Graphs, I

李昂生

Discrete Mathematics
U CAS
5, June, 2018

Outline

1. Definitions
2. Communication network
3. Graphs in the information age
4. Elements of graph theory
5. Exercises

General view

- The most **general mathematical model** for sciences
- **Networks** - the 21st century graph theory
- 丘成桐2017问题：**网络的公理化研究**

Nevertheless, we will need a new **network theory**

Definition

A graph is a pair of sets, i.e., $G = (V, E)$, such that

(i) V is the set of *vertices* or *nodes*,

(ii) E is the set of *edges* or *links* of two vertices.

- An edge (u, v) has two endpoints u, v.
- If the edges are ordered pairs of vertices, then G is called a *directed graph*. Otherwise, it is an *undirected graph*.
- An edge (u, v) is called a *self-loop*, if $u = v$.
  It reflects the *reflexivity* of relations.
- An edge $e = (u, v)$ may have a weight, in which case, the graph is called a *weighted graph*.

Graph as An Extension of Relations

We know that

(i) Relation is an extension of functions, and
(ii) Graphs are extensions of relations.

Questions

1) Why?
2) Are graphs well-defined mathematical model?

Classical Graphs

Classical graphs are the mathematical model for discrete and interesting systems that are usually

- Small-scaled
- Static
- Relatively regular
- Deterministic
- Representation of the laws of links

Advantages of Graphs

- Geometric intuition
- Combinatorial characteristics
- Algebraical study

Simple Graphs

Definition 1

A graph $G = (V, E)$ is called *simple*, if:

(i) there is no self-loop

(ii) there is no multiple edges.

Remark: Self-loops are important for CS

Multigraphs

A graph is called a multigraph, if it contains multiple edges.

This is a special case of weighted graphs.

Directed Graphs

Definition 2

A graph $G = (V, E)$ is called *directed*, if the edges E are ordered pairs.

For a directed graph G, an edge $e = (u, v)$ is an ordered pair. For an undirected graph G, an edge is an unordered pair, denoted $e = \{u, v\}$, meaning that there are two directions both from u to v and from v to u.

Elements of a Graph

How to understand a graph? A graph contains

- **Syntax**
- **Semantics**
- **Structural functions**

Fundamental Questions

(1) Representations
   - Algebraic representation
   - Geometric representation

   Key: The representations must support mathematical study of graphs.

(2) Mathematical properties of graphs

(3) Operations in graphs
   Key: Graph algorithms

Problems

(1) Characterisation of the graphs that support quick spreading of information and quick transportation of data
(2) Shannon 1953: What are the optimal communication networks?

We will talk one of the issues: the expander graphs

The Graphs in the Information Age

1) The web graph
2) Social networks
   – acquaintanceship graphs
   – friendship graphs
3) Influence graphs
4) Cooperation graphs
5) Call graphs
6) Citation graphs
7) Dependency graphs
8) Airline routes

The Graphs in the Information Age - Continued

9) Road networks
10) Protein interaction graphs
11) Computer networks
12) Mobile networks
13) Economics networks
14) Gene networks
15) Molecular topology
16) Physical systems

The Graphs in the Information Age - Continued

17) Networks that are naturally evolving
18) Engineering networks
19) Networking engineering
20) Computing systems etc

Basic Characteristics

- Large
- Dynamically evolving
- Sparse
- Laws
- Random variations
- Various interaction, communications and operations that occur in the networks

The Challenges of the New Graphs

1. What are the basic laws
2. Modelling
3. Dynamical complexity
4. Robustness and security
5. Concentration and convergence
6. New engineering
7. New economics and social sciences
8. New science in general including math, physics, CS and Information Science

Adjacency - notations

Let $G = (V, E)$ be a graph of n.

- Two vertices u and v in an undirected graph are called *adjacent* in G, if u and v are endpoints of an edge $e \in E$.
- If $e = (u, v)$ is an edge, we say that e is *incident with u and v*.

  For an edge $e = (u, v)$, we also say that
- e connects or links u and v
- u is a *neighbour of v*.

Neighbourhood

Definition 3

Let $G = (V, E)$ be a graph.

(1) For a vertex v, the *neighbourhood of v* is the set of all the vertices u such that u, v are adjacent.
We use $N(v)$, or $\Gamma(v)$ to denote the set of neighbours of v in G.

