PB165 – Grafy a sítě

Základní pojmy teorie grafů

Obsah předmětu

1. Základní pojmy teorie grafů
2. Stromy
3. Kostra grafu, algoritmy nalezení
4. Prohledávání grafu, hledání nejkratší cesty
5. Toky v sítích
6. Algoritmy směrování a přepínání
7. Peer-to-peer sítě

Proč právě grafy a sítě?

- Teorie grafů je důkladně rozpracována a nabízí mnoho využitelných algoritmů.
- Graf představuje velmi přímočarou reprezentaci sítě na všech úrovních OSI modelu, např.:
  - Síťové prvky a jejich fyzické propojení na nejnižší vrstvě,
  - síťové aplikace a TCP spojení mezi nimi na transportní vrstvě,
  - procesy distribuovaného výpočtu a jejich komunikace na aplikační vrstvě.
- Grafy nacházejí využití i v návrhu síťových protokolů a jejich formální verifikaci.

Pojem grafu

Graf je abstraktní pojem matematiky a informatiky užitečný pro modelování reálných objektů, situací. Intuitivně se graf skládá z:

- *Vrcholů (uzlů)*, znázorňovaých schematicky jako ,,body",
- *hran* spojujících vrcholy.

Co lze reprezentovat grafem?

- Mapu měst a silničního spojení,
- atomy v molekule a jejich vazby,
- vodovodní, elektrické sítě

a zejména

- počítačové sítě.

Označován také jako *jednoduchý* graf.

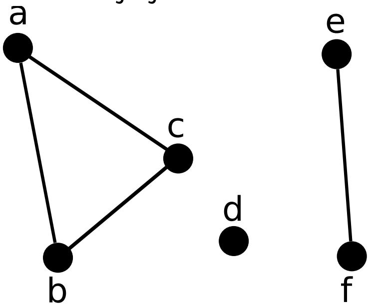
Formální definice grafu

Definice

Graf G je uspořádaná dvojice množin (V, E), kde

- *V je množina vrcholů a*
- *E je množina hran – dvouprvkových podmnožin V*

Vrcholy spojené hranou se nazývají *sousední*. Hrana se označuje jako *incidentní* k vrcholům, které spojuje.



$$V = \{a, b, c, d, e, f\}$$
$$E = \{\ \{a, b\}, \{a, c\}, \{b, c\}, \{e, f\}\ \}$$

Orientovaný graf

Hrany v grafu, jak byly definovány, spojují dva rovnocenné vrcholy. Takové grafy se také nazývají *neorientované*. U hran ovšem můžeme vyznačit směr, kterým vedou – hrany i graf se poté nazývají *orientované*.

Definice

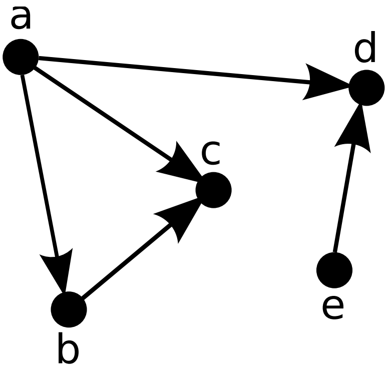
Graf, jehož hrany jsou uspořádané dvojice vrcholů, se nazýva orientovaný.

O hraně (u, v) říkáme, že vychází z vrcholu u a vstupuje do v. Graficky se orientace hrany označí šipkou ve směru, kterým hrana vede.

Orientovaný graf - příklady

Kde najdeme orientované grafy v sítích?

- Webové stránky – graf odkazů
- DNS – hierarchická struktura domén, serverů
- Směrování – graf cest paketů k cíli



$$V = \{a, b, c, d, e\}$$
$$E = \{ (a, b), (a, c), (b, c), (a, d), (e, d) \}$$

Multigraf

Definice grafu povoluje nejvýše 1 hranu mezi každou dvojicí vrcholů a požaduje, aby hrana spojovala různé vrcholy. Tato omezení odstraňuje multigraf:

Definice

Multigraf je graf, jenž nahrazuje množinu hran multimnožinou (smí obsahovat násobné prvky) a umožňuje existenci smyček – hran spojujícíh vrchol sám ze sebou.

Multigraf lépe odpovídá reálným fyzickým sítím, kde se často vyskytují redundantní linky.
Smyčky mohou znázornit např. loopback – rozhraní přijímající zprávy, které samo vysílá.

Ohodnocení grafu

Vrcholům a hranám je možné přiřadit např. číslo či barvu.

