tanulmányok 77/1978

MTA Számitástechnikai és Automatizálási Kutató Intézet Budapest

MAGYAR TUDOMÁNYOS AKADÉMIA
SZÁMITÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZETE

OPTIMALIZÀLÀSI FELADATOK CSOMAGKAPCSOLT SZÀMITÒGÈPHÀLÒZATOK TERVEZÉSÈNÈL

Irta:

SZÁSZNÉ TURCHÁNYI PIROSKA

Tanulmányok 77/1978.

A kiadásért felelős:

DR VÁMOS TIBOR

ISBN 963 311 062 9

ISSN 0324-2951

TARTALOMJEGYZÉK

oldal

1. BEVEZETÉS

A számitástechnika mind szélesebb körü alkalmazása és a számitógépek jobb kihasználtságának érdekében olyan lehetőségek megteremtésére van szükség, melyek révén egyszerüen, viszonylag olcsón, távolból, tehát nemcsak a számitógép közvetlen közelében lévő felhasználók számára hozzáférhető egy, vagy több számitógép.

E probléma megoldásában a time-sharing rendszer kifejlesztése volt az első nagy lépés: terminálhálózat kiépitése egy számitógép köré. Ezután merült fel az az igény, hogy több programfeldolgozást végző számitógépet kapcsoljanak össze oly módon, hogy egy felhasználó bármelyik számitógép, bármelyik szolgáltatását (speciális hardware és software, adatbázisrendszer) igénybe vehesse. Megkezdődött tehát a számitógéphálózatok kialakitása. A feldolgozást végző számitógépek ("erőforrások") egy önálló adatátviteli hálózatra (communication subnetwork) támaszkodnak, mely az adattovábbitást vonalkapcsolt (circuit/line switching), vagy csomagkapcsolt (message/packet switching) módon végzi, , következésképpen a teljes számitógéphálózat (computer communication network) osztott (distributed) jellegü.

Az ilyen számitógéphálózatok tervezése és üzemeltetése igen összetett és bonyolult feladat. Napjainkban a legnagyobb és leghatékonyabban müködő hálózat az Egyesült Államokbeli ARPANET, mely mintegy 100 feldolgozást végző számitógépet foglal magába, átfogja az USA egész területét, de Európában és Hawaiban lévő központok is tartoznak hozzá.

Magyarországon is megkezdődtek a számitógéphálózatok kialakitásával kapcsolatos tevékenységek, melyekben intézetünk is komoly szerepet kiván vállalni az akadémiai számitógéphálózat létrehozása révén.

E tanulmány célja, hogy a csomagkapcsolt hálózatban felmerülő optimalizálási problémák közül két alapvetőt ismertessen, és bemutassa az ezek megoldására az ARPANET-ben kidolgozott algoritmusokat; remélve, hogy ezzel intézetünk munkájához segitséget nyujt. Feltételezzük, hogy az olvasó számára a számitógéphálózatok problémaköre nem ismeretlen, de az érthetőség megkönnyitésére a Függelékben röviden összefoglaljuk az osztott hálózat felépitését és müködési elveit, illusztrációként pedig vázlatosan bemutatjuk a (csomagkapcsolt) ARPA számitógéphálózatot.

2. CSOMAGKAPCSOLT SZÁMITÓGÉPHÁLÓZATOK MATEMATIKAI PROBLÉMÁI

2.1. Általánosságban a matematikai problémákról

Matematikailag egy csomagkapcsolt számitógéphálózat egy véges gráffal reprezentálható, ahol

csucs: csomópont, azaz az adatátviteli hálózat egy kapcsolószámitógépe (ARPA-ban: IMP, vagy TIP) a hozzá tartozó terminálokkal és feldolgozást végző számitógép(ek)kel (ARPA-ban:HOST)

él: a csomópontokat összekötő csatornák.

Pontosabb definició csak a konkrét matematikai problémák leirásánál szükséges.

A hálózatban vagy teljes (program+adat) üzenetek (message switching), vagy egységnyi csomagok (packet-switching) keringenek. Egyszerüség kedvéért nem teszünk különbséget a két rendszer között, mindig üzenetekről fogunk beszélni.

Feltehetjük, hogy a gráf bármely csucsából bármely csucsába mehet üzenet. Az üzenetek érkezése, feladó, ill. rendeltetési helye, egy üzenet hossza: véletlen mennyiségek. Ezért a rendszer igen bonyolult. A probléma kezelhetősége érdekében tegyük fel, hogy az érkezések csucsonként egymástól független Poisson folyamatot alkotnak, a hosszak pedig exponenciális eloszlásuak.

A feladó-csomópontban jelentkező üzenetet el kell látni a továbbitáshoz szükséges információkkal (pl. üzenet eleje, vége, hossza, feladó- és rendeltetési hely), s ezt ugy kell végezni, hogy közben az üzenet bitjeinek száma ne nőjön jelentősen. Ez információelméleti probléma, mellyel eddig nem foglalkoztak. Az első ilyen témáju munka Gallager kézirata, melyben a fenti információk hosszára ad alsó becslést, s egyuttal olyan kódolási eljárást, mellyel jól megközelithető az optimum. [R.G.Gallager: Basic Limits on

Protocol Information in Data Communication Networks]

A rendszer müködésének elinditásához utkijelölő algoritmus szükséges, mely lehet determinisztikus, vagyis a feladóhelyen kijelölik a teljes utvonalat; és lehet adaptiv, amikor csomópontról csomópontra határozzák meg egy üzenet utját. Ha ez adott, kiszámitható egy üzenet hálózatban eltöltött idejének várható értéke, ami már igen bonyolult tömegkiszolgálási feladat [11] [12], hiszen a csucsokban az üzenetekből sorok keletkeznek. Ha ezt a problémát megoldottuk, következik az optimális utkijelölő algoritmus keresése, melynek a hálózatban eltöltött idő minimalizálása a feladata.

Tehát egy csomagkapcsolt számitógéphálózat tervezésekor, fejlesztésekor és üzemeltetésekor matematikailag lényegében a következő probléma komplexummal állunk szemben:
Minimalizálnunk kell

egy üzenet hálózatban töltött idejének várható értékét, azaz ennek megfelelő

optimális (determinisztikus v. adaptiv) utkijelölő algoritmust kell adnunk,

figyelembe véve bizonyos

költség előirásokat
vonalkapacitási és csomóponti tárolókapacitási korlátokat
megbizhatósági szintet

a következő hálózati paraméterek mellett:

topológiai paraméterek, pl.
feldolgozást végző számitógépek elhelyezkedése,
egy kapcsolószámitógéphez tartozó feldolgozást végző számitógépek és terminálok száma

továbbitási paraméterek, pl.
kapcsoló számitógépek üzenetkezelési sebessége
kapcsoló számitógépek tárolókapacitása
vonalkapacitások

üzenethossz-eloszlás
utvonalkijelölő algoritmus
üzenetforgalom-eloszlás (hálózati terhelés)
prioritási elv

A különböző paraméterek lehetnek adottak, de feladat lehet ezek meghatározása is aszerint, hogy tervezésről, fejlesztésről, vagy üzemeltetésről van szó. A problémakomplexum információelméleti, tömegkiszolgálási, matematikai programozási feladatok pontos megfogalmazását és megoldását kivánja. Most a matematikai programozási - hálózati folyamok - jellegü feladatok közül két olyat fogunk ismertetni, melyeket az ARPA számitógéphálózat müködtetéséhez oldottak meg. (A tömegkiszolgálás elméletének alkalmazásáról magyar nyelven a KFKI "Számitógéphálózatok" c. szemináriumi sorozatában jelent meg Záray Éva ismertetése).

2.2. Csomagkapcsolt hálózat matematikai programozási modellje

Legyen $G=(N,A,C)$ véges irányitatlan gráf,

ahol N a csucsok halmaza (csucsok száma: n)

A a csucsokat összekötő élek halmaza (élek száma: b) és

$C=(C_1, C_2, \ldots C_b) \geq 0$ az élekhez rendelt kapacitásértékekből alkotott vektor.

A számitógéphálózatot az üzenetforgalom (utvonalkijelölés, hálózatban töltött idő) szempontjából fogjuk vizsgálni, ezért elég az adatátviteli hálózatra szoritkoznunk, tehát a gráfban

csucs: kapcsolószámitógép (ARPA-ban IMP, TIP)

él: a csucsokat összekötő nagysebességü duplex csatornák

kapacitás: a csatornák mp-ként meghatározott mennyiségü bit szállitására képesek, és a kapcsolószámitógépek üzenetkezelési sebessége is figyelembe veendő.

Legyen $r_{ij} \geq 0$ az i-ik csucsból a j-ik csucsba továbbitandó üzenetek átlagos mennyisége (bit/sec) $i=1,\dots n$ $j-1,\dots n$

Ekkor a hálózat átlagos üzenetforgalmát a $\Gamma=((r_{ij}))$ $i=1,\dots n$ nem negativ elemü igénymátrix-szal és a $j=1,\dots n$

$$\gamma = \sum_i \sum_j r_{ij}$$

átlagos összforgalmi értékkel jellemezzük.

Ut a gráfban:

Az i-ik csucsból a j-ik csucsba vezető Πij utnak egymáshoz csatlakozó, hurokmentes élrendszert nevezünk. Ha az i-ik és j-ik csucs között több Πij ut létezik, az Γij igény eljuttatására több lehetőség van, és az utak közötti különbségtevés az utak hossza szerint történik. (Az ut hosszát esetenként másként definiáljuk. Igen egyszerü metrika, ha az utban lévő élek számát vesszük, de lényegesen bonyolultabb metrikát is definiálhatunk - ld.5.fejezet).

Folyam a gráfban:

Az r_{ij} igények eljuttatását egy $\Phi = \{f^{ij}_{k\ell}\}$ folyamfüggvénnyel irjuk le, ahol $f^{ij}_{k\ell}$ jelenti, hogy az i-ből j-be továbbitandó üzenetnek mekkora mennyiségét továbbitjuk a (k,l) élen. Tehát Φ a a következő feltételekkel jellemezhető:

$$(2.2.1) \quad \sum_{k=1}^{n} f^{ij}_{k\ell} - \sum_{m=1}^{n} f^{ij}_{\ell m} = \left.\begin{cases} -r_{ij} & \text{ha } \ell=i \\ +r_{ij} & \text{ha } \ell=j \\ 0 & \text{egyébként} \end{cases}\right] \quad \forall\ i,j$$

$$(2.2.2) \quad f^{ij}_{k,\ell} \geq 0 \qquad i,j,k,\ell=1,\ \dots\ n$$

A (2.2.1.), (2.2.2.) feltételeknek eleget tevő Φ folyamfüggvényt multicommodity folyamnak (m.c.f) nevezzük.

Egyszerübb, ha $f^{ij}_{k\ell}$ helyett f^{ij}_{t}-vel jelöljük az élenkénti értékeket: $t=1,\dots b$.

Legyen $\underset{\sim}{f}^{ij}=(f_1^{ij}, \ldots, f_b^{ij})$ az r_{ij} igényt i-ből j-be juttató <u>(i,j) elemi folyam.</u>

Ekkor az f_t^{ij} komponens ennek a t-ik élre jutó részét jelenti.

Legyen továbbá

$$\underset{\sim}{f} = \sum_{i=1}^{n} \sum_{j=1}^{n} \underset{\sim}{f}^{ij}$$ a hálózatbeli <u>teljes folyam</u>

Belátható, [8], hogy az

$F = \{\underset{\sim}{f} \mid \underset{\sim}{f} \text{ m.c.f.}\}$ halmaz konvex

és F extremális pontjai (az élek bármilyen metrizálása esetén!) "legrövidebb" folyamok.

