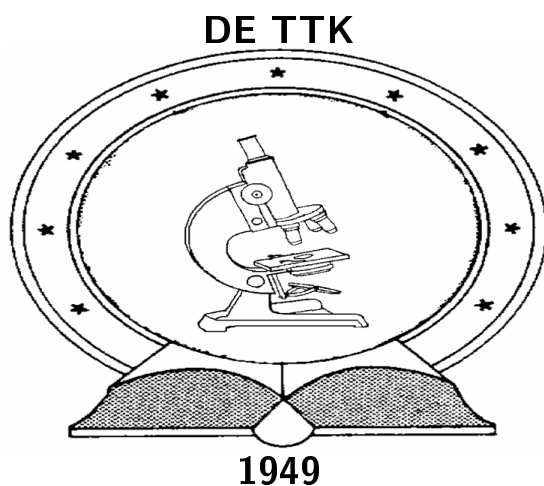




Új fizikai ismeretek bevezetése az oktatásba

Egyetemi doktori (PhD) értekezés

Horváth Árpád

Témavezető:

Dr. Trócsányi Zoltán
egyetemi tanár

DEBRECENI EGYETEM

Természettudományi Doktori Tanács

Fizikai Tudományok Doktori Iskolája

Debrecen, 2013

Ezen értekezést a Debreceni Egyetem Természettudományi Doktori Tanács Fizikai Tudományok Doktori Iskolája Részecskefizika programja keretében készítettem a Debreceni Egyetem természettudományi doktori (PhD) fokozatának elnyerése céljából.

Debrecen, 2013

Horváth Árpád

Tanúsítom, hogy Horváth Árpád doktorjelölt 2003-2013 között a fent megnevezett Doktori Iskola Részecskefizika programjának keretében irányításommal végezte munkáját. Az értekezésben foglalt eredményekhez a jelölt önálló alkotó tevékenységével meghatározóan hozzájárult. Az értekezés elfogadását javasolom.

Debrecen, 201

Dr. Trócsányi Zoltán
témavezető

Új fizikai ismeretek bevezetése az oktatásba

Értekezés a doktori (Ph.D.) fokozat megszerzése érdekében
a fizika tudományában.

Írta: Horváth Árpád okleveles matematika-fizika szakos tanár
Készült a Debreceni Egyetem Fizikai Tudományok Doktori Iskolája
Részecskefizika programja keretében

Témavezető:
Dr. Trócsányi Zoltán
egyetemi tanár

A doktori szigorlati bizottság:

elnök: Dr.
tagok: Dr.
Dr.

A doktori szigorlat időpontja: 201

Az értekezés bírálói:

Dr.
Dr.

A bírálóbizottság:

elnök: Dr.
tagok: Dr.
Dr.
Dr.
Dr.

Az értekezés védésének időpontja: 201

Tartalomjegyzék

1. Bevezetés	1
I. Részecskefizika az oktatásban	7
2. Mérések a CERN-sajátkezüleg honlappal	13
2.1. A méréshez kapcsolódó fontosabb fogalmak és tények	13
2.2. A detektorok felépítése, a DELPHI detektor	17
2.3. A CERN-sajátkezüleg honlap felépítése	25
2.4. A CERN-sajátkezüleg honlap fordítása és bővítése	25
2.5. A részecskecsaládok számának kiszámítása	26
2.6. A CMS detektor	28
2.7. CMS-sel kapcsolatos mérések a diákműhelyen	31
3. Részecskefizikai diákműhelyek	33
3.1. A részecskefizikai diákműhelyek története	33
3.2. A diákműhely kérdőíveinek kiértékelése	35
4. Oktatási és ismeretterjesztő segédanyagok	39
4.1. Közreműködésem a Wikipédiában	39
II. Összetett hálózatok	47
5. Bevezetés és alapfogalmak	49

6. Összetett hálózatok vizsgálata	65
6.1. A cxnet hálózatvizsgáló programcsomag	65
6.1.1. A fokszámeloszások kirajzolása	67
6.1.2. A networkx és igrph programcsomagok összehasonlítása .	70
6.1.3. A szoftvercsomag-hálózat létrehozása	75
6.2. A szoftvercsomag-hálózat tulajdonságai	79
6.2.1. Saját eredményeim	81
7. Adott tulajdonságú hálózatok létrehozása	93
7.1. A multifraktál-hálózatgeneráló módszer áttekintése	94
7.2. A valószínűségi mérték iterálása	94
7.3. A generáló mérték optimalizálása	99
7.4. A vizsgált céltulajdonságok	101
8. Az összetett hálózatok oktatása	105
8.1. A félév beosztása és követelményrendszere	105
8.2. Segédanyagok	108
8.3. Összetett hálózatok az informatikaversenyen	108
9. Összefoglalás és kitekintés	111
10. Summary and outlook	117

1. fejezet

Bevezetés

Az igazi tanár nem tanít a szó régi értelmében, hanem alkot. A saját mindig friss, mindig korszerű tudását alkotja, szüli újjá tanítványai lelkében. Így lesz művész, alkotó művész.

Öveges József

Doktori tanulmányaim során a fizika két olyan területének tanításával foglalkoztam, amelyekben az utóbbi két évtizedben jelentős új eredmények születtek. Az egyik a részecskefizika világa, amely az anyag tovább nem bontható alkotórészeinek keresésével és azok kölcsönhatásainak a vizsgálatával foglalkozik. A másik a hálózatok világa, amely olyan összetett rendszerek felépítésének és folyamatainak tanulmányozásáról szól, melyek sok hasonló alkotóelemből állnak, mint amilyen az emberekből felépülő szociális hálózat, vagy a weboldalakból felépülő Világháló.

Ahhoz, hogy megértsük, miért tartom érdemesnek ezeknek a területeknek a felhasználását a közoktatásban, szükséges a két témakör eredményei és a bennük felmerülő kérdések tárgyalása mellett az oktatásról is beszélni. Mivel matematika-fizika szakos tanárként végeztem és 14 éve a műszaki felsőoktatásban dolgozom, ezért a természettudományos tanárképzésre és a műszaki oktatásra van nagyobb rálátásom.

Sajnálatos fejlemény, hogy a természettudományos tanárképzésből az utóbbi évtizedben nagyon kevés tanár kerül ki, és helyezkedik el a szakmájában. Végzésem óta egy nagyságrenddel csökkent a fizikatanárként végzettek száma. Ráadásul a

velem együtt végzett fizikatanárok elég jelentős része sem tanári, hanem döntően informatikai területen dolgozik.

A ELTE Fizikai Intézet Professzori Tanácsa 2010-ben állást foglalt a természettudományos tanárképzés problémáiról, melyben kiemeli, hogy a 2009/2010-es évben 24 fő kezdte meg a természettudományos tanári képzést az egész országban, és fizikából mindössze négyen jelentkeztek a mesterképzés első évfolyamára. Eközben évente 6-700 természettudományt tanító tanár megy nyugdíjba, ami előbb utóbb tanárhiányt eredményez. A problémát részben a Bolognai-rendszer szerinti osztott képzésben, valamint a fő és mellékszakok óraszámainak alap- és mesterképzés közötti elosztásában látják [1]. A 2013/2014-es tanévtől kezdve a tanárképzés újra osztatlan lesz, és nagyobb arányú az államilag támogatott helyek száma is a tanár szakokon.

A Magyar Biológiai Tanárok Országos Egyesülete, a Kutató Tanárok Országos Szövetsége és a Kutató Diákok Országos Szövetsége közösen keresik a megoldásokat a természettudományok oktatásának problémáira. A felmérésükben megkérdezett diákok válaszaiból kiderül, hogy a diákok véleménye szerint a természettudományt tanító tanároknak rendkívül sok munkát kell végezniük szakmájuk társadalmi megbecsüléséhez képest. A felmérésük azt a régebb óta ismert tényt is megerősíti, hogy a természettudományos tantárgyak jóval kevésbé kedveltek a diákok között más tantárgyakhoz képest [2].

Nyilvánvalóan az oktatáspolitikában zajló változások lehetnek kedvező hatások, de ahhoz, hogy a diákok a fizikatanári vagy fizikusi pályát válasszák, szükség van arra is, hogy a diák úgy érezze, hogy a fizika érdekes tantárgy, ahol izgalmas kérdések várnak megválaszolásra. Úgy gondolom, hogy a részecskefizika kiváló lehetőséget nyújt a diákok érdeklődésének felkeltésére. A CERN 27 km kerületű Nagy hadronütköztető gyűrűje, amely Svájc és Franciaország alatt működik, jelenleg a legnagyobb részecskefizikai gyorsító mind a kísérlekben résztvevő kutatók számát, mind a luminozitás mértékét, mind a gyűjtött adatmennyiséget tekintve. Komolyabb szervezéssel ugyan, de akár diákokat is el lehet vinni oda, hogy megtekinthessék a kutatások helyszínét. Az új felfedezések, a magas technikai színvonalon megépített, technikai és informatikai újdonságokat is tartalmazó kísérleti és adatfeldolgozó berendezések, valamint a részecskefizika és a kozmológia megoldatlan problémái általában érdekli a diákokat, és általában a nem szakembereket is. A Higgs-bozon felfedezése idején jelentősen megemelkedett a Wikipédia kapcsolódó szócikkeinek a látogatottsága. Érdemes ezt az érdeklődést fenntartani úgy, hogy azt az iskolában tanultakkal összekapcsoljuk, ezzel érthetőbbé és izgalmasabbá téve mind az CERN-ről szóló híreket, mind a középiskolai tananyagot.

A részecskefizika kutatói komoly matematikai háttérrel rendelkező modelleket dolgoztak ki, közöttük a részecskefizika standard modelljét valamint a szuperszimmetriát illetve a gravitációt is magukban foglaló elméleteket. A standard modell számos kísérleti eredményt nagy pontossággal leír, egyik hátránya, hogy számos paraméterrel rendelkezik, amelyek értékére nem ad magyarázatot. Ilyen paraméterek többek között a kölcsönhatások csatolásának értékei.

A részecskefizikai kutatásokat a sok kutatót egyesítő nagyon összetett detektorokat használó kísérletek uralják. Mikor a doktori képzésbe bekapcsolódtam, a standard modellben szereplő részecskéket a Higgs-bozon kivételével felfedezték, a CERN Nagy elektron–pozitron ütköztető gyűrűjének (LEP-nek) kísérletei megmutatták, hogy a részecskének három családja van, és a neutrínóoszilláció léte is igazolásra került, a CERN-ben a Nagy hadronütköztető gyűrű (LHC) összeszerelése folyt. Ebben az időszakban, 2005-ben, indította útjára a CERN által szervezett Európai (pár éve Nemzetközi) Részecskefizikai Ismeretterjesztő Csoport a részecskefizikai diákműhelyeket, amelyek során a részecskefizikával, a CERN-nel és a CERN detektoraiiban lezajlott események vizsgálatával ismertetik meg a középiskolás diákokat. Ezek a mérések eleinte a LEP, 2011 óta pedig az LHC detektorainak adataival történnek. Minden nap végén négy-öt helyszín kapcsolódik össze videokonferencián, hogy a diákok a mért adatokat összevegyék, a kapott eredményt értelmezzék, és a CERN-i fizikusoknak feltegyék az őket érdeklő kérdéseiket.

Célom volt, hogy bekapcsolódjak a részecskefizika diákműhelyekbe, ennek segítségével meg tudjam mutatni középiskolás diákok számára, hogyan működnek a részecskefizikai kutatások, és általában hogyan működik a tudományos kutatás. Tervem volt az is, hogy bővítssem azoknak a jól rendszerezett magyar nyelvű segédanyagoknak a körét, amelyek segítik a felkészülést a diákműhelyre, az ott elsajátított dolgokat kiegészítik, és általában lehetővé teszik az elmélyülést a részecskefizikában a középiskolás diákok számára.

Munkahelyem az Óbudai Egyetem székesfehérvári Alba Regia Egyetemi Központja, ahol a Videotonban felhalmozódott tudásnak köszönhetően, elsősorban számítástechnikával kapcsolatos képzést hoztak létre jogelődünkben, az egykori Kandó Kálmán Villamosmérnöki Főiskolán. Az általam oktatott tárgyak döntő részt műszaki és informatikai tárgyak, valamennyi matematika és elvétve fizika. Mivel tanulóink a tudásuk jelentős részét az általános és középiskolából hozzák, ezért a műszaki és informatikai felsőoktatás eredményességét a természettudományos tárgyak óraszámának csökkenése és (hosszabb távon) a tanárképzés gondoljai is befolyásolják.

Saját tapasztalataim szerint az itt eltöltött 14 évben jelentősen csökkent a hozzánk érkező hallgatók matematikai tudása, amely jelentősen megnehezíti az arra épülő fizika és mérnöki tantárgyak oktatását. A mérnöki tárgyakban elég komoly nehézséget okoznak azok a részek, amely komplex mennyiségeket vagy matematikai analízist tartalmaznak, pedig az ezeket tárgyaló matematikakurzusok megelőzik ezeket a tárgyakat.

Hallgatóink több hálózattal is találkoznak tanulmányaik során. Sokan komoly szakértői lesznek a számítógépes hálózatoknak, és gyakorlatban is képesek ilyen hálózatokat megtervezni és létrehozni. Villamosmérnök hallgatóink tanulmányaik során megismerkednek a telefonhálózatokkal is. Informatikus hallgatóink, a számítástudományi ismereteik megalapozásaként jelentős gráfelméleti tudásra tesznek szert, amely jó alap ahhoz, hogy megismerkedjenek az összetett hálózatok megfigyelt tulajdonságaival, és a vizsgálatuk informatikai hátterével is.

Az összetett hálózatok olyan összetett rendszerek, amelyek rengeteg hasonló egyedből (például személyekből, internetes útválasztókból, weboldalakból, fehérjékből, idegsejtekből) épülnek fel úgy, hogy bizonyos egyedpárok között kapcsolat van valamilyen szempont szerint, más párok között pedig nincs. Gyakran az összetett viselkedést nem ezeknek az egyedeknek az összetettsége okozza, hanem az, hogy a köztük lévő kapcsolatok hogyan oszlanak el. Az ezredforduló adta meg azt a lehetőséget, hogy számos nagy méretű hálózatot feltérképezhessünk és tanulmányozhassunk. A tanulmányozott hálózatok között találhatók biológiai, információs, szociológiai és technikai hálózatok is. Ezek a hálózatok általában megegyeznek abban, hogy skálafüggetlenek, azaz a fokszámeloszlás negatív kitévős hatványfüggvényhez közeli, amit nem csak az átlagos fokszám jellemez, hanem az is, hogy a rengeteg átlagosnál kisebb fokszámú csúcs mellett néhány átlagosnál nagyságrendekkel nagyobb fokszámú csúcs is jelen van. A hálózatok elméletének gyakorlati alkalmazására is számos példát lehet már hozni a járványok terjedésének előrejelzésétől a védőoltások alanyainak kiválasztásán, az internetes protokollok tervezésén, a szervezetbeli fehérjék szerepének meghatározásán és a telefontarifák kialakításán keresztül a politikai választások során egy párt szavazóinak mozgósításáig.

A hálózatok tudományát és a gráfelméletet, bár jelentős részben azonos fogalomrendszert használnak, eltérő vizsgálati módszerek jellemzik. Míg a gráfelmélet – a matematika szokásainak megfelelően – a létező világról mit sem tudva, önmagában teljes, addig az összetett hálózatok tudománya döntően kísérleti tudomány, amely hálózatok mérhető tulajdonságait vizsgálja, és azoknak olyan mennyiségeit igyekszik meghatározni, amelyeknek szerepe van a hálózatok működésében, fejlődésében.

désében és a környezetével való kölcsönhatásban.

Olyan programcsomag kialakítását tűztem ki célul, amely informatikus és villamosmérnök hallgatók számára alkalmas hálózatok vizsgálata mellett hálózatokkal kapcsolatos algoritmusok implementálására is, nem csak TDK- vagy szakdolgozat, hanem egy tantárgy keretében is. Ennek részeként olyan hálózatok vizsgálatát terveztem megvalósítani, amelyek órán előállíthatóak, és amelyek tulajdonságai mellett a változások törvényszerűségei is felderíthetők. A programcsomag kialakítására azért volt szükség, hogy az említett tananyagban ne csak a hálózatok elméletét ismerhessék meg a hallgatók, képesek legyenek hálózatok tulajdonságait megvizsgálni, és a kapott eredményekből következtetéseket levonni a hálózat felépítéséről és működéséről, valamint képesek legyenek új problémák megoldására programot fejleszteni.

A fenti célok megvalósításának részeként programot írtam az Ubuntu és Debian Linux-terjesztések szoftvercsomag-hálózatának létrehozására. Mivel a szakirodalomban nem találtam ezeknek a hálózatnak olyan vizsgálatát, amely az összes fontos tulajdonságát megvizsgálta volna, ezért természetes volt, hogy ezzel is érdemes foglalkoznom.

Ha hálózatokon zajló folyamatokat (például fertőzésterjedést) vizsgálunk, vagy a hálózat működésének változását bizonyos hatások (például véletlen meghibásodások vagy célzott támadás) hatására, gyakran szükséges, hogy megvizsgáljuk, hogyan befolyásolják ezeket a hálózatok különböző jellemzői. Ezért terveztem olyan hálózatgeneráló program létrehozását, amely illeszkedik az előbb említett programcsomaghoz, és képes adott tulajdonságcsoporthal rendelkező hálózatok létrehozására.

Új fizikai területek tanítása azért is fontos, hogy lássák a diákok és hallgatók, a fizika nem egy lezárt terület, amilyenek 1900 körül gondolták. A témám oktatási jellege miatt igyekszem, ahol csak lehet, magyar nyelvű irodalmat is idézni. Bár a tudomány nemzetközi nyelve angol, a két téma megszerettetéséhez, véleményem szerint, legtöbb embernek az anyanyelvén keresztül vezet az út. Ha a tanulóban sikerült felkelteni az érdeklődést, és mélyebben el szeretne merülni a témában, úgyis szüksége lesz megismerni az angol szaknyelvet. A hálózatok tárgyalásakor az egyetemi órákon is magyarul használom a fogalmakat, de amennyire a programcsomagok használata indokolja, kitérek az angol megfelelőkre, és javaslom a beadandó programokban az angol szaknyelv használatát.

I. rész

Részecskefizika az oktatásban

Én is így próbálok csalás
nélkül szétnézni könnyedén.
Ezüstös fejszesuhanás
játsszik a nyárfa levelén.

József Attila

A könyvesboltokban és a Világhálón nem kevés olyan írás található, amelyben valaki azt állítja, hogy kitalált olyan elméletet, amely sokkal egyszerűbben és pontosabban képes magyarázni az ismert dolgokat, mint a ma elfogadott tudomány. Általában ezek a cikkek olyan összefüggéseket és kifejezéseket használnak, mint a többi tudományos cikk. Hogyan lehet akkor egy középiskolás diáknak megtudnia, hogy melyik az igaz?

Nem várhatjuk el, hogy a fejtegetésekben található logikai bukfenceket megtalálja, ahhoz esetleg kevés a tudása. Megtaníthatjuk viszont arra, hogy milyen módszerekkel igyekeznek a tudósok biztosítani, hogy „csalás nélkül tudjanak szétnézni” a világban. A vélemények ütköztetése a kutatóintézetekben, tudományos konferenciákon és tudományos cikkekben ezt a célt szolgálja.

A körülöttünk lévő világ működésének egy része jól megérthető középiskolás fizikatudással, más része csak kisebb-nagyobb egyszerűsítésekkel. Amellett, hogy a fizika alapvető összefüggéseit megtanulják a diákok, az is fontos, hogy megértsék a tudományos kutatás működését is. Erre is nagyon jó lehetőséget nyújtanak a 2005 óta évente megrendezett nemzetközi részecskefizikai diákműhelyek, amelyek rendezésébe volt szerencsém bekapcsolódni. Ezeken a műhelyeken középiskolás diákok vesznek részt, többnyire 11–12. osztályosok. A nap során, miután előadások keretében tisztázzák a szükséges tudnivalókat, a CERN-ben ténylegesen lejátszódtott fizikai folyamatok esetén beazonosítják, hogy milyen események zajlottak le, megszámlálják, hogy az egyes típusokból mennyi volt, az egyes mérőcsoportok eredményeit egy közös eredményként összesítik, és a nap végén több helyszínt videokonferenciával összekapcsolva megbeszélik, hogy a különböző helyen kapott eredmények összeegyeztethetőek-e egymással, ha van eltérés az eredményekben, az miből fakadhat. Ez a folyamat jól modellezi a kutatók munkáját, így lehetőséget teremt a tudományos megismerés folyamatának bemutatására. Az csak ráadás, hogy megismerkednek új fizikai fogalmakkal, illetve olyan, a tananyagban távol álló fogalmak kapcsolódnak össze, mint a részecskék mozgása mágneses térben, az anyag–energia egyenértékűség, a kvantummechanika statisztikus jellege, illetve a mérésiértékelés statisztikai fogalmai. További ráadások, hogy a diákműhely keretében megismerkednek hasonló érdeklődésű diákokkal, valamint – részben a

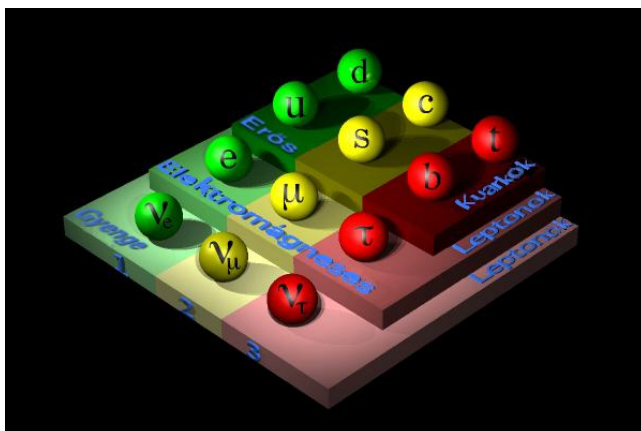
videokonferencián – a szakterület kutatóival is.

A részecskefizikai kísérletek összeállítása – főként annak magas technikai igényei miatt – egyre inkább sok ország közös vállalkozásaként valósul meg. Míg a részecskefizika iránti érdeklődés nagy, viszonylag távolra kell utazni, hogy a diák tényleg működő nagyenergiájú részecskegyorsítókat és részecskedetektorokat lásson. Bár van néhány diákcsoport, amelynek lehetősége volt meglátogatni az Európai Részecskefizikai Kutatóközpontot, a Genf mellett található CERN-t [3], és mindenképpen hasznos helyben látni, a részecskefizikai kutatás enélkül is közel hozható a középiskolás diákhhoz, és a felsőoktatásban tanuló hallgatókhoz.

Szerencsére a Világháló segítségével rengeteg hasznos információ érhető el a részecskefizikáról, többek között ismeretterjesztő cikkek, videók és oktatási segédanyagok. A Világhálón öröndetesen magas a magyar nyelvű színvonalas ismeretterjesztő anyagok száma, többek között a *Mindentudás Egyetemének* [4], az ELTE által szervezett *Atomoktól a csillagokig* előadássorozatnak [5], a *Fizikai Szemle* közkinccsé tett számainak [6] és a *CERN magyar nyelvű blogjának* [7] köszönhetően.

A Világhálón elérhető tudásanyag mellett fontosak a papíralapú könyvek és folyóiratok is. Ez részben átfedésben van az előző kategóriával, mivel ezek gyakran elektronikus formában is letölthetők szabadon vagy pénzért. Gyakran közöl részecskefizikával és CERN-nel kapcsolatos cikkeket a *Fizikai Szemle* és a *Természet Világa* [8].

Külön említeném az interaktív honlapokat. Egyszerűbb változatukra példa egy olyan ügyességi játék, amelyben egy elem polaritását kell megfelelő ütemben váltogatni, hogy a gyorsító a részecskét fel tudja gyorsítani. A később részletezendő *CERN-sajátkezüleg honlap* ennél lényegesen összetettebb kategóriába tartozik: egy teljes detektorrendszerben lehet vele nyomon követni ténylegesen mért eseményeket.



1.1. ábra. A részecskék családjai azokkal a kölcsönhatásokkal, amelyekben részt vesz. Amelyik részecske egy „dobogó” felett van, arra hat az a kölcsönhatás. A CERN-sajátkezüleg honlap egyik lefordított ábrája.

2. fejezet

Mérések a CERN-sajátkezüleg honlappal

2.1. A méréshez kapcsolódó fontosabb fogalmak és tények

A részecskefizika az anyag legalapvetőbb, tovább bonthatatlan építőköveivel – az úgynevezett elemi részecskékkel – foglalkozik. A mai tudásunk szerint az 1.1. ábrán látható részecskék az elemi részecskék, amelyek három családba sorolhatók. Az első részecskecsaládból felépíthető a körülöttünk lévő világ nagy része. Az atommagban található proton és neutron az u és d kvarkokból épül fel. Az atommag a körülötte lévő elektronokkal építi fel az atomot. Az atommagok bomlásakor és a csillagok energiatermelésében fontos szerepet játszik az elektronneutrínó (ν_e) is. A másik két család nagyon hasonlít az elsőre, a kvantumszámaik (elektromos és színtöltés, leptonszám, barionszám) egyeznek az első családban lévőkkel, csupán a részecskék tömege növekszik, ahogy a nagyobb sorszámú családok felé haladunk. Ilyen nagy fokú hasonlóság van az elektron, a müon (μ) és a tau részecske (τ) között, a hozzájuk tartozó neutrínók között, az u , s és b kvarkok között, valamint a d , c és t kvarkok között.

A részecskefizikai kísérletekben a részecskék kis tömege miatt a gravitációs kölcsönhatás nem játszik szerepet. A részecskefizika Standard modellje a következő három kölcsönhatás leírását foglalja magában [9]:

- az erős kölcsönhatást (az ábrán a legfelső dobogó), amely csak a kvarkok között hat, de ez tartja össze a kvarkokból felépülő neutronokat és protonokat is az atommagban, hasonlóan ahhoz, ahogy a semleges molekulák össze tudnak tapadni, mivel a töltésük nem egyenletesen oszlik el a térben.
- az elektromágneses kölcsönhatást (a középső dobogó), amely az elektromosan töltött részecskék között hat: az elektronra és nehezebb hasonmásaira, valamint a kvarkokra.
- a gyenge kölcsönhatást (legalsó dobogó), amely az összes általunk ismert anyagi részecskére hat.

A részecskefizika Standard modellje szerint az egyes kölcsönhatásokhoz is egy vagy több részecske tartozik, amely a kölcsönhatást (vonzást és taszítást) közvetíti. Az erőshöz a gluonok, az elektromágneseshez a fotonok, a gyenge kölcsönhatáshoz az elektromosan semleges Z bozon és az elektromosan töltött W^- és W^+ bozonok.

Minden anyagi részecskéhez tartozik egy olyan pár, amelyeknek három kvantumszáma (elektromos töltés, leptontöltés és barióntöltés) ellentétes előjelű, a többi tulajdonságuk – beleértve a tömeget is – megegyezik. A részecskéhez tartozó ilyen párt a részecske antirészecskéjének nevezzük. A CERN-ben az Antiproton-lassítónál (AD) találhatóak olyan kísérletek, ahol antihidrogént állítanak elő, azaz olyan atomot, amelyben a negatív elektromos töltésű antiproton körül kering anti-elektron, azaz pozitron. Az egyik ilyen jelentős magyar részvétellel működő kísérlet az ASACUSA, amelyben többek között az antihidrogén színeképét hasonlítják össze a hidrogénével, valamint olyan hélium színeképét, amelyben egy antiproton és egy elektron kering a héliumatommag körül (tehát a szokásos hélium egyik elektronját egy szintén negatív elektromos töltésű antiprotonnal helyettesítik), olyannal, amelyben az atom magja van antirészecskékből, és kívül proton és pozitron kering. Ezekkel a mérésekkel lehet tesztelni a CPT-szimmetria érvényességét [10].

A Nagy elektron–pozitron-ütköztető gyűrű (LEP) a CERN egyik részecskegyorsítója volt, amely 1989 és 2000 között elektronokat és pozitronokat gyorsított fel, majd több órán keresztül egymással szemben keringette azokat egymással ellentétes irányban. Ha a részecskék száma túlzottan lecsökkent, az ott található elektronokat és pozitronokat kiengedték az ütköztető gyűrűből, és újakat töltöttek be és gyorsítottak fel, azaz új futás kezdődött. Mivel az elektron és pozitron töltése ellentétes, ezért a keringetésükhöz ugyanaz a mágneses tér megfelelő, csak arra kell vigyázni, hogy az ellentétes irányba menő nyalábok a

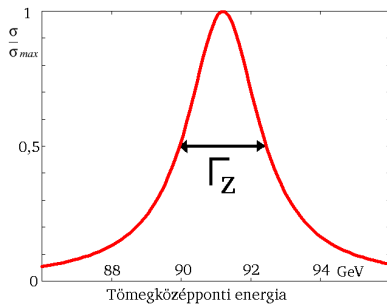
gyorsítóban ne keresztezzék egymás pályáját. A nyalábok egymás után haladó részecskecsomagokból álltak, csomagonként nagyjából 10^{11} részecskével. A csomagok kereszteződését a detektorok közepén kell előidézni megfelelő mágneses teret alkalmazva. A részecskecsomagok rengeteg elektronból illetve pozitronból álltak, úgyhogy a kereszteződés nagyjából olyan, mint amikor két szúnyograj áthalad egymáson. A LEP-ben átlagosan nagyjából 40000 ilyen csomag-találkozásból egyszer történt tényleges ütközés [11]. Mivel az elektron és a pozitron egymás antirészecskéi, ezért teljesen meg tudják egymás semmisíteni, ezt hívják elektron–pozitron szétsugárzásnak. Az elektron–pozitron szétsugárzás esetén általában a két részecske elektromágnesesen hat kölcsön, és fotonok viszik el az energiájukat, azonban a honlapon tanulmányozott esetekben az energia olyan különleges érték, amely a gyenge kölcsönhatás semleges részecskéjének, a Z bozonnak a nyugalmi tömegének megfelelő energiához (91,2 GeV [12]) nagyon közeli, így a fotonkeletkezéshez képest kivételesen nagyobb a valószínűsége, hogy a két részecske gyenge kölcsönhatással semmisüljön meg, és belőlük Z bozon keletkezzen:

$$e^- + e^+ \rightarrow Z \quad (2.1)$$

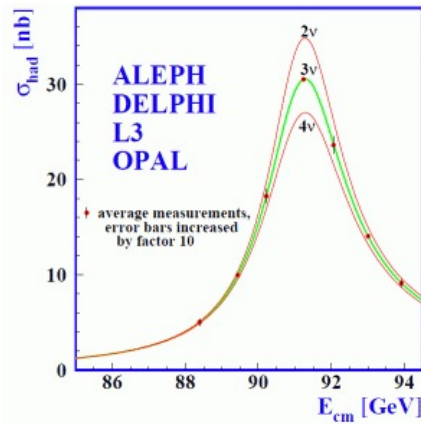
Mivel az elektron és a pozitron egymás antirészecskéi, a tömegük egyező. A gyorsítás során azonos energiára gyorsítják azokat, ami egyben azt is jelenti, hogy az ütközéskor a lendületvektoruk ellentett, tehát az összlendületük nulla. A Z bozon tehát nyugalomban keletkezik, és nagyon hamar elbomlik más részecskékre, maga a Z bozon közvetlenül nem is észlelhető. A jelenlétét viszont elárulja az, hogy a nyugalmi tömegének megfelelő energián megnövekszik a elektron–pozitron közötti kölcsönhatás valószínűsége, szakkifejezéssel élve megnövekszik a (2.1) egyenlettel leírt kölcsönhatás hatáskeresztmetszete, ahogy a 2.1a ábra mutatja. Az ábrán a fekete kétirányú nyíl ($\longleftrightarrow$) szemlélteti a kölcsönhatás egyik fontos jellemzőjét, a félértékszélességet: az ábrán látható görbe szélességét a maximális hatáskeresztmetszet felénél. A Γ félértékszélesség az – energia és az időtartam közötti határozatlansági relációnak köszönhetően – fordítottan arányos a keletkező részecske (itt a Z bozon) τ élettartamával [14, 11. oldal]:

$$\Gamma = \frac{\hbar}{\tau}. \quad (2.2)$$

A 2.1b ábra a LEP-es kísérletek mérési eredményeinek összesítését mutatja azokkal a görbékkel, amelyek akkor jönnének ki, ha 2, 3 illetve 4 neutrínófajta, és vele ugyanennyi részecskecsalád lenne. Az ábrán a mérési bizonytalanságot 10-szeresére növelték, hogy látható legyen. Egyértelműen látszik, hogy a mért pontok

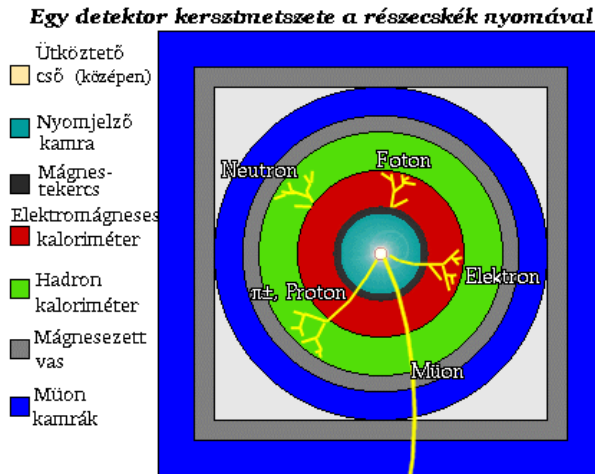


(a) A bomlási szélesség definíciója



(b) A LEP adatainak összesítésével nyert hatáskeresztmetszetek (piros pontok), a feltüntetett mérési hibák a valódi 10-szeresei, hogy egyáltalán láthatóak legyenek [13]. A folytonos vonalak azt mutatják, hogy 2, 3 illetve 4 neutrínó létezése esetén milyen görbe várható a Standard modellből

2.1. ábra. A Z bozon bomlási szélessége



2.2. ábra. Az egyes részecskék kölcsönhatása a detektorrétegekben

a középső görbére illeszkednek, tehát a mérés szerint nincs negyedik részecskecsalád.

