

Polunhakualgoritmit ja -järjestelmät

Rodion Efremov

Kandidaatintutkielma-aine
HELSINGIN YLIOPISTO
Tietojenkäsittelytieteen laitos

Helsinki, 27. marraskuuta 2014

Tiedekunta — Fakultet — Faculty		Laitos — Institution — Department	
Matemaattis-luonnontieteellinen		Tietojenkäsittelytieteen laitos	
Tekijä — Författare — Author			
Rodion Efremov			
Työn nimi — Arbetets titel — Title			
Polunhakualgoritmit ja -järjestelmät			
Oppiaine — Läroämne — Subject			
Tietojenkäsittelytiede			
Työn laji — Arbetets art — Level		Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Kandidaatintutkielma-aine		27. marraskuuta 2014	14
Tiivistelmä — Referat — Abstract			
<p>Haettaessa lyhimpiä polkuja verkossa joudutaan väistämättä ottamaan kohdeverkon ominaisuuksia huomioon, mikäli halutaan suoriutua tehtävästä pienimmässä mahdollisessa ajassa. Toisinaan on mahdollista määritellä heuristiikkafunktio, ja siten käyttää A*-perheen algoritmeja; toisinaan tätä ei voida tehdä, ei ainakaan tehokkaasti, jolloin jäljelle jää kaksisuuntaisuus ja/tai verkon esiprosessointi haun nopeuttamiseksi. Tämän dokumentin tarkoituksena on siis antaa (hyvin suppea) katsaus asiaan liittyviin algoritmeihin ja järjestelmiin.</p>			
Avainsanat — Nyckelord — Keywords			
verkot, lyhimmat polut			
Säilytyspaikka — Förvaringsställe — Where deposited			
Muita tietoja — Övriga uppgifter — Additional information			

Sisältö

1	Johdanto	1
2	Tavallisimmat algoritmit	1
2.1	Prioriteettijonon valinta	3
3	Kaksisuuntainen haku	4
3.1	Kaksisuuntainen Dijkstran algoritmi	5
3.2	Kaksisuuntainen A*	5
4	Kaikkien parien lyhimät polut	8
5	Ruudukkoverkko ja jump point -haku	9
6	Dijkstran algoritmi kaarivivuilla	9
6.1	Ositustekniikat	10
6.2	Tulokset	11
7	Polunhaku ja Multiple Sequence Alignment	
	-ongelma	11
7.1	Ongelman muotoileminen verkkona	12
7.2	Ratkaisu	12
8	Lyhimät polut ja rinnakkaisuus	12
	Lähteet	13

1 Johdanto

Polunhaku painotetuissa tai painottamattomissa verkoissa on perustavanlaatuisen ongelma, joka ei ole mielenkiintoinen vain itsessään, vaan myös tarvittavana alioperaationa muissa algoritmeissa. Esimerkiksi Edmond-Karpin algoritmi käyttää leveyssuuntaista hakua maksimivuo-ongelman ratkaisemiseksi [1]; multiple sequence alignment -ongelmaa on ruvettu viime vuosikymmeninä ratkomaan myös heuristisin polunhakualgoritmein.

Verkko G on pari (V, A) , jossa V on solmujen joukko ja $A \subset V \times V$ on (suunnattujen) kaarien joukko. Suuntaamatonta verkkoa $G' = (V, E)$ voidaan aina simuloida suunnatulla verkolla $G = (V, A)$ siten, että jokaista suuntaamatonta kaarta $\{u, v\} \in E$ kohti laitetaan A :han kaaret (u, v) ja (v, u) . (Suunnattu verkko on suuntaamattoman yleistys.) Polunhakua varten verkosta erotellaan kaksi solmua: lähtösolmu s ja maalisolmu t . Jatkossa, $n = |V|$ ja $m = |E|$; näin esimerkiksi leveyssuuntaisen haun aikavaativuus on $O(n + m)$. Polku on $\gamma_k = \langle u_0, u_1, \dots, u_k \rangle$, missä mikään solmu ei esiinny yhtä kertaa enempää, ja verkossa on kaari (u_i, u_{i+1}) jokaisella $i = 0, 1, \dots, k - 1$. Polkuun liittyvä kustannus on sen kaarien painojen summa, ja mitä tulee itse painoihin, niiden oletetaan olevan ei-negatiivisia. Ei-painotettujen verkkojen kohdalla jokaisen kaaren paino oletetaan olevan 1.

2 Tavallisimmat algoritmit

Keskustellessa polunhakualgoritmeista paras etenemissuunta lienee yleisimmistä tekniikoista ad-hoc ratkaisuihin. Optimaalin polun haku on intuitiivisinta painottomissa verkoissa, joissa niinkin helppo algoritmi kuten leveyssuuntainen haku (Algoritmi 2) on riittävä. Algoritmi käyttää kaksi tietorakennetta: saavutettujen solmujen FIFO-jonon (engl. *first-in, first-out*) ja hajautustaulun $\pi: V \rightarrow V$, joka assosioi jokaisen saavutetun solmun u kanssa se solmu v , josta haku on edennyt u :hun. Intuitio tämän takana on se, että leveyssuuntainen haun hakuvaruus etenee ”verkon taso” kerralla lähtien lähtösolmusta s , kunnes saavuttaa maalisolmun t , minkä jälkeen löydetty lyhin polku $\langle s = \pi(\pi(\dots)), \pi(\pi(t)), \pi(t), t \rangle$ voidaan rakentaa rutiinilla TRACEBACK-PATH (Algoritmi 1).

