Сетевое программирование

Лекция 6

Метрики центральности в сети. Интерпретация для телекоммуникационных сетей

Усіпов Н.М., PhD · КазНУ имени аль-Фараби

Цель и план лекции

Цель

Изучение основных мер центральности узлов и рёбер, умение вычислять их средствами NetworkX и правильно интерпретировать полученные значения при анализе телекоммуникационных сетей.

1. Введение
2. Основные меры центральности
3. Различие мер на примере
4. Интерпретация для телекоммуникационных сетей
5. Вычисление на взвешенных сетях
6. Связь мер центральности и выбор показателя
7. Устойчивость оценок и вычислительные аспекты

1 Введение

- Узлы одной сети неравноценны. Отказ маршрутизатора на границе сети затронет лишь небольшую группу абонентов, а отказ узла на магистрали может нарушить передачу данных в целом регионе.
- В телекоммуникационной сети важным может быть узел с большим числом подключений, узел, через который проходит много транзитного трафика, либо узел, расположенный близко ко всем остальным.

2 Основные меры центральности

Центральность по степени

Определяется числом непосредственных соединений узла, нормированным на максимально возможное значение:

$C_D(v) = deg(v) / (|V| - 1).$

Центральность по близости

Определяется через сумму расстояний от узла до всех остальных:

$C_C(v) = (|V| - 1) / \Sigma\, d(v, u).$

Центральность по посредничеству

Показывает, какая доля кратчайших путей проходит через данный узел:

$C_B(v) = \Sigma\, \sigma_{st}(v) / \sigma_{st},$

Центральность по собственному вектору

Важность узла определяется важностью его соседей. Значение пропорционально соответствующей компоненте главного собственного вектора матрицы смежности.

2 Основные меры центральности (продолжение)

PageRank

Вариант центральности для ориентированных сетей, предложенный создателями Google.

