Protok u mrežama

(*Flow in Networks*)

Problem: pokušajmo izračunati najveći mogući protok vozila i opterećenje pojedinih prometnica od Karlovca do Splita u vrijeme ljetnih gužvi.

Dva su glavna pravca: KA-Gospić-Zadar-Split (A1)

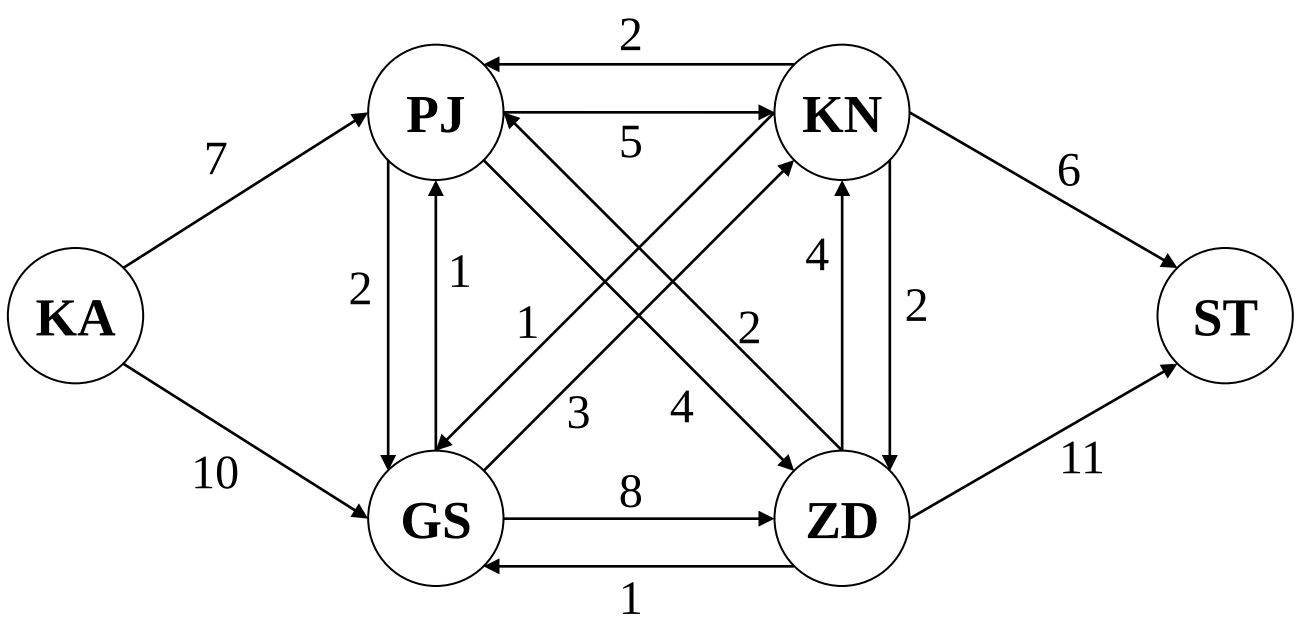
i KA-Plitvička jezera-Knin-Split (cesta D1).

Kapacitet (najveća moguća propusnost) pojedinog pravca modelirat će istodobno njegov stvarni kapacitet i vjerojatnost prolaska u određenom smjeru (poželjnost smjera).

Npr. neka je stvarni kapacitet pravca Zadar-Knin (i obratno) =4. Ako je Split odredište, a dolazi se iz Zagreba, put Zadar → Knin → Split nije najbolje rješenje (nije najkraći, dakle ni najpoželjniji), ali je svakako bolji od puta Knin → Zadar → Split. Zato ćemo smjeru Zadar-Knin pridijeliti kapacitet =4, a smjeru Knin-Zadar kapacitet =2.

Cjeloviti model:

	GS	PJ	ZD	KN	ST
KA	10	7	-	-	-
GS	-	1	8	3	-
PJ	2	-	4	5	-
ZD	1	2	-	4	11
KN	1	2	2	-	6



Sličan bi problem bilo i optimiranje npr. vodovodne ili elektroenergetske mreže. Ako je poznata struktura mreže, pitanje je koliko najviše vode (električne energije) možemo dovesti iz vodocrpilišta (elektrane) do potrošača te koliko je pritom opterećenje pojedinih opskrbnih pravaca, odnosno dijelova mreže.
Komunikacijske mreže su identičan primjer, itd.

Postavljeni problem nije trivijalan. Ipak, rješenje je jednostavno i intuitivno, ali tek ako znamo teorijska svojstva protočnih mreža (*flow networks*)...
U nastavku ćemo ih nazivati jednostavno mreže.

Mreža je povezani usmjereni graf $G=(V,E)$ u kojemu svi bridovi $(u,v) \in E$ imaju nenegativne težine $w_{uv} \geq 0$, tj. kapacitete $c(u,v) \geq 0$.

Dva su vrha posebna: polazište *s* (izvor; *source*) i
odredište *t* (slijev; *sink*).

Uvijek se pretpostavlja da je graf povezan u smjeru $s \to t$, tj. da je svaki vrh $v \in V$ na nekom putu od *s* do *t*.

Iz toga slijedi da je u mrežama uvijek $|E| \geq |V| - 1$.

Najvažniji pojam je protok u mreži kao cjelini, koji je intuitivno jasan. U našem primjeru to bi bio ukupni priljev vozila u Split.

No, sustavno (formalno) rješenje problema zahtijeva precizniju (matematičku) definiciju toka i protoka ...