Mitä tulee painotettuihin verkkoihin Edsger W. Dijkstra esitti vuonna 1959 kuuluisan polunhakualgoritminsa, joka toimii polynomisessa ajassa [2]. Algoritmin voidaan katsoa yhdistävän ahneuden (engl. *greedy algorithm*), dynaamisen ohjelmoinnin ja inkrementaalisen lähestymistavan. Saatuaan lähtösolmun s , algoritmi laskee lyhimpien polkujen puun lähtien solmusta s kunnes t joutuu *avoimeen listaan* (engl. *open list; search frontier*), ja sitä kautta *suljettuun listaan* (engl. *closed list; settled node list*), jolloin lyhin s, t -polku on löytynyt. Hart et al. esittivät vuonna 1968 kuuluisan A*-algoritminsa [4], joka – samoin kuten Dijkstran algoritmi – ylläpitää mm. kunkin saavutetun

Algoritmi 1: TRACEBACK-PATH(x, π, π_{REV})

```
1  $u = x$ 
2  $p = \langle \rangle$ 
3 while  $u$  is not nil do
4    $p = \langle u \rangle \circ p$ 
5    $u = \pi(u)$ 
   Kaksisuuntainen haku?
6 if  $\pi_{REV}$  is not nil then
7    $u = \pi_{REV}(x)$ 
8   while  $u$  is not nil do
9      $p = p \circ \langle u \rangle$ 
10     $u = \pi_{REV}(u)$ 
11 return  $p$ 
```

Algoritmi 2: BREADTH-FIRST-SEARCH(G, s, t)

```
1  $Q = \langle s \rangle$ 
2  $\pi(s) = \text{nil}$ 
3 while  $|Q| > 0$  do
4    $u = \text{DEQUEUE}(Q)$ 
5   if  $u$  is  $t$  then
6     return TRACEBACK-PATH( $u, \pi, \text{nil}$ )
7   for  $(u, v) \in G.A$  do
8     if  $v$  is not yet mapped in  $\pi$  then
9        $\pi(v) = u$ 
10      ENQUEUE( $Q, v$ )
11 return  $\langle \rangle$ 
```

solmun u g -arvon $g(u)$, joka on toistaiseksi pienin kustannus lähtösolmusta s solmuun u , ja joka on taattu olemaan pienin mahdollinen heti kun u poistuu avoimesta listasta. Erona on kuitenkin se, että A^* käyttää kunkin solmun u prioriteettinä sen f -arvoa, joka on $f(u) = g(u) + h(u)$, missä $h(u)$ on solmun u optimistinen (eli aliarvioitu) etäisyys maalisolmuun. Intuitio tämän järjestelyn takana on se, että A^* "tietää" mihin suuntaan haku on suunnattava, jota päästäisiin maalisolmuun, ainakin paremmin kuin Dijkstran algoritmi, jonka hakuavaruus kasvaa laajenevan pallon tavoin "kaikkiin suuntiin". A^* :n pseudokoodi on tasan sama kuin Dijkstran algoritmin (Algoritmi 3). Erotuksena riveillä 14 ja 18 $g(x)$:n sijasta on $f(x)$, jolle siis $f(x) = g(x) + h(x)$. Molemmat kutsuvat TRACEBACK-PATH-rutiinin, joka muodostaa lyhimmän polun edeltäjäpuusta (engl. *predecessor tree*) ajassa

Algoritmi 3: DIJKSTRA-SHORTEST-PATH(G, s, t, w)

Monikkosijoitus

```
1 OPEN, CLOSED,  $g, \pi = (\{s\}, \emptyset, \{(s, 0)\}, \{(s, \mathbf{nil})\})$ 
2 while |OPEN| > 0 do
3    $u = \text{EXTRACT-MINIMUM}(\text{OPEN})$ 
4   if  $u$  is  $t$  then
5     return TRACEBACK-PATH( $u, \pi, \mathbf{nil}$ )
6   CLOSED = CLOSED  $\cup \{u\}$ 
   Jokaisella solmun  $u$  lapsisolmulla  $x$ , tee...
7   for  $(u, x) \in G.A$  do
8     if  $x \in \text{CLOSED}$  then
9       continue
10     $g' = g(u) + w(u, x)$ 
11    if  $x \notin \text{OPEN}$  then
12       $g(x) = g'$ 
13       $\pi(x) = u$ 
14      INSERT(OPEN,  $x, g(x)$ )
15    else if  $g(x) > g'$  then
16       $g(x) = g'$ 
17       $\pi(x) = u$ 
18      DECREASE-KEY(OPEN,  $x, g(x)$ )
   Ei  $s, t$  -polkua verkossa  $G$ .
19 return  $\langle \rangle$ 
```

$\Theta(N)$, missä N on lyhimmän polun solmujen määrä. On huomattava, että A^* palautuu Dijkstran algoritmiin määrittelemällä $h(u) = 0$ jokaisella $u \in V$, sillä tuolloin $f(x) = g(x)$.

2.1 Prioriteettijonon valinta

Polkua haettaessa painotetussa verkossa joudutaan käyttämään prioriteettijonoja, jotka ovat tarpeellisia pitääkseen haut optimaaleina, ja joiden oletetaan tarjoavan ainakin seuraavat operaatiot:

1. INSERT(H, x, k) tallettaa solmun x sen prioriteettiavaimen k kera,
2. DECREASE-KEY(H, x, k) päivittää solmun x talletetun prioriteettiavaimen (pienemmäksi),
3. EXTRACT-MINIMUM(H) poistaa pienimmän prioriteetin omaava solmu,
4. MIN(H) palauttaa, muttei poista keosta solmun, jolla on pienin prioriteettiavain, ja

5. $\text{SIZE}(H)$ palauttaa jonossa olevien alkioiden määrän.

