

СНИЖЕНИЕ РАЗМЕРНОСТИ ЗАДАЧИ НАХОЖДЕНИЯ КРИТИЧЕСКИХ УЗЛОВ СЕТИ

Крыгин А. А.¹, Тарасова С. М.²
(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Одним из классов задач, решаемых при оценке устойчивости инженерной сети, являются задачи нахождения критических узлов. В ряде постановок эта задача формулируется как нахождение такого подмножества узлов заданной мощности (критических узлов), при выходе из строя которых всей сети будет нанесен максимальный ущерб. И наиболее частый способ оценки ущерба в такой постановке – определение количества связных пар узлов в сети с исключенными критическими узлами. Для таких узлов, которые соответствуют минимуму количества связных пар, требуются проведение дополнительных мер по повышению надежности и безопасности. Ряд методов решения задачи нахождения критических узлов использует сведение ее к эквивалентной задаче линейного программирования. Основной проблемой этого подхода является большая размерность задачи, и, как следствие, высокая вычислительная сложность ее решения. В работе проводится исследование различных характеристик вершин графовой модели сети, анализ значений которых позволит заранее установить факт принадлежности вершины к подмножеству критических или наоборот, к подмножеству некритических узлов. Благодаря этому можно сформировать дополнительные ограничения, снижающие размерность задачи линейного программирования и ее вычислительную сложность, что позволит находить критические узлы в инженерных сетях с большим количеством объектов за приемлемое время. В процессе исследования было решено множество различных подзадач, поэтому в работе описывается первая, подготовительная его часть.

Ключевые слова: устойчивость электрических сетей, поиск критических узлов, графовые модели сетей.

1. Введение

Наиболее простой математической моделью сети является ее представление в виде неориентированного и невзвешенного графа. Такая модель применима практически ко всем видам сетей и является базовой (начальной) в задачах оценки их устойчивости.

¹ Андрей Александрович Крыгин, к.т.н. (andreyakr@yandex.ru).

² Софья Михайловна Тарасова, техник (tarasva_sofia@mail.ru).

Объектам сети соответствуют вершины графа, а наличие какого-либо отношения (в зависимости от вида сети) между объектами соответствует ребрам. При этом все вершины считаются равноправными, отдельно не выделяются, например, вершины – источники ресурса и вершины-потребители.

Одним из классов задач, решаемых при оценке устойчивости сети, являются задачи нахождения критических ребер и узлов (Critical Link Disruptor (CLD) and Critical Node Disruptor (CND) optimization problems). Постановка этих задач заключается в нахождении такого подмножества узлов (или ребер), при выходе из строя которых всей сети будет нанесен максимальный ущерб.

В общем виде задачу нахождения критических узлов для графовой модели сети можно сформулировать так. Заданы связный, неориентированный граф без петель и кратных ребер $G(V, E)$, $|V| = n$, $|E| = m$, количество удаляемых вершин q и функция-критерий $D(G)$, областью определения которой является множество графов, а областью значений – вещественные числа. Обозначим W – подмножество V , $|W| \leq q$ и G_W – подграф графа G , получившийся при удалении из V подмножества W и инцидентных ребер. Тогда требуется найти такое подмножество W' , при котором $D(G_{W'})$ достигает своего минимума (или максимума, в зависимости от функции-критерия).

Чаще всего в качестве критерия $D(G)$ выбирают одну из трех функций:

- количество компонент связности (определяется максимум);
- размер наибольшей компоненты связности (определяется минимум);
- количество связных пар вершин (определяется минимум).

Наиболее часто в работах применяется третий критерий – количество связных пар вершин, который также используется и в данной работе.

Существуют и другие постановки задачи поиска критических узлов, например, задача нахождения критических узлов с учетом дистанции (Distance-Based Critical Node Problem)

[10, 11, 16, 21] и β -вершинная задача нахождения критических узлов (β -vertex disruptor problem) [9, 12, 17].

В качестве критерия $D(G)$ в постановке задачи нахождения критических узлов с учетом дистанции выбрано количество связанных пар вершин, расстояние между которыми не превосходит заданного параметра k . То есть минимизируется количество пар вершин, для которых в графе с удаленными q вершинами существует путь, содержащий не более k ребер. Во многих работах, посвященных этой задаче, используются различные методы сведения ее к эквивалентным задачам целочисленного линейного программирования.

В некотором смысле β -вершинную задачу нахождения критических узлов можно назвать обратной. Если граф сети после удаления q вершин остается связным, то количество связанных пар в нем равно $\frac{(n-q)(n-q-1)}{2}$. Обозначим M – количество связанных пар вершин в таком графе (связном или несвязном) и $\beta = \frac{2M}{(n-q)(n-q-1)}$. Тогда задачу можно сформулировать так: необходимо найти минимальное по мощности подмножество вершин, при удалении которых из графа $\beta < \beta_0$, где β_0 – заданный параметр. Эта задача тесно связана с «прямой» задачей; переход от одной задачи к другой в ее эквивалентной формулировке задачи линейного программирования, можно осуществить, поменяв местами целевую функцию и одно из ограничений.

Также активно исследуются задачи и для других критериев $D(G)$, например, в [13] рассматривается размер наибольшей компоненты связности.

1.1. Точные методы решения задачи нахождения критических узлов по критерию наименьшего количества связанных пар вершин

В [19] была доказана NP-полнота задачи нахождения критических узлов. На сегодняшний день наиболее эффективным методом нахождения строго оптимального (точного) решения задачи поиска критических узлов является сведение ее к эквивалентной

задаче линейного программирования (ЛП). За исключением метода полного перебора, каких-либо других точных методов решения этой задачи, включая алгоритмы на графах и сетях, не известно. Более того, в [18] доказана теорема о том, что нахождение приближенного решения, достаточно близкого к точному также является NP-трудной задачей, задана метрика «близости» и указана ее нижняя граница.

Рассмотрим постановку задачи ЛП для критерия минимального количества связных пар. Введем следующие булевские переменные:

$u_{i,j}$ – переменная, показывающая связаны ли вершины i и j . $u_{i,j} = 0$, если вершины i и j не связаны, $u_{i,j} = 1$, если вершины i и j связаны;

v_i – переменная, показывающая будет ли вершина i в оптимальном решении удалена из графа. $v_i = 0$, если вершина i не удаляется из графа, $v_i = 1$, если вершина i удаляется из графа.

