

## A CSOMAGTOVÁBBÍTÁS SKÁLÁZHATÓSÁGA: KORLÁTOK ÉS OPTIMUMOK

KÖRÖSI ATTILA, RÉTVÁRI GÁBOR

A számítógépes hálózatok egyik legalapvetőbb feladata az információ-továbbítás, ennek egy módja, amikor az útvonal mentén a köztes csomópontok döntenek el lépésről lépésre, hogy merre továbbítsák a csomagot. Éppen ezért egy elemi probléma, hogy ennek a megvalósításához mennyi információt kell eltárolni a hálózatban. Számos eredmény született már a témában, azonban azok jellemzően azt vizsgálták, hogy legrosszabb esetben minimálisan hány bitet kell eltárolni. Ebben a cikkben azonban egy új megközelítés révén ehelyett azt vizsgáljuk, hogy különböző konkrét hálózatokban mennyi információ szükséges. Mind analitikus, mind szimulációs vizsgálataink során azt tapasztaltuk, hogy a gyakorlatban elterjedt hálózatokban jellemzően lényegesen alacsonyabb a szükséges információmennyiség, mint a legrosszabb esetben.

### 1. Bevezetés

Napjainkban a számítógépes hálózat fogalma már nem jelent egyet az internettel, számos más formában is találkozhatunk velük, ilyenek például adatközpontok is. Bármilyen formájával találkozunk, mindenhol első számú feladata az információ továbbítása. Ennek két legegyszerűbb módja a forrásalapú csomagtovábbítás, mikor a csomag feladója előre eldönti a csomag útvonalát, azt belekódolja a csomagba, és a hálózatban a többi pont az útvonal mentén csak egyszerűen kiolvassa, hogy merre kell továbbítani. Illetve a köztes pontok alapján történő csomagtovábbítás, mely esetben a feladó a csomagba csupán a forrás és a cél csomópont azonosítóját írja a csomagra, és ezen információ alapján a köztes pontok maguk döntenek el, merre továbbítják a csomagot.

Mindkét módszernek megvan az előnye és a hátránya. A forrásalapú továbbításnál a feladónak ismernie kell a teljes hálózati topológiát, a köztes pontoknak azonban nem szükséges semmilyen számítást végezniük, mindent a feladó intéz. Ezzel szemben a köztes pont alapú megoldásnál a feladat terhe és az információtárolás is megoszlik a hálózatban, viszont figyelni kell arra, nehogy végtelen ciklusba kerüljön egy csomag.

Az internet az utóbbi mechanizmussal működik, és ebben a munkában a továbbiakban mi is csak ezzel fogunk foglalkozni. Attól függően, hogy milyen metrika szerint optimalizáljuk az útvonalainkat, a köztes pontoknak a soron következő pont megválasztásához esetleg elég csak a célállomást figyelembe venni, vagy a forrással is kalkulálnia kell. Például ha a legrövidebb utat keressük, akkor a forrástól teljesen független a további útvonal, azonban ha a legszélesebb útvonalat keressük, akkor az útvonal múltja is szempont lehet, konkrétan ha a legrövidebb legszélesebb útvonalon szeretnénk haladni, akkor számít a forrás is.

A továbbiakban az egyszerűség kedvéért legrövidebb útvonalakat fogunk keresni, így elég lesz a döntést a forrástól függetlenül, a cél csomópont függvényében meghozni. Ez a modellünkben azt jelenti, hogy minden csomópontnak le kell tárolnia egy  $n$  hosszú tömböt, ahol az  $i$ -edik elem az  $i$ -edik csomópont irányába adja a következő lépést. Ez a tömb felfogható egy  $n$  karakterből álló szövegnek is, ahol a karakterek a lehetséges következő lépések, szomszédok. Ezt a szöveget kell tömörítenünk, ha minimalizálni szeretnénk a csomagtovábbításhoz szükségesen letárolt információt.

Természetesen, már sokan vizsgáltak hasonló kérdéseket, a „compact routing” terület kimondottan azzal foglalkozik, hogy alsó-felső korlátokat adjon a csomagtovábbításhoz szükséges információ mennyiségéről. [4]-ben megmutatták, hogy adott  $d$ -re mindig létezik olyan  $n$  pontú, legfeljebb  $d$  fokú gráf, hogy  $\Omega(n^2 \log d)$  információt kell a gráfban tárolni. Ezt természetesen nem kívánjuk megcáfolni, de úgy gondoljuk, hogy a gyakorlatban előforduló gráfok esetében nem ilyen rossz a helyzet.