Operaatiot (4) ja (5) voidaan aina toteuttaa siten, että ne toimivat vakioajassa. Helpoin prioriteettijonorakenne (jatkossa vain ”keko”), jonka operaatiot (1) - (3) käyvät logaritmisessa ajassa, on binäärikeko. Tällaisella keolla Dijkstran ja A*-algoritmit käyvät kumpikin ajassa $O((m+n)\log n)$. Teoriassa edelläoleva ylläraja voidaan parantaa käyttämällä Fibonacci-kekoa, jonka lisäysoperaatio (1) käy eksaktissa vakioajassa, päivitysoperaatio (2) tasoitettussa vakioajassa, ja poisto-operaatio (3) tasoitettussa ajassa $O(\log n)$, jolloin haut voidaan suorittaa ajassa $O(m+n\log n)$. Huomaa, että kaikki tähän asti mainitut keot perustuvat vertailuihin, ja teoriassa enintään yksi operaatiosta INSERT tai EXTRACT-MINIMUM voi käydä eksaktissa tai tasoitettussa vakioajassa, ja toisen on käytävä ajassa $\Omega(\log n)$, koska muuten algoritmi 4 tällaisella keolla rikkoisi vertailuihin perustuvan lajittelemisen informaatioteoreettisen rajan, joka on $\Omega(n\log n)$.

Jos kuitenkin kaarien painot ovat kokonaislukuja, $O(m+n\log n)$ -rajaa voidaan parantaa: Mikkel Thorup esitti vuonna 2003 keon, jonka poisto-operaatio käy ajassa $O(\log \log \min n)$ ja muut operaatiot vakioajassa [9]. Jos kuitenkin kokonaislukupainot ovat väliltä $[0, N]$, poisto-operaatio voidaan suorittaa ajassa $O(\log \log \min\{n, N\})$. Nyt selvästi haun aikavaativuus tällaisella keolla on $O(m+n\log \log \min\{n, N\})$.

3 Kaksisuuntainen haku

Vaikka A* on tyypillisesti tehokkaampi kuin Dijkstran algoritmi, käyttämällä *kaksisuuntaista* hakua, voidaan päästää verrattavissa olevaan suorituskykyyn. Ajatus kaksisuuntaisuuden takana on se, että algoritmi kasvattaa kaksi hakupuuta: yhden normaaliin tapaan ja toisen maalisolmusta ihan kuin kaaret olisi ”käännetty” päinvastaiseen suuntaan, kunnes kaksi hakuavaruutta ”kohtaavat” keskellä. Nyt jos lyhin polku koostuu N kaaresta, ja verkon solmujen keskiarvoinen aste on d , tavallinen, eli yksisuuntainen haku vaatii ajan

$$\sum_{i=0}^N d^i,$$

kun kaksisuuntainen vaatii

$$2 \sum_{i=0}^{\lceil N/2 \rceil} d^i.$$

Ylläoleva pätee leveyssuuntaiseen hakuun sellaisenaan, ja painotetun haun kohdalla voidaan saada ylläraja kertomalla kunkin summan termin tekijällä $O(\log n)$, joka liittyy algoritmien käyttämään prioritetettijonon operaatioihin.

3.1 Kaksisuuntainen Dijkstran algoritmi

Ylläolevan analyysin nojalla, on selvä, että Dijkstran algoritmi hyötyy kaksisuuntaisuudesta, eikä edellytä minkäänlaista verkon esiprosessointia. Lisäksi, algoritmin vahvuutena suhteessa A^* :iin ei ole pelkästään verrattavissa oleva tehokkuus, vaan myös heuristiikkafunktion tarpeettomuus. μ on toistaiseksi lyhimmän polun kustannus, joka suorituksen alussa on ∞ . Kun algoritmi löytää toistaiseksi lyhimmän polun hakuavaruuksien kohdatessa, ”välisolmu” m ja sen implikoima kustannus μ päivitetään. Haku jatkuu siihen asti, kunnes molempien avointen listojen minimialkioiden kustannusten summa on vähintään μ .

Molemmat hakusuunnat käyttävät alirutiinin EXPAND, mutta eri parametrein: listat OPEN ja CLOSED kuuluvat saman suunnan dataan (suunta d), CLOSED₂ on vastakkaisen hakusuunnan \hat{d} suljettu lista, g, π ovat suunnan d kustannuskuvaus ja vanhempainkuvaus vastaavasti, m on toistaiseksi lyhimmän polun ”välisolmu”, jossa kaksi hakuavaruutta kohtasivat, μ on m :n implikoiman polun kustannus, $*$ on laajentumisoperaattori, jolle $*(u)$ antaa solmun u seuraajasolmut, ja kuvaus $w: A \rightarrow \mathbb{R}_{\geq 0}$ antaa kaarien painot. Parametrit otaksutaan tässä rutiinissa viiteparametreiksi, jolloin esimerkiksi μ ja m ovat päivitettävissä.

Rutiini UPDATE tarkistaa, että yhden hakusuunnan solmu on toisen suljetussa listassa, ja jos asia on niin, yrittää päivittää välisolmun ja siihen liittyvän toistaiseksi pienimmän kustannuksen μ ja se vaatii parametreikseen hakusuunnan d solmun x , suunnan \hat{d} CLOSED-listan, solmun x g -arvo g_x suunnassa d , suunnan \hat{d} g -kuvauksen, välisolmun m ja sen implikoiman kustannuksen μ . Tämänkin apurutiinin kohdalla parametrit otaksutaan viiteparametreiksi, jolloin niiden päivittäminen rutiinissa on mahdollista.