2.2. A detektorok felépítése, a DELPHI detektor

Az összetett részecskefizikai detektorok több rétegből állnak (2.2. ábra), ami miatt a hagymához szoktunk hasonlítani azokat. Ezeket a rétegeket a feladatuk szerint szoktuk felosztani, és esetleg minden réteg további alrétegeket is tartalmazhat.

Az ütközés pontjától kifelé a legbelső rétegnek az a feladata, hogy az ütközésből szétrepülő részecskéknél a nyomát minél pontosabban rögzítsék anélkül, hogy azok mozgási energiáját jelentősen lecsökkentsék. Ezekben a töltött részecskék hagynak csak nyomot. Ez a réteg általában több alrétegből áll, amelyek közül az ütközés pontjához közelebb lévők felbontása általában nagyobb. A közeli felbontása lehetővé teszi, hogy olyan gyorsan bomló részecske esetén, amely nem jut el a legbelső detektorig, a belőle keletkező másodlagos részecske nyomán észrevegyük, hogy nem pontosan az ütközési pont irányából jön, hanem onnan, ahol az eredeti részecske elbomlott.

A nyomjelző detektoroknak két nagy csoportját különböztetjük meg a működé-

si elvük alapján: míg korábban szinte kizárólagosak volt a huzalkamrák különböző típusai, jelenleg a félvezető alapú helymeghatározás terjedt el. A huzalkamrákban valamilyen gázban fémszálak, anódok, találhatóak. A huzalkamrákban mozgó részecskék ionizálják a gázt, elektronokat ütve ki belőlük. Az elektronok az anódra felfutva elektronikus jelet hoznak abban létre. A felfutás idejét is figyelembe véve viszonylag nagy térrészben kevesebb szál használata is elegendő a pontos helymeghatározáshoz, ezen az elven alapulnak a *időprojekciós* kamrák.

A *nyomjelző detektorok* másik nagy csoportja, a félvezető alapúak, úgy működnek, hogy a félvezetőben a töltött részecskék elektron-lyuk párt hoznak létre. Az elektronok és a lyukak szétválaszthatóak, így az elektródákon mérhető jel jön létre. A félvezető-detektorok legnagyobb előnye a nagy, $10\ \mu\text{m}$ körüli, felbontás.

A következő detektorrétegek a kaloriméterek. Ezek a részecskék bennük leadott energiáját mérik. Kétféle kaloriméter található az összetett részecskefizikai detektorokban. A belső az elektromágneses kaloriméter, amelyben elektronok, pozitronok és fotonok teljesen elnyelődnek, és az energiájuk mérhetővé válik. A korszerű elektromágneses kaloriméterek, amelyre a CMS elektromágneses kalorimétere is példa, szcintillátorokat tartalmaznak. Az elektromágneses kaloriméterrétegben az elektronokat és a pozitronokat az atommaggal való kölcsönhatás lefékezi, miközben azok fékezési sugárzást (fotonokat) bocsájtanak ki. A nagy energiájú fotonok ebben a rétegben párkeltéssel elektron-pozitron párt hoznak létre. $10\ \text{MeV}$ energia felett a párkeltés mellett a másik két folyamat, a fotoelektromos hatás és a Compton-szórás hatása elhanyagolható [14, 355. oldal]. Tehát egy bejövő részecske ilyen átalakulások sorozatát hozza létre, amelynek az eredménye egy nagy elektron-pozitron-foton-zápor, amelyből a kisebb energiájú fotonokat észlelik, amelyek számára a szcintillátor átlátszó [14, 354. oldal].

A hadronok (a kvarkokból felépülő részecskék) ugyan adnak le energiát az elektromágneses kaloriméterben is, de az nem csökkenti le jelentősen az energiájukat. A második kaloriméterréteg, a hadronkaloriméter, feladata az, hogy a hadronokat elnyelve, azok energiáját megmérje. A működésük hasonlít a DELPHI detektor elektromágneses kaloriméteréhez, csupán a fékezés döntő mértékben az erős kölcsönhatáson alapul, és sokkal vastagabb anyagréteg szükséges a hadronok lefékezéséhez, amelyre az olcsóbb acélt használják az ólom helyett.

A müonok ugyan minden eddigi rétegben adnak le valamekkora a energiát, a két kaloriméterréteg nem tudja megállítani azokat. A müon névbe ebben a bekezdésben az antimüont is beleérttem. A legkülső rész a müonkamrákból álló réteg, mely ugyan továbbra sem tudja megállítani a legnagyobb energiájú müonokat, de a müonok áthaladását észlelni tudja, lényegében a hadronkaloriméterrel azonos

elvek alapján, ezzel megkönnyítve a részecskeazonosítást, mivel eddig a rétegig az ütközés pontjától csak a müonok és neutrínók jutnak el. A neutrínók esetén pedig rendkívül kicsi az esély, hogy valamelyik réteggel kölcsönhassanak, így keletkezésükre csak a hiányzó energiából (lendületből) következtethetünk.

Fontos megemlíteni, hogy a detektorokban mágneses teret hoznak létre. Ez teszi lehetővé, hogy a töltött részecskék p lendületét meghatározzuk a pályájuk R görbületi sugarából. A legegyszerűbb esetben, a B mágneses indukciójú homogén mágneses térben, a mágneses indukcióvektorra merőlegesen halad a $q = ze$ töltésű (e az elemi töltés) m tömegű részecske, a rá ható erő nagysága

$$qvB = F_{cp} = m \frac{v^2}{R} \quad (2.3)$$

amiből a lendületre

$$p = mv = qBR = zeBR \quad (2.4)$$

összefüggést kapjuk. Bár a levezetés során mindenhol a fizika középiskolában tanult összefüggései szerepelnek, és nem vettük figyelembe a relativitáselméletet, a (2.4) képlet a relativisztikus sebességekre is érvényes azzal a megjegyzéssel, hogy abban az esetben a lendület definíciója kicsit másképp használatos:

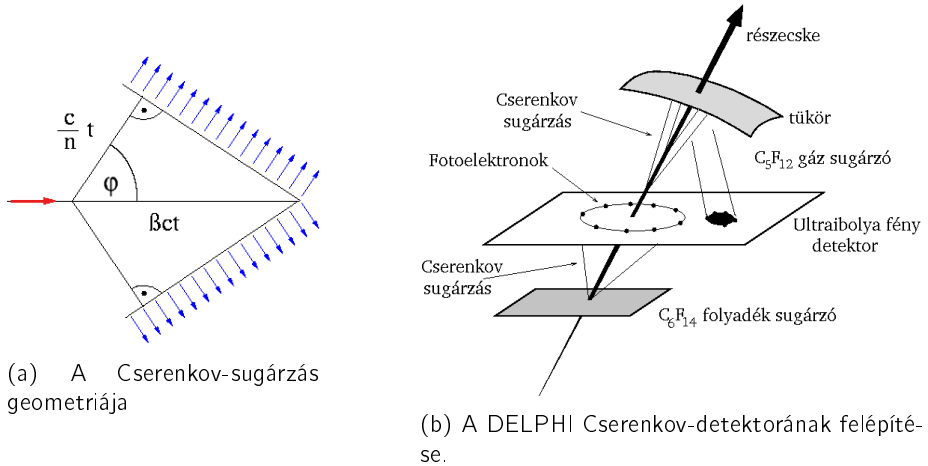
$$p = \frac{m}{\sqrt{1 - (\frac{v}{c})^2}} v = \gamma(v)mv, \quad \text{ahol} \quad \gamma(v) = \frac{1}{\sqrt{1 - (\frac{v}{c})^2}}. \quad (2.5)$$

A mágneses indukcióvektor ismert a detektor minden pontjában. Ha tudjuk, hogy a részecske töltése megegyezik az elemi töltéssel, vagy annak ellentettje, akkor a töltés előjele és a lendülete az eltérülés irányából és a pályasugárból meghatározható.

Az összetett részecskefizikai detektorokban további rétegek szolgálhatnak a részecskék azonosítására. A részecske sebessége – a repülési idő mérésén kívül – a Cserenkov-sugárzás kúpszögéből határozható meg. A Cserenkov-sugárzás akkor jön létre, ha a részecske egy közegben az adott közegbeli fénysebességnél nagyobb (de természetesen a vákuumbelinél kisebb) $\beta \cdot c$ sebességgel halad. A Cserenkov-sugárzás geometriáját a 2.3a. ábra mutatja, amiből a sebességet jellemző β számérték, és maga a sebesség is kiszámítható:

$$\beta = \frac{1}{n \cos \varphi}, \quad v = \frac{c}{n \cos \varphi}. \quad (2.6)$$

A sebesség és a lendület ismeretében a részecske tömege meghatározható, a részecske azonosítható. A DELPHI detektorban ezt a mérést végzi el a Cserenkov-



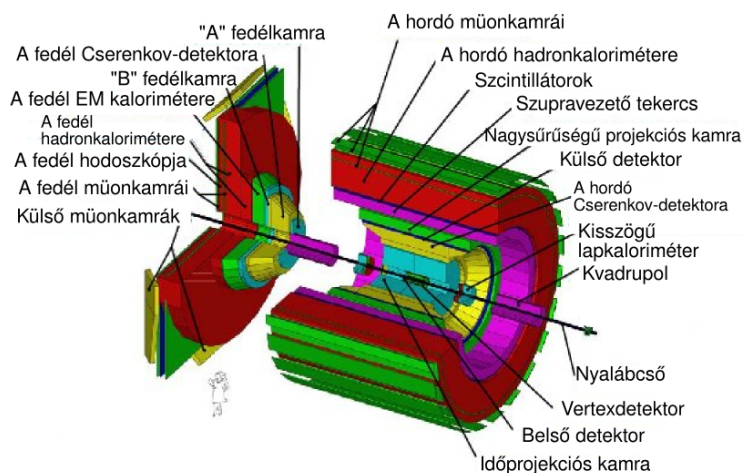
2.3. ábra. Ábrák a Cserenkov-sugáráshoz

detektor, amelynek a felépítését a 2.3b. ábra mutatja. A DELPHI Cserenkov-detektorában részben folyékony, részben légnemű közegben haladó részecske fényét ultraibolya fényre érzékeny detektorok érzékelik, és a detektorra vetülő ellipszisgyűrű adataiból meghatározható a részecskék sebessége.

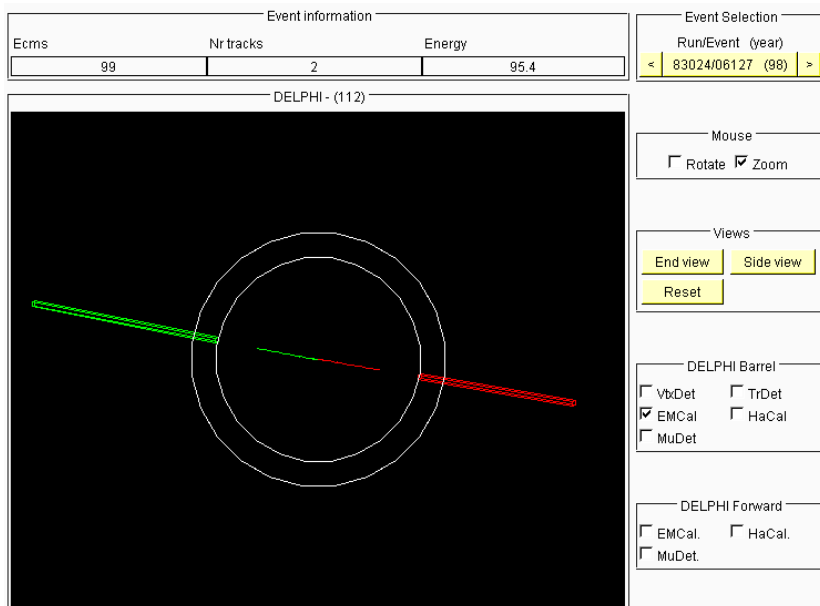
A DELPHI detektor felépítése a 2.4. ábrán látható [15]. A helymeghatározást végző rétegek a vertexdetektor, a belső detektor, az időprojekciós kamra és a külső detektor. Vertexnek nevezzük az ütközés pontját. Erről kapta a nevét a legbelső detektorréteg, mivel az a vertexhez közel található. Az időprojekciós kamra szintén a részecskék helyét határozza meg, de sokkal nagyobb térrészben, kisebb felbontásban.

Kaloriméterek és müonkamrák a hordórészben és a fedélrészben is találhatóak. A DELPHI hordórészének elektromágneses kaloriméterét nevezik nagysűrűségű projekciós kamrának. Ettől kifelé a hordórészben először a mágneses teret létrehozó szupravezető tekercs következik, majd a hadronkaloriméter és a müonkamrák.

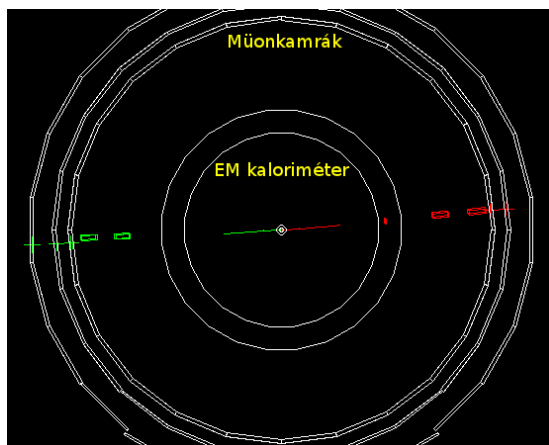
Az CERN-sajátkezüleg oldal eseménynézegetőjének ábrái láthatóak a 2.5–2.7. ábrákon. A 2.5. ábrán a teljes eseménynézegető látható. Ennek felső sávjában látható az ütközés energiája, a részecskenyomok száma, valamint a mért energia. Ha a mért energia jelentősen kisebb, mint a ütközés energiája, akkor feltételezhető, hogy olyan részecskeátalakulás történt, amelyben a neutrínó keletkezett, és az láthatatlanul elvitt nagy energiát.



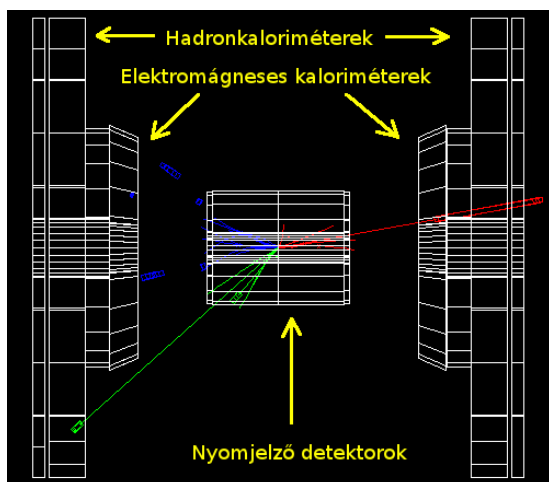
2.4. ábra. A DELPHI detektor felépítése. Jobb oldalon a hordórész, baloldalt a fedélrész látható. A detektor hordórészét mindkét végén fedélrész zárja le, amelyek közül az ábrán csak az egyik látható.



2.5. ábra. Elektronos esemény a CERN-sajátkezüleg honlap eseménynézegetőjén. A detektort a végéről látjuk, azaz az elektron és pozitron szemből illetve a hátunk mögül jön. Az ábrán a hordórész elektromágneses kaloriméterének körvonalai láthatóak.



2.6. ábra. Müonos esemény a CERN-sajátkezüleg honlap eseménynézegetőjén, a detektor vége felől nézve



2.7. ábra. A végállapotban 3 hadronzáport tartalmazó esemény a CERN-sajátkezüleg honlap eseménynézegetőjén. A detektort oldalról látjuk, azaz a nyálábok nyomvonalára merőlegesen látunk rá.

Jobbra fent található a futás sorszáma, azon belül az esemény sorszáma, valamint, hogy melyik évben történt az ütközés.

A jobb alsó sarokban lehet beállítani, hogy melyik detektorrészek körvonala látszszanak. Külön részben találhatóak a hordórészhez és a fedélrészhez tartozó detektorrétegek.

A látható eseményeket négy típusba soroljuk, amelyeknek a felismerési lehetőségét is megadom az alábbi felsorolásban. Az ötödik típusba azok a láthatatlan események tartoznak, ahol az ütközés után csak neutrínók keletkeznek.

1. Elektronos események esetén a Z bozonból egy elektron és egy pozitron (antielektron) keletkezik. Az elektron és pozitron egymással ellentétes irányba indul a lendületmegmaradás miatt. Mindkettő nyomot hagy a nyomjelző kamrákban, mivel töltöttek; az elektromágneses kaloriméterben teljesen elnyelődnek, amely az eseménynézegetőben nagy energialeadásban (hosszú hasáb) nyilvánul meg. (2.5. ábra)
2. A müonos események esetén a Z bozonból egy müon és egy antimüon keletkezik. Ezek a lendületmegmaradás miatt egymással ellentétes irányba indulnak. A belső detektorrétegek mindegyikével kis mértékben kölcsönhatnak, de a legnagyobb energiát a müonkamrákban adják le. A neutrínók kívül csak a müonok jutnak el a müonkamrákig. (2.6. ábra)
3. Tau-részecskés események esetén a Z bozonból egy tau és egy anti-tau részecske keletkezik. Ezek gyorsan elbomlanak, még azelőtt, mielőtt a nyomjelző detektorokba jutnának. Csak a bomlástermékeiket láthatjuk, amelyek (töltésmegmaradás miatt) 2, 4 vagy 6 töltött részecske nyomaként tűnnek fel. Ezek a másodlagos részecskék tehát a nyomjelző kamrákban nyomot hagynak, pályájukat itt a mágneses tér jelentősen meggömbíti.
4. Hadronzáporok esetén a Z bozonból két, egymással ellentétes irányba elinduló kvark keletkezik. Mivel a kvarkok kis energiákon nem lehetnek szabadok, ezért sok, nagyjából egy irányba haladó hadron keletkezik mindkét kvarkból, még mielőtt a legbelső detektorréteghez, a vertexdetektorhoz érnének. Ezt a jelenséget nevezzük hadronzápornak. Számuk kettőnél több is lehet, mert a kvarkok ütközben gluont bocsájthatnak ki, amely szintén hadronzáport hoz létre. A hadronzáporok sok hadronból állnak, amelyek a hadronkaloriméterben elnyelődnek, jelentős energiát adva le ott. A hadronok lehetnek elektromosan töltöttek vagy semlegesek, így a nyomjelző kamrákban vagy hagynak jelet, vagy nem. (2.7. ábra)

Azt, hogy egy eseménytípus milyen valószínűséggel fordul elő, az esemény elágazási arányának nevezzük.

2.3. A CERN-sajátkezüleg honlap felépítése

Az általam létrehozott *CERN-sajátkezüleg honlap* alapja a *Hands on CERN* angol nyelvű honlap 1-es verziója volt. Ezt a honlapot a Stockholmi Egyetemen dolgozó Erik Johansson professzor vezetésével készítették [16]. A honlap két fontos részből épül fel:

- Az *eseménynézegető* egy JAVA nyelven írt alkalmazás, amivel a CERN Nagy elektron–pozitron–ütköztető gyűrűjének DELPHI detektorában lezajlott eseményeket lehet megvizsgálni.
- Az *elméleti összefoglaló* részletezi az eseménynézegetővel elvégzendő mérés elméleti és kísérleti háttérét középiskolás szinten, valamint azokat a jellemzőket, amely alapján az eseményeket különböző csoportokba sorolhatjuk.

Az eseménynézegető tartalmaz tíz eseménycsoportban összesen $10 \cdot 100$ eseményt, amelyeket 1998-ban a LEP 91 GeV-es energiájú ütközéseiben rögzítettek. Az eseményeket csupán azért csoportosították 10 csoportba, mert a honlappal történő mérés tipikusan úgy zajlik, hogy 10 mérőpár dolgozik egy-egy számítógépnél, és mindegyik mérőpár egy száz eseményből álló eseménycsoportot kap feldolgozásra. Az eseménycsoportokban szereplő események statisztikailag helyesek olyan értelemben, hogy az egyes eseménytípusok elágazási arányait helyesen kapják meg a hallgatók. A mérés során a diákok feladata az, hogy az eseményeket csoportosítsák és meghatározzák az egyes elágazási arányokat.

2.4. A CERN-sajátkezüleg honlap fordítása és bővítése

Lefordítottam a teljes Hands on CERN honlapot. A honlapon található képek nagy részének feliratait is lefordítottam, és ahol indokoltnak láttam, további képanyaggal egészítettem ki a honlapot. A lefordított ábrák közé tartozik a 1.1., 2.2. és 2.1. ábra.

A CERN-sajátkezüleg honlap nem pusztán az angol változat fordítása [17]. A honlap ezen felül részletezi a témájához kapcsolódó magyar szakirodalmat is.

1943	Otto Stern német származású amerikai fizikus	1992 Nobel honlap 1992 Nobel díj  "részecske-detektorok ki- és továbbfejlesztéséért, különös tekintettel a többszálú proporcionalis számlálóra" "for his invention and development of particle detectors, in particular the multiwire proportional chamber" Georges Charpak lengyel születésű francia (1924 →)
1944	Isidor Isaac Rabi amerikai fizikus	
1945	Wolfgang Pauli osztrák fizikus	
1948	Patrick Maynard Stuart Blackett angol fizikus	
1949	Hideki Yukawa japán fizikus	
1950	Cecil Frank Powell angol fizikus	
1955	Polykarp Kusch német származású amerikai fizikus	
1957	Chen Ning Yang és Tsung-Dao Lee kínai fizikusok	

2.8. ábra. A Nobel-díj oldal egy részlete, és a Charpakról szóló rész

Kiegészült a honlap egy olyan lappal is, amelyen a részecskefizikai eredményért kiosztott Nobel-díjasok találhatók év szerint rendezve. Mindegyik évhez egy vagy két külön oldal is található, ahol a Nobel-díjasok arcképe látható, és a Nobel-díjak indoklása, esetenként kis kiegészítéssel (2.8. ábra). Minden ilyen oldalról el lehet jutni a Nobel-díj hivatalos oldalának megfelelő évhez tartozó lapjához is. Úgy gondoltuk, hogy az indoklás eredeti angol változatát is megtartjuk a honlapon, hiszen az a hivatalos, és nem árt, ha a tanulók az angol szakszavakat is megismerik.

2.5. A részecskecsaládok számának kiszámítása

Kidolgoztunk egy mérést középiskolás diákoknak az eseménynézegetőhöz, amellyel az események elágazási arányából kiszámítható a részecskecsaládok száma [18].

A méréshez szükséges a Z bozon tömege és bomlási szélessége (Γ_Z) is, melyeket a LEP-en az általunk tárgyalt méréstől függetlenül meghatároztak.

A kiinduló képlet, amelyet a középiskolás diákok számára nem részletezünk, a következő:

$$\sigma(e^+e^- \rightarrow x) = \frac{12\pi M_Z^2}{E_{CM}^2} \cdot \frac{\Gamma(Z \rightarrow e^+e^-) \cdot \Gamma(Z \rightarrow x)}{(E_{CM}^2 - M_Z^2)^2 + M_Z^2 \Gamma_Z^2}, \quad (2.7)$$

ahol x helyén a következő bomlástípusok állhatnak: e^+e^- , $\nu^+\nu^-$, $\tau^+\tau^-$, hadronos vagy láthatatlan (neutrínós). A képletben M_Z a Z bozon tömege, az E_{CM} a

tömegközépponti energia az ütközéskor, a $\Gamma(Z \rightarrow X)$ értékek az egyes bomlási módok járuléka a teljes bomlási szélességhez [19].

A diákok számára csupán azt mondjuk el, hogy a Standard modell szerint a különböző x eseménytípusok előfordulásának N_x várható száma a LEP-nél kiszámítható a következőképpen:

$$N_x = Nk\Gamma_e\Gamma_x. \quad (2.8)$$

A képletben az x -re a következő egyszerűsített jelöléseket használjuk:

x	eseménytípus
e	elektronos
μ	müonos
τ	taus
$hadron$	hadronzápor
n	nem látható, azaz neutrínós

Itt az N az összes észlelt látható esemény száma, k egy állandó, amely függ a Z bozon tömegétől és bomlási szélességétől, valamint az ütközés jellemzőitől: attól, hogy mennyi elektron és pozitron jön egymással szemben, mekkora a nyaláb keresztmetszete. Ennek értéke esetünkben $k = 5,964 \frac{1}{\text{GeV}^2}$. A további számolásokhoz érdemes kiszámítani az Nk szorzatot, amelyet a továbbiakban K -val jelölök. A Γ_x az x bomláshoz tartozó bomlási szélesség, az összes bomlástípusra összegezve a Γ_x -eket a Γ_Z -t kapjuk.

Az N_e ismeretében a Γ_e érték becsülhető, ennek ismeretében pedig a többi Γ_x érték is:

$$\Gamma_e = \sqrt{\frac{N_e}{K}} \quad \Gamma_x = \frac{N_x}{\Gamma_e \cdot K}. \quad (2.9)$$

Ezek bizonytalansága az alábbi képletek szerint becsülhető:

$$\Delta\Gamma_e = \frac{1}{2\sqrt{K}} \quad \Delta\Gamma_x = \frac{\Delta N_x}{\Gamma_e \cdot K} + \frac{N_x \cdot \Delta\Gamma_e}{\Gamma_e^2 \cdot K}. \quad (2.10)$$

Az $A_x = \frac{\Gamma_x}{\Gamma_Z}$ elágazási arányok, és azok ΔA_x bizonytalanságai ebből kiszámíthatóak a $\Gamma_Z = 2,495 \text{ GeV}$ -vel való osztással. A láthatatlan (neutrínós) események A_n arányát és annak ΔA_x bizonytalanságát a következőképp kapjuk:

$$A_n = 1 - (A_e + A_{mu} + A_{tau} + A_{hadron}) \quad (2.11)$$

$$\Delta A_n = \Delta A_e + \Delta A_{mu} + \Delta A_{tau} + \Delta A_{hadron} \quad (2.12)$$

A Standard modell alapján kiszámolható, hogy hányszor annyi az egyféle neutrínó keletkezésének a valószínűsége, mint elektron-pozitron páré. Erre 1,979-et kapunk.

A neutrínótípusok és ezzel a részecskecsaládok száma tehát kiszámolható, ha az összes láthatatlan esemény arányát elosztjuk az egyfajta neutrínó arányával:

$$N_\nu = \frac{A_n}{1,979A_e} \quad \Delta N_\nu = \frac{\Delta A_n}{1,979A_e} + \frac{A_n \cdot \Delta A_e}{1,979A_e^2}. \quad (2.13)$$

Például ezer eseménynél, ha

$$N_e = 45, \quad N_\mu = 46, \quad N_\tau = 25, \quad N_{hadron} = 884, \quad (2.14)$$

akkor

$$N_\nu = 3,284991, \quad \Delta N_\nu = 1,547977 \quad (2.15)$$

értékeket kapunk.

Összesen nagyjából 20 millió eseményből határozták meg a LEP-en, hogy ez az érték 3, pontosabban $2,994 \pm 0,012$ [12, 22. old.]. Írtam egy programot, amely az elágazási arányok ismeretében szimulálja a kapott eseménytípus-számokat adott teljes eseményszám esetén. Ezt többször lefuttatva adott eseményszámokra megvizsgálható, hogy mekkora teljes eseményszámnál milyen értékek várhatóak a részecskecsalád-szám bizonytalanságára.

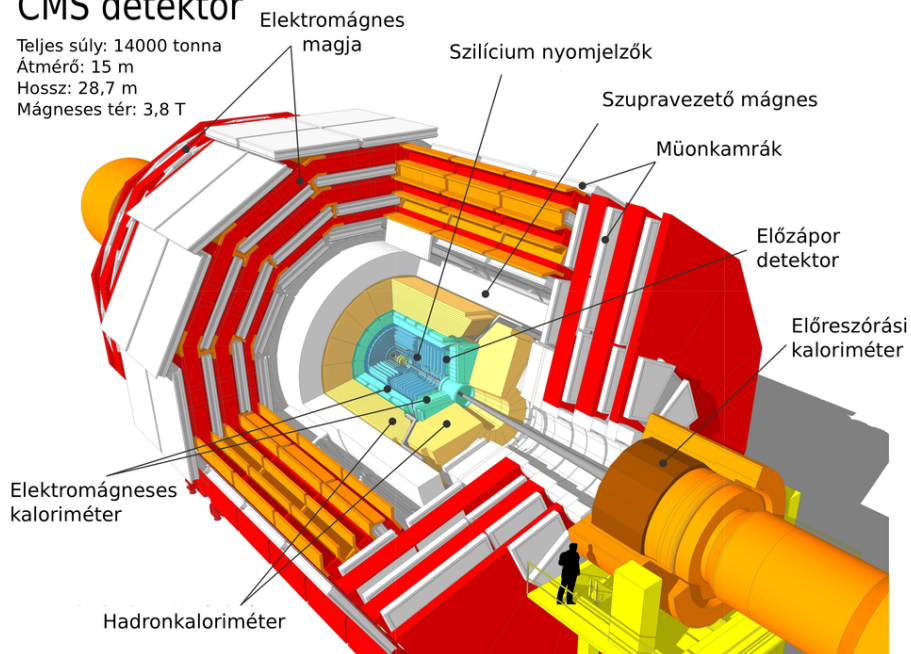
A lefuttatott szimulációk szerint száz esemény vizsgálata esetén a bizonytalanság nagysága nagyjából 6 körül van, de tíz feletti érték is könnyen előfordulhat, ha a véletlen úgy hozza, hogy nagyon kevés elektroneseeményünk lesz. Ezer esemény esetén másfél körül, tízezer esemény esetén fél körül, százezer esemény esetén 0,17, 20 milliónál 0,037 körül van a bizonytalanság nagysága. Ebből jól látszik, hogy igen nagyszámú esemény vizsgálata szükséges. A 20 milliós szimulációból látszik, hogy a (2.10) egyenletekből kapott bizonytalanság nagyságrendileg egyezik a LEP-nél kapott bizonytalansággal. A különbség részben abból adódhat, hogy a szimulációmnál nem vettem figyelembe azt, hogy a leptonok elágazási aránya közel egyező, csak a tömegük miatt van különbség.

2.6. A CMS detektor

A CERN jelenlegi legnagyobb gyorsítója a Nagy hadronütköztető gyűrű (LHC), ami a LEP alagútját hasznosítja újra. Működésének döntő részében protonok ke-

CMS detektor

Teljes súly: 14000 tonna
Átmérő: 15 m
Hossz: 28,7 m
Mágneses tér: 3,8 T



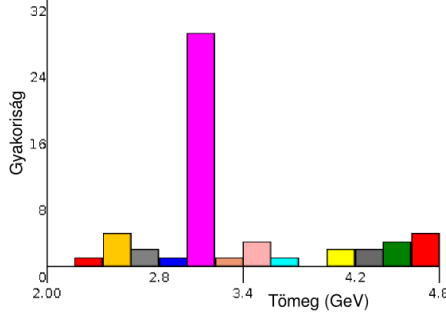
2.9. ábra. A CMS-nek, a Nagy hadronütköztető gyűrű egyik nagy általános célú detektorának a felépítése

ringenek benne egymással szemben. Mivel, szemben a LEP-pel, az LHC-ban szupravezető mágnesek tartják körpályán a protonokat, az energiafogyasztása nem növekedett: a villamosenergia-felhasználása nagyjából Debrecen városának felhasználásával egyezik.

Az LHC-ben két nagy általános célú detektor található, és néhány kisebb, specializáltabb. Az utóbbiak közül az ALICE főleg az ólom-ólom ütközésekre szakosodott. Az LHC-ben az ionok olyan nagy energiával ütköznek, hogy a kvarkok és a gluonok bennük rövid ideig szabad részecskéként viselkednek, úgynevezett kvark-gluon plazmát alkotnak. Ennek a tanulmányozása az ALICE feladata. A másik speciális detektor az LHCb, amelynek a feladata a CP-sértés paramétereinek részletes vizsgálata, hogy megértsük, miért van több anyag a Világegyetemben, mint antianyag.

A két nagy általános célú detektor egyike az ATLAS, amely a CERN fő telephelyének közelében található. Ezt rögtön a föld alatti üregében szerelték össze, ezért méretei nagyok lehetnek: 45 méter a hossza, és 25 méter az átmérője. A külső részében toroid alakú szupravezető mágnes hozza létre a mágneses teret. A másik nagy általános detektor, amelyben a Magyarország képviselőjében a CERN-ben dolgozó részecskefizikusok is részt vesznek, a Kompakt müonszolenoid (CMS). Azért kompakt, mert kisebb helyen több acélt tartalmaz, mint az ATLAS. A súlya 14000 tonna, nagyjából kétszerese az Eiffel-torony súlyának, döntő részét a benne található acél súlya teszi ki. Ennek az acélnak háromszoros szerepe van: szilárdan rögzíti az egyes detektorrészeket, hogy a mágneses tér hatására ne mozduljanak el, alakítja a szupravezető szolenoid által létrehozott mágneses teret, és a müonokkal kölcsönhatva, azok energiájának mérésében is részt vesz. A kis térfogat azért is volt fontos, mert a CMS-t eredetileg a földfelszínen szerelték össze tesztelésre, és csak utána eresztették le kerékszerű szeletenként végleges helyére.