Az $\underset{\sim}{f}$ teljes folyam nem azonos a Φ folyammal. A hálózatban felmerülő feladatok pedig Φ valamely $P(\Phi)$ függvényének optimalizálását jelentik. - Csak olyan feladatokkal fogunk foglalkozni (4.fejezet), melyben a $P(\Phi)$ függvény helyett a $P(\underset{\sim}{f})$ függvénnyel is dolgozhatunk.

A feladatokban rendszerint a folyamfüggvénynek a multi-commoditáson kivül más, u.n. addiciós feltételeknek is eleget kell tennie. Legegyszerübb a kapacitás feltétel:

$$f_t^{ij} \leqq C_t \quad t=1,\ldots b \quad \begin{matrix} i=1,..n \\ j=1,..n \end{matrix}$$

2.3. <u>Feladatcsoportositás a P(f)célfüggvény és az addiciós feltételek szerint</u>

2.3/a <u>Feltétel nélküli m.c.f. feladatok</u>

(azaz nincsenek addiciós feltételek)

α) <u>$P(\underset{\sim}{f})$ lineáris függvény</u>

Ilyen feladat pl. egy kapacitásmentes hálózatban a legrövidebb ut meghatározása. (Erre a feladatra számos megoldási módszer ismeretes, [8].)

β) <u>$P(\underset{\sim}{f})$ nem lineáris függvény</u>

Szeparábilis ($P(\underset{\sim}{f})= \sum_{i=1}^{\ell} P_i(Fi)$ és konvex függvény esetén, ha a hálózat kicsi, érdemes $P(\underset{\sim}{f})$ lineáris közelitésével az α) esetre visszavezetni a feladatot.

Differenciálható $P(\underset{\sim}{f})$ függvényre (tetszőleges méretü hálózatban) az érintősikokkal való közelitést alkalmazzák. Igy ismét lineáris feladatot kapnak, ahol az i-ik él hossza $\frac{\partial P}{\partial f_i}$

2.3.b. Optimalizálás az f -re vonatkozó addiciós feltételek mellett

α) $P(\underset{\sim}{f})$ lineáris függvény és az addiciós feltételek is lineárisak

A feladat megoldása általában a Danzig-Wolfe féle dekompoziciós módszerrel [2] történik, melyben lényegében 2.3/a. tipusu feladatsorozatokra vezetik vissza az eredeti feladatot.

β) $P(\underset{\sim}{f})$ nem lineáris függvény, az addiciós feltételek sem lineárisak

csak konvex függvény, és konkáv, nem negativitási feltételek esetén ismertek hatásos megoldási módszerek [2].

Az 5. fejezetben olyan módszert fogunk ismertetni, mely nemlineáris, differenciálható $P(\underset{\sim}{f})$ függvény feltétel nélküli optimalizálására alkalmas, de a 7. fejezetben vázolni fogjuk azt is, hogyan módositható ez a módszer egy diszkrét feladat megoldására.

$P(\underset{\sim}{f})$ leggyakrabban időfüggvény: annak az időnek várható értéke, melyet egy üzenet a hálózatban eltölt, mig a feladóhelytől a rendeltetési helyre jut. Az eljutást az $\underset{\sim}{f}$ (tulajdonképpen a Φ) folyam valósitja meg.

Egy üzenet hálózatban eltöltött idejének kiszámitását tárgyaljuk a következő (3.) fejezetben, de csak olyan mélységben, amennyire ezt a 4. fejezetben ismertetésre kerülő feladatok megkivánják.

3. EGY ÜZENET HÁLÓZATBAN ELTÖLTÖTT IDEJÉNEK KISZÁMITÁSA

Láttuk, hogy egy csomagkapcsolt hálózatban minden csomópontban különböző manipulációkat kell végezni egy üzenettel. Ennek következtében egy csomópontban (általában) több üzenet várakozik (tárolódik a memóriában), mig egynek folyik a feldolgozása. A teljes számitógéphálózat tehát egy igen bonyolult tömegkiszolgáló rendszert is jelent.

3.1. Vizsgáljunk először egy csomópontot, mint egycsatornás tömegkiszolgáló rendszert.

Legyen $A(x)$ az üzenetek beérkezése között eltelt idő eloszlásfüggvénye

$B(x)$ a csomópontbeli kiszolgálási idő eloszlásfüggvénye

λ az üzenetek érkezésének gyakorisága [messg/sec]

$\bar{t}$ egy üzenet átlagos kiszolgálási ideje [sec/messg]

$\frac{1}{\mu}$ az átlagos üzenethossz [bits/messg]

$C=\frac{1}{\mu\bar{t}}$ kapacitás: üzenetfeldolgozási, ill. továbbitási sebesség [bit/sec]

Matematikailag rendkivül leegyszerüsiti a rendszer leirását, ha feltesszük, hogy A és/vagy B exponenciális eloszlásuak. Ennek megfelelően több eset lehetséges:

<u>1. eset</u>

(i) Az érkezések között eltelt idők független, egyforma exponenciális eloszlások ($A(x)=1-e^{-\lambda x}$), tehát az érkezések Poisson folyamatot alkotnak.

(ii) A kiszolgálási idők független, egyforma exponenciális eloszlásuak $(B(x)=1-e^{-\frac{x}{t}})$

Ha a kiszolgálás az érkezés sorrendjében történik, azaz a a hálózatban nincs prioritása egy üzenetnek sem, akkor az átlagos kiszolgálási idő várható értéke elemi módon kiszámolható [10], [11] :

(3.1.1) $$T = \frac{\bar{t}}{1-\lambda\bar{t}} \quad \text{vagy} \quad T = \frac{1}{\mu C-\lambda}$$

ahol T az üzenetnek a kiszolgáló csomóponton eltöltött teljes idejét jelenti, amely két részre bontható: a kiszolgáló foglaltsága miatt esetleg szükséges várakozási időre, valamint a tényleges kiszolgálás idejére. Ennek a felbontásnak megfelelően a T teljes várakozási idő várható értéke igy irható:

(3.1.2) $$T = \frac{\lambda\bar{t}^2}{1-\lambda\bar{t}} + \frac{1}{\mu C}$$

2. eset

Az érkezések ismét Poisson folyamatot alkotnak, de a kiszolgálási idők tetszőleges eloszlásuak. Ekkor a Pollaczek-Hincsin formulából

(3.1.3) $$T = \frac{\lambda s}{2(1-\lambda t)} + \frac{1}{\mu C}$$

ahol $s=\int x^2 dP(X)$ a kiszolgálási idők második momentuma.

3. eset

A(x) is, B(x) is tetszőleges eloszlásfüggvények. Ekkor nem ismeretes zárt formula T meghatározására, csupán egy igen jó felső korlát [9]:

(3.1.4) $$T \leq \frac{\lambda(\sigma_1^2+\sigma_2^2)}{2(1-\lambda\bar{t})} + \frac{1}{\mu C}$$

ahol σ_1^2 ill. σ_2^2 az A(x) ill. B(x) eloszlások szórásnégyzetei.

3.2. Tekintsük most a teljes hálózatot (n csomóponttal és b éllel)

Poisson érkezési folyamat és exponenciális kiszolgálás esetén, ha az egyes csomópontok, mint tömegkiszolgálórendszerek függetlenek, egy üzenetre a hálózatban eltöltött idő kiszámitását a hálózat dekomponálásával - csomópontokra történő lebontásával - végezhetjük.
A csomagkapcsolt hálózat azonban másként müködik: a csomópontok egymástól nem függetlenek, s mivel egy üzenet hossza a továbbitás során nem változik, nem változik a csomópontonkénti kiszolgálási idő eloszlása sem. Ennek ellenére, ha az üzenetek hosszát minden egyes pontban valószinüségi változónak tekintjük, (ez sok pontból álló hálózatban, ahol egy csomópontra több irányból érkeznek és több irányba futnak ki üzenetek, nem is olyan irreális feltételezés) a függetlenség és Poissonitás feltételezésével elméleti uton számitott késési idő nagyjából megegyezik a szimulációs módszerekkel kapott késési idővel [10] , [12] .

Ennek alapján kiszámitható az egységnyi üzenet által a hálózatban töltött teljes idő T várható értéke:

Ha $\Gamma = ((r_{ij}))$ $i=1,\ldots n$ az igénymátrix, és

és $\gamma = \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij}$ a hálózat átlagos összforgalma,

akkor a csomópontonkénti lebontás alapján

(3.2.1) $$T = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{r_{ij}\, T_{ij}}{\gamma}$$

Az irodalomban az egyszerüség kedvéért a várakozási és kiszolgálási időből származó késést nem a csucsoknál, hanem az éleknél veszik figyelembe, és a csatornán való áthaladási időhöz adják hozzá. (Ez természetesen további durvitása a modellnek, hiszen pl. az éleket irányitatlanoknak veszik, igy egybe kell fogják a két irányból továbbitandó üzeneteket).

A T idő várható értéke ekkor:

(3.2.2) $$T = \sum_{i=1}^{b} \frac{\lambda_i}{\gamma}(T_i+p_i)$$

ahol λ_i az i-ik élre jutó átlagos üzenetmennyiség (mely a Γ mátrix és az utkijelölő algoritmus ismeretében egyszerüen kiszámitható)

p_i. egy üzenet áthaladási ideje az i-ik élen

$$T_i = \frac{\lambda_i/\mu C_i}{\mu C_i - \lambda_i} + \frac{1}{\mu' C_i}$$

(3.1.2) alapján $(\bar{t} = \frac{1}{\mu C_i})$ egy üzenet várakozási és tényleges kiszolgálási idejéből adódó késés az i-ik élen;

($\frac{1}{\mu}$ az üzenet hossza információ nélkül

$\frac{1}{\mu'}$ az üzenet hossza információkkal együtt.)

Szokás ezt a formulát még tovább finomitani, de erre most nem lesz szükségünk. [6], [10], [12]

Érdekes megemliteni azonban a következő formulát:

(3.2.3) $$T = \left(\sum_{i=1}^{b} \frac{\lambda_i}{\gamma} T_i^L\right)^{\frac{1}{L}}$$

ahol L alkalmasan választott hatványkitevő, azaz T az élenkénti késések hatványközepe, amely nagy L-ekre érzékenyebb az éleken előforduló legnagyobb késésekre.

A T idő ilyen módon történő számolása a következő esetben hasznos:

Előfordulhat, hogy egy bizonyos csatornán az átlagos késés nagyságrendekkel nagyobb, mint a többin, de ha a forgalom kicsi, ez a jelenség az előbbi módon számitott késési időre (3.2.2) nincs hatással, ennek következtében a felhasználó "rosszul jár" [6]. Hatványközép alkalmazásával ez a negativum csökkenthető.

4. DETERMISZTIKUS IRÁNYITÁSI- (UTKIJELÖLÉSI) ÉS CSATORNAKAPACITÁSTERVEZÉSI FELADATOK

A 2. fejezetben vázolt számitógéphálózatbeli matematikai problémák közül kettőnek a megoldását fogjuk ismertetni.

Az eddigi jelöléseknek megfelelően

n csomópontu,

b élü hälózatot vizsgálunk,

$\Gamma=((r_{ij}))$ $i=1,\ldots,n$ $j=1,\ldots,n$ igénymátrix esetén

$\underset{\sim}{C}=(C_1,\ldots,C_b)$ kapacitásvektorral

$\underset{\sim}{\lambda}=(\lambda_1,\ldots,\lambda_b)$ üzenet-érkezés gyakorisággal.