Тогда постановку задачи в терминах ЛП можно задать так:

$$\begin{aligned} \min \sum_{i,j \in V} u_{i,j} \\ u_{i,j} + v_i + v_j \geq 1, \forall (i,j) \in E \\ u_{i,j} + u_{j,w} - u_{w,i} \leq 1, \forall (i,j,w) \in V \\ u_{i,j} - u_{j,w} + u_{w,i} \leq 1, \forall (i,j,w) \in V \\ -u_{i,j} + u_{j,w} + u_{w,i} \leq 1, \forall (i,j,w) \in V \\ \sum_{i \in V} v_i \leq q. \end{aligned}$$

Целевая функция $\sum_{i,j \in V} u_{i,j}$ равна количеству всех связных пар вершин и она минимизируется.

Первая группа неравенств $u_{i,j} + v_i + v_j \geq 1$, называемая в литературе «неравенством треугольника», записывается для каждого ребра и задает структуру графа. Если соединенные ребром вершины i и j не удалены, то они всегда связны по ребру. Это утверждение эквивалентно вышеприведенному неравенству.

Для трех произвольных вершин i , j и w справедливо утверждение: если i связана с j , и j связана с w , то i связана с w . Это утверждение эквивалентно выражению: $u_{i,j} + u_{j,w} + u_{w,i} \neq 2$. Систему неравенств, эквивалентную этому выражению можно задать так:

$$\begin{aligned}u_{i,j} + u_{j,w} - u_{w,i} &\leq 1, \\u_{i,j} - u_{j,w} + u_{w,i} &\leq 1, \\-u_{i,j} + u_{j,w} + u_{w,i} &\leq 1.\end{aligned}$$

Эта группа неравенств записывается для всех троек вершин.

Последнее неравенство $\sum_{i \in V} v_i \leq q$ задает общее количество удаляемых вершин.

В [22] доказана теорема о том, что исходная постановка задачи и ее формулировка в виде задачи ЛП эквивалентны. В [17] показаны методы сведения задачи нахождения критических узлов к задаче ЛП для оставшихся двух критериев (количество компонент и размер максимальной компоненты) и доказана эквивалентность исходной задачи и задачи ЛП.

В большинстве случаев у задачи нахождения критических узлов существует несколько оптимальных решений. Если требуется найти все решения (или заданное количество решений), то поступают следующим образом. Будем обозначать буквой Q (с индексом или без) одно из решений задачи, т.е. подмножество вершин, мощность которого не превосходит q такое, что не существует другого подмножества R , $|R| \leq q$ и $D(G_Q) > D(G_R)$. Находят первое решение Q_1 и фиксируют значение функции-критерия $D_{opt} = D(G_{Q_1})$. После этого в систему неравенств добавляют неравенство, исключающее найденное решение: $\sum_{i \in Q_1} v_i < q$ и повторно запускают решение задачи ЛП. Если на k -й итерации значение $D(G_{Q_k})$ станет больше D_{opt} , то это свидетельствует о том, что все оптимальные решения найдены.

1.2. Методы, близкие к точным

Из постановки задачи ЛП видно, что наибольшую долю неравенств составляют тройки $u_{i,j} + u_{j,w} + u_{w,i} \neq 2$, число этих

троек равно $C_n^3 = O(n^3)$. И, например, для графа с 120-ю вершинами размер файла описания задачи ЛП занимает около 4-х Гб. За 2 недели расчета (процессор Intel Xeon E5-2620v4@2.1ghz, оперативная память 64 Гб, программа решения задачи ЛП GLPK) не было найдено ни одного решения для такой задачи. Использование вычислительных кластеров позволяет решать задачу ЛП за приемлемое время (несколько часов) для графов с $n < 200$ [19]. Для определения критических узлов у графов бóльшей размерности или для решения задачи за меньшее время появились эвристические методы поиска решений, близких к точным.

О количестве работ в этой области говорит тот факт, что появились работы, каталогизирующие, сопоставляющие и сравнивающие между собой уже имеющиеся методы. В [14] рассматривается несколько десятков методов решения задачи нахождения критических узлов в различных постановках. Обзорная работа [20] посвящена приближенным методам решения задачи для вышеописанных трех основных критериев.

В рассмотренных работах одна из основных схем решения задачи методами, близкими к точным, включает в себя два этапа. На первом этапе предлагается алгоритм нахождения «первоначального решения» из q вершин. Первоначальное решение – это подмножество из q вершин, которое не обязательно является точным решением, будем обозначать первоначальные и текущие решения буквой S . Далее это решение «обогащают» методом парного обмена.

В данном случае функция-критерий $D(G_S) = \sum_{i,j \in V \setminus S} u_{i,j}$, ищется минимум. Пусть S – текущее решение, в начале работы инициализируем S как первоначальное решение. Алгоритм парного обмена заключается в следующем:

1. Выбирается произвольная вершина $i \in S$.
2. Среди всех вершин из $V \setminus S$ определяется вершина j такая, что значение D будет минимальным, если в решении S вместо вершины i брать вершину j и $D(G_{S \cup \{j\} \setminus \{i\}}) < D(G_S)$. Если такой вершины нет, то переход к п.1.
3. Из S исключается i и добавляется j .

4. Выбирается следующая вершина из S .

5. Переход к п.2.

6. Алгоритм заканчивает работу по истечении заданного времени или когда для любой вершины $i \in S$ и любой вершины $j \in V \setminus S$ выполняется $D(G_{S \cup \{j\} \setminus \{i\}}) > D(G_S)$.

С помощью таких методов удалось добиться того, что в 40% случаев полученное решение совпадало с точным.

Скорость работы алгоритма и близость полученного решения к точному напрямую зависят от близости первоначального решения к точному решению. Хорошие результаты дают методы, в которых это решение получается с помощью исходной задачи ЛП, но с ослабленными условиями, что позволяет решать ее за приемлемое время.

В [18] используется тот факт, что решение целочисленной задачи ЛП во много раз медленней, чем решение той же задачи ЛП, но без ограничения на целые значения переменных. Моделирование показало, что, например, для графа из 35 вершин решение не целочисленной задачи занимает около 2–3 секунд, а ограничение на целочисленность увеличивает это время до 32 минут, т.е. более чем в 600 раз. Авторы [18] решают задачу ЛП, заменив условие $v_i \in \{0, 1\}$ на условие $0 \leq v_i \leq 1$ и аналогично для переменных $u_{i,j}$. После этого первоначальное решение набирается из переменных v_i , значения которых наиболее близки к единице.

Также есть работы, в которых используется следующий факт: время решения задачи ЛП возрастает с ростом q . Первоначальное решение получалось следующим образом. Требуемое количество раз решается задача ЛП для малого ($q = 1, 2$) количества удаляемых вершин, после чего полученные решения объединяются.

