Минимизация числа пересадок и ствол дерева

Б. Р. Френкин

В «Математическом просвещении» № 21 (с. 272) опубликована следующая задача 3:

Сеть железных дорог представляет собой дерево (неориентированный конечный связный граф без циклов). По сети проходят маршруты электричек, посредством которых можно проехать от любой станции до любой (возможно, с пересадками). Электричка может повернуть назад лишь на конечных станциях маршрута. Количество тупиков (висячих вершин, т. е. вершин степени 1) в данной сети равно n.

а) Каково наименьшее возможное количество маршрутов, если от любой станции до любой можно проехать, возможно с пересадками?

(*Фольклор*)

б) Каково наименьшее возможное количество маршрутов, если от любой станции до любой можно проехать не более чем с одной пересадкой?

(*Б. Р. Френкин*)

Ниже рассмотрена эта задача и связанные с ней результаты.

§ 1. Пересадки и тупики

Вначале обсудим пункт (а). Ясно, что в каждом тупике должен быть конец какого-то маршрута. А так как у маршрута два конца, то при чётном n получается не меньше $n/2$ маршрутов, а при нечётном — не меньше $(n+1)/2$. Эта величина равна $\lfloor (n+1)/2 \rfloor$, т. е. наибольшему целому числу, не превосходящему $(n+1)/2$. *Дальше мы для краткости обозначаем её через m.* Но всегда ли хватит такого количества маршрутов?

Ответ хорошо известен, и он положителен (см. ниже раздел «Решения»). Однако при этом может потребоваться много пересадок. Хотелось бы от них избавиться.

Чтобы в случае n тупиков можно было проехать между любыми двумя станциями без пересадки, потребуется, очевидно, $n(n-1)/2$ маршрутов.

Например, уже при 10 тупиках нужно 45 маршрутов, что многовато. Поэтому мы допустим одну пересадку, т. е. перейдём к пункту (б).

Ясно, что ответ пункта (б) «зажат» между m и $n(n-1)/2$. Однако он оказывается довольно неожиданным: достаточно m маршрутов — как в пункте (а)! Таким образом, если разрешить более одной пересадки, то выигрыша в количестве маршрутов не будет.

§ 2. Некоторые понятия и термины

Приведём пояснения для читателей, незнакомых с терминологией теории графов. Можно представлять граф как совокупность точек (*вершин*) и отрезков (*рёбер*), соединяющих некоторые вершины. Нам потребуются только *конечные* графы (с конечным количеством вершин и рёбер), причём *неориентированные* (на рёбрах не указано направление). И мы рассматриваем только *деревья*, т. е. графы, в которых между любыми двумя вершинами есть ровно один путь.

Количество рёбер, выходящих из данной вершины, называется *степенью* этой вершины. *Висячая вершина* графа — это вершина степени 1. Нетрудно убедиться, что *если в дереве больше одной вершины, то в нём не меньше двух висячих вершин*. Если висячих вершин ровно две, то дерево является *цепью*, т. е. в нём нет развилок. В этой заметке мы говорим о графах на языке дорог и станций, термин «висячие вершины» здесь неестествен, поэтому будем, как и выше, называть их *тупиками*.

Тупики — это «окраины» дерева. А что является его «серёдкой»? Нам этот вопрос интересен, поскольку пересадки целесообразно устроить где-то «посредине». Это понятие можно определить по-разному. Например: «*центр дерева*» состоит из таких вершин, для которых максимум расстояний до других вершин минимален. (Расстоянием до вершины считается количество рёбер в пути к этой вершине; если же каждому ребру приписана некоторая длина, расстоянием считается сумма длин рёбер в этом пути.) Нетрудно проверить, что центр всегда состоит из одной или двух соседних вершин.

Однако в нашей задаче существенны не все вершины дерева, а только тупики и развилки. Задача «не замечает» вершины степени 2. В то же время эти вершины влияют на положение центра! Видимо, потребуется другое понятие.