Egy üzenet hálózatban töltött idejét a

$$T = \sum_{i=1}^{b} \frac{\lambda_i}{\gamma}(T_i+p_i) = \sum_{i=1}^{b} \frac{\lambda_i}{\gamma}\left(\frac{1}{\mu C_i-\lambda_i} + p_i\right)$$

formulával számoljuk (ld. 3.1.1. és 3.2.2)

$$\text{Legyen} \quad \frac{\lambda_i}{\mu} = f_i \quad \text{és} \quad \mu p_i = p_i' \quad i=1,\ldots,b$$

vagyis f_i az i-ik élre (csatornára) jutó bitek átlagos száma, a $\underset{\sim}{f}$ teljes folyam i-ik komponense (ld. 2.2. fejezet)

Ekkor

$$T = \frac{1}{\gamma}\sum_{i=1}^{b} \frac{f_i}{C_i-f_i} + \frac{1}{\gamma}\Sigma f_i p_i'$$

4.1. <u>feladat</u>

<u>Optimális irányitás problémája</u>

Adott: $G=(N,A)$ hálózati topológia

$\underset{\sim}{C}$ kapacitásvektor

Γ igénymátrix

Minimalizálandó

(4.1.1.) $$T(\underset{\sim}{f}) = \frac{1}{\gamma}\sum_{i=1}^{b} \frac{f_i}{C_i-f_i} + \frac{1}{\gamma}\sum_{i=1}^{b} f_i p_i'$$

feltéve, hogy

(4.1.2) $\underset{\sim}{f}$ m.c. folyam

(4.1.3) $\underset{\sim}{f} \leqq \underset{\sim}{C}$ azaz $f_i \leqq C_i$ $i=1,\ldots,b$

4.2 feladat

Optimális irányitás és csatornakapacitástervezés additiv költség-kapacitásfüggvény esetén

Adott: $G=(N,A)$ hálózati topológia

Γ igénymátrix

$d(\underset{\sim}{C}) = \sum_{i=1}^{b} d_i(C_i)$ költség-kapacitásfüggvény

D költségkorlát

Minimalizálandó

$$T(\underset{\sim}{C},\underset{\sim}{f}) = \frac{1}{\gamma} \sum_{i=1}^{b} \frac{f_i}{C_i - f_i} + \frac{1}{\gamma} \Sigma p_i' f_i \tag{4.2.1}$$

feltéve, hogy

$$\underset{\sim}{f} \text{ m.c. folyam} \tag{4.2.2.}$$

$$\underset{\sim}{f} \leqq \underset{\sim}{C} \tag{4.2.3.}$$

$$d(\underset{\sim}{C}) \leqq D \tag{4.2.4.}$$

Ez utóbbi feladat megoldását $d(\underset{\sim}{C}) = \sum_{i=1}^{b} d_i C_i$ lineáris költség-kapacitás függvény esetén tárgyaljuk, ezért először nézzük meg, hogyan redukálódik ekkor a probléma.

A (4.2.1.) célfüggvényt először rögzitett $\underset{\sim}{f}$ mellett, $\underset{\sim}{C}$-ben minimalizáljuk, a Lagrange multiplikátoros módszerrel [7]. Ennek eredményeképpen

$$C_i = f_i + \frac{D - \sum_{i=1}^{b} f_i d_i}{d_i} \frac{\sqrt{f_i d_i}}{\Sigma \sqrt{f_j d_j}} \tag{4.2.5.}$$

melyet a (4.2.1.) célfüggvénybe visszahelyettesitve

$$T(\underset{\sim}{C},\underset{\sim}{f}) = T(\underset{\sim}{f}) = \frac{\left(\sum_{i=1}^{b} \sqrt{f_i d_i} \right)^2}{\gamma \left(D - \sum_{i=1}^{b} d_i f_i \right)} + \frac{1}{\gamma} \sum_{i=1}^{b} f_i p_i' \tag{4.2.6.}$$

Könnyen látható, hogy a (4.2.3.) - (4.2.4.) feltételek,

azaz $f_i \leqq C_i \qquad i=1,\dots,b$

$$\sum_{i=1}^{b} d_i C_i \leqq D \qquad \text{helyettesithetők a}$$

(4.2.7.) $$D - \sum_{i=1}^{b} d_i f_i \geqq 0 \qquad \text{feltétellel.}$$

Igy végül a következő feladathoz jutunk:

(4.2*) <u>feladat</u>

<u>Optimális irányitás és csatorna kapacitástervezés lineáris költség-kapacitás függvény esetén.</u>

Adott $G=(N,A)$ hálózati topológia

Γ igénymátrix

$d(\underset{\sim}{C}) = \sum_{i=1}^{b} d_i C_i$ költség-kapacitás függvény

D költségkorlát

Minimalizálandó

(4.2.*1) $$T(\underset{\sim}{f}) = \frac{\left(\sum_{i=1}^{b} \sqrt{f_i d_i}\right)^2}{\gamma\left(D - \sum_{i=1}^{b} d_i f_i\right)} + \frac{1}{\gamma} \sum f_i p_i'$$

feltéve, hogy

(4.2.*2) $\underset{\sim}{f}$ m.c. folyam

(4.2.*3) $D - \sum d_i f_i \geqq 0$

<u>Megjegyzések:</u>

1. A megengedett megoldások halmaza a (4.1.) és a (4.2*) feladatok mindegyikében konvex:

$$F_1 = F \cap \{\underset{\sim}{f} \mid \underset{\sim}{f} \leqq \underset{\sim}{C}\}$$

$$F_2 = F \cap \{\underset{\sim}{f} \mid D - \sum_{i=1}^{b} d_i f_i \geqq 0\}$$

ahol $F = \{\underset{\sim}{f} \mid \underset{\sim}{f}$ multicommodity folyam a Γ igénymátrix-szal$\}$.

Tehát az F_1, F_2 halmazok extremális pontjai legrövidebb m.c. folyamok [8].

2. Mind a (4.1), mind a (4.2*) feladat esetében, ha egy $\underset{\sim}{f}$ megengedett megoldás a feltételhalmaz határához közeledik, akkor a célfüggvény értéke végtelenhez tart, azaz a (4.1.1) célfüggvény a (4.1.3) feltételt, a (4.2*.1) célfüggvény a (4.2*.3) feltételt <u>büntetőfüggvényként</u> magában foglalja. Tehát, bármelyik feladatban ha valamilyen módon találtunk egy induló megengedett megoldást, a célfüggvény <u>feltételnélküli</u> minimalizálására térhetünk át.

3. Mivel a megengedett megoldások halmaza konvex, zárt, és korlátos, és mivel

   a (4.1.1) célfüggvény konvex,

   a (4.2*.1) célfüggvény kvázikonkáv (a kvázikonkávitás definicióját ld. 5. fejezet),

   azért mindkét feladatra igaz, hogy ha van megengedett megoldás, akkor van optimális megoldás is.

5. A "FLOW DEVIATION" – FD – MÓDSZER

5.1 Stacionaritás

Legyen $F = \{\underset{\sim}{f} \mid \underset{\sim}{f}$ m.c. folyam$\}$ és P folytonosan differenciálható függvény az F halmazon.

Definició:

$\underset{\sim}{f} \in F$ stacionárius pontja a P függvénynek, ha bármely kicsi $\delta\underset{\sim}{f}$ megváltoztatás esetén mindig teljesül $P(\underset{\sim}{f}+\delta\underset{\sim}{f}) \geqq P(\underset{\sim}{f})$

Mivel P lokális és globális minimumhelyei stacionárius pontok, szükséges és elégséges feltételt keresünk ahhoz, hogy $\underset{\sim}{f} \in F$ stacionárius pontja legyen P-nek.

Definició:

"Flow Deviation"

Legyen $\underset{\sim}{f} \in F$ rögzitett, $\underset{\sim}{v} \in F$ tetszőleges pont, és $0 \leqq \lambda \leqq 1$.

Tekintsük a következő konvex kombinációt:

$$\underset{\sim}{f}' = \underset{\sim}{f} + \lambda(\underset{\sim}{v} - \underset{\sim}{f}), \text{ ami } \in F, \text{ hiszen } F \text{ konvex.}$$

Legyen $0 < \lambda < \varepsilon$, ahol $\varepsilon \ll 1$. Ekkor a Lagrange-középérték tétel alapján

(5.1.1.) $$P(\underset{\sim}{f}') - P(\underset{\sim}{f}) = \lambda \nabla P(\underset{\sim}{g})\,(\underset{\sim}{v} - \underset{\sim}{f})$$

ahol

$$\underset{\sim}{g} = \underset{\sim}{f} + \Theta\lambda(\underset{\sim}{v} - \underset{\sim}{f}) \in F \quad , \quad 0 \leqq \Theta \leqq 1$$

és

$$\nabla P(\underset{\sim}{g}) = \left| \frac{\partial P}{\partial f_1}, \ldots, \frac{\partial P}{\partial f_b} \right)_{\underset{\sim}{f}=\underset{\sim}{g}}$$

vagyis a P függvény gradiense a $\underset{\sim}{g}$ pontban.

De mivel $0 < \lambda < \varepsilon$, és P folytonosan differenciálható:

$$(5.1.2.) \qquad P(\underset{\sim}{f}') - P(\underset{\sim}{f}) \approx \lambda \nabla P(\underset{\sim}{f})(\underset{\sim}{v}-\underset{\sim}{f}) = \lambda \sum_{k=1}^{b} \frac{\partial P}{\partial f_k}(v_k - f_k)$$

Jelöljük $\left.\frac{\partial P}{\partial f_k}\right|_{\underset{\sim}{f}} = \ell_k$ $(k=1,\ldots,b)$-vel a P függvény $\underset{\sim}{f}$ pontbeli parciális deriváltjait.

Mivel $\lambda > 0$, definició szerint $\underset{\sim}{f}$ a P függvénynek akkor és csak akkor stacionárius pontja, ha

$$(5.1.3.) \qquad \sum_{k=1}^{b} \ell_k(v_k - f_k) \geqq 0 \text{ bármely } \underset{\sim}{v} \in F \text{ esetén.}$$

Emlékezve $\underset{\sim}{f}$-nek a 2.2. fejezetben adott definiciójára, $\underset{\sim}{f}$ elég kicsi megváltoztatását elérhetjük ugy is, hogy csak egy (tetszőleges rögzitett) $\underset{\sim}{f}^{ij}$ elemi folyam-mennyiséget változtatunk meg, azaz a P függvény változtatását csak $\underset{\sim}{f}$ koordinátánkénti megváltoztatása esetén vizsgáljuk: azaz, ha $\underset{\sim}{v}$ a következőképpen áll elő $\underset{\sim}{f}$-ből:

$$\underset{\sim}{v}^{ij} = \begin{cases} \underset{\sim}{f}^{ij} & i \neq i_o,\ j \neq j_o \\ \underset{\sim}{v}^{i_o,j_o} & (i_o,j_o) \text{ tetszőleges, rögzitett } (i,j) \text{ pár} \end{cases}$$

akkor (5.1.3.) fennállása esetén igaznak kell lennie

$$(5.1.4.) \qquad \sum_{k=1}^{b} \ell_k(v_k^{ij} - f_k^{ij}) \geqq 0$$

egyenlőtlenségnek, mégpedig bármely $\underset{\sim}{v} \in F$ -re, hiszen (i_o,j_o) végigfut az összes (i,j) páron.

Forditott irányban az állitiás triviális, vagyis az (5.1.3.) és (5.1.4.) feltétel ekvivalens.

(5.1.3) és (5.1.4)-ből kapjuk:

(5.1.5.) $$\min_{\underset{\sim}{b} \in F} \sum_{k=1}^{b} \ell_k v_k \geq \sum_{k=1}^{b} \ell_k f_k$$

illetőleg

(5.1.6.) $$\min_{\underset{\sim}{v}^{ij} \in F^{ij}} \sum_{k=1}^{b} \ell_k v_k^{ij} \geq \sum_{k=1}^{b} \ell_k f_k^{ij}$$

De $\underset{\sim}{f} \in F$ illetőleg $\underset{\sim}{f}^{ij} \in F^{ij}$ miatt ez utóbbi két feltételnek egyenlőséggel kell teljesülnie.