Помимо описанной двухэтапной схемы решения существуют и другие подходы. Например, в [8] применяется необычный способ: условия задачи сначала усиливаются, на подмножество критических вершин накладываются дополнительные требования, а потом задача в новой постановке решается простым приближенным методом. Для этого водится новая метрика фрагментации

графа и далее формулируется смешанно-целочисленная задача квадратичного программирования, которая одновременно минимизирует количество связанных пар вершин и максимизирует фрагментацию графа. Эта задача приближенно решается с помощью жадного алгоритма. В [15] описывается способ обучения нейронной сети для определения критических вершин. Авторы [6] применяют генетический алгоритм и предлагают три различных подхода его начальной инициализации. В [23] используется мембранный эволюционный алгоритм.

Все рассмотренные в этом разделе алгоритмы являются эвристическими и не дают гарантии, что найденные решения являются оптимальными, а также не дают гарантированной оценки близости полученных решений к точному.

1.3. Постановка цели исследования.

В данном исследовании рассматриваются точные методы и методы, близкие к точным, в которых используется решение вышеописанной задачи ЛП для нахождения критических узлов сети по критерию наименьшего количества связанных пар вершин. Исследуется возможность сократить размерность задачи ЛП, используя характеристики вершин. К проведению этого исследования подтолкнуло следующее рассуждение. Рассмотрим некоторый граф, имеющий вершину степени 1. Очевидно, что удаление такой вершины никак не влияет на связность остальных вершин. Следовательно, за исключением одного случая, эта вершина не будет критической и в задаче ЛП можно заменить нулем переменную v_i , соответствующую этой вершине. Исключение происходит, когда при удалении q вершин все вершины графа G_Q имеют степень 0 или 1. То есть граф распадается на изолированные вершины и пары смежных вершин. Тогда при нахождении оптимального решения для $q + 1$ удаляемых вершин в него попадут вершины степени 1.

Целью исследования является подбор двух характеристик вершины, обладающих следующим свойством. Упорядочим все вершины графа по значению первой характеристики. Характеристика должна быть такая, что несколько первых вершин почти

всегда будут критическими и в задаче ЛП соответствующие переменные v_i можно заменить единицами. Аналогично последние несколько вершин, упорядоченные по значению второй характеристики, почти никогда не будут критическими и соответствующие переменные v_i можно заменить нулями. Кроме того эта пара характеристик должна максимально сокращать размерность задачи ЛП, среди всех других пар. Для найденных двух характеристик необходимо получить оценку количества вершин, для которых переменные v_i заменяются нулями и единицами. Будем называть такую пару характеристик *оптимальной* и задачу ЛП, в которой заменены нулями и единицами соответствующие переменные v_i , – *сокращенной*. Также будем считать, что в современных «решателях» задач ЛП первым стоит модуль оптимизации и уменьшения количества переменных. Тогда при добавлении в задачу ЛП равенств вида $v_i = 1$ или $v_i = 0$ будет происходить сокращение ее размерности.

Исследование получилось достаточно большим из-за того, что потребовалось решить ряд подготовительных задач. Возникла необходимость выбрать значения и обосновать этот выбор для двух параметров: количества критических вершин и доли решений. Только после этого стало возможным сформулировать корректную постановку исследуемой задачи. Общая схема исследования состоит из следующих пунктов:

1. Определение класса графов, для которых проводится исследование. Нет гарантии того, что для графов различных по своей природе инженерных сетей (например, транспортных и тепловых), искомая пара оптимальных характеристик будет одной и той же. Поэтому исследуемый класс графов нужно определить.

2. Определение количества критических вершин q . Для разных диапазонов значений q , по которым ищутся решения задачи ЛП, искомая пара оптимальных характеристик может различаться. Поэтому для постановки задачи необходимо выбрать этот диапазон и обосновать выбор.

3. Оценка доли решений. Предположим, что для заданного графа и количества критических вершин q имеется 100 решений,

т.е. 100 наборов из q вершин (одна и та же вершина может содержаться в нескольких наборах). И пусть вершина k содержится только в одном наборе. Тогда если объявить вершину k не критической, то будет найдено 99 решений. Возникает вопрос: допустима ли такая ситуация? Будем называть *долей решений* отношение количества решений в сокращенной задаче ЛП к количеству решений в исходной задаче. Тогда необходимо определить, какая доля решений допустима.

4. Постановка задачи.

5. Оценка верхней границы количества вершин, которые почти всегда являются критическими, и аналогично для вершин, которые почти никогда ими не являются. То есть оценка максимального допустимого количества равенств вида $v_i = 0$ и $v_i = 1$, которые можно добавить в задачу ЛП, не нарушая ограничений постановки, без учета значений каких-либо характеристик вершин.

6. Выбор оптимального алгоритма подбора характеристик. Предположим, что для заданного графа, q и двух пар характеристик для первой пары можно добавить 5 равенств вида $v_i = 1$ и 10 равенств вида $v_i = 0$, а для второй, соответственно, 4 и 15. Необходимо определить в каком из этих двух случаев сокращенная задача ЛП будет решена быстрее. Для этого нужно более детально оценить количество переменных и ограничений в сокращенной задаче ЛП для каждого случая.

7. Подбор оптимальной пары характеристик. Оценка для каждой оптимальной характеристики допустимого количества равенств вида $v_i = 0$ и $v_i = 1$, которые можно добавить в задачу ЛП и сравнение полученных значений с их верхними оценками.

8. Проверка результатов на отдельных, не рассматриваемых в исследовании примерах.

В данной работе описывается первая часть исследования, состоящая из выбора класса сетей, определении параметров, постановки задачи и оценки верхних границ (пункты 1–5).

2. Определение класса графов

Помимо указанной выше необходимости выбора класса графов для корректной постановки задачи справедливо следующее утверждение, которое накладывает дополнительное ограничение на выбираемый класс.

Утверждение 1. *В общем случае для произвольных сетей не существует характеристики, позволяющей сократить размерность.*

Доказательство. Введем понятие *полного* решения задачи нахождения критических узлов. Для заданного графа существуют две величины q_{begin} и q_{end} , такие, что для всех $q < q_{begin}$ удаление любых q вершин не нарушает связности оставшихся вершин. И существует такой набор из q_{end} вершин, что при их удалении оставшийся граф распадается на изолированные вершины и количество связных пар в нем равно нулю. Тогда полное решение задачи заключается в нахождении q_{begin} и q_{end} и всех решений для каждого $q_{begin} \leq q \leq q_{end}$.

Для доказательства утверждения достаточно привести пример задачи (т.е. граф и количество удаляемых вершин q) для которой не существует характеристики, сокращающей ее размерность

Рассмотрим следующий граф (рис. 1).