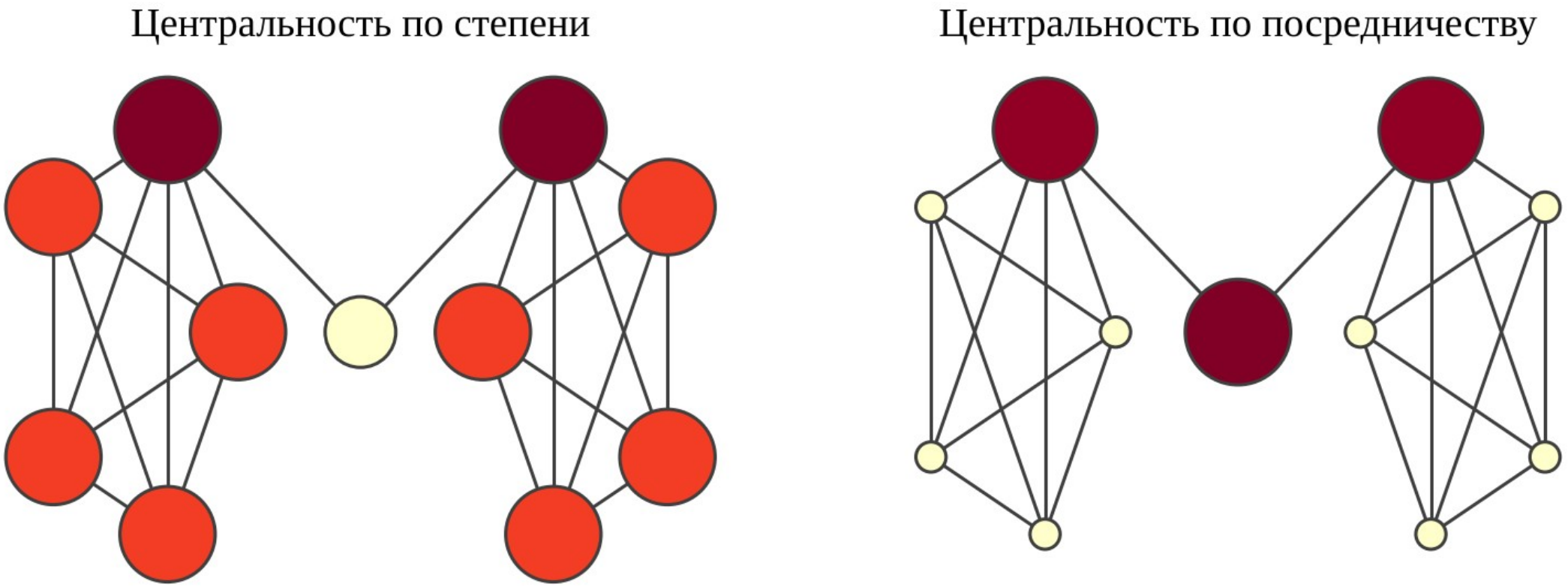
Центральность рёбер по посредничеству

Та же идея применительно к связям. Ребро с высоким значением является важным магистральным каналом, а в ряде алгоритмов служит признаком границы между кластерами сети.

3 Различие мер на примере

- Размер и цвет вершины отражают значение центральности.

3 Различие мер на примере (продолжение)



Сравнение центральности по степени и по посредничеству

3 Различие мер на примере (продолжение)

- Промежуточный узел имеет всего две связи, поэтому его центральность по степени мала.

3 Сравнение мер центральности

Листинг 6.1.

```
import networkx as nx

G = nx.barbell_graph(5, 1)      # две клики и связующий узел 5

deg = nx.degree_centrality(G)
clo = nx.closeness_centrality(G)
bet = nx.betweenness_centrality(G)
eig = nx.eigenvector_centrality(G, max_iter=1000)

print("Узел  Степень  Близость  Посредн.  Собств.")
for v in (0, 4, 5):
    print(f"{v:4} {deg[v]:8.3f} {clo[v]:9.3f} "
          f"{bet[v]:9.3f} {eig[v]:8.3f}")
```

```
Узел  Степень  Близость  Посредн.  Собств.
   0    0.400     0.400     0.000    0.305
   4    0.500     0.526     0.533    0.337
   5    0.200     0.556     0.556    0.164
```

4 Интерпретация для телекоммуникационных сетей

Высокая степень

Устройство обслуживает много соединений: требуются большой запас портов и производительности.

Высокая близость

Узел удобен для размещения общих сервисов, поскольку обеспечивает малое среднее число переходов к остальным узлам.

Высокое посредничество

Через узел идёт значительная часть транзитного трафика. Необходимы резервирование и мониторинг перегрузок.

Высокое значение собственного вектора

Узел находится в ядре сети, соединён с другими важными узлами.

Необходимо учитывать и ограничения. Центральность по посредничеству предполагает, что весь трафик идёт по кратчайшим путям, а все пары узлов обмениваются данными в одинаковом объёме.

4 Интерпретация для телекоммуникационных сетей (продолжение)

Расхождение мер

Наибольший интерес вызывают узлы с малой степенью и большим посредничеством: они легко остаются незамеченными при простом подсчёте подключений, но критичны для всей сети.

Необходимо учитывать и ограничения. Центральность по посредничеству предполагает, что весь трафик идёт по кратчайшим путям, а все пары узлов обмениваются данными в одинаковом объёме.

5 Вычисление на взвешенных сетях

- В NetworkX функции центральности поддерживают параметр `weight`. Важно понимать его смысл: при вычислении посредничества и близости вес трактуется как расстояние, поэтому можно использовать задержку или стоимость канала.
- Сложность вычисления посредничества алгоритмом Брандеса для невзвешенного графа составляет $O(|V||E|)$.

5 Центральность во взвешенной сети

Листинг 6.2.

```
import networkx as nx

G = nx.Graph()
G.add_weighted_edges_from([
    ("Алматы", "Астана", 12), ("Алматы", "Шымкент", 7),
    ("Шымкент", "Актобе", 14), ("Астана", "Караганда", 2),
    ("Астана", "Костанай", 7), ("Астана", "Актобе", 13),
    ("Актобе", "Атырау", 8)], weight="delay")

hop = nx.betweenness_centrality(G)
ms = nx.betweenness_centrality(G, weight="delay")
ebc = nx.edge_betweenness_centrality(G, weight="delay")

top = sorted(ms, key=ms.get, reverse=True)[:3]
for v in top:
    print(f"{v:10} по числу переходов: {hop[v]:.2f}, "
          f"по задержке: {ms[v]:.2f}")
u, v = max(ebc, key=ebc.get)
print("Самый загруженный канал:", u, "-", v)
```

5 Результат: Центральность во взвешенной сети

```
Астана     по числу переходов: 0.67, по задержке: 0.60
Актобе     по числу переходов: 0.43, по задержке: 0.33
Алматы     по числу переходов: 0.10, по задержке: 0.20
Самый загруженный канал: Алматы - Астана
```

Значения задержек в приведённом примере являются условными и служат только для демонстрации.