Tehát ahhoz, hogy valamely $\underset{\sim}{f} \in F$ pontról eldöntsük, stacionárius helye-e a P függvénynek, vagy sem, az $\{\ell_k\}$ metrikában legrövidebb folyamatot kell megkeresnünk, és ellenőriznünk a

(5.1.7.) $$\min_{\underset{\sim}{v} \in F} \sum_{k=1}^{b} \ell_k v_k = \sum_{k=1}^{b} \ell_k f_k$$

feltétel teljesülését.

Ha P szigoruan konvex függvény az F halmazon, akkor ott globális minimumhelye természetesen stacionárius pont, és megforditva.

5.2. Az FD módszer

Az FD módszernél a (5.1.7.) feltételből indulunk ki. Ha a feltétel teljesül, stacionárius pontban vagyunk. Ha nem teljesül az egyenlőség, akkor a következő geometriai megfontolás alapján járunk el:

A Lagrange középértéktétel alapján, ha az $\underset{\sim}{f}$ folyamot egyforma abszolutértékü vektorok irányában változtatjuk, akkor a P függvényt legjobban ugy tudjuk csökkenteni, ha a változás iránya egybeesik a negativ gradiens irányával. Előfordulhat azonban, hogy ebben az irányban indulva csak keveset tudunk az F halmazon belül haladni, és érdemesebb olyan irányt választani, amely a negativ gradiens "közelében" van, és az F halmaz jóval "terjedelmesebb" ebben az irányban. - Ezen meggondolás alapján azt a $\underset{\sim}{v} - \underset{\sim}{f}$ irányt választjuk, melyben $\underset{\sim}{v}$ a (5.1.7.) feltétel baloldalát minimalizáló vektor. Tehát az FD módszer a következő operátor ismételt alkalmazásából áll:

Definició:

$FD(\underset{\sim}{v},\lambda) : F \to F$ operátor $/F=\{\underset{\sim}{f}\,|\,\underset{\sim}{f}$ m.c.f$\}/$

$FD(\underset{\sim}{v},\lambda)\odot\underset{\sim}{f} = \underset{\sim}{f} + \lambda(\underset{\sim}{v}-\underset{\sim}{f}) = \underset{\sim}{f}'$

ahol $\underset{\sim}{v}$ az $\{\ell_k\}$ metrikában legrövidebb m.c. folyam és λ $(0\leqq\lambda\leqq 1)$ az az érték, melyre $P[(1-\lambda)\underset{\sim}{f} + \lambda\underset{\sim}{v}]$ minimális λ-ban.

Az FD módszer főbb lépései:

I. fázis

Meghatározunk egy $\underset{\sim}{f}^{o} \in F$ induló megoldást. (Ennek módja a konkrét feladattól függ)

II. fázis

1. Legyen $n = 0$

2. $\underset{\sim}{f}^{n+1} = FD(\underset{\sim}{v}^{n},\lambda^{n})\odot \underset{\sim}{f}^{n}$

ahol $\underset{\sim}{v}^{n}$ az $\{\ell_k^n = \frac{\partial P}{\partial f_k}\Big|_{\underset{\sim}{f}=\underset{\sim}{f}^n}$, $k=1,\dots,b\}$ metrikában legrövidebb m.c. folyam

$\lambda^n\colon \min\limits_{0\leqq\lambda\leqq 1} P[(1-\lambda)\,\underset{\sim}{f}^{n} + \lambda\underset{\sim}{v}^{n}]$

3. Ha $P(\underset{\sim}{f}^{n}) - P(\underset{\sim}{f}^{n+1}) < \varepsilon$

(vagy $\sum\limits_{k=1}^{b} \ell_k(f_k^n - v_k^n) < \varepsilon'$)

ahol $\varepsilon > 0$ (ill. $\varepsilon' > 0$) tetszőleges rögzitett hibakorlát akkor készen vagyunk

ellenkező esetben $n = n+1$, és a 2. lépéstől folytatjuk az eljárást.

5.3. A módszer konvergenciája

Tétel [5]

Ha a P függvény az F halmazon alulról korlátos, kétszer differenciálható

$$\frac{\partial P}{\partial f_k} \geqq 0 \quad \text{és} \quad \frac{\partial^2 P}{\partial f_j \partial f_k} \quad \text{felülről korlátos,}$$

továbbá P nem degenerált,

azaz bármely $\underset{\sim}{f}_1 \neq \underset{\sim}{f}_2$ stac. pont esetén $P(\underset{\sim}{f}_1) \neq P(\underset{\sim}{f}_2)$

akkor az FD módszer alkalmazható

P értéke minden FD iterációban csökken (nem nő)

a módszer stacionárius ponthoz vezet.

A P függvény első parciálisainak nem-negativitása annak a metrikának a nem-negativitását jelenti, mely szerinti legrövidebb folyamot egy FD iteráción belül meg kell határoznunk. A nem-negativitási feltétel zárja ki a negativ, ciklusok lehetőségét, s ez a feltétel a számitógéphálózat modelljeiben szereplő célfüggvényekre, mint látni fogjuk, teljesül.

Számunkra elég a Tételnek az a speciális esete, amikor P konvex függvény és az egyszerüség kedvéért a bizonyitást is csak erre az esetre végezzük el.
Tehát bebizonyitjuk:

Tétel

Ha P szigoruan konvex függvény az F halmazon, kétszer differenciálható és $\frac{\partial P}{\partial f_j \partial f_k}$ korlátosak, akkor az FD módszer P globális minimumhelyéhez konvergál.

Bizonyitás:

Feltéve, hogy találunk $\underset{\sim}{f}^{o}$ induló megengedett megoldást, legyen $\{\underset{\sim}{f}^{n}\}$ az FD módszer által generált sorozat. Mivel

$$P(\underset{\sim}{f}^{n}) - P(\underset{\sim}{f}^{n+1}) \geqq 0$$

azért a $\{P(\underset{\sim}{f}^{n})\}$ sorozat monoton nem növő, a konvexitás miatt alulról korlátos is, tehát

$$\exists \lim_{n\to\infty} P(\underset{\sim}{f}^{n}) = P(\underset{\sim}{f}^{*}) \quad \underset{\sim}{f}^{*} \in F$$

Állitás: f^{*} globális minimumhelye a P függvénynek, azaz $P(\underset{\sim}{f}^{*}) \leqq P(\underset{\sim}{f})$ bármely $\underset{\sim}{f} \in F$

Bizonyitás:

Tegyük fel, hogy a $\{P(f^{n})\}$ sorozat nem tart a P globális minimumához, azaz $\exists\ t > 0$ vagy $P(f^{n}) \geqq P(f^{min})+t$ bármely n-re

Ez szemléletesen azt jelenti, hogy a P függvényen nem jutunk lejjebb, mint a $P(\underset{\sim}{f}^{min})$ feletti t magasságban levő "sik-röviden" "t-sik" - által a P-ből kimetszett görbe. Ez a görbe korlátos és konvex, hiszen $P(\underset{\sim}{f})$ szigoruan konvex. Tetszőleges rögzitett $\underset{\sim}{f} \in F$ pontra, amelyre $P(\underset{\sim}{f})$ még a t-sikon, vagy a t-sik felett van, legyen

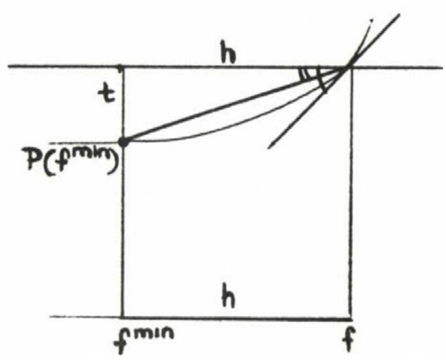
$$P(\lambda) = P(FD\odot \underset{\sim}{f}) = P[\underset{\sim}{f}+\lambda(\underset{\sim}{v}-\underset{\sim}{f})]$$

Ekkor ismét a Lagrange középértéktételből

$$P(\lambda) = P(0) + \lambda \left.\frac{dP}{d\lambda}\right|_{\lambda=0} + \frac{1}{2}\lambda^{2} \left.\frac{d^{2}P}{d\lambda^{2}}\right|_{\lambda=\xi}$$

ahol $\left.\frac{d^{2}P}{d\lambda^{2}}\right|_{\lambda=\xi} \leqq M<\infty$ a feltevések miatt, és $\left.\frac{dP}{d\lambda}\right|_{\lambda=0} = \sum_{k=1}^{b} \frac{\partial P}{\partial f_k}(v_k - f_k)$

Ha $\underset{\sim}{f} \in F$ olyan, hogy $P(\underset{\sim}{f})$ a t-sikon van, és h jelöli $\underset{\sim}{f}$ és $\underset{\sim}{f}^{min}$ távolságát, akkor a konvexitásból könnyen belátható, hogy $0 < \frac{t}{h_0} \leq \frac{t}{h} \leq \left.\frac{dP}{d\lambda}\right|_{\lambda=0}$



ahol $h_0 = \max h < \infty$

Tehát $\left.\frac{dP}{d\lambda}\right|_{\lambda=0} \leq - \frac{t}{h_0} = -\Delta < 0$ vagyis $\sum \frac{\partial P}{\partial f_k}(v_k - f_k) < 0$

minden $\underset{\sim}{f} \in F$ -re, melyre $P(\underset{\sim}{f}) = P(\underset{\sim}{f}^{min}) + t$ és következésképpen azokra is, melyekre $P(\underset{\sim}{f}) \geq P(\underset{\sim}{f}^{min}) + t$

Igy

$$P(\lambda) - P(0) \leq - \Delta\lambda + \frac{1}{2}\lambda^2 M = \frac{M}{2}\left(\lambda - \frac{\Delta}{M}\right)^2 - \frac{\Delta^2}{2M}$$

és a becslés egyenletes olyan $\underset{\sim}{f} \in F$ pontokban, melyekre $P(\underset{\sim}{f}) \geq P(\underset{\sim}{f}^{min}) + t$.

Tudjuk, hogy a λ lépéshossz nagysága $\min\limits_{0 \leq \lambda \leq 1} P(\lambda)$ meghatározásával történt, a fenti becslésben szereplő majoráló függvény minimumhelye pedig $\lambda = \frac{\Delta}{M} \geq 0$, és mivel Δ mindig választható ugy, hogy $\lambda = \frac{\Delta}{M} \leqq 1$ is teljesüljön, igaz lesz $P(\lambda) - P(0) \leqq -\frac{\Delta^2}{2M} = -\varepsilon < 0$.

Ez utóbbi megállapitás azt jelenti, hogy $P(\underset{\sim}{f})$ értéke minden egyes FD iterációban legalább ε-nal csökken, tehát

$$P(\underset{\sim}{f}^*) - P(\underset{\sim}{f}^n) = \sum_{\ell=n}^{\infty} P(\underset{\sim}{f}^{\ell+1}) - P(\underset{\sim}{f}^{\ell}) = -\infty$$

bármely n-re,

ami ellentmondás $\lim\limits_{n\to\infty} P(\underset{\sim}{f}^n) = P(f^*)$-gal. Q.E.D.

5.4 Megjegyzések

1. Ha P nem szigoruan konvex függvény az F halmazon, akkor olyan (heurisztikus) módszert kell találnunk a globális minimum meghatározásához, mely nem vizsgálja meg az összes lokális minimumhelyét a P függvénynek.
2. Igen egyszerü az FD módszer alkalmazása szigoruan konkáv vagy szigoruan kvázikonkáv P függvényre.

