

УДК 519.83

ББК 22.18

МОДИФИЦИРОВАННОЕ ЗНАЧЕНИЕ МАЙЕРСОНА ДЛЯ ОПРЕДЕЛЕНИЯ ЦЕНТРАЛЬНОСТИ ВЕРШИН ГРАФА

ВЛАДИМИР В. МАЗАЛОВ*

Институт прикладных математических исследований

Карельский научный центр РАН

185910, Петрозаводск, ул. Пушкинская, 11

Школа математики и статистики, Университет Циндао,

Институт прикладной математики провинции Шаньдун

Циндао 266071, Китай

e-mail: vmazalov@krc.karelia.ru

ВИТАЛИЯ А. ХИТРАЯ

Институт математики и информационных технологий

Петрозаводский государственный университет

185910, Петрозаводск, пр. Ленина, 33

e-mail: dobvitalia@yandex.ru

Для анализа структуры социальных сетей могут быть использованы методы кооперативной теории игр. Одним из таких методов является вычисление значений Майерсона как меры центральности вершин в графе. В этом случае в качестве характеристической функции используется число путей определенной длины в подграфах, соответствующих коалициям. В работе предложена модификация вектора Майерсона для случая, когда в рассмотрение включаются пути в графе, содержащие циклы. Показана эффективность данного подхода на ряде примеров.

©2019 В.В. Мазалов, В.А. Хитрая

* Работа частично поддержана грантом провинции Шаньдун (Shandong Province) "Double-Hundred Talent Plan" (No. WST2017009).

Ключевые слова: графы, пути с циклами, меры центральности, кооперативная игра, значение Майерсона.

Поступила в редакцию: 25.03.19 *После доработки:* 16.04.19 *Принята к публикации:* 10.06.19

1. Введение

Социальные сети дали толчок развитию новых теоретико-графовых методов сетевого анализа. Методы анализа социальных сетей применяются во многих других областях, таких как: экономика, физика, биология и информационные технологии.

Одним из основных инструментов в анализе социальных сетей являются меры центральности, которые соответствуют тому, насколько хорошо рассматриваемая вершина расположена на путях, которые соединяют остальные вершины. Это важно и для приложений в транспортных задачах, а также в вопросах распространения информации в социальных сетях. Одним из первых определений центральности было понятие "betweenness centrality" [6], которое было связано с кратчайшими путями, которые проходят через рассматриваемую вершину. Однако, рассмотрение только кратчайших путей вызывает серьезные возражения, поскольку информация может распространяться и не обязательно по кратчайшему пути. Поэтому появились меры центральности, основанные на случайных блужданиях (PageRank)[5], и другие меры центральности, основанные на вычислении обратной матрицы Лапласа. Последние имеют хорошую аналогию с электрическими сетями и законами Кирхгофа [3,4]. В последнее время появились работы, связанные с теоретико-игровыми мерами центральности [7,9]. В частности, мерой центральности может служить вектор Майерсона в кооперативной игре, в которой характеристическая функция представляет собой число простых путей фиксированной длины. В работах [2] и [8] доказано, что вектор Майерсона можно вычислить через число простых путей, проходящих через данную вершину. Для графов в виде деревьев предложен простой алгоритм [8], основанный на производящих функциях. Однако в общем случае найти число всех путей определенной длины, проходящих через данную вершину, непростая задача. Кроме того, число всех простых путей конечно. Имеет смысл также ввести в рассмотрение и циклы. Тогда число таких путей произвольной длины будет

бесконечно. Это может дать дополнительную информацию о структуре графа.

В данной работе мы предлагаем простой алгоритм нахождения числа появлений рассматриваемой вершины в путях фиксированной длины, которые содержат также и циклы. Затем это выражение мы используем в определении меры центральности, где вместо вектора Майерсона выступает его модификация. На ряде примеров мы демонстрируем эффективность данного подхода.

2. Значение Майерсона как мера центральности

Для определения центральности вершин в графе можно использовать теоретико-игровой подход. Определим на графе $G = (N, E)$, где N – множество вершин и E – множество ребер, кооперативную игру $\Gamma = \langle N, v \rangle$, $|N| = n$. В этой игре вершины являются игроками, а характеристическая функция $v(K)$, $K \subset N$ определяется как количество простых путей длины m в этом подграфе. Число $m = 1, 2 \dots$ фиксировано. Очевидно, что функция $v(K)$ является монотонной, т.е. $v(K_1) \leq v(K_2)$, если $K_1 \subset K_2$. Тогда для ранжирования вершин в графе можно использовать решение кооперативной игры в виде вектора Шепли-Майерсона. Напомним определение вектора Шепли [1]. Обозначим $\sigma = (\sigma(1), \dots, \sigma(n))$ произвольную перестановку игроков $1, \dots, n$. Предположим, что все перестановки σ равновероятны, и тогда вероятность каждой перестановки равна $1/n!$.

