

Opponensi vélemény

Jelasy Márk: *Gossip-based Protocols for
Large-scale Distributed Systems*

c. MTA doktori értekezéséről

Az értekezés nagyméretű elosztott számítógépes hálózatok "pletyka" alapú protokolljárásaival foglalkozik. Terjedelme 168 oldal, Bevezetést, 8 fejezetet és 145 tételből álló irodalomjegyzéket tartalmaz.

Az értekezés témája rendkívül aktuális az elosztott számítógéprendszerek nagy száma, méretei és eltérő szoftveres és fizikai megvalósításai miatt. Az értekezés központi irányítás nélküli egyenrangú résztvevőkből álló peer-to-peer (P2P) hálózatokat vizsgál.

Az értekezésben a szerző pletyka típusú algoritmusokat ad meg a hagyományos adatszórásán túlmutató problémák (globális számítási és mintavételezési feladatok, átfedőhálózatok építése, stb.) megoldására. A vizsgált feladatok közös jellemzője, hogy a résztvevők periodikus jelleggel információt cserélnek más résztvevőkkel, és ennek eredményeképpen módosítják az állapotukat. A különböző feladatokat ellátó pletyka algoritmusok modulként használhatók és akár komplett, több rétegből álló teljes rendszereket is megvalósíthatnak.

A szerző elméleti vizsgálatainak célja folyamatok tulajdonságait és dinamikáját szemléltető közelítő modellek alkotása és a javasolt algoritmusok viselkedésének a pontos leírása. A modellek pontosságát kísérleti módszerekkel, szimulációval ellenőrizte.

Az értekezésben javasolt eljárások tesztelésének legfontosabb eszköze a *Peersim* rendszer, amely egy nyilvános, a

<http://peersim.sourceforge.net/>

címről letölthető P2P szimulációs program (1.0.5 változat). A webhely jelentős mennyiségű kapcsolódó anyagot és információt tartalmaz a Peersim használatáról, amely a programrendszer jelentőségét és nemzetközi hatását jól mutatja (pl. 45 letöltés 2014-03-10 és 2014-04-16 között (6 nap), a kapcsolódó cikkek száma, stb.). A Peersim program fő fejlesztői *Jelasy Márk*, *Alberto Montresor*, *Gian Paolo Jesi* és *Spiros Voulgaris* voltak. Jelasy Márk több esetben olyan implementációkat is készített, amelyeket különböző valós teszhálózatokon vizsgált (pl. PlanetLab hálózat, DAS klaszter).

Az értekezés 1. fejezetében a szerző összefoglalja a legalapvetőbb fogalmakat, vizsgálati módszereket a vizsgálat módszereit és fontosabb eredményeit.

Az értekezés 2. fejezete a társ mintavételezéssel foglalkozik. A szerző azonosítja és motiválja a társ mintavételezést, mint önálló középhéteg szolgáltatást és javasol egy pletyka alapú megvalósítást, amelynek a lényege, hogy minden csúcs tárol egy kisszámú véletlen mintát a hálózat csúcsaiból. Ezek a minták egy véletlen fedőhálózatot definiálnak. A csúcsok úgy jutnak új mintákhoz, hogy az aktuálisan ismert szomszédokkal keverési lépéseket hajtanak végre folyamatosan, amelynek során egymás szomszédainak a segítségével frissítik a saját szomszédlistájukat. Az algoritmus paraméterezhető. A paraméterekkel folytonosan lehet állítani, hogy egyes tulajdonságok mekkora hangsúlyt kapjanak: a csúcsokon gyűjtött minták korrelálatlansága más csúcsok mintáival vagy a hibatűrő és önjavító képesség domináljon, vagy a kettő kombinációja jelenjen meg. Az algoritmust igen kimerítő jellegű (alapos) és méretű tesztelésnek vetette alá a szerző. Kísérletileg igazolta, hogy a keveredő véletlen átfedőhálózat rendkívül adaptív és robusztus, akár a korrelálatlan véletlen gráfok. A globális véletlenség vonatkozásában a szerző igazolta, hogy a rendszer drasztikusan különböző kezdőállapotokból ugyanabba a stabil konfigurációba konvergál, amelyben a paraméterek függvényében a vizsgált mutatók stabil értékeket vesznek fel.