A CMS felépítése a 2.9. ábrán látható. Belül a részecskék pályájának nagyon pontos mérésére szolgáló detektorok találhatók, amelyek félvezető alapúak. A következő réteg az ólomvolframát-kristályokból álló elektromágneses kaloriméter. Ez az anyag kisebb energiájú fotonok számára teljesen átlátszó, míg az ütközésben keletkező nagy energiájú fotonok számára áthatolhatatlan, azok a már említett módon elektron–pozitron–foton-záport hoznak létre. Az ólomvolframát-kristály radiációs hossza $X_0 = 0,89$ cm [20], ami azt jelenti, hogy az elektronok energiája fékezési sugárzás miatt ekkora távolságonként e-edrészére csökken. A CMS hadronkalorimétere kifelé haladva felváltva tartalmaz nagy sűrűségű anyagot (sárgarézet vagy acélt), és műanyag szcintillátort. A nyomjelző detektorok és a kaloriméterek belül találhatók a szupravezető szolenoidmágnesen. Ez a mágnes 4



2.10. ábra. A J/ψ mérésen kapott tömeghisztogram. A vízszintes tengelyen az invariáns tömeg található GeV egységben, a függőlegesen a talált események száma.

Tesla nagyságú mágneses teret hoz létre, amely 5 nagyságrenddel nagyobb, mint a Föld mágneses tere a felszínen.

A CMS műonkamráinak fejlesztésében a Debreceni Egyetem is részt vett. A kamrák relatív helyzetét folyton ellenőrizni kell, mivel a hőtágulás a mágneses tér, és a Hold gravitációs hatása miatt azok helyzete megváltozhat. A műonkamrákra LED-et szereltek, amelyek helyzetét kamerák vizsgálják folyamatosan, és a méréskor a mért áthaladási helyének meghatározásakor az eltéréseket figyelembe veszik [21].

2.7. CMS-sel kapcsolatos mérések a diákműhelyen

A CMS-sel kapcsolatos mérések során a DELPHI nézegetőjéhez hasonló eseménynézegetővel mérnek a diákok [22]. A részecskefizikai diákműhelyeken kétféle CMS-sel kapcsolatos mérés szerepelt eddig, mindkettőben szerepelt egy műonpár invariáns tömegének kiszámítása, ezért ezt a fogalmat részletezem először. A kiszámítását azzal a részecskefizikusok által használt egységrendszerrel írom fel, amelyben a c vákuumbeli fénysebesség értéke 1, így az m (invariáns) tömege egysége GeV/c^2 helyett GeV -ben, a p lendület egysége GeV/c helyett szintén GeV , azaz mindkettő egysége megegyezik az E energia egységével. Ezt az egységrendszert használva a műonpár invariáns tömege a következő lesz:

$$m_{\mu^+\mu^-}^2 = (E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2, \quad (2.16)$$

ahol a két müion energiája E_1 illetve E_2 , a lendületeik pedig $\vec{p}_1$ illetve $\vec{p}_2$.

A diákműhely egyik mérésében azokat az eseményeket kell kiválogatni a két müionos események közül, ahol J/Ψ részecske keletkezett. Egy táblázatban megtalálhatóak az eseményekben található két müion energiája és lendületeik három komponense. Azokat kell kiszűrni ezek közül, amelyek J/Ψ eseményre utaltak, és ebből el lehet készíteni a tömeghisztogrammot. A J/Ψ események közé sorolt eseményeket táblázatkezelővel gyorsan ki tudjuk választani, és a hisztogrammot valamilyen eszközzel kirajzoltathatjuk. A diákműhely szervezői erre internetes oldalt ajánlottak, ahol, ha feltöltjük a tömegértékeket, különböző binszélességekkel kirajzoltathatjuk a tömegeloszlást. A 2010-es diákműhelyen a 2.10. ábrán látható hisztogrammot kaptuk. Jól kiemelkedik rajta a $[3,0 \text{ GeV}; 3,2 \text{ GeV}[$ intervallum eseményeinek száma. A csúcs helye jól egyezik a J/Ψ tömeg $3,097 \text{ GeV}$ -es értékével [12, 86. old.].

A másik mérésnél olyan eseményeket kell kiválasztani, amelyben W vagy Z bozon keletkezett, meg kell vizsgálni háromféle eseménypár gyakoriságának arányát (W/Z , W^+/W^- , e/μ). Végül ábrázolni kell a Z mérésekből származó invariáns tömegek eloszlási grafikonját, hasonlóan a J/Ψ méréshez.

3. fejezet

Részecskefizikai diákműhelyek

3.1. A részecskefizikai diákműhelyek története

A CERN-sajátkezűleg oldalra építve szervezte meg először 2005-ben az Európai Részecskefizikai Ismeretterjesztő Csoport, az EPPOG, a részecskefizikai diákműhelyt. A szervezést azóta is ez a csoport szervezi, de azóta a csoport nemzetközivé vált, így a rövidítése IPPOG (International Particle Physics Outreach Group) lett [23].

A diákműhely 2005 óta minden év tavaszán kerül megrendezésre. Kezdetben két héten, újabban három héten keresztül naponta 5–10 egyetemen, illetve kutatóintézetben mérnek a diákok a CERN-sajátkezűleggel, vagy ahhoz hasonló alkalmazással. Miután az LHC első eredményei világot láttak, a diákműhely során olyan alkalmazásokat használunk, amelyek a CMS, az ATLAS és az ALICE detektorok eseményeit jelenítik meg. A mérést általában 20 diák végzi 10 mérőpárban, mindegyik mérőpár az adatok tizedrészét kapja meg feldolgozásra.

A diákműhelyek ajánlott napirendje alig változott az évek során. A délelőtt során a diákok bevezető előadásokat hallgatnak meg a részecskefizikáról, a gyorsítóról és a részecskedetektorokról és a CERN-ről. Az ebédszünet előtt az adott egyetem vagy kutatóintézet általában a saját laborjait mutatja be. Az ebéd után történnek a mérések. Általában a mérés kezdetekor lassabban haladnak a diákok, ilyenkor lehet szükség akár több segítőre is, akik az adott eseményeket be tudják sorolni a kívánt kategóriákba. A vége felé a diákok már sokkal gyorsabban haladnak. A gyorsabbak akár két adatsort is végig tudnak nézni. A videokonfe-

renciák során a CERN-es helyszín mellett általában 4-5 diákműhely-helyszín vesz részt, akik nem feltétlenül azonos, de összehasonlítható mérést végeztek el. Az LHC indulása előtti években a CERN-sajátkezüleg DELPHI-ből származó adatai mellett lehetőség volt egy másik honlappal az OPAL eseményeit csoportosítani, és a kapott elágazási arányok (branching ratio) összehasonlíthatóak voltak. Szintén összehasonlítható az elágazási arányokból kapható érték, amely az erős kölcsönhatás csatolására (α_s) adódott.

A videokonferenciát a CERN fizikusaiból alkotott párok vezették. Ők foglalták össze az egyes csoportok által mért eredményeket, és hasonlították össze az irodalmi adatokkal. A mérés értékelése után tőlük kérdezhettek a diákműhely résztvevői bármilyen, a nap során felmerült kérdést, amire a mérés helyszínén nem kaptak választ, vagy a videokonferencia során merült fel, akár a CERN-ben dolgozó fizikusok életmódjával kapcsolatos kérdéseket is. A videokonferencia során egy kvízzjáték is szerepel, amelyben szintén párokban dolgozó diákoknak kell a négy válaszlehetőség közül kiválasztani a helyeset.

A nap zárásakor van lehetőség az esetleges okleveleket és ajándékokat átadni a résztvevő diákoknak.

A diákműhely rendezvényein kezdettől fogva három magyar helyszín vesz részt,

- a Debreceni Egyetem,
- a Wigner Fizikai Kutatóközpont (korábban KFKI) Részecske- és Magfizikai Kutatóintézete,
- és az Óbudai Egyetem székesfehérvári Alba Regia Egyetemi Központja (korábban Budapesti Műszaki Főiskola, Számítógéptechnikai Intézete).

Mindhárom helyszínen 2013 lesz a kilencedik alkalom, hogy diákműhelyt szervezünk.

A székesfehérvári helyszín eseményeinek a szervezését, az előadások tartását és a mérések levezetését én végzem kezdettől fogva. Székesfehérváron általában a székesfehérvári gimnáziumok diákjai vettek részt, de nagy létszámban voltak jelen a dunaujvárosi Széchenyi István Gimnázium fizika fakultációs diákjai is, nem egy közülük két alkalommal is. Időnként Fejér megyén kívülről is érkeztek diákok. 2008-ban 5 fő vett részt a győri Révai Miklós Gimnáziumból. 2009-ben 6 fő vett részt a keszthelyi Vajda János Gimnáziumból és 4 fő a budapesti Szent István Gimnáziumból, sőt a keszthelyiek tanára, Farkas László, is részt vett a rendezvényen.

3.2. A diákműhely kérdőíveinek kiértékelése

A 2005-ös diákműhely nemzetközi szervezői megkértek minket, helyi szervezőket, hogy töltsünk ki egy kérdőívet a nap végén a diákokkal. A kérdések magyar fordítása megtalálható a 3.1. és 3.2. táblázatokban. Az elsőben főként a korábbi ismeretekről szóló kérdések vannak, a másodikban pedig a diákműhelyt értékelték a diákok. Debrecen és Székesfehérvár kérdőíveit dolgoztam fel. A legtöbb válasz esetén 1 és 5 közötti számértékkel tudtak válaszolni a diákok. Az előbb említett két táblázatban a válaszok átlaga található, külön az egyes helyszíneké [24].

A székesfehérvári és debreceni csoport értékei legtöbb helyen közel azonosak. Általában kevesebb, mint egy órát töltöttek naponta a résztvevő diákok 2005-ben a számítógép előtt. A tanórákon a legtöbb diák nem nagyon gyakran, de látott kísérleteket, viszont adatok kiértékeléséhez nem sok helyen használtak számítógépet. Legtöbben nagyjából egyetértettek azzal az állítással, hogy sokat tud fizikából, és (különösen a debreceni helyszín diákjai) úgy érezték, hogy többet szeretnének tudni fizikából. Eddig többet tudtak a radioaktivitásról, és keveset a kvarkokról, leptonokról, részecskefizikáról, részecske-detektorokról és részecskegyorsítókról, a nap során viszont az előzőek közül nem nagyon került szóba a radioaktivitás, de a többi témakörből sokat tanultak. Az előadásokat és a feladatokat érdekesnek találták, és úgy vélték, hogy lehetőségük volt kérdezni. Összességében tetszett a diákoknak a videokonferencia. Debrecenben inkább az előadás, Székesfehérváron inkább a mérési feladat aratott sikert. A diákműhely a kérdőív szerint nehézségben éppen ideális volt. A diákok megismerték, hogyan szervezik és kivitelezik a tudományos kutatást, és tájékozódhattak a fizika mindennapi élettel való kapcsolatáról is. Legtöbben úgy vélték, hogy a modern fizikának és benne a részecskefizikának nagyobb súllyal kellene szerepelnie a fizikaórákon.

Kérdés	Debrecen	Székesfehérvár
Mennyi időt töltesz a számítógép előtt? (1=több mint 1h/nap, 5=kevesebb mint 1h/hét)	2,05	1,31
Milyen gyakran kísérleteztek fizikaórán? (1=nagyon gyakran, 5=sosem)	3,36	3,4
Milyen gyakran használtok számítógépet az adatok kiértékeléséhez? (1=nagyon gyakran, 5=sosem)	4,21	4,31
Előismeretek (1=teljesen egyetértek, 5=egyáltalán nem értek egyet)		
Sokat tudok fizikából	2,84	2,87
Többet akarok tudni belőle	1,26	2,18
A fizikaórák jól előkészítettek a diákműhelyre	3,05	2,62
Szeretem a fizikaórákat	1,36	2,37
Az iskolai fizika kapcsolódik a mindennapi élethez	2,44	2,56
Az iskolai fizikaórán megtudom, hogy mi a szerepe a fizikának a modern technikai fejlesztésekben	2,26	2,5
Az iskolai fizikaóráról tudom, hogyan szervezik és kivitelezik a tudományos kutatási folyamatokat	3,31	2,87
Mielőtt ide jöttem a következő dolgokról (1=sokat, 5=nem) tudtam		
Radioaktivitás	2,31	2,93
Kvarkok és leptonok	3,73	4
Általában a részecskefizikáról	2,84	3,62
Részecskedetektorok	3,10	3,81
Részecskegyorsítók	2,42	3,68

3.1. táblázat. A debreceni és székesfehérvári diákok által adott válaszok átlagértékei az előzetes ismeretekkel kapcsolatban

Kérdés	Debrecen	Sz.fehérvár
Előadások (1=teljesen egyetértek, 5=egyáltalán nem értek egyet)		
Az előadásokat könnyű volt megérteni	1,5	2,93
Érdekesnek találtam az előadás témáját	1,33	2,25
Lehetőségünk volt kérdezni	1,33	2,06
Feladatok (1=teljesen egyetértek, 5=egyáltalán nem értek egyet)		
A másokkal való együttműködés fontos volt a probléma megoldásában	1,61	2,18
Úgy érzem, hogy kompetens vagyok a probléma megoldásában	2,33	2,68
Lehetőségünk volt kérdezni	1,22	2,19
Videokonferencia (1=teljesen egyetértek, 5=egyáltalán nem értek egyet)		
A videokonferencia technikai színvonala (1=jó, 5=rossz)	3,81	2,69
Jól tudtam követni az egész konferenciát	2,88	2,83
A videókapcsolat a többi ország hallgatóival bepillantást engedett számomra a nemzetközi kutatási együttműködések működésébe	3,05	2,76
Általános		
1. Hogyan tetszett a diákműhely? (1=nagyon, 5=egyáltalán nem)	1,64	2,43
2. Mi tetszett a legjobban a diákműhelyben (e=előadás, f= feladat, v=videokonferencia)	11e, 3f	10f, 3e, 3v
3. A diákműhely (1=túl könnyű, 5=túl nehéz) volt	3	3,37
4. A diákműhely után (1=kevésbé, 5=jobban) érdekel a fizika általában	4,11	3,53
5. Mit szeretnék másként? (1=teljesen egyetértek, 5=egyáltalán nem értek egyet)		
Több feladatot szeretnék mint előadást a diákműhelyen	3	3,06
Jobban szeretem azokat a programokat, mely több helyet ad a saját ötleteimnek	3,18	2,75
A diákműhely után többet tudok a részecskefizikáról	1,82	2,56
A diákműhely tájékoztatót a fizikának a modern műszaki fejlesztésekben betöltött szerepéről	2,11	2,37
A diákműhelyen megtanultam, hogyan szervezik és kivitelezik a tudományos kutatást	2,23	2,62
A modern fizikának mint a részecskefizikának nagyobb mennyiségben kellene szerepelnie az iskolai fizikaórán	1,70	2,75
A diákműhelyen hallott fizika megmutatta a kapcsolatot a mindennapi élettel	2,94	2,87
7. A diákműhelyfoglalkozáson (1=sokat, 5=nem) tanultam újat a következőkről		
Radioaktivitás	3,35	3,81
Kvarkok és leptonok	1,52	1,93
Általában a részecskefizikáról	1,70	2,5
Részecskedetektorok	1,82	2,31
Részecskegyorsítók	2,17	2,18

3.2. táblázat. A diákműhely értékelése a tanulók véleménye szerint

4. fejezet

Oktatási és ismeretterjesztő segédanyagok

4.1. Közreműködésem a Wikipédiában

A Wikipédia egy olyan enciklopédia, amelyet – bizonyos ésszerű korlátozások betartásával – mindenki szerkeszthet. A nagy lehetősége az oldalnak, hogy az egyes szócikkek közötti kapcsolatok könnyen kialakíthatóak (sokkal átláthatóbban, mint ha saját magunk szerkesztenénk egy saját sztatikus honlapot), és a kapcsolatok megléte könnyen ellenőrizhető, hiszen a hiányzó hivatkozások kék helyett piros színnel jelennek meg az oldalon. Ennek következtében a szerkesztő például anélkül indulhat neki, hogy létrehozza egy természettudós életrajzát, hogy külön meg kellene írnia a születőhelyéről, elvégzett egyetemével, a munkahelyeivel kapcsolatos szócikket, hiszen ezek egy része már jó eséllyel készen van. Gyakran vannak más nyelven készen levő jó szócikkek is, amelyeket könnyen át lehet emelni a magyarba, és lefordítani.

A Wikimedia Alapítvány a fenntartója a Wikipédiának és hasonló elven működő közösségi oldalaknak, mint a Wikibooks, a Wikiquote és a Wikispecies. Külön említést kíván ezek között a Wikimedia Commons oldal. A különböző nyelvű Wikipédiák és tárprojektek a képeket – ha csak lehet – egy közös tárolóban helyezik el, a Wikimedia Commons oldalon, ahonnan könnyedén be lehet illeszteni bármelyik nyelvű Wikipédia oldalára, és ahol meg lehet adni a kép fontosabb adatait és licencfeltételeit. Komoly megkötést jelent, hogy csak szabad licencű képek tölt-

hetők fel ide. Néhány esetben az angol oldalakon jogszerűen megjelenő képek nem tehetők fel más országok jogszabályai szerint a Wikipédiára, ezek a képek nem kerülhetnek fel a Wikimedia Commons képtárába. Ábráknál bátorítják az SVG nevű vektorgrafikus fájlformátum használatát. A vektorgrafikus fájlformátum jelentősen megkönnyíti a felirattal rendelkező fájlok fordítóinak életét, hiszen a vektorgrafikus formátumban a kép nem pixelenként tárolódik, hanem objektumok (kör, Bezier-görbék) álló görbe vonalak, szöveg) tárolódnak tulajdonságaikkal, és azok az objektumok bármikor könnyen törölhetők, vagy módosítható a vastagságuk, színük és egyéb jellemzőik. Ezt használtam ki például akkor, amikor a fotoelektron-sokszorozó, a proton-proton ciklus és a CNO-ciklus szócikkeinek ábráit lefordítottam.

Sajnos vannak bizonytalan esetek is. Korábban bizonyosnak tűnt, hogy a CMS által terjesztett képek szabadon felhasználhatóak, jelenleg viszont tisztázatlan, de reményt keltő a helyzet. 2010-ben David Barney, aki a CMS-együttműködést képviseli az IPPOG-ban [23], azt nyilatkozta, hogy a licencfeltételek felülvizsgálat alatt vannak. Levélben érdeklődtem tőle a jelenlegi helyzetről, amire Achintya Rao-tól, a CMS Communications Group tagjától, azt a választ kaptam, hogy nemrégiben a Wikimédia Alapítvány képviselőivel folytattak tárgyalásokat a CERN-ben, amelynek következtében egy munkacsoport alakult, amely összehangolva az érintettek igényeit várhatóan hamarosan új képicenc-politikát javasolnak. Addig is azt javasolja a Wikipédia felhasználóinak, hogy ne töröljünk semmilyen CMS-sel kapcsolatos képet a Wikipédiáról.

Sajnos licencfeltételek miatt lehet, hogy el kell távolítani azt az ábrát is, amely a Sudbury Neutrino Observatory felépítését mutatja, illetve el kellett távolítani azt, amelyik a Super-Kamiokande neutrínós eseményeit mutatja be.

Ugyan a szabad szerkeszthetőség miatt előfordulhat, hogy az érdemleges tartalmat valami teljesen mással cseréli fel valaki, azt azonban a Wikipédia lelkes járőrei hamar visszaállítják. A magyar Fizika portállal egyszer történt meg, hogy egy nem bejelentkezett szerkesztő teljesen lecserélte a tartalmát, de a tartalom még ugyanabban a percben visszaállításra került.

Az is előfordul időnként, hogy valaki rá akarta erőltetni a saját elméleteit a Wikipédiára, ilyenkor bizonyos IP-címeket letilthatnak az adminisztrátorok. Ez inkább filozófiai és vallási cikkeknel volt jellemző.

A Wikipédia szócikkeinek jobb felső sarkában van egy felirat, amelynek szövege vagy „Ellenőrzött lap”, vagy „ellenőrzésre váró változások”. Az utóbbi esetben viszonylag könnyen át lehet váltani az utolsó ellenőrzött változatra, és akár azt is megvizsgálni, mi az, ami azóta változott. Ez a megoldás a szócikkek minőség-

gét hivatott növelni. A szerkesztők kérhetik, hogy megkapják azt a jogot, hogy az általuk szerkesztett ellenőrzött lapok a szerkesztésük után is ellenőrzöttek maradjanak. Kevesebb szerkesztőnek van joga arra, hogy egy nem ellenőrzött lapot ellenőrzöttre állítson át.

A magyar Wikipédián jelentősen mértékben hozzájárultam a fizikával, főként részecskefizikával és annak kutatásával kapcsolatos cikkekhez. A magyar Wikipédia 2003-ban indult útjának, tevékenységem főként a magyar Wikipédia korai szakaszára – 2004 és 2010 közé – esik, ezért lehetőségem volt sok olyan részecskefizikával kapcsolatos szócikk létrehozására, amely a részecskefizika legalapvetőbb fogalmait dogozza fel, valamint lényegében egyedül hoztam létre a Fizika portált is a 2005-ös évben. Szerepem volt több szócikk-kategória kialakításában is, és néhány ábrával illetve ábrafordítással is hozzájárultam a részecskefizikai és összetett hálózatokkal kapcsolatos cikkeihez. A Wikipédia szócikkeinek száma azóta jelentősen megnövekedett, 2005 június környékén érte el a tízezret, 2011 második felében pedig a kétszázezret (4.1), tehát hat év alatt húszszorosára nőtt. Azok között a cikkek között, amelyekhez jelentősen hozzájárultam, szerepelnek

1. részecske- és magfizikai kutatóintézetek (CERN, DESY, ATOMKI, Brookhaveni Nemzeti Laboratórium, Super-Kamiokande, Sudbury Neutrínó Observatóriumi, Fermilab),
2. ezek gyorsítói és detektorai (LHC, CMS, ATLAS),
3. a gyorsítókkal, detektorokkal kapcsolatos alapfogalmak (részecskegyorsító, ciklotron, lineáris gyorsító, Cserenkov-effektus, szupravezetés, fotoelektron-sokszorozó),
4. a részecskefizika elméleti háttere (részecskefizika, alapvető kölcsönhatások, elektromos töltés, spin, Planck-állandó, kétneutrínó-kísérlet, elektronvolt, Standard modell, A mikrofizika története évszámokban)
5. a részecskék (proton, neutron, neutrínó, elektron, . . . , részecskefelfedezések évszámokban),
6. a részecskefizika kutatói (Szalay Sándor, Leon Lederman, Carlo Rubbia Wolfgang Pauli, Paul Dirac, Lise Meitner, Martinus Veltman, Albert Einstein, Wigner Jenő)
7. és a kutatókhoz kapcsolódó egyetemek (Princeton Egyetem, Chicagói Egyetem, Cambridge-i Egyetem, Heidelbergi Egyetem, Debreceni Egyetem).

8. Készítettem továbbá cikkeket a fizika tágabb területéről is (fizikai Nobel-díj, relativitáselmélet, speciális relativitáselmélet, általános relativitáselmélet, mikrohullámú kozmikus háttérsugárzás). A közreműködésem bővebb listája megtekinthető a Wikipédián [25].

4.1. táblázat. Néhány Wikipédia-cikk látogatottsága a 2013. január 19-ig tartó 90 napos időszakban a <http://stats.grok.se/> oldal szerint Az egy sorban található címek ugyanarra a (legelől szereplő) szócikkre mutatnak. A *-gal jelölt kidolgozottság esetén a szócikk vitalapján nincs megadva érték, saját magam soroltam be.

szócikk	látogatottság /90 nap	közreműködésem	kidolgozottság
Eötvös Loránd	20188	nincs	jól használható (4)*
Jedlik Ányos	10832	apró	kitüntetett (7)
Higgs-bozon	7805	közepes	vázlatos (2)
felületi feszültség	5745	nincs	kitüntetett (7)
Johannes Kepler, Kepler János	5158+39	jelentős	jól használható (4)
CERN	3010	jelentős	jól használható (4)
Portál:Fizika	4534	jelentős	-
fizikai Nobel-díj	4140	jelentős	bővítendő (3)
proton	2973	jelentős	bővítendő (3)
Simonyi Károly	2967	jelentős	bővítendő (3)
Nagy Hadronütköztető, LHC	2839+448	nagyon jelentős	bővítendő (3)
részecskefizika, nagy-energiájú fizika	2561+126	jelentős	-
alapvető kölcsönhatások, egyes számban	2331+304	jelentős	bővítendő (3)
neutrínó	2223	jelentős	bővítendő (3)
kozmikus mikrohullámú háttérsugárzás, m. k. háttérsugárzás	1427+373	jelentős	jól használható (4)
Cserenkov-effektus, Cserenkov-sugárzás	991+157	jelentős	vázlatos (2)
Marx György	976	nincs	vázlatos (2)*
MTA Atommagkutató Intézete, ATOMKI	304+284	közepes	-

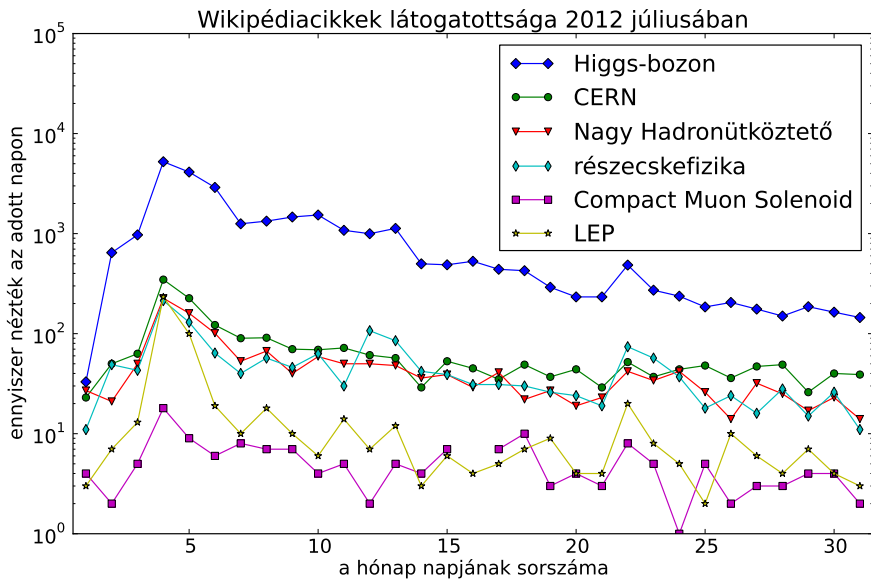
szócikk	látogatottság /90 nap	közreműködésem	kidolgozottság
Compact Muon Solenoid, CMS	373+109	jelentős	bővítendő (3)
Super-Kamiokande, Kamiokande	370+6	jelentős	jól használható (4)
Vicsek Tamás	338	nincs	jól használható (4)*
Leon Lederman	291	jelentős	vázlatos (2)
Eiffel-torony	18569	nincs	vázlatos (2)
Országház	17044	nincs	bővítendő (3)
Libegő	5524	nincs	jól használható (4)
Johann Sebastian Bach, Bach	12612+564	nincs	kitüntetett (7)
Tamási Áron	7726	nincs	vázlatos (2)
Fekete István	17545	jelentős	teljes (5)*
A varázsfuvola	6738	jelentős	-
Kányádi Sándor	6281	nincs	bővítendő (3)
Miklósa Erika	5462	nincs	-
Lackfi János	2189	nincs	bővítendő (3)
Lovász László	1794	apró	vázlatos (2)*
Orbán Viktor	30805	nincs	teljes (5)*
Horn Gyula	7695	nincs	bővítendő (3)*
Harry Potter	70947	nincs	színvonalas (6)
Debreceni Egyetem	3806	jelentős	jól használható (4)
Óbudai Egyetem	2757	apró	-
Princetoni Egyetem, Princeton Egyetem	1054+580	jelentős	bővítendő (3)
Budapest (1700000 fő)	61935	nincs	teljes (5)
Debrecen (208000 fő)	28121	kicsi	teljes (5)
Székesfehérvár (102000 fő)	15452	jelentős	jól használható (4)
Mór (14300 fő)	2835	nincs	születőben
Csákberény (1210 fő)	611	nincs	vázlatos (2)



4.1. ábra. A szócikkek számának alakulása a magyar Wikipédiában

A 4.1. táblázatban többnyire azokból a cikkekből válogattam, amelyekhez jelentősebb mértékben hozzájárultam. Összehasonlításképpen azonban feltüntettem néhány szócikket a politika, a kultúra és az egyetemek témakörében. A Higgs-bozon és pár kapcsolódó szócikk látogatottsági statisztikáiból egyértelműen kitalálható a bejelentés időpontja, ahogy az a júliusi eleji napi látogatottságokat tartalmazó 4.2. táblázatból és annak adatait grafikusán ábrázoló 4.2. ábrán is látható. A Higgs-bozon a CERN és Nagy hadronütköztető cikkekhez közel azonos látogatottsággal rendelkezett 1-jén. 2-ára mindegyik látogatottsága megnövekedett, ami vélhetően a Nature-beli hír hatásának köszönhető [26], de innentől a Higgs-bozon sokáig legalább egy nagyságrenddel megelőzi a másik két szócikket. Az angol változatokban hasonló jellegű lefutások találhatók, természetesen sokkal nagyobb látogatottsági értékekkel.

Az itt bemutatott adatokból jól látható, hogy a Wikipédia látogatottsága nagy, és egyes, a témával kapcsolatos érdekes bejelentések hatására nagyságrendekkel változhat. A Wikipédia tehát remek lehetőségekkel rendelkezik az ismeretterjesztés területén, ebből fakadóan viszont figyelemmel kell kísérni a szócikkeiben található tartalmakat, hiszen érthetetlen vagy téves fogalmazás esetén rossz hatással is lehet.



4.2. ábra. A Wikipédia három cikkének látogatottsága a 2012 júliusában. A függőleges skála logaritmikus

4.2. táblázat. Pár magyarországi Wikipédia-szócikk látogatottsága 2012 júliusának első napjaiban.

Nap	CERN	Higgs-boson	Nagy hadronütköztető
1	23	33	27
2	53	645	21
3	63	971	50
4	347	5232	226
5	226	4124	160
6	122	2896	101
7	90	1254	53

II. rész

Összetett hálózatok

5. fejezet

Bevezetés és alapfogalmak

Szemekkel az egész tengert halászod
s horoggal mit fogsz? Egynéhány halat.

Weöres Sándor

A legelemibb részecskék tulajdonságainak vizsgálata mellett egy ellentétes irányú kérdés is foglalkoztatja a fizikusokat. Hogyan lehet leírni azokat a rendszereket, amelyek nagyon összetettek? Hogyan tudjuk az egynéhány hal kifogása helyett az egész tengert halászni? Egyáltalán vannak-e közös törvényszerűségek ilyen esetben, azaz lehetőség van-e az egyes részek vizsgálata nélkül az egész rendszert jellemezni úgy, hogy a kapott jellemzőkkel meg tudjuk jósolni, hogyan fog változni a rendszer, vagy mi történik, ha az összetett rendszert külső hatás éri? Az esetek egy részében az összetett rendszer egy hálózattal írható le. Továbbiakban kizárólag ezekkel a rendszerekkel, az összetett hálózatokkal foglalkozom.

A legkorábbi tudományos hálózati vizsgálatok a szociológia területén történtek, ahol az emberek közötti ismeretségek hálózatát vizsgálják. Műszaki tudományok területén is számos hálózat található, mint például a számítógépekből, útvonalválasztókból (router), fémes vezetékekből és optikai szálakból álló Internet, valamint az Internetre épülő információmegosztó alkalmazás, a Világháló (World Wide Web). Az Internet és a Világháló tanulmányozása új lendületet adott a hálózatok vizsgálatának a századforduló éveiben. A biológiában és az orvostudományban a fehérjék kölcsönhatásának hálózata, a tápláléklánc, vagy a fertőzés terjedésének szempontjából az ismeretségi, közlekedési és – nemi betegségek esetén – a szexu-

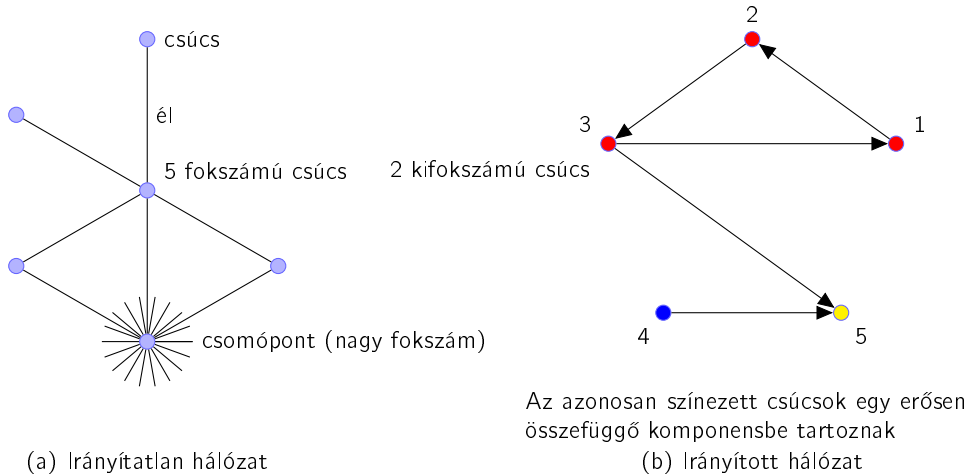
5.1. táblázat. Valódi hálózatok

hálózat	csúcsok	él van ha...	típus
ismeretségi hálózat	emberek	találkoztak	irányítatlan
Világháló	weboldalak	van köztük link	irányított
Internet	routerek	van vezeték közöttük	irányítatlan
cikkek hálózata	cikkek	hivatkozik a másokra	irányított
fehérjehálózat	fehérjék	közös kölcsönhatásban részt vesznek	irányítatlan
szavak hálózata	szavak	ha szerepelnek együtt a szinonimaszótárban	irányítatlan
matematikushálózat	matematikusok	írtak közös cikket	irányítatlan
színészek hálózata	színészek	szerepeltek közös filmben	irányítatlan

ális hálózat szerepe jelentős. Pár példa található hálózatokra az 5.1. táblázatban.