A következő fejezetben pontosan bemutatjuk a használt modellt. A 3. fejezetben ismertetünk pár elméleti eredményt, míg a 4. fejezetben szimulációs eredményekről és a további lehetséges kutatási irányokról ejtünk pár szót.

## 2. Modell

A továbbiakban a következő modellben fogunk dolgozni. Adott egy  $n$  pontú  $G$  gráf, amiben bármely két pont között definiáljuk az optimális útvonalat. Ezekről az utakról feltesszük, hogy minden  $v$  pont esetén a  $v$ -be menő optimális útvonalak uniója  $G$  egy feszítő fája, vagyis a köztes pontok számára az optimális utak csak a céltől függnek, ilyen például a legrövidebb utak rendszere. Ezen útvonalválasztó megvalósításához feltesszük, hogy minden  $u$  csomópontban van egy adatstruktúra, mely minden  $v$  csomópontra megadja, hogy  $u$ -nak melyik szomszédja felé kell továbbítani, amennyiben  $v$  a célpont. Feltehetjük, hogy a csomópontokat 1-től  $n$ -ig felszámozzuk, így a köztes pontokhoz tartozó struktúra lehet egy  $n$  hosszú tömb, mely  $i$ -dik eleme az  $i$  irányába menő út következő pontja.

Ezek után ez a tömb felfogható egy  $n$  hosszú szövegnek is, ahol a karakterek a szomszédokhoz tartozó azonosítók, mi pedig azt szeretnénk megállapítani, hogy

mennyi tárhelyre van szüksége ezeknek a szövegeknek. A szövegek tárolásigényének és tömörítésének irodalma elég bőséges, és nagy múltra tekinthet vissza. A fő eredmények viszonylag egyszerűek és széles körben ismertek.

Amennyiben a szövegre vonatkozó információk kimerül abban, hogy  $n$  hosszú és a karakterek  $\Sigma$  halmazának  $d$  az elemszáma, úgy értelemszerűen  $n \log d$  bitre van szükség. Amennyiben ismert az egyes karakterek eloszlása, úgy már tudunk entrópiára tömöríteni:  $H_0 = \sum_{c \in \Sigma} \frac{n_c}{n} \log \frac{n}{n_c}$ , ahol  $n_c$  darab  $c$  karakter van a szövegben, ez esetben  $nH_0 + o(n)$  bit szükséges [2]. Amennyiben struktúra is van a szövegben, úgy elképzelhető, hogy magasabb rendű entrópiára is tömöríthető:  $H_k = \sum_{q \in \Sigma^k} \sum_{c \in \Sigma} \frac{n_{qc}}{n} \log \frac{n}{n_{qc}}$ , ahol az  $n_{qc}$  azt mondja meg hányszor követi a  $q$ -t  $c$ . Itt is és a továbbiakban  $\log$  függvény alatt a kettes alapú logaritmust értjük.

### 3. Analízis

#### 3.1. Teljes gráf véletlen élsúlyokkal

Ebben a fejezetben azt a modellt vizsgáljuk meg, melyben a hálózat egy teljes gráf, azonban a pontok közötti élek hosszai független azonos exponenciális eloszlást követő valószínűségi változók. Mivel folytonos eloszlásúak, ezért azon elemi események halmaza, melyben van két pont, mik között létezik több legrövidebb út, null mértékű.

3.1. TÉTEL. *Az entrópia várható értéke  $\log e$ .*

*Bizonyítás.* A nulladrendű entrópia kiszámolásához, arra van szükségünk, hogy meg tudjuk mondani, melyik szomszédot hány célpont felé használja az adott pont. Ennek eloszlása megegyezik a pontból kiinduló legrövidebb utak fájában a gyökérről logó részfák méretének eloszlásával.

3.1. SEGÉDTÉTEL. *A legrövidebb utak fájának a gyökérből tekintett részfák méretének eloszlása megegyezik a kínai étterem asztalméret eloszlásával.*

*Bizonyítás.* A legrövidebb utak fáját megkaphatjuk Dijkstra algoritmusának segítségével, elindulunk az  $s$  pontunkból, ő lesz a fa gyökere, és megkeressük a belőle kiinduló legrövidebb élt, ez tartozzon a  $v_1$  ponthoz. A következő lépésben megkeressük azt a  $v_2$  pontot, aki a második legközelebb van  $s$ -hez.  $v_2$ -t vagy közvetlenül  $s$ -ből érjük el, vagy  $v_1$ -en keresztül.