Tok je funkcija f: V × V → R koja ima sljedeća tri svojstva:

1. ograničenost kapacitetom (*Capacity Constraint*)
   - za sve u,v∈V mora biti f(u,v) ≤ c(u,v)
2. neparnost (*Skew Symmetry*)
   - za sve u,v∈V mora biti f(u,v) = –f(v,u)
3. zadovoljava "zakon o očuvanju toka" (*Flow Conservation Property*)
   - za svaki vrh grafa (čvor mreže), osim za izvor i slijev, mora vrijediti da je izlazni tok jednak ulaznom toku (prvi Kirchofov zakon).

$$\forall u \in V - \{s,t\} \quad \sum_{v \in V} f(u,v) = 0$$

- zakon o očuvanju toka može se pisati i kao

$$\forall v \in V - \{s,t\} \quad \sum_{u \in V} f(u,v) = 0$$

Pozor! Definicijska svojstva toka unose slabo zamjetljivu, ali važnu razliku između tokova u mreži i transportnih modela kao u našem primjeru!

Npr. promet intenziteta 4 u smjeru Zadar-Knin i intenziteta 2 u smjeru Knin-Zadar nije tok u mreži jer $f(ZD,KN) \neq -f(KN,ZD)$. Ipak, ako je važno znati samo protok u mreži, a ne i opterećenje pojedinih smjerova, nejednak dvosmjerni promet se može modelirati funkcijom toka uvažavajući međusobna poništavanja već u modelu. Računamo kao da je promet jednosmjeran u smjeru ZD-KN, ali intenziteta $f(ZD,KN) - f(KN,ZD) = 2$ pa onda možemo postaviti $f(ZD,KN) = 2 = -f(KN,ZD)$ i zadovoljiti definicijska svojstva toka. Takvom transformacijom se svaki dvosmjerni promet može modelirati tokovima u mreži, što će omogućiti optimiranje mreže i izračunavanje najvećeg mogućeg protoka, ali više nećemo znati je li stvarni promet jednosmjeran ili dvosmjeran. Ako je to važno znati, treba mijenjati pristup problemu, tj. tok u mreži nije dobar model.

Protok ili jakost toka (*flow value*) mreže je ukupni izlazni tok iz izvora i označava se s $|f|$.

$$|f| = \sum_{v \in V} f(s,v)$$

Pozitivni ulazni tok u vrh ***v***: $\sum_{\substack{u \in V \\ f(u,v)>0}} f(u,v)$

Pozitivni izlazni tok iz vrha ***u***: $\sum_{\substack{v \in V \\ f(u,v)>0}} f(u,v)$

Protok ili neto-tok nekim vrhom

= pozitivni izlazni tok – pozitivni ulazni tok.

Zbog preglednosti zapisa kad se radi s tokovima u mreži, uvodi se **implicitno zbrajanje** (*implicit summation*). **Implicitno zbrajanje** je zapis u kojem jedan ili oba argumenta funkcije toka mogu biti podskupovi vrhova (čvorova) mreže, a tumači se kao rezultat funkcije nakon računanja (zbrajanja) sa svim mogućim parovima elemenata iz skupova argumenata. Argumenti se u tom zapisu označavaju velikim slovom ako su skupovi, a malim ako se radi o jednom vrhu. Na primjer, ako su X i Y neki podskupovi skupa vrhova V, onda zapis f(X,Y) znači

$$f(X,Y) = \sum_{u \in X} \sum_{v \in Y} f(u,v) \quad .$$

Zakon o očuvanju toka glasi: $\forall u \in V - \{s,t\}$, $f(u,V)=0$.

Implicitno zbrajanje omogućuje jednostavan zapis relacija među tokovima. Sljedeća lema navodi neka svojstva toka u mreži.

Lema 1: Ako je funkcija f tok u mreži $G=(V,E)$, vrijedi:

1. Za sve $X \subseteq V$ je $f(X,X) = 0$.
2. Za sve $X,Y \subseteq V$ je $f(X,Y) = -f(Y,X)$.
3. Za sve $X,Y,Z \subseteq V$ za koje je $X \cap Y = \{\ \}$,
   zbroj $f(X \cup Y, Z) = f(X,Z) + f(Y,Z)$
   i zbroj $f(Z, X \cup Y) = f(Z,X) + f(Z,Y)$.

Dokažite to za vježbu!

Kao primjer računanja s implicitnim zbrajanjem, dokazat ćemo intuitivno jasnu činjenicu koju smo do sada aksiomatski prihvaćali: ukupni izlazni tok iz izvora, tj. jakost toka $|f|$, jednak je ukupnom ulaznom toku (priljevu, dotoku) u slijev, što formalno (implicitnim zbrajanjem) zapisujemo kao $|f| = f(V,t)$.

Dokaz:

$|f| = f(s,V)$ //po definiciji

$= f(V,V) - f(V - s,V)$ //lema 1, točka 3

$= - f(V - s,V)$ //lema 1, točka 1

$= f(V,V - s)$ //lema 1, točka 2

$= f(V,t) + f(V,V - s - t)$ //lema 1, točka 3

$= f(V,t)$ //jer je $f(V,V - s - t) = 0$ zbog zakona o očuvanju toka. ■

Analiza mreža i formalno rješavanje problema protoka zahtijeva još uvođenje sljedeća tri pojma:

1. Preostala mreža ili ostatak mreže (*Residual Network*)
2. Dopunski put ili put proširenja (*Augmenting Path*)
3. Presjek mreže (*Cut*)

Preostala mreža (*Residual Network*) je mreža sastavljena od bridova čiji kapacitet još nije u potpunosti iskorišten, dakle od bridova koji mogu preuzeti na sebe dodatni tok.