A hálózatok – mint láthatjuk – valamilyen egyedekből állnak, melyek közül bizonyos egyedpárok között kapcsolat van valamilyen szempont szerint, más párok között nincsen. Az egyedeket csúcsoknak vagy vertexeknek nevezzük, a kapcsolatokat éleknek. A hálózatokat gyakran a gráfok szinonimájaként használják a szociológia, a fizika, a biológia és több más tudomány területén. Egy hálózatban az idő során új csúcsok és élek jelenhetnek meg, mások eltűnhetnek, ezért pontosabb az a felfogás, amely szerint a hálózatok gráfok időben változó sorozataként írhatóak le. A hálózatokra vonatkozó definíciókat (fokszám, átmérő, összefüggőség, párosság) gyakran a gráfelméletből veszik át, emiatt például, ha valaki az összefüggő vagy páros hálózatokról szeretne többet tudni, érdemes az összefüggő gráf, páros gráf címszavak után is keresnie.

A hálózatok pár alapfogalma látható az 5.1. ábrán. Egy *csúcs fokszámán* a hozzá kapcsolódó élvégpontok számát értjük, tehát egy *hurokél* – amely ugyanabban az élben végződik, mint amiből indul – kettővel növeli a csúcs fokszámát. Az *irányított hálózatokban* az élek két végpontja nem azonos szerepű a kapcsolat szempontjából. Például a Weboldalakon lévő linkek szempontjából egyáltalán nem mindegy, hogy melyik oldal hivatkozik melyikre. Az irányított élek egyik végén található csúcsot forrásnak, a másikat célnak nevezzük, és ábrázoláskor az él cél felé mutató végére nyilat rajzolunk. Irányított hálózatoknál a szokásos fokszám mellett kifokszámot és befokszámot is definiálhatunk: ebben az esetben csak az



5.1. ábra. A hálózatok pár alapfogalma

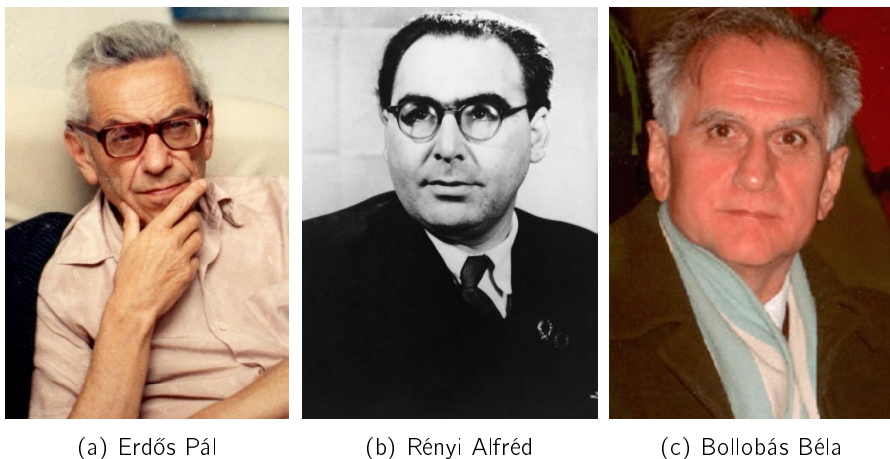
élek forrásnál, illetve célnál található végpontjait számoljuk.

A már említett hálózatok esetén azt tapasztalták, hogy vannak olyan csúcsok, amelynek nagy (nagyságrendekkel nagyobb) a fokszáma az átlagos fokszámhoz képest. Ezeket szokták csomópontnak nevezni. Mint látjuk, ez nem pontosan definiált fogalom, mégis jelentősége van. Ha teljesen véletlenszerűen húznánk meg az éleket a hálózatban, nem lennének csomópontok.

Két hálózatmodellt szokás a véletlen hálózatok modelljének nevezni. Mindkét modellben az egyik paraméter a csúcsok N száma. A $G(N, p)$ -vel jelölt modellben a második paraméter az élvalószínűség, amely megadja, hogy két tetszőleges csúcs között milyen valószínűséggel vezet él. Eszerint a modell szerint úgy generálhatunk véletlen hálózatot, hogy végigmegyünk az összes lehetséges csúcspáron, és kisorsoljuk, hogy lesz-e él közöttük. A csúcsok között p valószínűséggel húzunk élt, $1 - p$ valószínűséggel nem. Az élek számának várható értéke ebben a modellben:

$$M = p \binom{N}{2}. \quad (5.1)$$

A másik $G(N, M)$ -mel jelölt modellben az élek M száma a második paraméter. Ebben a modellben M -szer sorsolunk ki élvégpontokat egy-egy élnek. (Amennyiben már van olyan él, akkor újból sorsolunk.) Erdős Pál és Rényi Alfréd (5.2.



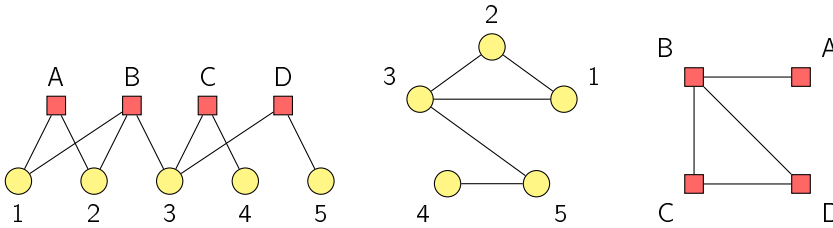
5.2. ábra. A véletlen hálózatok területének három magyar matematikusa

ábra) mindkét modellt vizsgálták, mégis az előbbi modellt szokás Erdős–Rényi-modellnek hívni. Mindkét modellben a csúcsok fokszámai az átlagos $2M/N$ érték közelében találhatók; nincsenek benne sem sokkal kisebb fokszámú csúcsok, sem csomópontok. A fokszámok eloszlása ezekben a modellekben Poisson-eloszlást követ.

A legtöbb hálózat tulajdonságai meglepő hasonlóságot mutatnak a fokszám-eloszlásuk szempontjából, annak ellenére, hogy különböző szakterületről származnak, hogy van közöttük a természetben létrejött és az ember által létrehozott hálózat, illetve, hogy az 5.1 utolsó három példája egy összetettebb, úgynevezett páros gráf projekciójának tekinthető. A páros gráf alatt olyan hálózatot értünk, amelynek kétféle típusú csúcsa van [27]. Az utolsó példát tekintve csak a színészeket tüntetjük fel a hálózatban, pedig a hálózat felépítésekor kezdetben azt lehet tudni, hogy melyik szereplők melyik filmen játszottak. A teljes információt ilyen esetekben egy páros gráf hordozza.

1. definíció (páros gráf). *Egy gráfot páros gráfnak nevezünk, ha a csúcsai feloszthatók két diszjunkt halmazba úgy, hogy minden él csak két különböző halmazbeli csúcs között halad.*

Az 5.3. ábra baloldali részábrája páros gráf. Az egyik halmazba tartoznak a négyzettel jelölt csúcsok, a másikba a körrel jelöltek.



5.3. ábra. Egy páros gráf és kétféle projekciója

A fenti példában a teljes információt egy olyan páros gráf hordozza, amelynél az egyik halmazba a filmek tartoznak, a másik halmazba a színészek, és össze van kötve egy film a színésszel, ha a színész szerepel a filmben. Minden páros gráfnál kétféle irányú projekció lehetséges. A már említett színészek hálózata mellett meg lehet alkotni a fenti páros gráfból a filmek hálózatát is, ahol azok a filmek vannak összekötve, amelyekben ugyanaz a színész szerepelt. Az 5.3. ábrán egy ilyen páros gráfot, és annak kétféle projekcióját láthatjuk. Tekinthejük úgy, mintha a szögletes csúcsok a filmek, a kerek csúcsok a színészek lennének a fent elmondott példában.

A szociológia területén folytattak először hálózatokkal kapcsolatos tudományos vizsgálatokat az ismeretségi hálózatok tulajdonságainak vizsgálatával, de ennek a hálózatnak a feltérképezése nagy méretű csoport (város, ország) esetén mindmáig nehéz feladat. Milgram amerikai szociológusnak az 1960-as években levéltovábbításon alapuló módszerrel viszont sikerült megbecsülnie az Egyesült Államok lakói közötti átlagos távolságot, és meglepően kis értéket, 5,5 körüli értéket kapott [28].

2. definíció (két csúcs távolsága). Egy hálózatban két csúcs távolságának azt a legkisebb egész számot nevezzük, ahány élen keresztül el lehet jutni az egyik csúcsból a másikba. (A legkisebb kitétel amiatt kell, mert lehet, hogy többféle útvonalon különböző számú élen keresztül lehet eljutni.) Ha két csúcs között nincs út, akkor végtelennek vesszük a távolságát.

A továbbiakhoz szükség lesz egy alapvető fogalomra, a fokszámoszlásra.

3. definíció (fokszámoszlás). A $p(k)$ fokszámoszlás egy olyan függvény, amely az egyes k fokszámokhoz hozzárendeli annak a valószínűségét, hogy egy véletlenszerűen kiválasztott csúcs k fokszámú, azaz

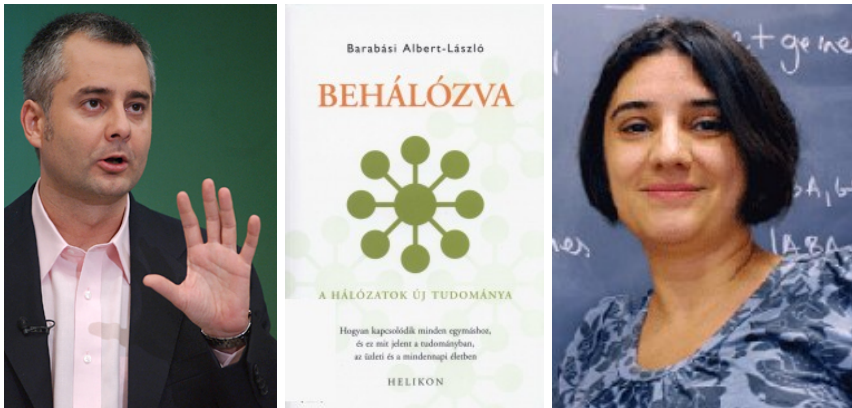
$$p(k) = \text{Prob}(\text{véletlen csúcs fokszáma} = k)$$

4. definíció (összegzett fokszámeloszlás). A $P(k)$ összegzett fokszámeloszlás egy olyan függvény, amely az egyes k fokszámokhoz hozzárendeli annak a valószínűségét, hogy egy véletlenszerűen kiválasztott csúcs fokszáma nagyobb vagy egyenlő mint k , azaz

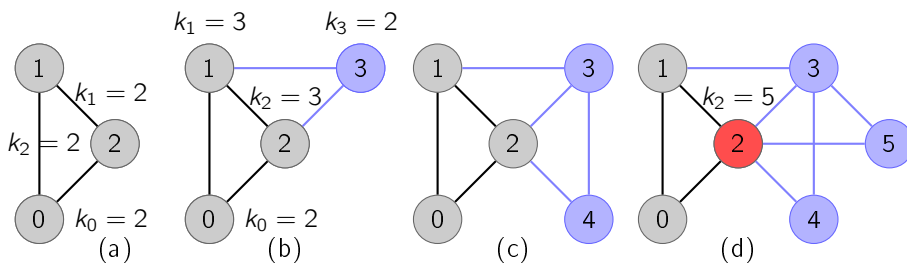
$$P(k) = \text{Prob}(\text{véletlen csúcs fokszáma} \geq k)$$

A legkorábbi hálózat, ahol jelentős mennyiségű adatot sikerült összegyűjteni a hálózat konkrét felépítéséről, a szakcikkek hivatkozási hálózata volt, amelyben a cikkek a csúcscok, és egyik cikktől a másikhoz él mutat, ha az előbbi hivatkozik az utóbbira. Ennél a hálózatnál Derek J. de Solla Price azt találta, hogy a fokszámeloszlás hatványfüggvényt követ és kidolgozott egy modellt, amelyben a fokszámeloszlás okát a ma népszerűségi kapcsolódásnak nevezett jelenségben találta meg [29, 30]. A népszerűségi kapcsolódás azt jelenti, hogy a hálózathoz kapcsolódó új csúcscok nem egyforma valószínűséggel kapcsolódnak a korábbiak mindegyikéhez, hanem azokhoz nagyobb valószínűséggel, amelyekhez már addig többen kapcsolódtak, azaz amelyek befokszáma nagyobb. A hatványfüggvény eloszláshoz egyenes arányosság szükséges a kapcsolódási valószínűség, és a befokszám között.

Barabási Albert-László és Albert Réka (5.4. ábra) 1999-ben a Világháló kis részét mint hálózatot tanulmányozta. Az adatokat kollégájuk, Hawoong Jeong programja gyűjtötte össze. Ebben szintén a hatványfüggvény eloszlást tapasztalták, és több más, akkor elérhető hálózatban is. Nem ismerték Price modelljét, de ahhoz hasonló modellt dolgoztak ki. A Barabási–Albert modell egy tetszőleges kis hálózatról indul ki, amelyhez minden egyes lépésben új csúcs jön létre, amely előre megadott állandó m számú éllel kapcsolódik m már meglévő csúcshoz úgy, hogy az egyes csúcscokhoz a fokszámukkal arányos a kapcsolódás valószínűsége. Az 5.5. ábrán egy olyan példa látható, amelynél a Barabási–Albert modell szerint fejlődik a hálózat, $m = 2$ érték mellett. Olyan hálózatról indulunk ki, amelyben minden csúcs fokszáma egyforma ((a) ábra), ezért ilyenkor a következő csúcs egyforma valószínűséggel kapcsolódik bármelyik kettő már meglévő csúcshoz. Az első lépés után, a (b) ábrán már nem lesznek egyformák a fokszámok. A lehetséges párok közül a legnagyobb valószínűséggel az 1-es és 2-es csúcshoz csatlakozik, de mivel csak kicsit kisebb a valószínűsége, hogy egy olyan, előre kiválasztott csúcspárhoz csatlakozzon, amelyek egyikének a fokszáma 3, a másiké 2, és ilyen csúcspárból



5.4. ábra. Barabási Albert-László, a Behálózva című könyve és Albert Réka



5.5. ábra. A hálózat fejlődése a Barabási–Albert modell szerint. Magyarázat a szövegben.

négy darab van, összességében nagyobb a valószínűsége, hogy a következő lépésben ilyen párhoz fog csatlakozni az új csúcs. Amelyik csúcs kezdetben több új kapcsolatot szerez, annak a fokszáma várhatóan továbbra is jelentősen fog növekedni, és emiatt alakulnak ki ebben a modellben a csomópontok. Ebben a modellben a korán jövőeknek előnyük van, nagyon kicsi a valószínűsége, hogy egy későn jövő csúcsból legyen csomópont.

Barabási Albert-Lászlótól és Albert Rékától ered az azóta elterjedt skálafüggetlen hálózat kifejezés. A skálafüggetlen hálózatok összegzett fokszámeloszlása is hatványfüggvényt követ abszolút értékekben eggyel kisebb kitevővel. Az összegzett fokszámeloszlás azért nagyon hasznos, mert úgy szabadul meg a nagy fokszámoknál látható statisztikai ingadozásoktól, hogy visszaállítható belőle az eredeti fokszámeloszlás. Mivel a skálafüggetlen hálózatoknak mind a fokszámeloszlása, mind az összegzett fokszámeloszlása hatványfüggvény, a skálafüggetlenség jól észrevehető, ha mindkét tengelyen logaritmikus skálát használva ábrázoljuk azokat.

Néhány valódi hálózat összegzett fokszámeloszlása a 5.6. ábrán látható. Az (e) kivételével minden ábrán mindkét tengelyen logaritmikus skála található. Ezt kivéve az összes hálózat skálafüggetlen hálózatnak tekinthető: a fokszámok legnagyobb két nagyságrendjében az eloszlás grafikonja közel egyenes.

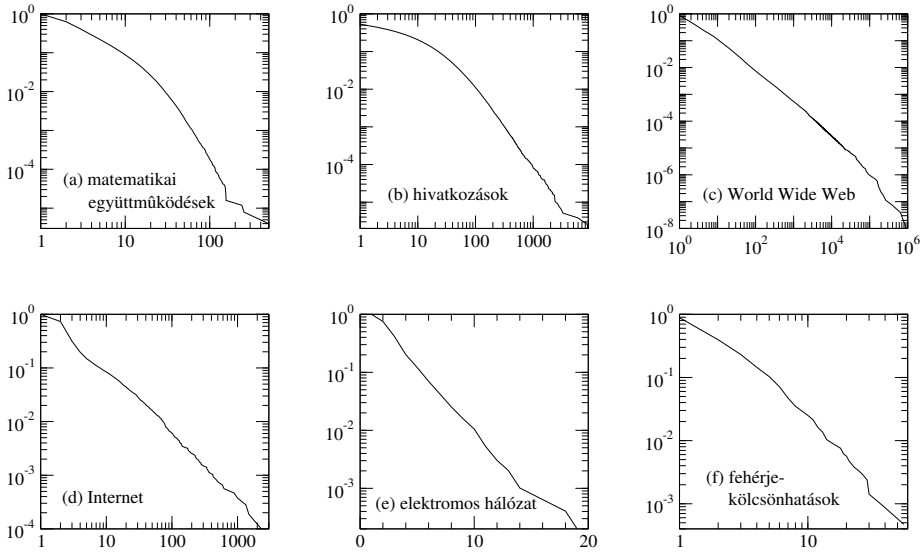
5. definíció (skálafüggetlen hálózat). *Skálafüggetlen hálózatoknak nevezzük azokat a hálózatokat, melyeknek a fokszámeloszlása hatványfüggvényt követ nagy fokszámok esetén:*

$$p(k) \sim k^{-\gamma} \quad (5.2)$$

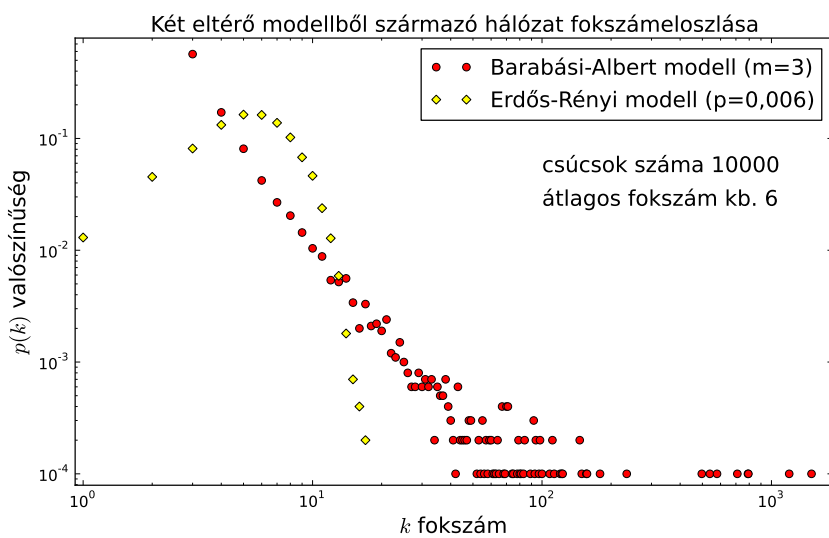
$$P(k) \sim k^{-(\gamma-1)} \quad (5.3)$$

Az 5.7 ábrán látható két fokszámgyakoriság az Erdős–Rényi modellből, illetve a Barabási–Albert modellből származó hálózatok fokszámgyakorisága. Mindkettőben 10000 a csúcsok száma és az átlagos fokszám nagyon jó közelítéssel 6, azaz az élek száma 30000 körüli. Jól látható, hogy a Barabási–Albert modellből származó fokszámeloszlásra illik a skálafüggetlen hálózat kifejezés, hiszen valóban nincsen jellemző fokszám: minél kisebb a fokszám értéke, annál több található belőle, és az átlagérték egyáltalán nem jellemző érték, szemben az Erdős–Rényi-modellből származó hálózatokkal.

A függőleges tengelyen valószínűség helyett darabszámot tüntettem fel, azaz fokszámeloszlás helyett fokszámgyakoriságot ábrázoltam, hogy a nagy értékeknél történő statisztikai ingadozás jobban érthető legyen. Adott fokszámú csúcsból



5.6. ábra. Hat hálózat összegzett fokszámeloszlása. A vízszintes tengelyeken a fokszám található (illetve a befokszám az irányított hivatkozási hálózat és a Világháló esetén). A függőleges tengelyen annak a valószínűsége található, hogy a hálózat egy véletlenszerűen választott csúcsának fokszáma nagyobb az adott k fokszámnál vagy egyenlő vele. Az egymásutáni gráfokon az alábbi hálózatok összegzett fokszámeloszlásai találhatóak: a matematikusok együttműködésének hálózata (akkor van él közöttük, ha írtak közös cikket), a cikkek közötti hivatkozások hálózata, a Világháló 300 milliós részhalmaza, az Internet a routerek szintjén, az USA nyugati részének elektromos hálózata, az élesztőben a fehérjék kölcsönhatása. Az ábra Newman cikkéből való, a pontos hivatkozások ott megtekinthetők. [31]



5.7. ábra. A Barabási–Albert-modellből és az Erdős–Rényi-modellből származó hálózatok foksámgyakoriságairól a cxnet-tel készített ábra

nyilván csak egész számú lehet. A nagy fokszámok esetén rengeteg olyan fokszám van, amekkora fokszámú csúcs nincs a hálózatban, és néhány olyan fokszám, amelyhez egy csúcs (egy csomópont) tartozik. A sima fokszámgyakoriság, illetve fokszámeloszlás ábrájából nehéz eldönteni, hogy nagy fokszámokra hatványfüggvényként viselkedik-e. Egyik lehetőség, hogy az összegzett fokszámeloszlást ábrázoljuk, ahogy a 6.10 ábrán láthatjuk. A másik lehetőség pedig a binelés, ami azt jelenti, hogy a fokszámok eloszlát úgy ábrázoljuk, hogy a fokszámok tengelyt szakaszokra osztjuk, és az egyes szakaszokhoz tartozó fokszámokhoz egyetlen valószínűséget rendelünk. Ez utóbbi majd a cxnet szoftvercsomag tárgyalásánál kerül bővebb kifejtésre.

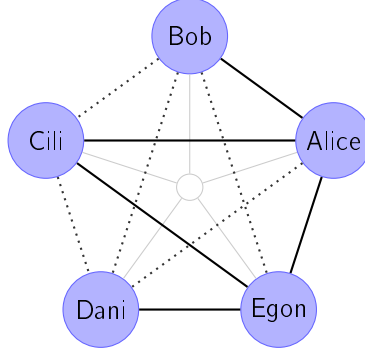
A hálózatok esetén fontos tulajdonság, hogy minden csúcsból el lehet-e érni minden másikat éleken keresztül, és ha nem, mekkora részekből áll. A következő fogalmak ezzel kapcsolatosak.

Az irányítatlan hálózat mindig felbontható részhálózatokra úgy, hogy a részhálózatok bármely pontjából el lehessen jutni bármely másik pontjába éleken keresztül. Ezeket a részhálózatokat (összefüggő) *komponenseknek* nevezzük. Irányított hálózatok esetén hasonlóan definiálhatóak a komponensek, azonban kétféle módon is. Amennyiben az összekötő útvonalnál nem vesszük figyelembe az élek irányát, *gyengén összefüggő komponensekről* beszélünk. Amennyiben fontos, hogy bármely csúcsból bármely csúcsba eljuthassunk az éleknek megfelelő irányba lépkedve az éleken, akkor *erősen összefüggő komponensekről* beszélünk. Egy irányítatlan hálózatot *összefüggőnek* nevezünk, ha egyetlen komponensből áll.

Az 5.1. ábrán szereplő irányított hálózat például egyetlen gyengén összefüggő komponensből áll, de három erősen összefüggőből. Az azonos színűek tartoznak egy erősen összefüggő komponenshez.

A valódi hálózatban általában nagy a csoportosodás, amely ismeretségi hálózatok esetén például azt jelenti, hogy az ismerőseim nagyobb valószínűséggel ismerik egymást, mint két véletlenszerűen kiválasztott személy a világon.

Ennek mérésére vezették be a csoportterősségi együtthatót, mely azt mutatja meg, hogy a szomszédok közötti lehetséges kapcsolatoknak mekkora hányada létezik a valóságban. Egy csúcs csoportterősségi együtthatója 0 és 1 közötti szám lehet, beleértve a határokat is. 0, ha a szomszédai között nincs egy él sem, és egy, ha az összes él létezik. Egy további példát láthatunk az 5.8. ábrán. Ezt az együtthatót irányítatlan hálózatok esetén szokták értelmezni. Irányított hálózat esetén úgy szokás csoportterősségi együtthatót számolni, hogy előbb irányítatlanul alakítják a hálózatot. Ismeretes, hogy az i -dik csúcs k_i szomszédja között



5.8. ábra. Egy példa egy csúcs csoportterösségi együttthatójára. Az eredeti csúcsot a középben lévő halvány kör jelképezi, a szomszédait a kék körök. A kék körök között összesen 10 csúcs lehetne, de csak 5 van (folytonos vastag vonalak), ezért a csúcs csoportterösségi együttthatója $5/10=0,5$.

maximálisan $\binom{k_i}{2} = \frac{k_i(k_i - 1)}{2}$ él futhat, ha részgráfként teljes gráfot alkotnak.

6. definíció (csúcs csoportterösségi együttthatója). Egy csúcs C_i csoportterösségi együttthatója

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \quad (k_i \geq 2) \quad (5.4)$$

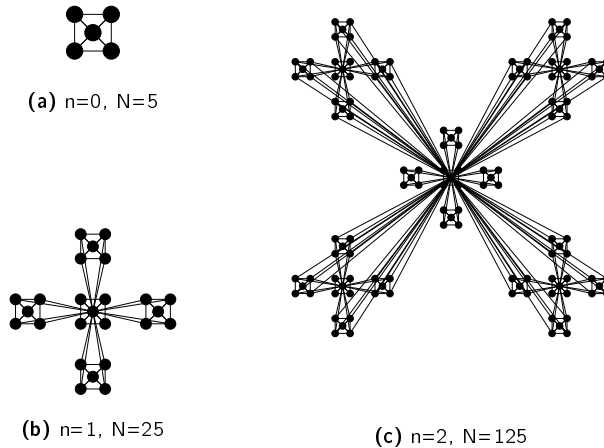
ahol E_i a csúcs szomszédjai közötti élek tényleges száma.

7. definíció (hálózat csoportterösségi együttthatója). A hálózat C csoportterösségi együttthatóját a csúcsok csoportterösségi együttthatóinak számtani közepeként értelmezzük:

$$C = \frac{1}{N_{[2,\infty[}} \cdot \sum_{\substack{i \in [1, N] \cap \mathbb{Z} \\ k_i \geq 2}} C_i \quad (5.5)$$

ahol $N_{[2,\infty[}$ a kettő vagy nagyobb fokszámú csúcsok számát jelenti.

A valódi hálózatok esetén általában a véletlen hálózatokénál jóval nagyobb csoportterösségi együttthatót találunk. Sok hálózatban a csoportterösségi együtttható a fokszám függvényében csökken, közel a fokszámmal fordított arányban. [32] Ezeket *hierarchikus hálózatoknak* hívjuk. Más esetekben a csoportterösségi együtttható

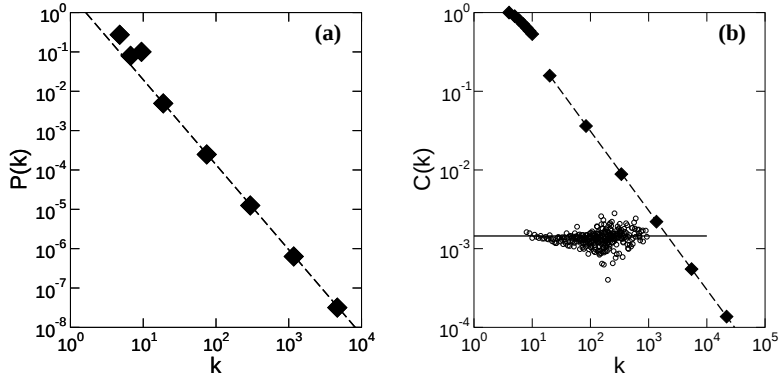


5.9. ábra. A hierarchikus modell pár lépése. [32]

független a fokszámtól. Ez utóbbi olyan hálózatokra jellemző, amelyeknél a földrajzi elhelyezkedés jelentősen befolyásolja a kapcsolat költségét. Például amíg a Világháló esetén nincs jelentősége, hogy másik kontinensen lévő oldalra hivatkozunk, addig az Internet esetén az óceán két felén fekvő két router összekötése óriási költség lenne.

Milyen képet képzelünk a hierarchikus hálózat mögé? Sok területen hierarchikus szerkezeten egy fa-szerkezetet értenek. Vegyük észre, hogy egy fa esetén minden csúcs csoportterősségi együtthatója nulla, hiszen ha egy csúcs két szomszédja össze lenne kötve, akkor kör lenne a hálózatban, ami ellentétben áll a fa definíciójával. A hierarchikus jelleget Barabási Albert-Lászlónak Ravasz Erzsébettel és Vicsek Tamással együtt kidolgozott modellje [32], a hierarchikus modell teszi elképzelhetővé, melyet a 5.9. ábrán láthatunk. A nulladik lépésben a hálózat egy öt csúcsú teljes gráfból áll. Minden lépésben négy másolatot készítünk a hálózatról, az eredetit középen tartva, és a középső (eredeti) hálózat középső csúcsait a négy másolat „szélső csúcsaival” összekötjük.

A hierarchikus modell tulajdonságait a 5.10. ábrán láthatjuk. Mint látjuk, a hierarchikus modell is skálafüggetlen hálózatot hoz létre, mint a Barabási–Albert



5.10. ábra. A hierarchikus modell tulajdonságai $n = 5^7$ esetén. A hierarchikus modell (a) is skálafüggetlen hálózatot hoz létre és (b) a csoportterösségi együtthatója is közel a fokszámmal fordítottan arányos. A (b) ábrán a Barabási–Albert modell adatai üres körrel szerepelnek. [32]

modell, de emellett teljesíti azt a feltételt is, hogy a csúcsok csoportterösségi együtthatói a fokszámmal közel fordítottan arányosak.

Az összetett hálózatok tudománya a valódi hálózatok vizsgálatát tűzte ki célul, amelyben a csúcsok többnyire fizikailag létező dolgok (tárgyak, élőlények, molekulák), és amelyek emiatt kísérletileg tanulmányozhatóak. A fizika szempontjából a cél olyan modellek kidolgozása, amelyekkel az összetett hálózatok változásainak törvényszerűségei leírhatóak, és a hálózatok jellemzőinek változásai bizonyos pontossággal megjósolhatóak, hasonlóan, ahogy Newton axiómái például az üstökösök pályáját (a bolygók hatását is figyelembe véve) kiszámíthatóvá tették. Ebben a munkában a mérhető mennyiségekkel és a modellekkel foglalkozunk. Érdeemes azonban megemlíteni, hogy a terület részben túljutott a fent említett célokon, és a hálózatok vezérelhetőségével kapcsolatban komoly eredmények születtek: irányított hálózatok esetén meg lehet határozni, hogy melyik csúcsokat kell vezérelni ahhoz, hogy a rendszer véges idő alatt tetszőleges állapotba eljuttatható legyen [33].

Az oktatás szempontjából öröndetes, hogy sok kiváló, a hálózatok tudományával kapcsolatos cikk érhető el szabadon a Világhálón, közöttük összefoglaló jellegű munkák is. Tájékozódás szempontjából ajánlom angolul értő tanárkollégáimnak a következő két cikket [31, 34]. A másodikban, Newman cikkében, számos korábban

tanulmányozott hálózat jellemzőit találhatjuk meg táblázatban összesítve.

A témáról olvasmányos magyar nyelvű könyvként Barabási *Behálózva* című könyve ajánlható akár matematikától idegenkedő hallgatóknak is [35].

6. fejezet

Összetett hálózatok vizsgálata

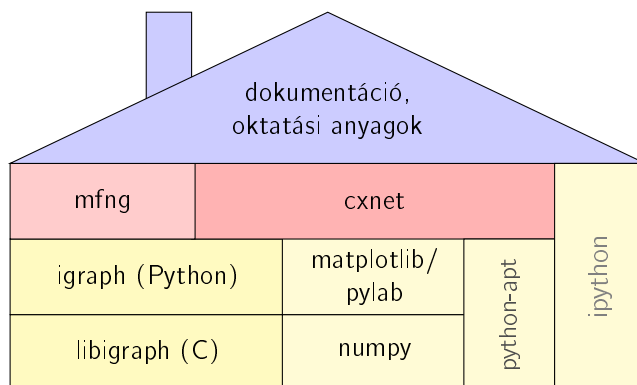
6.1. A cxnet hálózatvizsgáló programcsomag

Az oktatásban fontosnak tartom, hogy lehetőleg olyan szoftvereket használjak, amelyet a hallgató vagy diák szabadon használhat otthon is, könnyen megtanulható és a hallgató által továbbfejleszthető.