Laajentumisoperaattoreina (engl. *expansion operator*) kaksisuuntaisessa haussa on määriteltävä funktiot $e, e_{REV}: V \rightarrow \mathcal{P}(V)$, joille $e(u) = \{v \in V: (u, v) \in A\}$ ja $e_{REV}(u) = \{v \in V: (v, u) \in A\}$. Nyt siis $e(u)$ antaa solmun u lapsisolmut, ja $e_{REV}(u)$ sen vanhempainsolmut, jolloin e käytetään normaalissa ja e_{REV} käännettyssä haussa. Rutiini 7 määrittelee kaksisuuntaisen Dijkstran algoritmin pysähtymisehdon, joka on tarpeeksi vahva pitämään polut optimaaleina ja laskenta-ajan kohtuullisena.

3.2 Kaksisuuntainen A^*

Kaksisuuntaisen A^* :n saa aikaan muuttamalla algoritmin 5 rivillä 1 esiintyvä $g(x)$ $f(x)$:ksi ja muuttamalla TERMINATE-rutiinin ehto seuraavanlaiseksi:

$$\max\left(\min_{x \in \text{OPEN}} f(x), \min_{x \in \text{OPEN}_{REV}} f_{REV}(x)\right) \geq \mu,$$

joka on siis Ira Pohlin vuonna 1971 ehdotetun BHPA-algoritminsa pysähtymisehto [7].

Algoritmi 4: GENERIC-HEAP-SORT(S, H)

Tyhjennä keko
1 $H = \emptyset$
2 **for** $i = 1$ **to** $|S|$ **do**
 $S[i]$ on itsensä prioriteetti.
3 INSERT($H, S[i], S[i]$)
4 **for** $i = 1$ **to** $|S|$ **do**
5 $S[i] = \text{EXTRACT-MINIMUM}(H)$

Algoritmi 5: EXPAND($\text{OPEN}, \text{CLOSED}, \text{CLOSED}_2, g, g_2, \pi, \mu, m, *, w$)

1 $u = \text{EXTRACT-MINIMUM}(\text{OPEN})$
2 $\text{CLOSED} = \text{CLOSED} \cup \{u\}$
3 **for** $x \in *(u)$ **do**
4 **if** $x \in \text{CLOSED}$ **then**
5 **continue**
6 $g' = g(u)$
7 **if** $*$ **is** e_{REV} **then**
8 Käännetty haku.
9 $g' = g' + w(x, u)$
10 **else**
11 Normaali haku.
12 $g' = g' + w(u, x)$
11 **if** $x \notin \text{OPEN}$ **then**
12 $g(x) = g'$
13 $\pi(x) = u$
14 INSERT($\text{OPEN}, x, g(x)$)
15 UPDATE($x, \text{CLOSED}_2, g(x), g_2, \mu, m$)
16 **else if** $g(x) > g'$ **then**
17 $g(x) = g'$
18 $\pi(x) = u$
19 DECREASE-KEY($\text{OPEN}, x, g(x)$)
20 UPDATE($x, \text{CLOSED}_2, g(x), g_2, \mu, m$)

Algoritmi 6: UPDATE($x, \text{CLOSED}, g_x, g, \mu, m$)

1 **if** $x \in \text{CLOSED}$ **then**
2 $p = g_x + g(x)$
3 **if** $\mu > p$ **then**
4 $\mu = p$
5 $m = x$

Algoritmi 7: TERMINATE($\text{OPEN}, \text{OPEN}_{REV}, g, g_{REV}, \pi, \pi_{REV}, \mu, m$)

```

1 if  $g(\text{MIN}(\text{OPEN})) + g_{REV}(\text{MIN}(\text{OPEN}_{REV})) \geq \mu$  then
2   return TRACEBACK-PATH( $m, \pi, \pi_{REV}$ )
3 return nil

```

Algoritmi 8: BIDIRECTIONAL-DIJKSTRA-SHORTEST-PATH(G, s, t, w)

```

1 OPEN, CLOSED,  $g, \pi = \{s\}, \emptyset, \{(s, 0)\}, \{(s, \mathbf{nil})\}$ 
2 OPENREV, CLOSEDREV,  $g_{REV}, \pi_{REV} = \{t\}, \emptyset, \{(t, 0)\}, \{(t, \mathbf{nil})\}$ 
3  $\mu = \infty$ 
4  $m = \mathbf{nil}$ 
5 while  $|\text{OPEN}| > 0$  and  $|\text{OPEN}_{REV}| > 0$  do
6   if  $m$  is not nil then
7      $p = \text{TERMINATE}(\text{OPEN}, \text{OPEN}_{REV},$ 
8        $g, g_{REV},$ 
9        $\pi, \pi_{REV},$ 
10       $\mu, m)$ 
11     if  $p$  is not nil then
12       return  $p$ 
13   if  $|\text{OPEN}| < |\text{OPEN}_{REV}|$  then
14     EXPAND( $\text{OPEN},$ 
15        $\text{CLOSED},$ 
16        $\text{CLOSED}_{REV},$ 
17        $g, g_{REV}, \pi, \mu, m, e, w)$ 
18   else
19     EXPAND( $\text{OPEN}_{REV},$ 
20        $\text{CLOSED}_{REV},$ 
21        $\text{CLOSED},$ 
22        $g_{REV}, g, \pi_{REV}, \mu, m, e_{REV}, w)$ 
23 return  $\langle \rangle$ 