Рассмотрим игрока i . Представим, что коалиция будет сформирована при его появлении. Обозначим $P_\sigma(i) = \{j \in N : \sigma^{-1}(j) < \sigma^{-1}(i)\}$ множество игроков, которые прибыли до его появления в перестановке σ . Вклад игрока i в этой коалиции равен

$$m_i(\sigma) = v(P_\sigma(i) \cup \{i\}) - v(P_\sigma(i)).$$

Определение 2.1. *Значение Шепли есть средний вклад каждого игрока во всех возможных перестановках*

$$\varphi_i(v) = \frac{1}{n!} \sum_{\sigma} m_i(\sigma) = \frac{1}{n!} \sum_{\sigma} [v(P_\sigma(i) \cup \{i\}) - v(P_\sigma(i))], i = 1, \dots, n.$$

Для игр, определенных на графах, Майерсон модифицировал вектор Шепли. Майерсон предположил, что естественно рассматривать

несвязные коалиции как множество связанных компонент. Для каждой такой компоненты K ее выигрыш равен $v(K)$, и именно этот выигрыш распределяется между членами коалиции K . Распределение выигрыша производится по схеме Шепли. Майерсон [9] ввел аксиомы, которые однозначно определяют такое распределение выигрыша.

В работе [8] на основе аксиом Майерсона доказана теорема 2.1 для деревьев и в работе [2] для произвольных графов. Мы приводим здесь другое доказательство этого утверждения.

Теорема 2.1. *Пусть дан граф $G = \langle N, E \rangle$. Характеристическую функцию $v(K)$, $K \subset N$ определим как количество простых путей длины t в этом подграфе. Тогда значение вектора Майерсона для игрока i есть*

$$\varphi_i = a(i)/(t + 1),$$

где $a(i)$ есть число простых путей длины t , проходящих через вершину i .

Доказательство. Рассмотрим все перестановки σ с ненулевым вкладом игрока i . Тогда в перестановке $P_\sigma(i) \cup \{i\}$ обязательно должны входить вершины из пути длины t , где вершина i входит последней. Тогда вклад игрока i равен $+1$.

Для зафиксированного пути ровно в $n!/(t + 1)$ перестановках вершина i идет после всех вершин этого пути. Следовательно, для вершины i верно, что сумма вкладов вершины i по всем перестановкам есть

$$\frac{a(i)}{n!} \frac{n!}{t + 1} = \frac{a(i)}{t + 1}.$$

Из аддитивности вектора Майерсона следует утверждение. \square

Следствие 2.1. *Пусть характеристическая функция $v(K)$, $K \subset N$ определена как*

$$v(K) = \sum a_m r^m,$$

где a_m – количество простых путей длины t в этом подграфе. Тогда значение вектора Майерсона для игрока i есть

$$\varphi_i = \sum \frac{a_m(i)}{m + 1} r^m,$$

где $a_m(i)$ есть число простых путей длины m , проходящих через вершину i .

Таким образом, в качестве меры центральности вершины i в графе можно положить значение вектора Майерсона $Y = (Y_1, \dots, Y_n)$, где $Y_i = \varphi_i, i = 1, \dots, n$. В работах [2,8] была показана эффективность такого представления для оценки центральности вершин графов различного вида. Интерпретация такого представления характеристической функции и соответствующей меры центральности следующая.

В данной кооперативной игре каждая прямая связь – путь длиной 1 – приносит игрокам доход r , где $0 \leq r \leq 1$. Кроме того, игроки также извлекают выгоду из косвенных (непрямых) связей, но уже меньшую. За каждый путь длины 2 коалиция получает r^2 , за путь длины 3 получает r^3 и т. д. Так как любые две вершины могут быть связаны несколькими путями разной длины, то при построении характеристической функции учитываются только кратчайшие. А если две вершины связаны несколькими путями одинаковой длины, то учитываются все такие пути, при условии, что между этими вершинами нет другого меньшего пути.

Таким образом, распределение выигрыша между вершинами характеризует насколько эффективно расположена вершина в графе с точки зрения распространения информации в этой сети.

Пример 1. Граф из пяти вершин.

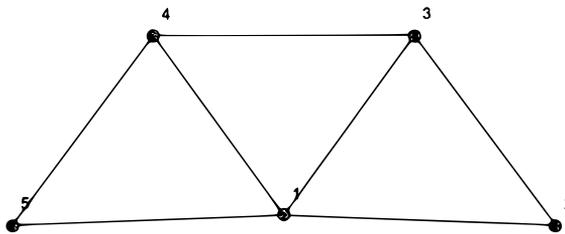


Рисунок 1. Граф из 5 вершин

Рассмотрим пути длины 3. Их всего 17:

1234, 1345, 1432, 1543, 2134, 2143, 2145, 2154, 2345,

2314, 2315, 3145, 3154, 3214, 3215, 3415, 4315.

Из них 16 содержат 1. Поэтому из Теоремы 2.1 следует

$$\varphi_1 = 16/4 = 4.$$

Если считать по определению через перестановки $\sigma = (\sigma(1), \sigma(2), \dots, \sigma(5))$, то игрок 1 получит +1, если войдет в коалицию четвертым или пятым.