Definice

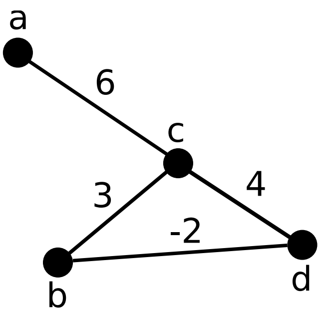
Přiřazení prvků konečné množiny vrcholům či hranám grafu nazýváme jejich ohodnocením.

Příklady ohodnocení na sítích:

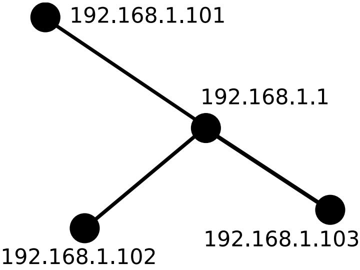
- Ohodnocení síťových zařízení L2 a L3 adresami (viz ARP protokol)
- Šířka pásma, latence, cena přenosu za jednotku dat jako ohodnocení linek – hran
- Ohodnocení vrcholů názvem stavu, hran typem zprávy při abstraktním návrhu protokolů (MSC[1])

[1] Message Sequence Charts

Ohodnocení – příklady



Hranově ohodnocený graf



Vrcholově ohodnocený graf

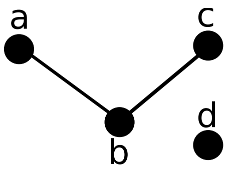
Stupeň vrcholu

Definice

Stupněm vrcholu v neorientovaném grafu nazýváme počet hran incidentních k vrcholu.

- Počet klientů připojených k Wi-Fi AP
- Počet uzavřených spojení spojovaného protokolu (např. TCP)

Stupeň vrcholu u značíme $deg(u)$.



$$deg(a) = deg(c) = 1$$
$$deg(b) = 2$$
$$deg(d) = 0$$

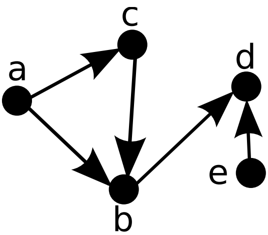
Stupeň vrcholu v orientovaném grafu

V případě orientovaného grafu rozlišujeme *vstupní* a *výstupní* stupeň.

Definice

Vstupním (výstupním) stupněm vrcholu u orientovaného grafu nazýváme počet hran vstupujících do, resp. vycházejících z, vrcholu u a značíme jej $deg^{+}(u)$, *resp.* $deg^{-}(u)$.

- V grafu odkazů mezi webovými stránkami reprezentuje vstupní vrchol počet odkazů na stránku vedoucích.



vrchol	deg^{-}	deg^{+}
a	2	0
b	1	2
c	1	1
d	0	2
e	1	0

Sledy v grafu

Definice

Sledem v grafu (neorientovaném grafu) nazýváme posloupnost vrcholů a hran

$$v_0, e_1, v_1, \ldots, e_n, v_n,$$

kde každá hrana e_i spojuje vrcholy v_{i-1}, v_i, resp. vede z v_{i-1} do v_i.

Sled v grafu je tedy „trasou", na které se mohou vrcholy i hrany opakovat. Se sledy se lze setkat i v reálných sítích:

- Cesta paketu sítí (některé směrovací algoritmy nezabraňují zacyklení paketu v průběhu výpočtu).
- Transakce v SIP protokolu (viz SIP ellipse).

Samostatný vrchol je také sledem.

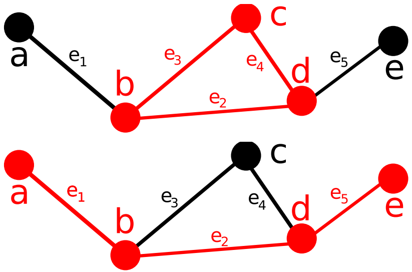
Cesty v grafu

Definice

Cesta v grafu je sled bez opakování vrcholů.

V cestě se v důsledku neopakování vrcholů nemohou opakovat ani hrany.

- Cesty definující směrování paketů mezi dvojicemi síťových prvků.



b, e_3, c, e_4, d, e_2, b je sledem v grafu, ale nikoliv cestou – vrchol b se opakuje.

a, e_1, b, e_2, d, e_5, e je sledem i cestou v grafu.

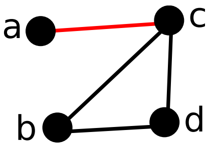
Souvislost grafu

Definice

Neorientovaný graf se nazývá souvislý, pokud mezi každými dvěma jeho vrcholy vede cesta.

