

ხელოვნური ინტელექტი 6G ქსელაბისთვის

სერგო შავგულიძე

ტექნიკურ მეცნიერებათა დოქტორი,
საქართველოს ტექნიკური უნივერსიტეტის პროფესორი

ნოდარ უღრელიძე

ტექნიკურ მეცნიერებათა დოქტორი,
კავკასიის უნივერსიტეტის პროფესორი



კავკასიის უნივერსიტეტი
CAUCASUS UNIVERSITY

სერგო შავგულიძე, ნოდარ უღრელიძე
ხელოვნური ინტელექტი 6G ქსელებისთვის

Sergo Shavgulidze, Nodar Ugrelidze
ARTIFICIAL INTELLIGENCE FOR 6G NETWORKS

წინამდებარე წიგნი წარმოგვიდგენს ახალ ტექნოლოგიებსა და კვლევის პერსპექტივებს AI-ზე დაფუძნებულ 6G მობილურ საკომუნიკაციო სისტემებში, ასევე ძირითად ტექნიკურ გამოწვევებს, დაკავშირებულს ასეთი სისტემების დანერგვასთან, რომელიც გაკეთებულია უახლეს პუბლიკაციებსა და სამეცნიერო მიღწევებზე დაყრდნობით. კერძოდ, განხილულია 6G ფიქტურ ქსელებში ფიზიკურ ფენასთან და ლინკის ფენასთან დაკავშირებული საკითხები, სადაც ML-ს შეუძლია მნიშვნელოვანი სარგებელი მოიტანოს; წარმოდგენილია ცხრა გამოწვევა, რომლებიც უნდა გადაიჭრას AI/ML და უსადენო კომუნიკაციების ინტერდისციპლინარული სფეროებით, განსაკუთრებული აქცენტით 6G უსადენო ქსელებზე; შემოთავაზებულია ინტელექტუალურ-ენდოგენური ქსელის კონცეფცია, რომელიც მიზნად ისახავს ქსელებში ცოდნის გრაფის შემოტანას და AI ტექნოლოგიების დანერგვას, რათა დაახასიათოს, შექმნას, გამოიკვილოს და გამოიყენოს ქსელის მრავალგანზომილებიანი სუბიექტური და ობიექტური ცოდნა; შემოთავაზებულია DL-ზე დაფუძნებული მეთოდების გამოყენება B5G/6G ქსელების დიზაინსა და ექსპლუატაციისთვის, რომელიც განსხვავდება ლიტერატურაში აღწერილი სტანდარტული, ბუფერის დახმარებით შემუშავებული რეტრანსლატორების შერჩევის მიდგომებისგან; წარმოდგენილია 6G ქსელებისთვის AI-ზე დაფუძნებული ქსელის ფენებად დაყოფის არქიტექტურა, რომელიც უზრუნველყოფს სინერგიას AI-ისა და ქსელის დაყოფას შორის, რითაც ხელს უწყობს ქსელის ინტელექტუალურ მართვას და მხარს უჭერს განვითარებად AI სერვისებს; ნაჩვენებია, თუ როგორ შეუძლია DML-ს ზოგადად, და კონკრეტულად FL-ზე დაფუძნებული სტრატეგიებით, წვლილი შეიტანოს 6G ხედვის რეალიზებაში და დაამყაროს ბალანსი კომუნიკაციასა და გამოთვლით შემზღვევებს შორის; განხილულია ფედერირებული DRL-ის მიდგომა ღია RAN-ის ფენებად დაყოფისთვის, რომლის მიხედვითაც MVNO-ები თანამშრომლობენ თავიანთ DRL-ზე დაფუძნებული RAN-ის დაყოფის მოდელების მუშაობის გასაუმჯობესებლად; მიმოხილულია გაყოფილი სწავლების პრინციპი და ნაჩვენებია მისი უსადენო პერიფერიულ ქსელებთან ინტეგრაცია; შესწავლილია AI-ზე ორიენტირებული 6G იმერსიული სერვისების რესურსების უზრუნველყოფის მიდგომა, Alimers-6G.

ვფიქრობთ, რომ მასალა სასარგებლო იქნება და დახმარებას გაუწევს საინფორმაციო ტექნოლოგიებისა და კომუნიკაციების დარგში მომუშავე სპეციალისტებს და აკადემიურ პერსონალს, ბაკალავრიატის მაღალი კურსის სტუდენტებს, მაგისტრანტებსა და დოქტორანტებს.

რედაქტორი — მარინა ჟღენტე

დიზაინი და დაკაბადონება — მია ფეიქრიშვილი

ნაშრომი შესრულებულია სსიპ — შოთა რუსთაველის ეროვნული სამეცნიერო ფონდის გრანტის (FR-22-4050) ფარგლებში.

დამტკიცებულია მონოგრაფიად კავკასიის უნივერსიტეტის კავკასიის ტექნოლოგიების სკოლის სარედაქციო კოლეგიის მიერ: 29.10.2024, ოქმი № 2.

© სერგო შავგულიძე, ნოდარ უღრელიძე, 2024.

© კავკასიის უნივერსიტეტის გამომცემლობა, თბილისი 2024
CAUCASUS UNIVERSITY PUBLISHING HOUSE, Tbilisi 2024

ISBN 978-9941-9935-0-3

www.cu.edu.ge

ყველა უფლება დაცულია.

დაუშვებელია ამ გამოცემის მთლიანი ან ნაწილობრივი რეპროდუქცირება გამომცემლის წერილობითი ნებართვის გარეშე.

სარჩევი

წინასიტყვაობა	5
თა30 1. ხელოვნური ინტელექტი 6G ქსელებისთვის: ტექნოლოგიების განვითარება და სტანდარტიზაცია . . .9	
1.1. შესავალი	9
1.2. ML ტექნიკის მიმოხილვა უსადენო ქსელებში	10
1.3. უსადენო ქსელები: კერძო შემთხვევების შესწავლა	13
1.4. ML-ის როლი სტანდარტიზაციაში	14
1.5. ღია გამოწვევები და საგზაო რუკა ML ტექნიკის დანერგვისთვის	18
1.6. პირველი თავის დასკვნა	20
თა30 2. ხელოვნური ინტელექტის და უსადენო კომუნიკაციების ცხრა გამოწვევა 6G ტექნოლოგიისთვის . . 21	
2.1. შესავალი	21
2.2. ძირითადი გამოწვევები ღრმა ნეირონულ ქსელებში უსადენო 6G კომუნიკაციებისთვის	22
2.3. განაწილებული ნეირონული ქსელები და განაწილებული სწავლება	25
2.4. სემანტიკური კომუნიკაციები	27
2.5. მეორე თავის დასკვნა	29
თა30 3. ინტელექტუალურ-ენდოგენური ქსელები: ინოვაციური ქსელის პარადიგმა 6G-ისთვის 30	
3.1. შესავალი	30
3.2. 6G ქსელების ინტელექტუალური მოთხოვნები	31
3.3. ინტელექტუალური მოთხოვნები 6G-ის ექსპლუატაციისა და ტექნიკური მომსახურებისთვის	33
3.4. ინტელექტუალურ-ენდოგენური ქსელები 6G-ისთვის	35
3.5. IEN გამოყენების შემთხვევა: გამტარუნარიანობის გაუმჯობესება საჰაერო საბაზო სადგურების მეშვეობით	39
3.6. მომავალი სამუშაოების ანალიზი	41
3.7. მესამე თავის დასკვნა	42
თა30 4. ღრმა სწავლებაზე ორიენტირებული და ბუფერზე დაფუძნებული კოოპერატიული ქსელები B5G/6G-ისთვის 43	
4.1. შესავალი	43
4.2. ტექნიკური გამოწვევები BACN ქსელებისთვის	45
4.3. DL-ის პოტენციური ტექნოლოგიები	47
4.4. DL-ით გაძლიერებული BACN-ებისთვის შემოთავაზებული სტრუქტურები	48
4.5. მომავალი კვლევითი მიმართულებები	52
4.6. მეოთხე თავის დასკვნა	54
თა30 5. AI-ზე დაფუძნებული ქსელის ფენებად დაყოფა 6G ქსელებისთვის 55	
5.1. შესავალი	55
5.2. AI-ზე დაფუძნებული ქსელის ფენებად დაყოფა	56
5.3. AI ფენებად დაყოფისთვის	59
5.4. ფენებად დაყოფა AI-ისთვის	63
5.5. კერძო შემთხვევის შესწავლა	65
5.6. მომავალი კვლევითი მიმართულებები	67
5.7. მეხუთე თავის დასკვნა	67

თაპი 6. განაწილებული მანქანური სწავლების გამოყენების ეფექტიანობა 6G ქსელებისთვის	68
6.1. შესავალი	68
6.2. კომპრომისი კომუნიკაციასა და გამოთვლებს შორის	69
6.3. განაწილებული სწავლება 6G-ისთვის	70
6.4. კერძო შემთხვევის შესწავლა: FL-ის გამოყენება სპექტრზე დინამიკური წვდომისთვის.	71
6.5. 6G-ში DML-ის გამოყენების მთავარი გამოწვევები	75
6.6. მეექვსე თავის დასკვნა	77
თაპი 7. ფედერირებული ღრმა განმტკიცებული სწავლება 6G ქსელებში ღია RAN-ის დაყოფისთვის.	78
7.1. შესავალი	78
7.2. DRL, როგორც RAN-ის ფენებად დაყოფის საშუალება.	80
7.3. FDRL -ზე დაფუძნებული კოოპერატიული RAN-ის ფენებად დაყოფა MVNO-ებს შორის	81
7.4. ექსპერიმენტის რიცხვითი შედეგები.	84
7.5. მეშვიდე თავის დასკვნა და სამომავლო სამუშაოები	88
თაპი 8. გაყოფილი სწავლება 6G პერიფერიულ ქსელებში	89
8.1. შესავალი	89
8.2. FL-ის და SL-ის შედარება	91
8.3. არქიტექტურული დიზაინი გაყოფილი პერიფერიული სწავლებისთვის	92
8.4. რესურსების კუთხით ეფექტიანი გაყოფილი სწავლების სტრუქტურები	93
8.5. რესურსების მართვა SL-ისთვის: ერთი ფიჭის პერსპექტივა	96
8.6. რესურსების მართვა SL-ისთვის: ქსელური ინჟინერიის პერსპექტივა	97
8.7. ღია პრობლემები და მომავალი კვლევითი მიმართულებები	99
8.8. მერვე თავის დასკვნა	100
თაპი 9. AI-ზე ორიენტირებული რესურსების უზრუნველყოფა 6G იმერსიული სერვისებისთვის	101
9.1. შესავალი	101
9.2. Almers-6G-ის სტრუქტურა	102
9.3. სისტემის მოდელი	105
9.4. რესურსების უზრუნველყოფა	107
9.5. ჰეტეროგენული აუქციონი	108
9.6. კომპიუტერული სიმულაციის შედეგები	110
9.7. მეცხრე თავის დასკვნა და სამომავლო კვლევითი სამუშაო	112
ბოლოთქმა	113
ლიტერატურა	114
აბრევიატურები და აკრონიმები.	120

წინასიტყვაობა

საერთაშორისო მობილური ტელეკომუნიკაციების (IMT) სისტემების დაპროექტება 2030 წლისთვის და შემდგომ პერიოდში მოითხოვს ახალ ტექნოლოგიურ გადაწყვეტილებებს. ხელოვნური ინტელექტი (AI), ეს არის კომპიუტერული მეცნიერების დარგი, რომელიც მიზნად ისახავს ინტელექტუალური კომპიუტერული მანქანის/პროგრამის შექმნას, რომელსაც ექნება უნარი მიაღწიოს ადამიანის ინტელექტის გაგებას. მეცნიერები მიიჩნევენ, რომ AI მნიშვნელოვან როლს შეასრულებს მომავალ სატელეკომუნიკაციო სისტემებში, სადაც ტექნოლოგიურმა ინოვაციებმა უნდა შეავსონ ერთმანეთი. ეს განსაკუთრებით მნიშვნელოვანია, რადგან IMT-2030-ის და შემდგომი პერიოდის ტექნოლოგიები შეიძლება განიხილებოდეს, როგორც ფართოდ გავრცელებული ზოგადი დანიშნულების სისტემა და არა მხოლოდ გაუმჯობესებული ტექნოლოგია, რაც გამოიწვევს რთულ ტექნიკურ დამოკიდებულებებს. მომავალი მობილური სისტემა უნდა იყოს უფრო ქმედუნარიანი და მას უნდა შეეძლოს უფრო მრავალფეროვანი სერვისების მხარდაჭერა, რაც გაზრდის ქსელის სირთულეს. AI-ის გამოყენება იქნება ბუნებრივი ყველგან, სადაც ეს იქნება საჭირო მომავალ ქსელში, მათ შორის ფიზიკური ფენის დიზაინის, რადიორესურსების მართვის, ქსელის უსაფრთხოების და აპლიკაციების გაუმჯობესების პროცესში, ასევე ქსელის არქიტექტურის შემუშავებისას, რასაც მიყვავართ მრავალფენიანი, ღრმა, ინტეგრირებული, ინტელექტუალური ქსელის დიზაინამდე. მომავალი ქსელი ასევე ხელს შეუწყობს განაწილებულ AI-ის, როგორც სერვისს უფრო ფართომასშტაბიანი ინტელექტუალურობისთვის. გარდა ძირითადი საკომუნიკაციო მახასიათებლების მნიშვნელოვანი განახლებისა, ბევრი მკვლევარი ხელოვნურ ინტელექტს მიიჩნევს მეექვსე თაობის (6G) ყველაზე მნიშვნელოვან შემადგენელ ნაწილად, რადგან თანამედროვე მანქანური სწავლების (ML) ტექნიკის გამოყენება მიჩნეულია როგორც საუკეთესო გადაწყვეტა ბევრ უკიდურესად ექსტრემალურ სიტუაციაში და რთულ სცენარში. ქსელის ინტელიგენტიზაცია იქნება ახალი ტენდენცია, რომელიც დაკავშირებულია ჰეტეროგენული მოწყობილობების ექსპონენციალურად მზარდი რაოდენობით შექმნილ გამოწვევებთან. თუმცა, სხვა სფეროებში ML-ის გამოყენებასთან შედარებით, ინტელექტუალური ქსელის მიმდინარე კვლევას ჯერ კიდევ დიდი გზა აქვს გასავლელი ავტომატურად კონფიგურირებული ფიჭური საკომუნიკაციო სისტემების რეალიზებისთვის. 6G-ში ამ ტექნიკის სრული გამოყენებისთვის უნდა მოგვარდეს სხვადასხვა პრობლემა საკომუნიკაციო სისტემის, ML არქიტექტურისა და გამოთვლების ეფექტიანობის თვალსაზრისით.

წინამდებარე წიგნში, რომელიც მიმოხილვითი ხასიათისაა, გრძელდება საუბარი 6G-ის თემატიკასთან დაკავშირებულ კვლევის შედეგებზე, რაც უახლეს პუბლიკაციებსა და სამეცნიერო მიღწევებზე დაყრდნობითაა მომზადებული. ამასთან, ეს წიგნი განსხვავდება ჩვენ მიერ ქართულ ენაზე გამოცემული, 6G-ისადმი მიძღვნილი წინა 3 წიგნისგან და ავსებს მათ. საქმე ისაა, რომ ბოლო პერიოდში ხელოვნურ ინტელექტზე ორიენტირებული 6G ქსელების შემუშავების მიმართულებით ძალიან ბევრი საინტერესო და სერიოზული საჟურნალო პუბლიკაცია გახდა ხელმისაწვდომი. გადაწყვეტით, ჩვენი აზრით საუკეთესო ნაშრომები შეგვეკრიბა (ყველა მათგანი მოცემულია წიგნის ბოლოს მითითებულ ლიტერატურაში), მოგვეხდინა მათი სისტემატიზაცია და ერთიან სტრუქტურაში გაერთიანება, ქართულ ენაზე წარმოგვედგინა ამ ნაშრომებში მოყვანილი ძირითადი შედეგები და დაგვეწერა ეს წიგნი. შებლუდული ფორმატის გამო, ავარჩიეთ მხოლოდ რამდენიმე პერსპექტიული მიმართულება, რომლებიც თავების მიხედვით შეჯამებულია შემდეგნაირად:

ცნობილია, რომ 6G ქსელები იქნება განსაკუთრებულად რთული, რაც მოითხოვს უფრო მეტ დროით, ღირებულებასთან დაკავშირებულ და მართვის ძალისხმევას. მეორე მხრივ, მობილური ქსელის ოპერატორები ითხოვენ, რომ ეს ქსელები იყოს ინტელექტუალური, თვითორგანიზებული და ხარჯთეფექტური საოპერაციო დანახარჯების (OPEX) შესამცირებლად. ML, რომელიც წარმოადგენს AI-ის ერთ-ერთ განშტოებას, იძლევა პასუხს ბევრ ასეთ გამოწვევაზე პრაგმატული გადაწყვეტილებების მიწოდებით, რომელსაც შეუძლია მთლიანად შეცვალოს უსადენო ქსელის ტექნოლოგიების მომავალი. რამდენიმე კერძო შემთხვევის შესწავლის მაგალითზე დაყრდნობით, **პირველ თავში** ჩვენ მოკლედ განვიხილავთ მნიშვნელოვან საკითხებს, დაკავშირებულს ფიზიკურ ფენასთან (PHY) და ლინკის ფენასთან ფიჭურ ქსელებში, სადაც ML-ს შეუძლია მნიშვნელოვანი სარგებელი

მოიტანოს. ასევე განვიხილავთ სტანდარტიზაციის აქტივობას უსადენო ქსელებში ML-ის გამოყენებასთან დაკავშირებით და სტანდარტების შემუშავებელი ორგანოების მზადყოფნის ვადებს ამ ცვლილებებთან ადაპტაციისთვის. თავის ბოლოს, წარმოვადგენთ უსადენო ტექნოლოგიებში ML-ის გამოყენებასთან დაკავშირებულ ძირითად პრობლემებს და განვიხილავთ შესაძლო გზებს ზოგიერთი მათგანის მოსაგვარებლად 6G უსადენო ქსელებში.

ბოლო წლების განმავლობაში, AI ტექნიკა, განსაკუთრებით ML, წარმატებით იქნა გამოყენებული სხვადასხვა სფეროში, რამაც განამტკიცა ფართოდ გავრცელებული მოსაზრება, რომ AI ითამაშებს მნიშვნელოვან როლს მომავალ უსადენო კომუნიკაციებში. ამ მისწრაფების განსახორციელებლად, **მეორე თავში** ჩვენ წარმოვადგენთ ცხრა გამოწვევას, რომლებიც უნდა გადაიჭრას AI/ML და უსადენო კომუნიკაციების ინტერდისციპლინარული სფეროებით, განსაკუთრებული აქცენტით 6G უსადენო ქსელებზე. კონკრეტულად, ეს თავი ახდენს გამოთვლების კუთხით ცხრა გამოწვევის კლასიფიცირებას AI-ის მიმართულებით, კერძოდ, ღრმა ნეირონულ ქსელებში, განაწილებულ ნეირონულ ქსელებსა და განაწილებულ სწავლებაში და სემანტიკურ კომუნიკაციებში.

6G ეპოქაში გათვალისწინებული ფუნქციური აპლიკაციები, რომლებიც წარმოდგენილია ჰოლოგრაფიული კომუნიკაციით, სრული სენსორული კომუნიკაციით, ინტელექტუალური ტრანსპორტით და ჭკვიანი წარმოებით, ვირტუალური და ფიზიკური ინტეგრაციის, იმერსიულობის, კონტექსტუალიზაციისა და პერსონალიზაციის თვალსაზრისით დიდ მოთხოვნებს აყენებს 6G ქსელების ინტელექტუალიზაციაზე. წესებზე დაფუძნებულ ალგორითმებზე შექმნილი ქსელის მოქმედი ამჟამინდელი მექანიზმი შეზღუდულია წინასწარ დადგენილი მკაცრი პირობებით. შესაბამისად, ძნელია დინამიკურად ადაპტირება მომხმარებელთა მუდმივად ცვალებად საჭიროებებთან და ქსელურ გარემოსთან. იმისათვის, რომ ქსელმა შეძლოს ეფექტიანად დააგროვოს და დამოუკიდებლად გამოიყენოს საოპერაციო გამოცდილება, და განახორციელოს ქსელის მუშაობის, მართვის და კონტროლის შესაძლებლობების დამოუკიდებელი ევოლუცია, **მესამე თავი** გვთავაზობს ინტელექტუალურ-ენდოგენური ქსელის კონცეფციას (IEN). IEN მიზნად ისახავს ქსელებში ცოდნის გრაფის შემოტანას და AI ტექნოლოგიების დანერგვას, რათა დაახასიათოს, შექმნას, გამოიკვლიოს და გამოიყენოს ქსელის მრავალგანზომილებიანი სუბიექტური და ობიექტური ცოდნა. მიღებული ცოდნის საფუძველზე, IEN-ს შეუძლია განახორციელოს სტერეო აღქმა, გადაწყვეტილების დედუქცია და დინამიკური რეგულირება; ასე რომ ქსელი შეიძლება შეიცვალოს საჭიროებისამებრ ნებისმიერი ახალი სერვისისთვის. ამ თავში მოცემულია IEN-ის ძირითადი ცნებები, მიზნები და მნიშვნელობები, მის ძირითად პრინციპებთან და ტექნოლოგიებთან ერთად და ნაჩვენებია, თუ როგორ რეაგირებს ის ინოვაციურად მიმდინარე ქსელის მრავალმხრივ გამოწვევებზე ინტუიციური გამოყენების შემთხვევაში.