(2) For a set A of vertices, define

$$N(A) = \cup_{v \in A} N(v).$$

This is also referred to as

$$\Gamma(A).$$

Degree

Definition 4

Let $G = (V, E)$ be a graph and $v \in V$. The *degree of v* is the number of edges that link to v.
We use $d(v)$ to denote the degree of v in G.
A vertex v is called *isolated*, if it has zero degree, i.e., $d(v) = 0$.

Theorem 5

Let $G = (V, E)$ be an undirected graph with m edges such that none of the edges is a selfloop. Then

$$\sum_{v \in V} d(v) = 2m.$$

Every edge contributes 2 to the total degree.
Note A selfloop of a vertex v contributes degree 1 to v.

Volume

Definition 6

Let $G = (V, E)$ be an undirected graph and $S \subseteq V$. The *volume* of S in G is defined by

$$\mathrm{vol}(S) = \sum_{u \in S} d(u). \tag{1}$$

Degree of Vertices of Weighted Graphs

Definition 7

Let $G = (V, E, W)$ be an undirected graph with weight function W from E to $\mathbb{R}^+$.
For a vertex $v \in V$, define the *degree* of v in G as follows:

$$d(v) = \sum_{e=(u,v)\in E} W(e). \tag{2}$$

Remark: Graphs may have both positive and negative weights. How can we deal with the graphs of this type?

Degree Priority

Definition 8

(Li, Pan) Given a graph $G = (V, E)$, suppose that the set of vertices V has a partition $\mathcal{P} = \{X_1, X_2, \cdots, X_N\}$ such that each module X_i is a natural group of G, in the sense that each vertex has a color, the vertices in every module X_i share the same color, and vertices in different modules have different colors.
For every vertex $v \in V$ and each possible j, define the j-th degree of v in G to be the j-th largest number of edges between v and the vertices of the same color. We use $d_j(v)$ to denote the j-th degree of v in G.
The *degree priority of v in G* is an l-tuple of the following form:

$$(d_1(v), d_2(v), \cdots, d_l(v)),$$

for some l.

Degrees in a Directed Graph

Given a directed graph $G = (V, E)$ and a vertex v, the *in-degree* of v in G is the number of edges arrival at v, denoted $d_{\text{in}}(v)$; the *out-degree* of v in G is the number of edges leaving from v in G, denoted $d_{\text{out}}(v)$.

Theorem 9

Let $G = (V, E)$ be a directed graph. Then

$$\sum_{v \in V} d_{\text{in}}(v) = \sum_{v \in V} d_{\text{out}}(v) = |E|. \tag{3}$$

Complete Graph K_n

- Denote K_n
- Simple
- For any $1 \leq i, j \leq n$, $i \neq j$, there is an (undirected) edge (i, j).

Cycle C_n

A cycle of length n for $n \geq 3$ is a graph of n nodes $v_1, v_2, \cdots, v_n$ having edges

- (v_i, v_{i+1}), $1 \leq i \leq n$, and
- (v_n, v_1).

Wheel W_n

For $n > 3$, the wheel W_n consists of a cycle C_{n-1} and a center vertex u such that for each vertex v of cycle C_{n-1}, there is an edge between u and v.

n-Cube Q_n

For each n, and n-dimensional hypercube or n-cube, written Q_n, consists of

- V consists of $(a_1, a_2, \cdots, a_n)$ where each $a_i = 0$ or 1.
- For $\alpha = (a_1, a_2, \cdots, a_n)$ and $\beta = (b_1, b_2, \cdots, b_n)$, there is an edge between α and β if and only if there is exactly one i such that $a_i \neq b_i$.

Bipartite Graphs

Definition 10

A graph $G = (V, E)$ is called *bipartite*, if:

(i) G is simple

(ii) there is a partition L and R of V such that all the edges of G are between nodes of L and the nodes of R.

Independent Set

Definition 11

Let $G = (V, E)$ be a graph and $S \subset V$, we say that S is an *independent set* of G, if there is no edges among vertices in S at all.

Independent set problem Find the independent set of maximum size in G.

Characteristics of Bipartite Graphs

Theorem 12

A simple graph is bipartite if and only if it is possible to assign one of two colores to each vertex of the graph such that no two adjacent vertices are assigned the same color.

Complete Bipartite Graphs

A *complete bipartite graph*, written $K_{l,r}$, is a graph that is partitioned into two subsets of l and r vertices X and Y such that

- both X and Y are independent sets, and
- for every two vertices $x \in X$ and $y \in Y$, there is exactly one edge between x and y.