Preostali kapacitet (*Residual Capacity*) je kapacitet bridova preostale mreže i po definiciji se računa kao razlika izvornog kapaciteta pojedinog brida i njegovog trenutačnog toka. Formalno, ako je $f(u,v)$ funkcija toka u mreži $G=(V,E)$, preostali kapacitet je

$$c_f(u,v) = c(u,v) - f(u,v).$$

S obzirom na definiciju, preostali kapacitet c_f nekog brida uvijek je pozitivan, a može biti i veći od izvornog kapaciteta.

$$f(u,v) \leq c(u,v) \quad \Rightarrow \quad c_f(u,v) = c(u,v) - f(u,v) \geq 0$$
$$f(u,v) < 0 \quad \Rightarrow \quad c_f(u,v) = c(u,v) - f(u,v) > c(u,v)$$

Preostali kapacitet veći od izvornog znači da je tok u bridu (u,v) negativan, odnosno da imamo tok u smjeru (v,u), pa povećanje kapaciteta iznad stvarne veličine odražava mogućnost da prvo poništimo negativni (suprotni) tok i onda još dodamo c(u,v) toka radi iskorištavanja cijelog raspoloživog kapaciteta.

Precizna formalna definicija preostale mreže temelji se na preostalom kapacitetu.

Definicija: Neka je polazna mreža G=(V,E) i tok u njoj f(u,v). Preostala mreža, za koju se kaže da je *inducirana* tokom f, je mreža $G_f = (V,E_f)$ čiji su bridovi $E_f = \{ (u,v) \in V \times V: \ c_f(u,v) > 0 \}$.

Definicija kazuje da su kapaciteti bridova preostale mreže jednaki preostalim kapacitetima, a znamo da su oni uvijek pozitivni. Štoviše, znamo da će bridovi preostale mreže uvijek biti ili bridovi polazne mreže (ako im kapacitet nije u potpunosti iskorišten) ili njima suprotno usmjereni bridovi (čim postoji tok), a mogu postojati samo ako postoji barem jednosmjerna veza dvaju vrhova u polaznoj mreži. Zaključujemo da preostala mreža može imati najviše $2 \cdot |E|$ bridova, tj. vrijedi $|E_f| \leq 2 \cdot |E|$.

Preostala mreža je opet mreža i to u čvrstoj vezi (matematičkom odnosu) s polaznom mrežom. Tu ćemo vezu iskoristiti za rješavanje problema protoka, stoga ju moramo precizno opisati i dokazati.

Lema 2: Neka je G=(V,E) mreža s izvorom *s* i slijevom *t*, a *f* tok u toj mreži. Nadalje, neka je G_f preostala mreža mreže G inducirana tokom *f* i neka je *f* ' tok u G_f. Tada je zbroj tokova $f + f'$, koji se definira relacijom $(f + f')(u,v) = f(u,v) + f'(u,v)$, također tok u mreži G, pri čemu je jakost tog toka $|f + f'| = |f| + |f'|$. □

Dokaz:

Prvo treba dokazati da je $f + f'$ valjana funkcija toka u G, dakle da zadovoljava definicijska svojstva, a potom i da vrijedi $|f + f'| = |f| + |f'|$.

Dokaz neparnosti:

$$
\begin{aligned}
(f + f')(u,v) &= f(u,v) + f'(u,v) \\
&= -f(v,u) - f'(v,u) \\
&= -(f(v,u) + f'(v,u)) \\
&= -(f + f')(v,u)
\end{aligned}
$$

Dokaz ograničenosti kapacitetom:

$$
\begin{aligned}
(f + f')(u,v) &= f(u,v) + f'(u,v) \leq f(u,v) + c_f(u,v) \\
&= f(u,v) + (c(u,v) - f(u,v)) = c(u,v)
\end{aligned}
$$

Dokaz podvrgavanja zakonu o očuvanju toka:

Za svaki $(u \in V) - \{s,t\}$ vrijedi

$$(f + f')(u,V) = f(u,V) + f'(u,V) = 0 + 0 = 0$$

Dakle, funkcija $(f + f')$ jest tok u G.

Dokaz da je $|f + f\,'| = |f\,| + |f\,'|$:

$$
\begin{aligned}
|f + f\,'| &= (f + f\,')(s,V) \\
&= f(s,V) + f\,'(s,V) \\
&= |f\,| + |f\,'| \qquad \blacksquare
\end{aligned}
$$

Dopunski put (*Augmenting Path*) je bilo koji jednostavni put (*simple path*) od *s* do *t* u preostaloj mreži G_f.

Iz definicije preostale mreže i leme 2 slijedi da u svakom bridu polazne mreže G koji se nalazi na dopunskom putu možemo povećati tok, a da ne premašimo (izvorni) kapacitet brida. Naravno, najveće moguće povećanje toka na nekom putu ograničeno je najmanjim kapacitetom na tom putu u preostaloj mreži.

Najveće moguće povećanje toka na dopunskom putu p nazivamo preostalim kapacitetom puta p i označavamo s $c_f(p)$. Formalno, $c_f(p)$ je

$$c_f(p) = \min\{\ c_f(u,v):\ (u,v) \in p\ \} \quad .$$

Lema 3: Neka je $G=(V,E)$ mreža, f tok u toj mreži i p dopunski put u mreži G_f induciranoj tokom f. Funkcija f_p: $V \times V \rightarrow R$ definirana s

$$f_p(u,v) = \begin{cases} c_f(p) & ; \quad (u,v)\ is\ on\ p \\ -c_f(p) & ; \quad (v,u)\ is\ on\ p \\ 0 & ; \quad else \end{cases}$$

je tok jakosti $|\, f_p \,| = c_f(p) > 0$ u G_f. □

Dokaz za vježbu.