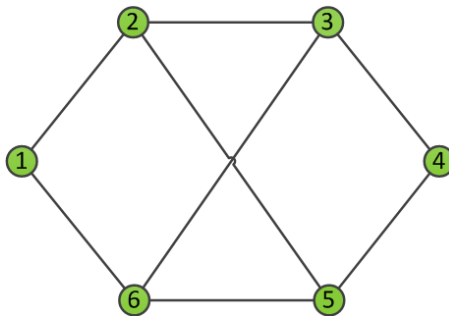


Рис. 1. Пример графа

Полное решение задачи для этого графа выглядит так. При удалении любой единичной вершины связность оставшихся вершин не нарушается. При удалении пары вершин (2, 6) или (3, 5) граф распадается на изолированную вершину 1 или 4 и оставшиеся три связных вершины. При удалении любой другой пары вершин связность оставшихся вершин не нарушается. При удалении тройки вершин (1, 3, 5) или (2, 4, 6) оставшиеся три вершины изолированы. При удалении любой другой тройки вершин связность оставшихся вершин не нарушается.

Покажем, что для этого графа нельзя сократить размерность задачи добавлением равенств вида $v_i = 1$ или $v_i = 0$, в которых вершины выбираются по значению некоторой характеристики.

В изображении данного графа имеются две оси симметрии (вертикальная и горизонтальная), следовательно, любые характеристики вершин 1, 4 совпадают и любые характеристики вершин 2, 3, 5, 6 также совпадают. То есть, с точки зрения выделения подмножества вершин для добавления равенств в задачу ЛП, вершины (1,4) и (2,3,5,6) неразличимы и добавляться могут только следующие группы равенств:

$$v_i = 0 \text{ или } 1; i \in (2, 3, 5, 6) \text{ и}$$

$$v_i = 0 \text{ или } 1; i \in (1, 4).$$

Перебирая эти варианты можно определить, что для $q = 3$ не существует группы равенств, добавление которых приводит к точным решениям.

Сформулируем условия, при которых размерность задачи ЛП будет уменьшаться. Вычислительная сложность современных алгоритмов решения задачи ЛП напрямую зависит от количества переменных и неравенств [4, 7]. Поэтому сокращение размерности будет происходить только при добавлении в задачу ЛП равенств вида $v_i = 1$ или $v_i = 0$. И с другой стороны, например, знание о том, что одна из трех вершин является критической, никак не позволит сократить размерность.

В графе, у которого есть теоретическая возможность снизить размерность задачи ЛП, должны быть вершины, которые почти всегда являются критическими, и вершины, которые по-

чти никогда не являются критическими. Определить это можно следующим способом. Рассмотрим для некоторого графа полное решение задачи нахождения критических узлов и определим для каждой вершины частоту ее появления – долю решений, в которых есть эта вершина. Построим следующую диаграмму: на оси абсцисс отрезок $[0; 1]$ разделен на большое количество (20–50) равных отрезков, обозначим эту ось f . Для каждого отрезка на оси ординат откладывается доля вершин (количество таких вершин, разделенное на количество всех вершин), частота которых лежит в диапазоне этого отрезка; обозначим эту ось h . Будем называть эту диаграмму «диаграмма частот появления вершин в решениях». Тогда у графа есть теоретическая возможность добавлять равенства в задачу ЛП, если на краях (в районе 0 и в районе 1) диаграммы имеются ненулевые столбцы.

В качестве исследуемого класса графов были выбраны укрупненные графы магистральных электрических сетей (сетей передачи энергии) государств СНГ, взятые из открытых источников. В [5] и [2] приведены такие схемы для Белоруссии и Таджикистана. Магистральные электрические сети, связывающие в единое целое отдельные регионы и их крупные источники, имеют замкнутую структуру, что хорошо согласуется с критерием наименьшего количества связных пар вершин. В отличие, например, от распределительных сетей, в которых важна связность источников и потребителей ресурса.

Рассматривалось порядка 20 сетей с количеством узлов 30–60. Для каждой сети было найдено полное решение задачи нахождения критических узлов, общее время расчета составило около 7 месяцев. Из-за такой большой продолжительности расчета не было возможности увеличить количество рассматриваемых графов или их размерность. К сожалению, все последующие выводы основаны на этой малой выборке.

На рис. 2 показаны типовые диаграммы частот появления вершин в решениях для выбранного класса сетей. Наличие ненулевых столбцов в районе единицы и нуля свидетельствует о том, что для рассматриваемых сетей есть теоретическая возможность

уменьшить размерность задачи ЛП.

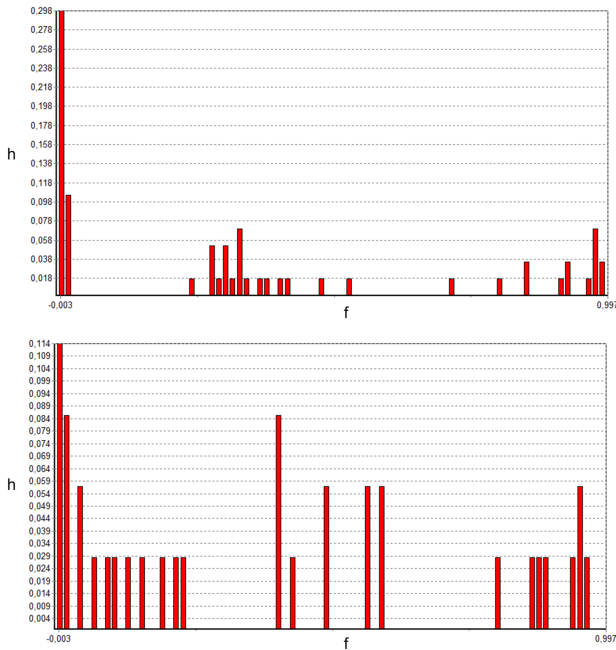


Рис. 2. Диаграммы частот появления вершин в решениях у электрических сетей

В дальнейшем характеристику частоты, с которой вершина встречается в полном решении, будем называть «частотой».

3. Оценка количества критических вершин

Рассмотрим два крайних случая: для заданного графа решается задача нахождения критических узлов при $q = 1$ и при всех $q \leq q_{end}$. Нет никаких гарантий того, что искомая оптимальная пара характеристик в этих случаях будет одной и той же. Будем считать, что задача решается для всех $q \leq q_0$, соответственно, необходимо оценить величину q_0 .

В рассматриваемом случае электрических сетей логично предположить, что после решения задачи ЛП для всех $q \leq q_0$

(нахождения всех решений для каждого q) строится объединение найденных критических вершин; обозначим $P(q_0)$ – объединение всех критических вершин, полученных при решении исходной задачи ЛП для всех $q \leq q_0$, $p = |P(q_0)|$ и $\bar{P}(q_0), \bar{p}$ – аналогичные величины для сокращенной задачи. Для узлов сети, соответствующих множеству $P(q_0)$, проводятся мероприятия, обеспечивающие повышенные требования по надежности и безопасности. Будем считать, что после проведения этих мероприятий вероятность выхода из строя узла, соответствующего вершине из множества $P(q_0)$, пренебрежимо мала. Учитывая, что на поддержание повышенного уровня надежности и безопасности объектов электросетей требуются значительные ресурсы, оценим величину p как функцию от количества вершин $p(n)$ и с ее помощью оценим величину $q_0(n)$.