```

Vuonna 2007 Taeg-Keun Whangbo ehdotti toisenlaisen kaksisuuntaisen heuristisen hakualgoritmin [10], joka olettaa, että kullakin solmulla on koordinaatit tasossa. h -arvon sijasta, määritellään

$$l(x) = \frac{(s-p) \cdot (x-p)}{|s-p|}, x \in \text{OPEN},$$

$$l_{REV}(x) = \frac{(t-p) \cdot (x-p)}{|t-p|}, x \in \text{OPEN}_{REV}.$$

Kun kaksi hakuavaruutta kohtaavat solmussa p , piirretään p :n kautta kulkeva suora Λ , joka on normaali sen suoran kanssa, joka kulkee solmujen s, t kautta. Nyt esimerkiksi jokaisella $x \in \text{OPEN}_{REV}$, $l_{REV}(x)$ on solmun x etäisyys Λ :sta, ja uusi pysähtymisehto on

$$L_{\min}^1 \leq \min_{x \in \text{OPEN}} (g(x) + l(x)),$$

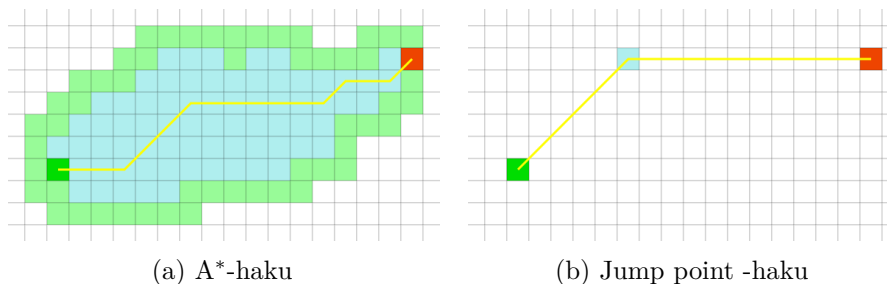
$$L_{\min}^2 \leq \min_{x \in \text{OPEN}_{REV}} (g_{REV}(x) + l_{REV}(x)),$$

missä L_{\min}^1 on lyhimmän s, p -polun kustannus normaalin haun puussa, ja L_{\min}^2 analogisesti lyhimmän p, t -polun kustannus vastakkainsuuntaisessa haussa. Whangbo raportoi tavallisen A^* :n vievän yhteensä 372 aikayksikköä laskettuna yhteen yli joukon hakuja. Samalla datalla BHPA vie 509 aikayksikköä, ja Whangbon variantti 209 aikayksikköä (huomaa BHPA:n vievän enemmän aikaa kuin tavallinen A^*).

4 Kaikkien parien lyhimvät polut

Toisinaan on annettu n solmua ja halutaan löytää lyhimvät polut kaikkien solmuparien välillä. Yksi tehokkaimmista algoritmeista on Floyd-Warshallin algoritmi, joka käy ajassa $\Theta(n^3)$, eikä sen toiminta riipu kaarien määrästä m . Ellei kohdeverkko ole täysi ($m = o(n^2)$), Johnsonin algoritmi saattaa olla parempi valinta, sillä Fibonacci-keolla edellämainittu käy ajassa $O(n^2 \log n + nm)$. Ajatus Johnsonin algoritmista on ensin tarkistaa, ettei verkossa ole sykliä, jonka kokonaispaino on negatiivinen, minkä jälkeen algoritmi ajaa n kertaa Dijkstran algoritmin jokaisesta solmusta, ja jokaisella kerralla hakee kokonaisen lyhimpien polkujen puun.

Yllä esitetyistä aikavaativuuksista ilmenee, etteivät asianomaiset algoritmit ole tarpeeksi tehokkaita jo n :n arvoilla ≥ 10000 . Jos kuitenkin ongelman koko sallii kaikkien parien algoritmin ajon, algoritmi palauttaa ”edeltäjä-matriisin” (engl. *predecessor matrix*), josta N :n solmun lyhin polku voidaan rakentaa ajassa $\Theta(N)$, mitä ei pysty parantamaan tuon enempää, ei ainakaan ilman edistynempää algoritmiikkaa. (Ja vaikka voisikin, polun tulostaminen ja/tai piirtäminen on jo vähintään $\Omega(N)$.)