Matching

Definition 13

Given a simple graph $G = (V, E)$ and a subset M of edges E,

1) We say that M is a *matching* of G, if there are no two edges in M that are incident with the same vertex.
2) We say that M is a *maximum matching* （最大匹配） of G, if M is a matching with the largest number of edges.
3) We say that M is a *complete matching* of G, if M is a matching of G and every vertex $v \in V$ is incident to one the the edges in M.

Maximal Matching

Definition 14

Let $G = (V, E)$ be a graph. A *maximal matching (极大匹配)* of G is a matching that cannot be enlarged by adding new edges. A *maximum matching* of G is a matching of maximum size among all the matchings in G.

Maximal Matching $\neq$ Maximum Matching

Consider:

- A path P_4 of 4 vertices
- A path P_6 of 6 vertices

M-Alternating Path

Definition 15

Given a matching M of a graph G, an *M-alternating path* is a path that alternates between edges in M and the edges outside M.

We say that an M-alternating path is an *M-augmenting path*, if the endpoints of the M-alternating path are unsaturated by M.

Given an M-augmenting path P, we can replace the edges of M in P with the other edges of P to obtain a new matching M' with one more edge.

Thus when M is a maximum matching, there is no augmenting path.

Symmetric Difference

Definition 16

For graphs G and H, the *symmetric difference* $G \triangle H$ is the subgraph of $G \cup H$ whose edges are the edges of $G \cup H$ appearing in exactly one of G and H.

We also use this notation for sets of edges; in particular, if M and M' are matchings, then $M \triangle M' = (M \setminus M') \cup (M' \setminus M)$.

Symmetric Difference of Matchings

Lemma 17

Every component of the symmetric difference of two matchings is a path or an even cycle.

Proof.

Let M and M' be matchings, and let $F = M \triangle M'$. Since M and M' are matchings, every vertex has at most one incident edge from each of them. Thus F has at most two edges at each vertex. Since $\Delta(F) \leq 2$ (the top degree), every component of F is a path or a cycle. Furthermore, every path or cycle in F alternates between edges of $M \setminus M'$ and edges of $M' \setminus M$. Thus each cycle has even length with an equal number of edges from M and from M'. □

Berge's Theorem

Theorem 18

(Berge, 1957) A matching M in G is a maximum matching in G if and only if G has no M-augmenting path.

Equivalently, we show that: G has a matching larger than M if and only if G has an M-augmenting path.
Clearly, an M-augmenting path can be used to generate a matching larger than M.

Proof - continued

For the converse, let M' be a matching in G larger than M, and let $F = M \triangle M'$. By Lemma 17, F consists of paths and even cycles, the cycles have the same number of edges from M and M'. Since $|M'| > |M|$, F must have a component with more edges of M' than of M. Such a component can only be a path that starts and ends with an edge of M'; thus it is an M-augmenting path in G.

Hall's Matching Theorem

Theorem 19

(Philip Hall, 1935) The bipartite graph $G = (V, E)$ with bipartition (V_1, V_2) has a complete matching from V_1 to V_2 if and only if $|\Gamma(A)| \geq |A|$ for all subsets A of V_1, where $\Gamma(A)$ is the set of all the neighbours of vertices of A.

Necessity Suppose that (V_1, V_2) form a complete matching between V_1 and V_2. Then for every $A \subseteq V_1$, there are at least $|A|$ neighbours of A in V_2.

Proof

Sufficiency
Suppose that (V_1, V_2) is a partition such that for any $A \subseteq V_1$, $|\Gamma(A)| \geq |A|$, where $V_1 \cup V_2 = V$ and $G = (V_1, V_2, E)$ is a bipartite graph.
We prove it by induction on $|V_1|$.
For $|V_1| = 1$. The result is obvious.
Suppose by induction that the result holds for $|V_1| = k$. Let $|V_1| = k + 1$. Consider two cases.
Case 1. For every j, $1 \leq j \leq k$ and for every $S \subseteq V_1$, if $|S| = j$, then

$$|\Gamma(S) \cap V_2| \geq j + 1. \tag{4}$$

Let $v_1 \in V_1$, $v_2 \in V_2$ be such that $(v_1, v_2) \in E$.
Let $V_1' = V_1 \setminus \{v_1\}$, $V_2' = V_2 \setminus \{v_2\}$, and $G' = (V_1', V_2', E')$, $E' = E \setminus \{(v_1, v_2)\}$.
By the assumption, and by the choices of G', for every $A' \subseteq V_1'$, $|\Gamma(A') \cap V_2'| \geq |A'|$. By the inductive hypothesis, there is a complete matching M' from V_1' to V_2' in G'.