Az összetett hálózatok vizsgálatát informatikus tanulóknak tanítom teljes féléves heti 2 órás tantárgy keretében, a villamosmérnök hallgatóknak a számítógépes laborgyakorlat keretén háromszor két óra jut erre a témakörre. Az informatikus hallgatók tisztaiban vannak az objektumorientált programfejlesztés alapjaival, a villamosmérnök hallgatók nagy része viszont a C programozási nyelvet ismeri.

Mindkét csoportnak és esetleg középiskolás diákok számára is a Python objektumorientált programozási nyelvet tartom legegyszerűbben használhatónak, többek között azért, mivel interaktív módon is lehet vele vizsgálatokat végezni: a Python-parancssorban egyes parancsokat leírva rögtön végrehajtnak azok, továbbá egyszerűbb programok írásakor nem kell ismerni az objektumorientált programozási módszertant [36]. Egy további nagyon fontos szempont, hogy a tudományos kutatást segítő átgondolt szoftvercsomagok állnak a programozó rendelkezésére [37]. Ez részben annak köszönhető, hogy a Python képes C-ben és FORTRAN-ban írt, sokszor tesztelt függvénykönyvtárakat felhasználni.

A cxnet programcsomaghoz Python nyelven két nagy tudású hálózatvizsgáló programcsomag állt rendelkezésemre: a networkx [38] és az igraph [39]. Kezdetben az előző szolgált alapul, de később részletezett okok miatt a fejlesztés során



6.1. ábra. A cxnet és az mfng programcsomagok alapjául szolgáló programcsomagok

inkább az igraph használatára helyeztem a hangsúlyt.

A két hálózatzvizsgáló programcsomag nagyon jól csinálja azt, ami a feladata: a hálózatok tárolását, és azokból adatok kinyerését, valamint a gráfok ábrázolását. Számomra viszont szükség van függvények (például fokszámeloszlások) ábrázolására is. Szerettem volna továbbá, hogyha olyan hálózatot tudnának vizsgálni a hallgatók, amely nem egy régen elmentett hálózat adata, hanem valami olyannak a térképe, amely ott van a közelükben, és a hálózatzvizsgáló program az éppen aktuális adatokat gyűjti össze. Ezért hoztam létre a cxnet programcsomagot a Python nyelvhez.

A cxnet és a később tárgyalásra kerülő mfng programcsomagok forráskódja szabadon letölthető és módosítható bárki számára, mivel BSD-licence alatt tettem közzé. A programcsomagok letölthetőek a GitHub-on található tárolójukból [40], továbbá a Git verziókezelő rendszer remek alapot nyújt ahhoz, hogy hallgatók az én kódomat továbbfejlesszék a maguk ízlése szerint, a GitHub oldalon közzétegyék, és ha tetszik nekem a kód, beleolvasszam a sajátomba [41].

A cxnet jelenleg két fontos lehetőséggel rendelkezik. Az első a hálózatok fokszámeloszlásának kirajzoltatása és skálafüggetlen esetben a $-\gamma$ hatványkitevő meghatározása. A második bizonyos Linux terjesztések szoftvercsomag-hálózatának meghatározása és kezelése.

A cxnet és a hozzá szükséges programok telepítése viszonylag egyszerű Ubuntu és Debian Linux-terjesztések alatt. Windows alatt legegyszerűbben úgy használ-

hatuk a cxnetet, ha az összetett hálózatok honlapról letölthető, az Ubuntu egyik változatát (Lubuntu) használó, előre telepített virtuális gépet használjuk, amely a VirtualBox alatt használható. Ehhez 6 Gigabájt körüli szabad hely szükséges.

6.1.1. A fokszámoszlások kirajzolása

A cxnet programcsomag kezdetben arra szolgált, hogy hálózatok fokszámoszlását lehessen kirajzoltatni a matplotlib programcsomag segítségével. A matplotlib pylab modulja olyan függvényeket definiál, amelyek a MATLABTM függvényeihez hasonlítanak: hasonló paraméterekkel rendelkeznek, de mivel Python nyelven íródtak, a paraméterkezelésük sokkal rugalmasabb.

Mind a matplotlib (pylab), mind a MATLAB mátrixokkal dolgozik. A függvény rajzoltatásakor első két paraméterként két sorvektort kell megadni, az egyik az x-koordinátákat, a másik az y-koordinátákat tárolja. A pylab a mátrixok kezelésében a numpy programcsomagra épül, amelyben definiálva vannak a trigonometrikus, exponenciális és más egyéb függvények, amelyeket sorvektorokra (array adattípusra) alkalmazva elemenkénti műveletvégzés történik, még hozzá gyorsabban, mintha Pythonban egyesével végeznénk a függvény kiértékelését minden elemre, ugyanis az elemenkénti kiértékelést már C-programban megírt függvény végzi.

A cxnetben a következő típusú fokszámoszlások rajzolhatók ki:

1. Összegzett fokszámoszlás

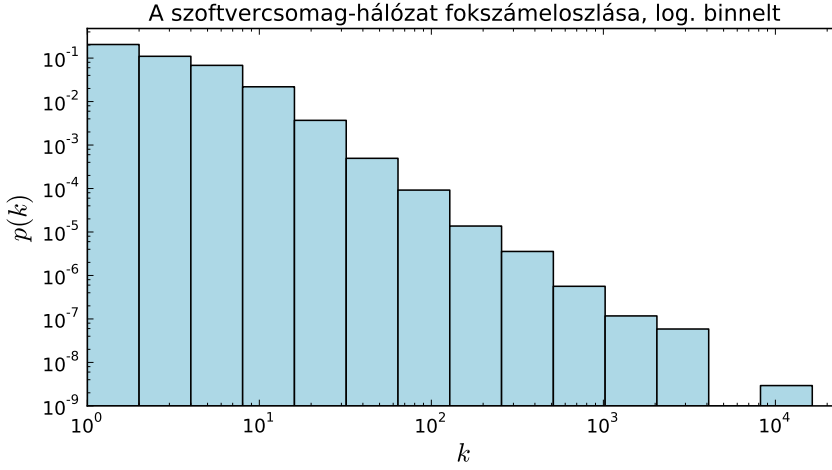
2. Binelt fokszámoszlások

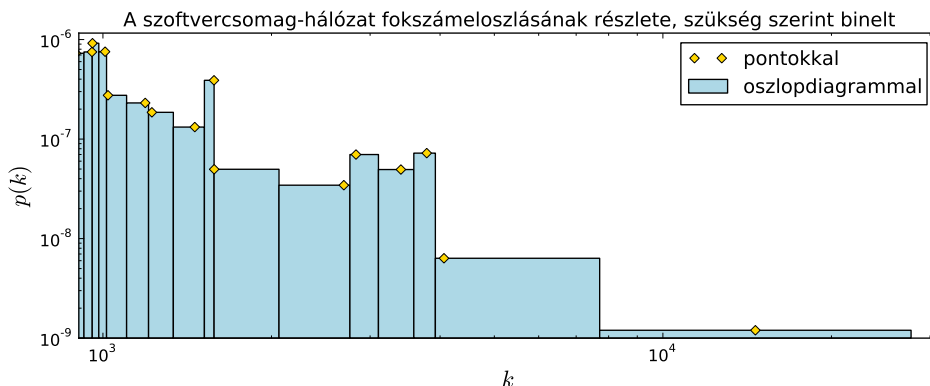
(a) logaritmikus bineléssel

(b) egységnyi hosszúságú binekkel

(c) szükségletnek megfelelő hosszúságú binekkel

A binelés lényegében azt jelenti, hogy ahelyett, hogy minden fokszámoszlásra ábrázolnám azt, hogy milyen valószínűséggel lesz egy véletlenszerűen kiválasztott csúcsnak a fokszáma az az érték, a fokszámoszlásokat csoportosítom. Az eloszlás grafikonjának fokszámoszlás-tengelyét felosztom intervallumokra, és azt a valószínűséget ábrázolom a függőleges tengelyen, hogy milyen valószínűséggel esik egy véletlenszerűen választott csúcs az adott intervallum egységnyi hosszúságú szakaszára. Az $[i, j[$ intervallumhoz tartozó $p_{[i, j[}$ valószínűséget úgy kapom meg, hogy azoknak a





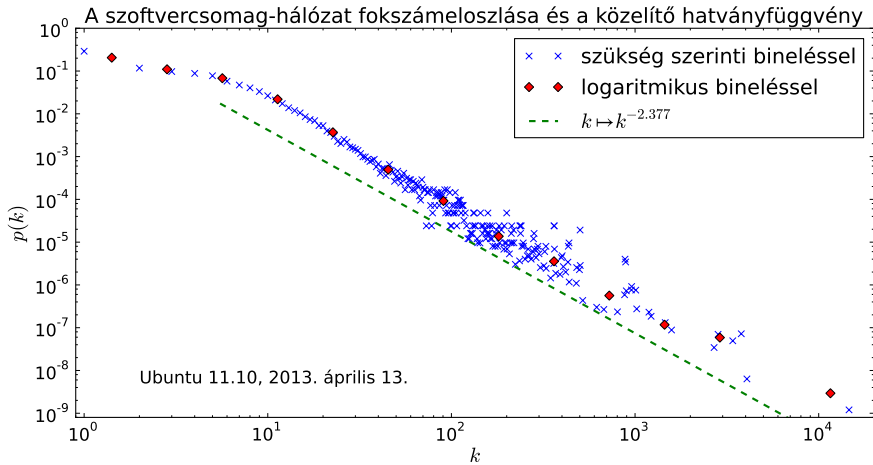
6.3. ábra. Egy fokszámeloszlás egy része szükség szerinti bineléssel ábrázolva

esetén úgy választom meg az intervallumhatárokat, hogy mindegyik $[d_i, d_{i+1}[$ intervallumba egyetlen olyan k_i fokszám essen, amihez van legalább egy csúcs, aminek az a fokszáma. Az intervallum határai itt általában nem egészek, hanem olyan módon választom, hogy ha a k_i és k_{i+1} egészekhez ($k_i < k_{i+1}$) tartozik vele egyező fokszámú csúcs, de a köztük lévő egyetlen egész számhoz sem, akkor a közöttük lévő d_i osztópontot (intervallum-végpontot) a két fokszám mértani közepeként számolom ki:

$$d_i = \sqrt{k_{i+1} \cdot k_i} \quad (6.3)$$

Ezzel majdnem minden osztópontot meghatároztam, csak a legkisebb fokszámhoz tartozó intervallum alsó határát és a legnagyobb fokszámhoz tartozó intervallum felső határát nem. Ezeket pedig úgy képzem, hogy legkisebb, illetve legnagyobb fokszám mértani közepe legyen a hozzá tartozó intervallum alsó és felső határának. Ezt a binelést és kétféle ábrázolási módszert használtam a 6.3. ábrán. A sárga gyémántjelek esetén a létező fokszámhoz rajzolta be a program a valószínűséget, az oszlopdigram szélessége jelzi az intervallum szélességét is, így jól nyomon követhető, hogy a logaritmikus ábrázolásnak és a mértani középnek köszönhetően az osztópontok vízszintesen felezik a két gyémántjellel jelölt pont vízszintes koordinátáját, illetve, hogy a legnagyobb fokszám felezi a felső intervallum két végpontját.

A cxnet legnagyobb valószínűség módszerével levezetett a 6.6. és a 6.7. képleteket használja a kitevő értékének és bizonytalanságának meghatározására. Ezek



6.4. ábra. Egy fokszámeloszlás egy része szükség szerinti és logaritmus bineléssel ábrázolva a közelítő hatványfüggvény meredekségével

a képletek a 6.2. fejezetben találhatók. Amennyiben nincs megadva minimális fokszám, a $k_{\min} = 5,5$ értékkel számol a program.

6.1.2. A networkx és igraph programcsomagok összehasonlítása

Python programnyelven két nagy tudású hálózatkezelő/elemező programcsomag létezik: a networkx és az igraph. A cxnet eredetileg a networkx-re épült, és bizonyos funkciói jelenleg is használhatóak azzal, azonban a cxnet minden lehetősége használható igraph alatt, és az új lehetőségek néha csak igraph-fal működnek. A networkx és az igraph is szabad forrású, szabadon felhasználható program. A networkx teljes mértékben Pythonban íródott, így a programkód minden operációs rendszeren működik, amelyre a Python telepíthető. Ezért választották a nagy tudású, Python nyelvet használó matematikai programcsomag, a SAGE részévé is.

Az igraph alapvető hálózatok létrehozására, módosítására alkalmas függvényeit C programozási nyelven valósította meg két magyar fejlesztő: Csárdi Gábor és Nepusz Tamás. Két nyelvhez írtak illesztőprogramot, az R statisztikai analízisre létrehozott nyelvhez és a Pythonhoz. Az igraph függvényei tehát C, C++, R és

Python programokban is használhatók.

Eredetileg a cxnet a networkx programcsomagon alapult, és egyes részei jelenleg is működnek vele. A fejlesztés egy részénél méréseket végeztem, hogy milyen gyorsan képes a networkx és az igraph hálózatok átmérőjét meghatározni, és emiatt a további fejlesztések alapjául az igraphot használtam.

8. definíció (hálózat átmérője). *A hálózat (gráf) átmérőjét úgy kapjuk, hogy az összes hálózatbeli csúcspárra megvizsgáljuk a csúcsok távolságát, és a kapott távolságok maximumát vesszük.*

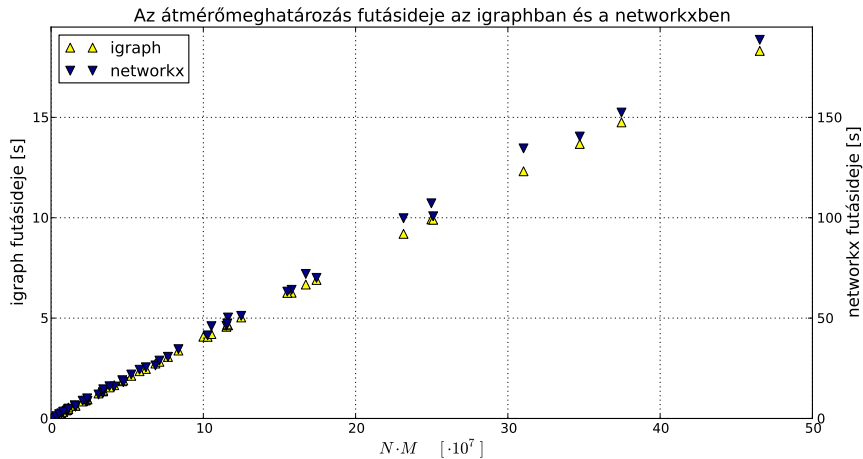
A nem összefüggő hálózat átmérőjét kétféleképpen szokás definiálni: az egyik esetben az átmérőt ilyenkor végtelennek veszik, másik esetben csak a véges távolságokat veszik figyelembe a maximum keresésekor.

Az előző definíciónak azt a változatát, amikor mindig véges az átmérő, átfogalmazhatjuk így is: a hálózat átmérőjének azt a minimális lépésszámot nevezzük, amennyi lépésszámmal a hálózat bármely csúcsából el tudunk jutni a hálózat tetszőleges másik csúcsába, amibe egyáltalán el lehet jutni.

Különösen az informatikus hallgatóim számára fontosnak tartottam, hogy a hálózatok kezelése mellett a hálózatok algoritmusainak hátterébe is belelássanak, hogy megértsék, hogy különböző módokon tárolhatunk hálózatokat, és ettől jelentősen függ, hogy a hálózat módosítása, illetve a hálózat adatainak meghatározása mennyi időbe telik. Mint látjuk majd, az igraph nagyon jó abban, hogy a hálózat adatait, például átmérőjét megtudjuk, viszont a hálózat módosítása viszonylag lassú a használt adatszerkezet miatt.

Az átmérő meghatározása nagy hálózatok esetén nagyon sok lépésből áll, mivel minden csúcsból meg kell keresnünk a tőle legtávolabb eső csúcsot. Ezt a csúcsot az úgynevezett *szélességi keresésnek* nevezett algoritmussal lehet megkeresni, ami legrosszabb esetben az élek számának megfelelő számú lépést igényel. Ha minden egyes csúcsból meg szeretnénk határozni a legtávolabbi csúcs távolságát, akkor tehát a szükséges lépések számát a legrosszabb esetben a csúcsok és élek számának szorzata adja meg. Az átmérő meghatározásához ezeknek a legnagyobb távolságoknak kell megkeresni a maximumát. A maximum megkeresésének ideje elhanyagolható a legnagyobb távolságok meghatározásához szükséges időhöz képest, tehát az átmérő-meghatározás futási idejének becslésekor nyugodtan felhasználhatjuk, hogy az jó közelítéssel a csúcsok és élek számának szorzatával arányos.

Érdemes megjegyezni, hogy az átmérő meghatározása jól párhuzamosítható,



6.5. ábra. Az átmérőmeghatározás futási ideje az $N \cdot M$ szorzat függvényében az igraph és a networkx programcsomag használatával véletlen hálózatok esetén. A két programcsomaghhoz két különböző futásidő-tengely tartozik, hogy a lineáris összefüggés az igraph esetén is jól látható legyen.

hiszen a különböző csúcsból kiinduló szélességi keresések különböző processzormagon/processzoron futhatnak.

Az átmérőmeghatározás futási idejét különböző csúcsszámú és élszámú véletlen hálózatok esetén vizsgáltam meg, és ábrázoltam az $N \cdot M$ szorzat (N a csúcsok száma, M az élek száma) függvényében. Azt kaptam, hogy kis szorzatértékekre különösen networkx használatakor függ a futásidő a csúcsok számától is, nem csak a szorzattól, nagyobb gráfok esetén viszont közel egyenes arányosságot kapunk. A mérések szerint nagy (100 millió – 500 millió közötti) szorzatértékekre a futásidők hányadosa jellemzően 10 körüli (9,5 és 11 közötti) értéket vesz fel, az igraph egy nagyságrenddel gyorsabban végzi el az átmérő kiszámítását, mint a networkx.

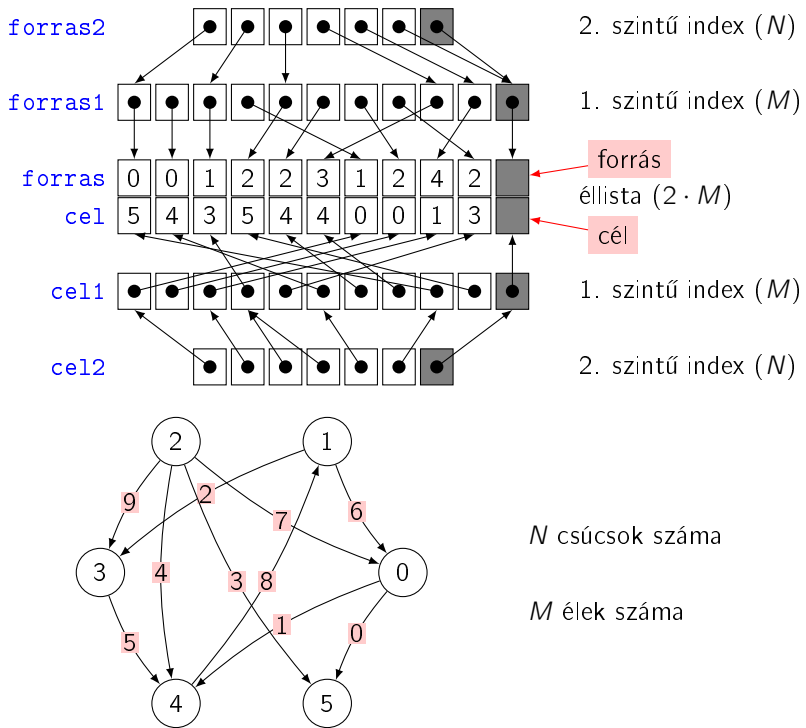
A sebességkülönbség nem csupán a C és a Python nyelvek sebességkülönbségéből adódik. Általában ugyanis egy C-ben írt program, amelyet az adott számítógép gépi kódjára fordítanak, jelentősen gyorsabb, mint a Pythonban írt, amely egy virtuális gép kódjára fordítódik le (hasonlóan a Java-hoz és a C#-hoz), és azt futtatja le a virtuális gép. Az igraph és a networkx jelentősen különböző adatszerkezetet

használ. A networkx külön adattípust használ az irányított és irányítatlan gráfokra.

Az `igraph` olyan adatszerkezetet használ a gráfok tárolására, amelynél a szomszéd csúcsokat nagyon könnyen meg tudjuk találni. Ehhez olyan éllistát tárol az `igraph` C-ben írt magja, amely kétszeresen indexelt (6.6. ábra). Az `igraph` a csúcsok és élek indexelését 0-val kezdi. Ha N csúcsa és M éle van a gráfnak, akkor a csúcsok indexei 0-tól $N - 1$ -ig futnak, az éleké 0-tól $M - 1$ -ig. Az éllista egy olyan lista, amelyben csúcspárok helyezkednek el. A továbbiakban az él egyik végén található csúcstól forrásnak, a másikon található célnak fogjuk nevezni, attól függetlenül, hogy irányított vagy irányítatlan gráfról van szó. Az `igraph` az élek forrásait és a céljait egy-egy külön tömbben tárolja. A 6.6. ábrán a 0-ás indexű él forrása a 0-ás indexű csúcs, a célja az 1-es indexű, tehát (amennyiben a gráf irányított) a 0-ás él a 0-ás csúcsból az 1-es csúcsba mutat, ahogy az ábrán alul láthatjuk, az 1-es indexű él a 0-ásból az 2-es csúcsba mutat, és így tovább. Az 1. szintű index egymás utáni elemei olyan sorrendben mutatnak az élek egyik végét képviselő csúcsokra, hogy azok monoton növekedőek legyenek. A 2. szintű index i -dik eleme mutat arra a legelső 1. szintű elemre, amely i értékre mutat. Az adatszerkezetben található még egy logikai érték, amely jelöli, hogy a gráf irányított, vagy nem.

A további példában vizsgáljuk először az irányított hálózatot. Az i . csúcs kifokszámát nagyon könnyen meg tudjuk határozni, csak a 2. szintű indexben ki kell vonni az $i + 1$ -dik elemből az i -diket. Ha a forrás 2. szintű indexét `forras2`-vel jelölöm (ahogy a kék felirat mutatja az ábrán), az indexet pedig a C-ben megszokott módon szögletes zárójelbe írjuk, akkor az i -dik csúcs kifokszámát a `forras2[i+1] - forras2[i]` képlettel kapjuk meg. A befokszám hasonlóan határozható meg, csak a `ce12` indexből. Ha a fokszámot szeretnénk meghatározni, akkor a be- és kifokszámot össze kell adni. Ugyanez az érték adja az irányítatlan eset fokszámát is.

Az i -dik csúcs ki-szomszédait (azokat a csúcsokat, amelybe az i -dik csúcsból él vezet) is könnyű meghatározni. A forrás 1. szintű indexében (ezt jelöljük `forras1`-gyel) végigmegyünk a `forras2[i]`-edik elemtől a `forras2[i+1]-1`-edikig. Az ezekben szereplő indexek fogják megadni azoknak az éleknek az indexeit, amelynél a forrás az i -dik csúcs. Csupán a célt tároló tömbben (`ce1`) kell megkeresni ugyanezen indexű elemeket, amik megadják a ki-szomszédok sorszámát, azokét, amelyek az i -dik csúcsból kifelé mutató élek céljánál találhatók. Hasonló módon megtalálhatóak a be-szomszédok is, irányítatlan esetben pedig a szomszédok a két halmaz egyesítésével.



6.6. ábra. A gráfok tárolására használt adatszerkezet az igrphban és az általa tárolt gráf

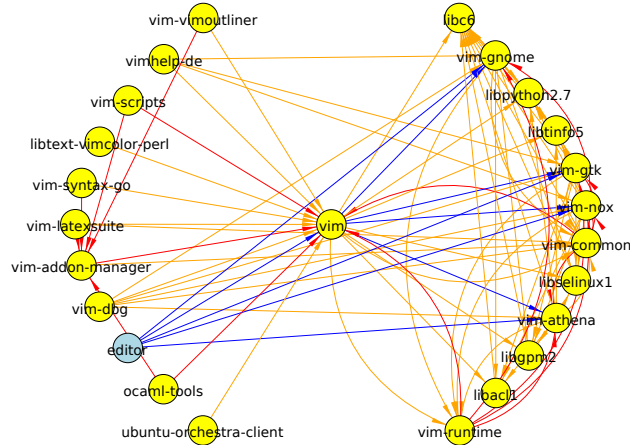
6.1.3. A szoftvercsomag-hálózat létrehozása

A cxnet második fő lehetőségeként a GNU/Linux operációs rendszer szoftvercsomag-hálózatát hozza létre és vizsgálatát segíti elő.

A GNU/Linux rendszereknek rengetegféle terjesztése van, melyek saját szoftvercsomag-rendszert alakítanak ki. Ezeknek a szoftvercsomagoknak két elterjedtebb formája van, melyek közül az egyik a Debian terjesztésben [42] kialakított *deb* csomagformátum. Ezeket a csomagokat optikai lemezzről (CD, DVD) vagy internetes tárolókból érhetjük el. A csomagok másik csomagoktól függhetnek, azaz azok nélkül nem vagy csak részlegesen képesek működni. Ezeket a függőségeket az apt csomagkezelő rendszer képes kezelni: képes többek között egy csomag telepítésekor a hozzá szükséges csomagokat megkeresni és azokkal együtt telepíteni, egy csomag eltávolításakor a rá épülő csomagokat is eltávolítani. A csomagok tehát egy irányított hálózatot alkotnak. A programban az élek irányát úgy definiáltuk, hogy azok a függő csomagtól mutatnak afelé, amelytől függ. A 6.7. ábrán a szoftvercsomag-hálózat egy kis részét láthatjuk, a Vim szövegszerkesztő csomagjának egy kis környezetét. Jobbra láthatóak azok a csomagok, amelyektől a vim függ, és balra azok, amelyek a vimtől függenek. Látható az ábráról, hogy a vim csomag telepítéséhez a Python programozási nyelv 2.7-es változata is (python2.7 csomag) szükséges, és a C programozási nyelv standard könyvtára (libc6 csomag) is, valamint a vim csomag is szükséges más csomagok telepítéséhez. Látható továbbá, hogy azok közül a csomagok közül is sok függ a libc6 csomagtól, amelyektől a vim függ.

A szoftvercsomag-hálózatban a csomagok kétfélek lehetnek. A legtöbb csomag, a vim is, valódi csomagok. Ezeket le lehet tölteni és telepíteni. A 6.7. ábrán ezeket a csomagokat piros kör jelöli. Van néhány virtuális csomag is, amilyen *editor* az ábrán. Ezekre hivatkozhatnak más csomagok valamilyen kapcsolatban. Emellett ezeket a csomagokat legalább egy valódi csomagnak szolgáltatnia kell. Az ábrán látható, hogy a vim, vim-gnome szolgáltatja ezt a csomagot, és néhány másik, amelyek nem láthatóak az ábrán, mint az emacs, a nano és az mcedit, tehát, ha egy csomag függőségeként listázza az *editor*-t akkor ezek közül akármelyik elegendő a függőség teljesítéséhez, és nem kell feltétlenül a csomagkarbantartó kedvenc szövegszerkesztőjét telepítenünk. A képen kék nyilak mutatnak a virtuális csomagoktól azok felé a csomagok felé, amelyek azt szolgáltatják.

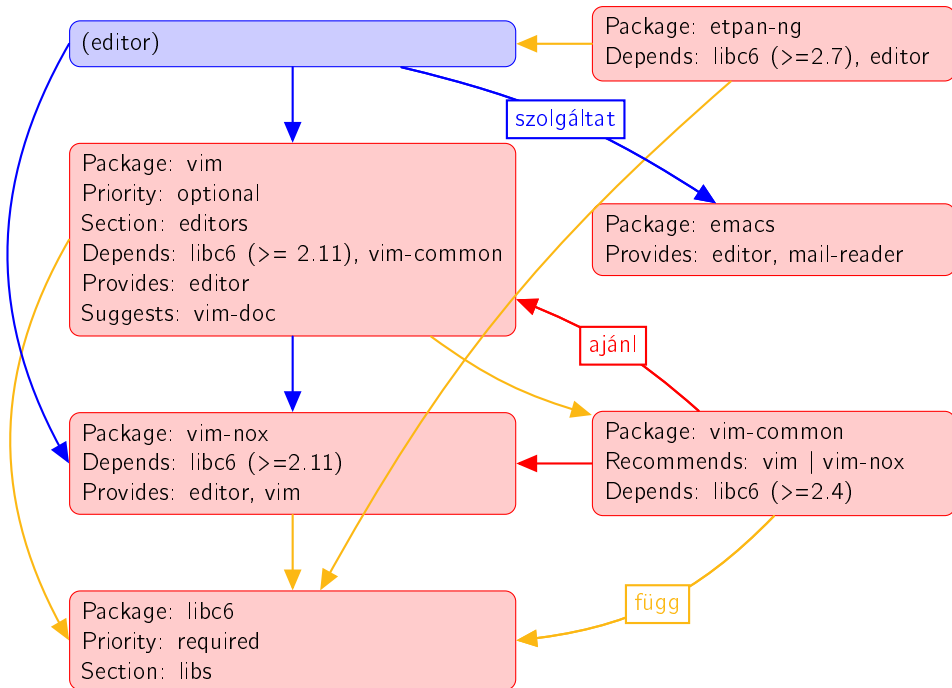
Eddig kétféle kapcsolatról volt szó: egy csomag függhet (depends) egy másik csomagtól, amit az ábrán narancssárga színű nyíl jelez, valamint egy csomag szolgáltat (provides) egy másik (általában virtuális) csomagot, amit pedig kék



6.7. ábra. A szoftvercsomag-hálózat egy része, a Vim szövegszerkesztő szomszédaival

nyíl jelöl. Van azonban a függőségen kívül néhány egyéb függőségjellegű kapcsolat, amely nem annyira kötelező, mint az igazi függőség. Egy csomag ajánlhat (recommends) más csomagot, ezek az ábrán piros nyílal vannak feltüntetve. Az ajánlott csomagok nem feltétlenül szükségesek a program működéséhez, de valamilyen lehetőséggel kibővítheti annak használatát, és általában az egyes programok dokumentációit is ajánlják a programok. Az utolsó, még kisebb erősségű kapcsolat, amikor egy csomag javasol (suggests) egy másikat [43].

A cxnet a szoftvercsomag-hálózat létrehozásakor a valódi és virtuális csomagokat, valamint a különböző függőségjellegű kapcsolatokat képes nagyjából 10–20 másodperc alatt feltérképezni a python-apt nevű szoftvercsomag segítségével [44], és akár fájlba is eltárolni ezeket későbbi vizsgálatok céljára. A szoftvercsomag-hálózatból egyszerű parancsokkal képes a legnagyobb (be-, ki- és sima) fokszerű csomagok listáját létrehozni, statisztikát készíteni a különböző típusú csomagokról illetve függőségjellegű kapcsolatokról, kirajzoltatni egy csomag szomszédságát a 6.7. Képes a cxnet arra is, hogy bizonyos típusú éleket vagy kapcsolatokat töröljön, így megvizsgálható annak a szoftvercsomag-hálózatnak a tulajdonsága is, amely csak a valódi éleket és közöttük csak a függőségeket tartalmazza. A cxnet



6.8. ábra. A szoftvercsomag-hálózat egy része a cxnet színeivel és a csomagok pár tárolt jellemzője. A virtuális csomagokhoz (pl. editor) nem tartozik semmilyen jellemző: azt, hogy mi szolgáltatja, és mi függ tőle, más csomagok jellemzőiből lehet meghatározni. A cxnetben szoftvercsomag-hálózat kirajzolásakor a valódi csomagokat piros, a virtuálisakat kék kör jelzi; a függőséget narancssárga, a szolgáltatást kék, az ajánlást (recommends) piros nyíl jelzi, a javaslatot (suggests) a cxnet jelenleg még nem tárolja.

már említett lehetőségét használva egyszerűen kirajzoltathatjuk a szoftvercsomag-hálózat fokszámoszlását, és berajzoltathatjuk a fokszámoszlást közelítő hatványfüggvény egyenesét is. Az itt leírtak mind a `cxnet` saját szolgáltatásai közé tartoznak. Mivel a `cxnet` `Network` osztálya az `igraph` `Graph` osztályából származik, ezért az előző lehetőségek mellett használhatjuk az `igraph` összes lehetőségét is: megvizsgálhatjuk a csúcsok fokszámaikat, felbonthatjuk a hálózatot erősen vagy gyengén összefüggő komponensekre, és minden fontosabb hálózati mennyiséget kiszámoltathatunk vele.

A 6.8 ábrán részletesebben látható a szoftvercsomag-hálózat, mint ahogy a `cxnet` ábrázolni képes. Ezen az ábrán feltüntettem az egyes csomagok bizonyos jellemzőit, amit a megértés szempontjából fontosnak tartottam. Minden valódi csomag esetén létezik például fontosság mező, de nem mindenhol tüntettem fel. Az ábra összevethető a `vim` csomag szomszédságát megjelenítő ábrával, az itt szereplő csomagok, az `emacs` kivételével, ott is szerepelnek. A `cxnet` nem tárolja az összes csomagjellemzőt, csak azokat, amelyeket fontosnak tartottam a hálózat jellemzőinek meghatározásához.