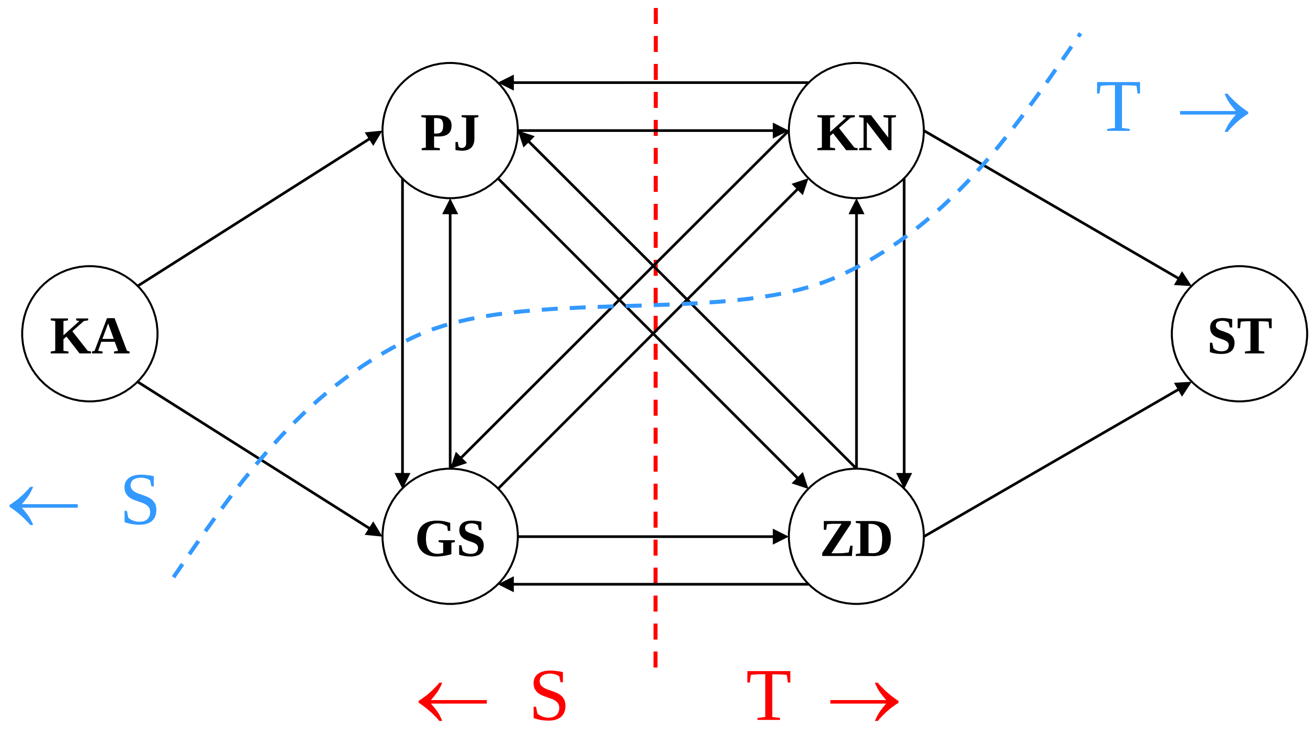
Sljedeći korolar potvrđuje intuitivno jasnu činjenicu da dodamo li f_p toku f dobivamo novi tok u G koji će biti veće jakosti od toka f, dakle bliži najvećem mogućem.

Korolar 1: Neka je G=(V,E) mreža, funkcija f tok u toj mreži, put p dopunski put u mreži G_f induciranoj tokom f i funkcija f_p definirana kao u lemi 3. Funkcija f' definirana s $f' = f + f_p$ je tok jakosti $|f'| = |f| + |f_p| > |f|$ u G. □

Dokaz: ovaj korolar izravno slijedi iz lema 2 i 3.

Presjek (*Cut*) **mreže** G=(V,E) su dva dijunktna podskupa S i T skupa vrhova V takvi da $s \in S$ i $t \in T$.

Dakle, vrijedi V = T + S. Presjek ćemo označavati s (S,T).



Definiraju se dva pojma:

protok ili neto-tok presjekom (S,T)

i kapacitet presjeka c(S,T).

Protok (neto-tok**) presjekom (S,T)** je tok *f*(S,T), dakle zbirni tok iz svih vrhova u S u sve vrhove u T, pri čemu se tokovi iz vrhova u T prema vrhovima u S oduzimaju (pribrajaju s negativnim predznakom).

Kapacitet presjeka (S,T) je zbirni kapacitet svih veza (bridova) iz S u T.

Bridovi iz T u S se ne pribrajaju kapacitetu presjeka jer to nema smisla kada je problem naći maksimalni tok iz izvora u slijev! Osim toga, kapaciteti bridova iz T u S bi trebali biti negativni, a kapaciteti su po definiciji pozitivne veličine.

Lema 4: Neka je G=(V,E) mreža s izvorom *s* i slijevom *t*. Protok $f(S,T)$ bilo kojim presjekom (S,T) u mreži jednak je protoku (jakosti toka) $|f|$ u cijeloj mreži, tj. za svaki presjek vrijedi $f(S,T) = |f|$. □

Dokaz:

$$\begin{aligned} f(S,T) &= f(S,V-S) = f(S,V) - f(S,S) \\ &= f(S,V) = f(s,V) + f(S-s,V) \\ &= f(s,V) = |f| \qquad \blacksquare \end{aligned}$$

Ovaj rezultat (lema 4) je vrlo važan za rješavanje problema protoka u mrežama jer potvrđuje intuitivno jasnu činjenicu da je protok u svim presjecima mreže isti te da je jednak jakosti toka (protoku) cijele mreže $|f|$.

Budući da su svi protoci u mreži isti i jednaki jakosti toka cijele mreže, zaključujemo sljedeće:

Korolar 2: Protok u mreži ograničen je odozgo kapacitetom bilo kojeg presjeka u njoj. □

Dokaz: po lemi 4 znamo da je | *f* | = *f*(S,T) za bilo koji presjek pa slijedi

$$\left|f\right| = f(S,T) = \sum_{u \in S} \sum_{v \in T} f(u,v)$$

$$\leq \sum_{u \in S} \sum_{v \in T} c(u,v) = c(S,T) \quad \blacksquare$$

(!) Izravna posljedica ovog korolara jest da je najveća moguća jakost toka u mreži ograničena najmanjim kapacitetom nekog presjeka.

