

Pletykaalapú tanulás együttműködésen alapuló mobiltelefonos hálózatban

Berta Árpád

Témavezető

Dr. Jelasity Márk

Számítógépes Algoritmusok és Mesterséges Intelligencia Tanszék

Informatika Doktori Iskola
Szegedi Tudományegyetem



Doktori értekezés összefoglaló

Szeged

2020

Bevezetés

Az elmúlt évtizedekben robbanásszerűen megnőtt a különböző mobil és okos eszközöknek a száma. Az élet legkülönbözőbb területein vált megkerülhetetlenné a jelenlétük. Használatuk során tömérdek adat keletkezik. Ezek kiemelkedően fontos nyersanyagai a legnépszerűbb szolgáltatásoknak, melyek mögött általában adatbányászati módszerek húzódnak. Ezekhez a legtöbbször központi helyre történő adatgyűjtés kapcsolódik. Ez a típusú adatkezelés viszont egyre gyakrabban személyiségi jogi, adatvédelmi aggályt vet fel és egyre nagyobb ellenállásba ütközik. Ennek köszönhetően egyre inkább előtérbe kerülnek az olyan megoldások, ahol a személyes adatunk a saját eszközünkön marad, de az adatfeldolgozás továbbra is elérhető elosztott algoritmusok segítségével.

Egy ilyen javasolt módszer a pletykaalapú tanulás [8]. A tézis során az ezzel kapcsolatban ismert nyitott kérdések megoldását jelöltük ki fő kutatási célként. Azért döntöttünk emellett, mivel ez ténylegesen egy teljesen elosztott megközelítés, nincs szükség semmilyen központi szerverre. A résztvevők a tanuló modelleket egymás között küldik és helyben végeznek el minden szükséges számítást. Mindez a skálázhatóságot jelentősen olcsóbbá teszi. Így megnyitva a lehetőséget a startupok, az NGO-k, a szövetkezetek és olyan helyi közösségek előtt, amelyek alacsony induló költségvetéssel rendelkeznek, hogy egy folyamatosan fejlődő, kollaboratív környezetben nyújtsanak intelligens szolgáltatásokat. Egy ilyen megfelelően kifejlesztett elosztott rendszer közcélokat is szolgálhat (pl. a közegészségügyben vagy a közoktatásban). Bár a mi figyelmünket elsősorban okostelefonos platformra fordítottuk, a pletykaalapú tanulás alkalmazható az okos mérőórák vagy a dolgok internetje (IoT) esetében is.

Egy ilyen megközelítés önmagában még nem garantálja az adatbiztonságot (privacy), de sokkal könnyebben elérhetővé teheti. Továbbá megadja az elméleti lehetőséget a differential privacy feltételeinek a teljesítésére. Ennek elérése érdekében a tézis során bemutatunk egy olyan véletlen séta menedzselő eljárást, amely képes alacsonyan tartani a privacy büdzsét, még hozzá oly módon, hogy egyetlen véletlen sétát tart fenn. Hosszú távú célunk egy olyan teljesen nyitott kollaboratív környezet létrehozása, ahol aki adatot biztosít a közösség számára, annak cserébe lehetősége van saját célra bányászni a közösség kollektív adatából. Egy ilyen környezet kialakításához elsődleges fontosságú egy olyan véletlen séta menedzselő eljárás, amely különböző független elosztott feladatokhoz kapcsolódó véletlen sétákat menedzsel. Felfoghatjuk ezt a feladatot úgy is, hogy a

1. táblázat. Kapcsolat az egyes tézispontok és a hozzájuk tartozó publikációk között (ahol a • jelöli azokat a publikációkat, amelyek fő eredményeit közöljük, míg a ◦ kapcsolódó publikációkat jelöl).