Величина $p(n)$ была оценена как $p(n) = \frac{n}{5}$ по следующим соображениям.

1. Будем считать, что количество объектов сети пропорционально площади, которую эта сеть охватывает. Количество потребителей также пропорционально этой площади, и на проведение вышеописанных мероприятий каждый из потребителей выделяет одинаковую сумму. Отсюда следует линейный вид $p(n)$.

2. Рассматривались различные государственные программы (касающиеся не только электросетей), в которых для достаточно большого множества объектов проводились мероприятия с целью улучшения показателей всего множества. Причем мероприятия заключались в улучшении, модернизации или замене отдельных объектов. В разных случаях доля таких объектов колебалась от 5 до 18%. Например, доля выселяемых по реновации домов в Москве составляет 6,5% [3], доля школ, попавших в программу обновления российских школ, – 18,3% [1].

3. Из-за небольшого количества как рассматриваемых сетей, так и вершин в каждой сети эта доля была увеличена до 20%.

Далее для каждого графа выполнялась следующая операция. С помощью последовательного перебора значений $q = 1, 2, \dots$ определялась величина \tilde{q} такая, что мощность $P(\tilde{q})$ не превосхо-

дит $\frac{n}{5}$, а мощность $P(\tilde{q} + 1)$ – превосходит. И среди всех графов определялся максимум величины $\frac{\tilde{q}}{n}$. Оказалось, что этот максимум также равен 0,2; это означает, что у одного из графов есть единственное решение задачи ЛП для найденного \tilde{q} и это решение включает в себя все вершины всех решений при $q < \tilde{q}$. В результате для дальнейшего исследования использовалась оценка $q_0 = \lceil n/5 \rceil$.

4. Оценка доли решений

Обозначим отношение количества решений в сокращенной задаче к количеству решений в исходной задаче μ , $0 \leq \mu \leq 1$. Будем называть решение исходной задачи *найденным*, если это решение также является решением сокращенной задачи, и *ненайденным* – в противном случае. Проведем оценку допустимых значений μ , используя следующее рассуждение.

Рассмотрим следующую ситуацию:

– предполагается, что количество узлов, которые могут одновременно выйти из строя, не превосходит q_0 ;

– в любом решении исходной задачи ЛП для количества вершин, не превосходящих q_0 (найденном или найденном), есть хотя бы одна вершина из множества $\overline{P(q_0)}$.

Тогда можно утверждать, что если одновременно выйдет из строя q_0 вершин, то эти вершины не являются решением задачи ЛП для q_0 вершин и никакое подмножество этих вершин также не является решением задачи ЛП для количества вершин, меньших q_0 . То есть любое подмножество вершин из множества $V \setminus \overline{P(q_0)}$ и мощностью не превосходящее q_0 не является критическим. Будем считать, что в этом случае электрическая сеть защищена в должной степени и из этого свойства оценим величину μ .

Пусть для некоторого q общее число всех решений равно d и количество найденных решений равно $\mu \cdot d$ (рис. 3). Обозначим среднюю частоту, с которой вершина встречается в решениях (рассматриваются только вершины, входящие хотя бы в одно из решений), за f . Пока будем считать, что частоты всех вершин

одинаковы и равны f . То есть во всех решениях встречается $f \cdot d$ экземпляров каждой вершины.

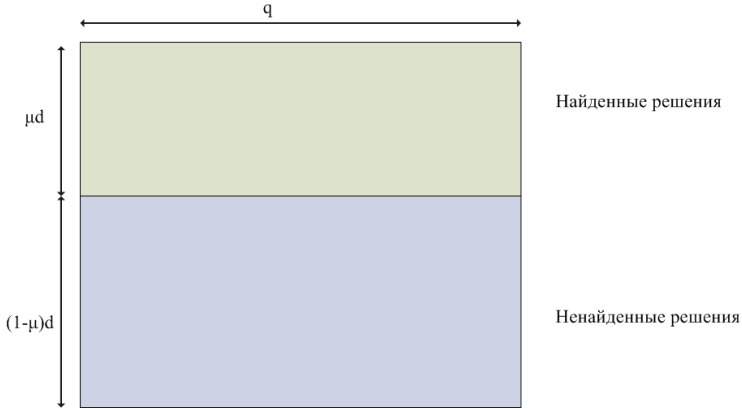


Рис. 3. Схема решений

По условию в каждом из ненайденных решений содержится хотя бы одна вершина из найденных решений. Для оценки μ удобно решать противоположную задачу. Предположим, что среди ненайденных решений существует решение R , ни одна из вершин которого не содержится в найденных решениях. Тогда все экземпляры вершин этого решения также находятся среди ненайденных решений. Из этого условия оценим снизу величину $1 - \mu$; тогда если выбрать μ таким, что $1 - \mu$ будет ниже этой оценки, то можно утверждать, что в любом ненайденном решении содержится хотя бы одна вершина из найденных решений.

Фактически нужно «упаковать» оставшиеся $(f \cdot d - 1) \cdot q$ экземпляров вершин в как можно меньшее количество решений так, чтобы все решения были различны. В каждом решении не может быть больше чем $q - 1$ вершин, совпадающих с вершинами из R . Соответственно, общее количество ненайденных решений можно оценить так:

$$(1 - \mu) \cdot d \geq 1 + \left[\frac{q \cdot (f \cdot d - 1)}{q - 1} \right].$$

$$\text{Тогда если } \mu > 1 - b_0, \text{ где } b_0 = \frac{1 + \left[\frac{q \cdot (f \cdot d - 1)}{q - 1} \right]}{d},$$

то не существует найденного решения R , ни одна из вершин которого не содержится в найденных решениях.

При различных значениях частот будем рассуждать аналогично. Предположим, что существует найденное решение R , ни одна из вершин которого не содержится в найденных решениях. Тогда нужно «упаковать» $\sum_{i \in R} f_i \cdot d - q$ оставшихся экземпляров вершин в как можно меньшее количество решений. И сравнение с предыдущей оценкой приводит к сравнению выражений $f \cdot q$ и $\sum_{i \in R} f_i$.

Результаты этого сравнения следующие: $f \cdot q$ всегда меньше $\sum_{i \in R} f_i$. На качественном уровне это объясняется тем, что величина f усредняется по всем вершинам, среди которых много вершин с малой частотой, а в выражении $\sum_{i \in R} f_i$ усреднение проводится только по q вершинам и доля вершин с малой частотой в них меньше. Количественно оценить величину $\eta = \frac{\sum_{i \in R} f_i}{f \cdot q}$ в общем виде не представляется возможным, поэтому было проведено моделирование. Типичный график зависимости величины η от номера решения N показан на рис. 4.