ბუფერზე დაფუძნებულმა კოოპერატიულმა ქსელებმა (BACN) მნიშვნელოვანი ყურადღება მიიპყრო მეხუთე თაობის შემდგომი პერიოდის (B5G) და მეექვსე თაობის (6G) საკომუნიკაციო სისტემებისთვის, კრიტიკულ სცენარებში მათი პოტენციური გამოყენების გამო. **მეოთხე თავი** იკვლევს ბუფერზე დაფუძნებული რეტრანსლაციის (სარელეო გადაცემის) სხვადასხვა ტიპური აპლიკაციის სცენარს B5G/6G ქსელებში, რათა ხაზი გაუსვას BACN-ის გამოყენების მნიშვნელობას. გარდა ამისა, წარმოდგენილია BACN-ში არსებული მნიშვნელოვანი ტექნიკური გამოწვევები, მათ შორის შეყოვნებაზე მკაცრი შეზღუდვები, მაღალი საიმედოობის საჭიროება, არხის მდგომარეობის შესახებ არასრულყოფილი ინფორმაციის (CSI) ქონა, გადაცემის უსაფრთხოება და ქსელის ინტეგრირებული არქიტექტურა. ამ პრობლემების გადასაჭრელად, შემოთავაზებულია ღრმა სწავლებაზე (DL) დაფუძნებული მეთოდების გამოყენება B5G/6G ქსელების დიზაინისა და ექსპლუატაციისთვის BACN-ით, რომელიც განსხვავდება ლიტერატურაში აღწერილი სტანდარტული, ბუფერის დახმარებით შემუშავებული რეტრანსლატორების შერჩევის მიდგომებისგან. კერძოდ, ჩვენ წარმოვადგენთ ორ შემთხვევას, რათა ვაჩვენოთ ცენტრალიზებული ღრმა განმტკიცებელი სწავლების (DRL) და დეცენტრალიზებული DRL-ის ეფექტიანობა ბუფერზე დაფუძნებულ არატერესტრიალურ ქსელებში. თავის ბოლოს, გამოვყოფთ სამომავლო კვლევის მიმართულებებს B5G/6G-ში, რომლებიც დაკავშირებულია BACN-ის გამოყენებასთან.

მოსალოდნელია, რომ 6G ქსელები იფუნქციონირებენ როგორც კოსმოსურ-საჰაერო-სახმელეთო ინტეგრირებული ქსელები, მათ ექნებათ მოწინავე ქსელური ვირტუალიზაცია და საყოველთაო ინტელექტი. **მეხუთე თავი** წარმოგვიდგენს 6G ქსელებისთვის AI-ზე დაფუძნებულ ქსელის ფენებად დაყოფის არქიტექტურას, რომელიც

უმრუნველყოფს სინერჯიას AI-ისა და ქსელის დაყოფას შორის, რითაც ხელს შეუწყობს ქსელის ინტელექტუალურ მართვას და მხარს დაუჭერს განვითარებად AI სერვისებს. AI-ზე დაფუძნებული გადაწყვეტილებები თავდაპირველად განიხილება ქსელის დაყოფის სასიცოცხლო ციკლის განმავლობაში, რათა ჭკვიანურად მოხდეს ქსელის ფრაგმენტების მართვა. შემდეგ შეისწავლება ქსელის დაყოფის გადაწყვეტილებები განვითარებადი AI სერვისების მხარდასაჭერად AI-ის ეგზემპლარების კონსტრუირებით და რესურსების ეფექტიანი მართვით. დაბოლოს, შესწავლილია კერძო შემთხვევა, რასაც მოჰყვება კვლევის ღია საკითხების განხილვა, რომლებიც დაკავშირებულია 5G ქსელებში AI-ის გამოყენებით ქსელის დაყოფასთან.

მობილური მოწყობილობების გამოთვლითი შესაძლებლობებისა და მეხსიერების ტევადობის მუდმივად გაუმჯობესებული სატელეკომუნიკაციო ქსელის პარადიგმების შესაბამისად, გაიზარდა კვლევითი ინტერესი განაწილებული მანქანური სწავლების (DML) ჩარჩოების შესამუშავებლად, რათა მიღწეულ იქნეს მახასიათებლების ის ძირითადი ინდიკატორები (KPI), რომლებიც მოსალოდნელია შემდეგი თაობის 5G ფიჭური ქსელებისთვის. პერიფერიულ გამოთვლებთან (Edge Computing) ერთად, ფედერირებული სწავლება (FL) გაჩნდა, როგორც DML არქიტექტურის არჩევანი დამკვიდრებულ უსადენო აპლიკაციებში. **შეექვსე თავი** წარმოგვიდგენს მიმოხილვას იმის შესახებ, თუ როგორ შეუძლია DML-ს ზოგადად, და კონკრეტულად FL-ზე დაფუძნებული სტრატეგიებით წვლილი შეიტანოს 5G ხედვის რეალიზებაში და დაამყაროს ბალანსი კომუნიკაციასა და გამოთვლით შებენებს შორის. როგორც პრაქტიკული გამოყენების შემთხვევას, ჩვენ განვიხილავთ მრავალაგენტიან განმტკიცებულ სწავლებას (RL) FL სტრუქტურაში სპექტრზე დინამიკური წვდომის (DSA) პრობლემის გადასაჭრელად და წარმოვადგენთ შეფასების შედეგებს. ასევე წარმოდგენილია თანამედროვე გამოწვევები DML მიდგომების გამოყენებისას 5G ქსელებში.

რადიოწვდომის ქსელის (RAN) ფენებად დაყოფა არის ძირითადი ელემენტი, რომელიც საშუალებას აძლევს არსებულ 5G და შემდეგი თაობის ქსელებს დააკმაყოფილოს სერვისების მოთხოვნა სხვადასხვა ვერტიკალში. თუმცა, ამ სერვისების მოთხოვნების ჰეტეროგენული ბუნება, RAN-ის შეზღუდულ რესურსებთან ერთად, RAN-ის დაყოფას ძალიან რთულს ხდის. მართლაც, გამოწვევა, რომელსაც აწყდებიან მობილური ქსელის ვირტუალური ოპერატორები (MVNO) არის თავიანთი RAN-ის დაყოფის სტრატეგიების სწრაფი ადაპტირება გარემოს შეზღუდვებისა და მომსახურების მოთხოვნების ხშირ ცვლილებებთან. DL ტექნოლოგია, როგორცაა DRL, სულ უფრო მეტად განიხილება RAN-ის დაყოფის ოპერაციების მართვისა და ექსპლუატაციის ავტომატიზაციის ძირითად საშუალებად. ამასთან, DRL მოდელების განზოგადების შესაძლებლობა RAN-ის დაყოფის მრავალფეროვან გარემოში შეიძლება შეიზღუდოს მათი ძლიერი დამოკიდებულების გამო იმ გარემოს მონაცემებზე, რომელშიც ისინი გადიან ტრენინგს. FL საშუალებას აძლევს MVNO-ებს გამოიყენონ ტრენინგის უფრო მრავალფეროვანი წყაროები DRL-ისთვის ამ მონაცემების შეგროვების მაღალი ღირებულების გარეშე სხვადასხვა RAN-იდან.

მეშვიდე თავში ჩვენ წარმოვადგენთ ფედერირებული DRL-ის მიდგომას ღია RAN-ის ფენებად დაყოფისთვის. ამ მიდგომით, MVNO-ები თანამშრომლობენ თავიანთი, DRL-ზე დაფუძნებული RAN-ის დაყოფის მოდელების მუშაობის გასაუმჯობესებლად. თითოეული MVNO ახორციელებს DRL-ის მოდელის ტრენინგს და აგზავნის მას აგრეგაციისთვის. აგრეგირებული მოდელი შემდეგ იგზავნება უკან თითოეულ MVNO-ში დაუყოვნებელი გამოყენებისა და შემდგომი ტრენინგისთვის. კომპიუტერული სიმულაციის შედეგები აჩვენებს შემოთავაზებული DRL მიდგომის ეფექტიანობას.

განაწილებული პერიფერიული გამოთვლითი რესურსების ზრდასთან ერთად, 5G მობილური ქსელი გადაიქცევა დაკავშირებული ინტელექტის ქსელად. ამ კუთხით, წინადადება FL-ის მობილურ პერიფერიაში ჩართვის შესახებ მნიშვნელოვანი ინტერესი გამოიწვია ბოლო წლებში. თუმცა, FL-ის განხორციელება მნიშვნელოვანი გამოწვევების წინაშე დგას, რადგან შეზღუდული რესურსების მქონე მასობრივმა საგანთა ინტერნეტის (IoT) მოწყობილობებმა ნაკლებად სავარაუდოა, რომ მხარი დაუჭირონ მოწყობილობაზე მოდელის ტრენინგს. ამან გამოიწვია გაცოფილი სწავლების (SL) გაჩენა, რაც სერვერებს საშუალებას აძლევს გაუმკლავდნენ ძირითადი ტრენინგის დატვირთვას, მონაცემთა კონფიდენციალურობის შენარჩუნებით. **მერვე თავში** ჩვენ წარმოვადგენთ SL-ის მოკლე მიმოხილვას და ვისაუბრებთ უსადენო პერიფერიულ ქსელებთან მისი ინტეგრაციის შესახებ. ჩვენ ვიწყებთ მორგებული 5G არქიტექტურის ილუსტრირებით, რომელიც მხარს უჭერს გაცოფილ პერიფერიულ სწავლებას (SEL). შემდეგ, ჩვენ განვიხილავთ SEL-ის დიზაინის კრიტიკულ საკითხებს, მათ შორის რესურსების

ეფექტიანი სწავლების ჩარჩოებს და რესურსების მართვის სტრატეგიებს ერთი პერიფერიული სერვერის ქვეშ. გარდა ამისა, ქსელის პერსპექტივიდან, ჩვენ ვაფართოებთ შესაძლებლობებს მრავალპერიფერიულ სცენარებზე, ვიკვლევთ შესაბამის თანამშრომლობას და მოდელის განთავსებას/მიგრაციას. დაბოლოს, ჩვენ განვიხილავთ ღია პრობლემებს SEL-ისთვის, მათ შორის კონვერგენციის ანალიზს, ასინქრონულ SL-ს და ეტიკეტების კონფიდენციალურობის დაცვას.

NG ტექნოლოგიების მიღწევებიდან გამომდინარე დიდი ინტერესი გაჩნდა NG იმერსიული სერვისების მიმართ. NG იმერსიული სერვისები მოიცავს სხვადასხვა ტარებად მოწყობილობას, რათა უზრუნველყონ ვირტუალური სცენების მაღალი ხარისხის აღქმა მომხმარებლებისთვის. სერვისის მომხმარებლების (SU) და სერვისის პროვაიდერების (SP) აქტიური მონაწილეობა განაპირობებს ტარებადი მოწყობილობების სწრაფ გავრცელებას. განვითარებულმა ტექნოლოგიებმა, როგორცაა ღრუბლოვანი გამოთვლები და პერიფერიული გამოთვლები, ასევე ხელი შეუწყო NG იმერსიული სერვისების ბაზრის სწრაფ განვითარებას, რაც შესაძლებელს გახდის საყოველთაოდ იმერსიულ სერვისებს. თუმცა, ტარებადი მოწყობილობების გავრცელებამ გამოიწვია სერიოზული გამოწვევები, მათ შორის რესურსების იერარქიული უზრუნველყოფა, სერვისებსა და რესურსებს შორის დროის მიხედვით დამოკიდებულებები და რესურსების არაერთგვაროვანი მოთხოვნები. ამ ხარვეზის შესავსებად, **მეცხრე თავი** გვთავაზობს AI-ზე ორიენტირებულ NG იმერსიული სერვისების რესურსების უზრუნველყოფის მიდგომას, Almers-NG-ის, დიდი და პატარა რეგიონების პერსპექტივის შესაბამისად. დიდ რეგიონში ჰეტეროგენული რესურსებია გამოყოფილი SU-ების აღქმის გამოცდილების მოთხოვნების დასაკმაყოფილებლად. რესურსების უზრუნველყოფის პრობლემას წყვეტს ლიაპუნოვის ოპტიმიზაციის ალგორითმი, რომელიც ეფუძნება კონტექსტით გაცნობიერებულ იმერსიულ სწავლებას. მცირე რეგიონში კარგად შემუშავებული ბლოკჩეინზე დაფუძნებული ორმაგი ჰოლანდიური აუქციონის (BDDA) მექანიზმი გამოიყენება ჰეტეროგენული რესურსების შესატყვისობების და ფასების დასადგენად. დაბოლოს, წარმოდგენილია საილუსტრაციო სიმულაციები შემოთავაზებული სქემის ეფექტიანობის საჩვენებლად.

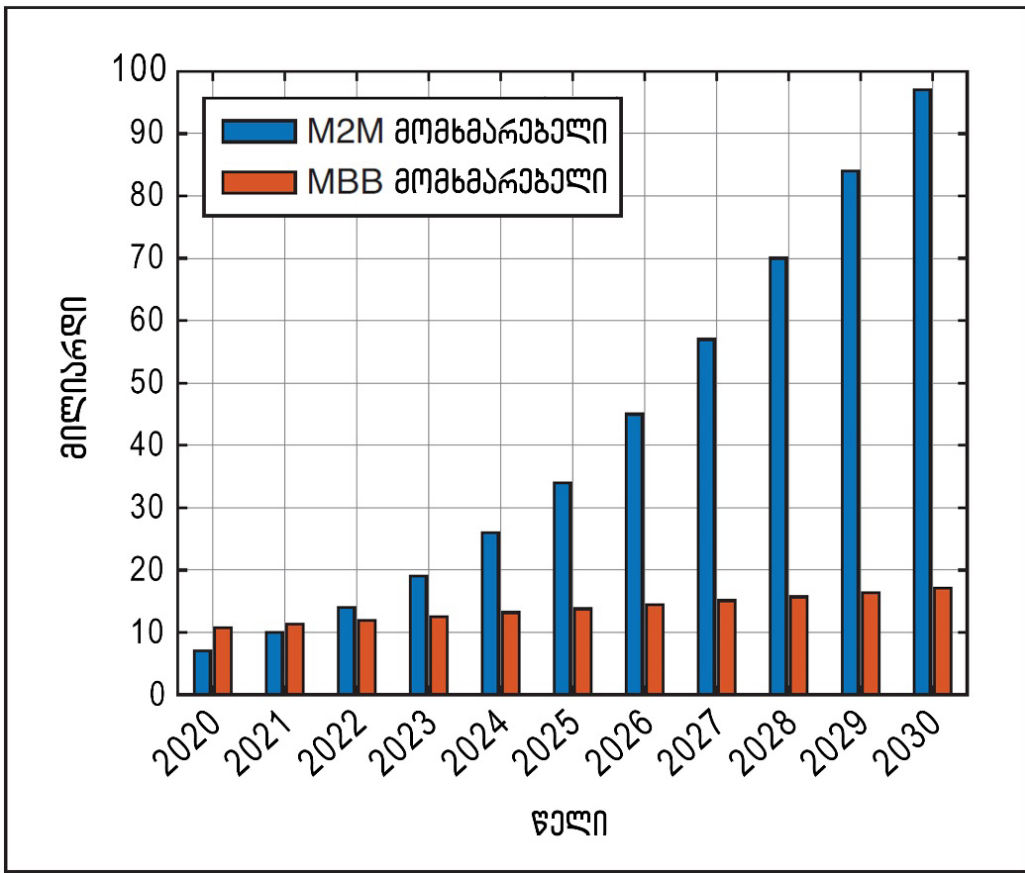
დამატებით აღვნიშნავთ, რომ წიგნის ბოლოს წარმოდგენილია გამოყენებული ინგლისურენოვანი აბრევიატურებისა და აკრონიმების მნიშვნელობები და მათი ქართული თარგმანები.

ვფიქრობთ, რომ მასალა სასარგებლო იქნება და დახმარებას გაუწევს საინფორმაციო ტექნოლოგიებისა და კომუნიკაციების დარგში მომუშავე სპეციალისტებს, აკადემიურ პერსონალს, ბაკალავრიატის მაღალი კურსის სტუდენტებს, მაგისტრანტებსა და დოქტორანტებს.

თავი 1 ხელოვნური ინტელექტი 6G ქსელებისთვის: ტექნოლოგიების განვითარება და სტანდარტიზაცია

1.1. შესავალი

გლობალური ფიჭური ტრაფიკის უპრეცედენტო ზრდა (როგორც ნაჩვენებია ნახ. 1.1-ზე) და მონაცემთა გადაცემის სიჩქარე უზარმაზარი მოთხოვნები იქცა დიდ გამოწვევად, რამაც უსადენო ინდუსტრია მიიყვანა შემდეგ მეექვსე თაობამდე, სახელად 6G. 6G-ის ეპოქა აერთიანებს ციფრულ, ფიზიკურ და ბიოლოგიურ სამყაროებს, რათა გააუმჯობესოს ადამიანების გამოცდილება და კეთილდღეობა. 6G იფუნქციონირებს ტერაჰერცულ (ტჰც) სიხშირეებზე (0.1–10 ტჰც); აქედან გამომდინარე, 6G სასარგებლოა მრავალჯერადი გამოყენების შემთხვევაში სამრეწველო აპლიკაციებში, რაც უზრუნველყოფს მონაცემთა გადაცემის მაღალ სიჩქარეს (≈ 1 ტბიტ/წმ) და უზრუნველყოფს საგნების ინტერნეტის ეფექტიან მუშაობას და ქსელის გაზრდილ დაფარვას. ხელოვნური ინტელექტი (AI) და მანქანური სწავლება (ML) გზას გაუხსნის ტჰც კომუნიკაციებს სხვადასხვა ფენაზე, მაგალითად, არხის მახასიათებლების მოპოვებისა და შეფასების მხარდაჭერას და მოდულაციის კლასიფიკაციას ფიზიკურ (PHY) ფენაში. ანალოგიურად, ლინკის ფენაში, სხივის ფორმირების დიზაინსა და არხების განაწილებას შეუძლია გამოიყენოს ML. ტჰც სისტემებში არხი შეიძლება მნიშვნელოვნად იცვლებოდეს მიკრომეტრული მასშტაბით, რაც გამოიწვევს არხის შეფასების სიხშირის და შესაბამისი ზედნადების დიდ ზრდას. ML ალგორითმებს შეუძლიათ ამ პრობლემის გადალახვა, მაგალითად, არხების პროგნოზირების გაუმჯობესებული ტექნიკის გამოყენებით.



ნახ. 1.1. გლობალური მობილური აბონენტების შეფასება მანქანა-მანქანასთან (M2M) და მობილური ფართობლოვანი (MBB) კომუნიკაციებისთვის 2020 წლიდან 2030 წლამდე

ბოლო დროს, მეხუთე თაობის (5G) სწრაფად მზარდმა გამოყენებამ წარმოშვა მრავალი გამოწვევა, მათ შორის ქსელის არქიტექტურის ძალიან დიდი სირთულე, დაბალი შეყვანება, მაღალი ღირებულება, ენერჯის გაზრდილი მოხმარება და ჰიბრიდული გრძელვადიანი ევოლუციის (LTE) ახალი რადიოს (NR) გამოყენება, რამაც მიგვიყვანა სირთულეებამდე ქსელის ოპტიმიზაციის პროცესში. ასეთ რთულ სცენარში ყურადღება გამახვილებულია ქსელის ინტელექტზე, რადგან ის მნიშვნელოვან როლს შეასრულებს ისეთი რთული პრობლემების გადაჭრაში, როგორებიცაა ქსელის თვითაღდგენა, თვითოპტიმიზაცია და თვითკონფიგურაცია.

მომავალი ქსელები გახდება „კოგნიტური“ ისე, რომ მრავალი ასპექტი, როგორცაა სპექტრის ზონდირება/გაზიარება, ფენებად დაყოფა, რადიორესურსების მართვა და მობილურობის მენეჯმენტი, დაფუძნებული იქნება ML-ზე. გარდა ამისა, მოსალოდნელია, რომ ML ფუნდამენტურ გავლენას მოახდენს 5G რადიონტერფეისზე და ის შემუშავებული იქნება ML-ის მხარდასაჭერად. ბოლოდროინდელ კვლევებში აღნიშნული მიმართულებით შესწავლილია რამდენიმე პრობლემა, მაგალითად, შემოთავაზებულია სხვადასხვა საგზაო რუკა 5G-ისთვის, მაგრამ ისინი ნაკლებად განიხილავენ ML-თან მიმართებაში არსებულ ტექნიკურ გამოწვევებს და მის აპლიკაციებს, სტანდარტიზაციის ვადებს და ML-ის 5G-ში გამოყენებასთან დაკავშირებულ სხვა საკითხებს.