Ako je najmanji kapacitet gornja granica, naslućujemo da je najveći mogući tok u mreži zapravo jednak najmanjem kapacitetu nekog presjeka. To potvrđuje Max-flow Min-cut teorem.

Teorem: Max-flow Min-cut
Ako je G=(V,E) mreža s izvorom *s* i slijevom *t*, a *f* tok u njoj, sljedeće su tvrdnje ekvivalentne, tj. ili su sve (istodobno) istinite ili sve neistinite:

1. Tok *f* je najveći mogući tok u G (dakle tok za koji je jakost toka, tj. protok najveći mogući).
2. U preostaloj mreži G_f nema dopunskih puteva.
3. Jakost toka $|f|$ je jednaka kapacitetu c(S,T) nekog presjeka u G. □

Dokaz: treba dokazati da vrijedi (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1)

(1) $\Rightarrow$ (2): Recimo da (1) nije dovoljan uvjet za (2), tj. da iz (1) ne slijedi nužno i (2). Dakle, neka je (1), ali neka postoji dopunski put *p* u G_f. No, tada je prema korolaru 1 tok $f+f_p$ (f_p definirana u lemi 3) također tok u G i to veći od $|f|$, što je u suprotnosti s pretpostavkom da je *f* najveći mogući tok.

(2) $\Rightarrow$ (3): Ako u G_f nema dopunskog puta, znači nema puta od *s* do *t* i možemo razdijeliti mrežu na dva dijela (podskupa) S i T, od kojih jedan (S) sadrži izvor, a drugi (T) slijev. Za sve bridove iz S u T sigurno je $f(u,v) = c(u,v)$ jer u protivnom bi postojao dopunski put. Skupovi S i T čine presjek (S,T) pa znamo da je $|f| = f(S,T)$, a zbog $f(u,v) = c(u,v)$ je i $f(S,T) = c(S,T)$ pa je $|f| = c(S,T)$.

(3) $\Rightarrow$ (1): Po korolaru 2 znamo da je $|f| \leq c(S,T)$ za bilo koji presjek (S,T). Protok $f(S,T)$ presjekom (S,T) koji dijeli mrežu kada nema dopunskog puta upravo je jednak c(S,T), a po lemi 4, jakost toka $|f|$ je jednaka protoku bilo kojeg presjeka, pa tako i protoku $f(S,T)$. Dakle, kada je $|f| = f(S,T) = c(S,T)$ protok je najveći mogući. ■

Max-flow Min-cut teorem je osnova Ford-Fulkersonove metode za rješavanje problema protoka u mrežama. Za tu osnovnu ideju kažemo da je metoda, a ne algoritam jer se može provesti na razne načine koji vode u algoritme različitih složenosti.

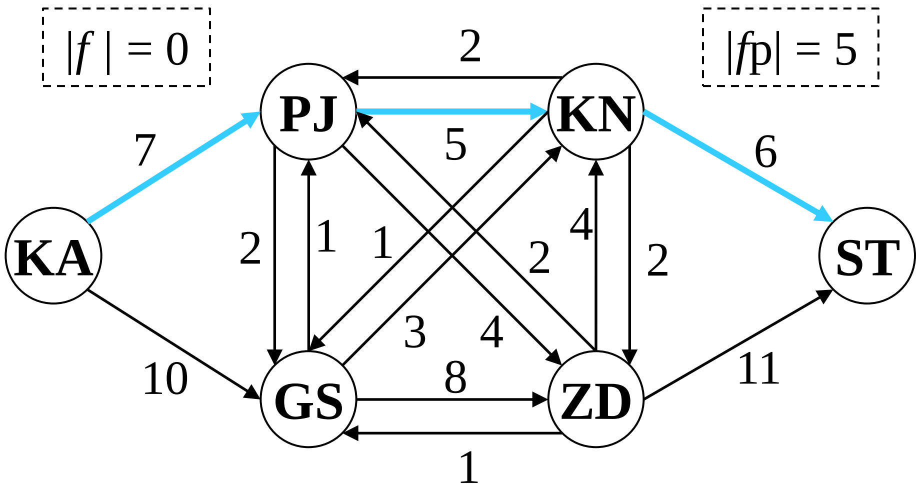
Ford-Fulkersonova metoda

inicijalizacija: *tokovi svih bridova* f(u,v) = 0
`while` *u mreži postoji dopunski put* `p`
    uvećati sve tokove na putu `p` *za* f_p

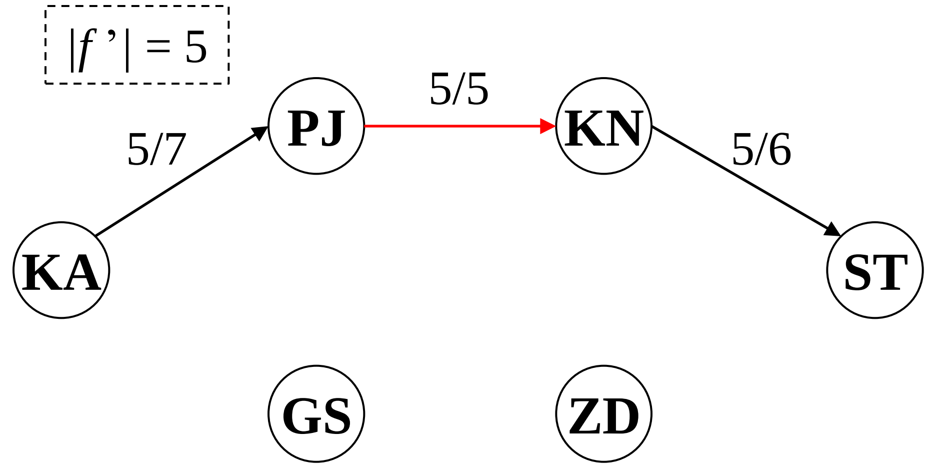
Napomena: f_p je funkcija definirana u lemi 3, dakle preostali kapacitet na putu *p*, odnosno najveće moguće povećanje toka na tom putu.