§ 3. «Удобные» вершины и ствол

Из каждой вершины дерева в каждый тупик идёт ровно один путь, и он начинается с некоторого ребра, смежного с этой вершиной. Будем для краткости говорить, что это ребро *ведёт* в данный тупик, а он с данной вершиной *соединён через это ребро*. Например, если сама вершина — тупик, то единственное выходящее из неё ребро ведёт во все остальные тупики.

Нас интересует, наоборот, «серёдка» дерева. Поэтому будем называть вершину *удобной*, если каждое смежное с ней ребро ведёт не более чем в m тупиков (напомним, что m — наибольшее целое, не превосходящее $(n+1)/2$). Остальные вершины назовём *неудобными*.

Упражнение 1. В любом ли конечном дереве существуют удобные вершины?

Упражнение 2. Докажите, что при $n > 3$ множество удобных вершин либо состоит из одной вершины, либо является цепью, у которой степени концов не меньше 3. Остальные её вершины (если они есть) имеют степень 2, с тем исключением, что при нечётном $n > 3$ одна вершина может иметь степень 3, и тогда к ней примыкает ещё одна цепь (уже без развилок). При $n \leqslant 3$ все вершины дерева — удобные.

Совокупность всех удобных вершин и соединяющих их рёбер, с учётом такого её строения, назовём *стволом*. В задаче 21.3(б) оказывается важен именно ствол.

§ 4. Оптимальные сети

Совокупность маршрутов будем называть *оптимальной сетью*, если между любыми станциями можно проехать, сделав не более одной пересадки, причём количество маршрутов минимально возможное. Как отмечено в конце § 1, это количество равно m. При чётном n это означает, что маршруты оптимальной сети попарно соединяют тупики. Если n нечётно, то $m-1$ маршрутов попарно соединяют $n-1$ тупиков, а оставшийся маршрут выходит из n-го тупика и может дойти до какого-либо другого тупика (который тогда принадлежит двум маршрутам), а может и не дойти.

Упражнение 3. (а) Докажите, что для любой оптимальной сети найдётся вершина ствола, через которую проходят все её маршруты.

(б) Назовём маршрут *полным*, если он содержит все вершины ствола, и *неполным* в противном случае. Каковы возможные количества неполных маршрутов в оптимальной сети: а) при чётных n, б) при нечётных n?

В заключение рассмотрим задачу, в некотором смысле противоположную. Сосредоточение пересадок на стволе имеет и положительную сторону (упрощается организация проезда), и отрицательную (возникает «узкое место»). Попробуем рассредоточить пересадки.

Упражнение 4. (а) Приведите пример дерева и оптимальной сети маршрутов на нём, где все пересадки можно осуществить вне ствола.

(б) Докажите, что в любом примере для п. (а) ствол состоит из единственной вершины степени 3.

Решения

Задача 21.3(а). *Ответ:* $m = \lfloor (n+1)/2 \rfloor$, т. е. $n/2$ *при чётном* n *и* $(n+1)/2$ *при нечётном.*

Вначале разобьём тупики на пары произвольным образом и соединим маршрутом тупики одной пары. При нечётном n один тупик останется без пары — соединим его с одним из остальных тупиков. Количество маршрутов будет как в ответе, но они не обязательно охватывают всё дерево.