В первом случае в половине случаев его положительный вклад будет в 6 перестановках и в другой половине в двух. Во втором случае, его вклад будет в 16 перестановках.

Таким образом,

$$Y_1 = \frac{1}{5} \left(\frac{1}{2}6 + \frac{1}{2}2 \right) + \frac{1}{5}16 = 4.$$

3. Модификация значения Майерсона

Определим центральность вершины как значение вектора Майерсона в кооперативной игре. Как показано выше, его значение можно вычислить как число появлений этой вершины во всех простых путях фиксированной длины k . Однако, вычисление этого числа непросто. В данном разделе мы приводим точное выражение для числа появлений каждой вершины в путях фиксированной длины, включая также и циклы.

Теорема 3.1. *Суммарное число появлений вершины i в путях длины k может быть вычислено по формуле*

$$s_i(k) = \sum_{j=1}^n a_{ij}^{(k)} + \sum_{l=1}^n \left[a_{li} \sum_{j=1}^n a_{ij}^{(k-1)} + a_{li}^{(2)} \sum_{j=1}^n a_{ij}^{(k-2)} + \dots + a_{li}^{(k)} \right], \quad (3.1)$$

где $A^{(k)}$ – матрица смежности графа, возведенная в степень k .

Доказательство. Покажем вначале, что число появлений вершины i в путях длины k , начинающихся в вершине l , включая циклы, $s_i^{(l)}(k)$ имеет вид:

$$s_i^{(l)}(k) = \delta_{\{l=i\}} \sum_{j=1}^n a_{lj}^{(k)} + a_{li} \sum_{j=1}^n a_{ij}^{(k-1)} + a_{li}^{(2)} \sum_{j=1}^n a_{ij}^{(k-2)} + \dots + a_{li}^{(k)}, \quad (3.2)$$

где $\delta_{\{l=i\}}$ – индикатор события $\{l = i\}$.

Рассмотрим вершину i и все пути длины k в графе G . Если путь начинается в вершине i , то вершина i встречается в этом пути не менее одного раза. Число путей длины k , начинающихся в вершине i можно вычислить с помощью матрицы A^k , сложив все элементы в строке i . Это соответствует первому слагаемому в формуле (3.2).

Вершина i может встретиться в пути длины k несколько раз. Так если вершина i встретилаь в пути длины k на следующем шаге, то она будет встречаться хотя бы раз во всех путях длины $k - 1$. Это соответствует второму слагаемому в формуле (3.2). Рассуждая по индукции, получаем, что если вершина i встретилаь в пути длины k на шаге m , $m = 1, \dots, k$, то число ее появлений во всех путях далее будет равно числу путей длины m из начальной вершины до вершины i (это $a_{ii}^{(m)}$), умноженному на число путей длины $k - m$ из вершины i (это число $\sum_{j=1}^n a_{ij}^{(k-m)}$).

Складывая (3.2) по всем l , получим выражение (3.1). □

Пример 2. Справедливость формулы для подсчета $s_i^{(l)}(k)$ проиллюстрируем с помощью примера.

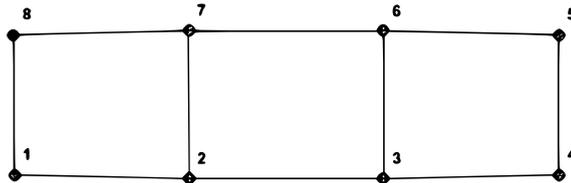


Рисунок 2. Граф из 8 вершин

Рассмотрим граф G из 8 вершин (рис. 2) с матрицей смежности A :

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Будем искать все возможные пути длины 4, проходящие через вершину 1. На рис. 3 изображено дерево путей, начинающихся в вершине 1.

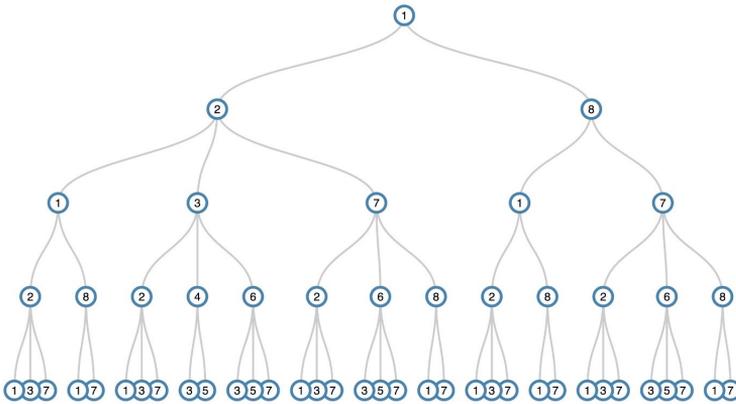


Рисунок 3. Дерево путей, начинающихся в вершине 1

Поскольку все пути, начинающиеся в вершине 1, содержат эту вершину хотя бы один раз, необходимо вычислить общее число путей длины 4, начинающихся в вершине 1. Эту величину можно вычислить с помощью суммы элементов первой строки матрицы