Az értekezés 3. fejezete az "átlagszámítás" feladatával foglalkozik. Itt az alapfeladat egy elosztott, dinamikus rendszermodellben olyan számítások elvégzése az elosztott csúcsok felett, amelyek aggregálják a helyi adatokat. A szerző egy teljesen elosztott algoritmust javasol adat aggregációra, amely az átlag, egyéb középértékek, minimum, maximum, magasabb momentumok kiszámítására (és a hálózat méretének meghatározására is) képes. Az algoritmus alapötlete (Algorithm 10) igen egyszerű és szellemes: csúcspárokhoz tartozó értékek átlagolása és az eredeti értékek felülírása az átlaggal, iteratívan. A szerző igazolja, hogy a közelítések varianciája a hálózatban exponenciálisan csökken, és egyúttal megadja a pontos konvergencia sebességét is. Az algoritmus viselkedését jellemzi abban az esetben is, ha üzenetvesztés, vagy csúcsok kiesése is lehetséges. Az algoritmus implementált változatát kiegészíti egy újraindító mechanizmus, amely adott számú iteráció után egy újabb „korszakot” indít. Ez az aszinkron megoldás biztosítja a rendszer robusztusságát változó csúcshalmaz és más hibák esetében is. Az algoritmust eredményesen tesztelte a PlanetLab teszhálózatban is (<http://www.planet-lab.org/>).

Az értekezés 4. fejezete az elosztott hatványiterációt vizsgálja és egy

aszinkron algoritmust javasol nemnegatív élsúlyú, erősen összefüggő átfedő-hálózatok domináns sajátvektorának a meghatározására tetszőleges pozitív domináns sajátérték esetére. Az algoritmus Lubachevskynek és Mitra - a $\pi = \pi P$ ($\pi > 0$, $P \geq 0$ irreducibilis mátrix) sajátérték feladat megoldására szolgáló - aszinkron iterációs algoritmusának a módosítása pletyka alapú pletyka alapú elosztott normalizálással (aggregációval) abból a célból, bármilyen pozitív domináns sajátérték esetén használható legyen. Itt az aggregációnak két lehetséges célja van: a konvergencia biztosítása, ill. a vektor hosszának a normalizálása. Az algoritmus a PageRank értékek számítására is használható. Az algoritmust szimulált és valós hálózatokon is kiértékelte. A bemutatott teszteredmények szerint bizonyos esetektől eltekintve az algoritmus robusztus és konvergál.

Az 5. Átfedőhálózatok szeletelése című fejezet erőforrások allokációját vizsgálja. A konkrétan vizsgált feladat a hálózat szeletelése, ami a hálózat csúcsainak osztályozása adott képességek szerint. A feladat nehézségét az jelenti, hogy a rendelkezésre álló erőforrások eloszlása nem ismert, így lokálisan nem dönthető el, hogy egy adott csúcs melyik osztályba tartozik. A probléma megoldására javasolt eljárás a következő: Tegyük fel hogy az i csúcs x_i erőforrással rendelkezik. Minden csúcs egyenletes eloszlásból vesz egy r_i mintát. Az algoritmus ezeket a mintákat rendezi az x_i értékek mentén. A rendezés úgy zajlik, hogy minden csúcs cserепartnereket keres, akikkel az r értékeket kicserélve a rendezettséget növelni tudja. A rendezés után - az r_i értékeket ismerve - lehet lokális döntéseket hozni az osztályba sorolásról. A szerző megmutatja a feladat kapcsolát az átlagolással. Az 5.4.1. állításban igazolja, hogy a rendezés során a helyes indextől vett abszolút távolság várható értékben átlagolódik egy sikeres cserét követően. Tehát a sikeres cserék sorozatát tekintve az algoritmus erre a rendezetlenségi mértékre nézve átlagolásként viselkedik. A kapcsolódó szimulációs kísérletek, amelyekben a hibátűrést, a csúcsszám változás hatását is vizsgálta, támogatják az elméleti eredményeket.