Пусть через некоторую станцию A не проходит никакой маршрут. Она не является тупиком (все тупики охвачены). Проведём из A пути по двум разным рёбрам и продолжим эти пути до каких-то тупиков B и C (рис. 1). Эти тупики соединены маршрутами с какими-то тупиками D и E соответственно. Отменим маршруты BD и CE и введём маршруты BC и DE. Новая сеть маршрутов охватывает всё, что охватывали прежние маршруты, и при этом содержит станцию A. Если остались станции вне сети маршрутов, повторяем сделанный шаг, и т. д. Рано или поздно все станции будут охвачены, при этом количество маршрутов не изменится. □

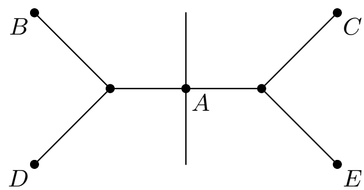

Рис. 1. К задаче 21.3(а)

Задача 21.3(б). *Ответ:* m, т. е. $\lfloor (n+1)/2 \rfloor$ — как в пункте (а). Причём *все пересадки можно устроить на одной станции, и для этого подходят все удобные станции* (определение см. в § 3) *и только они.*

Пусть A — некоторая удобная станция. Проведём пути из A во все тупики и назовём эти пути *лучами*. Занумеруем их. При этом лучи, выходящие из A по одному и тому же ребру, будем нумеровать подряд.

Вначале пусть n чётно. Если номера двух лучей отличаются на $n/2$, то лучи выходят из A по разным рёбрам. Действительно, если они проходят по одному ребру, то по тому же ребру проходят все лучи с промежуточными номерами. Но тогда по этому ребру проходит более $n/2$ лучей, что невозможно, поскольку A — удобная станция. Значит, любые два луча с разницей номеров $n/2$ можно объединить в маршрут. Всего таких маршрутов получится $n/2$, они проходят через A и образуют оптимальную сеть (определение см. в § 4).

Пусть теперь n нечётно. Если по какому-то ребру из A выходит ровно $(n+1)/2$ лучей (такое ребро может быть только одно), то один из этих лучей временно «забудем». Если по любому ребру выходит меньше лучей, то «забудем» произвольно выбранный луч. Останется чётное количество лучей, причём по любому ребру из A выходит не больше $(n-1)/2$ лучей. Поэтому их можно объединить в маршруты попарно, как выше. Объявим «забытый» луч отдельным маршрутом (можно произвольно продолжить его за вершину A). Мы получили оптимальную сеть, и все её маршруты проходят через A.

Если же некоторая станция B — неудобная, то устроить там «всеобщую пересадку» в оптимальной сети маршрутов невозможно. В самом деле, некоторое ребро e, смежное с B, ведёт более чем в m тупиков. Но оптимальная сеть состоит из m маршрутов, и если все они проходят через B, то какие-то два луча, содержащие e, принадлежат одному и тому же маршруту, что невозможно.

Осталось недоказанным только одно: существование удобной станции. Это тема упражнения 1. □

Упражнение 1. *Ответ: да.* Решение. Выберем некоторый тупик. Из него выходит единственное ребро. Если из его конца какое-то ребро ведёт больше чем в m тупиков, пойдём по нему. В следующей вершине поступим так же, и т. д. Мы ни в какой момент не повернём назад. Действительно, в момент поворота по каждую сторону от пройденного ребра находилось бы больше половины всех тупиков, что невозможно.

Поскольку мы не можем повернуть назад, продолжать путь до бесконечности или замкнуть его, то рано или поздно остановимся. Но это означает, что мы пришли в удобную вершину. □

Упражнение 2. Случай $n \leqslant 3$ очевиден. Пусть $n > 3$ и в стволе (множестве удобных вершин) нашлись две различные вершины A и B. Пройдём по единственному пути из A в B. Пусть нам встретилась неудобная вершина C. Какое-то выходящее из неё ребро e ведёт в более чем m тупиков. Предположим, что мы пришли в C по этому ребру (рис. 2а). Тогда следующее ребро пути, если пройти по нему в обратном направлении, тоже ведёт более чем в m тупиков. Поэтому следующая вершина тоже окажется неудобной, и т. д. вплоть до вершины B. Но тогда B не принадлежит стволу — противоречие. Если же мы пришли в C по некоторому ребру f, отличному от e (рис. 2б),