$$A^4 = \begin{pmatrix} 9 & 0 & 9 & 0 & 4 & 0 & 12 & 0 \\ 0 & 18 & 0 & 9 & 0 & 16 & 0 & 12 \\ 9 & 0 & 18 & 0 & 12 & 0 & 16 & 0 \\ 0 & 9 & 0 & 9 & 0 & 12 & 0 & 4 \\ 4 & 0 & 12 & 0 & 9 & 0 & 9 & 0 \\ 0 & 16 & 0 & 12 & 0 & 18 & 0 & 9 \\ 12 & 0 & 16 & 0 & 9 & 0 & 18 & 0 \\ 0 & 12 & 0 & 4 & 0 & 9 & 0 & 9 \end{pmatrix}.$$

Тогда $\sum_{j=1}^8 a_{1j}^{(4)} = 34$. На следующем шаге могут встретиться вершины 2 и 8. На втором шаге опять может встретиться вершина 1, таких путей два. После этого остается еще два шага. Количество путей длины 2, начинающихся из вершины 1 равно $\sum_{j=1}^8 a_{1j}^{(4-2)} = 5$. Таким образом, вершина 1 встретится еще $2 \cdot 5 = 10$ раз. Наконец, вершина 1 может встретиться в конце пути длины 4. Число таких путей можно найти как $a_{1,1}^{(4)} = 9$.

Складывая по формуле (3.1), получим $34 + 10 + 9 = 53$. Именно столько раз встретится вершина 1 в путях длины 4, начинающихся из вершины 1.

Пример 3. Пусть матрица смежности для графа G из 4 вершин имеет вид

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Выпишем все пути длины 2, содержащиеся в данном графе:

121, 123, 124, 212, 232, 234, 242, 243, 321,

323, 324, 342, 343, 421, 423, 424, 423, 434.

Вычислим количество появлений вершин в путях длины 2. Очевидно, что $s_1(2) = 7, s_2(2) = 19, s_3(3) = 14, s_4(2) = 14$.

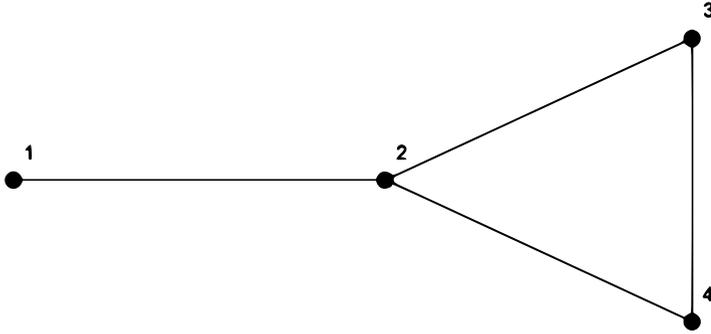


Рисунок 4. Граф из 4 вершин

Воспользуемся приведенной формулой. Для этого необходим квадрат матрицы смежности.

$$A^2 = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 3 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}.$$

$$s_1(2) = \sum_{j=1}^4 a_{1j}^{(2)} + \sum_{l=1}^4 \left[a_{l1} \sum_{j=1}^4 a_{lj} + a_{l1}^{(2)} \right] = 3 + [1 + 1 \cdot 1 + 1 + 1] = 7,$$

$$s_2(2) = \sum_{j=1}^4 a_{2j}^{(2)} + \sum_{l=1}^4 \left[a_{l2} \sum_{j=1}^4 a_{lj} + a_{l2}^{(2)} \right] = 19,$$

$$s_3(2) = \sum_{j=1}^4 a_{3j}^{(2)} + \sum_{l=1}^4 \left[a_{l3} \sum_{j=1}^4 a_{lj} + a_{l3}^{(2)} \right] = 14,$$

$$s_4(2) = \sum_{j=1}^4 a_{4j}^{(2)} + \sum_{l=1}^4 \left[a_{l4} \sum_{j=1}^4 a_{lj} + a_{l4}^{(2)} \right] = 14.$$

Вершина под номером 1 является висячей, соответственно она реже появляется в путях, что подтверждено наименьшим значением $s_1(2)$. Вершина 2 связывает висячую вершину с петлей, что предполагает ее появление в большом количестве путей и, следовательно, наибольшее значение $s_2(2)$. Вершины 3 и 4 расположены симметрично, для них значения $s_3(2)$ и $s_4(2)$ совпадают.

Теорема 3.2. Число путей длины k (включая циклы) R_k в графе есть

$$\frac{\sum_{i=1}^n s_i(k)}{k+1},$$

где $s_i(k)$ – суммарное число появлений вершины i в путях длины k , n – число вершин в графе.