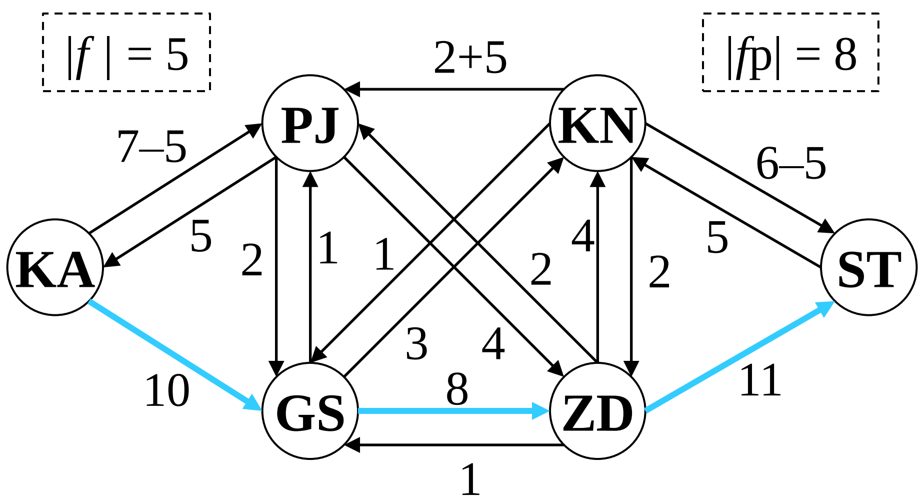
Primjer: rješavanje polaznog problema

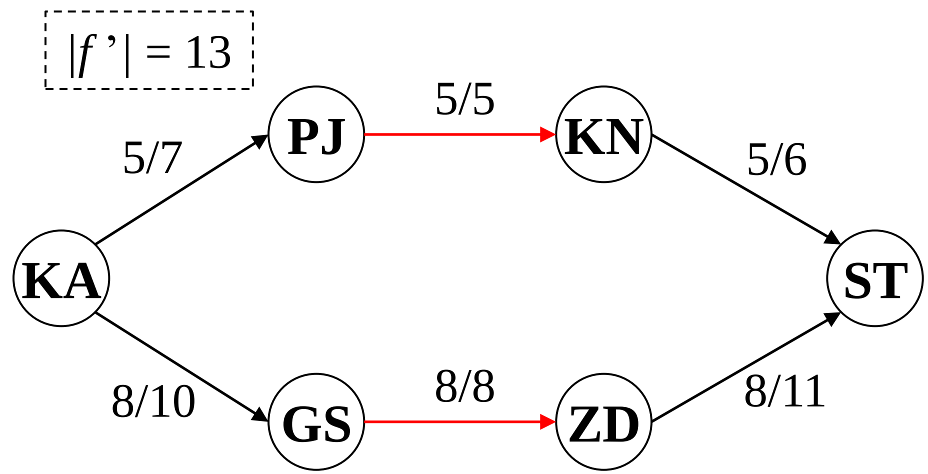
preostala mreža

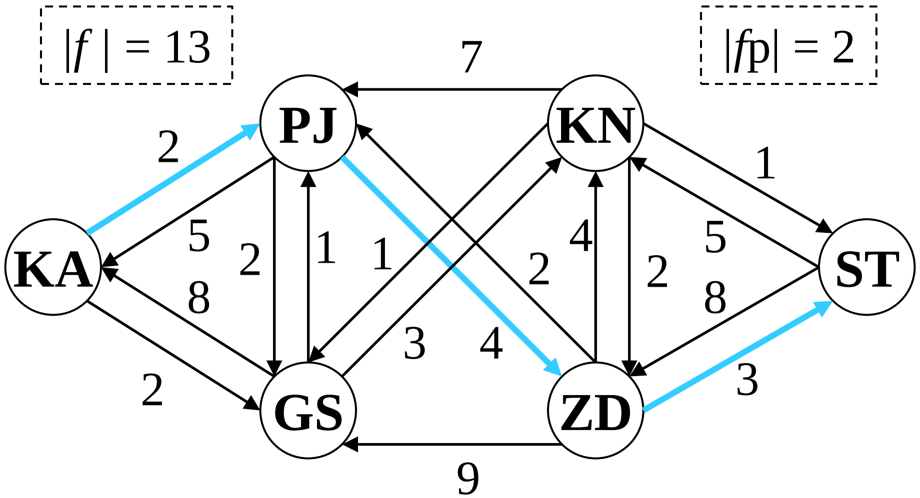


tok: $f' = f + fp$

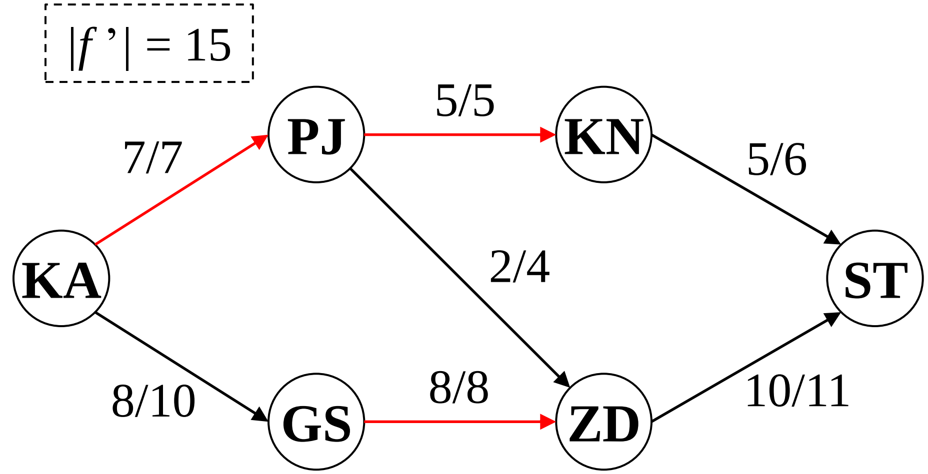