რეკონფიგურირებადი ინტელექტუალური ზედაპირები (RIS) და არაორთოგონალური მრავალჯერადი წვდომა (NOMA) არის ორი ძირითადი ტექნოლოგია 5G-ისთვის. RIS-ებს შეუძლიათ ელექტრომაგნიტური ტალღების ხელახალი ფორმირება; ამიტომ, ისინი სასარგებლოა ინფორმაციის მიწოდებისთვის იქ, სადაც დაბრკოლებები ბლოკავს დანიშნულების ადგილს. RIS შეიძლება ინტეგრირებული იყოს ML-თან, რაც მას საშუალებას აძლევს მიიღოს გარემოსდაცვითი ინფორმაცია სხვადასხვა სენსორის კონფიგურაციით, ხოლო ML-ს შეუძლია ინტელექტუალურად შეისწავლოს დინამიკური პარამეტრები, რაც ამცირებს RIS-ზე დაფუძნებული ქსელების გამოთვლით ღირებულებას. ანალოგიურად, NOMA არის პერსპექტიული წვდომის ტექნიკა 5G-ისთვის. ML-ით გაძლიერებულ NOMA-ზე დაფუძნებულ ქსელებში, gNodeB (gNB) კვანძებს შეუძლიათ ჰკვიანურად განსაზღვრონ თავისი მართვის პოლიტიკა და გააუმჯობესონ გადაწყვეტილების მიღების უნარი.

დღეს არსებული ქსელები იყენებენ მოდელზე დაფუძნებულ მეთოდებს ქსელის სხვადასხვა ფუნქციის (NF) ოპტიმიზაციისთვის, ჩართული პროცესის კარგი მახასიათებლების უზრუნველყოფით. თუმცა, ეს მოდელები შეიძლება იყოს ძალიან რთული იმისთვის, რომ განხორციელდეს რეალურ ვადებში, ან შეიცავდეს აბსტრაქციის მაღალ დონეს ზოგად გარემოში ფუნქციონირებისთვის. ამის საპირისპიროდ, ML-ზე დაფუძნებულ გადაწყვეტილებებს შეუძლიათ ადაპტირება რეალურ დროში (RT) სცენარის ცვლილებებთან და ლოკალიზებულ მახასიათებლებთან, შეისწავლის რა კონკრეტულ გარემოს გადამცემების გარშემო. ამ თავში მნიშვნელოვანია ორი მხარე:

ა. ჩვენ შევისწავლით ზემოხსენებულ პრობლემებს ინდუსტრიული პერსპექტივიდან გამომდინარე და ვუჩვენებთ არსებულ „უფსკრულს“ ჩატარებულ კვლევებსა და პრაქტიკულ საჭიროებებს შორის.

ბ. ჩვენ განვიხილავთ სტანდარტიზაციის აქტივობებს ML-ის ადაპტაციის კონტექსტში უსადენო კომუნიკაციების სხვადასხვა ასპექტში, მაგალითად, არხის პარამეტრების მოპოვებასა და პოზიციონირებაში. გარდა ამისა, ჩვენ წარმოვადგენთ ძირითად საკითხებს და შესაძლო კვლევით მიმართულებებს უსადენო ქსელებში ML-ის გამოყენებასთან დაკავშირებით.

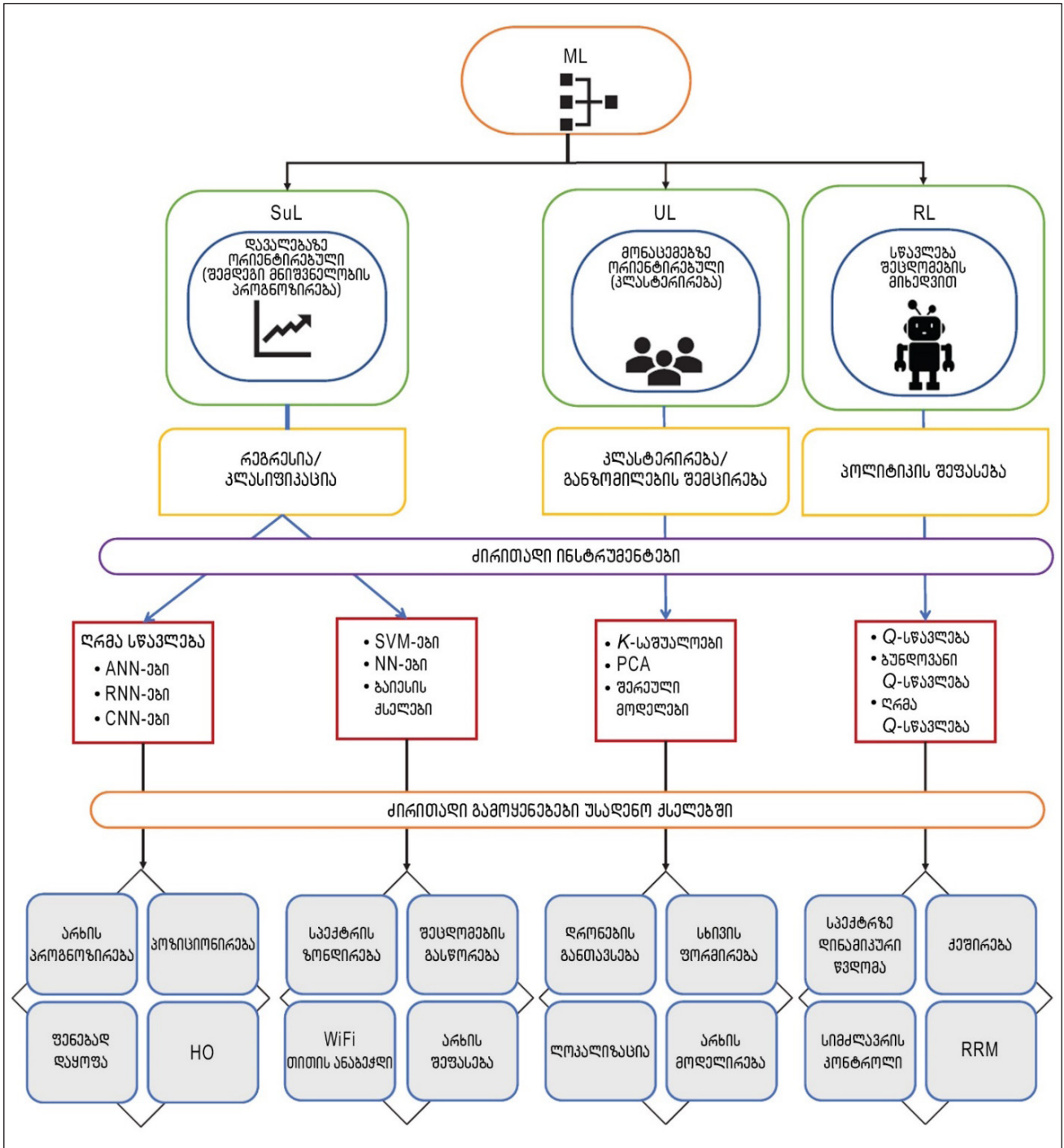
1.2. ML ტექნიკის მიმოხილვა უსადენო ქსელებში

ML არის მანქანური ტრენინგის პროცესი მონაცემების საშუალებით, პროგრამირების აშკარა გამოყენების გარეშე. ზოგადად რომ ვთქვათ, ML შედგება სამი პარადიგმისგან: ზედამხედველობითი სწავლება (SuL), ზედამხედველობის გარეშე სწავლება (UL) და განმტკიცებული სწავლება (RL). ყველა ამ პარადიგმას აქვს ტრენინგის/გამოკვლევის ფაზა, რომელიც ახდენს სწავლების ალგორითმის ოპტიმიზაციას. ყოველივე ეს მოგვიანებით შეიძლება გამოყენებულ იქნეს წინასწარმეტყველების და ექსპლუატაციის ფაზაში, რომელიც დასკვნას იძლევა უცნობი შემავალი მონაცემების მიხედვით. როგორც ნაჩვენებია ნახ. 1.2-ზე, ჩვენ მოკლედ ვაჯამებთ მათ უსადენო ქსელებში გამოყენების შემთხვევების წარმოდგენით.

SuL იყენებს ეტიკეტირებულ მონაცემთა ბაზას იმ (ფარული) ფუნქციის შესასწავლად, რომელიც, მაგალითებზე დაყრდნობით, ასახავს შემავალ მონაცემებს მოსალოდნელ გამოძვალ მონაცემებში. სტანდარტული

მეთოდები, რომლებიც გამოიყენება ზედამხედველობით სწავლებაზე დაფუძნებული პრობლემების გადასაჭრელად არის ხელოვნური ნეირონული ქსელები (ANN), დამხმარე ვექტორული მანქანები (SVM), ბაიესის ქსელები, რეკურენტული ნეირონული ქსელები (RNN) და კონვოლუციური ნეირონული ქსელები (CNN).

UL არ სარგებლობს ეტიკეტირებული მონაცემებიდან; ამის ნაცვლად, ტრენინგი ეფუძნება არაეტიკეტირებულ მონაცემთა ნაკრებს. K-საშუალოები და ძირითადი კომპონენტების ანალიზი არის ორი ძირითადი ინსტრუმენტის მაგალითი, რომლებიც გამოიყენება, შესაბამისად, კლასტერირებისა და განზომილების შემცირებისთვის.



ნახ. 1.2. ML პარადიგმების, ძირითადი ინსტრუმენტების და აპლიკაციების მიმოხილვა უსადენო ქსელებში: ANN — ხელოვნური ნეირონული ქსელები; RNN — რეკურენტული ნეირონული ქსელები; CNN — კონვოლუციური ნეირონული ქსელები; SVM — დამხმარე ვექტორული მანქანები; PCA — ძირითადი კომპონენტების ანალიზი; RRM — რადიორესურსების მართვა; HO — ჰენდოვერი.

RL არ არის დაფუძნებული ტრენინგზე; უფრო მეტიც, აგენტი/გადაწყვეტილების მიმღები სწავლობს და იღებს გადაწყვეტილებებს ონლაინ რეჟიმში, რაც ახდენს კუმულაციური ჯილდოს მაქსიმიზაციას. RL სასარგებლო მართვის პრობლემებში, სადაც აგენტი ადაპტირდება გარემო პირობების ისეთ ცვალებადობასთან, როგორცაა აპლინგში სიმძლავრის კონტროლი.

სხვადასხვა სფეროში ML-ის მნიშვნელოვანი უპირატესობების გათვალისწინებით, მისი გამოყენება ასევე განიხილება უსადენო ქსელებში, კომუნიკაციის თითქმის ყველა ფენაში. აქ ჩვენ ყურადღებას ვამახვილებთ მის გავლენაზე რადიოწვდომის ქსელებზე (RAN), განსაკუთრებით PHY ფენაზე და ლინკის ფენაზე. ზოგიერთი კერძო შემთხვევის შესწავლა, რომელიც დაფუძნებულია ML ინსტრუმენტებზე (იხ. ნახ. 1.2), მოგვიანებით იქნება წარმოდგენილი 1.3 პარაგრაფში.

PHY-ში ოპტიმიზაციის მრავალი პრობლემა არაამოზნეცილია, მაგალითად, ჯამური სიჩქარის მაქსიმიზაცია. ML არის მძლავრი ინსტრუმენტი ასეთი არაამოზნეცილი ოპტიმიზაციის პრობლემების კარგი გადაწყვეტილებების მოსაძებნად. სწავლების მოწინავე ალგორითმებზე დაყრდნობით, 6G ქსელები იძლევა შემდეგ ძირითად უპირატესობებს ML-ის გამოყენებით:

- ML შეიძლება ეფექტიანი იყოს ქსელის რთული პრობლემების გადასაჭრელად. 6G ქსელები გახდება უფრო რთული, ქსელის მრავალი ტოპოლოგიის, ფიჭური მომხმარებლების რაოდენობის უზარმაზარი ზრდის, მონაცემთა გადაცემის სიჩქარეზე განსაცვიფრებელი მოთხოვნების, რთული რადიოინტერფეისის, ქსელის კოორდინაციის ფართო მეთოდების გამო და ა. შ. 6G ქსელების მნიშვნელოვანი სირთულის გათვალისწინებით, ოპტიმალური მახასიათებლების მქონე გადაწყვეტილებების შემუშავება თითქმის შეუძლებელია ML-ის გარეშე.
- ML-ს შეუძლია მნიშვნელოვანი როლი ითამაშოს მოდელის დეფიციტის პრობლემების მოგვარებაში. ამჟამინდელი ფიჭური ქსელები ემორჩილება მათემატიკურ გამოყვანებს; მაგალითად, ინფორმაციის თეორია იძლევა ჩაკეტილი ფორმის გამოსახულებებს სხვადასხვა ამოცანისთვის, როგორცაა შენონის (Claude Shannon) თეორემა. თუმცა, 6G ქსელების თანდაყოლილი სირთულე ხელს უშლის ჩაკეტილი ფორმის ანალიზური გამოსახულებების გამოყენების შესაძლებლობას, რაც შეიძლება გამოწვეული იყოს, მაგალითად, არხის ან ქსელის მოწყობილობებში არსებული არაწრფივობით. ML გვთავაზობს არაწრფივობასთან გამკლავების ეფექტიან გზას, სასარგებლო ფორმით შესაძლო გადაწყვეტილებების მიწოდებით.
- ML-ს შეუძლია გაუმკლავდეს ალგორითმის დეფიციტის პრობლემებს. თანამედროვე ფიჭურ ქსელებში, ბევრი ოპტიმალური ალგორითმი, მიუხედავად იმისა, რომ კარგად არის აღწერილი, არაპრაქტიკულია განსახორციელებლად. მრავალშესასვლელიანი და მრავალგამოსასვლელიანი (MIMO) სისტემების მაგალითის გათვალისწინებით, სადაც ცნობილია ოპტიმალური გადაწყვეტილებები (მაგალითად, „დასვრილი ქალაქის“ კოდირება), ისინი შეუმჩნეველია ისეთი ხაზოვანი გადაწყვეტილებების სასარგებლოდ, როგორცაა, მაგალითად, წრფივი მინიმალური საშუალო კვადრატული შეცდომა (MSE). იმედი გვაქვს, რომ ML-ს შეუძლია გზა გაუხსნას უფრო ეფექტიანი და, ამავე დროს, პრაქტიკული გადაწყვეტილებების დანერგვას.

ML გამოიყენებოდა სხვადასხვა PHY პრობლემის შესასწავლად და მიუხედავად იმისა, რომ ეს კვლევები არ არის ამომწურავი, ზოგიერთი საინტერესო სფერო მოიცავს შემდეგს: CNN გამოყენებული იყო მოდულაციის კლასიფიკაციისათვის და შემუშავებული იყო RNN-ზე დაფუძნებული მასიური MIMO-ს არხის პროგნოზირების სისტემა, რომელსაც შეუძლია იმუშაოს არხის მატრიცის როგორც შესაფასებელ, ასევე შეკუმშულ (დაქვანტულ) ვერსიებთან.

1.3. უსადენო ქსელები: კერძო შემთხვევების შესწავლა

შემდეგ პარაგრაფებში წარმოვადგენთ გამოყენების სამ შემთხვევას ინდუსტრიულ უსადენო ქსელებში ML ტექნიკის გამოყენების დემონსტრირებისთვის. ამ გამოყენების შემთხვევებისთვის მისადაგებული ML ინსტრუმენტები გამოსახულია ნახ. 1.2-ზე.

მომხმარებლის მოწყობილობის პოზიციონირება:

მომხმარებლის მოწყობილობის (UE) მაღალი სიზუსტით პოზიციონირება მე-3 თაობის პარტნიორობის პროექტის (3GPP) კვლევების ერთ-ერთი მთავარი მიმართულებაა, მისი 15-ე გამოშვების (Release 15) შემდგომ პერიოდში. დღევანდელ ფიჭურ ქსელებში UE პოზიციონირების დასადგენად გამოიყენება სხვადასხვა ჩამოსვლის კუთხეზე (angle-of-arrival) და ჩამოსვლის დროზე (time-of-arrival) დაფუძნებული მეთოდები. ყველა ეს მეთოდი მოითხოვს ტრიანგულაციის ტექნიკას UE-ის ადგილმდებარეობის დასადგენად და მათი გამოყენებისას ადგილი აქვს დროითი სინქრონიზაციის შეცდომებს.

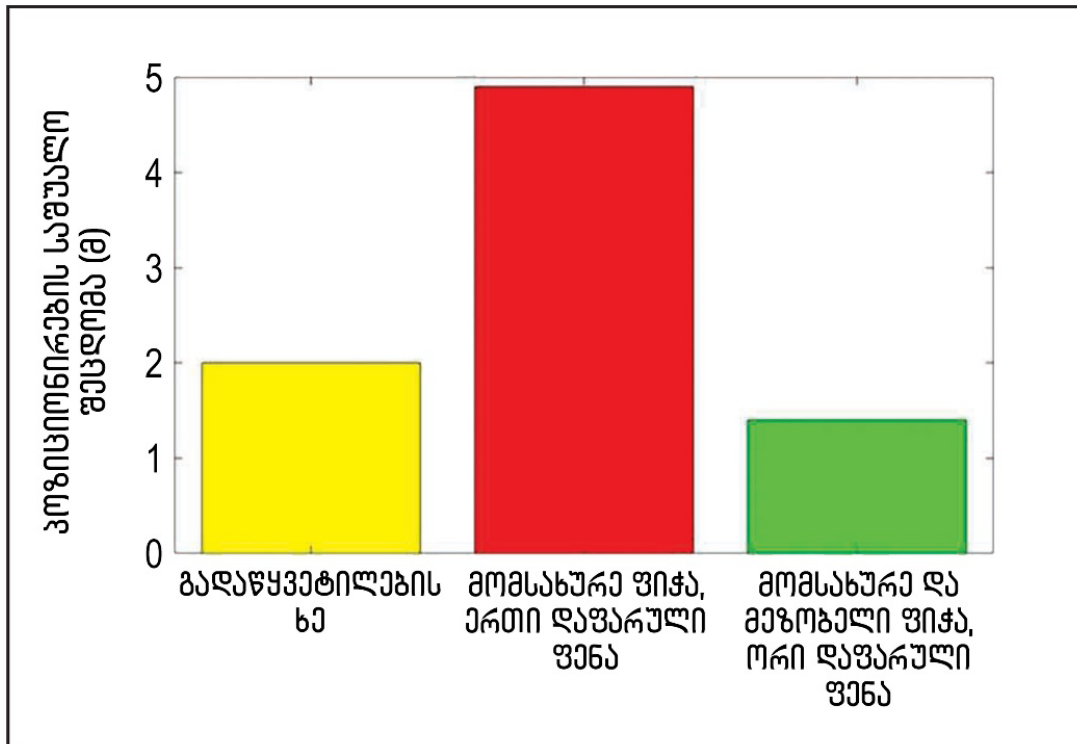
შესწავლილი იქნა UE-ის პოზიცია რადიოსიხშირული (RF) თითის ანაბეჭდის და ML-ის ორი მეთოდის გამოყენებით, ღრმა სწავლებისა (DL) და გადაწყვეტილების ხის მეშვეობით, შენობებს გარეთა სცენარისთვის. მომსახურე ფიჭის საცნობარო სიგნალის მიღებული სიმძლავრე (RSRP) და მეზობელი ფიჭის RSRP მნიშვნელობები გამოყენებული იყო როგორც მახასიათებელი სიდიდეები ღრმა ნეირონული ქსელის (DNN) ტრენინგისათვის. როგორც ნახ. 1.3-ზეა ნაჩვენები, DNN-ისთვის მიღწევა თითქმის 5 მსიზუსტე, როდესაც მხოლოდ ოთხი მომსახურე ფიჭის RSRP მნიშვნელობა და შესაბამისი სხივების იდენტიფიკატორები განიხილება, როგორც შემავალი მახასიათებელი სიდიდეები. შედეგი უმჯობესდება თითქმის 1 მ-მდე, როდესაც კიდევ ორი RSRP მნიშვნელობა უძლიერესი მეზობელი ფიჭიდან და მათი შესაბამისი ფიჭების და სხივების იდენტიფიკატორები ემატება შემავალი ფუნქციების კომპლექტს. გადაწყვეტილების ხის შემთხვევაში, DNN-ზე ნაკლებად რთული ალგორითმი, უბრუნველყოფს დაახლოებით 2 მ სიზუსტეს, როდესაც გამოყენებულია როგორც მომსახურე, ასევე მეზობელი ფიჭის სხივების მონაცემები, როგორც შემავალი მახასიათებელი სიდიდეები. DNN-დან მიღებული თითქმის 1 მ საშუალო სიზუსტე შედარებადია ტრადიციული მეთოდებით მიღწეული სიზუსტის დონესთან ტრიანგულაციის საჭიროების გარეშე და ადგილი არ აქვს სიგნალის დროითი სინქრონიზაციის პრობლემებს.