Definició:

a) A P függvény kvázikonkáv az $\hat{\underset{\sim}{f}} \in F$ pontban, ha minden olyan $\underset{\sim}{f} \in F$-re, melyre $P(\underset{\sim}{f}) \geqq P(\hat{\underset{\sim}{f}})$ fennáll:

$$P(\hat{\underset{\sim}{f}}) \leqq P[\hat{\underset{\sim}{f}} + \lambda(\underset{\sim}{f} - \hat{\underset{\sim}{f}})] \quad \text{ahol} \quad 0 \leqq \lambda \leqq 1$$

és

$$\hat{\underset{\sim}{f}} + \lambda(\underset{\sim}{f} - \hat{\underset{\sim}{f}}) \in F,$$

ha F konvex, egyébként az $F \wedge \{\underset{\sim}{f} - \hat{\underset{\sim}{f}}\}$ halmazra vonatkozik a feltétel.

b) A P függvény kvázikonkáv az F halmazon, ha minden $\underset{\sim}{f} \in F$ pontban kvázikonkáv.

c) Ha $\underset{\sim}{f} \neq \underset{\sim}{\hat{f}}$, $\lambda > 0$ és $P(\underset{\sim}{\hat{f}}) < P[\underset{\sim}{\hat{f}} + \lambda(\underset{\sim}{f} - \hat{f})]$

akkor a függvényt szigoruan kvázikonkávnak nevezzük.

Szigoruan kvázikonkáv függvény lokális minimumhelyei az F halmazon az F extremális pontjai, azaz "legrövidebb" m.c. folyamok, ennek következtében az FD iterációkban a lépéshossz: $\lambda = 1$ (vagyis mindig F extremális pontjain - legrövidebb m.c. folyamokon haladunk).

6. AZ OPTIMÁLIS IRÁNYITÁS PROBLÉMÁJÁNAK MEGOLDÁSA FD MÓDSZERREL

6.1 A megoldhatósághoz szükséges feltételek teljesülése

A 4. fejezetben már ismertettük a feladatot:

Adott $G(N,A)$ hálózati topológia (n csucs, b él)

$\underset{\sim}{C} = (C_1, \ldots, C_b)$ kapacitásvektor

$\Gamma = ((v_{ij}))_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ igénymátrix

mellett meghatározandó

$\underset{\sim}{f} \leqq \underset{\sim}{C}$ m.c. folyam,

melyre

$$T(\underset{\sim}{f}) = \frac{1}{\gamma} \sum_{i=1}^{b} \frac{f_i}{C_i - f_i} + \frac{1}{\gamma} \sum_{i=1}^{b} f_i p_i'$$

minimális.

Láttuk, hogy a $\underset{\sim}{f} \leqq \underset{\sim}{C}$ feltételt a célfüggvény büntető függvényként magában foglalja, ezért ha találtunk egy $\underset{\sim}{f} \leqq \underset{\sim}{C}$ megengedett megoldást, akkor ebből inditva az FD iterációkat, nem kell külön feltételként figyelembevenni a kapacitáskorlátot.

Mivel $T(\underset{\sim}{f})$ konvex függvény, és a feltételhalmaz konvex halmaz, ha a feladatnak van $\underset{\sim}{f}$ megengedett megoldása, akkor van $\underset{\sim}{f}^*$ optimális is, és erre az $\underset{\sim}{f}^*$ opt. megoldásra $f_i^* < C_i \quad i=1,\ldots,b$ azaz $\exists \varepsilon > 0$, melyre

$\underset{\sim}{f}^* \in F_\varepsilon = \{\underset{\sim}{f} \mid \underset{\sim}{f}$ m.c. folyam $f_i \leqq C_i - \varepsilon \quad i=1,\ldots,b\}$

Ez azt jelenti, hogy elég kis ε esetén elég az F_ε halmazon alkalmaznunk az FD módszert. Ezen viszont a $T(\underset{\sim}{f})$ függvény kielégiti az (5.3.) -beli konvergenciafeltételeket:

$\underset{\sim}{T}(f)$ kétszer differenciálható $\underset{\sim}{f} \in F_\varepsilon$ $T(\underset{\sim}{f})$ alulról korlátos függvény (konvex)

$$\frac{\partial T}{\partial f_k} = \frac{1}{\gamma}\left(\frac{C_k}{(C_k-f_k)^2} + P_k'\right) \geq 0 \qquad \underset{\sim}{f} \in F_\varepsilon$$

$$\frac{\partial T}{\partial f_i \partial f_k} \begin{cases} 0 & \text{ha } i \neq k \\ \frac{1}{\gamma}\frac{2C_k}{(C_k-f_n)^3} < \infty & \text{ha } i=k \text{ és } \underset{\sim}{f} \in F_\varepsilon \end{cases}$$

6.2 Algoritmus induló megoldás meghatározására

Induló megoldásnak a Γ igénymátrixnak megfelelő olyan $\underset{\sim}{f}$ m.c. folyamot (röviden: Γ-folyamot) kell keresnünk, amely a kapacitásfeltételeket is kielégiti.

Olyan Γ-folyamot, amelyről nem kivánjuk meg a kapacitásfeltételek teljesülését, könnyű találnunk. Az alábbiakban vázolt módszer olyan iteráció, mely Γ-folyamok olyan $\underset{\sim}{f}^0, \underset{\sim}{f}^1, \ldots$ sorozatát határozza meg, melyek durván szólva, "egyre közelebb visznek" a kapacitásfeltételek teljesüléséhez.

1. Kiindulásként tekintsük az $\underset{\sim}{f} \equiv 0$ pontban számolt legrövidebb folyamatot, azaz az

$$\ell_k = \left.\frac{\partial T}{\partial f_k}\right|_{f_k=0} = \frac{1}{\gamma}\left(\frac{1}{C_k} + p_k'\right)$$

metrikának megfelelő legrövidebb $\underset{\sim}{f}'$ Γ-folyamot. Legyen $n = 0$.

2. Ha már ismert az $\underset{\sim}{f}^n$ Γ- folyam, legyen

$$\sigma^n = \max_k \frac{f_k^n}{C_k}$$

3. Az iteráció befejeződik, ha $\sigma^n < 1$

4. Ha $\sigma^n \geq 1$, akkor legyen $N_{n+1} = (1+\varepsilon(1-\sigma^n))/\sigma^n$ ahol $0<\varepsilon<1$ alkalmasan rögzitett konstans, és legyen

$$\underset{\sim}{f}^{n+1} = \frac{1}{N_{n+1}} \text{ FD} \odot (N_{n+1}\underset{\sim}{f}^{n})$$

ahol az FD operátorhoz a $\underset{\sim}{v}$ folyamot és λ lépséhosszt a szokott módon határozzuk meg.

Mivel $N_{n+1} \cdot \sigma^{n} < 1$, az $N_{n+1} \cdot \underset{\sim}{f}^{n}$ folyam, mely $(N_{n+1} \cdot \Gamma)$-folyam, kielégiti a kapacitásfeltételeket. Heurisztikusan belátható, hogy ha egy, a kapacitásfeltételeket kielégitő folyamra alkalmazzuk az FD operátort, akkor az eredményül kapott folyam általában az előzőnél kevésbé használja ki a kapacitásokat, ami annak mellékeredménye, hogy az uj folyam a $T(\underset{\sim}{f})$ célfüggvény értékét csökkenti. Ezért formálisan általában az várható, hogy $\sigma^{n+1} < \sigma^{n}$. Az uj $\underset{\sim}{f}^{n+1}$ folyam pedig nyilván Γ-megengedett.

5. Az iteráció befejeződik, ha $\sigma^{n} < 1$, azaz találtunk egy <u>megengedett megoldást</u>,
illetve, ha megfelelően rögzitett Θ és δ pozitiv türési állandók mellett <u>egyszerre</u> teljesül, hogy

$$\left| \sum_{k=1}^{b} \ell_k \, (v_k - N_{n+1} \cdot f_k^{n}) \right| < \Theta$$

és

$$\left| \sigma^{n} - \sigma^{n+1} \right| < \delta$$

ez esetben az adott Θ, δ türés mellett a feladatnak <u>nincs</u> <u>megengedett megoldása.</u> Egyébként a 2-ik lépéstől folytatjuk az eljárást.

Ha meghatároztunk egy induló megengedett megoldást, az (5-ik fejezetben ismertetett) FD módszer II. fázisa következik, mely ennél a feladatnál elvezet az optimális megoldáshoz.

[5] szerint a még 21 csomópontból és 26 csatornából álló (az algoritmus során a csatornák számának kétszeresét kell venni, az irányitatlanság miatt) ARPA hálózatra az FD módszerrel $r_{ij} \equiv r = 1.187$ kbit/sec igénymátrix esetén T=0.2406 sec volt a minimális idő. A FORTRAN nyelvü program IBM 360/91 számitógépen 30 sec alatt futott le.

7. AZ FD MÓDSZER GYORSITÁSA AZ OPTIMÁLIS IRÁNYITÁS ÉS CSATORNAKAPACITÁS–TERVEZÉS PROBLÉMÁJÁNAK MEGOLDÁSA

7.1 Nem-elágazó folyamok, "nagy és kiegyensulyozott" hálózat

Definició:

az $\tilde{f} \in F$ m.c. folyamot nem-elágazónak nevezzük, ha $\tilde{f}^{ij}$ elemi folyam egy uton folyik minden (i,j) párra.

Adódhat olyan feladat, melyben feltételként szerepel a megoldás nem-elágazó tulajdonsága. Más feladatokban egy nem-elágazó folyam igen jó közelitése lehet az optimális megoldásnak, s nyilván jelentősen gyorsul az FD algoritmus, ha a feladatot eleve nem-elágazó folyamokra korlátozzuk.

Jelölje F^{nb} a nem elágazó m.c. folyamok halmazát:

$$F^{nb} = \{f \mid \tilde{f} \in F, \ \tilde{f} \text{ nem elágazó}\}$$

A nem-elágazási feltételt diszkrét változók bevezetésével irhatjuk fel: Minden (i,j) párra ki kell zárni azokat a lehetséges $\tilde{f}^{ij}$ elemi folyamokat, melyek az (i,j) utak kombinációira oszlanak el. Ez már tizes nagyságrendü csomópontból álló hálózatban is olyan nagyszámu diszkrét feltételhez vezet, hogy diszkrét programozási módszerrel a feladat megoldása igen nehézkessé válik. Folytonos módszerek pedig az F^{nb} diszkrét halmazon nyilván nem alkalmazhatók.

Bevezetjük a "nagy és kiegyensulyozott" hálózat fogalmát, és megmutatjuk, hogy egy ilyen hálózatban az FD módszer speciális egyszerüsitésével haladhatunk csak nem-elágazó folyamokon a nem-elágazó optimális megoldás felé, és ezzel az optimumot igen jól tudjuk közeliteni.

A nagy és kiegyensulyozott hálózat jellemzése érdekében néhány mérőszámot definiálunk:

Legyen $$r = \frac{1}{(n-1)n} \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij}$$

a hálózatbeli átlagos igény,

és legyen $$m = \max_{i,j} \frac{r_{ij}}{r}$$

Nyilván $r \geqq \min_{i,j} r_{ij}$ és $m \geqq 1$,

és egyenlőség csak $\Gamma = ((r_{ij} \equiv r))_{i \neq j}$ igénymátrix esetén lehetséges.

Legyen $$\bar{p}' = \frac{\sum_{j=1}^{n} \sum_{j=1}^{n} r_{ij} p'_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij}}$$

ahol p'_{ij} az i-ik csomópontból a j-ik csomópontba vezető legrövidebb ut hossza (ut hossza: a benne levő ivek száma)

azaz $\bar{p}'$ az átlagos uthossz abban az esetben ha $\underset{\sim}{f}^{ij}$ a legrövidebb (i,j) elemi folyam /minden (i,j) párra/, vagyis r_{ij} a legrövidebb uton jut el i-ből j-be /minden (i,j) párra/.