A `cxnet` a szoftvercsomag-hálózat esetén többek között a következő adatokat tárolja (zárójelben a `cxnet`-ben használt jellemzőnév):

1. A teljes hálózathál:

- (a) A szoftvercsomag-adatok frissítésének időpontját: az `apt` csomagkezelő rendszer képes az internetes tárolókból az éppen fent levő csomagok jellemzőit letölteni. A `cxnet` az utolsó ilyen frissítés dátumát tárolja, nem azt, amikor a `cxnet` a hálózatot feltérképezte. (`update_time`)
- (b) A tárolók adatait: azt, hogy melyik tárolók vannak beállítva az `apt` számára, hogy csomagjellemzőket és csomagokat töltsön le azokról. (`sources_list`)
- (c) A processzor-architektúrát: az Ubuntu a 32 bites PC-k esetén az `i386` jelölést, a 64 bites PC-k esetén az `amd64` jelölést használja. (`architecture`)

2. Az egyes csomagok (csúcsok) esetén:

- (a) a csomag nevét (`name`)
- (b) a csomag típusát, amely valódi vagy virtuális lehet (`type`)
- (c) a csomag fontosságát (`priority`)

- (d) a csomag milyen szakaszba tartozik (section)
 - (e) a csomag csomag rövid összefoglalóját (summary)
 - (f) a csomag változatát (version)
 - (g) a csomag letöltendő (tömörített) méretét (filesize)
3. a csomagok közötti függőségszerű kapcsolatok (élek) esetén:
- (a) a kapcsolat típusát (type)

Ezek az adatok a hálózat fájlba mentésekor mind megőrződnek. A `cxnet` képes olyan Linuxon is futni, amelyen nincsen grafikus felület telepítve, így képes időzített feladatként adott időközönként elmenteni egy-egy állapotot.

6.2. A szoftvercsomag-hálózat tulajdonságai

A szoftvercsomag-hálózatok vizsgálatával kapcsolatban az általam talált legkorábbi cikk LaBelle és Wallingford 2004-es cikke [45]. Ebben a cikkben megvizsgálták a Debian Linux és a FreeBSD szoftvercsomag-hálózat be- és kifokszám-eloszlását és csoportterősségi együttthatóját. A fokszám kitevő illesztéssel történő meghatározásához a legkisebb négyzetek módszerét használták, de a kapott pontokra nem alkalmaztak semmilyen binelési módszert, emiatt a hatványfüggvény-eloszlás kitevőjére a ténylegesnél jóval kisebb értéket kaptak, a Debian esetén $\gamma_{out} = 2,33$ és $\gamma_{in} = 0,9$ értéket. A Debian szoftvercsomag-hálózatának csoportterősségi együttthatójára 0,52-es értéket kaptak.

Bár már Newman 2003-as cikkének táblázatában is szerepelnek a szoftvercsomag-hálózat adatai, de azt, hogy milyen módon mérték meg a fokszámeloszlások kitevőit, nem sikerült megtalálnom ebben a cikkben [31]. A táblázatban a szoftvercsomagok hálózatára a következő adatok szerepelnek: 1439 csúcsa és 1723 éle van, a be- és kifokszámeloszlás hatványkitevője $\gamma_{out} = 1,4$ és $\gamma_{in} = 1,6$, a csoportterősségi együttthatója pedig 0,082. A fenti táblázat minden sorában van egy olyan oszlop, amely olyan cikkekre hivatkozik, amelyben az adott hálózatról van szó [46]. A hivatkozott cikkből kiderül, hogy a GNU/Linux operációs rendszerről van szó, amely általában a Debiant szokta jelenteni, de a fokszámeloszlás mérési módszerére ebben sincs utalás. Ebben a cikkben a szoftvercsomag-hálózat egyéb adatai szerepelnek (a csúcsainak száma és az úgynevezett asszortativitása).

Maillart és társai 2008-as cikkükben több 2002 és 2008 közötti Debian verzió fokszámeloszlását hasonlították össze [47]. Megállapították, hogy minden esetben

a fokszámok több mint négy nagyságrendjében teljesül a hatványfüggvény eloszlás. A fokszámeloszlás hatványkitevőjét a legnagyobb valószínűség (maximum likelihood) módszerével állapították meg, és a mi cikkünkben leírtakhoz hasonlóan ábrázolták a kapott hatványkitevő-értéket és annak szórását a figyelembe vett fokszámok k_{min} alsó határa függvényében, de a cikkünkben nem állapítottak meg statisztikus és szisztematikus bizonytalanságot. Ellenőrizték viszont, hogy a népszerűségi kapcsolódás szerint fejlődik a hálózat, azaz a nagyobb fokszámú csúcsokhoz (szoftvercsomagokhoz) több új csúcs csatlakozik. Cikkük 2/A. ábrája alapján az egyenes arányosság valószínűsíthető.

Sousa és munkatársai 2009-es cikkükben [48] szintén a Debian terjesztés szoftvercsomag-hálózatát vizsgálja. Egységnyi hosszúságú binekkel ábrázolja a fokszámgyakoriságokat, és a befokszámok esetén $\gamma_{in} = 2$ értéket kap a kitevőre, a kifokszám-eloszlást inkább hatványfüggvény jellegűnek találja. Az illesztés pontos menetét nem tárgyalja, az ábrából úgy tűnik, mintha kis fokszámokra illesztene csak.

Gradowski és munkatársai 2011-es cikkükben a Debian 6.0-ás verziójának szoftvercsomag-hálózatát vizsgálják vélhetően 2011-es állapotát [49]. Az illesztéshez logaritmikus binelést használtak, az illesztés módját nem írja le a cikk. A befokszámot hatványfüggvénynek találták $-\gamma_{in} = -2$ kitevővel, a kifokszámeloszlást pedig olyan függvénynek, amely 20-nál kisebb és annál nagyobb értékekre is hatványeloszlást követ, de jelentősen különböző kitevőkkel, kis fokszámokra $-\gamma_{out,small} = -2$ nagy fokszámokra $-\gamma_{out,large} = -5$ kitevőértékekkel. A cikkben a csoportterösségi együttható számolása más módon történik, mint a fentebb említett.

A csoportterösségi együtthatót irányított hálózatokra értelmezi Barrat és munkatársainak cikke alapján [50]. Barrat és munkatársai irányított súlyozott hálózatokra értelmezést adnak a csoportterösségi együtthatóra, ezt egyszerűsíti a cikk súlyozatlan esetre. Eszerint az irányított hálózat i -dik csúcsának csoportterösségi együtthatója.

$$C_{i,dir} = \frac{1}{k_i(k_i - 1)} \sum_{j,k} a_{ij}a_{ik}a_{jk}, \quad (6.4)$$

ahol az a_{ij} értéke 1, ha irányított él vezet az i -dik csúcsból a j -dik csúcsba, máskülönben 0. A hálózat csoportterösségi együtthatóját ezután – az általunk irányítatlan esetben alkalmazott módszerhez hasonlóan – úgy kapják, hogy a kettő és nagyobb fokszámú csúcsokra átlagolják az egyes csúcsok csoportterösségi együtthatóját, és így módon $C_{dir} = 0,306$ értéket kapnak. A cikk megállapítja továbbá, hogy a hálózat hierarchikus, de a csoportterösségi együtthatók átlagának kifokszám-függésére nem -1 -es, hanem -2 -es hatványkitevőt kaptak. Az ábrából úgy tűnik, hogy csak

6.1. táblázat. A fokszámmillesztések adatai

eloszlás	kitevő	intervallum	$\chi^2/\text{d.o.f.}$
befokszám	-1,79	$3 \leq k_{in} \leq 11114$	0,2%
kifokszám	-2,88	$3 \leq k_{out} \leq 161$	7,2%
kifokszám	-1,87	$3 \leq k_{out} \leq 29$	0,4%
kifokszám	-4,79	$30 \leq k_{out} \leq 161$	3,4%
fokszám	-2,11	$5 \leq k \leq 11115$	0,2%

a 30-nál nagyobb kifokszámokat veszik figyelembe.

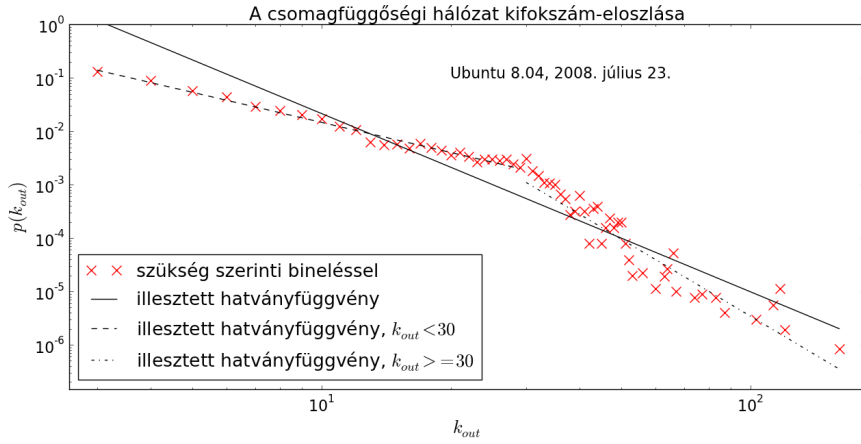
6.2.1. Saját eredményeim

A szoftvercsomag-hálózatot először 2008-ban tanulmányoztam az [51]. A konferencia, amin az eredményeket bemutattam, egy hónappal Maillart és társai által írt cikk megjelenése előtt zajlott, így az ő eredményeikről nem tudtam.

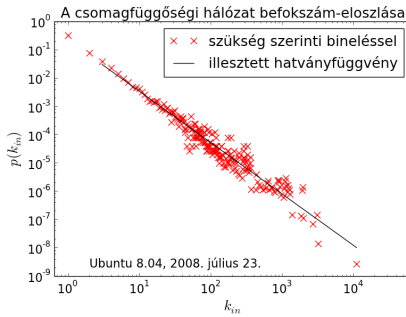
A *deb*-csomagokat használó Ubuntu terjesztés [52] 8.04-es verzióját vizsgáltam meg, a 2008. júliusi állapotot. A *cxnet* elődjével hoztam létre a szoftvercsomag-hálózatot, amelyben csak valódi csomagok, és a kapcsolatok közül csak a függőségek szerepeltek. A teljes hálózat 25518 csúcsból áll, 121304 élt tartalmaz, így az átlagfokszáma 9,50. A *libc6* nevű csomag a legnagyobb fokszámú és egyben legnagyobb befokszámú csúcs, a 11115-ös fokszámmal illetve 11114-es befokszámmal. A legnagyobb kifokszám értéke 161. Megállapítottam, hogy – sejtésemmel ellentétben – nem teljesen körmentes, hanem tartalmaz 657 olyan párt, amelyek között oda és vissza is mutat él.

Újabban megvizsgáltam ugyanezt a hálózatot a komponensek szempontjából is. A teljes hálózat 25518 csúcsból áll, a legnagyobb gyengén összefüggő komponens 23773 csúcsot, a csúcsok 93,2%-át tartalmazza. 24700 erősen összefüggő komponens tartalmaz, ezek közül a három legnagyobb komponensben 48, 11 és 10 csúcs található.

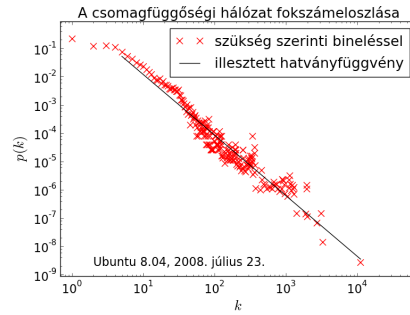
Az [51] cikkben a hálózat fokszámeloszlását a legkisebb négyzetek módszerével határoztam meg úgy, hogy a fokszámeloszlásra előbb a szükség szerinti binelési módszert alkalmaztam. Ahogy a 6.9. ábrán is látható a befokszám-eloszlást a teljes tartományban egyetlen hatványeloszlás láthatóan jól leírja, a kifokszám-eloszlás viszont két szakaszra bontható, harmincnál nagyobb és kisebb fokszámkok esetén különböző hatványkitevők írják le az eloszlást. Az illesztés jóságát az egy szabad-



(a) Kifokszám-eloszlás



(b) Befokszám-eloszlás



(c) Sima fokszámeloszlás

6.9. ábra. Az Ubuntu szoftvercsomag-hálózatának fokszámeloszlásai 2008 júliusán és az azokat közelítő hatványfüggvények. A piros $\times$ -ek jelölik a szükség szerinti bineléssel kapott eloszlást a fekete egyenesek pedig a közelítő hatványfüggvényeket.

sági fokra eső χ^2 értékkel jellemeztem. A számszerű értékek a 6.1. táblázatban találhatók.

A cikkben a csomagok között definiáltam egy hasonlósági mértéket is a beszomszédok alapján, azaz azoknak a csomagoknak alapján, amelyek az adott csomagtól függenek. Az i -dik csomag beszomszédainak halmazát jelölje B_i . A halmaz számosságát az abszolútérték-jellel jelölve az i -dik és j -dik csúcs *beszomszéd-hasonlóságát* a következőképpen definiáltam:

$$s_{ij} = \frac{|B_i \cap B_j|}{|B_i \cup B_j|} \quad (6.5)$$

Úgy gondoltam, hogy azok a csomagpárok érdekesek egy hálózatban, amelyek mindegyikének nagy a befokszáma, és a beszomszédok nagy része megegyezik, hiszen akkor a két csomag képességeit általában együtt használják. A fejlesztők és csomagkarbantartók számára ez azt jelentheti, hogy a két csomagot érdemes lehet összevonni. A 0,9-nél nagyobb beszomszéd-hasonlósággal rendelkező csomagpárok közül azok, amelyeknél a nagyobb befokszámú csomag befokszáma 50 felett van, a 6.2. táblázatban találhatók. A táblázat alján szereplő csomagpárok 6 olyan csomagból állnak, amelyek közül bármelyik kettőt kiválasztva a befokszámuk több mint 90%-a közös.

Később az Ubuntu Jaunty kódnevű 9.04-es verziójának szoftvercsomag-hálózatát vizsgáltam meg, annak 2009. november 3-ikái állapotát [53]. Ebben is csak valódi csomagok és függőségek szerepeltek.

Ebben a hálózatban a csúcsok és az élek száma $N = 27554$, $M = 126540$. A hálózat nem összefüggő, benne 25663 gyengén összefüggő komponens van, a csúcsok 93,14%-a tartozik a legnagyobb komponenshez. A legnagyobb gyengén összefüggő komponens átmérője 13, azaz benne bármely két csúcs elérhető egymásból legfeljebb 13 egymáshoz csatlakozó élen keresztül. A hálózat előállítás után meghatározható a hálózat összegzett fokszámeloszlása és a legnagyobb valószínűség módszerével meghatározható a kitevő $\gamma > 0$ ellentettje és annak szórása [54]:

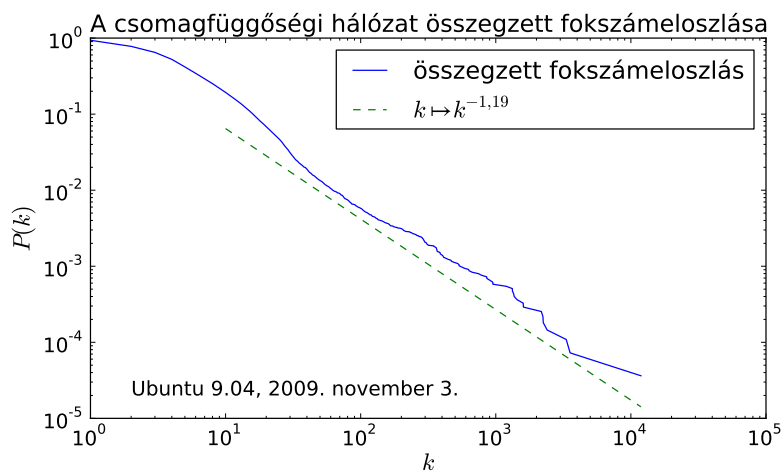
$$\gamma = 1 + N \left[\sum_{k_i > k_{\min}} \ln \frac{k_i}{k_{\min}} \right]^{-1}, \quad (6.6)$$

$$\sigma = \frac{\gamma - 1}{\sqrt{N}} \quad (6.7)$$

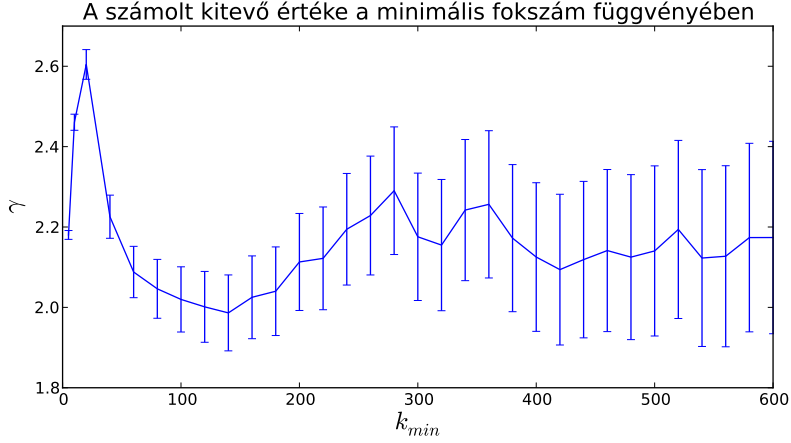
A képlet folytonos hatványfüggvény-eloszlásra vonatkozik, de kis módosítással jól használható a diszkrét hatványfüggvény eloszlásra [55]. A pontosabb közelítés ér-

6.2. táblázat. Csomagpárok $s_{ij} > 0,9$ $N_{ij,max} > 50$ értékkel. ($N_{ij,max} = \max(q_i, q_j)$ a két csomag befokszámának maximuma.)

1. csomag	2. csomag	$N_{ij,max}$	s_{ij}
libatk1.0-0	libpango1.0-0	1280	0,917
libcairo2	libpango1.0-0	1280	0,904
libatk1.0-0	libcairo2	1226	0,907
libice6	libsm6	1134	0,990
libxrandr2	libxcursor1	906	0,903
libgl1-mesa-glx	libgl1	363	0,922
libxcomposite1	libxdamage1	332	0,938
libqt4-core	libqt4-gui	310	0,916
libglu1-mesa	libglu1	287	0,982
kde-icons-oxygen	kdebase-runtime	174	0,982
kde-icons-oxygen	kdebase-runtime-data	173	1
kdebase-runtime	kdebase-runtime-data	174	0,982
libesd0	libesd-alsa0	108	0,990
libblas3gf	libblas.so.3gf	72	0,958
roxen2	roxen	66	0,924
liblapack3gf	liblapack.so.3gf	61	0,950
libgnustep-base1.14	gnustep-base-runtime	70	0,971
gnustep-gpbs	gnustep-back0.12	59	0,983
libgnustep-gui0.12	gnustep-gui-runtime	65	0,953
gnustep-gpbs	gnustep-gui-runtime	62	0,935
gnustep-back0.12	libgnustep-gui0.12	65	0,907
gnustep-back0.12	gnustep-gui-runtime	62	0,920



6.10. ábra. A szoftvercsomag-hálózat **összegzett fokszámeloszlása** (folytonos). Szaggatott vonallal a legnagyobb valószínűségű fokszámhoz tartozó $-1,19$ kitevőű **hatványfüggvény**. $\left(\gamma = 2,19 \pm 0,14 \begin{smallmatrix} +0,1 \\ -0,12 \end{smallmatrix}\right)$



6.11. ábra. A fokszámeloszlás kitevőjének abszolútértékére a legnagyobb valószínűség módszerével kapott értékek és bizonytalanságuk a figyelembe vett legkisebb fokszám függvényében.

dekében a figyelembe venni kívánt legkisebb egész értéknél féllel kisebbet szoktunk írni a képletbe, azaz a

$$\gamma = 1 + N \left[\sum_{k_i > k_{\min}} \ln \frac{k_i}{k_{\min} - 0,5} \right]^{-1} \quad k_{\min} \in \mathbb{Z} \quad (6.8)$$

képlettel számolunk.

A képlet akkor adja pontosan a kitevőt és a szórását, ha az értékek ténylegesen hatványfüggvényből származnak. Esetünkben nem bizonyítottam a hatványfüggvény jelleget. Az eloszlásfüggvény pontos elemzése mélyebb statisztikai ismereteket igényel [55].

Más esetekben érdemesnek tartom ábrázolni a γ érték függését a minimális fokszámtól ($k_{\min}$) a 6.11. ábrán látható módon. Az ábra szerint a σ statisztikai bizonytalanság kis $k_{\min}$ értékekre csábítóan kicsi a sok figyelembe vett értéknek köszönhetően, azonban ez nem jelenti a kitevő igazi bizonytalanságát. Mint láthatjuk, a kitevő értéke csak 240 körül kezd beállni egy állandó értékre. A kitevőt csak a 240-nél nagyobb értékekre érdemes értelmezni. A kitevő bizonytalanságánál pedig érdemes figyelembe venni a statisztikai ingadozást jellemző σ mellett a vál-

tozó γ értéket is: a szisztematikus bizonytalanságot. $k_{\min}=240$ esetén $\gamma = 2,19$, $\sigma = 0,14$ a (6.8) és (6.7) képletekből. A szisztematikus bizonytalanságot a 240-nél található értéktől való eltérésekből számolhatjuk.

$$240 \leq k_{\min} \leq 600 \text{ esetén } 2,07 \leq \gamma \leq 2,29 \quad (6.9)$$

Lefelé a legnagyobb eltérés 0,12, felfelé 0,10, tehát

$$\gamma = 2,19 \pm 0,14 \begin{matrix} +0,1 \\ -0,12 \end{matrix} \quad (6.10)$$

Az ennek a kitevőnek az összegzett eloszlás használatakor $-\gamma + 1 = -1,19$ kitevő felel meg. Az ehhez tartozó meredekséget az összegzett eloszlás ábráján ábrázolva ellenőrizhetjük, hogy az összegzett eloszlás valóban közelítőleg párhuzamosan fut a $k^{-1,19}$ hatványfüggvénnyel.

A hálózatban megkereshetjük a legnagyobb be- és kifokszámú csúcsokat (csoomagokat), és a fokszámukat összevethetjük az átlagos befokszámmal, illetve a vele azonos átlagos kifokszámmal. A befokszámra három, a kifokszámra kettő időpontbeli állapotot vet össze a 6.3. táblázat, kivételesen az eddig vizsgáltnál újabb, 2012. július 10-ei állapot sorrendjében. A táblázatban látható, hogy az egyes csomagok fokszáma az idő során általában növekedett, csak némelyiknél csökkent, esetenként a csomagok helyezése is jelentősen megváltozott. A továbbiakban visszatérek a 2009-es változat vizsgálatára.

Az átlagos befokszám: élek száma/csúcsok száma = $M/N = 4,592$. Ehhez képest a legnagyobb befokszám 11868, ami jelentősen nagyobb.

A csoporterősségi együttható vizsgálatához irányítatlanná kell tennünk a hálózatot, mert azt irányítatlan hálózatokra értelmeztük. A hálózat csoporterősségi együtthatója 0,307796. Vessük össze ezt a különböző modellekből kapott értékekkel! Ha a hálózat véletlen gráf lenne, akkor az élvalószínűség és így a csoporterősségi tényező értéke

$$p = \frac{M}{\binom{N}{2}} = \frac{2M}{N(N-1)} = C = 0,000333 \quad (6.11)$$

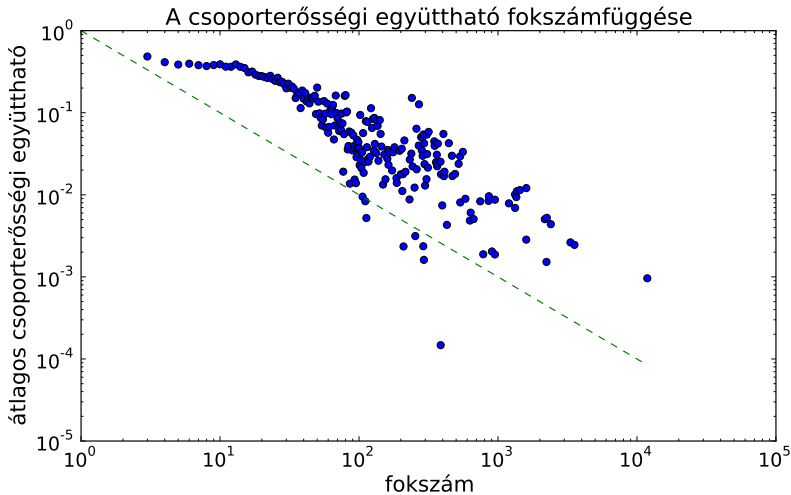
lenne. A Barabási–Albert modell szerint fejlődő hálózatokban az átlagos fokszám az m paraméter kétszerese, ezért a szoftvercsomag-hálózat átlagos fokszámának felét véve és kerekítve kapjuk a szoftvercsomag-hálózattal összevethető modell paraméterét. Mivel az átlagos fokszám $\langle k \rangle = 2M/N = 9,1849$, az $m = 5$ választás megfelelő. Az $N = 27554$ és $m = 5$ paraméterek mellett lefuttatva a modellt, egy 137745 élű hálózatot kapok, melynek csoporterősségi együtthatója 0,003223.

6.3. táblázat. Az Ubuntu legnagyobb ($k_{in} > 1000$) befokszámú csomagjai 2012. 07. 10-én (11.10-es Ubuntu verzió), k_{in} : az akkori fokszám, k_{in}' : az 2009. 11. 03-ai befokszám (9.04-es verzió), k_{in}'' : a 2008. 07. 18-ai (8.04-es verzió) fokszámai. (¹ A libqtcore4 csomag korábban libqt4-core névre hallgatott)

k_{in}	k_{in}'	k_{in}''	csomag	leírás
14565	11866	10748	libc6	C standard könyvtár
4056	3319	2803	libstdc++6	C-fordító könyvtárai
3779	3548	2970	libgcc1	Libraries of C compiler
3397	2229	1842	perl	Perl programnyelv
2827	1595	1122	python	Python programnyelv
2684	2243	1818	libglib2.0-0	GLib könyvtár
1571	2170	3172	libx11-6	A grafikus felület könyvtára
1457	947	945	dpkg	deb csomagkezelő
1435	1571	1275	libgtk2.0-0	GTK+ grafikus felhasználói felület könyvtára
1217	553	108	libqtcore4 ¹	Qt 4 magmodul
1189	2399	1963	zlib1g	tömörítő könyvtár
1018	666	447	python-support	a Python-modulok kezelésére

6.4. táblázat. A legnagyobb kifokszámú csomagok ($k_{\text{out}} \geq 90$) 2012. július 11-én. k_{out} : kifokszám 2012. július 11-én, k_{out}' : kifokszám 2009. november 3-án. Ahol a k_{out}' helyén kötőjel (–) van, ott korábban nem volt azonos nevű csomag.

k_{out}	k_{out}'	csomag	leírás
210	–	ubuntu-sugar-remix	Oktatóprogramok
200	113	ubuntu-desktop	Ubuntu a GNOME asztali környezethez
199	74	xubuntu-desktop	Ubuntu az XFCE asztali környezethez
176	111	ubuntustudio-desktop	Multimédiakészítő-eszközök
176	66	kubuntu-desktop	Ubuntu a KDE asztali környezethez
144	133	ichthux-desktop	Keresztény asztali környezet
130	–	libmono-cil-dev	Fejlesztőfájlok Monohoz
119	–	kubuntu-full	A Kubuntuhoz tartozó összes csomag
106	103	texlive-full	Az összes TeX Live csomag telepítéséhez
102	2	med-bio	Biológiával kapcsolatos csomagok
97	83	libjifty-perl	Programkönyvtár a Jifty webkeretrendszerhez
90	–	lubuntu-desktop	Ubuntu az LXCE asztali környezethez



6.12. ábra. A csoporterősség fokszámfüggése a vizsgált szoftvercsomag-hálózatban (pontok). A szaggatott vonal a -1 kitevőjű hatványfüggvény.

Látható, hogy a két modellben 2–3 nagyságrenddel kisebb értékek adódnak a csoporterősségi együtthatóra. A hierarchikus modell $C_h = 0,743$ -as csoporterősségi együtthatójával [32] a mi hálózatunké nagyságrendileg egyezik.

Az egyes fokszámokra az adott fokszámú csúcsok átlagos csoporterősségi együtthatóját ábrázoltuk fokszám szerint a 6.12. ábrán. A szaggatott vonal a -1 -es kitevőjű hatványfüggvénynek felel meg. Az eloszlás nem mond ellent a más hálózatokban tapasztalt fokszámfüggésnek, a szoftvercsomag-hálózat hierarchikusnak tekinthető.

A 2012-es cikkemben már nem csak a függőségekről és valódi csomagokról, hanem az ajánlások, szolgáltatások számáról, és a virtuális csomagokról is tudtam adatokat írni [56]. Az Ubuntu 11.10-es verziója 2012. június 15-én 36685 (88,7 %) valódi csomaggal és 4686 (11,3 %) virtuális csomaggal rendelkezett. Ezek között 170185 kapcsolat volt: 150249 (88,3 %) függőség, 13077 (7,7 %) ajánlás és 6859 (4,0 %) szolgáltatás. A hálózatban voltak többszörös élek: 566 élt (0,3 %) szükséges eltávolítani ahhoz, hogy ne legyenek többszörös élek. Jellemző eset a többszörös élek esetén, hogy egy csomag függ egy másiktól, de újabb változatát ajánlja. Amennyiben csak a valódi csomagokat és függőségeket hagytam meg,

nem volt benne többszörös él.

7. fejezet

Adott tulajdonságú hálózatok létrehozása

A multifraktál alapú hálózatgeneráló módszer (MFNG) lehetőséget ad arra, hogy lényegében tetszőleges előre megadott tulajdonságeggyüttessel létrehozzunk hálózatot. Adott fokszámsorozathoz egyszerűen lehet hálózatot készíteni, amelynek az a fokszámsorozata. (Természetesen csak akkor, ha van ilyen hálózat.) Egymás után kell ugyanis kisorsolni az egyes élek végpontjait, azok közül a csúcsok közül, amelyek még nem rendelkeznek a megadott fokszámmal. Ha már minden csúcs a kellő fokszámmal rendelkezik, akkor az algoritmus végetér. Adott fokszámeloszlású hálózat készítése tehát egyszerű. Jelentősen nehezebb azonban olyan hálózat készítése, amelynek nem csak a fokszámeloszlására szabunk ki feltételt, hanem a csoportterősségi együttthatójára is megszabunk valamilyen értéket. Az MFNG ígéretes módszer arra, hogy ilyen tulajdonságeggyüttésekkel rendelkező hálózatok létrehozása is lehetővé váljon. Először adott átlagfokszámot, majd adott fokszám-eloszlást céloztam meg.

A módszert Palla Gergely, Vicsek Tamás és Lovász László dolgozták ki [57]. A szerzők megvalósították C++ nyelven. Lehetőségem volt a lefordított fájlt kipróbálni, azonban csak megadott céltulajdonság generálására volt vele lehetőség, másféleképp kicsit átírva újra kellett volna fordítaniuk. Hogy ne legyek zárt forráskódú programhoz kötve, ezért valósítottam meg Python nyelven az MFNG módszert, és alkalmaztam adott fokszámeloszlású hálózatok létrehozására [58, 59].

7.1. A multifraktál-hálózatgeneráló módszer áttekintése

A multifraktálon alapuló hálózatgeneráló algoritmusban egy generáló mérték paramétereit változtatgatjuk úgy, hogy a hozzá tartozó hálózat tulajdonságai egyre közelebb kerüljenek egy adott tulajdonsághoz. A 7.1. ábrán követhető a folyamat, hogyan kaphatjuk meg a generáló mértékből a hálózatot. Először iterációval előállítjuk az élvalószínűségi mértéket (1-es nyíl az ábrán), majd az élvalószínűségi mérték felhasználásával létrehozuk a hálózatot (2-es nyíl). Az első lépés teljesen determinisztikus, azaz mindig ugyanazt az élvalószínűségi mértéket kapjuk, ha viszont több hálózatot hozunk létre az élvalószínűségi mértékből, azok tulajdonságai eltérhetnek egymásétól.

Mind a generáló mérték, mind az élvalószínűségi mérték egy úgynevezett *valószínűségi mérték*, csak a generálásban betöltött szerepük különbözik. Módszereinkhöz a valószínűségi mértéket az $[0, 1] \times [0, 1]$ egységnégyzeten értelmezzük a következőképpen. Mind az x , mind az y tengely $[0, 1]$ intervallumát felosztjuk m nem szükségképpen egyenlő részre. Az osztópontokat a következő módon jelöljük a továbbiakban:

$$d_0 = 0, d_1, d_2, \dots, d_{m-1}, d_m = 1 \quad (7.1)$$

A két tengelyen ugyanazokat az osztópontokat választjuk az egyszerűség kedvéért. Ez a felosztás az egységnégyzetet m^2 téglalapra osztja.