ML-ით მხარდაჭერილი პროაქტიული მობილურობა:

შეუფერხებელი და ეფექტიანი მობილურობისთვის, კარგად ოპტიმიზებულმა ქსელმა უნდა შეამციროს გადაცემის პროცესში ჰენდოვერების (HO) რაოდენობა და თავიდან აიცილოს როგორც HO, ასევე რადიოკავშირის გაუმართაობა. ახალი მიდგომა იყენებს ML-ზე დაფუძნებულ ალგორითმებს, რომლებიც უბრუნველყოფენ პროაქტიულ და UE-ისთვის სპეციფიკურ მობილურობას gNB -ში. შედარებით მარტივი მიდგომა რადიოგაზომვების ML-ზე დაფუძნებული შემფასებლის შემუშავება, როგორცაა მომსახურე და მეზობელი ფიჭების RSRP, გარკვეული მინიმალური სიზუსტით და გარკვეულ დროით ინტერვალში. რადიოგაზომვები ტრადიციულად სრულდება UE-ის მხარეს და შედეგების შესახებ ეცნობება მომსახურე gNB -ს (ან gNB -ის ცენტრალიზებულ ბლოკს), რადიორესურსების მართვის კონკრეტული კონფიგურაციების მიხედვით. ML-ზე დაფუძნებული პროგნოზირების მიზნებისთვის, RSRP-ის დროითი კვალისთვის ან საცნობარო სიგნალის მიღებული ხარისხისთვის, მნიშვნელობები უნდა შეგროვდეს UE-ში და/ან მომსახურე gNB -ში.

მაგალითად, RSRP მნიშვნელობების შეგროვებული დროითი მწკრივები გამოიყენება ML-ზე დაფუძნებული პროგნოზირების სისტემის შემავალ მონაცემებად, რომელიც UE-ზე და/ან მომსახურე gNB -ზე უბრუნველყოფს საკმარისად ზუსტად შეფასებული RSRP მნიშვნელობების ერთობლიობას მოცემულ შემდგომ დროით ინტერვალში. ეს სიგნალის შეფასებები შემდეგ გამოიყენება HO-ს შესაძლო პირობების პროგნოზირებადი შეფასებისთვის, რამაც შეიძლება მოგვაწოდოს პროაქტიული გაზომვის ანგარიშები UE-დან და/ან პროაქტიული HO მოქმედებებიდან მომსახურე gNB -ზე. ეს ორი საფეხური მეორდება მოცემული დროითი პერიოდულით, მაგალითად, დისკრეტობის სიხშირით, სიჩქარით და RSRP-ის გაზომვების შემავალი მნიშვნელობების დროითი გაფილტვრით. ალტერნატიულად, ეს ნაბიჯები შეიძლება ასევე გააქტიურდეს მომსახურე gNB -ის მიერ, როდესაც დაკმაყოფილებულია ტრაფიკის ან მობილურობის მომსახურების ხარისხის (QoS) გარკვეული პირობები.

აღწერილი ML-ზე დაფუძნებული მობილურობის ალგორითმი შეიძლება განხორციელდეს UE-ში, gNB -ში ან ორივეში, თითოეულ კვანძში არსებული ML-მხარდაჭერის შესაძლებლობების მიხედვით. გარდა ამისა, ეს



ნახ. 1.3. UE-ის პოზიციის შედარება როგორც DNN, ასევე გადაწყვეტილების ხის ტექნიკისთვის. სისტემის დონის ქსელის პარამეტრები მოიცავს რვა ადგილს, 110 მ დამორებით და 28 გპც გადამტანი სიხშირით.

მექანიზმი შეიძლება ინტეგრირებული იყოს თვითორგანიზებულ ქსელურ გადაწყვეტილებებში, რათა მოხდეს მობილური კომუნიკაციების საიმედოობის ოპტიმიზაცია.

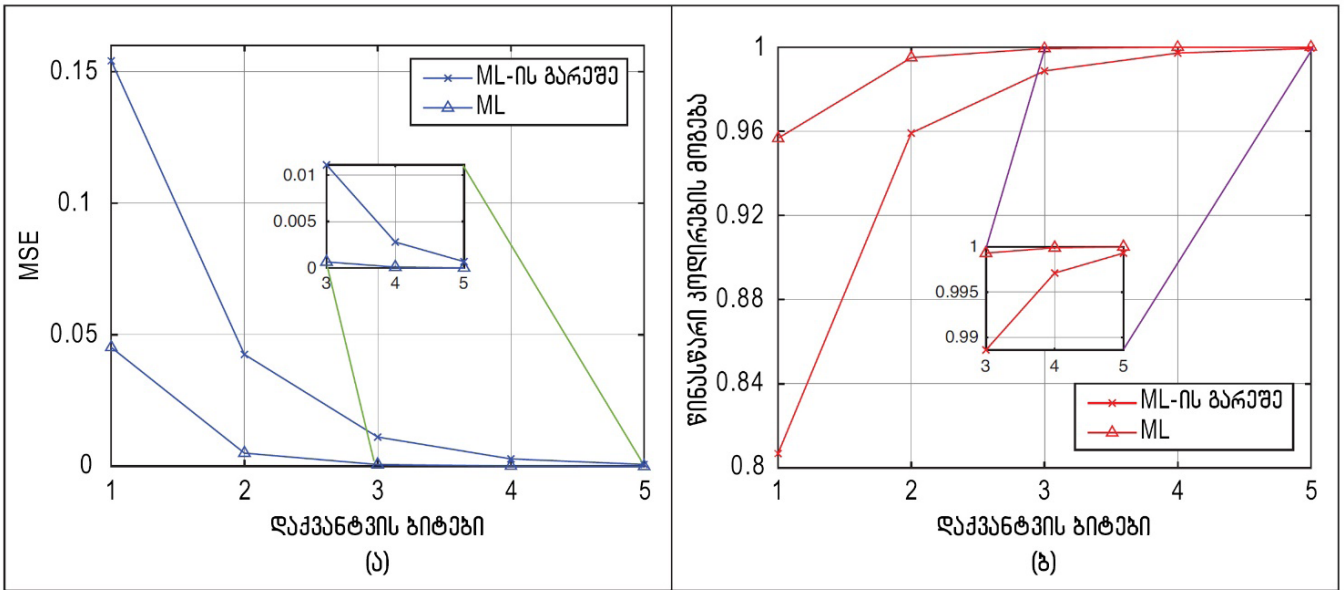
არხის მდგომარეობის შესახებ ინფორმაციული უკუკავშირი:

არხის მდგომარეობის შესახებ ინფორმაციული (CSI) უკუკავშირი დაუნლინკის არხზე მთავარი გამოწვევაა 3GPP-ის მე-17 გამოშვებისთვის და შემდგომ პერიოდში. ამჟამად, CSI-ის სიზუსტეზე გავლენას ახდენს სტანდარტის მიერ დაწესებული გაზომვების შეკუმშვა.

განვიხილოთ ორი RNN-ზე დაფუძნებული ტყუპი არხის პროგნოზირების სისტემა gNB-ზე და UE-ზე. წარსული CSI გამოიყენება საკომუნიკაციო სისტემის ორივე ბოლოში RNN-ის ტრენინგისთვის. UE-ის უკუკავშირი ფასდება პროგნოზირებულ არხთან მიმართებაში. ნახ. 1.4 ასახავს MSE-ის, ფაქტობრივ და მოპოვებულ არხს შორის gNB-ზე, და წინასწარი კოდირების მოგებას, როდესაც დაქვანტის ბიტების სხვადასხვა რაოდენობა გამოიყენება UE-დან CSI-ის მისაწოდებლად. შედეგები შედარებულია ML-ით და მის გარეშე CSI-ის უკუკავშირისთვის. როგორც ნახ. 1.4-დან ჩანს აშკარა სარგებელი აქვს ML-ის გამოყენებას. ჩვენ გვჯერა, რომ ML-ზე დაფუძნებული გადაწყვეტილებები გააუმჯობესებს არსებულ მახასიათებლებს სიგნალის ზედნადების გაზრდის გარეშე.

1.4. ML-ის როლი სტანდარტიზაციაში

ML-ის პოტენციური 5G-ისთვის ფართოდ იქნა აღიარებული ლიტერატურაში და აპლიკაციები სტანდარტადაც კი იქცა მაღალ დონეზე, როგორცაა ქსელი და უსაფრთხოება. 3GPP-მ მე-15 და მე-16 გამოშვებებში შემოიტანა სპეციფიკაცია, სახელად ქსელის მონაცემთა ანალიზური ფუნქცია (NWDAF), როგორც 5G ძირითადი არქიტექტურის ნაწილი. NWDAF პასუხისმგებელია ქსელის ანალიზის მიწოდებაზე NF-ის მოთხოვნის შემთხვევაში. მონაცემები გროვდება აპლიკაციის ფუნქციის (AF) მეშვეობით და ხდება მათი ექსპლუატაცია, ადმინისტრირება, ტექნიკური მომსახურება და განთავსება მონაცემთა საცავში. სპეციფიკაციები ასევე ეხება ავტომატიზაციისა და მონაცემთა შეგროვების მიზნით ურთიერთთანამშრომლობის საკითხებს.



ნახ. 1.4. MSE-ის მახასიათებლები (ა) და წინასწარი კოდირების მოგება (ბ). განიხილება 2×1 MIMO კონფიგურაცია და RNN შედგება ერთი ფარული ფენისგან.

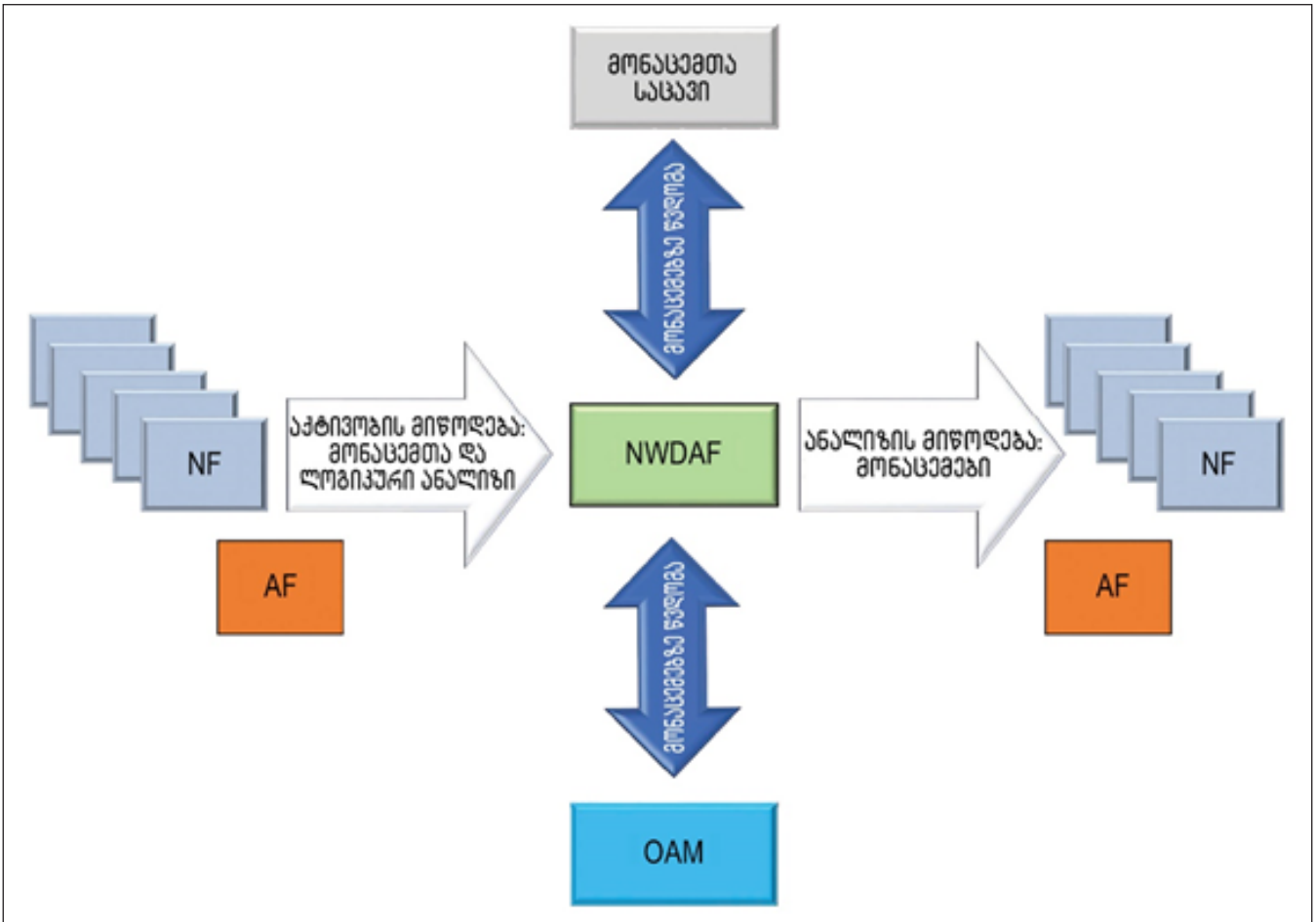
3GPP NWDAF სტრუქტურა 5G სისტემებისთვის გამოსახულია ნახ. 1.5-ზე. ეს ავტომატიზაცია აძლევს ქსელის პროვაიდერებს შესაძლებლობას განათავსოს და შეამოწმოს ინი-RT (არარეალური დროის) სიტუაციაში ML-თან დაკავშირებული გამოყენების შემთხვევები. ნახ. 1.5-ზე, შიდა ინტერფეისები აგროვებს მონაცემებს სხვადასხვა ქსელის წყაროდან, სადაც კომუნიკაცია ხდება არსებული სერვისზე დაფუძნებული ინტერფეისების გამოყენებით. გარე ინტერფეისები უზრუნველყოფს ანალიზზე დაფუძნებულ და ალგორითმულ გადაწყვეტილებებს AF-ისა და NF-ისთვის.

რაც შეეხება PHY ფენას, აქ ML ტექნიკა ჩამორჩება რიგი საკითხების გამო. პირველ რიგში, PHY იყენებს აბსტრაქციებს და მათემატიკურ მოდელს, რომლებიც მიღებულია ფიზიკური რეალობისა და ელექტრომაგნიტური პრინციპებიდან. სანამ ასეთი მოდელები ზუსტად აღწერს რეალურ სამყაროს, არ არის საჭირო ML. თუმცა, პრაქტიკაში, მოდელები და ფიქსირებული ალგორითმები არაეფექტიანია სწრაფად ცვლადი და ჰეტეროგენული გარემოს პირობებში. მაგალითად, იმავე არხის პარამეტრების მოპოვების სქემის გამოყენება CSI-ის მისაღებად ლებტოპიდან, რომელიც იმყოფება gNB-ისთან პირდაპირი ხედვის ხაზზე (LoS), პლანშეტიდან სწრაფ მატარებელში ან მობილური ტელეფონიდან, რომელიც სწრაფად მოძრაობს ძალიან მჭიდროდ დასახლებულ ზონაში, შეიძლება არ იყოს ოპტიმალური. შესაბამისად, ინტელექტუალური ტექნიკის სტანდარტიზაციის მცდელობებმა უკვე მოიპოვა იმპულსი და მიუხედავად იმისა, რომ 3GPP მზად არის დაიწყო კვლევები ML-ის დანერგვაზე, ღია RAN (O-RAN) იქნება ML-ისთვის მშობლიური, რომელიც განსაზღვრავს RAN-ის ინტელექტუალურ კონტროლერს (RIC) და რომელიც გააუმჯობესებს RAN-ის ზოგიერთ ფუნქციას.

3GPP-მა დაიწყო ML-ის გამოყენების შედეგების შესწავლა პირველ ფენაზე და შეთანხმებული იქნა ML-ის საკვლევო თემები NR რადიოინტერფეისისთვის. RAN-1 სამუშაო ჯგუფის კვლევების შემდეგ, პროტოკოლის ასპექტები შეისწავლება RAN-2-ში და შემდეგ თავსებადობისა და ტესტირების ასპექტები განიხილება RAN-4 სამუშაო ჯგუფში. ბოლო ეტაპზე, სექციები შეაჯამებენ PHY-ისთვის ML მეთოდების სტანდარტიზაციის სტატუსს, როგორც 3GPP-ის, ასევე O-RAN-ის კუთხით.

CSI უკუკავშირი:

CSI უკუკავშირი დაუნლინკის არხისთვის 3GPP მე-17 გამოშვებაში არის რთული პრობლემა, რომელშიც UE-ზე დაფუძნებული სხივის შერჩევა მოჰყვება CSI-ის საცნობარო სიმბოლოების (RS) ტრენინგს და წინასწარი კოდირების მატრიცის ინდექსის (PMI) ანგარიშს, და ბოლოს, დემოდულაციის საცნობარო სიგნალს (DMRS) და, წინასწარ კოდირებული არხის შემდგომ შეფასებას. ზოგადად, სხივის შერჩევა მიზნად ისახავს ლინკის საკმარისად ძლიერი ბიუჯეტის შექმნას UE-სა და gNB-ს შორის. CSI-RS გამოიყენება არხის ზუსტი შეფასებისთვის, რომელიც



ნახ. 1.5. 5G ქსელის ავტომატიზაციის განზოგადებული სტრუქტურა 3GPP მე-16 გამოშვებაში, რომელიც აჩვენებს, რომ NWDAF-ს უნდა შეეძლოს შეაგროვოს ოპერატორის ოპერატიული, ადმინისტრაციული და ტექნიკური მონაცემები

შემდეგ მიეწოდება gNB-ს წინასწარი კოდერის შესამუშავებლად; საბოლოოდ, DMRS არის წინასწარ კოდირებული საპილოტე სიგნალები, რომლებსაც UE-ები იყენებენ კოპერენტული დემოდულაციის განსახორციელებლად. ამჟამად, თითოეული ეს ფაზა იქმნება წინასწარ დადგენილი წესების შესაბამისად, რაც მცირე ადგილს ტოვებს ინტელექტუალური ქცევისთვის. იდეა არის ის, რომ ML-ს შეეძლოს ყოველი ფაზის გაუმჯობესება სხვადასხვა გზით. სხივების შერჩევა შეიძლება გაუმჯობესდეს სხივების ინტელექტუალური კორელაციით UE-ების პოზიციასთან ან იდენტურობასთან. ეს საშუალებას მისცემს gNB-ს ჭკვიანურად შეარჩიოს სხივები, რითაც თავიდან აიცილოს უხეში შერჩევა. CSI-RS შეიძლება გაუმჯობესდეს საპილოტე სიგნალების და PMI უკუკავშირის შეკუმშვით, სპეციალურად გამოყოფილი ML კომპრესორების დახმარებით. გარდა ამისა, არხების პროგნოზირების ტექნიკა შეიძლება გამოყენებულ იქნეს CSI უკუკავშირის საბაზისო ხაზის წინასწარ დასაყენებლად. სხვა ასპექტები, რომლებიც შეიძლება გაუმჯობესდეს, მოიცავს საპილოტე სიგნალების სიხშირეს როგორც CSI-RS-ში, ასევე DMRS-ში, სიმძლავრეს და სინქრონიზაციას, და CSI-RS პორტის შერჩევას.

RS-DMRS:

უხეშად რომ ვთქვათ, DMRS არის RS, რომელიც გამოიყენება არხის შეფასებისთვის კოპერენტული დემოდულაციის განსახორციელებლად. არხის სწორი შეფასება ასეთი პილოტ-სიგნალების გამოყენებით დიდ გავლენას ახდენს მახასიათებლებზე ბიტზე შეცდომის ალბათობის და, შესაბამისად, ბლოკზე შეცდომის ალბათობის თვალსაზრისით. ML-ის როლი ასეთ დომენში ორმხრივია. პირველი, ის შეიძლება გამოყენებულ იქნეს არხის შეფასების მუშაობის გასაუმჯობესებლად. მეორეც, ML-ს შეუძლია უზრუნველყოს DMRS-ების უფრო გონივრული პოზიციონირება მათი სიხშირის შესამცირებლად, და, შესაბამისად, შეამციროს ზედნადები ხარჯები 6G-ში.

პოზიციონირება:

ზუსტი პოზიციონირება არის ერთ-ერთი ასპექტი, რომელიც აჩვენებს უდიდეს გაუმჯობესებას LTE-ის ისეთი პარამეტრებისთვის, როგორებიცაა დაკვირვების შედეგად განსაზღვრული დროითი სხვაობა ჩამოსვლისას და აპლიკაციის ჩამოსვლისას დროითი სხვაობა, რომლებიც განსაზღვრულია 3GPP მე-9 გამოშვებაში. 5G-ის სხვადასხვა ასპექტი იძლევა UE-ის ზუსტი პოზიციონირების საშუალებას, როგორცაა ანტენის ელემენტების დიდი რაოდენობა 5G-ზე, გადაცემა მილიმეტრული ტალღების დიაპაზონში და ქსელის მკვირივი განლაგება. თუმცა, მეთოდები, რომლებიც დაფუძნებულია როგორც ჩამოსვლის კუთხეზე, ასევე ჩამოსვლის დროზე, არაეფექტიანია, როდესაც განიხილება არაპირდაპირი ხედვის სცენარები, ინტერფერენციით შეზღუდულ სცენარებში. მოსალოდნელია, რომ ML ტექნიკა (იხ. ნახ. 1.2) დაგვეხმარება პოზიციის გაუმჯობესებაში არხის დიაგრამების გამოყენებით, რითაც შეისწავლის UE-ის სავარაუდო პოზიციას ანგარიშის საფუძველზე (როგორცაა CSI ანგარიში და ზონდირების საცნობარო სიგნალის რუქები) და მოახდენს ინფორმაციის ერთად მულტიპლექსირებას, რომელიც შეიცავს პოზიციონირების ინფორმაციას, მაგრამ ძნელად გამოსაყენებელია კლასიკური გზით.