Definice

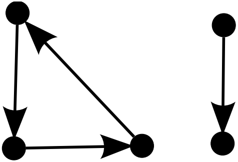
Nahradíme-li všechny hrany orientovaného grafu G neorientovanými a získáme-li tak souvislý graf, G je slabě souvislý.
Orientovaný graf je silně souvislý, pokud mezi každými dvěma jeho vrcholy vedou cesty v obou směrech.

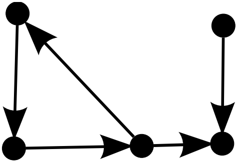


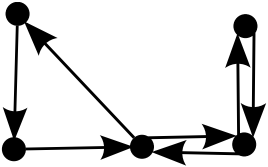
Tento graf je souvislý; po odebrání vyznačené hrany by souvislý nebyl, odebrání jedné z nevyznačených hran by jeho souvislost zachovalo.

Souvislost grafu – příklady

- Internet na fyzické vrstvě tvoří souvislý graf. (?)
- Internet na IP vrstvě tvoří slabě souvislý (ovšem silně nesouvislý) orientovaný graf (adresy za NAT).
- Orientovaný graf webových stránek není silně ani slabě souvislý.







Grafy na obrázcích jsou:

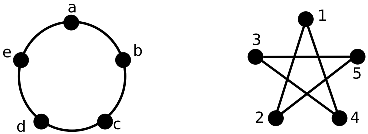
1. Nesouvislý
2. Slabě souvislý
3. Silně souvislý

Grafy, lišící se např. nakreslením, označením vrcholů a hran, ohodnocením, se nemusejí lišit svojí strukturou – mohou být isomorfní.

Definice

Isomorfismus mezi grafy G, H je bijektivní zobrazení vrcholů, které zachovává hrany – tj. pokud vede v grafu G hrana mezi vrcholy u, v, pak v grafu H vede hrana mezi vrcholy $f(u), f(v)$. Pokud mezi grafy G, H existuje isomorfismus, nazývají se isomorfní.

Isomorfismus (jelikož je relací ekvivalence) tak definuje třídy grafů, které lze považovat za totožné.



Isomorfismus grafů

- Co musejí isomorfní grafy splňovat?
  - Mít stejný počet vrcholů.
  - Mít stejný počet hran.
  - Jejich vrcholy musejí mít stejné stupně.
- V mnoha případech je snadné dokázat, že grafy isomorfní nejsou pomocí těchto (a některých dalších) invariantů. Jsou-li tyto invarianty shodné, je nutno vyloučit všechny možné isomorfismy.
- Důkaz isomorfie dvou grafů vyžaduje přímo nalezení konkrétního isomorfismu mezi těmito grafy.

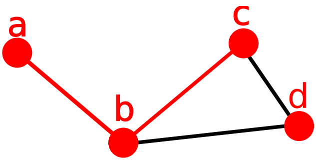
Podgrafy

Definice

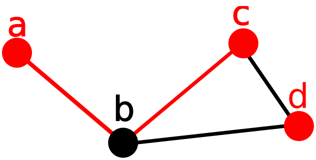
Graf H je podgrafem grafu G, pokud platí následující podmínky:

1. *Vrcholy grafu H tvoří podmnožinou vrcholů grafu G.*
2. *Hrany grafu H tvoří podmnožinou hran grafu G.*
3. *Hrany grafu H mají oba vrcholy v H.*

Graf G je poté nadgrafem grafu H.



Vyznačené vrcholy a hrany tvoří podgraf.



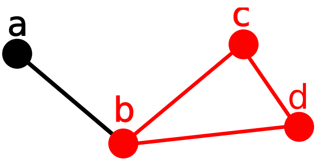
Vyznačené vrcholy a hrany netvoří podgraf.

Indukované podgrafy

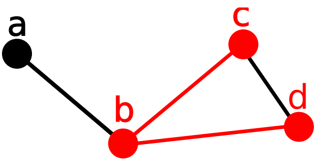
Definice

Podgraf H se nazývá indukovaný, pokud obsahuje všechny hrany, které mezi jeho vrcholy vedou v nadřazeném grafu G.

- Lokální síť je indukovaným podgrafem Internetu na fyzické vrstvě.
- Samostatné vrcholy libovolného grafu tvoří jeho podgraf.



Vyznačené vrcholy a hrany tvoří indukovaný podgraf.