Ennek a keresésnek egy szemléletes módja, ha az élekre véletlen exponenciális idejű csörgőórákat képzelünk el, melyek csörgési ideje az élek hosszát jelenti. Ennek megfelelően, első lépésben elindítjuk az órákat az  $s$ -ből kiinduló éleken. Az elsőnek megcsörrenő él végpontja lesz  $v_1$ , ebben a pillanatban elindítjuk az órákat a  $v_1$ -ből kiinduló éleken is, miközben az  $s$ -ből kiindulókon még mindig futnak az órák. Ekkor nyilván a következőnek megszólaló él végpontja lesz  $v_2$ , és attól függően,

hogy  $s$ -ből vagy  $v_1$ -ből indult az él, oda kötjük be a fában. Ezt az eljárást folytatva megkapható a legrövidebb utak fája.

Az exponenciális eloszlás örökifjú tulajdonságát felhasználva megtehetem, hogy amikor az első él megcsörren, és elindítom az órákat a  $v_1$ -ből kiinduló éleken, akkor az  $s$ -ből kiinduló többi élen újraindítom az órákat. Ezzel az eloszlásokon nem változtattam, azzal a megkötéssel, hogy nyilván hozzá kell adnom az élhosszakhoz az újraindítás időpontját is. Ez azt is jelenti, hogy amennyiben minket csak az érdekel, hogy melyik pont hova kötődik be a fában, úgy azt tapasztaljuk, hogy ugyanannyi eséllyel fognak megszólalni az  $s$ -ből kiinduló élekre rakott órák, mint a  $v_1$ -ből kiinduló élekre rakottak.

Tehát az új pontok a fa régebbi pontjaihoz egyenletes eloszlással kötődnek be, a részfák a méretükkel arányos valószínűséggel bővülnek. Új részfa akkor keletkezik, ha az új pont  $s$ -hez kötődik, ami  $v_k$  esetén  $1/k$  valószínűséggel következik be. Ez pedig pontosan a kínai étterem növekedési modellje.  $\square$

Az entrópia eloszlásának kiszámításához szükség lenne a részfák méretének eloszlására is, mi azonban megelégszünk a várható entrópiával, ahhoz pedig elég a különböző méretű részfák várható darabszáma.

### 3.2. SEGÉDTÉTEL. Várhatóan $1/k$ darab $k$ méretű részfa lesz.

Ebből könnyen számolható a várható entrópia egy pontban:  $H_0 = \sum \frac{n_i}{n} \log \frac{n}{n_i}$ , ahol  $n_i$  az  $i$ -edik szomszédhoz tartozó részfa mérete. Ez átírható a következő alakba:  $H_0 = \sum X(k) \frac{k}{n} \log \frac{n}{k}$ , ahol  $X(k)$  a  $k$  méretű részfák száma.  $X(k)$ -ről viszont már tudjuk, hogy értéke várhatóan  $1/k$ , ez alapján  $H_0$  várható értéke is számolható:

$$\begin{aligned} \mathbf{E}H_0 &= \sum \mathbf{E}X(k) \frac{k}{n} \log \frac{n}{k} = \sum \frac{1}{k} \frac{k}{n} \log \frac{n}{k} = \frac{1}{n} \sum \log \frac{n}{k} = \frac{1}{n} \log \frac{n^n}{n!} \\ &\sim \log n - \log n + \log e = \log e, \end{aligned}$$

hiszen  $\log n! \sim n \log n - n \log e$ .  $\square$

Ebből az is kiszámolható, hogy egy pont várhatóan  $\ln n$  pont felé fog továbbítani, ami várhatóan  $n \log \ln n$  bitet jelentene csomópontonként, ha nem vesszük figyelembe az eloszlást. Az entrópia segítségével viszont várhatóan  $n \log e$  környékére tömöríthetők a csomópontokhoz tartozó útvonalválasztó tömbök.

### 3.2. Szabályos gráfstruktúrák egységnyi élsúlyokkal

Az előző példával ellentétben ebben a részben olyan gráfokról esik szó, melyekben az élek egységnyi hosszúak, ebből következően két pont között jellemzően több legrövidebb út is van, ami azt is jelenti, hogy nem egyértelmű a következő pont számtalan célpontra vonatkozólag. Ha véletlenszerűen választanánk ezek közül,

akkor jellemzően egyenletes lenne a következő pontok eloszlása is, illetve a magasabb rendű entrópiával se mennénk sokra. Ezért a továbbiakban erre figyelni kell, hogy ügyesen válasszuk meg a következő pontokat.

