться теперь на основе более широкого множества дуг $E^{(1)} \supset E$. Это множество, в свою очередь, приводит к возможности возникновения на втором году периода I' другого множества дуг $E^{(2)} \supset E^{(1)}$ и т.д. В результате прогнозируемое множество связей должно зависеть не только от состава дуг E, но и от вновь возникших дуг E', что приводит к необходимости построения внутренне согласованной процедуры прогноза.

Метод мультиотношений дает естественную возможность для реализации такой процедуры. Действительно, центральным для этого метода является представление исходной информации в виде некоторого мультиотношения на множестве $\bar{E}$, используемого для построения матрицы переходов (12). Эта матрица входит в систему уравнений (14) или (15) однородным образом: умножение всех ее элементов на одно и то же число не влияет на вид скалярного критерия. Это означает, что в качестве значений мультиотношения можно использовать не только целые, но и рациональные числа, а поскольку множество рациональных чисел плотно в R^1, то и любые действительные числа. Пользуясь этой возможностью, переопределим мультиотношение (10) следующим образом. Введем сначала в него характеристическую функцию $E(l)$ множества E

$$E(l) = \begin{cases} 1, & l \in E \\ 0, & l \notin E. \end{cases}$$

Для этого достаточно переписать через $E(l)$ функцию $\rho(l)$. Именно, согласно (8), при $l = cd$

$$\rho(cd) = \sum_{a \in A} E(ac)E(ad). \qquad (19)$$

Как изменится вид этой функции, если мы учтем возможность появления в графе $G = (A, E)$ новых дуг? Согласно методу мультиотношений, численно эта возможность оценивается скалярным критерием $\sigma(l)$ (см. (14) или (15)). На его основе введем $\xi(l)$ – "функцию существования" дуги l, положив при $l = cd$

$$\xi(cd) = \begin{cases} 1, \text{ если } cd \in E \\ 0, \text{ если } c = d \\ \sigma(cd), \text{ если } cd \in \bar{E}. \end{cases} \qquad (20)$$

Дугу $l \in L(A)$, для которой $0 < \xi(l) < 1$, назовем *виртуальной*, если же $\xi(l) = 1$, то l – дуга в графе G.

Наличие виртуальных дуг приводит к изменению мультиотношения (10), которое мы запишем теперь в виде

$$\mu(l, l') = \Delta + p(l) - p(l'), \qquad (21)$$

где

$$p(cd) = \sum_{a \in A} \xi(ac)\xi(ad) \qquad (22)$$

является новым выражением для функции $\rho(cd)$ из (19), в которой вместо характеристической функции $E(l)$ реальных дуг графа G вставлена функция $\xi(l)$, допускающая существование виртуальных дуг.

Выражения (8), (10), (21) показывают, что в новой схеме функция ближайшего соседства $f_a(l)$ заменена функцией $g_a(l)$, принимающей при $l = cd$ значение

$$g_a(cd) = \xi(ac)\xi(ad)$$

и отличной от нуля для элементов c и d, не являющихся ближайшими соседями элемента a. Например, для части графа G, представленной на рис. 1, она равна $g_a(cd) = \sigma(ad)$, а для рис. 2 $g_a(cd) = \sigma(ac)\sigma(ad)$.

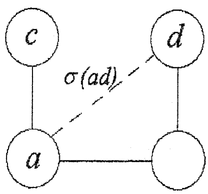

Рис. 1

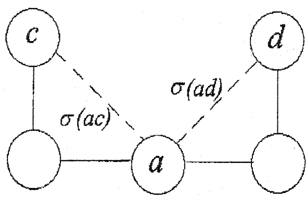

Рис. 2

Матрицу переходов $\tilde{T} = (\tilde{t}_{ll'})_{l,l' \in \bar{E}}$ для "мультиотношения" (21) определим аналогичным (12) способом,

$$\tilde{t}_{ll'} = \begin{cases} \mu(l', l), \text{ если } l' \neq l \\ \sum\limits_{l'' \in \bar{E} \setminus \{l\}} \mu(l, l''), \text{ если } l' = l. \end{cases} \tag{23}$$