	Tézis 1	Tézis 2	Tézis 3
P2P 2014 [3]	•	◦	◦
TIST 2016 [7]	•	◦	◦
ICCGI 2017 [9]	•		
SCN 2018 [5]	•		
DAIS 2019 [11]	•		
PDP 2016 [1]	◦	•	
JOWUA 2016 [6]	◦		•
PDP 2017 [2]	◦		•
ESANN 2014 [4]		◦	
IJASO 2018 [10]	◦		

hálózat minden résztvevőjéhez kapcsolódik egy véletlen séta, így összességében a hálózat mérete ($O(n)$) nagyságrendű sétaszám lehetséges. Erre a problémára mutatunk be egy olyan véletlen séta menedzselő eljárást, amely a lehető leggyorsabb továbbhaladást biztosítva a lehető leghatékonyabb módon, a lehető legtávolabbi életben tudja tartani a sétákat. Ahhoz, hogy még pontosabb képet kapjunk, hogyan is működne egy pletykaalapú protokoll a valóságban, még realizisztikusabbá kell tennünk a szimulációkon alapuló kiértékeléseket. A tézisben bemutatunk egy okostelefonos trace-t, amely általunk összegyűjtött adatokon alapul. Ennek során felhasználói viselkedéseket figyeltünk meg, amelyek hálózati képességekkel vannak kiegészítve. A tézisben bemutatott összes eredményünket az okostelefonos trace-en alapuló szimulációkkal értékeltük ki. Munkánk során kiemelt figyelmet fordítottunk a pletykaalapú tanulás hatékonyabbá tételére. A legnagyobb korlátot a kommunikáció magas költsége okozza, ezért a lehető leghatékonyabb kommunikációra kell törekedni, a küldendő csomagok méretét csökkenteni. Az elosztott tanulás esetében ezek a csomagok legfőképpen a tanuló modellt jelentik, amelyek dimenziója jellemzően magas. Erre a problémára több különböző elosztott dimenziócsökkentő eljárást mutattunk be a tézis során.

A tézis eredményeinek összefoglalása

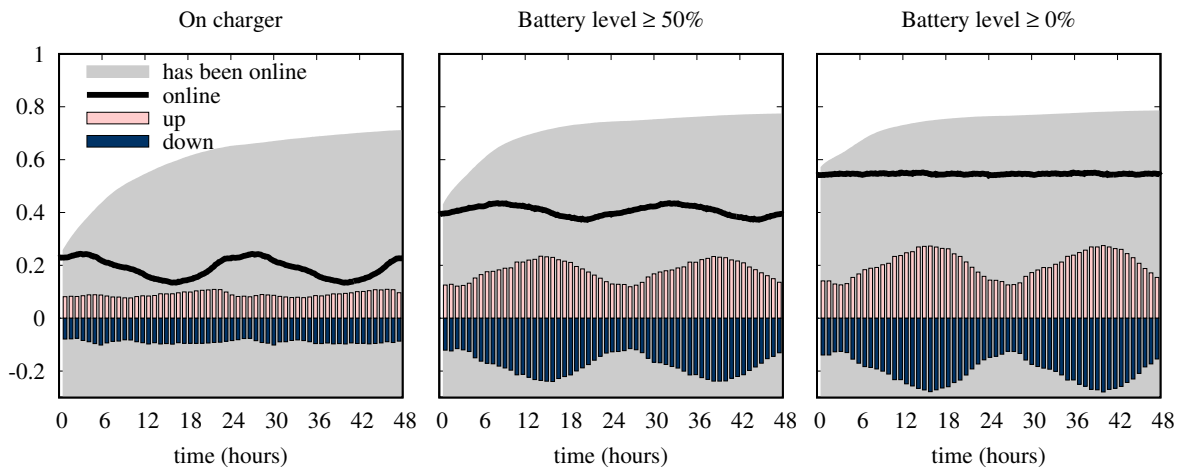
A tézis fő célkitűzéseként számos, változatos okostelefonos pletyka alapú tanulási problémára mutattunk be robusztus módszereket annak érdekében, hogy a pletykaalapú tanulás még kedve-

zőbb alternatívája legyen a jelenleg alkalmazott, központi szerveren történő adatbányászatnak. Az alábbiakban röviden összefoglaljuk az eredményeket tézispontokra bontva. Ennek során azokat az eredményeket emeltük ki, amelyek elérésében a tézis szerzőjének a szerepe meghatározó fontosságú volt. Az 1. táblázat részletesen bemutatja, hogy mely tézis ponthoz kapcsolódóan mely publikációk tartalmazzák az alapvető, új eredményeket (●-tal jelölve). A táblázat továbbá tartalmazza azt is, hogy az adott tézis ponthoz mely publikáció kapcsolódik még mint felhasznált korábbi kapcsolódó eredmény (○-rel jelölve). Az 1. tézisben bemutatjuk az okostelefonos traceszel kapcsolatos eredményeket, amelyek a disszertáció 3. és 6. fejezetében is megtalálhatóak. A 2. tézisben található dimenziócsökkentő eljárásokkal kapcsolatos eredmények a disszertáció 4. fejezetében vannak részletesebben kifejtve; a 3. tézisponthoz pedig a disszertáció 5. fejezete kapcsolódik.