Kuva 1: Polkusymmetria

5 Ruudukkoverkko ja jump point -haku

Ruudukkoverkko (engl. *grid graph*) on suuntamattoman verkon erikoistapaus, ja sen rakenne voidaan määritellä siten, että kullakin solmulla on kahden kokonaisluvun koordinaatti, eli solmujoukko on $\{(x, y): 1 \leq x \leq w, 1 \leq y \leq h\}$, missä w on ruudukkoverkon leveys ja h on sen korkeus. Nyt kun on annettu kaksi solmua (x_1, y_1) ja (x_2, y_2) , jos vain yksi koordinaateista eroavat yksikön verran, kyseessä on vaaka- tai pystysuuntainen kaari ja sen painoksi asetetaan 1. Toisaalta, kun molemmat koordinaatit eroavat yksikön verran, kyseessä on vino kaari, jonka painoksi asetetaan $\sqrt{2}$. On selvää, että jo leveyssuuntainen haku on optimaali tällaisella verkolla, sillä aina kun se etenee vinottain, esim. solmusta (x, y) solmuun $(x + 1, y - 1)$, se ohittaa kahden kaaren siirron solmun $(x + 1, y)$ tai $(x, y - 1)$ kautta, joka pidentäisi lyhimmän polun pituuden kahdella yksiköllä $\sqrt{2}$ sijaan. Heuristinen haku on kannattavaa tällaisella verkkotyypillä, sillä heuristiset funktiot kuten euklidinen (tai Manhattan-metriikka, ellei vinoja kaareja sallita) on helppoa toteuttaa ilman esiprosessointia. Toisaalta, juuri tällaisella verkkotyypillä A* kärsii polkujen ”symmetriasta”, kuten kuvasta 1a ilmenee. Jos siis solmujen s ja t molemmat koordinaatit eroavat, A* kasvattaa suunnikkaanmuotoisen hakuavaruuden solmusta s solmuun t , sillä niiden välissä on monta saman kustannuksen omaavaa polkua. Kuten yleensä erikoistapauksiin rajoittuessa, polunhaku ruudukolla voidaan tehdä paljon tehokkaammin. Vuonna 2011 Harabor ja Grastien esittivät ”jump point search” -nimisen A*:n muunnelman, joka karsii pois polkusymmetriat ja etenee ”hyppien” yli monen solmun siinä, missä muut algoritmit etenevät jokaisen välisolmun kautta [3].

6 Dijkstran algoritmi kaarivivuilla

Vuonna 2007 Möhring et al. esittivät mielenkiintoisen tavan nopeuttaa Dijkstran algoritmi [6]. Ajatuksena on osittaa suunnatun verkon $G = (V, A)$

solmujoukko osioihin V_1, \dots, V_p siten, että

$$\bigcup_{i=1}^p V_i = V,$$

ja $V_i \cap V_j = \emptyset$ jokaisella $i \neq j$. Ositus voidaan toteuttaa kuvauksella $r: V \rightarrow \{1, \dots, p\}$. Kustakin osiosta V_i puhuttaessa, sen ”rajasolmut” (engl. *boundary nodes*) määritellään joukkona

$$B_i = \{v \in V_i : \exists(u, v) \in A \text{ siten että } r(v) \neq r(u)\}.$$

Lisäksi, järjestelmä liittää jokaiseen verkon kaareen ”kaarivipuvektorin” (engl. *arc-flag vector*), jota voidaan toteuttaa p :n bitin bittivektorina. Nyt jokaisella osiolla V_i järjestelmän esiprosessointialgoritmi ajaa ”takaperin” tavallisen Dijkstran algoritmi kustakin rajasolmusta $b \in B_i$, ja asettaa tuloksena syntyvässä lyhimpien polkujen puussa jokaisen kaaren a kohdalla a :n vipuvektorin $r(b)$:n bitin päälle. Tuloksena syntyvässä järjestelmässä, haettaessa polkua solmuun t , nopeutettu Dijkstran algoritmi voi karsia kaikki ne kaaret, joiden vektorin $r(t)$:s bitti ei ole päällä, ainakin niin kauan kunnes haku pääsee samaan osioon solmun t kanssa. Tekniikka voidaan siis nähdä tasopainoilevan tavallisen Dijkstran algoritmin ($V_1 = V$) ja kaikkien parien algoritmin (kukin solmu on osio) välillä.

6.1 Ositustekniikat

Kuten ylläolevasta kävi ilmi, ”kaarivipu”-Dijkstra vaatii verkon solmujen osituksen, minkä jälkeen joudutaan esiprosessoimaan koko verkko. Koska esiprosessoinnin aika riippuu lineaarisesti kaikkien osioiden kaikkien rajasolmujen yhteenlasketusta määrästä, jälkimmäisen minimointi on toivottavaa.

Mikäli on annettu kunkin solmun koordinaatit tasossa, helpoin tapa osioida verkko (”ruudukointi”) on jakaa pienin, kaikki solmut sisältävä suorakulmio w sarakkeeseen ja h riviin. Tämä ei kuitenkaan ole vaillo ongelmia: esimerkiksi viidenkymmenen neliökilometrin osio pääkaupunkiseudulla sisältäisi paljon enemmän infrastruktuuria kuin jokin samankokoinen alue Kainuun maakunnassa. Asia voidaan parantaa käyttämällä ”nelipuita” (engl. *quad-tree*): koko suorakulmio jaetaan neljään, samankokoiseen suorakulmioon, minkä jälkeen jaetaan jälkimmäiset, ja niin edelleen pysäyttäen jaon niiden suorakulmioiden kohdalla, joissa on enintään κ solmua (κ annetaan nelipuu-algoritmilta parametrina). Tämä ottaa solmujakauman jo paremmin kuin ruudukointi, mutta ei niin hyvin kuin kd -puu (engl. *kd-tree*), joka lajittelee kaikkien solmujen listan ensin esimerkiksi x -koordinaattien perusteella, poimii mediaanialueen x -koordinaatin x_{mid} , ja implisiittisesti jakaa koko listan kahteen osalistaan V_{\leq} ja $V_{>}$, missä $V_* = \{x \in V : x * x_{mid}\}$, minkä jälkeen lajitellaan V_{\leq} ja $V_{>}$, mutta jo y -koordinaattien perusteella, ja jako pysähtyy

niiden solmujoukkojen kohdalla, joissa on enintään κ solmua (tässäkin κ on kd -puulle annettu parametri).