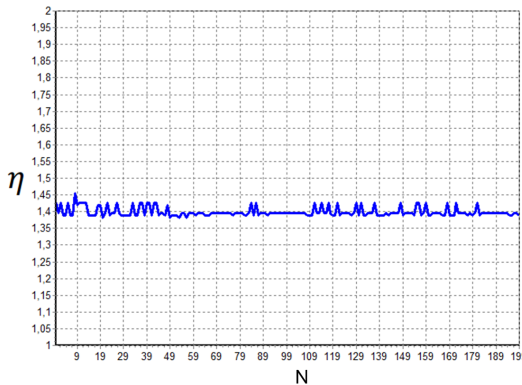


Рис. 4. График зависимости η от номера решения

Графики представляют собой незначительные колебания около горизонтальной линии. В зависимости от графа и количества удаляемых вершин величина η_0 , вокруг которой происходят колебания, меняется от 1,05 до 1,8, и для всех рассматриваемых графов и всех решений $\eta \geq 1,05$.

Окончательно величина μ была оценена как $\max(1 - b)$, где
$$b = \frac{1 + \left[\frac{1,05 \cdot q \cdot (f \cdot d - 1)}{q - 1} \right]}{d},$$
 f – средняя частота, d – количество решений, q – количество удаляемых вершин и максимум брался по всем рассматриваемым графам и всем $q \leq q_0(n)$. Было получено значение $\mu \approx 0,5$, т.е. в результате в задачу ЛП можно добавлять ограничения до тех пор, пока количество решений не сократится вдвое.

5. Постановка задачи

Суммируя результаты, полученные в предыдущих разделах, можно дать следующую постановку задачи.

Рассматривается задача нахождения критических узлов укрупненных магистральных электрических сетей по критерию наименьшего количества связных пар вершин, которая решается точными или приближенными методами с помощью сведения ее к эквивалентной задаче линейного программирования. При этом задача решается для всех $q \leq \lceil n/5 \rceil$, где q – количество удаляемых вершин в графе сети и n – количество вершин графа. Требуется разработать метод снижения размерности этой задачи ЛП. Если для заданного графа и q с помощью сокращенной задачи ЛП будет найдено не меньше половины решений, то такой результат считается приемлемым.

Так как каких-либо определенных свойств графов магистральных электрических сетей неизвестно (даже не имеющее отношения к задаче свойство безмасштабности в последние годы ставится под сомнение), то решение предполагает часто используемый подход, заключающийся в следующем. Разрабатывается эвристический метод и при необходимости уточняются (опреде-

ляются, аппроксимируются) параметры метода на некоторой выборке графов указанного класса сетей. Далее на другой выборке определяется точность метода. Идея предлагаемого метода заключается в подборе двух характеристик вершин, свойства которых описаны в разделе «Постановка цели исследования». Предлагаемый метод также является эвристическим и не дает гарантии того, что не существует графа электрической сети, для которого в сокращенной задаче ЛП будет найдено меньше половины решений.

6. Оценка количества вершин, которые почти всегда являются критическими

Исследование этого вопроса можно разделить на два этапа:

1. Определить для каждого из 20 рассматриваемых графов и для каждого $q \leq \lceil n/5 \rceil$, существует ли такие подмножества вершин, которые почти всегда являются критическими. Найти подмножество, обладающее следующими свойствами (будем обозначать это подмножество A и $a = |A|$): доля решений, содержащих A , больше 0,5, и доля решений, содержащих любое подмножество вершин мощности $a + 1$, меньше 0,5;

2. Оценить величину a для произвольного графа электрических сетей.

6.1. Определение подмножеств критических вершин

Для нахождения подмножеств A наибольшей мощности использовался метод полного перебора, который осуществлялся следующим образом. Для каждого графа и для каждого $1 < q \leq \lceil n/5 \rceil$ перебирались все возможные комбинации вершин, количество вершин в комбинации r изменялось в диапазоне $1 \leq r \leq q - 1$. Для каждой комбинации вычислялась частота ее вхождения в решения, и для каждого r определялась группа комбинаций, частота которых превышала $\mu = 0,5$. В результате для каждого графа и каждого q была получена группа подмножеств A .

6.2. Оценка величины $a = |A|$

Эта подзадача заключается в оценке a для произвольного графа магистральной электрической сети. по имеющимся группам подмножеств A . Так как в распоряжении имелись только группы подмножеств A , то использовались методы аппроксимации, что, конечно, не в лучшую сторону сказалось на точности оценки.

Очевидно, что a зависит от q , какой-либо другой зависимости a , например, от количества вершин и ребер графа, выявлено не было. Оказалось, что лучше оценивать зависимость от q величины $q - a$. На рис. 5 представлен график этой зависимости.

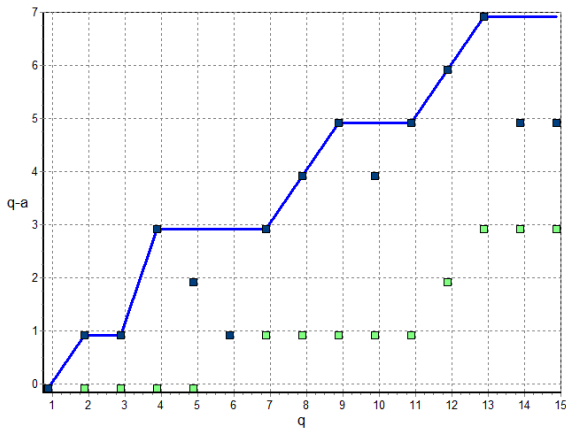


Рис. 5. График зависимости $q - a$ от q

Зелеными точками показаны усредненные по всем рассматриваемым графам значения величины $q - a$, а темно-синими – максимальные (среди всех рассматриваемых графов) значения этой величины. Из графика видно, что средние значения $q - a$ не убывают, следовательно, логично предположить, что и максимальные значения будут не убывать, а убывание в некоторых областях на графике вызвано недостаточным количеством рассматриваемых графов электрических сетей, которое, в свою очередь, вызвано длительным периодом решения для них задачи нахождения

дения критических узлов. Синим цветом показан график неубывающих максимальных значений $q - a$, для которого и строится аппроксимация.

Обозначим как $f^+(q)$ функцию, аппроксимирующую этот график, и чтобы определить ее вид, также рассмотрим график зависимости $\frac{q - a}{q}$ от q , изображенный на рис. 6; раскраска точек аналогична предыдущему рисунку.