Доказательство.

$$\begin{aligned} \sum_{i=1}^n s_i(k) &= \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij}^{(k)} + \sum_{l=1}^n \left[a_{li} \sum_{j=1}^n a_{ij}^{(k-1)} + a_{li}^{(2)} \sum_{j=1}^n a_{ij}^{(k-2)} + \dots + a_{li}^{(k)} \right] \right) = \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij}^{(k)} + \sum_{i=1}^n \sum_{j=1}^n \left(a_{li} \sum_{j=1}^n a_{ij}^{(k-1)} + a_{li}^{(2)} \sum_{j=1}^n a_{ij}^{(k-2)} + \dots + a_{li}^{(k)} \right) = \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij}^{(k)} + \sum_{i=1}^n \sum_{l=1}^n a_{li}^{(k)} + \\ &+ \sum_{i=1}^n \sum_{l=1}^n \left(a_{li} \sum_{j=1}^n a_{ij}^{(k-1)} + a_{li}^{(2)} \sum_{j=1}^n a_{ij}^{(k-2)} + \dots + a_{li}^{(k-1)} \sum_{j=1}^n a_{ij} \right). \end{aligned}$$

Здесь $\sum_{i=1}^n \sum_{j=1}^n a_{ij}^{(k)}$ – сумма элементов матрицы $A^{(k)}$, т.е. число путей длины k , включая циклы, в графе G . Тогда полученное выражение можно переписать как

$$2R_k + \sum_{i=1}^n \sum_{l=1}^n \left(a_{li} \sum_{j=1}^n a_{ij}^{(k-1)} + a_{li}^{(2)} \sum_{j=1}^n a_{ij}^{(k-2)} + \dots + a_{li}^{(k-1)} \sum_{j=1}^n a_{ij} \right).$$

Поскольку для степеней матриц выполняется свойство $A^m \cdot A^n = A^{m+n}$, выражение в скобках представляет собой сумму $(k-1)$ слагаемых, равных R_k , т.е.

$$\sum_{i=1}^n s_i(k) = (k+1) \cdot R_k.$$

□

Таким образом, если в качестве выигрыша игроков рассматривать число путей длины k в графе, включая и циклы, т.е. величину

$$\frac{\sum_{i=1}^n s_i(k)}{k+1},$$

то естественно распределить выигрыш между игроками в виде

$$\left(\frac{s_1(k)}{k+1}, \frac{s_2(k)}{k+1}, \dots, \frac{s_n(k)}{k+1} \right).$$

Это похоже на распределение Майерсона в Теореме 2.1, поэтому данный способ распределения будем называть модифицированным значением Майерсона.

Определение 3.1. *Центральностью k -го порядка вершины i будем называть число появлений вершины i в путях длины k , включая циклы.*

Введем вектор $\sigma(k)$ – вектор центральностей вершин графа G , i -ая компонента которого равна $\sigma_i(k) = \frac{s_i(k)}{k+1}$, $i = 1, \dots, n$. Для графа из примера 3 запишем вектора центральности для путей длины 2, 3, 4 и 5:

$$\sigma(2) = \{2.3; 6.3; 4.67; 4.67\},$$

$$\sigma(3) = \{4; 14; 10; 10\},$$

$$\sigma(4) = \{9; 30.6; 22.2; 22.2\},$$

$$\sigma(5) = \{17.67; 66.3; 48; 48\}.$$

Полученные результаты можно сравнить с векторами Майерсона $\varphi(k)$, где i -ая компонента – вклад игрока – вершины i в выигрыш гранд-коалиции, в случае когда характеристическая функция представляет собой число путей длины k в коалиции [2]

$$\varphi(2) = \{2.3; 6.3; 4.67; 4.67\},$$

$$\varphi(3) = \{4.67; 12.67; 10.3; 10.3\},$$

$$\varphi(4) = \{10.8; 26.8; 23.17; 23.17\},$$

$$\varphi(5) = \{23.5; 55.5; 50.5; 50.5\}.$$

Видно, что полученные значения различаются незначительно (а для путей длины 2 и вовсе совпадают), при этом порядок вершин сохраняется, что позволяет использовать полученный вектор для оценки центральностей вершин, не вычисляя количество путей в графе, как этого требует оценка с помощью вектора Майерсона. Кроме того, стандартный подход к его вычислению предполагает довольно трудоемкий процесс перебора всех возможных коалиций.

4. Частные случаи

4.1. Пусть граф G – звезда из n вершин, в центре звезды расположен игрок 1 (рис. 5).