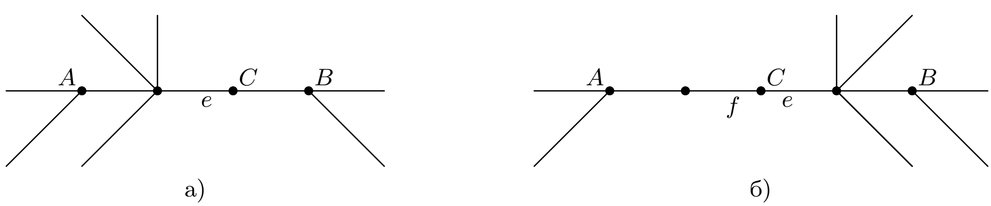

Рис. 2. К упражнению 2

то f тоже ведёт более чем в m тупиков, поэтому предыдущая вершина пути — неудобная, и т. д. вплоть до вершины A — снова противоречие.

Мы доказали, что совокупность удобных вершин и соединяющих их рёбер связна и, значит, является деревом, которое мы обозначим T. Пусть T содержит более одной вершины, и пусть C и D — два его тупика. Если, скажем, C является тупиком в исходном дереве, то единственное выходящее из него ребро ведёт в $n-1$ тупиков. Но оно не может вести в более чем $\lfloor (n+1)/2 \rfloor$ тупиков, поэтому $n \leqslant 3$. Значит, при $n > 3$ вершина C не является тупиком в исходном дереве. Если её степень равна 2, то обе соседние с ней вершины — удобные, что невозможно. Следовательно, в вершине C имеется развилка. То же верно и для D (рис. 3).

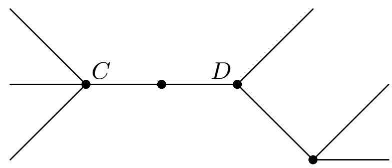

Рис. 3. К упражнению 2

Пройдём по пути, соединяющему C с D. Пусть n чётно. Первое ребро пути ведёт не более чем в $n/2$ тупиков, и каждое последующее ребро ведёт в не большее количество тупиков, чем предыдущее. Если последнее ребро пути ведёт менее чем в $n/2$ тупиков, то в обратную сторону оно ведёт более чем в $n/2$ тупиков, что невозможно, так как D — удобная вершина. Значит, каждое ребро пути из C в D ведёт ровно в $n/2$ тупиков, поэтому в промежуточных вершинах нет развилок. Как следствие, ствол является цепью. С каждой стороны от него расположено по $n/2$ тупиков.

Пусть теперь n нечётно. Аналогично получаем, что первое ребро пути ведёт не более чем в $(n+1)/2$ тупиков, а последнее — не менее чем в $n-(n+1)/2=(n-1)/2$ тупиков. Эти величины отличаются на 1, поэтому возможна лишь одна промежуточная развилка F, где от пути из C в D отходит *отросток* — путь, ведущий в единственный тупик (и значит, не имеющий развилок). Любая вершина отростка, кроме F, не принадлежит стволу, поскольку ребро, выходящее из неё в сторону F, ведёт в $n-1>m$ тупиков (рис. 4). Разумеется, отростка может и не быть (рис. 5). В обоих случаях ствол является цепью, с одной стороны от которой расположено $(n-1)/2$

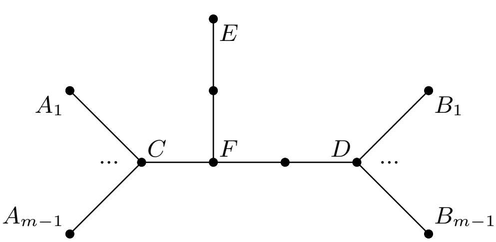

Рис. 4. К упражнениям 2, 3(б)

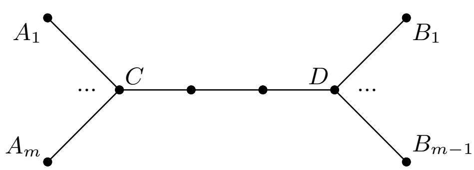

Рис. 5. К упражнениям 2, 3(б)