Neljäs tapa, jota Möhring et al. ovat tarkastelleet, on vuonna 1998 kehitetty METIS [5], joka ei edes tarvitse solmukoordinaatteja. Järjestelmä toimii siten, että syöteverkosta G_i muodostetaan verkko G_{i+1} , joiden suhde on sellainen, että verkossa G_i yhdistetään ”tiheästi” yhdistetyt solmujoukot yhdeksi solmuksi verkossa G_{i+1} . Alunperin syötetään verkko $G_0 = G$, jolloin saadaan G_1 , minkä jälkeen syötetään samaan algoritmiin G_1 ja saadaan G_2 ja niin jatketaan kunnes saadaan tarpeeksi pieni verkko G_m . Kun G_m on osioitu, ”laajennetaan takaperin” verkot G_{m-1}, G_{m-2}, \dots kunnes päästään takaisin alkuperäiseen verkkoon $G_0 = G$, joka on osioitu.

6.2 Tulokset

Koska kaksisuuntainen haku on mahdollista myös kaarivipujärjestelmässä, asia vaatii vain sen, että kuhunkin kaaren liitetään kaksi vipuvektoria, yksi kutakin hakusuuntaa varten. Tällaisella algoritmilla Möhring et al. raportoivat nopeutuksen suhteessa tavalliseen Dijkstran algoritmiin olleen yli 500 noin yhden miljoonan solmun ja 2.5×10^6 kaaren verkolla.

7 Polunhaku ja Multiple Sequence Alignment -ongelma

Multiple sequence alignment -ongelmassa on annettu κ sekvenssiä yli aakkoston Σ , kustannusmatriisi $c: \Sigma^2 \rightarrow \mathbb{Z}$ ja ”välisakko” (engl. *gap penalty*), joka liittyy merkkiin -. Aakkostona on tyypillisesti 20 aminohappoa, ja tarkoituksena on laittaa eri sekvensseissa välimerkit ”-” siten, että jokainen (mahdollisesti) pidennetty sekvenssi sisältää saman määrän merkkejä, ja koko ”linjaus” (engl. *alignment*) omaa optimaalin kustannuksen. Jos M on $\kappa \times L$ -matriisi, jossa kukin rivi on tietty sekvenssi välimerkkeineen, kunkin sarakkeen s kustannus määritellään olevan

$$\mathfrak{C}(s) = \sum_{1 \leq i < j \leq \kappa} c(M_{i,s}, M_{j,s}),$$

ja koko linjauksen kustannus on

$$\sum_{i=1}^L \mathfrak{C}(i).$$

Lisäksi, määritellään $c(-, -) = 0$, ja aina kun jompikumpi merkki $a \in \Sigma$ ja toinen on väli -, käytetään merkkiparin kustannuksena em. välisakko. (Näin siis optimaalissa linjauksessa ei ole sarakkeita, joiden jokainen merkki on -.)

7.1 Ongelman muotoileminen verkkona

Kun on annettu sekvenssit S_1, \dots, S_κ , kukin niistä asetetaan κ -ulotteisen ”hilan” (engl. *lattice*) akseleiksi. Kukin solmu voidaan ajattella omistavan koordinaatit (x_1, \dots, x_κ) , ja kun siitä siirrytään seuraavaan solmuun (y_1, \dots, y_κ) , kukin y_i on joko x_i tai $x_i + 1$, jolloin niiden koordinaattien y_i kohdalla, jotka eroavat koordinaatista x_i , luetaan i :nnen sekvenssin y_i :des merkki, ja muiden kohdalla ei lueta mitään, vaan laitetaan samaan sarakkeeseen välimerkki -. Selvästi, jokaisella hilaan ”sisäsolmulla” on tasan $2^\kappa - 1$ lapsisolmuja, ja vaikka hilaverkko on syklieton, jo muutaman sekvenssin hilaa ei voida säilyttää muistissa eksplisiittisesti, vaan solmut joudutaan generoimaan vanhempainsolmuja laajennettaessa (sekvenssien pituus on sadan merkin suuruusluokkaa). Nyt ongelma palautuu reitinhakuongelmaksi, jossa on löydettävä lyhin polku hilaan solmusta $(0, \dots, 0)$ solmuun $(|S_1|, \dots, |S_\kappa|)$.

7.2 Ratkaisu

Yoshizumi et al. raportoivat A^* :n pystyvän käsittelemään enintään seitsemän sekvenssia, sillä sen avoimet ja suljetut listat kasvavat liian isoiksi. He ehdottavat PEA^* nimistä algoritmia (engl. *Partial Expansion A^**), jonka voidaan ajatella uhraavan hieman aikaa pitääkseen listat pienempinä, jolloin algoritmi pystyy käsittelemään 8 sekvenssia [11]. Algoritmi lajittelee avoimen listan solmut ei f -, vaan F -arvojen perusteella, missä $F(u)$ on solmun u ”ei-lupaavien” solmujen pienin f -arvo. Myös on annettu ei-negatiivinen katkaisuarvo C ; (engl. *cutoff value*). Lapsisolmu pidetään ”lupaavana”, jos sen f -arvo ei ole suurempi kuin vanhempainsolmunsa F -arvon ja C :n summa. Alunperin kunkin solmun F -arvo on sen f -arvo. Lisäksi, jos solmulla u on ei-lupaavat lapsisolmut, u laitetaan takaisin avoimeen listaan.

Kaikki kaikkiaan, C :n arvolla 50, PEA^* vähentää muistintarpeen 87%, vaikka käyttää vain 20% enemmän aikaa kuin A^* hilassa, jossa solmun ja sen lapsen f -arvot eroavat enintään 396 yksikköä. Lisäksi, PEA^* :n suhde A^* :iin on se, että edellinen palautuu jälkimmäiseksi, kun $C = \infty$.