Minden egyes résztéglalaphoz egy p_{ij} valószínűséget rendelünk szimmetrikusan,

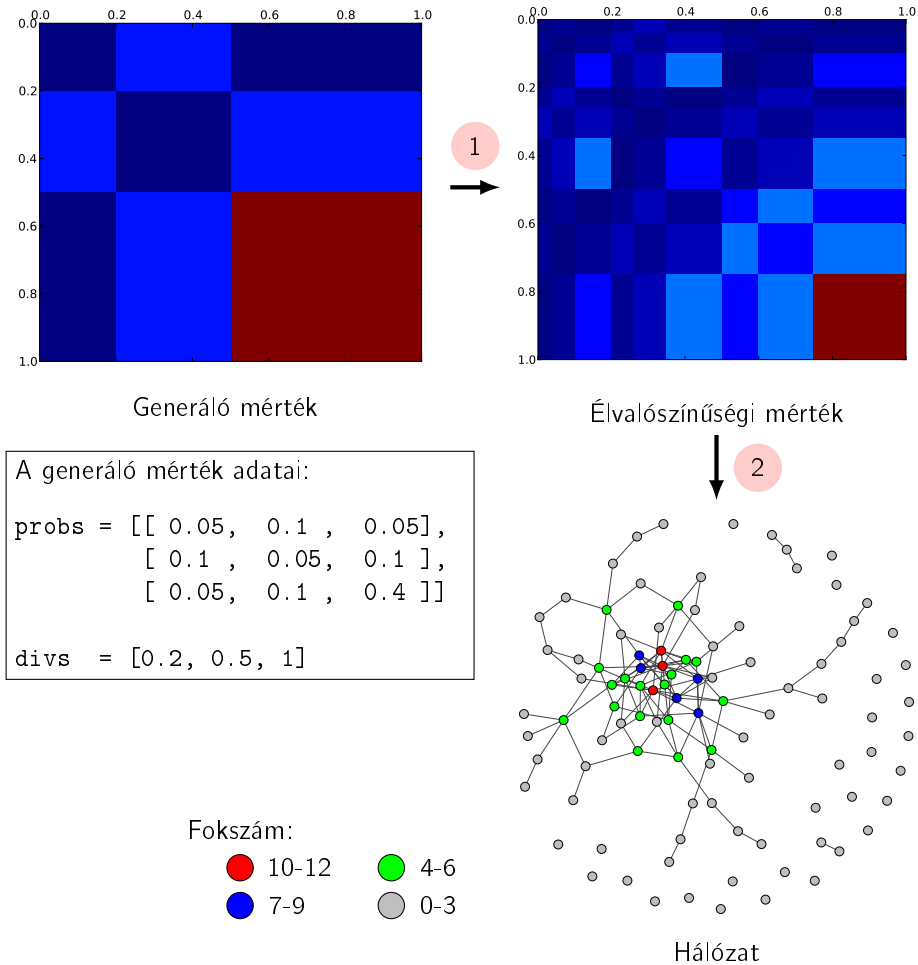
$$p_{ij} = p_{ji}, \quad \sum_{i,j=0}^{m-1} p_{ij} = 1. \quad (7.2)$$

Az origóhoz legközelebbi valószínűséget jelöljük p_{00} -al, a szemközti sarokban lévőt pedig $p_{m-1,m-1}$ -gyel.

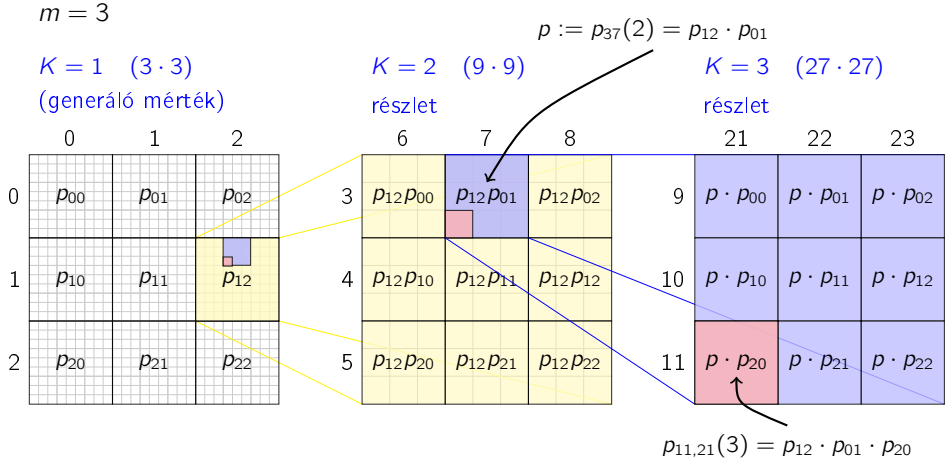
A Python nyelv és annak programcsomagunk által használt numpy csomagja az indexelést 0-ával kezdi, tehát a mátrix „bal felső” elemének mindkét indexe nulla. Ezt a jelölésmódot használom az értekezés többi képletében is.

7.2. A valószínűségi mérték iterálása

Mint már említettem, a hálózatot nem közvetlenül a generáló mértékből hozzuk létre, hanem előbb iterálással létrehozuk az élvalószínűségi mértéket. Ebben a fejezetben részletesen leírom, mit jelent az iterálás.



7.1. ábra. A hálózat létrehozása egy generáló mértékből: Balra egy generáló mérték látható 0,2 és 0,5 osztópontokkal, alatta a pontos adataival. Jobbra fent az iterált mérték látható, amelyet kapcsolódási valószínűségi mértékként használva, a jobbra lent található hálózatot kapjuk. Jól látható, hogy a generáló mérték minden egyes téglalapja helyett az iteráltban 3·3 téglalap van. A hálózat csúcsait a fokszám alapján színeztem.



7.2. ábra. Az iterált mérték valószínűségeinek kiszámítása $m = 3$, $K = 3$ esetre.

A generáló mérték K -dik iterációja során szintén téglalapokra osztjuk fel az egységnégyzetet, ahogy a generáló mértéknél, de $m^K \cdot m^K$ téglalapra. Definíció szerint az első iteráláskor ($K = 1$) magát a generáló mértéket kapjuk.

A $K > 1$ esetben az osztópontokat úgy kapjuk, hogy a $K - 1$ -dik iteráltban minden egyes szakaszt továbbosztunk m részre oly módon, hogy a részintervallumok hosszai arányosak legyenek az eredeti generáló mérték intervallumainak hosszával.

A K -dik iterált $p_{ij}(K)$ valószínűségeit következőképp lehet kiszámítani

$$p_{ij}(K) = \prod_{q=1}^K p_{i_q j_q}, \quad (7.3)$$

ahol

$$i_q = \left\lfloor \frac{i \bmod m^{K-q+1}}{m^{K-q}} \right\rfloor. \quad (7.4)$$

A $(i \bmod d)$ jelölés az i/d egész osztás maradékát jelöli, $\lfloor x \rfloor$ pedig az x egész részét (legnagyobb egész számot, ami nála kisebb). A 7.4 egyenlethez hasonló igaz j_q indexekre is, csupán minden i -t j -re kell cserélni.

Hogy jobban megértsük, mi is történik az iterálás során, a továbbiakban úgy érdemes a K -dik iteráltra tekinteni, mint ami a $K = 1$ -es iteráltból, azaz a generáló

mértékből $K-1$ iterációs lépésben keletkezik. A továbbiakban először egy iterációs lépést veszünk szemügyre. Az $m = 3$, $K = 3$ esetet a 7.2. ábra mutatja be, a továbbiakban általánosságban elmondottakat konkrét esetre ezen az ábrán lehet követni.

A K -dik iteráció valószínűségi mátrixát úgy kapom a $K-1$ -dik iteráltéból, hogy a $K-1$ -dik iterált mátrixának minden eleme helyett $m \cdot m$ elemet írok, melynek következtében a sorok és az oszlopok száma is m -szeresére változik a következő módon.

A $K-1$ -dik iterált i, j -dik eleme a korábbi jelöléssel $p_{ij}(K-1)$. Az összes létező i és j indexekre $p_{ij}(K-1)$ helyére a következő elemeket írom:

$$\begin{array}{cccc}
 p_{ij}(K-1)p_{00} & p_{ij}(K-1)p_{01} & \dots & p_{ij}(K-1)p_{0,m-1} \\
 p_{ij}(K-1)p_{10} & p_{ij}(K-1)p_{11} & \dots & p_{ij}(K-1)p_{1,m-1} \\
 \vdots & \vdots & \ddots & \vdots \\
 p_{ij}(K-1)p_{m-1,0} & p_{ij}(K-1)p_{m-1,1} & \dots & p_{ij}(K-1)p_{m-1,m-1}
 \end{array} \quad (7.5)$$

Ahelyett, hogy igazolnám, hogy a (7.4) képlet alkalmazásakor tényleg a fenti módon helyettesítődnek az elemek, ha a $K-1$ -dik iterációról a K -ra lépünk, csupán a már említett $m = 3$, $K = 3$ eset egyik mátrixelemére a $p_{11,21}(3)$ -ra mutatom be az egyezést. Ekkor $i = 11$ indexre a képletből a következő értékek jönnek ki a szorzatbeli indexekre:

$$\begin{aligned}
 i_1 &= \left\lfloor \frac{11 \bmod 3^3}{3^2} \right\rfloor = \left\lfloor \frac{11}{9} \right\rfloor = 1 \\
 i_2 &= \left\lfloor \frac{11 \bmod 3^2}{3} \right\rfloor = \left\lfloor \frac{2}{3} \right\rfloor = 0 \\
 i_3 &= \left\lfloor \frac{11 \bmod 3}{1} \right\rfloor = \left\lfloor \frac{2}{1} \right\rfloor = 2
 \end{aligned} \quad (7.6)$$

$j = 21$ indexre pedig:

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
i_1	0									1									2								
i_2	0			1			2			0	1	2			0			1			2						
i_3	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2

$$i = 11 \Rightarrow i_1 = 1, i_2 = 0, i_3 = 2$$

$$p_{11,21} = p_{12} \cdot p_{01} \cdot p_{20}$$

$$j = 21 \Rightarrow j_1 = 2, j_2 = 1, j_3 = 0$$

7.3. ábra. Az i_1, i_2 és i_3 értékek alakulása $m = 3, K = 3$ esetben a 7.4. képletben különböző i értékek esetén. Az $i = 11$ és $i = 21$ eset színezéssel kiemelve.

$$\begin{aligned}
 j_1 &= \left\lfloor \frac{21 \bmod 3^3}{3^2} \right\rfloor = \left\lfloor \frac{21}{9} \right\rfloor = 2 \\
 j_2 &= \left\lfloor \frac{21 \bmod 3^2}{3} \right\rfloor = \left\lfloor \frac{3}{3} \right\rfloor = 1 \\
 j_3 &= \left\lfloor \frac{21 \bmod 3}{1} \right\rfloor = \left\lfloor \frac{0}{1} \right\rfloor = 0
 \end{aligned} \tag{7.7}$$

Így a keresett valószínűségimátrix-elem:

$$p_{11,21}(3) = p_{i_1 j_1} \cdot p_{i_2 j_2} \cdot p_{i_3 j_3} = p_{12} \cdot p_{01} \cdot p_{20} \tag{7.8}$$

Ezeket az eredményeket összehasonlítva a 7.2. és a 7.3. ábrákkal egyezést tapasztalunk.

Miután megértettük, mi történik, lássunk egy teljes indukciós bizonyítást arra, hogy a valószínűségek összege valóban egy marad az iteráció során. Mivel az első iterált maga a generáló mérték, és az K index nélküli p_{ij} jelölést a generáló mérték valószínűségi mátrixára tartottuk fenn, ezért $p_{ij} = p_{ij}(1)$, emiatt az állítás $K = 1$ esetre (7.2) felhasználásával nyilvánvalóan igaz:

$$\sum_{i=0}^{m-1} \sum_{j=0}^{m-1} p_{ij}(1) = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} p_{ij} = 1 \tag{7.9}$$

Igazoljuk, hogyha a fenti állítás $K = n$ -re igaz, abból következik az is, hogy $K = n + 1$ -re is igaz. Ha ez sikerül, akkor a teljes indukció szerint minden pozitív egész számra igaz lesz. Tegyük fel tehát, hogy $K = n$ -re igaz az állítás:

$$\sum_{i=0}^{m^n-1} \sum_{j=0}^{m^n-1} p_{ij}(n) = 1 \quad (7.10)$$

Ebből $K = n + 1$ -re is következik az állítás, hiszen

$$\begin{aligned} \sum_{i=0}^{m^{n+1}-1} \sum_{j=0}^{m^{n+1}-1} p_{ij}(n+1) &= \sum_{i'=0}^{m^n-1} \sum_{j'=0}^{m^n-1} \left(\sum_{i=0}^{m-1} \sum_{j=0}^{m-1} p_{ij} \cdot p_{i'j'}(n) \right) = \\ &= \sum_{i'=0}^{m^n-1} \sum_{j'=0}^{m^n-1} p_{i'j'}(n) \underbrace{\left(\sum_{i=0}^{m-1} \sum_{j=0}^{m-1} p_{ij} \right)}_1 = 1 \end{aligned} \quad (7.11)$$

Az utolsó átalakítást a (7.2) összefüggés és a (7.10) indukciós feltevés teszi lehetővé.

7.3. A generáló mérték optimalizálása

Az optimalizáláshoz szükség van egy energiafüggvényre (egy nemnegatív függvényre), amely a generáló mérték jóságát méri, azaz azt, hogy mennyire van közel a belőle az élvalószínűségi mértéken keresztül előállítható hálózat tulajdonsága a megcélzott tulajdonsághoz.

Szükséges, hogy eldöntsük az intervallumok számának m értékét, és az iterációk számának K értékét, valamint a hálózat csúcsainak N számát. Az általam írt program kezdetben egyforma hosszúságú intervallumokra osztja a tengelyeket, és egyforma valószínűségértékeket használ. A kezdeti generáló mértékhez tartozó hálózat tehát egy Erdős–Rényi-modellből származó hálózat, amelynél a csúcsok száma N , az élvalószínűség pedig

$$p = \frac{1}{m^K \cdot m^K} = m^{-2K}, \quad (7.12)$$

tehát a kiinduló foksámeloszlás Poisson-eloszlás.

A program minden lépésben vagy áthelyez egy osztópontot, vagy megváltoztat egy valószínűséget és végigoszt a valószínűségek összegével, hogy a valószínűségek összege továbbra is 1 maradjon. Ezt a kétféle változtatást váltogatja: egyik után mindig a másik jön.

Minden lépés után megvizsgáljuk a generáló mértékhez tartozó új E^* energiaértéket. Ha ez kisebb, mint az utoljára elfogadott generáló mérték E energiája, akkor az új generáló mérték adatait és energiáját megjegyezzük, és ezzel folytatjuk az optimalizálást. Azért, hogy ne maradjunk benne helyi energiaminimumokban, akkor sem mindig utasítjuk el az új generáló mértéket, ha annak energiája nagyobb, mint az utoljára elfogadott E energia, hanem

$$P(T) = \exp\left(-\frac{E^* - E}{T}\right), \quad (7.13)$$

valószínűséggel elfogadjuk, és $1 - P(T)$ valószínűséggel elutasítjuk. A $T > 0$ paraméter játssza a (energia egységeken mért) hőmérséklet szerepét. Ha a hőmérsékletet az optimalizálás során lassan csökkentjük, az lehetővé teszi, hogy az generáló mértéknek sikerül kiszabadulnia a helyi energiaminimumokból. Minél kisebb a hőmérséklet, annál több változás kerül elutasításra, és a hálózat tulajdonságai egyre jobban közelítenek az előírt tulajdonsághoz.

A programban beállítható a kezdeti T_0 hőmérséklet, az annál kisebb T_{limit} hőmérséklet-határ, amelynél az optimalizálás megszakad, és a lépések száma. Ebből kiszámítja azt az egynél kisebb értéket, amennyivel két lépésenként szoroznia kell a hőmérsékletet, hogy az előírt lépésszámmal elérjünk az előírt hőmérséklet-határhoz. A program minden hőmérséklet-értéken két változtatást csinál, egyik esetben osztópontot módosít, a másik esetben valószínűségeket.

A futtatás során csak a generáló mérték iterációját kell elvégezni, mert az abból létrehozható hálózatok számszerűsíthető tulajdonságának várható értékei kiszámíthatóak anélkül, hogy hálózatot hoznánk létre. Az iteráció és az energiaérték kiszámítása sok számolással jár, de jelentősen kevesebb futásidőt igényel, mint ha hálózatot hoznánk létre, és annak a tulajdonságait vizsgálnánk. Van továbbá egy még fontosabb érv is, hogy ne hozzuk létre a hálózatot. Ha az élvalószínűségi mértékből hálózatot hozunk létre, akkor annak számszerűsíthető tulajdonságai egy adott értékhez – a várható értékhez – közel vannak, de általában nem egyeznek azzal. Ha egy hálózatot generálok az élvalószínűségi mértékből, akkor annak a tulajdonsága lehet, hogy közelebb lesz a megcélzott értékhez, mint az eddig elfogadott legjobb élvalószínűségi mértékhez tartozó érték, annak ellenére, hogy az új élvalószínűségi mérték nem jobb, mint a korábban elfogadott. Két élvalószínűségi mérték közül azt nevezem jobbnak, amelyhez tartozó hálózatok tulajdonságainak várható értéke közelebb esik a megcélzott tulajdonsághoz.

7.4. A vizsgált céltulajdonságok

Mielőtt a vizsgált céltulajdonságokra rátérhetnénk, szükséges megemlíteni, hogyan lehet az élvalószínűségi mértékből kiszámítani a belőle létrehozott hálózatok várható átlagos fokszámát és várható fokszámeloszlását.

A hálózat $p(k)$ fokszámeloszlása egy olyan függvény, amely a k fokszámokon értelmezett, és megadja azt a valószínűséget, hogy egy véletlenül kiválasztott csúcs fokszáma k . Ezt a következőképpen lehet kiszámítani a K iterációval kapott élvalószínűségi mérték $p_{ij}(K)$ valószínűségeiből és d_i osztópontjaiból [57]:

$$p(k) = \sum_{i=0}^{m^K} p_i(k) \ell_i, \quad (7.14)$$

ahol

$$p_i(k) = \frac{\langle k_i \rangle^k}{k!} e^{-\langle k_i \rangle}, \quad (7.15)$$

$$\langle k_i \rangle = N \sum_{j=0}^{m^K-1} p_{ij}(K) \ell_j, \quad (7.16)$$

A $\langle k_i \rangle$ érték az i -dik intervallumhoz tartozó csúcs fokszámának várható értéket, az $\ell_i = d_{j+1} - d_j$ pedig a j -dik intervallum hossza.

A program futtatása során eleinte adott átlagfokszámot céloztam meg. Az átlagfokszám kiszámítható a 7.16. képletből, az energiafüggvényt pedig a következőképpen definiálhatjuk:

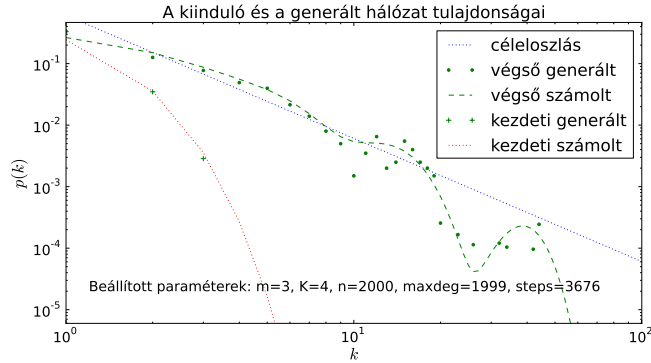
$$E = \frac{|\langle k \rangle^* - \langle k \rangle|}{\langle k \rangle}, \quad (7.17)$$

ahol $\langle k \rangle^*$ a pillanatnyi átlagfokszám, a $\langle k \rangle$ pedig a célba vett átlagfokszám.

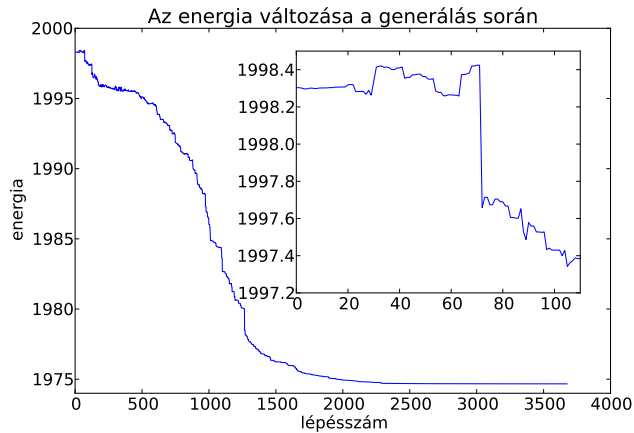
A későbbiekben adott (többnyire hatványfüggvénnyel leírható) fokszámeloszlást céloztam meg

$$E = \sum_{k=k_{min}}^{k_{max}} \frac{|p^*(k) - p(k)|}{p(k)}, \quad (7.18)$$

ahol $p^*(k)$ a pillanatnyi élvalószínűségi mérték, a, a $p(k)$ pedig a célba vett fokszámeloszlás. Az optimalizációhoz a szokásos értékek (kezdeti hőmérséklet, hőmérséklet-határ, lépésszám) mellett meg kell adni a megcélzott függvényt, valamint a minimális és maximális fokszámot, amelyek közötti fokszámokra összehasonlítjuk a két fokszámeloszlást.

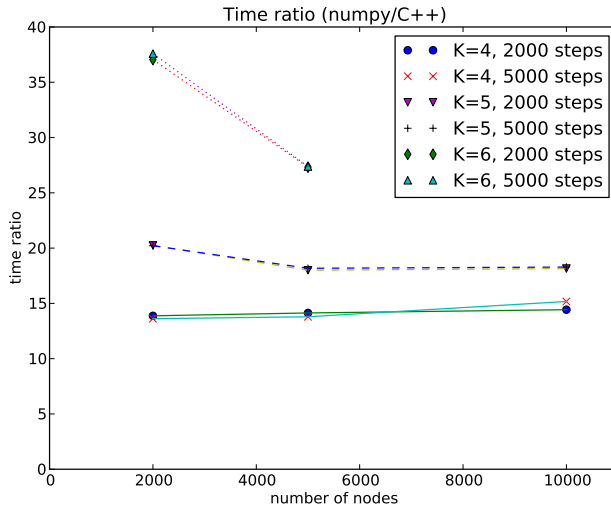


(a) Fokszámeloszlások

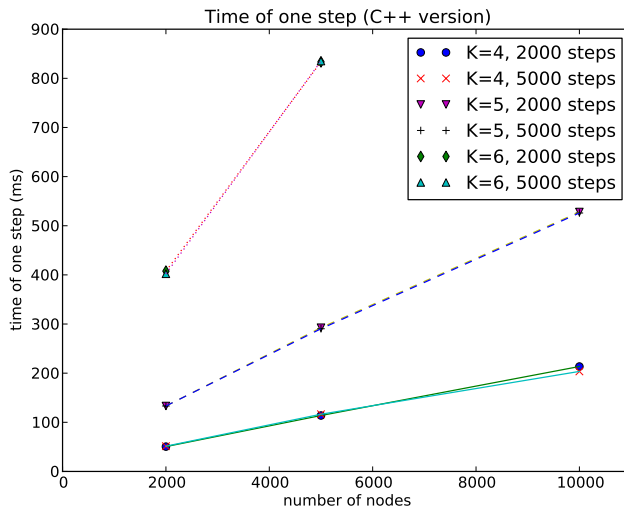


(b) Az energia változása

7.4. ábra. Egy 2000 csúccsal rendelkező hálózathoz generálásához tartozó grafikonok. A megcélzott fokszámeloszlás egy -2 -es kitevőjű hatványfüggvény volt. Az (a) részában a megcélzott fokszámeloszlás kék pontozott vonallal látható, a kiinduló generáló mértékből számolt fokszámeloszlás piros pontozott vonallal, a legutoljára elfogadott generáló mértékhez tartozó fokszámeloszlás pedig szaggatott vonallal. A köröcskével jelölt eloszlásértékek az utoljára elfogadott mértékből generált hálózathoz tartozó eloszlás szükség szerinti bineléssel. A (b) részában az utoljára elfogadott mérték energiaértékei láthatóak az optimalizáció során a lépésszám függvényében. Ezen az ábrán a kezdeti változások felnagyítva is láthatóak. Jól látszik, hogy az energia nem csak csökkenhet az optimalizáció során.



(a) A tisztán Pythonban írt, és a C++-kódot is tartalmazó kódok futásiidejének a hányadosa



(b) A C++-kódot is tartalmazó változat egy lépésre eső futásiideje

7.5. ábra. Az MFNG hálózatgeneráló program két változatának futási idejét jellemző grafikonok

A programnak azt a részét, amelynek a futása a Python változatban a legtöbb időt vett igénybe C++ nyelven újraírtuk, és a Python verzióhoz illesztettük. A 7.5a. ábrán jól látható, hogy a vizsgált esetekben több mint tízszeresére nőtt a futási sebesség. A növekedés különösen nagy K iterációs szám esetén jelentős. A tényleges egy lépésre eső futási időket a C++-kódot is tartalmazó esetben a 7.5b. ábra mutatja.

Tapasztalataim szerint az MFNG változatos hálózatok előállítására alkalmas, azonban nehezen közelít a kezdeti foksámeloszlástól távoli eloszlásokat. A 7.4a. ábrán látható, hogy az eredeti foksámeloszláshoz képest sokkal közelebb került a megcélzott eloszláshoz. Ehhez hasonló közelítést már a kisebb iterációs és lépésszám esetén is sikerült elérnünk. A sebesség növelése lehetővé tette, hogy nagyobb iterációs számot és több lépést használjunk, azonban ezek nem javítottak jelentősen a közelítés minőségén.

8. fejezet

Az összetett hálózatok oktatása

Munkahelyemen, az Óbudai Egyetemen kidolgoztam műszaki informatikus hallgatóinknak egy tantárgyat „Összetett hálózatok vizsgálata” címmel. Ezt 2012 őszén már másodjára tanítottam. A tantárgyat úgy terveztem, hogy számítógépes laborban lehessen tartani, ezért komoly időt igényelt a fejlesztéshez használt eszközök elsajátítása. Az informatikus hallgatók korábban C# nyelvet tanultak, tehát ismerős nekik az objektumorientált programfejlesztési paradigma. A beadandó feladatoknál is gyakran ez alapján dolgozunk, már meglevő osztályokból származtatnak osztályokat a hallgatók, amelyhez új tagfüggvényeket hoznak létre. A feladatok része az is, hogy a tagfüggvények működésének tesztelésére teszteket (unit test) írjanak. Az órán gyakran használjuk az ipython interaktív Python-parancsértelmezőt, amellyel a hálózatok fontosabb tulajdonságainak vizsgálatát, a függvények és hálózatok kirajzoltatását könnyen el tudjuk végezni.

8.1. A félév beosztása és követelményrendszere

Az egyes hetek tervezett témái a következők voltak:

1. A Python-nyelv, verziói, alapértelmezett shell és ipython, range függvény, szeletek, for ciklus, szöveges fájlok kezelése.

A Vim használata: Python-szkriptek írása Vim-mel. A `.vimrc` fájl. Snip-Mate

2. Python: logikai hamis értékek, listával dolgozó függvények (`all`, `any`, `max`, `min`, `sum`), függvénydefiniálás, egyszerűbb tesztek írása.
3. A hálózatok vizsgálata és kirajzolása a Python `igraph` moduljával. Gráfelméleti alapfogalmak ismételése (egyszerű gráf, teljes gráf fokszáma, komponens, fa élszáma, kapcsolat az átlagfokszám, a csúcsszám és élszám között).

Megjegyzés: A diákok a harmadik-negyedik héten egy-egy beadandó feladatot kapnak. Például meg kell tudni határozni egy hálózatban a komponensek számát, létrehozni szöveges fájlban tárolt éllistából hálózatot.

4. A hálózatok kirajzolásának finomhangolása, csúcs- és élattribútumok, fokszám, szomszédok, adott tulajdonságú csúcsok és élek keresése

Osztályok létrehozása a Pythonban, tesztek létrehozása a `unittest` modullal

Megjegyzés: Az `igraph`-ban a hálózatok kirajzolásának tulajdonságait (szín, méret) él- és csúcstarttribútumokkal szabályozhatjuk, de ezek szolgálhatnak a csúcsok és élek adatainak, élek súlyának tárolására is.

5. Az Internet, mint hálózat vizsgálata (csúcs- és élszám, átlagfokszám, maximális és minimális fokszám, átmérő)

Függvényábrázolás Pylabbal. `linspace`, `plot`, `semilog(y|x)`, `loglog`, `legend`, `title`, `xlabel`, `ylabel`, `grid`, TeX-es feliratok. Grafikonok mentése.

Megjegyzések: A pylab függvényei nagyjából megegyeznek a MATLAB függvényeivel. Az itt tanultakkal együtt tanuljuk meg azt is, hogy hogyan érdemes a függvényeket ábrázolni, hogy észrevegyük azok hatványfüggvény, illetve exponenciális jellegét.

Ezen a héten már elkezdődik a beadandó feladatok ellenőrzése is. Amennyiben a kurzus létszáma megengedi, a hallgatók minden egyes feladat megoldását és tapasztalatait röviden bemutatják az órán.

6. Az Erdős–Rényi modell (véletlen gráf), a véletlen gráfok komponenseinek vizsgálata.

Adott hálózatával egyező csúcsszámú és élszámú véletlen hálózat generálása.

7. A Price-modell és a Barabási-Albert modell. Az összetett hálózatok fokszámeloszlása és összegzett fokszámeloszlása. A fokszámeloszlás létrehozása és ábrázolása.
8. A szoftvercsomag-hálózat vizsgálata. Összehasonlítása az eddigi modellekkel. Az összegzett fokszámeloszlás. Skálafüggetlen hálózatok.
9. Gráfelméleti adatstruktúrák (szomszédossági és illeszkedési mátrix, szomszédossági tömb, láncolt szomszédossági lista). Hálózatok tárolása fájlokban (éllista, graphml, gml, pickle, dot, gml).
10. A fokszámeloszlás simításának módszerei: a beosztás megváltoztatása (bin smearing). A kitevő becslése.
11. A multifraktál hálózatgenerátor (MFNG) elméleti háttere és használata.
12. A csoportterösségi együttható, annak fokszámfüggése. A hierarchikus modell.
13. A hálózatok sérülékenysége: véletlen hibák és célzott támadás.
14. Áttekintés, értékelés.

A 4. héten a Python programnyelv alapjait és a Vim szövegszerkesztő használatát kértem számon, két alkalommal az összetett hálózatok elméletéből írtam teszteket a Moodle elektronikus oktatási keretrendszeren, valamint egy beadandó feladatot kellett bemutatni.

Igyekeztem mennél több olyan beadandó feladatot is kitalálni, hogy lehetőleg a beadandók eredményei azelőtt mutassanak valamilyen, a hálózatok elméletében ismert eredményt, mielőtt tanórán kimondanánk az arra vonatkozó állítást. Egyik ilyen lehetőség az Erdős–Rényi modellben az p élvalószínűség növelésekor létrejövő fázisátmenet helyének vizsgálata, amelynél a sok különböző kis hálózat kezd egy nagygyá összeállni [60]. Meg lehet vizsgálni, hogy hogyan függ a csúcsok számától a fázisátmenet helye, valamint milyen gyorsan lépnek fel változások az egyes hálózat-jellemzőkben, mint amilyen a legnagyobb komponens relatív mérete a teljes hálózathoz viszonyítva, a komponensek száma, vagy az átmérő.

Egy másik érdekes feladat annak a vizsgálata, hogy mennyiben különbözik az Erdős–Rényi modellből és a Barabási–Albert modellből származó hálózatok ellenálló képessége véletlen hibákkal és célzott támadásokkal szemben [61].

Informatikus hallgatók számára fontosnak tartom gráfokkal kapcsolatos alapvető függvények megírását még akkor is, ha azt már az igraph szerzői többnyire C

nyelven létrehozták. Ilyenek az egyes hálózatmodelleket megvalósító függvények, vagy egy hálózat adott komponenseihez tartozó csúcsok megkeresése, amelyhez csak azt a függvényt használhatjuk fel, amely egy csúcs szomszédaival tér vissza.

A beadandó feladatnál a hallgatók gyakran párokban dolgoztak, főként olyan feladatoknál, amelyek két jól elkülöníthető részre tagolódtak.

A tesztelés számonkérésnél kihasználtam a Moodle-tesztfeladatok fajtáinak sokszínűségét [62], így a szokásos feleletválasztós tesztek mellett számításos feladatok és esszékérdések is gyakran előfordultak. A Moodle-rendszer arra is lehetőséget adott, hogy a segédanyagokat, a tárgyhoz készített kis videóimat átlátható formában egy helyről tudják elérni a hallgatók, illetve a beadandó feladatokat feltöltsék oda a hallgatók.

A félév vége felé a hallgatók képesek voltak egyszerűbb programok megírására, az `igraph`, a `cxnet` és a `pylab` szoftvercsomagok használatával, a kapott eredményeket grafikon formájában elmenteni.

8.2. Segédanyagok

A tantárgyhoz többféle segédanyag tartozik. A tárgyhoz írtam egy segédletet, amely témánként tartalmazza a téma során feldolgozandó elméleti fogalmakat, a programozáshoz szükséges ismereteket, valamint a beadandó feladatokat esetlegesen némely fogalom részletesebb kifejtésével.

Az általam kialakított összetett hálózatok weboldalról [63] elérhető a fenti segédlet, és több olyan, általam felvett videóanyag, amely bemutatja, hogyan lehet a hálózatok vizsgálatát elvégezni a Python programnyelv programcsomagjaival segítségével. Ezek olyan képernyőfelvételek, amelyet hangalámondással kísérek magyarázatként. A felhasznált utasítások mindegyiknél egy külön Python-forrásfájlban megtalálhatóak. Ezeket a videókat a hallgatók egy része megnézte, és hasznosnak találta.

8.3. Összetett hálózatok az informatikaversenyen

Az Óbudai Egyetem székesfehérvári Alba Regia Egyetemi Kara 2012 őszén harmadszorra rendezte meg középiskolás diákok számára az Alba Regia Informatikaversenyt [64], amely két fordulóból áll, egy internetes selejtezőből, és a karunkon rendezett döntőből.