тупиков, а с другой — $(n-1)/2$ при наличии отростка или $(n+1)/2$ при его отсутствии. □

Упражнение 3. (а) Доказываемое утверждение означает, что любая оптимальная сеть строится по некоторой вершине ствола, как в решении задачи 21.3(б). Достаточно показать, что через некоторую вершину дерева проходят все маршруты сети, поскольку такая вершина заведомо принадлежит стволу. Но здесь можно воспользоваться известным фактом: *если по дереву проходит несколько путей и любые два пересекаются, то найдётся вершина, общая для всех путей* (докажите!).

Можно и не использовать этот общий факт. Пусть S — вершина ствола, не принадлежащая некоторому маршруту μ из данной оптимальной сети (рис. 6). Тогда существует (единственный) путь, соединяющий S с μ. Пусть он начинается с ребра e и заканчивается в вершине T. Рёбра, выходящие из S и отличные от e, ведут суммарно не менее чем в $m-1$ тупиков. Маршруты, выходящие из этих тупиков, должны пересекаться с μ согласно § 4. Поэтому все они различны и проходят через T. Так как их количество не меньше $m-1$ (на самом деле и не больше), то вместе с μ они составляют всю сеть. Таким образом, все маршруты сети проходят через T.

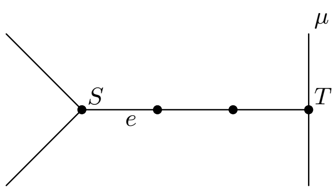

Рис. 6. К упражнению 3(а)

(б) *Ответ:* 0 *при чётном* n; 1 *или* 2 *при нечётном* n *и наличии отростка* (см. решение упр. 2); 0 *или* 1 *при нечётном* n *и отсутствии отростка*. Пусть ствол содержит более одной вершины. Если n чётно, то согласно решению упр. 2 с каждой стороны от ствола расположено $n/2$ тупиков, а согласно § 4 оптимальная сеть попарно соединяет тупики. Если соединены два тупика с одной стороны от ствола, то соединены и два тупика с другой стороны. Но соответствующие маршруты не пересекаются, а это невозможно согласно § 4. Значит, каждый маршрут оптимальной сети соединяет тупики с разных сторон от ствола и потому полон (содержит весь ствол).

Пусть теперь n нечётно и имеется отросток (рис. 4). Пусть тупики $A_1, \ldots, A_{m-1}$ расположены со стороны вершины C, а тупики $B_1, \ldots, B_{m-1}$ — со стороны вершины D. Маршруты A_1E, B_1E, A_2B_2, $\ldots$, $A_{m-1}B_{m-1}$ образуют оптимальную сеть, причём два маршрута A_1E и B_1E неполны, а остальные полны. Количество неполных маршрутов здесь можно уменьшить до 1: заменим, например, маршрут A_1E на A_1B_1. До нуля это количество уменьшить не удастся: маршрут, содержащий E, не может содержать весь ствол.

Покажем, что больше двух неполных маршрутов здесь быть не может. Рассмотрим какую-нибудь оптимальную сеть на данном дереве. Проходящий по отростку маршрут μ начинается (в обозначениях рис. 4) в E и либо заканчивается в F, либо идёт от F в направлении одного из концов ствола, например C. Так как маршрут μ пересекается с остальными, то из тупиков $B_1, \ldots, B_{m-1}$, расположенных за вершиной D, исходят $m-1$ маршрутов сети, содержащих F. Вместе с μ они составляют всю сеть. Тогда до $m-1$ тупиков $A_1, \ldots, A_{m-1}$, расположенных за вершиной C, доходит не менее $m-1$ маршрутов. Если туда доходят все маршруты или не доходит только μ, то только μ неполон. Если же туда не доходит какой-то другой маршрут, то он может быть полным или неполным, маршрут μ заведомо неполон, а остальные полны. Таким образом, количество неполных маршрутов равно 1 или 2.