1. tézis: Okostelefonos trace

Valós felhasználói viselkedéseken alapuló trace-t javasoltunk, amelynek segítségével lehetőségünk van az eszközök hálózati jelenlétében történő változást szimulálni. A trace-hez az általunk fejlesztett Android alkalmazás által gyűjtött adatok révén jutottunk. Az alkalmazás alapvető hálózati információkat szolgáltat a felhasználóknak, ők pedig cserébe elérhetővé teszik az adataikat a kutatási céljainkra. A trace több, mint ezer különböző felhasználótól lett összegyűjtve. Ezeket egynapos (illetve egynapos átfedéssel kétnapos) időszakokra bontottuk, ami összesen körülbelül 40 000 időszakot jelent. Minden idő szeletet egy felhasználónak feleltethetünk meg, így egy virtuális nap (vagy két nap) hosszú szimulációt tudunk végrehajtani. Amikor több felhasználó van, mint időszak, akkor visszatevés nélkül újra mintavételezhetjük az eredeti időszakokat. A bemutatott trace-nek köszönhetően sokkal realiztikusabb tudjuk kiértékelni az elosztott hálózati protokollokat, ahogy azt a tézis további részein majd láthatjuk.

A szimulációnak szabad paramétere lehet a töltési állapot, tehát az, hogy megköveteljük-e, hogy az eszköz töltőn legyen, vagy elvárunk-e egy minimális töltöttségi szintet. Az ezzel kapcsolatos két nap hosszú időszakokból álló trace-en történt megfigyeléseket az 1. ábrán láthatjuk. Egy hálózaton belül különböző szimulációs forgatókönyveket hajthatunk végre az alapján, hogy milyen követelményeket szabunk meg az elérhetőséggel kapcsolatban. Amennyiben telefon- és



1. ábra. Az idő függvényében a jelenlegi és valaha volt online felhasználók hányada. Az oszlopok az adott időszakban fel- vagy lecsatlakozó felhasználók hányadát jelzik (a lecsatlakozásokat negatívként).

felhasználóbarát szolgáltatást szeretnénk nyújtani, akkor az lesz a feltétel, hogy az eszköz a hálózati elérhetősége mellett legyen töltőn is. Ehhez képest, ha a kiértékelés szempontjából kritikus a fel- és lekapcsolódások száma, akkor semmilyen töltési állapottal kapcsolatos feltételt nem szabunk meg, ami sokkal dinamikusabb szimulációt eredményez. Így ebből a szempontból egy legrosszabb eset forgatókönyvet hajthatunk végre.

Az elmúlt években a legtöbb bemutatott módszerünket ezzel az okostelefonos trace-szel értékeltük ki. Így jártunk el a dimenziócsökkentő protokollok és a véletlen sétákat menedzselő protokollok esetében is. Ezekon kívül bemutattunk egy teljesen elosztott mini-batch gradiens alapú módszert, amely ideiglenesen önszerveződő csoportokat hoz létre a hálózatban. A csoportok kollaboratívan számolják ki a gradienst. Ezek a csoportok úgy alakulnak ki, hogy egy gyökérből induló k hosszú törzsszel rendelkező binomiális fa épül az overlay hálózaton. Ebben a munkában kiemelkedően fontos volt bemutatni, hogy praktikusán megvalósítható-e egy ilyen overlay hálózat feletti mini-batch gradiens lépés. Az okostelefonos trace-en végrehajtott empirikus vizsgálataink alapján teljesen világossá vált, hogy csak nagyon kis részben volt sikertelen az ilyen faépítési próbálkozás.

Ezen felül bemutattunk egy időben inhomogén Markov-modellt, amely egy alternatívája lehet a fel- és lecsatlakozások szimulációjának, és amely szintén az okostelefonos alkalmazás által gyűjtött adatokon alapul. Ennek elérése érdekében megvizsgáltuk az elérhetőségi állapotok hosszának a feltételes valószínűségi eloszlásait az előző állapot hossza illetve az aktuális napszak függvényé-

ben. A megfigyelt eloszlás újramintavételezésével szintetikus trace-t hoztunk létre a felhasználók viselkedéséről. A modellt többféleképpen validáltuk. Többek között összevetettük a valós trace-szel és azzal megegyező hálózati terjedési mintákat figyelhetünk meg.