A 6. fejezetben általános célú átfedőhálózatot létrehozó algoritmust (T-MAN) javasol, amely csak társ mintavételezést használ. A javasolt megoldás lényege, hogy a társ mintavételezésnél látott módszerhez hasonlóan a csúcsok rendszeresen kicserélik egymással a szomszédlistájukat, így gyűjtve össze azokat a szomszédokat, amelyeket az adott topológia megkövetel. A topológia egy rangsoroló függvényvel van meghatározva, amely minden i csúcs szempontjából bármely csúcshalmazt rangsorol abból a szempontból, hogy az adott halmazon belül i számára mennyire kívánatos mint szomszéd. A

rendszer inicializálása véletlen szomszédokkal történik, amihez kell a társ mintavételező szolgáltatás is. A T-MAN protokoll 5 paraméterrel rendelkezik, amely lehetővé teszi egy sor feltevés empirikus vizsgálatát, ill. az eljárás megfelelő beállítását. Az algoritmus elemzése azt mutatja, hogy a legtanácsosabb a legközelebbi ismert szomszédot választani a csere céljára, kombinálva egy tabulistával, amelyen az elmúlt néhány iteráció szomszédai szerepelnek. A helyben tárolt, összegyűjtött szomszédok számát nem kell korlátozni, mert a hálózat méretében logaritmikus tárhelyigény lép csak fel. A tesztelések azt támasztják alá, hogy a kívánt topológiát a hálózat méretének logaritmusával arányos időn belül előállítja az algoritmus. Ezen felül az algoritmus hibatűrése is kiváló. A T-MAN algoritmus robusztus az üzenetvesztésre és késleltetésre, valamint a csúcsok távozására.

Az értekezés 7. fejezete a T-MAN algoritmus alkalmazása a 2001-ben bevezetett, hash-tábla bázisú CHORD átfedőhálózat gyors hidegindítására véletlen hálózattól kiindulva. A T-MAN algoritmussal rendezett gyűrűt hoz létre és közben a húrokat is létrehozza, felhasználva, hogy a meglátogatott szomszédok éppen olyan eloszlással rendelkeznek a csúcstól való távolságot tekintve, mint a keresett húrok és így nagy valószínűséggel további költségek nélkül majdnem minden húr-hely betölthető. Az algoritmust szimulációval értékelte, amely azt mutatja, hogy a létrehozott elosztott hash tábla minősége megfelel a követelményeknek, és a rendezett gyűrű létrehozásával megegyező költséggel létrehozható.

Az értekezés utolsó "*Towards a generic bootstrapping service*" című 8. fejezetében általános átfedőhálózatok hidegindításával foglalkozik, abból kiindulva, hogy a korábbi alkalmazások felfogható egymásra épülő középréteg szolgáltatásnak és a fedőhálózat konstruálás ideillesztése újabb funkciókat és alkalmazásokat tenne lehetővé. Itt javasol egy P2P architektúrát két fő komponenssel: társ mintavételezés, hidegindítás. Végül adaptálja a T-MAN algoritmust bármely prefix alapú elosztott hash tábla gyors hidegindítására, amelyet tesztel is a PeerSim szimulációs rendszer segítségével.

Az értekezés gondosan megírt és jól olvasható. Kivételt talán csak az alkalmazott algoritmus leírások jelentenek összehasonlítva az eredeti cikkekben szereplő algoritmus leírásokkal, amelyek könnyebben érthetőek. Kifogásként még megjegyzem, hogy a (4.4) képlet korrekt, de nem elegáns. A képlet számlálójában szereplő mennyiség skalár, tehát ott nem kell norma jelölést alkalmazni.

Összefoglalás

Jellasi Márk jelentős alkotó és fejlesztő munkát végzett P2P hálózatok különféle - elméleti és gyakorlati szempontból is - fontos problémáinak a megoldása körében. Az általa javasolt pletyka alapú protokoll eljárások sok esetben igen szellemesek és a problémakör mély ismeretéről tanúskodnak. A javasolt algoritmusokat elméleti és gyakorlati szempontból egyaránt vizsgálta. A javasolt algoritmusok jóságát, hatékonyságát és hasznosságát jelentős mennyiségű, részletes és alapos, mondhatni kimerítő alapossgal elvégzett számítógépes szimulációval is alátámasztotta. Az értekezésben foglalt eredmények nagyon jelentős nemzetközi hatással bírnak. A szerző a témakör egyik nemzetközileg legismertebb vezető kutatója. Az itt bemutatott eredmények egy része máris megtalálható külföldi egyetemi kurzusok anyagában is. A disszertációban foglalt eredményeket (téziseket) elfogadom, a nyilvános vita kitűzését és a mű elfogadását (az MTA doktora cím odaítélését) javaslom.

Budapest, 2014. március 17.

Dr. Galántai Aurél
az MTA doktora