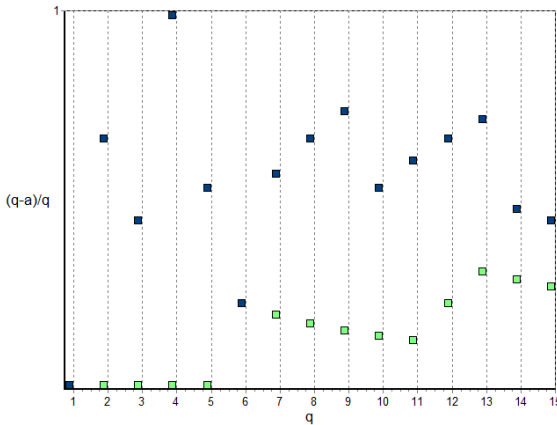


Рис. 6. График зависимости $\frac{q - a}{q}$ от q

Функция $f^+(q) \approx q - a$ не может иметь порядок строго выше q , так как в противном случае $a \approx q - f^+(q) \rightarrow -\infty$. Если бы $f^+(q)$ имела порядок строго ниже q , то на рис. 7 это проявилось бы в немонотонном убывании как средних, так и максимальных значений $\frac{q - a}{q}$, а этого не наблюдается. Поэтому логично аппроксимировать величину $q - a$ функцией порядка q . На рис. 7 коричневой линией показана аппроксимация величины $q - a$ функцией $f^+(q) = 2 \lfloor \frac{q}{4} \rfloor + 1$.

В итоге будем оценивать величину a (наибольшее количество добавляемых в задачу ЛП ограничений вида $v_i = 1$) следующим образом: $a = q - 2 \lfloor q/4 \rfloor - 1$.

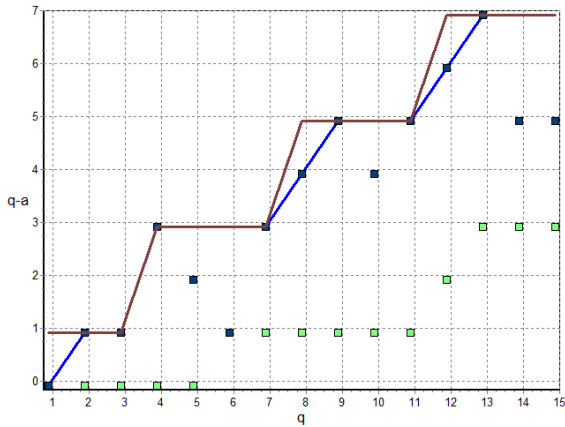


Рис. 7. График зависимости $q - a$ от q и его аппроксимация

7. Оценка количества вершин, которые почти никогда не являются критическими

Аналогично предыдущему разделу исследование этого вопроса можно разделить на два этапа:

1. Определить (для каждого q) подмножество вершин, которые никогда не являются критическими (будем обозначать это подмножество B и $b = |B|$). Найти подмножество, обладающее следующими свойствами: доля решений, не содержащих B , больше 0,5, и доля решений, не содержащее любое подмножество вершин мощности $b + 1$, меньше 0,5.

2. Оценить и по возможности спрогнозировать значение величины $b = |B|$;

7.1. Определение подмножеств B

Аналогично подмножествам A подмножества B также находились методом полного перебора, при этом для каждого графа и каждого q отдельно выделялись подмножества вершин, не входящих ни в одно решение. Это существенно сократило время перебора.

7.2. Оценка величины $b = |B|$

Очевидно, что b зависит и от q и от n . На рис. 8 представлен график зависимости $b(q, n)$.

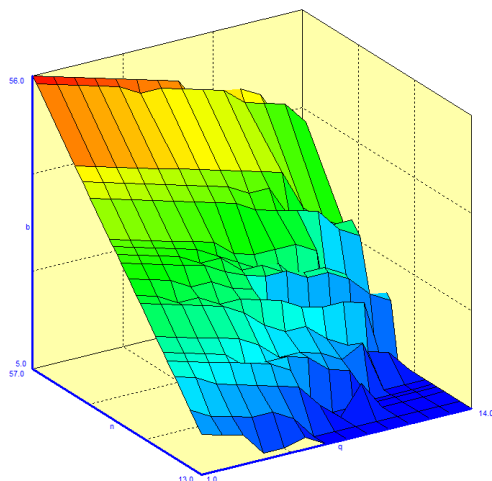


Рис. 8. График зависимости b от q и n

Этот график можно охарактеризовать как «плоскость с отклонениями». Причем отклонения увеличиваются при возрастании q и убывании n , т.е. в области $q \geq \lceil n/5 \rceil$, в которой количество критических узлов превышает полученную в разделе 4 оценку. Было установлено, что в области $q \leq \lceil n/5 \rceil$ выполняется соотношение $b \leq n - q$, причем в большинстве случаев $b = n - q$. Это объясняется тем, что при малых q количество решений мало и равенство будет в случае единственного решения задачи ЛП. В итоге будем оценивать величину b значением $n - q$.

8. Заключение

Рассматривается задача нахождения критических узлов инженерной сети – набора вершин, при удалении которых из графа сети, количество связных пар вершин в оставшемся графе будет

минимально. Один из наиболее часто используемых методов нахождения точного решения этой задачи или решения, близкого к точному, заключается в сведении ее к эквивалентной задаче линейного программирования (ЛП). Для этих методов проводится исследование возможности сокращения размерности задачи ЛП следующим путем. Необходимо подобрать две характеристики вершин таких, что первые несколько вершин, упорядоченных по значению первой характеристики, почти всегда являются критическими и последние несколько вершин, упорядоченных по значению второй характеристики, почти никогда не являются критическими. Если такие характеристики возможно подобрать, то в задаче ЛП можно заменить переменные, отвечающие за то, является ли данная вершина критической или нет, значениями 1 и 0 соответственно, тем самым сократив размерность задачи.

В работе представлена подготовительная часть исследования и получены следующие результаты, конкретизирующие задачу и сокращающие дальнейший перебор.

1. Показано, что в общем случае (для произвольного графа) вышеописанные характеристики вершин подобрать невозможно. Поэтому вместо всех возможных графов рассматривались графы магистральных электрических сетей, для которых показана теоретическая возможность существования таких характеристик.

2. Из общих практических соображений показано, что достаточно рассматривать наборы критических вершин, мощность которых не превышает пятую часть всех вершин.

3. При решении сокращенной задачи возможно будет найдено не все множество решений, т.е. не все наборы критических вершин. Также из общих соображений показано, что допустимо сокращать размерность до тех пор, пока количество найденных решений больше половины всех решений исходной задачи.

4. Сформулирована корректная постановка задачи.

5. Получены верхние оценки количества вершин, которые почти всегда являются критическими и почти никогда не являются критическими.