Az adatgyűjtő alkalmazás megjelenése óta több millió adatrekord gyűlt össze. Ilyen mértékű adatmennyiség esetén természetes a zaj megjelenése az adatban, továbbá az is, hogy inkonzisztens bejegyzések érkeznek. Ezért fontos volt, hogy figyelmet fordítsunk a megfelelő adattisztításra. Javasoltunk egy módszert, amely javítani tudja a hibás NAT méréseket. Bemutattuk, hogy mik azok az ajánlások, amelyek az évek alatt az adatgyűjtéssel kapcsolatban tapasztalt tanulságokból levonhatóak. Továbbá kiegészítettük az adatgyűjtést tényleges peer-to-peer (P2P) kapcsolat-kiépítés próbálkozás mérésével. Mára a trace tartalmaz minden lokálisan, az eszközön megfigyelhető lényeges információt, mint a hálózati elérhetőség vagy a töltöttségi állapot, ezen felül pedig a STUN mérés mellett a közvetlen P2P kapcsolat mérési eredménye is elérhető. Ezek együttes figyelembe vétele mellett megbecsülhető a kapcsolódás sikeressége. Ezekről jól érthető, tömör leírást és alapvető statisztikákat is közöltünk.

Főbb eredmények:

- Valós okostelefon-felhasználói szokásokon alapuló trace a le- és felkapcsolódások szimulációjára.
- Az okostelefonos trace jellemzése és alapvető statisztikák bemutatása.
- Időben inhomogén Markov-modell a le- és felkapcsolódások szimulációjára.
- Egy k hosszú törzsszel rendelkező binomiális fa építésén, és egy biztonságos összegző protokollon alapuló decentralizált mini-batch módszer empirikus kiértékelése okostelefonos trace-en alapuló overlay hálózaton.
- Adattisztító eljárás hibás NAT mérések javítására.
- Kapcsolódó publikációk: [3, 5, 7, 9, 11]

Árpád Berta, Vilmos Bilicki, and Márk Jelasity. Defining and understanding smartphone churn over the internet: a measurement study. In *14-th IEEE International Conference on Peer-to-Peer Computing (P2P)*, pages 1–5, Sept 2014

István Hegedűs, Árpád Berta, Levente Kocsis, András A. Benczúr, and Márk Jelasity. Robust decentralized low-rank matrix decomposition. *ACM Transactions on Intelligent Systems and Technology*, 7(4):62:1–62:24, May 2016

Zoltán Szabó, Vilmos Bilicki, Árpád Berta, and Zoltán Richárd Jánki. Smartphone-based data collection with Stunner using crowdsourcing: Lessons learnt while cleaning the data. In *The Twelfth International Multi-Conference on Computing in the Global Information Technology (ICCGI 2017)*, pages 28–35, Jul 2017

Gábor Danner, Árpád Berta, István Hegedűs, and Márk Jelasity. Robust fully distributed mini-batch gradient descent with privacy preservation. *Security and Communication Networks*, 2018:15, 2018

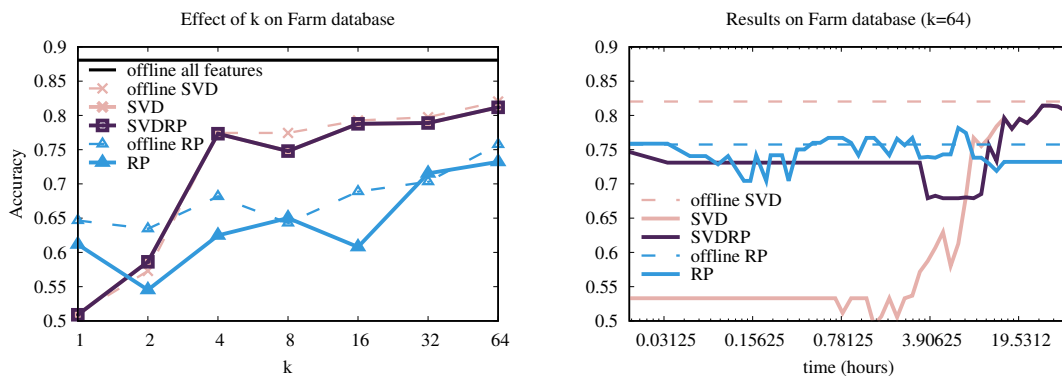
Zoltán Szabó, Krisztián Téglás, Árpád Berta, Márk Jelasity, and Vilmos Bilicki. Stunner: A smart phone trace for developing decentralized edge systems. In José Pereira and Laura Ricci, editors, *Proceedings of the 19th International Conference on Distributed Applications and Interoperable Systems (DAIS 2019)*, pages 108–115, Cham, 2019. Springer International Publishing