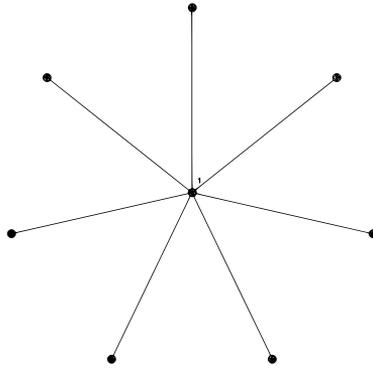


Рисунок 5. Звезда

Тогда матрица смежности примет вид

$$A = \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Для того, чтобы воспользоваться формулой (3.1), требуется возведение матрицы A в степень k . В зависимости от четности k получим два вида матриц. Возводя матрицу в четную степень, имеем:

$$A^k = \begin{pmatrix} (n-1)^{\frac{k}{2}} & 0 & 0 & \dots & 0 \\ 0 & (n-1)^{\frac{k-2}{2}} & (n-1)^{\frac{k-2}{2}} & \dots & (n-1)^{\frac{k-2}{2}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & (n-1)^{\frac{k-2}{2}} & (n-1)^{\frac{k-2}{2}} & \dots & (n-1)^{\frac{k-2}{2}} \end{pmatrix}.$$

В случае нечетной степени матрица примет вид:

$$A^k = \begin{pmatrix} 0 & (n-1)^{\frac{k-1}{2}} & (n-1)^{\frac{k-1}{2}} & \dots & (n-1)^{\frac{k-1}{2}} \\ (n-1)^{\frac{k-1}{2}} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (n-1)^{\frac{k-1}{2}} & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Применив формулу (3.1), получим выражения для центральных вершин графа G . Для четных значений k :

$$\begin{aligned} s_1(k) &= (n-1)^{\frac{k}{2}} \left(\frac{k}{2}n + 1 \right), \\ s_i(k) &= (n-1)^{\frac{k-2}{2}} \left(\frac{k+2}{2}n - 1 \right), \end{aligned}$$

для нечетных значений:

$$\begin{aligned} s_1(k) &= (k+1)(n-1)^{\frac{k+1}{2}}, \\ s_i(k) &= (k+1)(n-1)^{\frac{k-1}{2}}. \end{aligned}$$

Стоит отметить, что для нечетных длин путей выполняется соотношение

$$\frac{s_1(k)}{s_i(k)} = n - 1,$$

так же, как и для вектора Майерсона.

4.2. Цепь. Пусть граф G – цепь (путь) из n вершин. Соответствующая матрица смежности A :

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

В том случае, если длина пути k меньше $\frac{n}{2}$, начиная с $(k+1)$ -ой вершины в цепи и вплоть до $(n-k)$ -ой устанавливается стабильное

значение центральности порядка k . Для вычисления этого значения можно воспользоваться формулой

$$s_i(k) = 2^k(k + 1), \quad i = k + 1 \dots n - k.$$

5. Вычислительные эксперименты

5.1. Муравьиная колония

В качестве исходных были использованы данные швейцарских биологов, полученные в ходе исследования [11]. Изучалось поведение муравьиной колонии в случае, когда в колонию попадают зараженные особи. Рассматриваемая колония состоит из 105 муравьев.

Группа ученых фиксировала время контакта между парами муравьев до и после заражения, в результате чего были получены графы взаимодействий, где вершины - это помеченные муравьи, а ребра между вершинами имеют веса равные суммарному времени контакта в течение всего эксперимента. Далее, поскольку метод вычисления вектора центральностей предполагает работу с матрицей инцидентности, было выбрано пороговое значение времени взаимодействия, равное $0,15 \cdot T_{max}$, где T_{max} - максимальное время взаимодействия между муравьями в колонии. Все ребра, веса которых оказались меньше данного порогового значения в матрице инцидентности были помечены 0, остальные 1. Полученные графы изображены на рис.6, рис.7. Для всех вершин этих графов вычислены значения центральностей.

На графах круглые вершины соответствуют муравьям «nurses» (n), треугольные - муравьям «foragers» (f), с помощью звезды отмечена матка колонии («queen» (q)). Вершины, отмеченные квадратами, соответствуют особям, которые подвергнутся заражению. Стоит отметить, что заражению подвергались только муравьи «foragers».

В табл. 1 представлены значения мер центральности 5 порядка для некоторых вершин графа до заражения. Наибольшее значение центральности наблюдается у вершины 102, данная вершина соответствует матке колонии. Также интерес представляют значения мер центральности для вершин, которые будут заражены, их значения представлены в табл. 2.

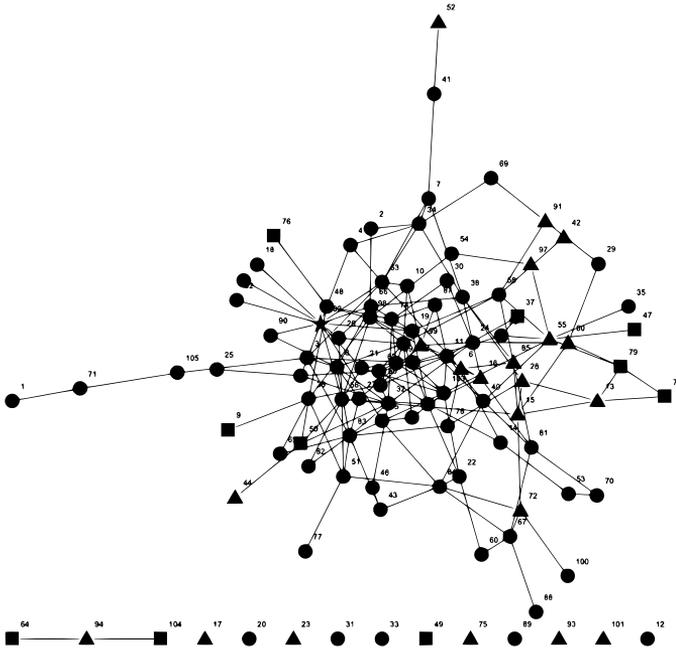


Рисунок 6. Граф взаимодействий до заражения

Таблица 1.