Vyznačené vrcholy a hrany netvoří indukovaný podgraf.

Jak efektivně uložit graf v paměti počítače či aktivního síťového prvku?

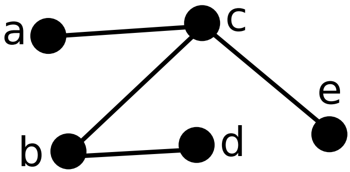
Vrcholy označíme čísly $1 \dots n$. Pro uložení hran máme dvě základní možnosti:

- *Matice sousednosti*
- *Seznamy sousedů*

Matice sousednosti

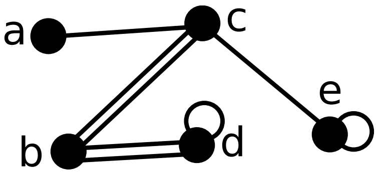
- Matice E o rozměrech $n \times n$ pro n vrcholů grafu
- $E_{ij} = 1$ pokud hrana (i, j) patří do grafu
- $E_{ij} = 0$ jinak
- Pro neorientovaný graf je symetrická, pro orientovaný nemusí

Matice sousednosti



	a	b	c	d	e
a	0	0	1	0	0
b	0	0	1	1	0
c	1	1	0	0	1
d	0	1	0	0	0
e	0	0	1	0	0

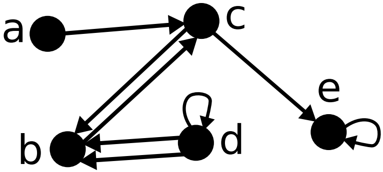
V této podobě je matice sousednosti vhodná jen pro reprezentaci jednoduchého grafu. Multigraf lze reprezentovat maticí sousednosti, jejíž prvky udávají počet hran mezi každými dvěma vrcholy:



	a	b	c	d	e
a	0	0	1	0	0
b	0	0	2	2	0
c	1	2	0	0	1
d	0	2	0	1	0
e	0	0	1	0	1

Obdobně lze ukládat jednoduchý hranově ohodnocený graf.

Matice sousednosti pro orientovaný multigraf



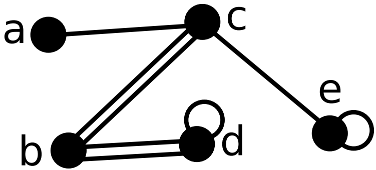
$$\begin{array}{c|ccccc} & a & b & c & d & e \\ a & 0 & 0 & 1 & 0 & 0 \\ b & 0 & 0 & 1 & 0 & 0 \\ c & 0 & 1 & 0 & 0 & 1 \\ d & 0 & 2 & 0 & 1 & 0 \\ e & 0 & 0 & 0 & 0 & 1 \end{array}$$

Seznamy sousedů

Pro každý vrchol existuje samostatný seznam sousedů, s nimiž je tento spojen hranou (či do nich vede orientovaná hrana).

- Lze implementovat pomocí 2 jednorozměrných polí:
  - V jednom jsou uloženy všechny seznamy za sebou, seřazené podle čísla vrcholu
  - Druhé uchovává indexy, na kterých začínají v prvním poli sousedé každého vrcholu
- Násobné hrany v multigrafu jsou zadány násobným uvedením vrcholu v seznamu sousedů
- Pro „řídké" grafy (výrazně méně než n^2 hran) jsou seznamy sousedů výhodnější z hlediska paměťové náročnosti než matice sousednosti. Takových grafů je mezi sítěmi většina.

Seznamy sousedů – příklad



Pole indexů do seznamu sousedů:

a	b	c	d	e
1	2	6	10	13

Seznam sousedů:

1	2	3	4	5	6	7	8	9	10	11	12	13	14
c	c	c	d	d	a	b	b	e	b	b	d	c	e

Cvičení

1. Nakreslete všechny neisomorfní grafy na:
   1. 3 vrcholech
   2. 4 vrcholech
   3. 5 vrcholech
2. Dokažte, že platí $\sum_{y \in V(G)} deg(y) = 2|E(G)|$ pro neorientovaný graf.
3. Mohou dva grafy, orientovaný a neorientovaný, mít stejné matice sousednosti a zároveň stejné počty hran? Jak takové grafy vypadají?
4. Uvažme orientovaný graf, který reprezentuje relace na množině vrcholů tak, že hrana spojuje právě ty prvky, které jsou spolu v relaci. Jakou strukturu bude mít graf reprezentující:
   1. tranzitivní relaci
   2. relaci ekvivalence