Как и прежде, эта матрица имеет постоянную строчную сумму, равную $2\Delta(|\bar{E}) - 1)$. Значит, искомый скалярный критерий определяется из системы уравнений (15), в которой T заменено на $\tilde{T}$,

$$\sum_{l \in \bar{E}} \sigma(l) = 1, \quad \sigma\tilde{T} = 2\Delta(|\bar{E}| - 1)\sigma. \tag{24}$$

Однако, в отличие от (15), система уравнений (24) является нелинейной. Найдем приближенное решение этой системы, считая, что

$$\sigma(l) = \sigma_0(l) + \sigma_1(l), \tag{25}$$

где $\sigma_0(l)$ удовлетворяет системе (15), т.е.

$$\sum_{l \in \bar{E}} \sigma_0(l) = 1, \quad \sigma_0 T = 2\Delta(|\bar{E}) - 1)\sigma_0. \tag{26}$$

Слагаемое $\sigma_1(l)$ в (25) считаем малой добавкой к $\sigma_0(l)$. Для нее с точностью до линейных по $\sigma_0(l)$ членов (отбрасывая члены порядка $\sigma_1(l)\sigma_0(l)$, $\sigma_1(l)\sigma_0^2(l)$) из (24) вытекает следующая система уравнений

$$\sum_{l \in \bar{E}} \sigma_1(l) = 0, \quad \sigma_1 T = 2\Delta(|\bar{E}| - 1)\sigma_1 + \phi_0, \tag{27}$$

где ϕ_0 – некоторый вектор, зависящий от структуры графа G и поведения $\sigma_0(l)$, получаемого решением системы (26).

Утверждение 5. Система уравнений (27) имеет единственное решение.

Доказательство. Поскольку матрица T неразложима, то в силу теоремы Фробениуса-Перрона [6] у нее существует единственный (с точностью до умножения на константу) положительный собственный вектор, отвечающий максимальному собственному значению $\lambda_0 = 2\Delta(|\bar{E}| - 1)$. Отсюда вытекает, что ранг матрицы $T - \lambda_0 \cdot \mathbf{1}$ ($\mathbf{1}$ – единичная матрица) равен $(|\bar{E}| - 1)$. Значит, решение неоднородной системы уравнений

$$\sigma_1(T - \lambda_0 \cdot \mathbf{1}) = \phi_0$$

определяется с точностью до значения одной из компонент $\sigma_1(l)$. Величина этой компоненты находится из первого уравнения в (27). *Q.E.D.*

Из утверждения 5 вытекает, что скалярный критерий $\sigma(l)$ в (25) находится единственным образом и отличается от критерия $\rho(l)$, отвечающего за упорядочение по числу ближних соседей. Заметим, что величина добавки $\sigma_1(l)$ в этот критерий зависит от "виртуальных" связей между вершинами графа $G = (A, E)$ (см. рис. 1 и 2) и приводит, вообще говоря, к возможности образования дуг, не имеющих общих соседей в графе G.

Литература

1. Newman M.E.J. The Structure and Function of Complex Networks. SIAM Review, 45: 167-256, 2003.

2. Liben-Nowell D., J. Kleinberg. The Link Prediction Problem for Social Networks. Proceedings of the Twelfth Annual ACM International Conference on Information and Knowledge Management (CIKM'03), November 2003, pp. 556-559.

3. Левченков В.С. Аксиоматический подход к самосогласованному выбору. ДАН, 1993, т.330, N2, 173-176.

4. Левченков В.С. Два принципа рациональности в теории выбора: Борда против Кондорсе. - М.: Издательский отдел ф-та ВМиК МГУ, 2002.

5. Левченков В.С. Элементы эргодической теории с приложениями к проблемам выбора. II. Приложение эргодической теории к задачам выбора. - М.: ВМиК МГУ, 1997.

6. Гантмахер Ф.Р. Теория матриц. М.: Наука, 1967.