6 Связь мер центральности и выбор показателя

- Меры центральности часто коррелируют, однако различия между ними содержат важную информацию.
- Для ориентированных сетей каждая мера требует дополнительного уточнения. Близость определяется для входящих и исходящих расстояний отдельно.

$$H(v) = \sum_{u \neq v} 1 / d(u, v).$$

- Для учёта реальной нагрузки разработаны разновидности посредничества. Мера потока учитывает не только кратчайшие пути, а центральность случайного блуждания предполагает, что данные могут двигаться по нескольким альтернативным маршрутам.

7 Устойчивость оценок и вычислительные аспекты

- Значения центральности могут значительно изменяться при небольшой модификации сети.
- Также важен порядок вычислений. Степень находится за время O(|V| + |E|), близость требует поиска путей от каждой вершины, а точное посредничество на сети с миллионом узлов выполнить затруднительно.
- Применение центральности при управлении сетью включает: выбор узлов для размещения средств мониторинга, определение приоритетов технического обслуживания и резервирования, выбор мест подключения новых абонентов.

8 Разбор примера

1. Вычислим показатели для звезды из пяти узлов (центр 0, листья 1–4).
2. Центральность по степени: для центра 4/(5−1) = 1, для листа 1/4 = 0,25.
3. Для центра сумма расстояний равна 4, близость 4/4 = 1.
4. Для листа расстояние до центра равно 1, до трёх остальных листьев – по 2, сумма 7, близость 4/7 ≈ 0,571.
5. Через центр проходят кратчайшие пути всех пар листьев: их C(4,2) = 6.
6. Нормирующий множитель (n−1)(n−2)/2 = 6, поэтому посредничество центра равно 1, а для листьев оно равно 0.
7. Следовательно, все меры указывают на центр как на единственный важный узел.

8 Разбор примера (продолжение)

8 В реальных топологиях такого единодушия обычно нет, поэтому требуется сравнение нескольких мер.

9 Вопросы для обсуждения на занятии

Резервирование

Какие три узла вы защищали бы в первую очередь при ограниченном бюджете и по какой мере?

Веса

Как изменится оценка узлов при переходе от числа переходов к задержке?

Нагрузка

Какими данными необходимо дополнить анализ, чтобы центральность отражала фактический трафик?

Контрольные вопросы

1. Для чего используются меры центральности?
2. Как определяются центральность по степени и центральность по близости?
3. Что показывает центральность по посредничеству? Приведите её формулу.
4. В чём заключается идея центральности по собственному вектору и PageRank?
5. Почему узел с небольшой степенью может оказаться критическим для сети?
6. Какие меры целесообразно использовать при выборе места для сервера и при поиске узких мест?
7. Какие допущения лежат в основе центральности по посредничеству?
8. Почему пропускную способность нельзя непосредственно использовать как вес при вычислении посредничества?

Список литературы

1. Freeman L.C. A set of measures of centrality based on betweenness // Sociometry. – 1977. – Vol. 40, No. 1. – P. 35–41.

2. Freeman L.C. Centrality in social networks: conceptual clarification // Social Networks. – 1978. – Vol. 1, No. 3. – P. 215–239.

3. Bonacich P. Factoring and weighting approaches to status scores and clique identification // Journal of Mathematical Sociology. – 1972. – Vol. 2, No. 1. – P. 113–120.

4. Brandes U. A faster algorithm for betweenness centrality // Journal of Mathematical Sociology. – 2001. – Vol. 25, No. 2. – P. 163–177.

5. Brin S., Page L. The anatomy of a large-scale hypertextual Web search engine // Computer Networks and ISDN Systems. – 1998. – Vol. 30, No. 1–7. – P. 107–117.

6. Newman M. Networks. – 2nd ed. – Oxford: Oxford University Press, 2018. – 780 p.

7. NetworkX Documentation [Электронный ресурс]. – URL: https://networkx.org/documentation/stable/