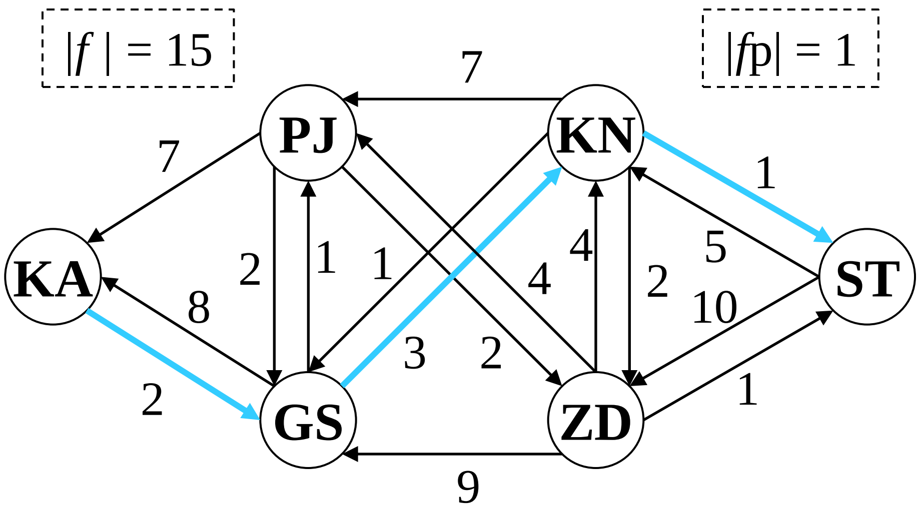




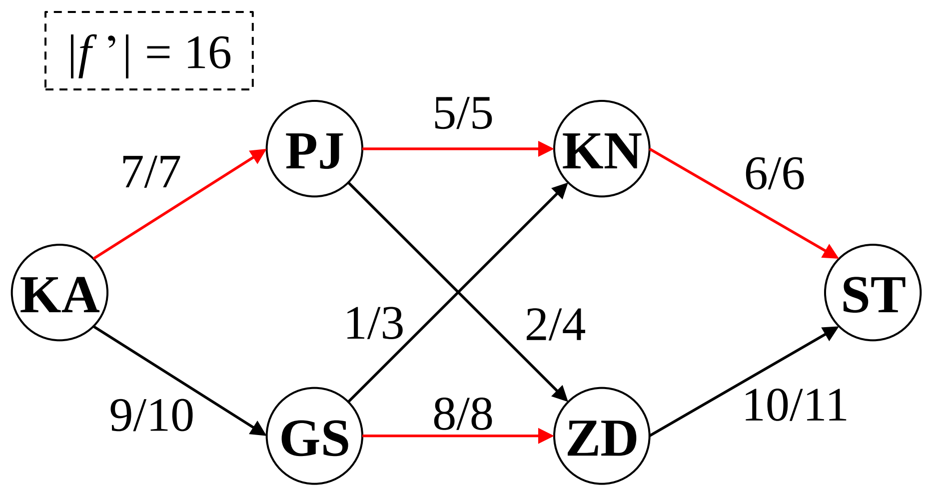
|f | = 13
|fp| = 2
PJ
KN
KA
ST
GS
ZD
7
2
5
2
1
1
2
4
2
5
1
8
8
2
3
4
3
9



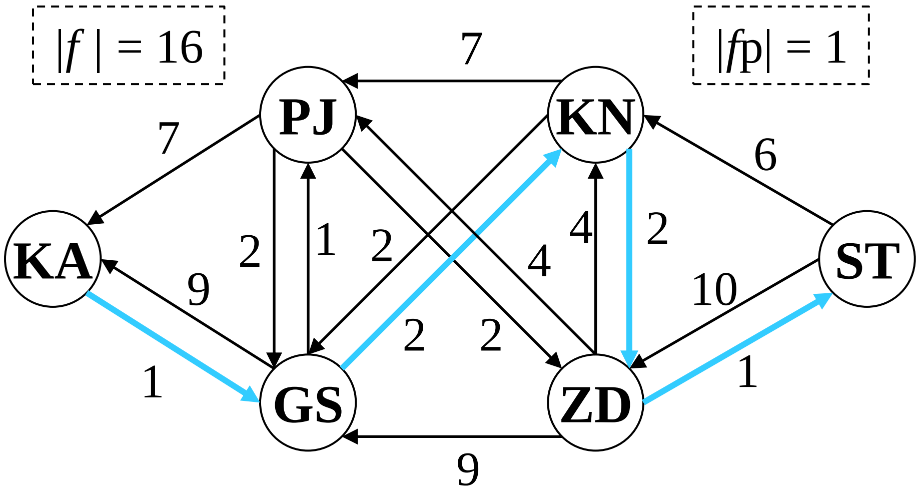
|f ’| = 15
PJ
KN
KA
ST
GS
ZD
5/5
7/7
5/6
2/4
8/10
8/8
10/11