Предположим теперь, что отросток отсутствует (рис. 5). Тогда с одной стороны от ствола находится m тупиков, а с другой $m-1$. Ясно, что возможна оптимальная сеть из полных маршрутов. Пусть в оптимальной сети некоторый маршрут ν неполон. Если он соединяет два тупика, то они расположены с одной стороны от ствола. Все маршруты, исходящие из тупиков с другой стороны ствола, пересекаются с ν и поэтому содержат весь ствол. Вместе с ν они составляют всю сеть. Если же каждый неполный маршрут содержит лишь один тупик, то, как мы видели в § 4, такой маршрут единствен. Таким образом, если n нечётно и отростка нет, то либо неполный маршрут один, либо таких нет совсем. $\square$

Упражнение 4. (а) Пусть ствол состоит из одной вершины S степени 3 (отметим, что в этом случае тупиков больше трёх, так как иначе ствол совпадал бы со всем деревом, см. упр. 2). Согласно упражнению 3, в оптимальной сети все маршруты проходят через S. Пусть каждый из них соединяет два тупика (это заведомо верно при чётном n, см. начало § 4). Тогда любые два маршрута имеют общее ребро, смежное с S, и пересадку между ними можно устроить на другом конце этого ребра.

(б) Пусть все пересадки в оптимальной сети можно устроить вне ствола. Тогда $n > 3$ (иначе ствол совпадал бы со всем деревом). Вначале предположим, что ствол состоит из одной вершины S степени выше трёх (рис. 7). Согласно упр. 3(а) все маршруты сети проходят через S. Если один из них кончается в S (что возможно при нечётном n, см. § 4), то продолжим его произвольным образом. Ясно, что сеть останется оптимальной и пересадки будут возможны вне ствола.

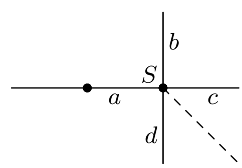

Рис. 7.
К упражнению 4(б)

Пусть некоторый маршрут №1 проходит через рёбра a и b, смежные с S. К S примыкают ещё два ребра c и d (и, возможно, ещё какие-то рёбра). По ребру c проходит некоторый маршрут №2. Пересадка с него на маршрут №1 вне ствола возможна, если маршрут №2 проходит через одно из рёбер a, b — например, через a.

Предположим, что некоторый маршрут №3 не содержит ребро a. Так как с него можно пересесть вне ствола на маршруты №1 и №2, он содержит рёбра b и c. По ребру d проходит некоторый маршрут №4. С него можно пересесть вне ствола на маршрут №1 (аналогично на маршрут №2), поэтому маршрут №4 содержит ребро a или b (аналогично получаем a или c). Это означает, что маршрут №4 содержит ребро a. Но если он проходит через a и d, то с него нельзя пересесть вне ствола на маршрут №3 — противоречие.

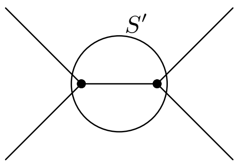

Рис. 8. К упражнению 4(б)

Таким образом, все маршруты проходят через ребро a. Но тогда его конец, отличный от S, принадлежит стволу — противоречие.

Пусть теперь ствол содержит более одной вершины. Заменим его одной вершиной S' (рис. 8) и преобразуем маршруты естественным образом. Стволом нового дерева является S', оптимальная сеть остаётся оптимальной, и по-прежнему все пересадки можно сделать вне ствола. По доказанному, вершина S' должна иметь степень 3. Но на каждом конце прежнего ствола была развилка, поэтому S' имеет степень не ниже 4 — противоречие. $\square$

Борис Рафаилович Френкин, МЦНМО
frenkin@mccme.ru