Tetszőleges $\underset{\sim}{f} \in F$ esetén, ha p_{ij} az $\underset{\sim}{f}^{ij}$ elemi folyamhoz tartozó uthossz, az átlagos $\bar{p}$ uthossz:

$$\bar{p} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} p_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij}}$$

Vizsgáljuk meg, mit mondhatunk a hálózat élein a teljes folyam és (i,j) elemi folyam arányáról /minden $(i,j)_{i\neq j}$ párra/: azaz vizsgáljuk az

$$\frac{f_k}{f_k^{ij}} \quad \text{értékeket } k = 1,\ldots,b$$

illetőleg ezek $\frac{1}{b}\sum_{k=1}^{b}\frac{f_k}{f_k^{ij}} = \overline{\left(\frac{f_k}{f_k^{ij}}\right)}$ átlagát:

Mivel $f_k^{ij} \leqq m.r$

$$\overline{\left(\frac{f_k}{f_k^{ij}}\right)} \geqq \frac{1}{b\cdot m\cdot r}\sum_{k=1}^{b} f_k = \frac{1}{b\cdot m\cdot r}\sum_{i=1}^{n}\sum_{j=1}^{n} r_{ij}p_{ij} =$$

$$= \frac{1}{b\cdot m\cdot r}\, r(n-1)n\cdot\bar{p} \geqq$$

$$\geqq \frac{n(n-1)}{b\cdot m}\,\bar{p}' = \frac{1}{\eta}$$

(7.1.2) *Definició:* A $G=(N,A)$ n számu csoportból, b számu élből álló hálózatot, a $\Gamma = ((r_{ij}))_{\substack{i=1,\ldots,n\\ j=1,\ldots,n}}$ igénymátrixnak megfelelő m.c. folyamokkal, nagy és kiegyensulyozott hálózatnak nevezzük, ha $\eta \ll 1$.

Ilyen hálózatban ugyanis $\frac{\bar{f}_k}{f_k^{ij}} \geqq \frac{1}{\eta}$ miatt átlagosan egy-egy élen az (i,j) elemi folyam elhanyagolható a teljes folyamhoz képest.

FD módszer nagy és kiegyensulyozott hálózatban

Legyen $\underset{\sim}{f} \in F^{nb}$, $\Gamma = ((r_{ij}))$ igénymátrix

$\{\ell_k = \frac{dP}{df_k}\Big|_{\underset{\sim}{f}}$ $k=1,\dots,b\}$ a szokásos metrika

$\P_{ij}$ az $\underset{\sim}{f}^{ij}$-hoz tartozó ut

$\P'_{ij}$ az $\{\ell_k\}$ metrikában legrövidebb (i,j) ut

$\underset{\sim}{v}$ az $\{\ell_k\}$ metrikában legrövidebb folyam

Legyen az FD operátor olyan, hogy λr_{ij} mennyiséget $(0 \leqq \lambda \leqq 1)$ "tesz át" a $\P_{ij}$ utról a $\P'_{ij}$ utra.

Vizsgáljuk most a $P(\lambda) = P[\underset{\sim}{f}\,(1-\lambda) + \underset{\sim}{v}\,\lambda]$ függvényt.

$$P(\lambda) = P(0) + \lambda \sum_{k=1}^{b} \ell_k (v_k - f_k) + \sigma[\lambda(\underset{\sim}{v} - \underset{\sim}{f})]$$

ahol $\sigma[\lambda(\underset{\sim}{v} - \underset{\sim}{f})]$ tartalmazza $P(\lambda)$ Taylor sor szerinti kifejtésében a magasabbrendü tagokat.

Ha a hálózat nagy és kiegyensulyozott, az $f_k^{ij} \ll f_k$ $k=1,\dots b$

Mivel $\underset{\sim}{f}$ és $\underset{\sim}{v} \in F^{nb}$, a $(v_k - f_k)$ különbség infinitezimális.

Az 5. fejezetben láttuk, hogy $\sum_{k=1}^{b} \ell_k (v_k - f_k) \leqq 0.$

Ha $\sum_{k=1}^{b} \ell_k (v_k - f_n)$ elegendően negativ ahhoz, hogy

$\sigma[\lambda(\underset{\sim}{v} - \underset{\sim}{f})]$ elhanyagolható legyen, akkor

$$\min_{0 \leq \lambda \leq 1} P(\lambda) = P(1)$$

azaz az FD operátorral $\underset{\sim}{f}$-ről teljesen áttérünk $\underset{\sim}{v}$-re:

$$FD \odot \underset{\sim}{f} = \underset{\sim}{v}$$

s erről tudjuk, hogy nem-elágazó folyam: $\underset{\sim}{v} \in F^{nb}$.

Ha viszont $\sum_{k=1}^{b} \ell_k (v_k - f_n) \approx 0,$ akkor

$\lambda_{min} < 1$, azaz a $FD \odot \underset{\sim}{f} = \underset{\sim}{f}(1-\lambda_{min}) + \underset{\sim}{v} \cdot \lambda_{min}$ folyam már

elágazó. Ugyanakkor $\sum_{k=1}^{b} \ell_k (v_k - f_n) \approx 0$ azt jelenti, hogy

az optimális megoldás közelében vagyunk (ld. 5. fejezet!), vagyis ekkor az $\underset{\sim}{f}$ nem-elágazó folyam előre rögzithető hibakorlát szerinti közelitése az optimális, de elágazó m.c. folyamnak.

A "nem-elágazó FD módszer" főbb lépései

I. fázis: A konkrét feladattól függő módon meghatározunk egy megengedett (induló) megoldást.

II. fázis:

Indulás az I. fázis végén kapott $\underset{\sim}{f}^{o}$ m.c. folyammal.

Az iteráció n-edik lépése:

1. Meghatározzuk az $\left\{ \ell_k = \frac{\partial T}{\partial f_k}\Big|_{\underset{\sim}{f}=\underset{\sim}{f}^n} \quad k=1,\dots,b \right\}$

metrikában legrövidebb Π_{ij}' $\quad i=1,\dots,n \quad j=1,\dots,n$

utakat.

2. $\underset{\sim}{g} = \underset{\sim}{f}^n$ és minden (i,j) párra elvégezzük a következőket.

2.a) $\underset{\sim}{g}$ folyamot ugy változtatjuk, hogy a $\underset{\sim}{g}^{ij}$ mennyiséget a $\P'_{ij}$ utra tesszük át.

2.b) Ha a kapott folyam megengedett, és egyuttal a célfüggvény értéke is csökken, ezzel a folyammal folytatjuk az eljárást 2.a) lépéstől a következő (i,j) párra.

Ha nem-megengedett folyamhoz jutottunk, vagy a célfüggvény nem csökken, az előző folyammal folytatjuk az eljárást /2.a) lépéstől a következő (i,j) párra/.

3. Ha már minden (i,j) párra sor került, két eset lehetséges:

$\underset{\sim}{g} \equiv \underset{\sim}{f}^n$ változatlan maradt, ekkor $\underset{\sim}{f}^n$ tovább nem javitható nem elágazó megoldás, az eljárás véget ért.

$\underset{\sim}{g} \not\equiv \underset{\sim}{f}^n$, tehát találtunk nem-elágazó, $\underset{\sim}{f}^n$-nél jobb megoldást, ezzel inditjuk az (n+1)-ik iterációt.

A nem-elágazó folyamok száma véges, igy a módszer mindenképpen véges lépésből áll.

A 6-ik fejezet végén emlitett hálózatban (21 csomópont 26·2=52 csatorna)

$$\eta = \frac{52}{21\cdot 20} \cdot \frac{1}{\bar{p}'} < \frac{52}{21\cdot 20\cdot 21} \ll 1$$

a nem-elágazó FD módszer ugyanazon az IBM 360/91 számitógépen már 4 mp alatt igen jól megközelitette a "pontos" FD módszerrel (30 sec alatt) kapott optimumot: $T_{min} = 0.2438$ ($T_{min} = 0.2406$)

7.2 Az optimális irányitás és kapacitástervezés probélmájának megoldása lineáris költség-kapacitás függvény esetén

A feladatot a 4-ik fejezetben ismertettük, és láttuk, hogy a kétváltozós $T(\underset{\sim}{C},\underset{\sim}{f})$ célfüggvényt először rögzitett $\underset{\sim}{f}$ mellett $\underset{\sim}{C}$-re optimalizálva, a feladat következő alakjához jutunk!

Adott $G=(N,A)$ hálózati toplógia

Γ igénymátrix

$d(\underset{\sim}{C}) = \sum_{i=1}^{b} d_i C_i$ költség-kapacitás függvény

és D költség korlát mellett meghatározandó az az $\underset{\sim}{f} \in F$ m.c. folyam, melyre

$$T(\underset{\sim}{f}) = \frac{1}{\gamma} \frac{\left(\sum_{i=1}^{b} \sqrt{f_i d_i}\right)^2}{D - \sum_{i=1}^{b} f_i d_i} + \frac{1}{\gamma} \sum_{i=1}^{b} f_i p_i'$$

minimális.

A célfüggvény az $\underset{\sim}{f}$ m.c. folyamra tett $\sum_{i=1}^{b} f_i d_i$ D feltételt büntetőfüggvényként magában foglalja.

Hasonlóan a (4.1) feladat célfüggvényéről a 6. fejezetben tett meggondoláshoz, a $T(f)$ függvényt elég az

$$F_{2,\varepsilon} = \{f \mid \underset{\sim}{f} \text{ m.c. folyam, } \sum f_i d_i \quad D-$$

halmazon vizsgálni, ahol $\varepsilon > 0$ alkalmasan rögzitett konstans.

Ekkor $T(\underset{\sim}{f})$ alulról korlátos és kétszer differenciálható függvény az $F_{2,\varepsilon}$ halmazon,

$$\frac{\partial T}{\partial f_k} = \frac{1}{\gamma}\left(\frac{\Sigma\sqrt{f_j d_j}}{D - \Sigma f_j d_j}\right)\sqrt{\frac{d_k}{f_k}} + \frac{1}{\gamma}\left(\frac{\Sigma\sqrt{f_j d_j}}{D - \Sigma f_j d_j}\right)^2 d_k + \frac{1}{\gamma} p_k'$$

$$\frac{\partial T}{\partial f_k} \geqq 0 \qquad \underset{\sim}{f} \in F_{2,\varepsilon}$$

kiszámolva a második parciálisokat, adódik $\dfrac{\partial T}{\partial f_i f_k}$ végessége

$$i=1,\dots,b$$
$$k=1,\dots,b$$

tehát az $F_{2,\varepsilon}$ halmazon az FD módszer alkalmazható.

Megjegyzés:

$$\lim_{f_k \to 0} \frac{\partial T}{\partial f_k} = \infty$$

ami azt jelenti, hogy ha egy FD iterációban valamely k-ra $f_k=0$, azaz valamely élen a folyam zérussá redukálódik, ennek az élnek a hossza az $\{\ell_k\}$ metrikában végtelen, ezért a további FD iterációban sem nőhet zérus fölé az f_k folyam. A célfüggvénynek az a tulajdonsága jól felhasználható, ha a hálózat topológiájának megtervezése a feladatunk: adott csomópontok mellett a teljes gráfból indulunk és az FD módszerrel állapitjuk meg, mely éleket hagyunk el a gráfból. [5], [10].

A $T(\underset{\sim}{f})$ függvény (4.2*.1.) alakját lineáris költség-kapacitás függvény esetén kaptuk. Bebizonyitható, hogy ekkor de tetszőleges konkáv $d_i(C_i)$ függvények esetén is, $T(\underset{\sim}{f})$kvázikonkáv lesz az $F_2 = \{\underset{\sim}{f} \mid \underset{\sim}{f}$ m.c. folyam, $\Sigma f_i d_i \leqq D\}$ halmazon. [5].

Tehát, mint ahogy az 5-ik fejezetben már emlitettük, az FD módszer nagyon leegyszerüsödik: $\lambda = 1$ lesz az egyes iterációkban, vagyis mindig extremális pontokon $\equiv$ legrövidebb folyamokon haladunk. A legrövidebb folyamok viszont nem-elágazóak vagyis egy nem-elágazó FD módszerrel dolgozunk.