მობილურობის გაუმჯობესება:

5G-ში ფიჭების ხშირმა არჩევამ და RSRP-ის ხშირმა გამოვლამ შეიძლება გავლენა მოახდინოს UE-ების ბატარეის მუშაობის ხანგრძლივობაზე. გარდა ამისა, დატვირთვის დაბალანსების ალგორითმებს შეუძლიათ მოიხმარონ ინტელექტუალური ტექნიკა, რომელიც გამოიყენებს UE-ის სპეციფიკურ არხს, მოძრაობა-ტრაექტორიას და ტრაფიკის მოთხოვნის პროგნოზებს. გარდა ამისა, ისეთი სცენარები, როგორცაა სწრაფი მატარებლები ან არატერესტრიალური ქსელები, გამოწვევას შეუქმნის HO და პირობით HO ოპერაციებს. გათვალისწინებული ახალი სცენარი, ამჟამინდელ 3GPP მე-17 გამოშვებასთან შედარებით, მოიცავს UE-ის სპეციფიკურ, ML-ზე დაფუძნებული პროგნოზირების ალგორითმების გამოყენებას, რომლებიც შექმნილია ფიჭების ხელახალი შერჩევისა და HO-ის წარუმატებლობის გასაუმჯობესებლად, რითაც გაუმჯობესდება საერთო QoS.

ML მონაცემთა შეგროვების სტანდარტიზაცია:

3GPP-მ დაიწყო მუშაობა მონაცემთა შეგროვებაზე ML ალგორითმების გასაშვებად 5G ქსელებში. ასეთი კვლევის სფერო მოიცავს მექანიზმების იდენტიფიცირებას, რომლებიც აგროვებენ მონაცემებს ქსელიდან დრაივ-ტესტის სტრუქტურის მინიმიზაციის ან შემდგომი გაუმჯობესების გზით. გარდა ამისა, კვლევა ფოკუსირებული იქნება მანქანური სწავლების მოდელების განთავსებაზე, როგორც ტრენინგის, ასევე დასკვნის მიზნებისთვის ქსელის სხვადასხვა ერთეულზე სხვადასხვა გამოყენების შემთხვევისთვის, ასევე ნებისმიერი ახალი ინტერფეისის იდენტიფიცირებაზე, რომელიც საჭიროა მონაცემების მოდელებზე გადასატანად.

ფედერირებული სწავლების მოდელების კრებული:

ML მოდელებზე დაფუძნებული ტრენინგი და პროგნოზები შექმნის დამატებით დატვირთვას ქსელებზე, რომლებიც უკვე გადასცემენ დიდი მოცულობის მონაცემებს. აქედან გამომდინარე, მნიშვნელოვანია მოდელების ტრენინგისა და დასკვნის გავლენის შეფასება ქსელის ტრაფიკზე, განსაკუთრებით ფედერირებული სწავლებისთვის (FL), სადაც UE-ები იმოქმედებენ როგორც განაწილებული ჰოსტები. ადგილობრივად დატრენინგებული მოდელების შეგროვების შეყოვნება შეზღუდულია FL-ით და ქსელის ლინკები უნდა აკმაყოფილებდეს შეყოვნების ბიუჯეტს. ეს განსაკუთრებით რთულია თანამედროვე ქსელებში, სადაც UE-ის საკუთარი QoS მოთხოვნები უკვე მაღალია და FL მოდელების სწავლება და მონაცემთა შეგროვება დამატებით დატვირთვას შექმნის ქსელში. ანალოგიურად, გაყოფილი დასკვნა, სადაც UE-ები ერთმანეთთან ურთიერთობენ საერთო დასკვნის შესასრულებლად, იწვევს ქსელის ტრაფიკის გაზრდას. კვლევები 3GPP მე-18 გამოშვებაში ფოკუსირებული იქნება ზემოაღნიშნულ საკითხებზე, რათა მხარი დაუჭიროს ML/FL მოდელების ტრენინგს და დასკვნების შემუშავებას უსადენო ლინკებში.

O-RAN RIC:

O-RAN ალიანსი მიზნად ისახავს შექმნას RAN, რომელიც არ არის გაყიდვების კუთხით სპეციფიკური და რომელსაც აქვს ჩაშენებული მხარდაჭერა ML-ისთვის, როგორც ავტომატიზაციისა და საოპერაციო დანახარჯების დაზოგვის საშუალება. O-RAN ალიანსს აქვს განსაზღვრული ინტერფეისები ინფორმაციის გაცვლისთვის პროტოკოლის სტეკში. ამ მიზნით, ქსელის ავტომატიზაციის O-RAN არქიტექტურა მოიცავს ML-ზე დაფუძნებულ RIC-ს ორივე სცენარისთვის, ანუ non-RT და RT შემთხვევებისთვის. non-RT RIC-ში, ML ალგორითმის ტრენინგი ხორციელდება ქვედა ან უფრო მაღალი ფენებიდან შეგროვებული მონაცემების გამოყენებით. თუმცა, სწავლის პროცესი ნელი რჩება; ამიტომ, მას უწოდებენ non-RT RIC-ს. მოგვიანებით, მოსწავლე გადაყვანილია RT RIC-

ში, რომელიც იყენებს RT-ის მიერ მოპოვებულ მონაცემებს გადაწყვეტილების მისაღებად ონლაინ რეჟიმში. გარდა ამისა, non-RT-ის ფუნქციონირება მოიცავს პოლიტიკის მართვას და უფრო მაღალი დონის პროცედურების ოპტიმიზაციას. ამიტომ, RAN-ს ან ძირითად ქსელს შეუძლია განათავსოს ასეთი მექანიზმი შეგროვებულ მონაცემებზე დაყრდნობით.

1.5. ღია გამოწვევები და საგზაო რუკა ML ტექნიკის დანერგვისთვის

მონაცემთა ხელმისაწვდომობა და ბენჩმარკინგი:

მიუხედავად იმისა, რომ ML არის პოტენციური ტექნოლოგია და გამაუმჯობესებელი საშუალება შემდეგი თაობის უსადენო ქსელებისთვის, არსებობს მის პრაქტიკულ გამოყენებასთან დაკავშირებული რამდენიმე გამოწვევა, რომელიც განხილულია ამ პარაგრაფში. უსადენო ქსელების ერთ-ერთი მთავარი გამოწვევა არის მონაცემთა ხელმისაწვდომობა. მონაცემთა ხელმისაწვდომობა ეხება საერთო და მიღებულ მონაცემთა ნაკრების იდენტიფიცირების პრობლემას (მაგალითად, არხის რეალიზაციას) ML ალგორითმების ტესტირებისა და ბენჩმარკინგის მიზნით (განვმარტავთ, რომ გამოთვლით ტექნიკაში, ბენჩმარკი ანუ საორიენტაციო ნიშანი არის კომპიუტერული პროგრამების ან სხვა ოპერაციების გაშვება ობიექტის შედარებითი მახასიათებლის შესაფასებლად, ჩვეულებრივ, სტანდარტიზებული ტესტებისა და ცდების სერიის მეშვეობით). ასეთ პრობლემას გადამწყვეტი მნიშვნელობა აქვს სტანდარტიზაციისთვის, სადაც ჩვეულებრივ, ალგორითმების და წინადადებების ტესტირება ხდება დადასტურებული საბაზისო ფიზიკური მოდელების (მაგალითად, ურბანული მაკროფიჭების/ მიკროფიჭების არხის მოდელების), შეფასების მეთოდოლოგიებისა და დაკალიბრებული სიმულატორების გამოყენებით. სხვა სფეროებისგან განსხვავებით, ფიჭურ ქსელებს არ გააჩნიათ სტანდარტული მონაცემთა ნაკრები ML ალგორითმის ტრენინგისა და ბენჩმარკინგის განხორციელებისთვის. აქედან გამომდინარე, სინთეზური ან პროგრამული უზრუნველყოფის მიერ გენერირებული მონაცემთა ნაკრები არის უპირატესი მნიშვნელობის ML ალგორითმ(ებ)ის ტრენინგისა და ბენჩმარკინგისთვის, ასევე წინადადებების და სტანდარტული ალგორითმების შეფასების საერთო მეთოდოლოგიაზე შეთანხმებისთვის.

უსადენო ქსელებში მახასიათებლების ძირითადი ინდიკატორების განსაზღვრა კიდევ ერთი გადამწყვეტი ამოცანაა ML სტანდარტიზაციისთვის. აუცილებელია მეტრიკის ნაკრების შემუშავება ML ალგორითმებისა და მათი მახასიათებლების კლასიფიკაციისა და რანგირებისთვის. კლასიკური მიდგომები, როგორცაა გამტარუნარიანობა და სიგნალის ფარდობა ინტერფერენცია-პლუს-ხმაურთან, შეიძლება არ იყოს საკმარისი, რადგან ამ მნიშვნელობების მცირე გაუმჯობესება შეიძლება მოხდეს სირთულის მნიშვნელოვანი ზრდისა და ენერჯის მოხმარების გაზრდის ფასად.

ML-ის და ML-ისგან განსხვავებული გადაწყვეტილებების შერჩევა:

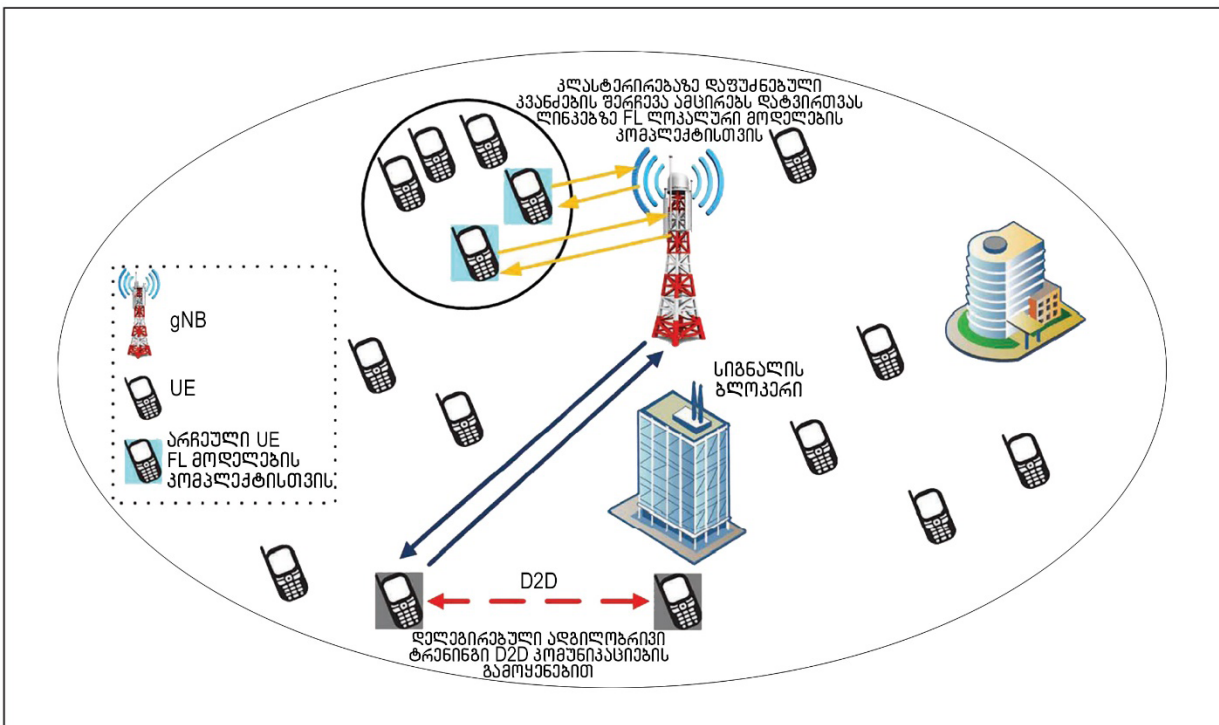
ML-ის ინსტრუმენტები განიხილება, როგორც იმპლემენტაციაზე ორიენტირებული ინსტრუმენტები და არა სტანდარტებთან დაკავშირებული ასპექტი. ამის მიღმა იდეა ემყარება იმ ფაქტს, რომ თითოეულ გამყიდველს (ვენდორს) აქვს თავისუფლება ეფექტიანად განახორციელოს სტანდარტის თითოეული ასპექტი, მანამ, სანამ გარე ინტერფეისები დაცულია. ამის მარტივი მაგალითი მოცემულია CSI უკუკავშირში, სადაც UE-მ უნდა აირჩიოს კონკრეტული PMI, მაგრამ სტანდარტი არ იძლევა რაიმე კონკრეტულ გზას ამ შერჩევის გასაკეთებლად. თუმცა, ცოტა ხნის წინ, იდეამ ML-ზე დაფუძნებული შეტყობინებების გაცვლის და მახასიათებლების შესახებ, რომლის მიღწევაც მხოლოდ ML-ით შემუშავებულ ალგორითმს შეუძლია, გზა გაუხსნა ML ალგორითმების სტანდარტიზაციას. ეს იძლევა რამდენიმე პრობლემის მიმართულებით მუშაობის დაწყების შესაძლებლობას; მაგალითად, დააწესებს თუ არა სტანდარტი სპეციფიკურ ML ჩარჩოს, რომელიც მოახდენს მინიმალური მახასიათებლისა და განხორციელების სტრუქტურის კლასიფიცირებას, თუ ის შორს დარჩება განხორციელებისგან? რაც შეეხება ნეირონულ ქსელებს (NN), ჯერ კიდევ ღია საკითხია, ჰიპერპარამეტრები დარჩება თუ არა კონკრეტული გამყიდველის შეხედულებისამებრ, თუ ისინი განსაზღვრული იქნება სტანდარტით.

ML ალგორითმების სირთულე:

ბატარეის შეზღუდული მუშაობის, მეხსიერების, გამოთვლითი შესაძლებლობების და კომუნიკაციის შეზღუდული გამტარუნარიანობის გათვალისწინებით ფიჭური ქსელის უმეტეს ობიექტებში, ML მოდელის ღირებულებასა და ეფექტიანობას შორის ურთიერთგაცვლა ფუნდამენტური საკითხია. კიდევ ერთი საკითხია სიჩქარის/დროის საფეხურები, რომლითაც უნდა შესრულდეს ტრენინგი და მიღებულ იქნეს დასკვნა. მაშინ როდესაც მყარ-სადენიან gNB-ებს აქვთ საკმარისი გამოთვლითი სიმძლავრე რთული ML ალგორითმების გასაშვებად, UE-ის უწევს გამკლავება ბატარეის, გათბობის და მკაცრი სირთულის შეზღუდვებთან. ასეთი საკითხების შესაძლო გადაწყვეტილებები მოიცავს გონივრულად და სათანადოდ გაყოფილი ML ალგორითმების განხორციელებას UE-ის მხარესა და gNB-ის მხარეს შორის.

კომუნიკაციით გაცნობიერებული FL:

ტრადიციული ML მოდელები მხარს უჭერენ ცენტრალიზებულ სწავლებას. UE-ისგან დიდი რაოდენობით ტრენინგის მონაცემების შეგროვების ისეთი სირთულეების გამო, როგორებიცაა, მაგალითად, კონფიდენციალურობის საკითხები და გამტარუნარიანობის შეფერხება, FL გამოჩნდა, როგორც პერსპექტიული გადაწყვეტა. FL-ში ტრენინგი ტარდება განაწილებულად, ქსელის მოწყობილობებზე, რომლებსაც უწოდებენ ლოკალური მოდელის ჰოსტებს, ხოლო აპლიკაციის სერვერი ქსელის მხარეს მოქმედებს, როგორც ცენტრალური ჰოსტი ადგილობრივი მოსწავლეების მიერ გადაცემული ლოკალური მოდელების აგრეგაციისთვის. როგორც წესი, აპლიკაციის სერვერის ჰოსტი აერთიანებს მოდელებს მხოლოდ მაშინ, როდესაც განახლებები ხელმისაწვდომია ყველა ადგილობრივი მოსწავლისგან, რასაც სინქრონული მოდელის გადაცემა ეწოდება. თუმცა, ეს ძალზე არაეფექტიანია უსადენო ქსელებში, სადაც საკომუნიკაციო ლინკები არაპროგნოზირებადია, ადგილობრივი მოსწავლეები (UE-ები) ენერგიით შეზღუდულია და აქვთ საკუთარი QoS მოთხოვნები. ასინქრონული მოდელების კომპლექტის გამოყენება არის ყველაზე ეფექტიანი გადაწყვეტა FL-ისთვის უსადენო ქსელებში, სადაც UE-ების ქვევგუფი შეირჩევა ლოკალური მოდელის განახლებისთვის მოდელების შეგროვების თითოეულ რაუნდში. თუმცა, UE-ის შერჩევა თითოეულ რაუნდში რთული პრობლემაა, რადგან UE-ები შეზღუდულია ენერგიით და ქსელის გამტარუნარიანობა მწირია, რაც ხელს უშლის ლოკალური მოდელების შეგროვებას ყველა UE-დან დამოუკიდებლად და თანაბრად შეგროვებული მონაცემების წარმოსაჩენად. ეს მექანიზმები, როგორც წესი,



ნახ. 1.6. მოდელების კომპლექტი FL-ისთვის უსადენო ქსელში, როდესაც ზოგიერთ UE-ს აქვს დიდი ბლოკირება და იყენებს D2D კომუნიკაციას მოდელის გადასატანად. კლასტრზე დაფუძნებული UE-ის შერჩევა არის კიდევ ერთი გამოსავალი მოდელების ასინქრონული შეგროვებისთვის, რომელიც აკმაყოფილებს ქსელის QoS მოთხოვნებს.

არის გამყიდველის საკუთრება, მაგრამ სტანდარტიზაციას ჯერ კიდევ სჭირდება რამდენიმე საერთო მექანიზმის განსაზღვრა მოდულების ეფექტიანი შეგროვებისთვის. როგორც ნაჩვენებია ნახ. 1.6-ზე, UE კლასტერიზება და ლოკალური მოწყობილობებს შორის კომუნიკაცია (D2D) ასინქრონული მოდულების შეგროვებისთვის არის შესაძლო გადაწყვეტილებები ქსელური კომუნიკაციის შესამცირებლად და საჭიროებს სტანდარტიზაციის მხარდაჭერას.

ML ტექნოლოგიის სტაბილურობა და ადაპტაციის უნარი:

ML ალგორითმები, რომლებიც გამოიყენება უსადენო ქსელებში, უნდა იყოს ადაპტიური, რადგან მათ მოუწევთ გაუმკლავდნენ პარამეტრებს, რომლებიც დინამიკურად იცვლება. კერძოდ, NN-ის წონა შეფასებულია ონლაინ ტრენინგის მონაცემების საფუძველზე. თუმცა, ეს მიდგომა შეიძლება არ იყოს გამოყენებული უსადენო ქსელებში და კონკრეტულად სტანდარტში, სადაც კოორდინაცია უნდა მოხდეს სხვადასხვა ოპერატორის მფლობელობაში მყოფ და სხვადასხვა მომწოდებლის მიერ წარმოდგენილ ერთეულებს შორის და სადაც სწრაფი რეაგირების საჭიროებამ შეიძლება გამოიწვიოს ამა თუ იმ გადაწყვეტილების მიღება. შესაძლო გადაწყვეტილებები მოიცავს: წინასწარ დატრენინგებულ ან ნაწილობრივ დატრენინგებულ NN-ს (ანუ NN-ს, მოცემული საწყისი წერტილით); ღრუბელზე დაფუძნებულ, ჩამოტვირთვად მონაცემთა ნაკრებს NN ტრენინგისთვის; და კოდურ წიგნზე დაფუძნებულ NN-ს, რომელშიც გამოიყენება სხვადასხვა NN-ის კოდების წიგნი და შეთანხმებულია ეNB-სა და UE-ებს შორის. კიდევ ერთი დაკავშირებული პრობლემა არის მოძველებული ML მოდელის გამოვლენა დასკვნის მაღალი შეცდომით და მისი შეცვლა. მოძველებული მოდელის ახალი მოდელით ჩანაცვლება შემდგომ იწვევს შეყვანებას. ამრიგად, უნდა არსებობდეს პროექტიული მექანიზმი, რომელიც ახდენს ML მოდელის ადაპტაციას ქსელის პირობებთან ისე, რომ NF-ები განიცდიან მახასიათებლების მინიმალურ დანაკარგებს.