Az 1. táblázat tartalmazza az ide vonatkozó eredményeinket. Fákat,  $d$  dimenziós hiperkockákat, négyzetrácsokat és tóruszokat vizsgáltunk. Az  $I$  oszlop azt mutatja, mennyi a szükséges várható információ, ha véletlenszerűen választunk a legrövidebb utak közül. A  $H_0$  oszlop a nulladrendű entrópiát jelöli, ahol az értékeket úgy kaptuk meg, hogy a lehetséges következő pontok között felállítunk egy prioritást, és ez által lesznek gyakoribb meg ritkább szomszédok a tömbben. Továbbá a  $H_1$  oszlop jelöli az elsőrendű entrópiát, itt figyelünk a csomópontok megfelelő számozására is. Fák esetében ez elérhető mélységi számozással, míg a többiek esetén ez elérhető a koordinátázásból adódó felsorolással. Végül az utolsó oszlopban az irodalomban ismert korlátok szerepelnek.

1. táblázat. Entrópia különböző speciális struktúrájú gráfokra

| Gráfosztály         | $I$   | $\bar{H}_0$        | $\bar{H}_1$              | Compact routing          |
|---------------------|---|--------------------|--------------------------|--------------------------|
| $n$ pontú fa        | $\frac{\log n}{n} \longleftrightarrow \frac{\log e}{2}$ | $\frac{\log e}{2}$ | $2 \frac{\log n}{n}$     | $\frac{\log n}{n}$ [3]   |
| $d$ dim. hiperkocka | $\log d$  | 2                  | $d \frac{\log n}{n}$     | $d \frac{\log n}{n}$ [5] |
| $d$ dim. rács       | $\log(2d)$  | $\frac{\log e}{2}$ | $d \frac{\log n}{n}$     | $d \frac{\log n}{n}$ [5] |
| $d$ dim. tórusz     | $\log(2d)$  | 1                  | $(d+1) \frac{\log n}{n}$ | $d \frac{\log n}{n}$ [5] |

#### 4. Szimulációs eredmények, távlati célok

Az elméleti eredmények azt sugallják, hogy már ezzel az egyszerű módszerrel is a minimálisához közeli eredmények érhetőek el. Sajnálatos módon a fentiek felül más gráfokra egyelőre nem sikerült analitikus korlátokat találnunk. Ezek a kezdeti eredmények elég ígéretesnek tünnek ahhoz, hogy szimulációs vizsgálat alá vonjunk a gyakorlatban előforduló gráfokat. Az elméleti példákban a szabályos struktúrákhoz természetesen adódott a csomópontok megfelelő felszámozása, azonban a gyakorlatban szükségünk volt valamilyen heurisztikára, amivel csökkenthettük a magasabb rendű entrópiákat. Többféle módszerrel is próbálkoztunk, jellemzően azt próbáltuk elérni, hogy a közeli pontok közeli sorszámot kapjanak, így a távolabbi pontok következő pont tömbjében azonos karakterekből álló részsorozatok jöhetnek létre, amiknek jellemzően kicsik a magasabb rendű entrópiáik.

Az 2. táblázatban különböző AS-gráfokra vonatkozó eredmények láthatók, az AS-gráfok az internet autonóm rendszerei (Autonomous System) [1]. Az eredményekből látható, hogy pusztán az eloszlás figyelembevételével már könnyen felére csökkenthető a szükséges tárhely, továbbá magasabb rendű entrópia figyelembevételével még az is jelentősen tovább csökkenthető.

2. táblázat. AS-gráfok címzési méretei

| Gráf    | $n$    | $m$     | $\bar{I}$ | $\bar{H}_0$ | $\bar{H}_1$ |
|---------|--------|---------|-----------|-------------|-------------|
| 3-core  | 12 905 | 235 102 | 3         | 1,6         | 1,02        |
| 5-core  | 5328   | 186 570 | 3,95      | 2,4         | 1,56        |
| 7-core  | 3729   | 169 594 | 4,41      | 2,51        | 1,66        |
| 9-core  | 3006   | 158 898 | 4,90      | 2,73        | 1,53        |
| 17-core | 1685   | 128 718 | 5,55      | 2,99        | 1,33        |

A továbbiakban szeretnénk folytatni az analitikus vizsgálatokat, több gráfosztályra is szeretnénk analitikus korlátokat adni az entrópiaértékekre, illetve szeretnénk találni egy jó felszámozást, ami általános gráfokra kellően csökkenti a magasabb rendű entrópiákat.