2. tézis: Dimenziócsökkentő módszerek

Olyan robusztus és hatékony elosztott dimenziócsökkentő eljárásokat mutattunk be, amelyek olyan rendszermodellben képesek elfogadható teljesítményt nyújtani, ahol minden résztvevőnél a hálózatban mindössze egyetlen adatbejegyzés található. Az elosztott tanuló algoritmusok esetében kritikus a nyers adat mérete, ugyanis ez, ha túl nagy, akkor hatalmas kommunikációs költséget von maga után, ami pedig a legrosszabb esetben akár teljesen ellehetetlenítheti az elosztott tanulást. Éppen ezért ebben a környezetben különösen fontos a dimenziócsökkentő módszerek alkalmazása. Jelöljük a továbbiakban d -vel az eredeti dimenziószámot és k -val az elérni kívánt csökkentett dimenziószámot, ami szabad paraméter, általában egy kis szám.

Az első javasolt módszerünk megkeresi a lehető legjobb véletlen projekciós mátrixot. Ez azon az ötleten alapul, hogy véletlen projekciós mátrixokat nagyon olcsó generálni. Emellett az elosztott kiértékelése is nagyon könnyen megvalósítható. Kommunikációs szempontból elegendő mindössze a generáláshoz szükséges véletlen seed és a kiválasztott, rögzített tanuló modell továbbküldése. Ez $O(k)$, lényegében elhanyagolható mértékű csomagméretet eredményez. Ennek köszönhetően rengeteg különböző véletlen mátrix dimenziócsökkentő képessége értékelhető ki. Ezzel szemben az elosztott szinguláris érték felbontás (SVD – Singular Value Decomposition, visszonyítási módszer) esetében a teljes projekciós mátrix utazik a hálózaton. Így ebben az esetben $O(k \cdot d)$ csomagmérettel számolhatunk (a d akár milliós nagyságrendű is lehet).

A bemutatott algoritmusokat részletesen kiértékeljük és az elért eredményeiket összevetettük



2. ábra. Dimenziócsökkentő módszerek összehasonlítása accuracy mérték alapján a második szimulált nap végén a k függvényében (bal oldalt) és accuracy-változása az idő függvényében, rögzített $k = 64$ mellett (jobb oldalt).

az SVD eredményeivel. A vizsgálatainkat a fentebb részletezett valós okostelefonos trace segítségével valósítottuk meg. Az eredmények alapján arra jutottunk, hogy a javasolt véletlen projekciós mátrixok között szelektáló módszerünk előnyösebb, amennyiben nagyon rövid időn belül szükségünk van egy közel elfogadható minőségű dimenziócsökkentésre. Ezzel szemben az SVD módszer esetében a projekciós mátrix sokkal lassabban konvergál ahhoz az állapothoz, amely az SVD-vel elérhető legjobb accuracyt eredményezi a vizsgált feladaton. Ugyanakkor, ha ez idővel megtörténik, akkor valamivel jobb teljesítményre képes, mint ami véletlen projekció segítségével elérhető.

Ezeket felül bemutattunk egy további hibrid módszert, amely egyesíti az előnyeiket a véletlen projekciós módszernek és az SVD-nek. Azt javasoltuk, hogy párhuzamosan fusson a két módszer. Mivel az SVD esetében egy nagyságrenddel nagyobb a csomagméret, ezért a kihasználható sáv-szélesség 99%-ában ezt forgalmazzuk és 1%-nyi sáv-szélességet rendelünk a javasolt véletlen projekciós módszerhez. Dimenziócsökkentéshez kezdetben a véletlen projekciós mátrixot használjuk. Az SVD projekciós mátrix sorai szekvenciális sorrendben konvergálnak, és ekkor páronként ortogonálisok válnak. A páronkénti ortogonalitás mérhető, így amikor már vannak ilyen sorok, akkor létrehozható egy olyan mátrix, amely az SVD konvergált sorait használja, a maradék részét pedig feltöltjük az aktuálisan ismert legjobb véletlen mátrix soraival. Attól kezdve, hogy az SVD minden sora bekonvergált, már kizárólag az SVD projekciós mátrixát használjuk a dimenziócsökkentésre. Arra is rávilágítottuk a figyelmet, hogy ez hibrid módszer minden vizsgált időpontban kiemelkedő teljesítményre képes.