1.6. პირველი თავის დასკვნა

მოტივირებული ML ალგორითმების გამოყენების შესაძლებლობებით, ჩვენ წარმოვადგინეთ ML ტექნოლოგიის მიმოხილვა, რომელიც გამოიყენება 5G-ის შემდგომი პერიოდის (B5G) და 6G უსადენო ქსელებში. გარდა ამისა, ჩვენ განვიხილეთ ML-ზე დაფუძნებული გადაწყვეტილებების ძირითადი როლი ინდუსტრიის და სტანდარტიზაციის პერსპექტივიდან. ჩვენ ასევე ხაზგასმით აღვნიშნეთ ML ტექნიკის დანერგვის პრაქტიკული გამოწვევები უსადენო ქსელებში და თუ როგორ გაუმკლავდეთ მათ. non-RT და უფრო მაღალი ფენის, ML-ზე დაფუძნებული გადაწყვეტილებები უკვე შეიძლება გამოყენებულ იქნეს დღევანდელ ქსელებში. RT ML გადაწყვეტილებების დანერგვა PHY/გარემოს წვდომის კონტროლისთვის 6G ქსელებში არის შემდეგი დიდი გამოწვევა კვლევითი საზოგადოებისთვის. ჩვენ გვჯერა, რომ ამ გამოწვევების დაძლევა, როგორც კვლევის, ასევე სტანდარტიზაციის დონეზე, გზას გაუხსნის უსადენო კომუნიკაციების შემდგომ თაობას იყოს ეფექტიანი და მდგრადი.

თავი 2 ხელოვნური ინტელექტის და უსადენო კომუნიკაციების ცხრა გამოწვევა 6G ტექნოლოგიისთვის

2.1. შესავალი

გასული საუკუნის 80-იანი წლების ბოლოს, გამოთვლითი ნეირონული ქსელების ალგორითმების ძირითადი ანსამბლის ფორმირება თითქმის დასრულებული იყო, თუმცა დღესდღეობით აღნიშნულ პერიოდთან შედარებით, ნეირონული ქსელების ზომა გაცილებით დიდია. სულ რაღაც 30 წლის წინ, საძიებო კვლევის სტადიაში მყოფი NN, როგორც წესი, შედგებოდა მხოლოდ რამდენიმე ნეირონისგან, რომლებიც კონფიგურირებული იყო მრავალ ფენაში. მას შეეძლო მხოლოდ შეზღუდული რაოდენობის ტრენინგის მონაცემების დამუშავება გამოთვლითი რესურსების შეზღუდვისა და სირთულის გამო, რაც თავის მხრივ ზღუდავდა მის შესაძლებლობებს, მახასიათებლებს და გამოყენებადობას. 2000-იანი წლების დასაწყისისთვის რამდენიმე მეცნიერი, კერძოდ, ბენგიო (Yoshua Bengio), ჰინტონი (Geoffrey Hinton) და ლეკუნი (Yann LeCun), ჯიუტად აგრძელებდა მუშაობას ამ მიმართულებით და მათ შეძლეს ახალი აზროვნების გამომუშავება, რამაც გამოიწვია DL-ის უზარმაზარი წარმატება და საბოლოოდ მიგვიყვანა მათ მიერ ძალზე პრესტიჟული ტიურინგის (Alan Turing) სახელობის პრიზის მოგებაზე 2018 წელს.

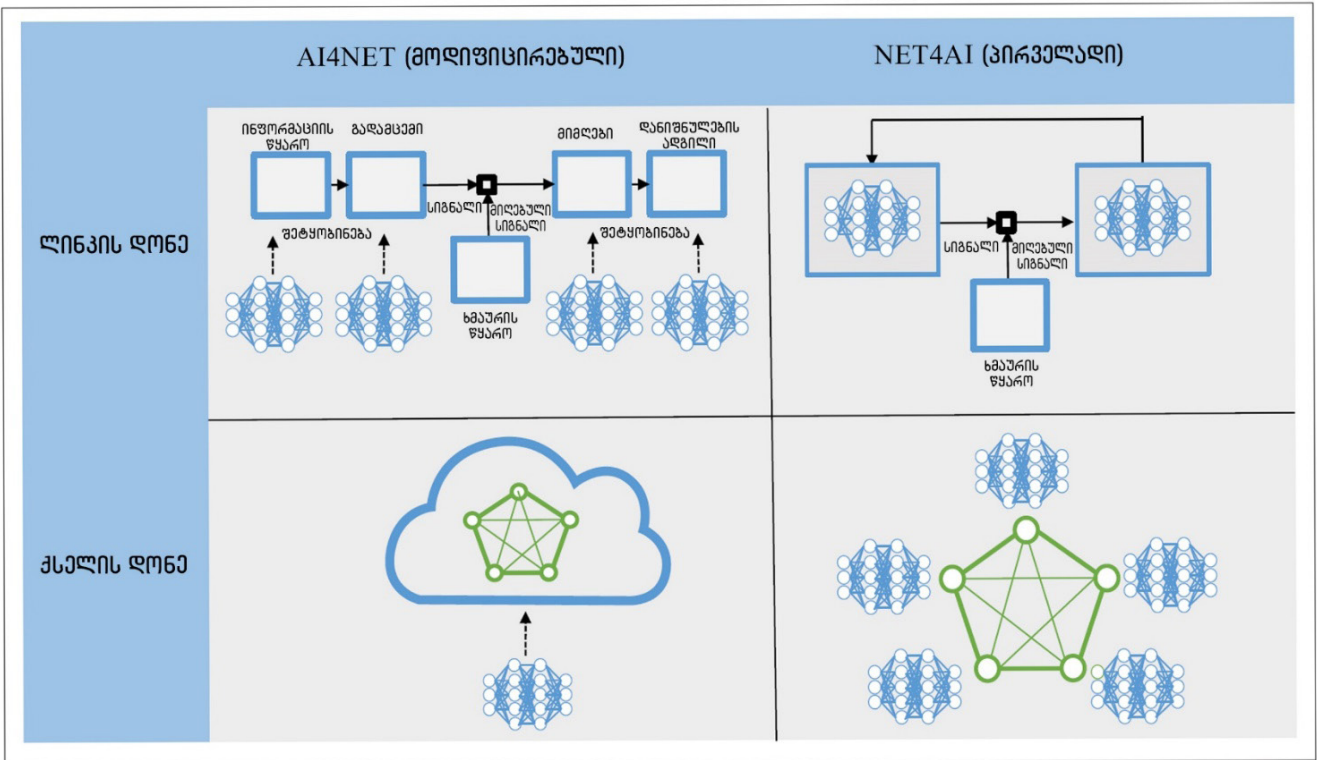
გამოთვლითი სიმძლავრე გადამწყვეტ როლს თამაშობს AI-ის განვითარებაში. მურის (Gordon Moore) კანონის შესაბამისად, გამოთვლითი სიმძლავრე დაახლოებით 200 000-ჯერ გაიზარდა 30 წლის წინანდელთან შედარებით, რამაც 100 000-ჯერ გააუმჯობესა ML-ის უნარი. თითოეული ფენა შედგება მილიონობით ან თუნდაც ათობით მილიონი ნეირონისგან ასობით მილიონი პარამეტრით. ამ რამდენიმე თავდადებული მეცნიერის ძალისხმევით, საკმაოდ უხეში გამოთვლებისა და დიდი მონაცემების წყალობით, DL-მა გადააჭარბა ადამიანის ტვინის დონეს მრავალ აპლიკაციაში.

აღსანიშნავია, რომ ჯერ კიდევ 30 წლის წინ, მსოფლიოს წამყვანი უნივერსიტეტების დოქტორანტები ელექტრონიკის ინჟინერებსა და კომპიუტერულ მეცნიერებაში სწავლობდნენ საგნებს ნეირონული ქსელების შესახებ. მაგრამ იმ პერიოდში ML-ის მიმართულებით ჯეროვანი ცოდნის მიღების შემთხვევაშიც კი, სამსახურის პოვნა საკმაოდ რთული იყო. აქედან გამომდინარე, წარმატებული დოქტორანტების ინტერესი გადატანილი იყო უსადენო კომუნიკაციებზე და ისინი მონაწილეობას იღებდნენ პერსპექტიულად მიჩნეულ კვლევებში პირველი თაობიდან (1G) მეხუთე (5G) თაობამდე.

მომდევნო 10-20 წლის განმავლობაში ჩვენი საერთო ამოცანაა ხელი შევუწყოთ ML ტექნოლოგიების სხვადასხვა აპლიკაციის ინტელექტუალური კომუნიკაციების საშუალებით. ახლა, ჩვენ გვინდა განვმარტოთ AI-ისა და უსადენო ქსელების დამაკავშირებელი რამდენიმე კონცეფცია, მათ შორის, AI საკომუნიკაციო ქსელისთვის (AI4Net) და საკომუნიკაციო ქსელი AI-ისთვის (Net4AI). ნახ. 2.1-ზე ნაჩვენებია მათი კორელაცია და განსხვავება შესაბამისად, ლინკისა და ქსელის დონეზე. მაგალითისთვის ავიღოთ ლინკის დონე. რამდენიმე წლის წინ, ჩვენ ფოკუსირებული ვიყავით AI-ის გამოყენებაზე გადაცემის პროცესში ინტელექტისა და მახასიათებლების გასაუმჯობესებლად და საკომუნიკაციო სისტემების სირთულის შესამცირებლად, რაც აღწერილი იყო სხვადასხვა ლიტერატურაში და რომელსაც AI4Net ეწოდება.

მომავალში, უსადენო კომუნიკაციები გახდება AI-ის ინტეგრირებული ნაწილი მახასიათებლების ოპტიმიზაციისთვის, გარდა მონაცემთა შეგროვებისა და გადაცემისა, რაც არის Net4AI. ჩვენ გვჯერა, რომ AI4Net ძირითადად დაკავშირებულია 5G და 5.5G ტექნოლოგიებთან, ხოლო Net4AI არის მთავარი საყრდენი 6G-ისა და მომავალი უსადენო ქსელებისთვის. 6G-ისთვის და მის ფარგლებს გარეთ, AI გამოყენებული იქნება ყველგან ჩვენს ყოველდღიურ ცხოვრებაში და საკომუნიკაციო ქსელები გადასცემს და შეაგროვებს ML-ის მიერ მოთხოვნილ მონაცემთა უზარმაზარ რაოდენობას. ამრიგად, AI მოახდენს რეგულაციას 6G-ში და ასევე იქნება ფუნდამენტური ბიზნესი 6G-ისთვის. AI 6G-ში ბევრად მეტია, ვიდრე უბრალოდ ინტერნეტზე დაფუძნებული გამორჩეული აპლიკაციები. 6G-ისა და მის ფარგლებს გარეთ, ჩვენ განსაკუთრებულ ყურადღებას გავამახვილებთ Net4AI-ზე.

უხეში გამოთვლები ხელს შეუწყობს AI-ის განვითარებას გარკვეული პერიოდის განმავლობაში. ახალი შეხედულებები, ახალი გამოგონებები და სხვა ინოვაციები დიდწილად ეყრდნობა AI-ის გამოთვლით სიმძლავრეს. უფრო მეტიც, AI-ის წარმატება და პოპულარიზაცია მომავალ უსადენო ქსელებში დაკავშირებული იქნება დიდი



ნახ. 2.1. AI4Net-ის ან Net4AI-ის შედარება შესაბამისად, ლინკის დონეზე და ქსელის დონეზე

მონაცემების გაუთავებელი გადაცემისა და შეგროვების პროცესთან. სამწუხაროდ, არც უხეში გამოთვლები და არც უსასრულო რაოდენობის მონაცემები არ არის მდგრადი. აქედან გამომდინარე, ჩვენ გამოვყოფთ ცხრა გამოწვევას, რომელიც უნდა მოგვარდეს 6G-ისთვის, რომლებიც ძალიან კრიტიკულია და შემდგომში დეტალურად იქნება აღწერილი. თუ არ იქნა თეორიული გარღვევა დაკავშირებულ სფეროებში, ჩვენ მაინც სიბნელეში მოგვიწევს ხეტიალი. იმედია, AI-ისა და 6G უსადენო ქსელების ურთიერთდაკავშირება სტიმულს მისცემს ახალი თეორიების შექმნის პროცესს მომავალი უსადენო ქსელებისთვის.

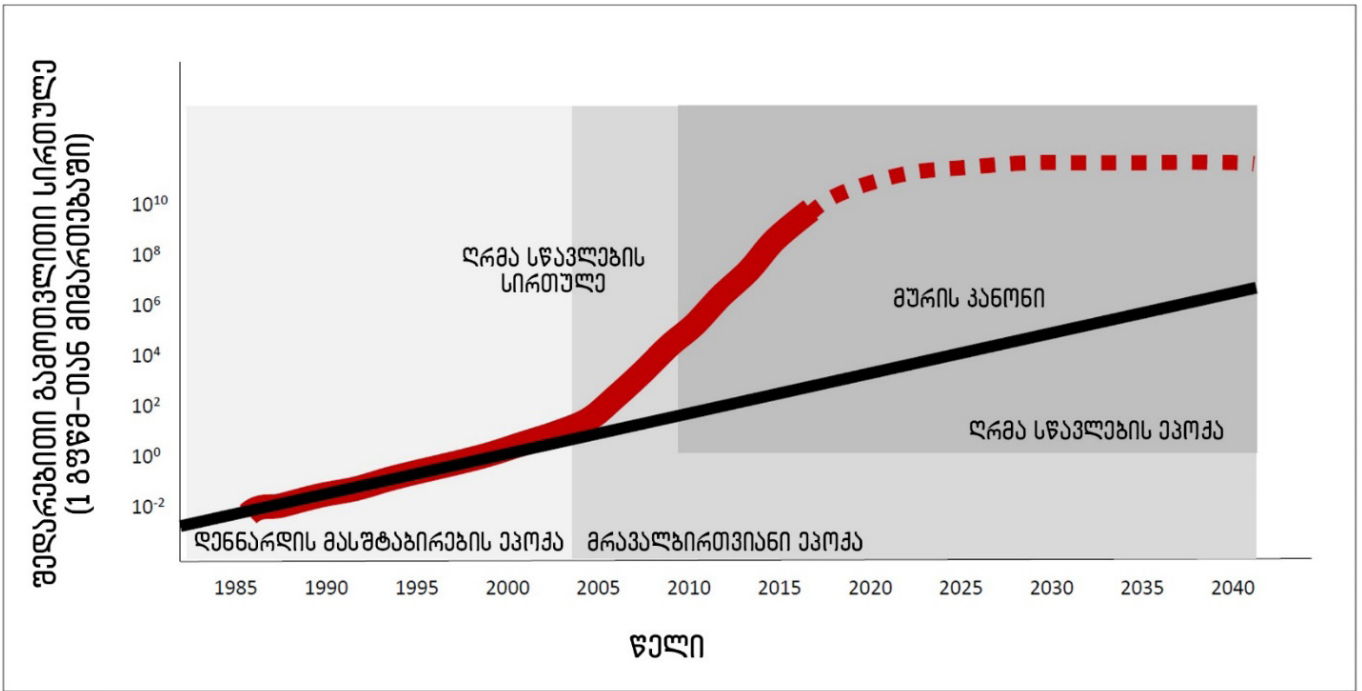
ამ თავში განვიხილავთ ცხრა ფუნდამენტურ სამეცნიერო გამოწვევას კომუნიკაციებსა და AI-ში, განსაკუთრებული აქცენტით DL და უსადენო კომუნიკაციების ინტერდისციპლინურ სფეროებზე, განსაკუთრებით ისეთ საკითხებზე, რომლებიც უზრუნველყოფენ 6G ქსელების წარმატებულ ფუნქციონირებას. კერძოდ, ისინი მოიცავენ ზოგად გამოწვევებს ღრმა ნეირონულ ქსელებში უსადენო კომუნიკაციებისთვის, განაწილებული DL და განაწილებული ნეირონული ქსელებისთვის და DL-ში ჩართული სემანტიკური კომუნიკაციებისთვის.

2.2. ძირითადი გამოწვევები ღრმა ნეირონულ ქსელებში უსადენო 6G კომუნიკაციებისთვის

ამ პარაგრაფში ჩვენ განვიხილავთ ხუთ ფუნდამენტურ გამოწვევას ML-ში, რომლებიც მჭიდრო კავშირშია 6G უსადენო კომუნიკაციებთან.

გამოწვევა 1: DL-ის გამოთვლითი კრიზისი

მასაჩუსეტსის ტექნოლოგიური ინსტიტუტის გუნდის კვლევის საფუძველზე, ნახ. 2.2 ადარებს ML-ის სირთულეს მურის კანონთან სხვადასხვა ეპოქაში. დენარდის (Robert Dennard) მასშტაბირების (Dennard scaling) ეპოქაში, გამოთვლითი შესაძლებლობები უმჯობესდება კომპიუტერის ტაქტური სიხშირის გაზრდით, ხოლო AI გამოთვლების ენერჯის მოხმარება თითქმის მიჰყვება მურის კანონს. ის იწყებს ზრდას უფრო სწრაფად, ვიდრე მურის კანონი მრავალბირთვიან ეპოქაში. დღეს ჩვენ შევდით DL-ის ეპოქაში. AI-ში უხეში გამოთვლების მოთხოვნის გამო, ენერჯის მოხმარება ბევრად აღემატება მურის კანონის ზრდის ტემპს, მინიმუმ 10^5 -ით. ამიტომ



ნახ. 2.2. გამოთვლითი სირთულე სხვადასხვა ეპოქაში, სადაც გფშმ წარმოადგენს გიგაფლოპს წამში

მდგრადი განვითარების უზრუნველსაყოფად აუცილებელია ენერჯის მოხმარების გათვალისწინება. წინააღმდეგ შემთხვევაში, ჩვენი ძალისხმევა ამ სფეროში შეიძლება უშედეგო იყოს.

უფრო მეტიც, ML-ის გამოყენებისას შეცდომის ალბათობის 5 პროცენტამდე შესამცირებლად, გამოთვლითი ღირებულება ძალიან დიდია, რაც მოითხოვს 100 მილიარდ აშშ დოლარს და 10 კვინტილიონ პროგრამულ ინსტრუქციას. ამიტომ, გამოთვლითი სირთულის შემცირება ხდება მთავარი ამოცანა მომავალი AI-ისთვის. მაშინაც კი, თუ ზოგიერთი ტექნიკა, როგორცაა გასხვლა, დაბალგანზომილებიანი შეკუმშვა, დაქვანტვის ნაკლები დონეები ან მცირე DNN, შეიძლება გამოყენებულ იქნეს პრობლემის შესამსუბუქებლად, AI-ის გამოთვლითი კრიზისი კვლავ გრძელვადიანი საინჟინრო პრობლემაა მომდევნო ათწლეულებში.

როგორც ქვემოთ დავინახავთ, განაწილებული ნეირონული ქსელები და განაწილებული სწავლება შესაძლებელს ხდის გამოთვლითი დატვირთვის გადანაწილებას სხვადასხვა მოწყობილობასა თუ ნეირონულ ქსელს შორის. ამიტომ, ზემოაღნიშნული გამოწვევების გადასაჭრელად, პირველ რიგში უნდა ჩამოვყალიბოთ გამოთვლითი ტექნიკასა და კომუნიკაციაში ენერჯის მოხმარების შეფასების მეთოდოლოგიების ნაკრები, რათა ხელი შევუწყოთ სწორი გადაწყვეტილებების შერჩევას.

გამოწვევა 2: გრადიენტის გაქრობა DL-ში

გრადიენტის გაქრობის პრობლემა ეს არის მოვლენა, რომელიც წარმოიქმნება ღრმა ნეირონული ქსელების ტრენინგის დროს, როდესაც ქსელის განახლებისთვის გამოყენებული გრადიენტები ხდება უკიდურესად მცირე ან „ქრება“, რადგან ისინი გადადიან გამომავალი ფენებიდან უფრო წინა ფენებზე.

უკუმპართულებით გავრცელების (უკუგავრცელების) გრადიენტის იტერაცია ღრმა ნეირონული ქსელების ტრენინგისას არის ძალიან რთული საკითხი აპარატურის ფუნქციონირების აჩქარებისთვის, რაც გავლენას ახდენს უსადენო გადაცემის შეყოვნებაზე და დიდი მნიშვნელობა აქვს გამჭოლი (ანუ ბოლოების დამაკავშირებელი) კომუნიკაციისთვის.