A (4.2*) feladat megoldásakor két problémánk van még:

1) Induló megengedett megoldás meghatározása

2) A módszer véges ugyan, de csak lokális minimumhoz vezet

[5]-ben mindkét problémára igen heurisztikus megoldást javasolnak:

I. fázis:

Induló megoldás keresése:

a) a hálózat éleihez tetszőleges, de egyenlő hosszt rendelünk.

b) az igy kapott metrikában meghatározzuk a legrövidebb Γ--megengedett m.c. folyamot

c) Ha erre a folyamra $D - \sum_{i=1}^{b} f_i d_i > 0$ is teljesül, akkor

F_2-megengedett folyamunk van, indulhat a II. fázis.

Ellenkező esetben a folyamot ejtjük, és másik metrikával folytatjuk az eljárást az a) lépéstől.

II. fázis - lokális opt.

Az I. fázisból kapott megoldással indulunk.

Az n-ik iteráció főbb lépései:

1) $\underset{\sim}{f}^{n+1} = FD \odot \underset{\sim}{f}^{n}$

ahol

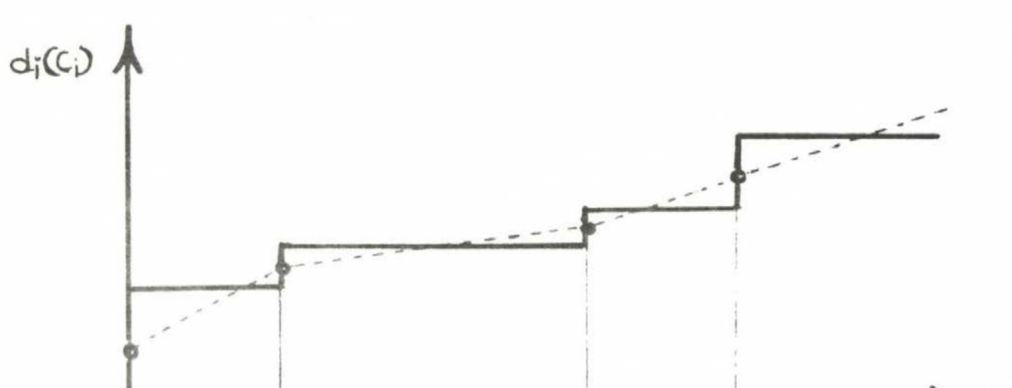
$\underset{\sim}{f}^{n+1}$ az $\left\{ \ell_k^n = \frac{\partial T}{\partial f_k} \Big|_{\underset{\sim}{f}=\underset{\sim}{f}^n} \quad k=1,\dots,b \right\}$ metrikában meghatározott legrövidebb folyam

2) Ha $T(\underset{\sim}{f}^{n+1}) < T(\underset{\sim}{f}^n)$, következik az (n+1)-ik iteráció. Ellenkező esetben lokális minimumhoz értünk, igy a II. fázis véget ért.

II. fázis - szuboptimális megoldás

A kapott lokális minimum függ az I. fázis által adott, induló megoldástól. Ezért az I. és II. fázist egymás után elég sokszor végrehajtjuk, és a véletlen módon generált induló megoldásokból kapott lokális minimumok minimumát fogadjuk el u.n. szuboptimális megoldásnak.

Az ARPA hálózatban felmerült (4.2) tipusu feldatban a költség-kapacitásfüggvény diszkrét (lépcsős) függvény volt, melyet szakaszonként lineáris függvénnyel közelitettek.



A kapott lokális optimumot ezért még hozzá kellett igazitani az eredeti költség függvényhez. Ezt ugy végezték, hogy lokálisan optimális megoldásban kapott élenkénti C_i kapacitásértékeket a lépcsősfüggvény ugráshelyeinek megfelelően meg-

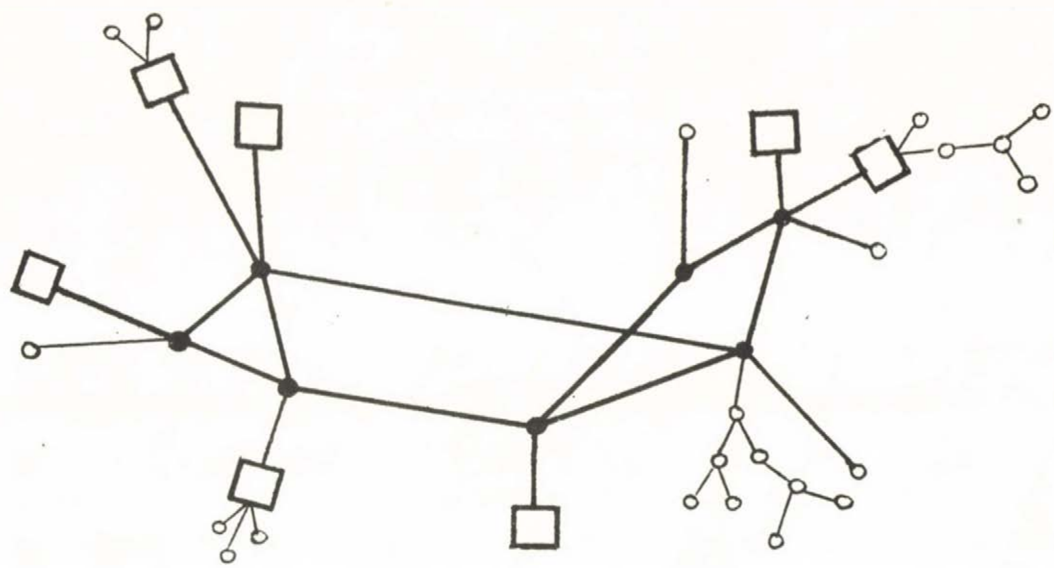
növelték (a költség két szomszédos ugráshelye közötti intervallumban nem változik!), és igy ujabb $\underset{\sim}{f}$ megoldást határoztak meg FD algoritmussal.
Természetesen nem állithatjuk, hogy ez utóbbi megoldás (lokálisan) optimális, de nem is definiálhatjuk pontosan, mit értsünk lépcsős függvény esetén lokális opitmumon.
[5]-ben a szerzők szerint eljárásuk nem rosszabb, mint más, hasonló esetben alkalmazott heurisztikus technikák.

FÜGGELÉK

F.1. Osztott számitógéphálózatok általános jellemzése

Egy osztott számitógéphálózat funkcionálisan két részből áll: a programfeldolgozást végző "erőforrás" számitógépekből és a hozzájuk tartozó helyi és távoli terminálokból, valamint az összeköttetést megvalósitó adatátviteli hálózatból (communication/data network), mely utóbbin - tágabb értelemben - nemcsak a technikai berendezéseket, hanem minden olyan hardware-t, software-t és matematikai módszert értünk, melyek az adat-továbbitást hivatottak biztositani minél nagyobb megbizhatósággal és különféle szempontok szerint optimális költséggel.

Sematikusan igy ábrázolható egy osztott számitógéphálózat:



1. ábra

Fekete korongokkal és vastag vonalakkal jelöltük az adatátviteli hálózat részeit: ennek csomópontjai a kapcsolószámitógépek, (switching-computer) élei nagysebességü (50-200 kbit/sec) csatornák.
A feldolgozó ("erőforrás") számitógépek (világos négyzetekkel jelölve) a kapcsoló számitógépeken keresztül kötődnek az adatátviteli hálózathoz, szintén nagysebességü csatornákkal. A feldolgozó számitógépekhez terminálok (világos körök) is tartozhatnak, de egy terminál közvetlen összeköttetésben is állhat az adatátviteli hálózattal - "kis"sebességü vonalak segitségével.

A teljes számitógéphálózatban bizonyos csomópontokban üzenetek (programok/adatok v. ezek fix hosszuságu része) jelentkeznek megadott csomópontban lévő számitógéphez való eljuttatás, abbeli feldolgozás céljából. Igy egy üzenet szempontjából egy csomópont lehet "feladóhely" (ez terminál v. számitóközpont), továbbitóhely (kapcsoló számitógép), és rendeltetési hely (számitóközpont).

Az adatátviteli rendszer alapján a hálózat lehet vonalkapcsolt (circuit/line switching), vagy csomagkapcsolt (message/packet-switching)

Vonalkapcsolt hálózatban a kapcsoló számitógépek nem rendelkeznek tárolókapacitással, igy egy üzenet eljuttatása a feladóhelyről a rendeltetési helyre csak ugy lehetséges, ha a két hely közötti teljes vonal szabad. A feladóhely egy speciális jelet bocsájt ki, amely valamilyen uton eljut a rendeltetési helyre, miközben lefoglalja az utat alkotó összes csatornát. Ezután a rendeltetési helyről visszajelzés érkezik, s megkezdődhet a tényleges üzenettovábbitás. A felhasznált csatornák csak akkor szabadulnak fel, amikor a továbbitás teljesen befejeződött, függetlenül attól, hogy az adott uton az üzenet melyik csatornán tart éppen.

A másik fajta adatátviteli rendszerben minden kapcsoló számitógép képes üzenet-tárolásra. Aszerint, hogy az üzenet egy teljes

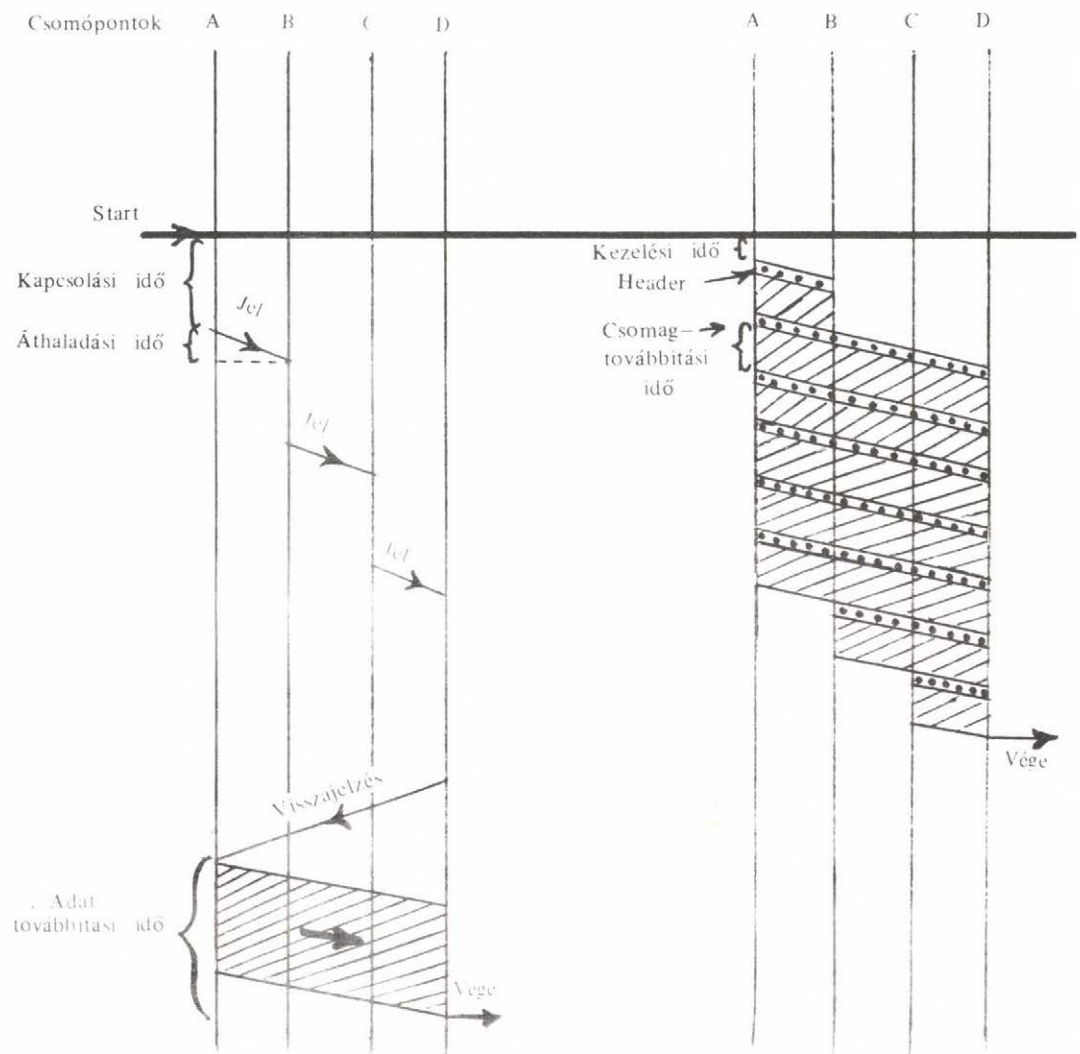
program/adathalmaz, vagy ennek csak fix hosszusägu része(egy "csomag"), az angol nyelvü irodalomban "message-switching" ill. "packet-switching" elnevezést visel a hálózat, Magyarországon a "csomagkapcsolt" név van elterjedőben. (A csomagokra bontás javitja a rendszer karakterisztikáját - megbizhatóságát, gyorsaságát, stb.)
Csomagkapcsolt hálózatban egy üzenet továbbitásához adott időben csak egy csatorna lefoglalása szükséges. Az üzenet először a feladóhelytől a legközelebbi kapcsoló pontig halad, s mikor ide teljesen megérkezett, hibaellenőrzés után meghatározzák, milyen utvonalon, azaz mely kapcsoló pontokon keresztül haladjon a rendeltetési hely felé. Ha két kapcsolópont közti csatorna foglalt, az üzenet vár (tárolják), sőt sorbaáll, s ez igy folytatódik, mig a rendeltetési helyére nem ér.