i	102 (q)	32 (n)	59 (n)	66 (n)	83 (n)	6(f)	74 (f)
$\sigma_i^{pre}(5)$	19919,3	15484,6	9289	6727,3	12944	5424	3743,6

Таблица 2.

i	8	36	46	48	49	63	72	75	78	103
$\sigma_i^{pre}(5)$	396,6	934,6	397,6	0	3979	4	423,6	237,3	1442,6	4

После заражения структура графа заметно меняется. Здоровые муравьи минимизируют общение с зараженными муравьями. При этом изолированные зараженные муравьи образуют свою структуру, в которой у некоторых вершин значения мер центральности возрастают, что может быть связано с тем, что они становятся лидерами в своем сообществе.

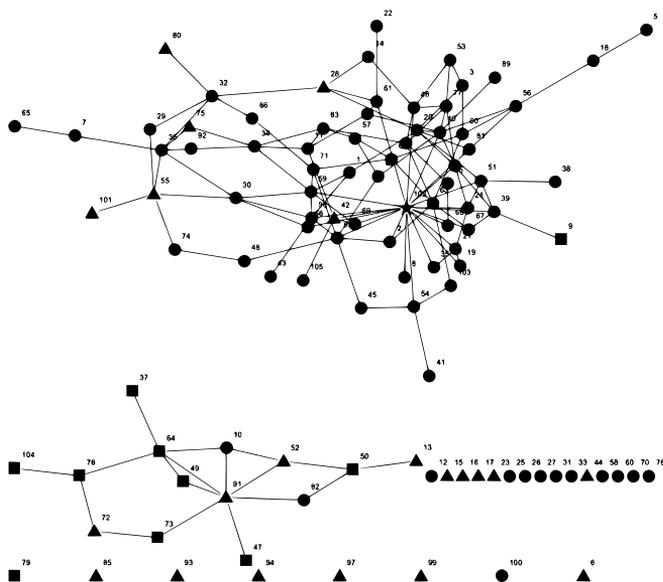


Рисунок 7. Граф взаимодействий после заражения

Меры центральности вершин 5 порядка для графа после заражения представлены в табл. 3 и 4.

Таблица 3.

i	102 (q)	32 (n)	59 (n)	66 (n)	83 (n)	6 (f)	74 (f)
$\sigma_i^{post}(5)$	61601,6	1159,3	11511,6	5301,3	1425,3	0	334,3

Вершина - матка и в этом случае обладает наибольшей мерой центральности.

Таблица 4.

i	8	36	46	48	49	63	72	75	78	103
$\sigma_i^{pre}(5)$	396,6	934,6	397,6	0	3979	4	423,6	237,3	1442,6	4
$\sigma_i^{post}(5)$	615,3	113,6	150,3	378	217,3	882	231,6	238,3	0	43,6

5.2. Метрополитен

В качестве объекта вычислительного эксперимента использовался граф, построенный на основе схемы Санкт-Петербургского метрополитена (рис. 8). Оценим меры центральности вершин в этом графе для путей длины 5 и 10.

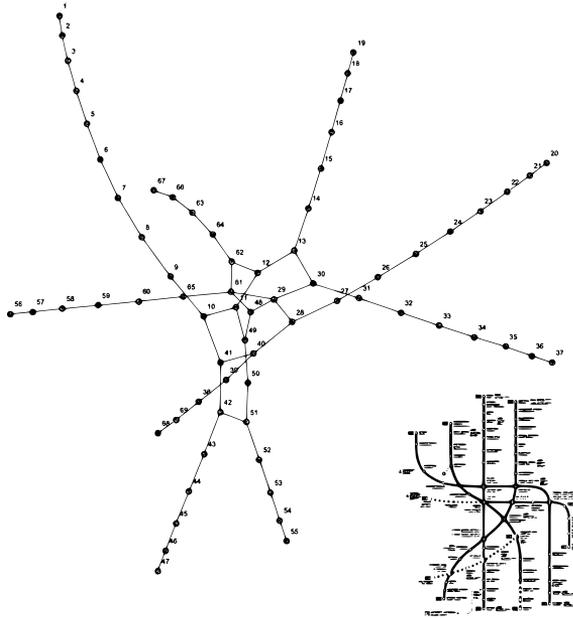


Рисунок 8. Граф Петербургского метрополитена

Наименьшие значения мер центральности наблюдаются у станций Девяткино, пр. Ветеранов, Парнас, Купчино, Рыбацкое, ул. Дыбенко, Комендантский проспект, Международная, Беговая. Все эти станции являются конечными.