Főbb eredmények:

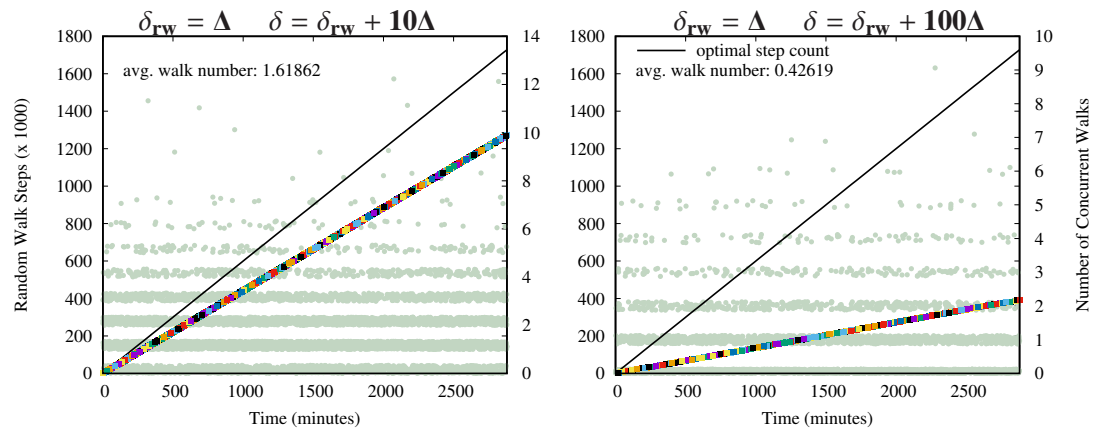
- Egy olyan elosztott algoritmus, amely a legjobb véletlen projekciós mátrixot keresi.
- Egy hibrid algoritmus, amely egyesíti az előnyeit a véletlen projekciós módszernek és az SVD-nek.
- Az elosztott SVD algoritmus dimenziócsökkentéssel kapcsolatos kiértékeléseknek a bemutatott algoritmusokkal való összevetése.
- Kapcsolódó publikáció: [1]

Árpád Berta, István Hegedűs, and Márk Jelasity. Dimension reduction methods for collaborative mobile gossip learning. In *2016 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP)*, pages 393–397, Feb 2016

3. tézis: Véletlen séták menedzselése

Ebben a tézispontban két véletlenséta-menedzselő hálózati szolgáltatásra tettünk javaslatot két jelentősen eltérő feladatra. Ami közös mégis bennük, hogy mindkettő elsősorban a sztochasztikus gradiens módszer (SGD – stochastic gradient descent) elosztott implementációjában a modellek tanítása során szükséges véletlen sétákat menedzseli. Így ezzel mindkettő illeszkedik a pletykaalapú tanulási paradigmához. Továbbá mindkettővel kapcsolatban azonos elvárások teljesülnek, mégpedig hogy a séták legyenek agilisek, hosszú-életűek és hatékonyak. A módszereket a fentebb részletezett okostelefonos trace-en értékeltük ki.