E hálózat másik neve, mely müködéséről többet árul el: "raktározd és továbbitsd" - store and forward, röviden S/F - hálózat.

Látjuk tehát, hogy itt az üzenetkezelés több adminisztrációt kiván a hálózat minden pontjában, mint a vonalkapcsolásos rendszerben, hiszen pl. az üzeneteket el kell látni a feladó- és rendeltetési helyre vonatkozó információval, ("header") ami - különösen packet-switching esetén - növeli az üzenet hosszt; ugyancsak packet-switching esetén a teljes üzenet (message) csomagokra bontása, majd a rekonstruálás is többlet feladat. Óriási előny viszont, hogy a csatornák állandóan továbbitásra kész állapotban vannak, egy csatornán egymás után haladhatnak különböző feladó- és rendeltetési helyü üzenetek.

Nem dönthető el egyértelmüen, melyik adatátviteli rendszer jobb. Két számitógéphálózat nagyon különbözhet egymástól a felhasználók mennyiségi és minőségbeli igényeiben, igy a forgalom nagyságában, időbeli eloszlásában, stb. Már egyetlen számitógéphálózat is lehet ebben az értelemben változó jellegü.

Mégis, az eddigi tapasztalatok alapján, nem tul hosszu üzenetek esetén csomagkapcsolt hálózatban rövidebb a továbbitási idő, és különösen packet-switching esetén megbizhatóbb a rendszer, igy jobb a feldolgozást végző számitógép kihasználtsága is. Tehát mind a felhasználók, mind az üzemeltetők számára a csomagkapcsolt rendszer előnyösebb.



2. ábra

Üzenettovábbitás /a/ vonalkapcsolt hálózatban

/b/ csomagkapcsolt hálózatban
/packet switching/

F.2 NÉHÁNY SZÓ AZ ARPA SZÁMITÓGÉPHÁLÓZATRÓL

Napjainkban a legnagyobb müködő számitógéphálózat az Egyesült Államok csaknem egész területét átfogó ARPANET, melyet az Advanced Research Project Agency teremtett meg (többek közt a Szovjetunió ürkutatási kisérleteinek hatására) az USA Nemzetvédelmi Minisztériumának megbizásából.

Tervezése 1967-ben kezdődött, és müködésének hajnalán, 1969. végén 4 kapcsolócsomópontból állt. A fejlődés ütemének érzékeltetésére néhány további adat:
1972 szeptemberében 34, egy évvel később 39 kapcsolópont müködött. 1975. juniusában pedig 57 csomópontból állt az adatátviteli hálózat, mintegy 100, földrajzilag igen változatos elhelyezkedésü "erőforrás" számitógép (- Európában és Hawaiban lévő számitóközpontok is vannak közöttük -) összekapcsolását valósitva meg.

Nem célunk részletesen ismertetni a jelenlegi ARPA hálózatot, csupán azokat a jellemzőket tekintjük át, melyek a felmerülő matematikai problémák megértését segitik elő. Az ARPANET-tel foglalkozó irodalom meglehetősen gazdag, magyar nyelven igen jó a KFKI "Számitógéphálózatok" c. szemináriumi sorozatában Harangozó József ismertetése.

Az ARPA hálózat osztott és csomagkapcsolt. Topológiailag a hálózat osztottsága azt jelenti, hogy (az adatátviteli hálózatban) bármely két csomópont között legalább két, fizikailag szeparált ut létezik. - Az egymástól igen sok tekintetben (gyártó cég, specializált software, adatbázis-rendszer) inkompatibilis erőforrás számitógépek - közös néven HOST-ok mindegyike 100 kbit/sec sebességü vonalon keresztül egy-egy kisteljesitményü kapcsolószámitógéppel, az IMP-pel (Interface Message Processor) áll közvetlen összeköttetésben. Az IMP-ek egymással 50 kbit/sec duplex csatornákkal kapcsolódnak.

Az ARPANET korai változatában az adatátviteli hálózat Honeywell 516 tipusu IMP-ekből (12 K/core memória, 0.96 msec ciklusidő) és a már emlitett 50 kbit/sec duplex csatornákból állt. 1975-ben már két 230.4 kbit/sec vonal is müködött Californiában, Hawaival 50 bit/sec szputnyik valósitotta meg az összeköttetést. Jelentősebb azonban a kapcsolószámitógépekben történt változás: a rendszer hatékonyságának növelése érdekében meg kellett teremteni annak lehetőségét, hogy a felhasználó közvetlenül a hálózatba kapcsolódva léphessen összeköttetésbe az általa kiválasztott HOST-tal. Ha a felhasználó a tőle nagy távolságban lévő HOST-ot ugy éri el, hogy egy hozzá közeli HOST termináljához kapcsolódik, s ez utóbbi HOST közvetiti tovább az eredetileg kiválasztott HOST-nak, akkor a HOST-ok egyrészt gyakran csak egyszerü továbbitó szerepet töltenek be, tehát kihasználtságuk minősége romlik, másrészt a vonalkoordinálást (terminál-karakterekből üzenet átalakitás és viszont) kell végezniük, azaz munkájuk ésszerütlenül megnő.

Ezért egy uj csomóponti számitógéptipust fejlesztettek ki, a Terminál IMP-et, vagy röviden TIP-et, mely tehát közvetlen terminál kapcsolatot biztosit az adatátviteli hálózattal.
A TIP-ek Honeywell 316 (28 K/core, 75 usec/character) kis számitógépek.

Az ARPA hálózatban előforduló logikai és fizikai összeköttetések fajtái: HOST-HOST, terminal-HOST, terminal-TIP, HOST-IMP, TIP-IMP, IMP-IMP.

Tekintsük át a HOST-ok közötti üzenetforgalom mechanizmusát:

A feladóhely (HOST, vagy terminal) az üzenetet, a rendeltetési helynek megfelelő információval ellátva, saját IMP-jének adja át. Az üzenet bitjeinek számát az IMP memóriakapacitása korlátozza: az ARPANET-ben 8063 bit a felső határ. Az IMP hibaellenőrzés után megváltoztatja a feladótól kapott üzenet formátumát, sebességét és egyenlő (1008 bit) hosszuságu részekre szabdalja, melyeket még "megfejel" a hálózat számára a továbbitáshoz szükséges információkkal. Igy keletkezik a "csomag" (packet). Az egyes csomagok egymástól függetlenül jutnak el a rendeltetési hely IMP-jéhez, ahol ez az IMP összeállitja az eredeti üzenetet, és miután a rendeltetési HOST-hoz továbbitotta, jelzi a feladó HOST-nak az üzenet megérkezését.Csak ezután alakulhat ki ujabb logikai ut a két HOST között, nehogy az IMP-ben az üzenetek összetorlódjanak. A csomagok az adatátviteli hálózatban IMP-ről IMP-re vándorolnak. Az ut minden IMP pontján hibaellenőrzést végeznek, s ha egy csomag meghibásodott, az előző IMP-ben tárolt másolat alapján megismétlik a küldést. Természetesen az is előfordulhat, hogy egy vagy több csomag "elvész"" a hálózatban IMP hiba miatt , ekkor a cél-IMP ezt jelzi a feladónak, és az egész üzenet küldését megismétlik a feladó IMP-jében tárolt másolat alapján.

Érdekességként megemlitjük, hogy egy IMP a csomagfeldolgozás mellett más feladatokat is ellat:

- ellenőrzi a hálózat állapotát (HOST és IMP hibák megállapitása, lehetőség szerint korrekció)

- statisztikai adatokat gyüjt (a tervezők számára, módszereik helyességének, alkalmazhatóságának ellenőrzéséhez) a következőkről:

  a várakozó sorok hossza
  az üzenetek feladó- és rendeltetési helye 0.5 mp-ként

  a feldolgozott üzenetek száma
  a HOST-okkal lebonyolitott forgalom 10 mpként
  az üzenetek megérkezésének ideje

IRODALOM

[1] G. Cantor - M. Gerla:
Optimal Ronting in a Packet-switched Computer Network
/IEEE Transactions on Computers, Vol /C-23, No 10, 1974, 1062 - 1068/

[2] G.B. Danzig:
Linear Programming and Extensions
/Princeton University Press, 1963/

[3] H. Frank - W. Chou:
Routing in Computer Networks
/Networks 1, 1971, 99-112/

[4] H. Frank - R.E. Kahn - L. Kleinrock:
Computer Communication Network Design - Experience with Theory and Practice /AFIPS Conference Proceedings S CC 1972, AFIPS Press 1972, 255-270/

[5] L. Fratta - M. Gerla - L. Kleinrock:
The Flow Deviation Method: An Approach to Store - and - Forward Communication Network Design
/Networks 3, 1973. 97-133/

[6] L. Fratta - U. Montanari:
Analytical Techniques for Computer Networks /Analysis and Design. /Computer Architectures and Networks, North Holland Publ. Company 1974, 155-185/

[7] G. Hadley: Nonlinear and Dynamic Programming
/Addison - Wesley, 1964/

[8] T.C. Hu: Integer Programming and Network Flows
/Addison - Wesley, 1969/

[9] I.F.C. Kingman:
Some Inequalities for the quene GI/G/1. /Biometria, 1962, 315-324/

[10] L. Kleinrock:
Analytic and simulation Methods in Computer Network Design. /AFIPS Conference Proceedings SJCC 1970, AFIPS Press 1970, 569-579/

[11] L.Kleinrock: Queuing Systems, Volume 1: Theory /Wiley, 1975/

[12] L. Kleinrock: Queuing Systems, Volume 2: Computer Applications /Wiley, 1976/

[13] S.M. Ornstein - F.E. Heart - W.R. Crowther - H.K. Rising - S.B. Russel - A. Michel:
The Terminal IMP for the ARPA Computer Network
/AFIPS Conference Proceedings SJCC 1970, AFIPS Press 1972, 243-254/

[14] Harangozó József: Az ARPA számitógéphálózat
/KFKI Számitógép-hálózatok 3. szemináriumi füzet/

[15] Manigáti Csaba:"Számitógépes hálózatok modellezésének néhány matematikai problémája"
/KFKI Számitógéphálózatok 7. szemináriumi füzet/

[16] Záray Éva: Számitógépes hálózatok modellezésének néhány matematikai problémája I.
/KFKI Számitógép-hálózatok 6. szemináriumi füzet/