Mindhárom évben én alkottam meg az adatbázissal kapcsolatos feladatokat, és ezekben összetett hálózatok adatait használtam fel. A diákok számára készítettem két fájlt: az egyikben a névfeloldás volt, azaz a csúcsok sorszámahoz meg lehetett találni a hozzá tartozó nevet, a másikban pedig az élek voltak megadva a csúcsokat azonosító sorszámpárral megadva. Ezeket a szöveges fájlokat kellett beolvastatni a Microsoft Access vagy OpenOffice.org Base adatbázis-kezelő programokba, és kilistáztatni a legnagyobb fokszámú csúcsok neveit, létrehozni a fokszámgyakoriságokat, majd a fokszámgyakoriságokat táblázatkezelőbe importálva ábrázolni kellett azt az adatok jellegének megfelelő grafikonon. A feladatot az nehezítette, hogy a fokszám meghatározásához a kapcsolat mindkét oldala felől meg kellett keresni, hány másikkal kapcsolódik, és a két értéket összegezni.

Az eddigiek során a következő hálózatok szerepeltek a versenyen: a szoftvercsomag-hálózat; a Nyomorultak című regény szereplőinek hálózata az együtt előfordulással, mint éllel; az Egyesült Államok egyetemi focicsapatainak hálózata, ahol azok között volt él, akik egymással játszottak adott évadban; az asztrofizika kutatóinak hálózata, melyben akkor köt össze két kutatót él, ha az arxiv.org adatbázisában szerepel közösen írt cikkük.

A döntőbe került diákok az adatok beimportálásával a táblák létrehozását többnyire sikeresen megoldották, a további feladatok esetén viszont jelentős különbség volt a résztvevők pontszámaiban. A feladatok tehát jó képességű diákok számára megoldhatóak volt, de nem túlzottan egyszerűek.

9. fejezet

Összefoglalás és kitekintés

Doktori értekezésemben bemutattam az elméleti és kísérleti részecskefizika tanításával valamint az összetett hálózatok vizsgálatával és tanításával kapcsolatos eredményeimet.

A fizikai elméletek megértéséhez nagyfokú matematikai tudás kell. Ha középiskolások, vagy olyan hallgatók számára tanítunk részecskefizikát, akik nem fizikusnak készülnek, mindig komoly kihívás, hogy pontosan és mégis érthetően beszéljünk a témáról. Erre legalkalmasabbak a kísérletek, hiszen azok esetén hétköznapi ismert fogalmak (hely, mozgás, eltérülés a mágneses térben, elektromos jel) szerepelnek. A nemzetközi részecskefizikai diákműhely és a hozzá kapcsolódó eseménynézegető programok lehetőséget teremtenek arra, hogy középiskolási diákok megismerjék a részecskefizikai kutatásokat, általában a tudományos kutatás módszereit, a részecskefizikai jelenségek statisztikus jellegét, a nagyszámú és pontos mérés szerepét az eredmény bizonytalanságának csökkentésében, a mérési eredmények kiértékelésének módját, a tapasztalatok összevetésének szerepét, és azt, hogy az alapkutatások nélkül hosszabb távon nem lehetnek jelentős műszaki újítások sem. Az alapelveket megértve a diákok kedvet kaphatnak a téma mélyebb megismeréséhez, és fizikusként vagy más terület kutatóiként hasznosítani tudják az előbb felsorolt elveket és módszereket; döntéshozóként nagyobb rálátással tudnak dönteni kutatások támogatásáról, nemzetközi kutatóintézetekhez történő csatlakozásról, hétköznapi emberként pedig könnyebben el tudnak igazodni az igazi és az áltudományok között.

A diákműhely foglalkozások segítésére létrehoztam a Hands on CERN honlap magyar változatát CERN-sajátkezüleg néven, magyar feliratú ábrákat is készítet-

tem hozzá, kibővítettem a részecskefizika területén szerzett Nobel-díjak táblázatával és leírásával, illetve a téma diákoknak való irodalmával. Ezt a honlapot a nemzetközi részecskefizikai diákműhelyeken felhasználtuk. Az oldal tartalmaz egy eseménynézegetőt, amellyel a CERN DELPHI detektorában lezajlott események vizsgálhatóak meg. Az eseménynézegetővel a diákok meg tudják határozni, hogy milyen részecskék keletkeztek az elektron–pozitron-ütközés során, és az eseményeket csoportosítani tudják, meghatározva az egyes eseménytípusok elágazási arányát. Kidolgoztam egy olyan mérést, amellyel a honlapon mérhető elágazási arányokból az elemi részecskék családjainak számát számolhatják ki a hallgatók.

A részecskefizikai diákműhelyt először 2005-ben, a fizika évében szervezte meg az Európai Részecskefizikai Ismeretterjesztő Csoport. Magyarország kezdettől részt vesz minden tavasszal a diákműhelyen: minden évben a Debreceni Egyetem, a Részecske és Magfizikai Kutatóintézet és az Óbudai Egyetem székesfehérvári Alba Regia Egyetemi Központja vett és vesz részt remélhetően még sokáig a mérésekben. A székesfehérvári helyszínen én szervezem a diákműhely eseményeit a kezdetek óta, én tartom a bevezető előadásokat és végzem a mérés irányítását is. A 2013-as alkalom immár a kilencedik alkalom lesz, amikor húsz diák látogat el a központunkba, és ott CERN-es eseményeket elemeznek. A 2005-ös részecskefizikai diákműhelyen a diákokkal kérdőívet töltöttünk ki. A székesfehérvári és a debreceni helyszín válaszait feldolgozva arra következtethetünk, hogy a diákműhely érdekli a diákokat, új ismerettel gazdagodnak a nap során és a végén jobban érdekli őket a fizika, mint előtte.

A 2013-as diákműhely után öt héttel szerveztem a diákoknak egy előadást, melyen Horváth Dezsőtől hallhattak előadást a Higgs-bozonról. Ez lehetőséget teremtett arra is, hogy az előadás előtt kérdőívvel felmérjem, mennyi maradt meg a diákműhelyen hallottakból. A kérdőív 14 kérdést tartalmazott a részecskefizika elméletével és kísérletével kapcsolatban. Némelyik megválaszolásához tényekre kellett emlékezni, másikaknál összefüggéseket kellett érteni. Mindegyik válasz után 0 és 5 közötti skálán kellett értékelni két dolgot: mennyire biztos a válaszában, illetve mennyire tudta volna a választ a diákműhely előtt. A kérdőívek tanulsága szerint a diákműhely során új dolgokat tanultak meg, illetve egyes területeken biztosabbá vált a tudásuk.

A diákműhelyhez kidolgozott eseménynézegetők hozzáértő tanárok kezében akár középiskolai szakkörök munkaeszközüvé is válhat. Hozzáértő tanárok pedig egyre többen vannak a CERN által 1998 óta szervezett különböző tanárprogramok révén, amelyek 2006 óta már nem csak angol nyelven, hanem a tagországok anyanyelvén is elérhetőek, így tőlünk is minden évben negyven fizikatanár látogat

ki, és néznek körül a CERN-ben, ismerkednek meg az ott folyó kutatásokkal, azok elméleti és informatikai hátterével, valamint mérik meg a víz forráspontját a Mont Blanc szomszédságában található Aiguille de Midi csúcson, 3800 méter magasan.

Az összetett hálózatok kutatása fiatalabb terület, mint az anyag legalapvetőbb építőköveinek és kölcsönhatásaiknak a kutatása. Az Internet, a technika, az életudományok területén számos hálózat található, amelyek felépítésének és fejlődésének tanulmányozása hasznos technológiák kifejlesztését tette lehetővé eddig is, és vélhetően ezután is. Ilyen, teljesen vagy részben megvalósult, vagy reménytelinek látszó eredmények a fertőzések terjedésének előrejelzése, a munkahelyeken híreket hatékonyan terjesztő személyek megtalálása és felhasználása, az egyes fehérjék és más anyagok élő szervezetbeli szerepének felderítése, az agy működésének tisztázása, a telefonhálózatok jobb kihasználása, hatékonyabb Internetes protokollok kidolgozása, a programkódok minőségének számszerűsítése, a reklámok célcsoporthoz juttatása.

Hogy a hálózatokat könnyebben tanulmányozhassuk a hallgatóimmal, létrehoztam a cxnet hálózatvizsgáló programcsomagot. Ez Ubuntu és Debian Linux operációs rendszerek alatt képes rövid idő alatt létrehozni a tárolókban található szoftvercsomagok függőségi hálózatát, így helyben létrehozott, friss adatokat tanulmányozhatunk. A cxnet képes hálózatok fokszámoszlását ábrázolni különböző módokon: különböző binelési módszerekkel és összegzett fokszámoszlásként is, skálafüggetlen fokszámoszlások esetén képes a fokszámoszlás kitevőjének meghatározására is.

Tananyagot dolgoztam ki az Összetett hálózatok vizsgálata című, számítógépes laborban tartott kurzushoz alapképzésbeli mérnök informatikus szakos hallgatóknak. A tárgy keretében hallgatóim megismerkedtek az összetett hálózatok fontosabb jellemzőivel és modelljeivel. Megtanulták a hálózatok tulajdonságainak megállapítását és programok írását különböző összetett hálózatokkal kapcsolatos feladatok megoldására. Ezekhez a feladatokhoz a cxnet és az igrph programcsomagokat használtuk fel.

Részletesen elemeztem az Ubuntu Linux operációs rendszer szoftvercsomag-hálózatának tulajdonságait. Megállapítottam, hogy ez a hálózat skálafüggetlen, a fejlődésében szerepet játszik a népszerűségi kapcsolódás, hierarchikus és nagy csoportterősségi együttthatóval rendelkezik.

Az általam feltérképezett és más forrásból elérhető hálózatokat felhasználtam a munkahelyem által harmadik éve szervezett Alba Regia Informatika Verseny adatbázissal kapcsolatos feladatainak létrehozásához is. Ezeket a feladatokat a felkészültebb diákok sikeresen oldották meg.

Ahhoz, hogy tetszőleges tulajdonságegyüttessel rendelkező hálózatot hozhassak létre, optimalizáción alapuló eljárás tűnt megfelelőnek. Az egyik lehetőség a multifraktálokra alapuló hálózatgeneráló algoritmus (MFNG). Python programnyelven elsőként valósítottam meg az MFNG algoritmust. Adott átlagfokszámú, majd adott fokszámeloszlású hálózatok előállítását céloztam meg a programmal, amelyeket sikeresen közelített meg a szoftver. A Python programnyelven írt program futásának döntő részét kitevő számításokat átírtuk C++ nyelvre, amivel jelentős gyorsulást értünk el. Ötvözni tudtuk tehát a két nyelv előnyeit: a Python kényelmes használatát a rugalmas függvényparaméterezés és a magasszintű adatstruktúrái következtében, valamint a C++ nyelv sebességét.

Azért, hogy a részecskefizika és a hálózatok témájában könnyebben tudjanak tájékozódni a diákok és a hallgatóim, jelentősen bővítettem a Wikipedia magyar oldalát részecskefizikai, fizikatörténeti témákkal, fizikusok életével, egyetemekkel és kutatóintézetekkel kapcsolatos oldalakkal, a fizika portállal, összetett hálózatokkal kapcsolatos szócikkekkel, valamint ábrákkal. Ezek az oldalak máig látogatottak.

A munka az értekezéssel nem ért véget. Több lehetőség van az eddigi munkáim továbbfejlesztésére. A továbbiakban ezek közül sorolok fel pár lehetőséget.

A cxnet program – ahogy a 6.1. szakaszban említettem – képes olyan operációs rendszeren is futni, amelyen nincs grafikus felület. Ez lehetőséget adott nekem arra, hogy az egyik kiszolgálógépen, amely nem túl nagy leterheltségű web-kiszolgálóként is működik, beállítsam úgy, hogy az Ubuntu 12.04-es verziójának állapotáról óránként mentést készítsen, így 2012. július 12. és 2013. február 10. között már több mint 5000 pillanatfelvételt készített. A kapott adathalmaz nagy előnye, hogy egyetlen verzióról kapunk olyan információsort, amelyben a beállított tárolók nem változnak, így összehasonlíthatóak. Ez az adatsor lehetőséget ad számomra vagy a hallgatóim számára, hogy meghatározzuk ezen hálózatok esetén a változások törvényszerűségeit: tesztelhetjük, hogy itt is érvényes-e a népszerűségi kapcsolódás [65], illetve vannak-e más fontos törvényszerűségek a fejlődésben. Az eltűnő és megjelenő csomagok meghatározására és a megváltozott tulajdonságú csomagok felderítésére már elkezdtem írni egy programot. Mivel a már mentett hálózatok száma nagy, az összes esetén csak a betöltésük minimum 7 órát igényel, ezért a tervem az, hogy egyszer végighaladva a hálózatokon, a lehető legtöbb adatot nyerje ki a további vizsgálatokhoz.

A cxnet programcsomag szerveren futtatva lehetőséget teremthet olyan honlap létrehozására is, amelyen hallgatók vagy diákok hálózatokat vizsgálhatnak. A honlapot a Pythonban írt Django web-keretrendszer segítségével oldanám meg, amely lehetővé teszi összetett dinamikus honlapok viszonylag egyszerű fejlesztését.

Terveim között van egy ilyen honlap létrehozása a hozzá kapcsolódó középiskolai tananyag kidolgozásával együtt.

Az MFNG-ről szóló kiegészítő anyagai között megtalálható, hogyan lehet kiszámolni az élvalószínűségi mértékből a hálózat csoportterösségi együttthatóját [57, Supporting Information]. Az általam készített mfng Python-csomag lehetővé teszi, hogy egyszerre két tulajdonságot is megadjunk, ekkor a két tulajdonságból származó energiák összegével hajtja végre az optimalizálást. Például megcélozhatunk egy skálafüggetlen fokszámeloszlást és adott csoportterösségi együttthatót egyszerre. Így olyan lehetőséghez jutunk, amelyet nem lehet olyan egyszerű módszerekkel generálni, mint amilyen a 7. fejezet elején említett módszer, amely adott fokszámeloszlású hálózat létrehozására képes. Érdemes lenne ezt a lehetőséget is kipróbálni.

A CMS honlapja részletes, diákok számára is követhető leírást ad a CMS detektor felépítéséről [66]. A CERN-es mérés honlapja leginkább a méréshez szorosan kapcsolódó dolgokat részletezi [22]. Nagyon hasznos lenne a diákok előzetes felkészülése és utólagos tájékozódása szempontjából egy olyan honlap létrehozása, amely az előző két oldalon található ismeretek mellett elmagyarázza magyar nyelven a méréshez kapcsolódó részecskefizikai alapismereteket és a kísérleti részecskefizika általános elveit is. A CERN-sajátkezűleg honlap átdolgozásával viszonylag könnyen lehetne ilyen oldalt létrehozni, de van egy jobb lehetőség is. Az ATLAS kísérlet honlapja megoldotta a dolgot: a Z bozonhoz kapcsolódó méréshez kidolgozott leírások ezeket az alapismereteket tartalmazzák nem csak angol, hanem más nyelveken is [67]. A CMS-sel kapcsolatos diákműhely-kísérletek leírásához az ott található oldalak egy részét át lehetne emelni, és a CMS részletesebb leírásával kiegészíteni, majd ezeket a részeket is lefordítani más nyelvekre.

10. fejezet

Summary and outlook

In my dissertation, I have presented my results on introducing particle physics and complex networks into the teaching curriculum at advanced high school and BSc level.

Understanding the theories of physics requires a high level knowledge of mathematics. Teaching particle physics for high school students and university students with non-physics major, it is challenging to speak exactly and comprehensible. For this reason the most suitable way is to show and discuss experiments, because they use concepts the students are familiar with from their everyday life (position, momentum, changing speed in magnetic field, electric signal). The international particle physics masterclasses and the event viewers of their home pages make it possible for the students to get familiar with methods of scientific research, the statistical nature of the phenomena of particle physics, the need for large number and precise measurements to reduce the uncertainty of the results, the method of evaluation of the results, the role of discussion of observations, the importance of basic research in significant technological innovations and vice versa. Understanding the fundamental principles, high school students may take a fancy become absorbed in this field and they may learn how to appreciate the difference between real and pseudoscience.

To help the students in their work I created a Hungarian version of the Hands on CERN home page, drew a table with the Nobel laureates of the field of particle physics. On this home page I listed Hungarian articles and books suitable for high school students as well. The students used this home page to prepare to the particle physics masterclasses. With the event viewer included in the Hands on

CERN home page students can analyze events collected in the DELPHI detector at CERN, and they can identify the particles emerging from the electron and positron colliding in the middle of the detector. The events can be grouped into event types determining the branching ratios of event types. I worked out a measurement, to measure the number of particle generations based on the branching ratios.

The particle physics masterclass was organized first time in 2005, in the world year of physics, by the European (now International) Particle Physics Outreach Group. Hungary participated in every spring from the beginning with three sites (Debrecen University, Research Institute for Particle and Nuclear Physics and Óbuda University). On the Óbuda University site I organize the local events of the masterclass, introduce the students into the field of the theoretical and experimental particle physics, help them in carrying out the measurements. In 2013 it will be the ninth occasion of receiving 20 high school students in our university centre to investigate events collected by the detectors at CERN. In 2005 we asked the students to fill a questionnaire form. From the analysis of the answers from two Hungarian sites I concluded that they were interested in these events, they collected previously unknown information about particle physics and they became more interested in the physics than they had been before.

The event viewers used in the masterclasses may become useful tools for study circles at high schools. Thanks to the teacher programs at CERN since 1998 in English, and since 2006 in Hungarian as well, the number of the teachers knowing more about particle physics is steadily growing.

In every summer since 2006 forty Hungarian high school teachers visit CERN to learn about the research carried out there and its theoretical and information technological background.

The research of the complex networks is younger than that of the most basic building blocks of the matter and their interactions. There are lots of networks in the fields of the technology, in the life sciences and in the Internet. Investigating these networks can lead to developing useful technologies like forecasting the spreading of diseases, clearing up the roles of given proteins and gene sequences in the living organizations and the function of the brain, making the most of the telephone networks, working out more efficient Internet protocols, quantifying the quantity of the program codes and sending the advertisements to target audience.

I created the cxnet network-analyzer program package to make easier investigating networks. This package is able to create the whole software dependency network of the Ubuntu and Debian Linux distributions in tens of seconds, so we can analyze locally created fresh data. The cxnet is able to plot the degree dis-

tributions of the networks with several methods: using several binning methods or as cumulative distribution. For scale-free networks it can be used to determine the exponent of the power-law function.

I worked out a course for the students of computer engineering about the investigation of complex networks. During the course the students become familiar with the most important properties and models of complex networks. They learn to determine these properties for networks stored in files and write program codes to solve problems related to complex networks with the `cxnet` and `igraph` program packages.

I analyzed the properties of the software package dependency network of the Ubuntu Linux distribution. I pointed out that it is scale-free, hierarchical, it has large clustering coefficient, and its evolution is driven by the preferential attachment.

I used networks generated by the `cxnet` package and some others to prepare problems for the informatics competition organized by the Óbuda University in the last three years. These exercises are about using relational databases to obtain degree distributions. The students have to load the network into the Microsoft Access or OpenOffice.org Base from two files containing the names of the packages for every identifier number, and the edge list as the pairs of identifier numbers respectively. Only the most talented competitors could solve the complete problems.

For creating network with given property group I used a method based on optimization: the multifractal network generator (MFNG) method developed by Palla et al. In this case we optimize a probability measure—the so called generating measure—which is iterated during the process in order to get the link probability measure. From the link probability measure one can calculate the expected value of the properties of the network, such as the average degree, the degree distribution, the clustering coefficient, without the actual generation of the network. For the optimization we need to define an energy function as the function of the generating function in such a way, that the closer the expected properties of the network to the given target properties are, the smaller the energy. We need to minimize the energy in a simulated annealing process.

I was the first to implement the MFNG algorithm in Python programming language. I targeted first a given average degree, then a given degree distribution for the generated network. We have rewritten the part of the program that made the most of the running time into C++. This made the program run an order of magnitude faster. Combining the codes written in C++ and Python we have been

able to use the advantages of both programming language: the flexible usage of function parameters and the high level data structures in Python, and the speed of C++.

I improved the articles of the Hungarian Wikipedia in the fields of particle physics and networks significantly. The articles I created or improved were articles of particle physics, history of physics, biography of physicists, universities, research institutions and complex networks and the physics portal. These pages are visited regularly.

My work is not finished with this dissertation. I envisage several options to improve my results.

The cxnet program can run on a Linux operating system without any graphical user interface, which allowed me to create a snapshot of the software package dependency network of the 12.04 version of the Ubuntu Linux distribution every hour on a server machine, so I have more than 5000 snapshots (February 2013). This data set have the advantage that these snapshots used the same repositories to fetch the package information, so the snapshots are comparable. These snapshots make for me and my students the opportunity to determine the laws of the changes in this network: testing, whether the preferential attachment is true for this network and whether there are any other laws of the evolution of the network. I started to write a program to find the packages appeared, disappeared or changed in the network. As the number of the saved network is large, I planned to write a program that saves as many useful information about the changes as it can as the program process the networks. I do not want to go cycles through the networks too many times, because just to load all of the networks without any investigation needs more than 7 hours.

My MFNG Python package allows me to give two properties simultaneously so the optimization will be run with the sum of the energies calculated from the properties. So we get an opportunity to create networks with more given properties. In the Supporting Information of the article about the MFNG there is an equation to calculate the clustering coefficient from the iterated generating measure. So we can choose to create a network with scale-free property with a given exponent and with a given clustering coefficient simultaneously.

Köszönetnyilvánítás

Köszönöm Trócsányi Zoltánnak a javasolt feladatokat, a cikkeimhez adott tanácsokat, valamint a lehetőséget az EPPOG-ban való részvételre.

Köszönöm Nepusz Tamásnak, hogy az igrph-fal kapcsolatos kérdéseimre adott részletes és hasznos válaszait, és azt hogy kifejlesztették az igrph-ot.

Köszönöm Orlik Iván Péternek, hogy segítségemre volt az MFNG felgyorsításában és az értekezés hibáinak keresésében.

Köszönöm Sükösd Csabának és Jarosievits Beátának, amiért évente megszervezik a magyar tanárok részvételét az pár napos magyar nyelvű CERN-es fizikatanárprogramban, amelyek a keretében 2007-ben én is kijutottam a CERN-be.

Köszönöm Orosz Tamásnak és Lakner Józsefnek, hogy állandóan figyelemmel kísérte és sürgette az értekezés és a folyóiratcikkek megírását. Köszönöm az Óbudai Egyetemnek és azon belül az Alba Regia Egyetemi Központnak, hogy támogatta a tanulmányaimat, és lehetővé tette az Összetett hálózatok tárgy létrehozását és a diákműhelyek rendezését.

Köszönöm Erik Johanssonnak hogy a részecskefizikai diákműhelyeket útjára indította; a hallgatónak és a diákműhely résztvevőinek a kitartásukat és a kérdéseiket, amellyel engem is a témák alaposabb átgondolására ösztönöztek; a diákok tanárainak, hogy elküldték hozzám a tanítványaikat; Jancsó Gábornak, hogy összefogta a diákműhelyekről szóló híradásainkat.

Köszönöm Füzi Péternek, Babindai Zsoltnak és Balogh Zoltánnak, hogy a diákműhely videokonferenciáinak kivitelezésében segítettek.

Végül, de nem utolsó sorban, köszönöm a családomnak, hogy türelemmel viselték a doktori képzéssel járó elfoglaltságaimat.

Irodalomjegyzék

- [1] „Végveszélyben a természettudományos tanáképzés, mta-hírek, 2010.”
http://mta.hu/mta_hirei/vegveszelyben-a-termeszettudomanyos-tanarkepzes-91011.
- [2] „Mi lehet a megoldás a természettudományos tanárhiányra? napi gazdaság.” http://www.napi.hu/magyar_gazdasag/mi_lehet_a_megoldas_a_termeszettudomanyos_tanarhianyra.540038.html.
- [3] „A CERN honlapja.” <http://cern.ch>.
- [4] „Mindentudás Egyeteme.” <http://mindentudas.hu>.
- [5] „Az atomoktól a csillagokig.” <http://www.atomcsill.elte.hu>.
- [6] „Fizikai Szemle.” <http://www.kfki.hu/fszemle>.
- [7] „CERN blog.” <http://cernblog.wordpress.com>.
- [8] „Természet Világa.” <http://www.termeszetvilaga.hu>.
- [9] S. Weinberg, *A Model of Leptons*, *Phys.Rev.Lett.* **19** (1967) 1264–1266.
- [10] „ASACUSA honlap.”
<http://ad-startup.web.cern.ch/AD-Startup/Science/ASACUSA.html>.
- [11] G. Bencze, *A CERN részecskegyorsítói, Természet Világa, III. különszám* (2000).
- [12] D. Groom *et. al.*, *Particle Physics Booklet*. Springer, 2000.

- [13] S. Schael *et. al.*, [ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group], *Precision electroweak measurements on the Z resonance*, *Phys.Rept.* **427** (2006) 257–454, [[hep-ex/0509008](#)].
- [14] D. H. Perkins, *Introduction to High Energy Physics*. Cambridge University Press, Apr., 2000.
- [15] P. Aarnio *et. al.*, [DELPHI Collaboration], *The DELPHI detector at LEP, Nuclear Instruments and Methods in Physics Research Research: Accelerators, Spectrometers, Detectors and Associated Equipment* **303** (1991) 233–276. CERN-PPE 90174, LAL 90-82.
- [16] K. E. Johansson, C. Nilsson, J. Engstedt, and A. Sandqvist, *Astronomy and particle physics research classes for secondary school students*, *American Journal of Physics* **69** (May, 2001) 577.
- [17] „CERN-sajátkezüleg honlap.”
<http://www.erek.uni-obuda.hu/fizika/cern-sajatkezuveg>.
- [18] Á. Horváth, *Lássuk a részecskéket!*, *Fizikai Szemle* **LV** (2005), no. 8 261–264.
- [19] A. Pich, *The Standard Model of Electroweak Interactions*, . Comments: Based on lectures given at the 6th CERN-Fermilab Hadron Collider Physics Summer School (Geneva, 2011), 2010 European School of HEP (Raseborg, Finland), 2010 Int. School on Astroparticle Physics (Zaragoza, Spain), 2010 IDPASC School (Sesimbra, Portugal) and 2009 Int. Summer School and Conference on HEP (Mugla, Turkey, 2009). 52 pages, 38 figures.
- [20] Q. Ingram, *The Lead Tungstate Electromagnetic Calorimeter of CMS*, Tech. Rep. CMS-CR-2006-002, CERN, Geneva, Dec, 2005.
- [21] Z. Szillási, *Kísérleti eszközök fejlesztése a nagyenergiájú fizika számára*. PhD thesis, Debreceni Egyetem Természettudományi Kar, Debrecen, 2007.
- [22] „A CMS-sel kapcsolatos diákműhely-mérések magyar honlapja.”
<http://cms.physicsmasterclasses.org/cmshu.html>.
- [23] „A nemzetközi részecsefizikai ismeretterjesztő csoport (ippog) honlapja.”
<http://ippog.web.cern.ch>.

-
- [24] Á. Horváth, *Részecskefizikai diákműhely, Fizikai Szemle* **LV** (2005), no. 8 292–296.
- [25] „Wikipédiás közreműködésem.”
<http://hu.wikipedia.org/wiki/Szerkesztő:Harp/Közreműködésem>.
- [26] M. Chalmers, „Physicists find new particle, but is it the Higgs?.”
<http://www.nature.com/news/physicists-find-new-particle-but-is-it-the-higgs-1.10932>.
- [27] B. Bollobás, *Modern Graph Theory*. Springer-Verlag, New York, 1998.
- [28] J. Travers and S. Milgram, *An experimental study of the small world problem, Sociometry* **32** (1969) 425–443.
- [29] D. J. de S. Price, *Networks of scientific papers, Science* **149** (1965) 510–515.
- [30] D. J. de S. Price, *The general theory of bibliometric and other cumulative advantage processes, J. Amer. Soc. Inform. Sci.* **27** (1976) 292–306.
- [31] M. E. J. Newman, *The structure and function of complex networks, SIAM Review* **45** (JUN, 2003) 167–256, [cond-mat/0303516].
- [32] E. Ravasz and A. Barabasi, *Hierarchical organization in complex networks, Physical Review E* **67** (Feb, 2003) [cond-mat/0206130].
- [33] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabasi, *Controllability of complex networks, Nature* **473** (May, 2011) 167–173.
- [34] R. Albert and A.-L. Barabasi, *Statistical mechanics of complex networks, Reviews of Modern Physics* **74** (Jan, 2002) 47–97.
- [35] A.-L. Barabási, *Behálózva: A hálózatok új tudománya*. Helikon Kiadó, Budapest, 2008.
- [36] M. Summerfield, *Programming in Python 3: A Complete Introduction to the Python Language*. Addison-Wesley Professional, 2010.
- [37] „Python Scientific Lecture Notes.” <http://scipy-lectures.github.com/>.
- [38] A. Hagberg, D. Schult, and P. Swart, „NetworkX: High productivity software for complex networks.” <http://www.webcitation.org/5aMSYmJUi>, 2004–.

- [39] G. Csárdi and T. Nepusz, *The igraph software package for complex network research*, *InterJournal Complex Systems* (2006) 1695.
- [40] „Tárolóim a GitHub oldalon, közöttük a cxnet és az mfng.”
<https://github.com/horvatha>.
- [41] J. Loeliger, *Version Control with Git*. O'Reilly, 2009.
- [42] „Debian homepage.” <http://www.debian.org/>.
- [43] „Basics of the debian package management system.”
http://www.debian.org/doc/manuals/debian-faq/ch-pkg_basics.en.html.
- [44] „python-apt documentation.” <http://apt.alioth.debian.org/python-apt-doc/>.
- [45] N. LaBelle and E. Wallingford, *Inter-package dependency networks in open-source software*, 2004.
- [46] M. E. J. Newman, *Mixing patterns in networks*, *Phys. Rev. E* **67** (Feb, 2003) 026126.
- [47] T. Maillart, D. Sornette, S. Spaeth, and G. von Krogh, *Empirical tests of Zipf's law mechanism in open source Linux distribution*, *Phys. Rev. Lett.* **101** (2008) 218701.
- [48] O. F. de Sousa, M. A. de Menezes, and T. J. P. Penna, *Analysis of the package dependency on Debian GNU/Linux*, *Journal of Computational Interdisciplinary Sciences* **1** (Mar., 2009) 127–133.
- [49] T. Gradowski, M. Mrowinski, and R. Kosinski, *The dependency network in free operating system*, in *Acta Physica Polonica B, Proceedings Supplement*, vol. 5, 2012.
- [50] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, *The architecture of complex weighted networks*, *Proceedings of the National Academy of Sciences of the United States of America* **101** (Mar 16, 2004) 3747–3752, [cond-mat/0311416].
- [51] Á. Horváth and Z. Trócsányi, *Complex networks in the higher education*, in *Proceedings of the 6th International Symposium Intelligent Systems and Informatics* (A. Szakál, ed.), pp. 1–6, 2008.

-
- [52] „Ubuntu honlap.” <http://www.ubuntu.com/>.
- [53] Á. Horváth and Z. Trócsányi, *Complex networks in the curriculum of computer engineers*, in *IEEE Proceedings of the 8th International Symposium on Applied Machine Intelligence and Informatics*, (Herľany, Slovakia), 2010.
- [54] M. E. J. Newman, *Power laws, Pareto distributions and Zipf's law*, *Contemporary Physics* **46** (2005) 323.
- [55] A. Clauset, C. R. Shalizi, and M. E. J. Newman, *Power-law distributions in empirical data*, 2007.
- [56] Á. Horváth, *The software package dependency networks of some linux distributions*, in *4th IEEE International Conference on Nonlinear Science and Complexity: NSC 2012*. (I. J. Rudas, ed.), (Budapest, Hungary), 2012.
- [57] G. Palla, L. Lovász, and T. Vicsek, *Multifractal network generator*, *Proceedings of the National Academy of Sciences* **107** (2010), no. 17.
- [58] Á. Horváth and Z. Trócsányi, *Multifractal network generator with igrph*, *AIS 2010, Symposium on Applied Informatics and Related Areas* (2010).
- [59] Á. Horváth, *Creating networks with given degree distribution with the multifractal network generator*, in *AIS 2010, Symposium on Applied Informatics and Related Areas* (G. Györök, ed.), (Székesfehérvár, Hungary), 2010.
- [60] P. Erdős and A. Rényi, *On the evolution of random graphs*, *Publication of the Mathematical Institute of the Hungarian Academy of Sciences* **5** (1960) 17–61.
- [61] R. Albert, H. Jeong, and A.-L. Barabasi, *Error and attack tolerance of complex networks*, *Nature* **406** (2000).
- [62] „Moodle, Quiz module.” http://docs.moodle.org/24/en/Quiz_module.
- [63] „Összetett hálózatok oldal @ AREK.” <http://django.arek.uni-obuda.hu/cxnet>.
- [64] „Az Alba Regia Informatikaverseny honlapja.” <http://infoverseny.arek.uni-obuda.hu>.

- [65] H. Jeong, Z. Nédá, and A. L. Barabási, *Measuring preferential attachment in evolving networks*, *EPL (Europhysics Letters)* (Feb., 2003) 567+.
- [66] „CMS detector design.” <http://cms.web.cern.ch/news/cms-detector-design>.
- [67] „Az ATLAS Z bozonos diákműhely-mérés honlapja.”
<http://atlas.physicsmasterclasses.org/en/zpath.htm>.