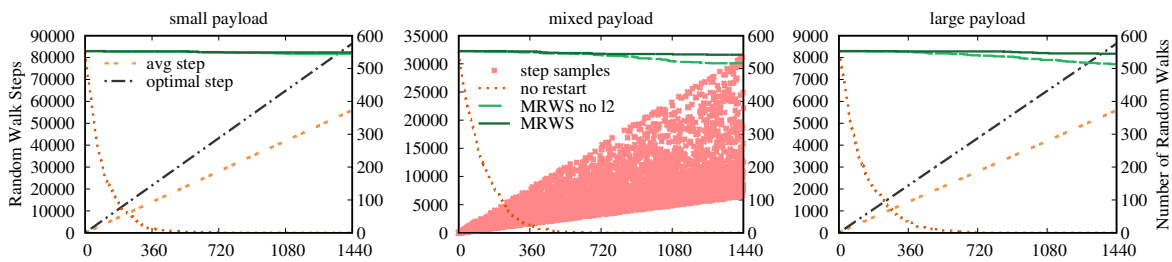
Az első feladat esetében az a célunk, hogy az egész hálózatban lehetőleg minél kevesebb, legjobb esetben egyetlen véletlen séta legyen. Ez azért fontos, mert így az SGD módszer esetén a privacy büdzsé alacsony marad. Ezzel pedig elérhető a differential privacy, amennyiben mindemellett az SGD modell a megfelelő mértékű zajjal lett terhelve. Olyan egyetlen véletlenséta-menedzselő eljárást mutattunk be erre a feladatra, amely azt használja ki, hogy minden érintett csomópont fenntartja az utolsó lokálisan ismert állapotot, és emellett megosztják pletykaalapú módon egymással a véletlen séta haladását. Ez még abban az esetben is egy relatív kis méretű közös állapotnak számít, ha a véletlen séta állapota egyébként túl nagy. Ezeknek az ismeretében minden résztvevő el tudja dönteni, hogy a nála lévő állapottal szükséges-e az újraindítás. A legnagyobb valószínűséggel az a résztvevő fogja újraindítani az eltűnő modellt, amelynél a legközelebbi utolsó állapot megtalálható.



3. ábra. Az általunk javasolt egyetlen véletlen sétát menedzselő eljárás kiértékelése 5% mesterséges üzeneteldobás mellett. A véletlen séták számát pontokkal jelöltük. Ezenkívül négyzetekkel jelöltük a legtovább eljutott lépésszámot, eltérő árnyalat más véletlen sétához tartozik.

Bemutattuk, hogy egy okostelefon-hálózatban teljesülnek a fent vázolt követelmények. Kijelenthető az, hogy a javasolt módszer robusztus a fő paraméterére, az újraindítási időzítőre δ , ami alapján a lokális döntés megszületik. Amint a 3. ábrán láthatjuk, még a valóságtól elrugaszkodott, extrém forgatókönyv, 5% mesterséges üzeneteldobás mellett is elfogadható teljesítményre képes. Ebben az esetben érzékenyebb a δ -ra, de ésszerű beállítás mellett jó kompromisszum érhető el hatékonyság és agilitás között.

A második probléma esetében egy olyan környezetben oldottuk meg a véletlen séta menedzselését, ahol minden felhasználó számára elérhető az a szolgáltatás, hogy egyedi véletlen sétát indítson a hálózatban. Ebben az esetben a hálózat méretével nagyságrendben azonos darabszámú ($O(n)$) véletlen séta menedzselésének igényére lehet számítani. Erre a problémára mutattunk be egy olyan módszert, amely három konceptuális szinttel rendelkezik, annak eldöntésére, hogy szükséges-e egy séta újraindítása. Az első, legalacsonyabb szinten egy lokális mechanizmus működik, ahol az a csomópont, amely befejezte a séta továbbítását, még megfigyeli, hogy a következő lépés is sikeres lesz-e. Amennyiben nem kap visszaigazolást a továbbküldésről, életbe lépteti az újraindítást. Akkor lépünk a következő szintre, ha a megfigyelő csomópont kiesik, ekkor a küldő felkér egy másik szomszédot a megfigyelésre. Ugyanakkor ez a második szintű megfigyelő hiba esetén nem rendelkezik a séta legaktuálisabb végállapotával, így egy broadcast üzenettel továbbítja az újraindítási igényt. Minden csomópont csak néhány, paraméter által rögzített számú korábbi állapotot tud megjegyezni. Ez a szint hasonlóan működik egy δ időzítővel, mint amit az egyetlen



4. ábra. Az általunk javasolt módszer eredményei kis, vegyes és nagy csomagmérettel, átlagos online résztvevővel azonos számú séta menedzselése esetén

véletlen séta esetében bemutattunk. Előfordulhat, hogy e két szint közül egyik sem jár sikerrel, ezért a véletlen séta kihalna. Ebben az esetben ezt elosztott módon már nem tudjuk elkerülni. Ekkor lépünk a harmadik szintre, ahol központi módon a séta tulajdonosa, amint érzékelt a leállást, végrehajtja az újraindítást.