წარსულში შემუშავებული იყო პატჩინგი (patching), ანუ შესწორების მრავალი მეთოდი, რათა გადაჭრილიყო პრობლემები, რომლებიც დაკავშირებულია შეცდომების უკავარცხლებასთან გრადიენტული დაღმართის იტერაციულ ალგორითმებში. განვმარტავთ, რომ გრადიენტული დაღმართი არის ოპტიმიზაციის ალგორითმი, რომელიც გამოიყენება ნეირონული ქსელის მუშაობის გასაუმჯობესებლად, ქსელის პარამეტრებში შესწორებების შეტანით, რომ განსხვავება ქსელის პროგნოზირებად და ქსელის რეალურ/მოსალოდნელ მნიშვნელობებს შორის

(მოხსენიებული, როგორც დანაკარგი) იყოს რაც შეიძლება მცირე. განვიხილოთ ზოგიერთი მეთოდი:

- სიგმიოიდური ფუნქციის გაჯერება, როგორცაა გასწორებული წრფივი ერთეულის (ReLU) ფუნქცია, იწვევს გრადიენტის გაქრობას. ReLU ფუნქციის უარყოფითი ნახევარდერძი არის მკვდარი ზონა, სადაც გრადიენტი მიდის ნულამდე. ამიტომ, ReLU-ს გაუონილი ვერსია (LeakyReLU) და პარამეტრული ReLU (PReLU) იქნა შემოთავაზებული ReLU-ს ჩანაცვლებისთვის.
- გრადიენტისა და წონის განაწილების სტაბილურობის გასაძლიერებლად, ექსპონენციალური LU (ELU) და მასშტაბირებული ELU (SELU) იქნა გამოყენებული ღრმა ნეირონულ ქსელებში.
- ძნელია გრადიენტების გადალახვა, თუ ნეირონული ქსელი ძალიან ღრმაა. აქედან გამომდინარე, გამოიგონეს მაგისტრალური ქსელი, რომელშიც პარამეტრებიც კი არის გამოტოვებული და მათ ნაცვლად გამოიყენება ნარჩენი ქსელები (ResNets).
- პაკეტურ ნორმირებას (BatchNorm) შეუძლია ნეირონული ქსელის საშუალო მნიშვნელობის და დისპერსიის ნორმალიზება.
- ამოვარდნას (DropOut), რომელიც იგნორირებას უკეთებს ზოგიერთ ერთეულს/ნეირონს ტრენინგის დროს, შეუძლია გაზარდოს ხმაური გრადიენტულ ნაკადში. ამიტომ, მას შეუძლია შეამსუბუქოს ML-ის არასასურველი ქცევა — ზედმეტად მორგება და შეამციროს განზოგადების შეცდომა.
- RNN-ში გრადიენტი ზოგჯერ არასტაბილურია. ამიტომ ფართოდ გამოიყენება გრძელი მოკლევადიანი მეხსიერების მქონე (LSTM) ქსელი, რაც შეიძლება კიდევ უფრო გაუმჯობესდეს კარიბჭის შემცველი რეკურენტული ერთეულების დამატებით.
- იენსენ-შენონის დივერგენცია გენერაციულ-შეჯიბრებით ქსელში (GAN) იწვევს გრადიენტების გაქრობას ან არაზუსტობას, რაც შეიძლება მოგვარდეს ვასერშტეინის GAN-ის (WGAN) გამოყენებით.

ზემოთ ჩამოთვლილია შესწორების მხოლოდ რამდენიმე მეთოდი, რომელიც ვერ ახერხებს პრობლემის სრულად მოგვარებას.

კოლმოგოროვ-არნოლდის (Andrey Kolmogorov, Vladimir Arnold) თეორემის (რომლის მიხედვითაც თითოეული მრავალცვლადიანი უწყვეტი ფუნქცია შეიძლება წარმოდგენილი იყოს როგორც ერთი ცვლადის უწყვეტი ფუნქციების სუპერპოზიცია) ხელახალი გადახედვით, შესაძლოა მოხდეს ერთსაფეხურიანი ან იტერაციული ალგორითმის შემუშავება DL-ში იტერაციული გრადიენტული დაღმართის საკითხის გადასაჭრელად. DL-ში უკუგავრცელების ბუნებისა და ქცევის შესახებ სიღრმისეული კვლევის მეშვეობით და ზოგიერთი თეორიული მიდგომის განვითარებით, ჩვენ შესაძლოა შევამსუბუქოთ ან მთლიანად გადავჭრათ ეს პრობლემა. რა თქმა უნდა, ეს გრძელვადიანი მიზანია და ჩვენ მოუთმენლად ველით დიდ გარღვევას.

გამოწვევა 3: ღრმა ნეირონული ქსელების მეხსიერების ტევადობა

ვარაუდობენ, რომ ნეირონული ქსელების მეხსიერების ფუნქცია სრულად იქნება გამოყენებული მომავალ კომუნიკაციებში. მაგალითად, სემანტიკური გადამცემი და მიმღები საჭიროებს მეხსიერების უზარმაზარ ტევადობას საბაზისო ცოდნის შესანახად. ზოგადად, რაც უფრო დიდია ნეირონული ქსელის მეხსიერების ტევადობა, მით უფრო მაღალია კომუნიკაციის ეფექტიანობის გაუმჯობესების პოტენციალი. ჩნდება ბუნებრივი კითხვა: რამდენი ტიპის ობიექტის ან მოვლენის დამახსოვრება შეუძლია ნეირონულ ქსელს ან რა არის ნეირონული ქსელის მეხსიერების ტევადობა?

ლიტერატურაში ნაჩვენებია, რომ ჰოპფილდის (John Hopfield) ქსელის მეხსიერების ტევადობა n ნეირონით არის $C \approx n/(2 \log_2 n)$. თუმცა, მეხსიერების ტევადობა ზოგადი ღრმა ნეირონული ქსელებისთვის ჯერჯერობით არ არის ცნობილი.

კიდევ ერთი მნიშვნელოვანი კითხვა შეეხება მეხსიერების ტევადობისა და გამოთვლითი სირთულის ურთიერთდამოკიდებულებას. არის ეს წრფივი, ექსპონენციალური ან სხვა სახის? ასეთ ურთიერთდამოკიდებულებას დიდი მნიშვნელობა აქვს საკომუნიკაციო სისტემებში მიმღების შემუშავებისა და შეფასებისთვის.

გამოწვევა 4: DL-ის დამოკიდებულება დიდ მონაცემებზე

ღრმა ნეირონული ქსელების ტრენინგისთვის საჭირო დიდი მონაცემები ხშირად გროვდება უსადენო კომუნიკაციებით. აქედან გამომდინარე, კრიტიკული ხდება, თუ როგორ უნდა ჩავატაროთ ნეირონული ქსელების ტრენინგი ეფექტიანად და სწრაფად და, შესაბამისად, შევამციროთ მოთხოვნა დიდ მონაცემებზე და უსადენო კომუნიკაციებზე. მაშინაც კი, თუ მეხსიერების ტევადობასა და ღრმა ნეირონული ქსელების ზომას შორის კავშირი ვერ კიდევ გაურკვეველია, როგორც ეს მე-3 გამოწვევაშია მითითებული, მეხსიერების დიდი ტევადობის მქონე

ნეირონული ქსელი, როგორც წესი, შეიცავს დიდი რაოდენობის პარამეტრებს და საჭიროებს მონაცემების უზარმაზარ რაოდენობას მოდელის ტრენინგისთვის.

გარკვეული საწყისი სამუშაო ჩატარდა კომუნიკაციის სფეროში დომენის ცოდნის გამოყენების მიზნით, რათა შემცირდეს მოთხოვნილება ტრენინგის მონაცემებზე, მაგალითად, მოდელზე დაფუძნებული DL საკომუნიკაციო სისტემებისთვის. რთული ამოცანაა დომენის ცოდნის გამოყენების, დიდ მონაცემებზე მოთხოვნების, კომუნიკაციის მახასიათებლების და სისტემის სირთულის დაბალანსება. ასევე მნიშვნელოვანია ზოგადად ვიცოდეთ მოცემული საკომუნიკაციო აპლიკაციებისთვის საჭირო მონაცემების მინიმალური რაოდენობა.

გამოწვევა 5: დინამიკური, აკრეციული და მეტასწავლება

ამჟამად, AI-ის მოდელები გადიან ტრენინგს სპეციალური სწავლების ჰიპოთეზის მიხედვით, რომელიც გულისხმობს, რომ გარემო (ტრენინგის მონაცემების სტატისტიკა) სტატიკურია, სულ მცირე, ტრენინგის პერიოდში. თუმცა, რეალური საკომუნიკაციო სცენარები, განსაკუთრებით მობილური კომუნიკაციებისთვის, მუდმივად იცვლება. დინამიკური DL-ის ძირითადი მახასიათებლები უსადენო კომუნიკაციებისთვის მოითხოვს ძირფესვიან გამოკვლევას.

გამოწვევა პოტენციურად შეიძლება გადაიჭრას აკრეციული სწავლებისა (accretory learning) და მეტასწავლების განვითარებით. ადამიანთა აკრეციული სწავლების თეორია შემოგვთავაზებს კოგნიტურმა ფსიქოლოგებმა 40 წელზე მეტი ხნის წინ. ახლახან შემუშავდა აკრეციული სწავლების კონცეფცია კლასიფიკაციისთვის. კერძოდ, ტრადიციული ML ატრენინგებს კლასიფიკაციის მოდელს ფიქსირებული ამოცნობის კომპლექტებით, რაც ვერ ხერხდება მონაცემთა ნაკრებისთვის ახალი ფუნქციებით. აკრეციულ სწავლებას შეუძლია პრობლემის მოგვარება ღრმა ნეირონული ქსელების გაფართოებით, აკრეციის, დაბუსტების და რესტრუქტურის ურთიერთქმედების გზით. უფრო მეტიც, მეტასწავლება, რომელიც ასევე ცნობილია როგორც „სწავლება სწავლისთვის“, ბოლო დროს ხდება აქტიური კვლევის სფერო. მისი განვითარებით, ერთ დღეს შეიძლება მიღწეულ იქნეს დინამიკური სწავლება საკომუნიკაციო სისტემებისა და ქსელებისთვის.

2.3. განაწილებული ნეირონული ქსელები და განაწილებული სწავლება

დღესდღეობით მსოფლიოში დაახლოებით შვიდი მილიონი საბაზო სადგური და ექვსი მილიარდი მობილური ტერმინალია და ეს რიცხვები სულ უფრო იზრდება. შეუძლებელია ყველა მონაცემის შეგროვება სხვადასხვა საბაზო სადგურიდან და ტერმინალიდან ყველაზე თანამედროვე სერვერზეც კი ცენტრალიზებული დამუშავებისთვის, შენახვის უზარმაზარი მოთხოვნილების, მაღალი გამოთვლითი სირთულის, ენერჯის მაღალი მოხმარების და კონფიდენციალურობის საკითხების გამო. ამიტომ, განაწილებული ნეირონული ქსელები და განაწილებული სწავლება მომავალში უაღრესად მნიშვნელოვანი გახდება.

გამოწვევა 6: კოლექტიური სწავლება უსადენო კომუნიკაციების მონაცემებისთვის

კოლექტიური სწავლება გულისხმობს, რომ მრავალი სერვერი აგროვებს მონაცემებს და ერთობლივად ატრენინგებს ML მოდელს. ის მიეკუთვნება განაწილებული სწავლების ზოგად კატეგორიას, მაგრამ განსაკუთრებული დამოკიდებულება აქვს კომუნიკაციებთან. კოლექტიური სწავლების არქიტექტურა და მახასიათებლები დამოკიდებულია როგორც AI-ზე, ასევე კომუნიკაციებზე. AI-ისა და უსადენო კომუნიკაციების ინტეგრირებულ არქიტექტურაში, პირველადი პრობლემებია მონაცემების სეგმენტირება და კომუნიკაციისა და გამოთვლების გამიჯვნა, სადაც მონაცემები ძირითადად გამოიყენება ML მოდელის ტრენინგისთვის, ხოლო კომუნიკაციის მოდელი გამოიყენება მონაცემთა/პარამეტრების გადაცემისთვის. აქ ჩვენ განსაკუთრებულ ყურადღებას ვაქცევთ კომუნიკაციის საკითხებს. კერძოდ, როგორ დავყოთ სეგმენტებად ნეირონული ქსელი და მისი ამოცანები კოლექტიური სწავლებისთვის, არის გადაუჭრელი პრობლემა, რომელიც ასევე ღრმა თეორიული საკითხია. ზოგიერთი დაკავშირებული კვლევითი პრობლემა შემდეგია:

გამოწვევა 7: უსადენო კომუნიკაციებში ჩართული ფედერირებული სწავლება

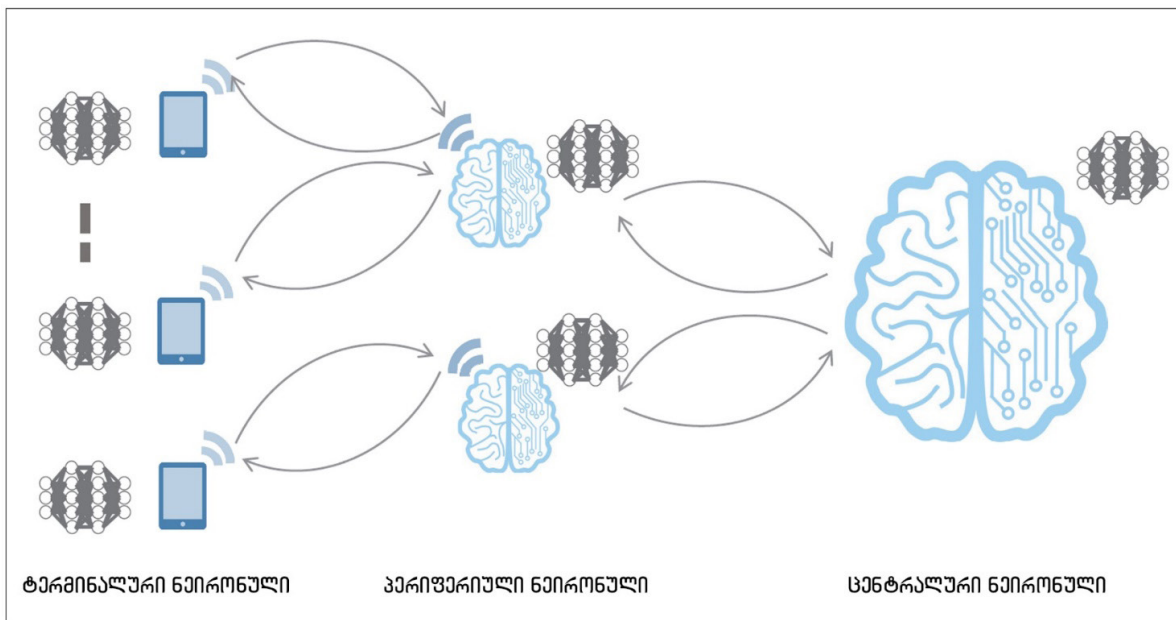
ფედერირებული სწავლების გამოყენება, რომელიც განაწილებული სწავლების სახეობაა, შესაძლებელი ხდება უსადენო კომუნიკაციებით. მე-ნ გამოწვევაში განხილული კოლექტიური სწავლებისგან განსხვავებით, ფედერირებული სწავლება უფრო მეტად ეყრდნობა კომუნიკაციებს. როგორც ნაჩვენებია ნახ. 2.3-ზე, მრავალი კლიენტი/ტერმინალი ერთობლივად ატრენინგებს ერთსა და იმავე მოდელს ფედერირებულ სწავლებაში და ადგილი აქვს მხოლოდ პარამეტრების გაცვლას მრავალ ტერმინალს შორის, და არა მონაცემების, ამოცანების

და ნეირონული ქსელების სეგმენტირებას, როგორც ეს კოლექტიურ სწავლებაშია. ინფორმაციის გაცვლა ტერმინალსა და პერიფერიულ ნეირონულ ქსელებს შორის დამოკიდებულია უსადენო კომუნიკაციებზე. იგივე ხდება პერიფერიულ ნეირონულ ქსელებსა და ცენტრალურ ქსელებს შორის. ამიტომ, გადაცემის შეცდომა მნიშვნელოვნად მოქმედებს განაწილებულ/ფედერირებულ სწავლებაზე.

წარსულში უსადენო კომუნიკაციების და AI-ის სფეროები დამოუკიდებლად შეისწავლებოდა ოპტიმიზაციის სხვადასხვა მიზნით. კერძოდ, საკომუნიკაციო სისტემები, როგორც წესი, შექმნილია მონაცემთა გადაცემის სიჩქარის მაქსიმალურად გაზრდისა და ბიტზე შეცდომის ალბათობის შესამცირებლად, ხოლო AI-ის მახასიათებლები, როგორც წესი, იზომება ამოცნობის სიზუსტით და რეპროდუცირებული მონაცემების ხარისხით. ზოგადად, AI-ისა და უსადენო კომუნიკაციების ურთიერთქმედება და მათ შორის კავშირი ახალი ტენდენციაა, რომელიც აქამდე არასდროს ყოფილა გამოკვლეული.

მაგალითისთვის ავიღოთ ღრმა ნეირონული ქსელის სწავლების პროცესი. მაშინაც კი, თუ კოეფიციენტის შეცდომის გავლენა პირველ რამდენიმე იტერაციაში ტრივიალურია, ის ხდება ძალიან კრიტიკული და მგრძობიარე ბოლო რამდენიმე ფენისათვის ჩატარებულ ბოლო რამდენიმე იტერაციისას. ამრიგად, ნეირონული ქსელის სხვადასხვა იტერაციასა და სხვადასხვა ფენაზე, მოთხოვნები უსადენო გადაცემის ხარისხზე მრავალფეროვანია. ანალოგიური პრობლემა მქვანდება პროგნოზირების (ანუ, ვარაუდით დასკვნის გამოტანის) მოდელებში.

შეცდომის მგრძობიარეობასთან დაკავშირებული ზემოაღნიშნული პრობლემის გარდა, არსებობს მრავალი სხვა გამოწვევა. მაგალითად, როგორ შევკუმშოთ მოდელის პარამეტრები საკომუნიკაციო რესურსების დაზოგვის მიზნით და როგორ მოვახდინოთ პარამეტრების საიმედო აგრეგაცია ცენტრალურ სერვერზე, რათა აღმოიფხვრას კომუნიკაციის შეცდომები და შეფერხებები. ეს გამოწვევები წარმოადგენს ინოვაციურ შესაძლებლობებს 5G-ისთვის. სასურველია თეორიული საფუძვლიდან გავარკვიოთ კავშირი ML მოდელის შეცდომებსა და უსადენო გადაცემისას წარმოქმნილ სხვადასხვა შეცდომას/ნაკლოვანებას შორის, რათა გამარტივდეს მათი ერთობლივი ოპტიმიზაციის პროცესი.



ნახ. 2.3. განაწილებული სწავლება

2.4. სემანტიკური კომუნიკაციები

სემანტიკური კომუნიკაციები პოსტ-შენონის ეპოქაში წარმოშობილი პრობლემაა. დაახლოებით 70 წლის წინ, შენონთან თანავტორობით გამოცემულ ნაშროში, ვივერმა (Warren Weaver) კომუნიკაციები სამ დონედ გაანაწილა, რომელიც მოიცავს სიმბოლოების გადაცემას, სემანტიკურ გაცვლას და სემანტიკური გაცვლის ეფექტს. როდესაც შენონმა შექმნა ინფორმაციის თეორია, მან ყურადღება გაამახვილა მხოლოდ სიმბოლოების გადაცემის მათემატიკურ თეორიაზე. კერძოდ, წყაროში ინფორმაციის რაოდენობა განსაზღვრულია მისი ენტროპიით, ხოლო საკომუნიკაციო არხის გამტარუნარიანობა დადასტურებულია, რომ არის მაქსიმალური ერთობლივი ინფორმაცია არხის შესასვლელსა და გამოსასვლელს შორის, სადაც ორივე იზომება ბიტებით. ზოგადი მათემატიკური ინსტრუმენტების არარსებობის გამო სემანტიკისა და გადაცემული შინაარსის მნიშვნელობის მოდელირებისა და ანალიზისთვის, წარსულში მიღწეული იყო შეზღუდული პროგრესი, მიუხედავად იმისა, რომ მრავალი გამოჩენილი მკვლევარი შეეხო ამ სფეროს.