Большие значения мер центральности у пересадочных станций, наибольшие у станций Садовая, Спасская, Сенная площадь. Согласно статистике Петербургского метрополитена [<http://kommet.ru/stats>] на данных станциях наблюдается наибольший ежедневный и ежемесячных пассажиропоток. Хотя, согласно этой же статистике, достаточно большой пассажиропоток и на конечных станциях спальных районов.

Таблица 5. Значения центральности

Станция	$s_i(5)$	$\sigma_i(5)$	$s_i(10)$	$\sigma_i(10)$
Девяткино	44	7,33	1 588	144,36
Проспект Ветеранов	44	7,33	1 784	162,18
Парнас	44	7,33	1 598	145,27
Купчино	44	7,33	1648	149,82
Рыбацкое	44	7,33	2 154	195,82
Комендантский проспект	44	7,33	1 992	181,09
Улица Дыбенко	46	7,67	2 927	266,09
Международная	46	7,67	4 281	389,18
Беговая	46	7,67	3 563	323,91
Спасская	1 988	331,33	964 437	87 676,09
Садовая	2 578	429,67	1 224 429	111 311,73
Сенная площадь	2 810	468,33	1 342 204	122 018,55

6. Заключение

В работе предложено новое определение мер центральности вершин в графе, основанное на модификации значения Майерсона в кооперативной игре, где в качестве характеристической функции выступает число простых путей определенной длины. Ранее было показано, что значение Майерсона для данной вершины в такой игре можно найти как число появлений этой вершины в простых путях определенной длины. При этом, циклы исключались. В данной работе предлагается модифицировать меру центральности, включив в рассмотрение также и циклы. Найдено простое аналитическое выражение для такой меры центральности. Разработан алгоритм, реализующий данный метод. Эффективность метода проиллюстрирована с помощью ряда примеров.

СПИСОК ЛИТЕРАТУРЫ

1. Aumann R., Myerson R. *Endogenous formation of links between players and coalitions: an application of the Shapley value*. In: *The Shapley value*, Cambridge University Press, 1988. P. 175–191.

2. Avrachenkov K., Kondratev A.Yu., Mazalov V.V. and Rubanov D.G. *Network partitioning as cooperative games* // Computational social networks. 2018. V. 5. No. 11. P. 1–28.
3. Avrachenkov K.E., Mazalov V.V., Tsynguev B.T. *Beta Current Flow Centrality for Weighted Networks* // Proceedings of CSoNET. 2015. LNCS. V. 9197. P. 216–227.
4. Brandes U., Fleischer D. *Centrality measures based on current flow* // Proceedings of the 22nd annual conference on Theoretical Aspects of Computer Science. 2005. P. 533–544.
5. Brin S., Page L. *The anatomy of a large-scale hypertextual Web search engine* // Computer Networks and ISDN Systems. 1998. V. 30, no. 17. P. 107–117.
6. Freeman L.C. *A set of measures of centrality based on betweenness* // Sociometry. 1977. V. 40. P.35–41.
7. Mazalov V., Chirkova J. *Networking games*. Academic Press, 2019.
8. Mazalov V.V., Trukhina L.I. *Generating functions and the Myerson vector in communication networks* // Disc. Math. and Appl. 2014. V. 24(5). P. 295–303.
9. Michalak T.P., Aadithya K.V., Szczepanski P.L., Ravindran B., Jennings N.R. *Efficient computation of the Shapley value for game-theoretic network centrality* // J Artif Intell Res. 2013. V. 46. P. 607–650.
10. Myerson R.B. *Graphs and cooperation in games* // Math. Oper. Res. 1977. V. 2. P. 225–229.
11. Stroeymeyt N., Grasse A.V., Crespi A., Mersch D.P., Cremer S., and Keller L. *Social network plasticity decreases disease transmission in a eusocial insect* // Science. 2006. V. 362. P. 941–945.

MODIFIED MAYERSON VALUE FOR DETERMINING
THE CENTRALITY OF GRAPH VERTICES

Vladimir V. Mazalov, Institute of Applied Mathematical Research of Karelian Research Centre of RAS, School of Mathematics and Statistics, Qingdao University, Institute of Applied Mathematics of Shandong, Dr.Sc., professor (vmazalov@krc.karelia.ru),

Vitaliya A. Khitraya, Petrozavodsk State University (dobvitalia@yandex.ru).

Abstract: To analyze the structure of social networks, we can use methods of cooperative game theory. One of such methods is based on the calculation of the Myerson values as a measure of the centrality of the vertices in the graph. In this case, the number of paths of a certain length in the subgraphs corresponding to the coalitions is used as the characteristic function. The paper proposes a modification of the Myerson value for the case when the paths in the graph containing cycles are included in the consideration. The effectiveness of this approach is shown on a number of examples.

Keywords: networks, paths with cycles, centrality measure, cooperative game, Myerson value.