### Köszönetnyilvánítás

Jelen munka részben az OTKA FK123957 támogatással készült, továbbá a szerzők szeretnék megköszönni a BME TMIT tanszékének támogatását.

### Hivatkozások

- [1] *The CAIDA AS Relationships Dataset (2013-08-01)*, <http://www.caida.org/data/active/as-relationships>.
- [2] COVER, T. M. AND THOMAS, J. A.: *Elements of information theory*, Wiley-Interscience (1991). DOI: [10.1002/0471200611](https://doi.org/10.1002/0471200611)
- [3] FRAIGNIAUD, P. AND GAVOILLE, C.: *Routing in Trees*, in: *ICALP '01*, pp. 757-772 (2001). DOI: [10.1007/3-540-48224-5\\_62](https://doi.org/10.1007/3-540-48224-5_62)
- [4] GAVOILLE, C. AND PÉRENNÈS, S.: *Memory Requirement for Routing in Distributed Networks*, in: *ACM PODC*, pp. 125-133 (1996). DOI: [10.1145/248052.248075](https://doi.org/10.1145/248052.248075)
- [5] VAN LEEUWEN, J. AND TAN, R. B.: *Interval Routing*, *Comput. J.*, Vol. **30** No. **4**, pp. 298-307 (1987). DOI: [10.1093/comjnl/30.4.298](https://doi.org/10.1093/comjnl/30.4.298)



Kőrösi Attila 2007-ben szerezte meg diplomáját matematikából, majd doktorandusz lett a BME-TMIT tanszékén, és csatlakozott a HSN Laborhoz, ahol kamatoztatva gráf- és valószínűségszámítás-elméleti ismereteit, különböző kutatói projekteken vett részt. Kezdetben P2P-segédletű Video-on-Demand rendszerek hatékonyságán dolgozott. Később főleg hálózatokkal foglalkozott, azokon belül is főleg hálózatok kialakulásával, illetve útvonalkereséssel. Jelenleg az akadémiai világtól némileg eltávolodva, tőzsdei előrejelzésekkel foglalkozik.

KŐRÖSI ATILA

MTA-BME Információs Rendszerek Kutatócsoport  
korosi@tmit.bme.hu



Dr. Rétvári Gábor 1999-ben szerzett villamosmérnöki oklevelet a Budapesti Műszaki és Gazdaságtudományi Egyetem (BME) Villamosmérnöki és Informatikai Karán. PhD fokozatát 2007-ben kapta meg. Doktori disszertációjának témája a szolgáltatásminőség alapú útvonalválasztás és a hálózati folyamatok elméletének távközlési alkalmazásai. Jelenleg tudományos főmunkatárs és kutatási témavezető a BME Villamosmérnöki és Informatikai Kar Távközlési és Médiainformatikai Tanszékén. Kutatási területe a kommunikációs hálózatok és informatikai rendszerek matematikai mo-

dellezése, a felhő-alapú számítástechnika hálózati vonatkozásai és a programozható hálózati eszközök architektúrája, analízise és optimalizálása.

RÉTVÁRI GÁBOR

MTA-BME Információs Rendszerek Kutatócsoport  
retvari@tmit.bme.hu

ON THE SCALABILITY OF HOP-BY-HOP PACKET ROUTING:  
TIGHT BOUNDS AND OPTIMAL ADDRESS SPACES

ATTILA KŐRÖSI, GÁBOR RÉTVÁRI

Routing in large-scale computer networks today is built on hop-by-hop routing: packet headers specify the destination address and routers use internal forwarding tables to map addresses to next-hop ports. In this paper we take a new look at the scalability of this paradigm.

First, we give a model that translates problems related to routing scalability to the language of information-theory. Forwarding tables are modeled as sequential strings, admitting tight memory requirement characterizations using Shannon's entropy measures and standard data compression techniques.

Contrary to previous work, our analysis is not of worst-case nature, but gives verifiable and realizable memory requirement characterizations even when subjected to concrete topologies and routing policies.

Using pure information-theoretic arguments, we are able to reproduce most of the space bounds obtained in the compact routing literature using piecemeal addressing schemes and analysis.

Our evaluations suggest that in most practically important cases significant memory savings can be attained by forwarding table compression over optimal address spaces. As far as we are aware of, this is the first study to link computer network scalability to information theory.

*Keywords:* shortest path routing, routing table entropy.

*Mathematics Subject Classification* (2000): 68R10, 05C12.