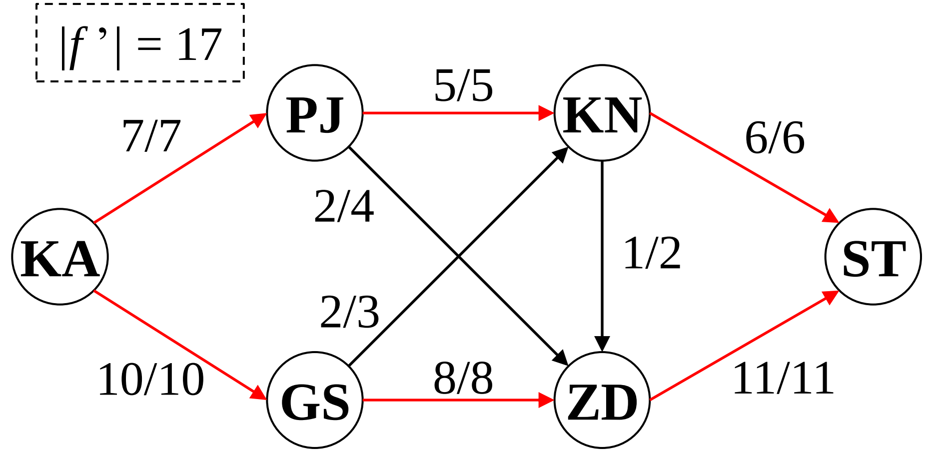


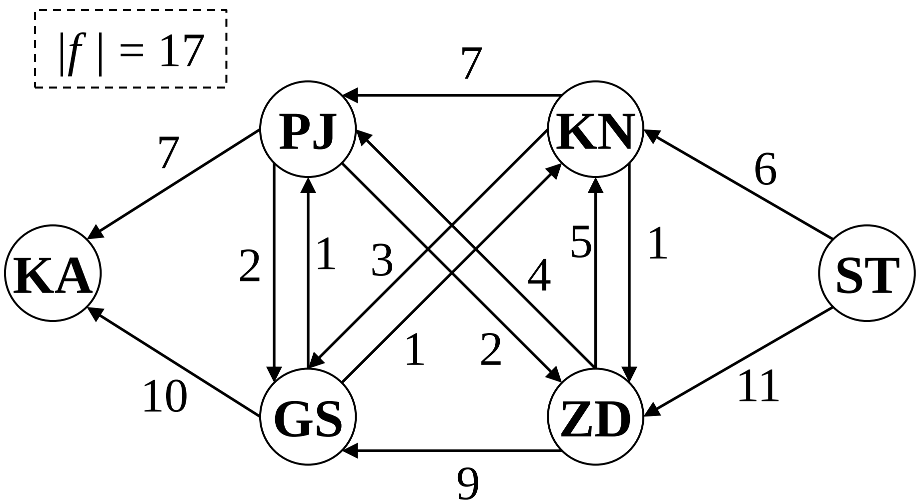
|f | = 15
|fp| = 1
PJ
KN
KA
ST
GS
ZD
7
7
1
2
1
1
4
4
2
5
8
10
3
2
2
1
9



|f ’| = 16
PJ
KN
KA
ST
GS
ZD
5/5
7/7
6/6
1/3
2/4
9/10
8/8
10/11

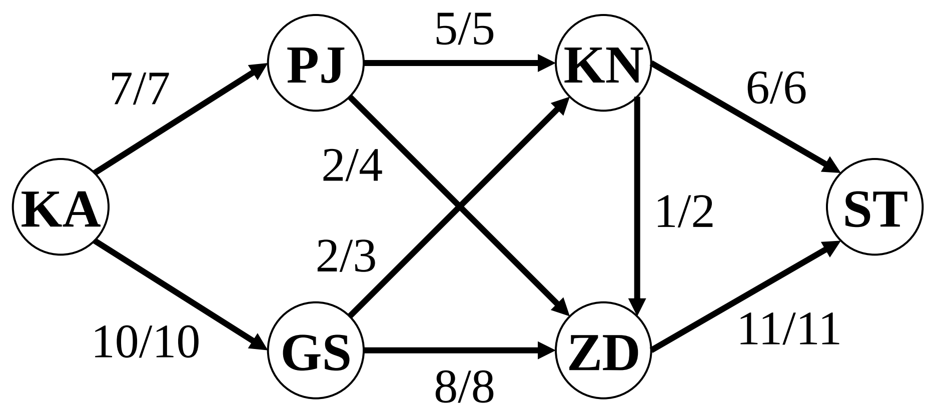






Više nema dopunskog puta!

Iskorištenje prometnica:



Složenost Ford-Fulkersonove metode, odnosno algoritma izvedenog iz nje, ponajviše ovisi o načinu na koji se traži dopunski put u preostaloj mreži.
Algoritam čak može i divergirati ako su kapaciteti bridova iracionalni brojevi (vidi Cormen...).

U svakom slučaju,

$$T(FF) = T(\text{inicijalizacija}) + f_m \cdot T(\text{dopunski put}) \ ,$$

gdje je $f_m = | f |_{max}$, jer `while` petlja se u najgorem slučaju obavlja f_{max} puta.

Inicijalizacija zahtijeva $\Theta(E)$ koraka, a tražimo li dopunski put DFS ili BFS algoritmom, složenost jedne potrage bit će $T(\text{dopunski put}) = O(V + E)$. Budući da je u mrežama uvijek $E \geq V - 1$, vrijeme $T(\text{dopunski put}) = O(E)$ ili ukupno

$$T(FF) = \Theta(E) + f_m \cdot O(E) = O(f_m \cdot E).$$

Najbolji rezultat postiže se Edmonds-Karpovim algoritmom. Edmonds-Karpov algoritam je algoritam izveden iz Ford-Fulkersonove metode koji za traženje dopunskih puteva koristi BFS, pri čemu mrežu "promatra" kao bestežinski graf (svi bridovi imaju jednaku težinu).

Za taj se algoritam dokazuje gornja granica složenosti $O(V \cdot E^2)$ (vidi Cormen...).

Ako mreža ima više ulaza i/ili izlaza, uvode se “umjetni” superizvor (*supersource*) i superslijev (*supersink*), čime se problem svodi na već riješenu situaciju s jednim izvorom i jednim slijevom.