8 Lyhimmät polut ja rinnakkaisuus

Toistaiseksi lyhimpien polkujen haku ei ole juuri antanut paljon aihetta rinnakkaistamiseen. Vuonna 1998 Meyer ja Sanders esittivät Δ -stepping-nimisen algoritminsa, joka asettaa kunkin saavutetun solmun u omaan ”koriin” (engl. *bucket*) numero i aina, kun $g(u) \in [(i-1)\Delta, i\Delta)$, jolloin kukin säie käsittelee vain osan kaikista koreista. 2^{19} solmun verkolla, jonka keskiarvoinen asteen 3, Meyer ja Sanders raportoivat peräkkäisen (engl. *sequential*) version olleen 3.1 kertaa nopeampi kuin ”optimoitu” Dijkstran algoritmi, ja 16 suorittimen hajautetussa järjestelmässä nopeutus 9.2 on mitattu suhteessa peräkkäiseen Δ -stepping-algoritmiin. On huomattava,

että algoritminsa toiminta riippuu Δ :n arvosta, ja Meyer et al. ehdottavat arvon $\Delta = 4/d$, missä d on keskiarvoinen solmun aste.

Rinnakkaistamiseen liittyvien käytännön ongelmista huolimatta, myös kaksisuuntaisen A^* :n variantti nimeltään NBA* sai rinnakkaisen version: PNBA* käyttää kaksi säiettä, kukin omaa hakusuuntaa varten, ja sisältää suhteellisen vähän synkronoinnin tarvetta [8]. Esimerkiksi, 15-palapelillä (engl. *15-puzzle*), NBA* löysi lyhimmän 58 siirron polun noin 2.5 kertaa nopeammin kuin A^* , ja PNBA* oli noin tasan kaksi kertaa nopeampi kuin edellinen.

On huomattava, että rinnakkaistaessa algoritmeja, ei ole mahdollista saada mielivaltaisen suuria nopeutuksia jo Amdahlin lain nojalla, jonka mukaan maksimaalinen nopeutus on

$$\frac{1}{(1 - P) + \frac{P}{N}},$$

missä N on suorittimien määrä ja $P \in (0, 1]$ on sen laskennan suhteellinen osuus, jota voidaan tehdä rinnakkain, eikä P ole koskaan 0, sillä jokaisessa rinnakkaisessa laskennassa joudutaan luomaan säikeet, mikä on ainakin osittain peräkkäinen operaatio. Ottaen raja-arvon N :n kasvaessa rajatta, saadaan maksimaalinen nopeutus $1/(1 - P)$.

Lähteet

- [1] Cormen, Thomas H., Leiserson, Charles E., Rivest, Ronald L. ja Stein, Clifford: *Introduction to Algorithms*, luku 26, sivu 727. The MIT Press, 3 painos, 2009, ISBN 0262033844, 9780262033848.
- [2] Dijkstra, Edsger W.: *A note on two problems in connexion with graphs*. Numerische Mathematik, 1:269–271, 1959.
- [3] Harabor, Daniel ja Grastien, Alban: *Online Graph Pruning for Pathfinding on Grid Maps*. Teoksessa *25th National Conference on Artificial Intelligence. AAAI*, 2011.
- [4] Hart, Peter E., Nilsson, Nils J. ja Raphael, Bertram: *A formal basis for the heuristic determination of minimum cost paths*. IEEE Transactions on Systems, Science, and Cybernetics, SSC-4(2):100–107, 1968.
- [5] Karypis, George ja Kumar, Vipin: *A Fast and High Quality Multilevel Scheme for Partitioning Irregular Graphs*. SIAM J. Sci. Comput., 20(1):359–392, joulukuu 1998, ISSN 1064-8275. <http://dx.doi.org/10.1137/S1064827595287997>.
- [6] Möhring, Rolf H., Schilling, Heiko, Schütz, Birk, Wagner, Dorothea ja Willhalm, Thomas: *Partitioning Graphs to Speedup Dijkstra's Algorithm*.

- J. Exp. Algorithmics, 11, helmikuu 2007, ISSN 1084-6654. <http://doi.acm.org/10.1145/1187436.1216585>.
- [7] Pohl, Ira: *Bi-directional search*. Machine Intelligence 6, sivut 127–140, 1971.
 - [8] Rios, Luis Henrique Oliveiral ja Chaimowicz, Luiz: *PNBA*: A Parallel Bidirectional Heuristic Search Algorithm*. 2011.
 - [9] Thorup, Mikkel: *Integer Priority Queues with Decrease Key in Constant Time and the Single Source Shortest Paths Problem*. Teoksessa *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*, STOC '03, sivut 149–158, New York, NY, USA, 2003. ACM, ISBN 1-58113-674-9. <http://doi.acm.org/10.1145/780542.780566>.
 - [10] Whangbo, Taeg Keun: *Efficient Modified Bidirectional A* Algorithm for Optimal Route-Finding*. Teoksessa *New Trends in Applied Artificial Intelligence, 20th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2007, Kyoto, Japan, June 26-29, 2007, Proceedings*, sivut 344–353, 2007. http://dx.doi.org/10.1007/978-3-540-73325-6_34.
 - [11] Yoshizumi, Takayuki, Miura, Teruhisa ja Ishida, Toru: *A* with Partial Expansion for Large Branching Factor Problems*. Teoksessa *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence*, sivut 923–929. AAAI Press, 2000, ISBN 0-262-51112-6. <http://dl.acm.org/citation.cfm?id=647288.721436>.