Proof - continued

Case 2. Otherwise.
In this case, there is a $j \leq k$ such that for some $S \subseteq V_1$, $|S| = j$, and $|\Gamma(S) \cap V_2| < j + 1$ & $\geq j$, hence $= j$.
Suppose that $s_1, s_2, \cdots, s_j$ are the list of S and that $t_1, t_2, \cdots, t_j$ are the corresponding neighbours of $s_1, s_2, \cdots, s_j$.
Let $T = \Gamma(S) \cap V_2$. Then $T = \{t_1, t_2, \cdots, t_j\}$.
Let $H' = (S, T)$.
Then for any $A \subseteq S$, $|\Gamma(A) \cap V_2| \geq |A|$. Therefore, $\Gamma(A) = \Gamma(A) \cap (\Gamma(S) \cap V_2) = \Gamma(A) \cap T$.
This shows that $|\Gamma(A) \cap T| = |\Gamma(A)| \geq |A|$. By the inductive hypothesis, there is a complete matching from S to T, denoted by M_1.

Proof - continued

Let $G' = (V_1 \setminus S, V_2 \setminus T)$.
By the assumption, for any $B \subseteq V_1 \setminus S$, $|\Gamma(B) \cap V_2| \geq |B|$, noting that $\Gamma(S) = \Gamma(S) \cap V_2 = T$ and $\Gamma(B) \cap T = \emptyset$.
Therefore,

$$\Gamma(B) \cap (V_2 \setminus T) = \Gamma(B) \cap V_2,$$

which contains at least $|B|$ elements.
By the inductive hypothesis, there is a complete matching M_2 from $V_1 \setminus S$ to $V_2 \setminus T$.
Let $M = M_1 \cup M_2$. Then M is a complete matching from V_1 to V_2.

Induced subgraph

Given a graph $G = (V, E)$ and a subset $S \subset V$, the *induced subgraph of S in G* is the graph with vertices S and edges of E whose two endpoints are both in S.

Operations

There are many operations for graphs and in graphs:

- Edge contraction
- removal of vertices or edges
- graph union
- graph reduction
- interaction
- communication
- transportation
- virus spreading

Applications - Local Area Networks

- star
- cycle
- star $+$ cycle
- new structures, optimal structures? what are the principles? (Big challenge)

Applications - Parallel Computation

Sequential vs **Parallel**
Basic structures of parallel computation:

- K_n
- Path
- Grid
- Q_n, n-cube

Requirements: Most works are done by individual processors, and the interactions among the different processors are small. What is the mathematical measure for these intuitive requirements? (A big challenge)

Open Questions

1. What are the principles for the network of computing systems?
2. What is the correct model of cloud computing?

Data Structure

Classic data structures:

- linear ordering
- grid
- trees
- table

Open Questions

1. What is the principle of classical data structure? Why?
2. What is the principle for the structure of big data?

Exercises - 1

(1) For any simple graph of 5 vertices, either there is a clique of size 3, or there is an independent set of size 3.

(2) Let $G = (V, E)$ be a connected, undirected multigraph with n vertices. A cut C of G is a set of edges of G whose removal results in G being disconnected. The minimum cut problem is to find the cut of the least number of edges in G. Let k be the size of the minimum cut of G. Show that

0.1 The least degree of vertices of G is at least k
0.2 The number of edges in G is at least $\frac{kn}{2}$.

Define the operation of contraction as follows: Pick an edge uniformly at random and merge the two endpoints of the edge together.
Parallel edges are always kept.

(3) An edge contraction does not reduce the size of the minimum cut.

Exercises - 2

(4) For the min-cut problem. We run the edge contraction until the resulted graph has only two vertices, which gives rise a cut of G, that is, the parallel edges between the last two vertices.
Let C be a minimum cut of G.
For every i, $1 \leq i \leq n-2$, let $\mathcal{E}_i$ be the event that: not picking an edge of C at the i step of the execution of the contractions.
Show that

$$\Pr \mathcal{E}_1] \geq 1 - \frac{2}{n}.$$

(5) Show that

$$\Pr[\cap_{i=1}^{n-2} \mathcal{E}_i] > \frac{2}{n^2}.$$

谢谢！