Литература

1. *Минпросвещения России.* – Режим доступа: <https://edu.gov.ru/modernization> (дата обращения: 22.05.2024).
2. *Перспективы применения комплексов альтернативной энергии на примере республики Таджикистан.* – Режим доступа: <https://research-journal.org/archive/10-64-2017-october/perspektivy-primeneniya-kompleksov-alternativnoj-energii-na-primere-respubliki-tadzhikistan> (дата обращения: 21.05.2024).
3. *Программа реновации. Итоги и планы на 2024 г.* – Режим доступа: <https://www.sobyain.ru/programma-renovatsii-itogi-i-planu-na-2024>(дата обращения: 22.05.2024).
4. ХАЧИЯН Л.Г. *О точном решении систем линейных неравенств и задач линейного программирования* // Вычисл. матем. и матем. физ. – 1982. – №22(4). – С. 999–1002.
5. *Электроэнергия в Беларуси: распределение – статьи энергетической тематики.* – Режим доступа: https://energobelarus.by/blogs/Energy_dis-senting_opinion/19/ (дата обращения: 21.05.2024).
6. BECZI E., GASKO N. *Approaching the bi-objective critical node detection problem with a smart initialization-based evolutionary algorithm* // Peer Computer Science. – 2021. – Vol. 7. – P. 750–758.
7. BONNANS J.F., GILBERT J.C., LEMARECHAL C. et al. *Numerical optimization: Theoretical and practical aspects.* – Springer Berlin–Heidelberg, 2006. – Vol. 51. – P. 494–495.
8. CHEN W., JIANG M., JIANG C. et al. *Critical node detection problem for complex network in undirected weighted networks* // Physica A: Statistical Mechanics and its Applications. – 2020. – Vol. 538. – P. 11–45.

9. DINH T.N., XUAN Y., THAI M.T. et al. *On New Approaches of Assessing Network Vulnerability: Hardness and Approximation* // IEEE ACM Trans. on Networking. – 2012. – Vol. 20(2). – P. 609–619.
10. GLORY U.A., ARULSELVAN A., AKARTUNALI K. et al. *Efficient methods for the distance-based critical node detection problem in complex networks* // Computers and Operations Research. – 2021. – Vol. 131. – P. 108–121.
11. HOOSHMAND F., MIRARABRAZI F., MIRHASSANI S.A. *Efficient Benders decomposition for distance based critical node detection problem* // Omega. – 2020. – Vol. 93. – P. 16–31.
12. LALOU M., KHEDDOUCI H. *Network Vulnerability Assessment Using Critical Nodes Identification* // Int. Symposium on Networks, Computers and Communications (ISNCC). – 2023. – P. 1–6.
13. LIU C., GE G., ZHANG Y. *Identifying the cardinality-constrained critical nodes with a hybrid evolutionary algorithm* // Information Sciences. – 2023. – Vol. 642. – P. 24–41.
14. MEGZARI A., PRAVIJA RAJ P.V., OSAMY W. *Applications, challenges, and solutions to single- and multi-objective critical node detection problems: a survey* // J. Supercomput. – 2023. – Vol. 79. – P. 19770–19808.
15. MUNIKOTI S., DAS L., NATARAJAN B. *Scalable graph neural network-based framework for identifying critical nodes and links in complex networks* // Neurocomputing. – 2023. – Vol. 422. – P. 211–221.
16. SALEMI H., BUCHANAN A. *Solving the Distance-Based Critical Node Problem* // INFORMS Journal on Computing. – 2022. – Vol. 34(3). – P. 1309–1326.
17. SHEN S., SMITH J.C., GOLI R. *Exact interdiction models and algorithms for disconnecting networks via node deletions* // Discrete Optimization. – 2012. – Vol. 9. – P. 172–188.

18. SHEN Y., NGUYEN N.P., XUAN Y. et al. *On the Discovery of Critical Links and Nodes for Assessing Network Vulnerability* // IEEE/ACM Trans. on Networking. – 2013. – Vol. 21. – P. 963–973.
19. THAI M.T., DINH T.T., SHEN Y. *Hardness and Approximation of Network Vulnerability* // Handbook of Combinatorial Optimization. – 2013. – Vol. 5. – P. 1631–1666.
20. UGURLU O., AKRAM N., AKRAM V.K. *Critical nodes detection in IOT-based cyber-physical systems: Applications, methods, and challenges* // Emerging trends in IoT and integration with data science, cloud computing, and big data analytics. – 2022. – Vol. 2022. – P. 226–239.
21. VEREMYEV A., PROKOPYEV O.A., PASILIAO E.L. *Critical nodes for distance-based connectivity and related problems in graphs* // Networks. – 2015. – Vol. 66. – P. 170–195.
22. WALTEROS J.L., VEREMYEV A., PARDALOS P.M. et al. *Detecting critical node structures on graphs: A mathematical programming approach* // Networks. – 2018 – Vol. 73. – P. 48–88.
23. XU Y., GUO P. *MEA-CNDP: A Membrane Evolutionary Algorithm for Solving Biobjective Critical Node Detection Problem* // Computational Intelligence and Neuroscience. – 2021 – Vol. 2021. – P. 101–118.

REDUCTION THE DIMENSIONALITY OF THE TASK OF FINDING CRITICAL NODES IN THE NETWORK

Andrey Krygin, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand.Sc., (andreyakr@yandex.ru).

Sofia Tarasova, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, technician (tarasva_sofia@mail.ru).

Abstract: One of the classes of problems solved in the assessment of the stability of an engineering network is the problem of finding critical nodes. In many formulations, this problem is posed as finding a subset of nodes of a given cardinality (critical nodes) such that the failure of which would cause maximum damage to the entire network. And the most common way to assess the damage in such a formulation is to determine the number of connected node pairs in the network with excluded critical nodes. For such nodes that correspond to the minimum number of connected pairs, additional measures are required to increase reliability and safety. Several methods of solving the problem of finding critical nodes use reducing it to an equivalent linear programming problem. The main problem of this approach is the large size of the problem, and consequently, the high computational complexity of its solution. The work conducts research on various characteristics of vertices of a graph model of a network, the analysis of the values of which will allow determining in advance the fact of belonging to the subset of critical or, conversely, to the subset of non-critical nodes. Thanks to this, it is possible to form additional constraints that reduce the dimensionality of the linear programming problem and its computational complexity, which will allow finding critical nodes in engineering networks with a large number of objects in an acceptable time. During the research, a large number of different subproblems were solved, so the work describes only the first, preparatory part of it.

Keywords: resilience of electrical networks, search for critical nodes, graph models of networks.

УДК 519.1

ББК 31.27-082

DOI: 10.25728/ubs.2024.111.05

*Статья представлена к публикации
членом редакционной коллегии А.А. Печниковым.*

Поступила в редакцию 08.04.2024.

Дата опубликования 30.09.2024.