Az általunk vizsgált összes forgatókönyv esetében a hibák jelentős hányadával az első, lokális szinten megbirkóztak a hálózat résztvevői. A hibák kis része eszkalálódott a második szintre. Ebben az esetben az egyidejű broadcast üzenetek darabszáma és mérete lényegében elhanyagolható mértékű volt. Ez annak köszönhető, hogy a vizsgált esetekben egyszerre, egy időben kevés hiba került át a második szintre. Másrészt ezeknek az üzeneteknek relatívan kisebb a mérete, mint a véletlen séta állapotának. A teljes kiértékelés során csupán néhány esetben jutottunk el a harmadik, központi irányítású szintre. Ezenkívül demonstráltuk a véletlen séták sebességét is, amely nagyon közel van az optimálishoz. Egy kivonat látható ezekről az eredményekről a 4. ábrán.

Főbb eredmények:

- Egy teljesen elosztott módszer egyetlen véletlen séta menedzselésére.
- Egy robusztus módszer $O(n)$ véletlen séta menedzselésére olyan hálózatban, ahol n résztvevő van.
- Kapcsolódó publikációk: [2, 6]

István Hegedűs, Árpád Berta, and Márk Jelasity. Robust decentralized differentially private stochastic gradient descent. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*, 7(2):20–40, June 2016

Árpád Berta and Márk Jelasity. Decentralized management of random walks over a mobile phone network. In *2017 25th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP)*, pages 100–107, Mar 2017

Hivatkozások

- [1] Árpád Berta, István Hegedűs, and Márk Jelasity. Dimension reduction methods for collaborative mobile gossip learning. In *2016 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP)*, pages 393–397, Feb 2016.
- [2] Árpád Berta and Márk Jelasity. Decentralized management of random walks over a mobile phone network. In *2017 25th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP)*, pages 100–107, Mar 2017.
- [3] Árpád Berta, Vilmos Bilicki, and Márk Jelasity. Defining and understanding smartphone churn over the internet: a measurement study. In *14-th IEEE International Conference on Peer-to-Peer Computing (P2P)*, pages 1–5, Sept 2014.
- [4] Árpád Berta, István Hegedűs, and Róbert Ormándi. Lightning fast asynchronous distributed k-means clustering. In *22th European Symposium on Artificial Neural Networks, ESANN 2014*, pages 99–104, 2014.
- [5] Gábor Danner, Árpád Berta, István Hegedűs, and Márk Jelasity. Robust fully distributed mini-batch gradient descent with privacy preservation. *Security and Communication Networks*, 2018:15, 2018.

-
- [6] István Hegedűs, Árpád Berta, and Márk Jelasity. Robust decentralized differentially private stochastic gradient descent. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*, 7(2):20–40, June 2016.
- [7] István Hegedűs, Árpád Berta, Levente Kocsis, András A. Benczúr, and Márk Jelasity. Robust decentralized low-rank matrix decomposition. *ACM Transactions on Intelligent Systems and Technology*, 7(4):62:1–62:24, May 2016.
- [8] Róbert Ormándi, István Hegedűs, and Márk Jelasity. Gossip learning with linear models on fully distributed data. *Concurrency and Computation: Practice and Experience*, 25(4):556–571, 2013.
- [9] Zoltán Szabó, Vilmos Bilicki, Árpád Berta, and Zoltán Richárd Jánki. Smartphone-based data collection with Stunner using crowdsourcing: Lessons learnt while cleaning the data. In *The Twelfth International Multi-Conference on Computing in the Global Information Technology (ICCGI 2017)*, pages 28–35, Jul 2017.
- [10] Zoltán Szabó, Vilmos Bilicki, Árpád Berta, and Zoltán Richárd Jánki. Smartphone-based data collection with Stunner, the reality of peer-to-peer connectivity and web real-time communications using crowdsourcing: Lessons learnt while cleaning the data. *International Journal On Advances in Software*, 11(1-2):120–130, 2018.
- [11] Zoltán Szabó, Krisztián Téglás, Árpád Berta, Márk Jelasity, and Vilmos Bilicki. Stunner: A smart phone trace for developing decentralized edge systems. In José Pereira and Laura Ricci, editors, *Proceedings of the 19th International Conference on Distributed Applications and Interoperable Systems (DAIS 2019)*, pages 108–115, Cham, 2019. Springer International Publishing.