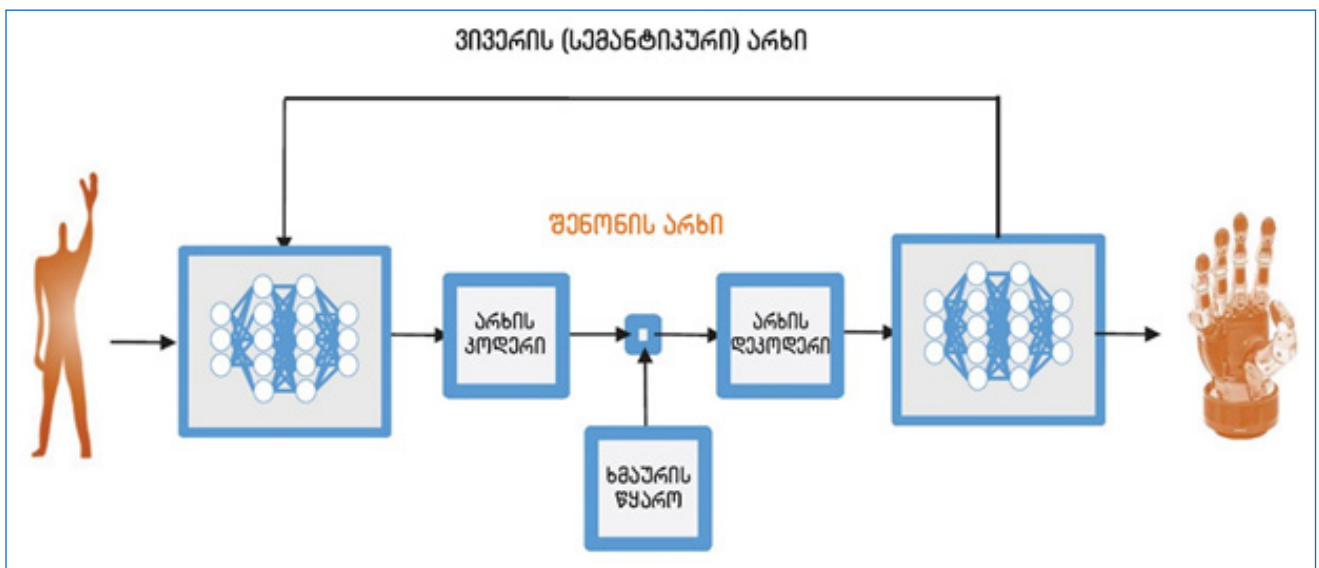
DL-ისა და მისი აპლიკაციების ბოლოდროინდელი მიღწევები საშუალებას გვაძლევს გადავხედოთ ვივერის მეორე და თუნდაც მესამე დონის მოდელს, სადაც მიზანია არა ბიტების თანმიმდევრობის გადაცემა, არამედ შინაარსის მნიშვნელობა. ქვემოთ აღნიშნულ პროცესს ვუწოდებთ სემანტიკურ კომუნიკაციებს. კომუნიკაციის წინასწარი არქიტექტურა ნაჩვენებია ნახ. 2.4-ბე, სადაც შენონის არხი ჩამოყალიბებულია შიგნით და ვივერის სემანტიკური არხი აგებულია გარეთ. ამ არქიტექტურის მიხედვით, სამეცნიერო კითხვები მოიცავს შემდეგ საკითხებს:

- როგორ გავზომოთ სემანტიკა?
- რა არის სემანტიკური საკომუნიკაციო სისტემების ზღვარი?
- არსებობს თუ არა შენონის ინფორმაციის თეორიის ანალოგიური თეორია?
- როგორ დავნერგოთ და განვახორციელოთ სემანტიკური კომუნიკაციები ეფექტიანად?

სემანტიკურმა კომუნიკაციებმა დიდი ყურადღება მიიპყრო გასული რამდენიმე წლის განმავლობაში და ლიტერატურაში წარმოდგენილია გარკვეული საწყისი კვლევითი შედეგები ამ სფეროში. ჩვენ გვჯერა, რომ უახლოეს მომავალში კიდევ უფრო მეტი სამუშაო ჩატარდება სემანტიკური კომუნიკაციების თეორიისა და პრაქტიკული რეალიზაციის მიმართულებით.

გამოწვევა 8: სემანტიკური კომუნიკაციების საფუძველი

კოლმოგოროვმა, რომელიც იყო გასული საუკუნის დიდი მათემატიკოსი, გადაამუშავა შენონის ინფორმაციის თეორია და კიდევ უფრო გააფართოვა ინფორმაციის თეორიის მათემატიკური საფუძველები. მისი გამოკვლევის მიხედვით, ავტომორფიზმის დიდი კლასის სისტემების ეკვივალენტობის საკმარისი და აუცილებელი პირობაა,



ნახ. 2.4. შენონისა და ვივერის არხები ტიპური კომუნიკაციებისა და სემანტიკური კომუნიკაციებისთვის

რომ მათ ჰქონდეთ ერთნაირი ენტროპია. კოლომგოროვმა ასევე გაავრცელა შენონის ენტროპიის მათემატიკური საფუძვლები ტოპოლოგიისა და ალგებრის ახალ სფეროებში.

ასევე განსაკუთრებით მნიშვნელოვანია შენონის შემდგომი პერიოდის კომუნიკაციების ან სემანტიკური კომუნიკაციების საფუძველი. ენტროპია და გამტარუნარიანობა შენონის ინფორმაციის თეორიაში განისაზღვრება სტატისტიკური ალბათობის საფუძველზე, რომელიც შეიძლება არ იყოს გამოყენებული სემანტიკური კომუნიკაციებისთვის.

სემანტიკური კომუნიკაციების მათემატიკური საფუძვლის შესაქმნელად, სტატისტიკური ალბათობა ჩანაცვლებულია შეტყობინების ლოგიკური ალბათობით, სადაც შესაბამისად განისაზღვრება სემანტიკური წყარო, სემანტიკური ხმაური, სემანტიკური ენტროპია, სემანტიკური შესაძლებლობები და სემანტიკური კოდირება. უფრო მეტიც, სემანტიკური წყაროს კოდირების თეორემა და სემანტიკური არხის კოდირების თეორემა, რომელიც ანალოგიურია შენონის ინფორმაციის თეორიაში წარმოდგენილი თეორემებისა უკვე ჩამოყალიბებული და დამტკიცებულია.

სემანტიკური კომუნიკაციების საფუძვლების შესაქმნელად სასურველია ავაშენოთ ყოვლისმომცველი ჩარჩო, რომელიც განსაზღვრავს მის ოპტიმალურ სტრუქტურას და უზრუნველყოფს სახელმძღვანელო მითითებებს განხორციელებისთვის. თუმცა, სემანტიკისა და გადაცემული შინაარსის მნიშვნელობის მოდელირებისა და ანალიზის მათემატიკური ინსტრუმენტი ჯერ კიდევ არ არის შემუშავებული, რაც ართულებს ისეთივე სემანტიკური თეორიის შემუშავებას, როგორც შენონის ინფორმაციის თეორიაა. მეორე მხრივ, სემანტიკური კომუნიკაციების განხორციელება დიდწილად დამოკიდებულია AI-ზე. აქედან გამომდინარე, აზრი აქვს AI-ზე დაფუძნებული თეორიის ან სემანტიკური კომუნიკაციების საფუძვლების შემუშავებას.

გამოწვევა 9: სემანტიკური საკომუნიკაციო სისტემების სტრუქტურა

ჯერ კიდევ შენონის ფუნდამენტურ შრომებში, გაკეთდა ყველაზე ადრეული განცხადება ბუნებრივი ენის AI-ის შესახებ. ინგლისური ენის შესახებ დევეის (John Dewey) სტატისტიკურ შედეგებზე დაყრდნობით, შენონი ცდილობდა ინგლისური წინადადებების რეკონსტრუქციას შემთხვევითი წესით აღებული ინგლისური ასოებიდან, ალბათობისა და სტატისტიკური მეთოდების გამოყენებით სემანტიკური კომუნიკაციების გამოსაკვლევად. იმ დროს შენონმა ნაწილობრივ უკან დაიხია ალბათობის ფუნდამენტური თეორიიდან, რათა მხარი დაეჭირა ენტროპიის კონცეფციისთვის ინფორმაციის თეორიაში, რომელიც რეალურად უგულვებლყოფდა სემანტიკური კომუნიკაციების მნიშვნელოვან ელემენტებს. ჰაფმანის კლასიკური წყაროს კოდირებისა და ღრმა ნეირონული ქსელის მეთოდები 2018 წელს შედარდა ტექსტის გადაცემისთვის. აღმოჩნდა, რომ ღრმა ნეირონული ქსელის გამოყენებით გაუმჯობესება საკმაოდ შეზღუდულია, რამაც კიდევ ერთხელ დაადასტურა შენონის ინფორმაციის თეორიის სამართლიანობა.

კრიტიკული საკითხია შემდეგი: ინფორმაციის აღდგენა სემანტიკურ კომუნიკაციებში უნდა მოხდეს შესაბამისი ბიტების ზუსტი აღდგენით თუ პირდაპირი სემანტიკური აღდგენის გზით. ეს შეიძლება ჩაითვალოს სტრუქტურულ პრობლემად სემანტიკურ კომუნიკაციებში, როგორც ამას მიუთითებენ შენონი და ვივერი. აქ ჩნდება კითხვა, უნდა გამოვიყენოთ თუ არა ზოგადი DL ნეირონული ქსელი სემანტიკური კომუნიკაციების ტრენინგისა და სწავლებისთვის, თუ უნდა გამოვიყენოთ შემდგომში კომუნიკაციის სხვადასხვა სტრუქტურული დონის შესასწავლად.

ცოტა ხნის წინ, შეიქმნა DL-ზე დაფუძნებული სემანტიკური კომუნიკაციის სისტემა ტექსტის გადაცემისთვის, სახელად DeepSC (ანუ, ღრმა სემანტიკური კომუნიკაციები). ტრანსფორმატორისა და დანაკარგების ახალი ფუნქციის შემუშავების დახმარებით, DeepSC მიზნად ისახავს სისტემის გამტარუნარიანობის მაქსიმიზაციას და სემანტიკური შეცდომების მინიმიზაციას წინადადებების მნიშვნელობის აღდგენის გზით, ნაცვლად ბიტზე ან სიმბოლოზე წარმოქმნილი შეცდომების გასწორებისა, როგორც ესა ტრადიციულ კომუნიკაციებშია. სემანტიკური კომუნიკაციების მახასიათებლების სწორად შესაფასებლად, შემოთავაზებულია ახალი მეტრიკა, სახელად წინადადებების მსგავსება.

ტრადიციულ საკომუნიკაციო სისტემებთან შედარებით, რომლებიც არ ითვალისწინებენ სემანტიკური ინფორმაციის გაცვლას, DeepSC უფრო მდგრადია დროში ცვალებადი არხებისთვის და შეუძლია მიაღწიოს ორენოვანი შეფასების შესწავლის (BLEU) უკეთეს მახასიათებლებს, განსაკუთრებით სიგნალისა და ხმაურის თანაფარდობის (SNR) დაბალი მნიშვნელობებისათვის, როგორც ნაჩვენებია ნახ. 2.5-ზე. უკვე პირველი შედეგები

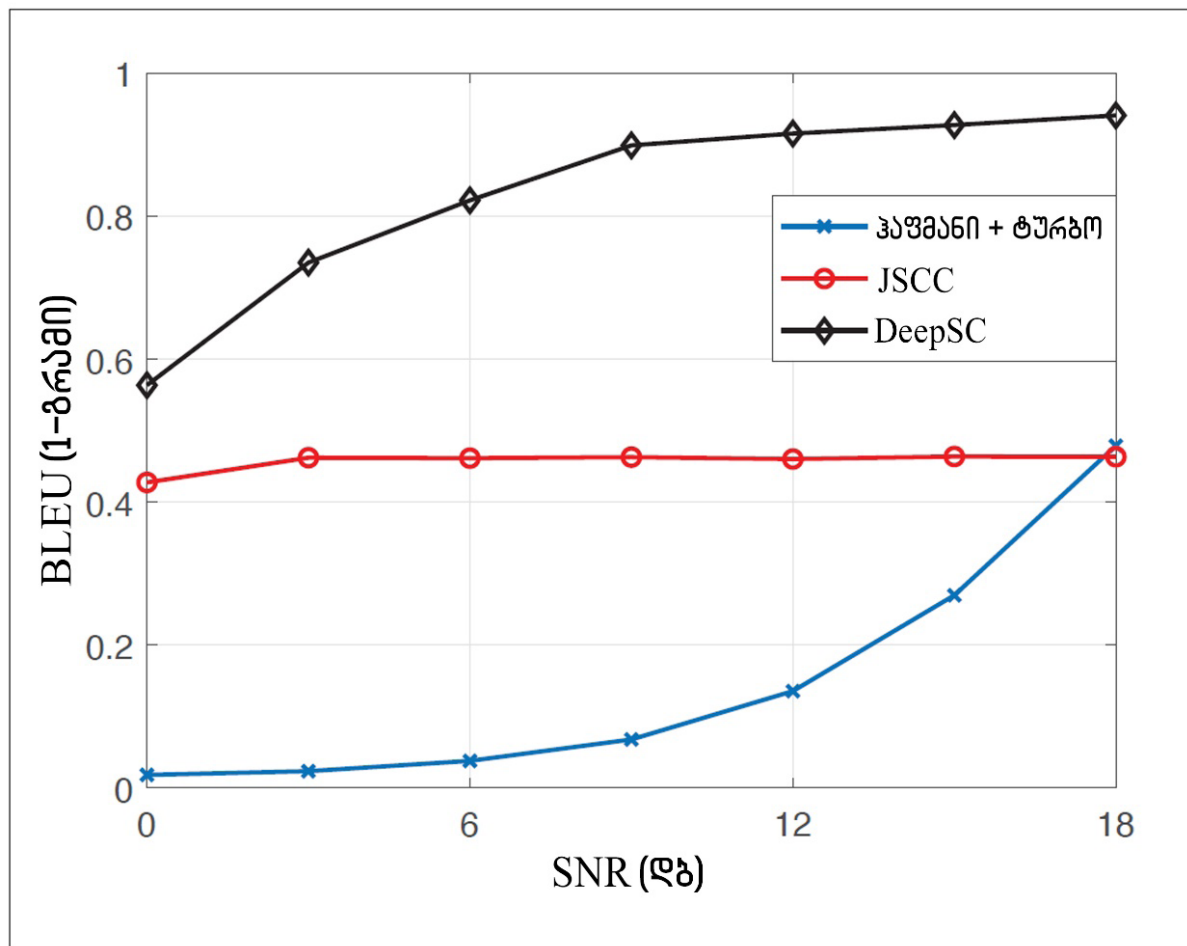
გვიჩვენებს სემანტიკური კომუნიკაციების ეფექტიანობას. მოსალოდნელია, რომ ჩატარდება მეტი კვლევა ამ სფეროში, როგორცაა სემანტიკური კომუნიკაციები ვიდეოს, გამოსახულებისა და ხმის გადაცემისთვის.

2.5. მეორე თავის დასკვნა

ამ თავში განვიხილეთ ცხრა გამოწვევა 6G-ის წარმატებული ფუნქციონირების უზრუნველსაყოფად, მათ შორის:

- გამოთვლითი პრობლემა DL-ში (გამოწვევა 1).
- ღრმა სწავლების პროცესში კომუნიკაციასთან დაკავშირებული გამოწვევები (გამოწვევები 2-5).
- განაწილებული სწავლების არქიტექტურული საკითხები კომუნიკაციების საშუალებით (გამოწვევები 6 და 7).
- გამოწვევები DL-ით მხარდაჭერილი სემანტიკური კომუნიკაციის დროს (გამოწვევები 8 და 9).

ჩვენ დარწმუნებული ვართ, რომ კვლევის ეს მდიდარი სფერო, უახლოეს მომავალში, შეიმუშავებს უფრო პერსპექტიულ გადაწყვეტილებებს ამ გამოწვევების გადასაჭრელად, რაც ღრმა გავლენას მოახდენს მომავალი კომუნიკაციების ევოლუციურ განვითარებაზე.



ნახ. 2.5. სემანტიკური საკომუნიკაციო სისტემის ეფექტიანობა: DeepSC-ის შედარება არსებულ სტანდარტებთან, წყაროსა და არხის ერთობლივ კოდირებასთან (JSCC) და (ჰაფმანის) წყაროს და (ტურბო) არხის კოდირების სისტემასთან

თავი 3 ინტელექტუალურ-ენდოგენური ქსელები: ინოვაციური ქსელის პარადიგმა 6G-ისტვის

3.1. შესავალი

5G ტექნოლოგიის განვითარებამ გარკვეულწილად შეამსუბუქა საგნების ინტერნეტის (IoT) და ინდუსტრიული ქსელების წინაშე არსებული გამოწვევები. 5G უზრუნველყოფს ქსელის სამი ტიპის სერვისს: გაძლიერებულ მობილურ ფართობოლოვან (eMBB) კომუნიკაციებს, ულტრა საიმედო დაბალი შეყოვნების კომუნიკაციებს (URLLC) და მასობრივ მანქანური ტიპის კომუნიკაციებს (mMTC), რომლებიც უზრუნველყოფენ აუცილებელ ქსელურ ინფრასტრუქტურას განვითარებადი ვირტუალური/მომატებული რეალობისთვის, ავტონომიური მართვისთვის, ჯკვიანი ქსელებისთვის და IoT აპლიკაციებისთვის.

5G ძირითადად ფოკუსირებულია ზოგიერთი კონკრეტული სცენარის და ძირითადი აპლიკაციების გადაუდებელი საჭიროებების დაკმაყოფილებაზე ვერტიკალურ ინდუსტრიებში. ამის შემდეგ, ქსელის შესაძლებლობების გაფართოება მაღალსიხქარიანი მობილური კავშირის მისაღწევად ყველაფერთან, ნებისმიერ დროს და ნებისმიერ ადგილას, და შემდგომი მხარდაჭერა სენსორული ურთიერთდაკავშირების, ჰოლოგრაფიული გადაცემის, ციფრული ტყუპების (digital twins) და ა. შ., გახდა მობილური საკომუნიკაციო ქსელების განვითარების ახალი ხედავ. მომავალი 6G საკომუნიკაციო სისტემების წინაშე აღმოჩნდება გამოწვევები მონაცემთა ულტრა მაღალი სიხქარის (4.3 ტბიტი/წმ-მდე), ულტრა დაბალი შეყოვნების (1 მწმ-ზე ნაკლები) და ულტრა მაღალი საიმედოობის (მტყუნების შანსი ერთი მილიონიდან) მიღწევაში. მაგალითად, ზუსტი ინდუსტრიული წარმოება მოითხოვს ბოლოების დამაკავშირებელ (გამჭოლ) შეყოვნებას მკწმ-ებსა და მწმ-ებს შორის, ტაქტილური ინტერნეტი მოითხოვს გამჭოლ შეყოვნებას დაახლოებით 1 მწმ-ს, ხოლო ზოგიერთი ჯკვიანი ქსელის სერვისი მოითხოვს 99.9999 პროცენტის საიმედოობას. გარდა ამისა, პროცესები, როგორცაა ჰოლოგრაფიული შეხვედრა, დისტანციური ჰოლოგრაფიული ქირურგია და მრავალსენსორული კომუნიკაცია მოითხოვს არა მხოლოდ ულტრა დაბალ შეყოვნებას, არამედ მონაცემთა გადაცემის მაღალ სიხქარეს ტერაბიტამდე წამში. 6G ქსელებს სასწრაფოდ ესაჭიროება მექანიზმი მახასიათებლების სხვადასხვა მოთხოვნის დასაკმაყოფილებლად, რაც უფრო მაღალ მოთხოვნებს უყენებს ქსელის ინტელექტუალური კონფიგურაციის შესაძლებლობებს.

ვირტუალური და ფიზიკური ინტეგრაციის, იმერსიის, კონტექსტუალიზაციის, პერსონალიზაციისა და ყოვლის-მომცველი ბიზნესმოთხოვნების გათვალისწინებით, ისევე როგორც კოსმოსურ-საჰაერო-სახმელეთო-სა-ზღვაო ინტეგრირებული საკომუნიკაციო ქსელის მოთხოვნების გათვალისწინებით, 6G ქსელს უნდა შეეძლოს საჭიროებისამებრ შეიცვალოს ნებისმიერი ახალი სერვისის დასაკმაყოფილებლად, რაც ძალიან მაღალ მოთხოვნებს უყენებს 6G ქსელების ინტელექტს. არსებული ქსელების სამუშაო პარადიგმები, ძირითადად მკაცრ წესებზე დაფუძნებული ალგორითმებით, რთულია დინამიკურად მოერგოს მომხმარებლის საჭიროებებსა და ქსელური გარემოს მუდმივ ცვლილებებს. უფრო მეტიც, ქსელის მუშაობის გამოცდილების ეფექტიანად დაგროვება შეუძლებელია, რაც ზღუდავს ქსელის შესაძლებლობების მუდმივ გაუმჯობესებას. ანუ, ამჟამინდელი მოქმედი პარადიგმების პირობებში, ქსელებს არ გააჩნიათ თვითგანვითარების შესაძლებლობა. ნებისმიერი განახლება და გაუმჯობესება უნდა ეყრდნობოდეს მრავალ პროფესიონალურ სამუშაოს. ეს მიუღებელია 6G ქსელებისთვის, უპრეცედენტო გეოგრაფიული მასშტაბით და ოპერაციული სირთულით. ქსელში „ინტელექტუალური გენების“ დამატება ქსელებში ინტელექტუალური და თვითგანვითარების შესაძლებლობების შესაქმნელად, ზემოაღნიშნული პრობლემების გადაჭრის მნიშვნელოვანი გზაა.

ამჟამად, AI ტექნოლოგიები, რომლებიც წარმოდგენილია ღრმა ML-ითა და ცოდნის გრაფით, სწრაფად ვითარდება. განვმარტავთ, რომ ცოდნის გრაფი ფორმალურად წარმოადგენს სემანტიკას ერთეულებისა და მათი ურთიერთობების აღწერით. ცოდნის გრაფმა შეიძლება გამოიყენოს ონტოლოგიები, როგორც სქემის ფენა. ამით ის იძლევა ლოგიკურ დასკვნას არაცხადი ცოდნის მოსაპოვებლად, იმის ნაცვლად, რომ დაუშვას მხოლოდ მოთხოვნები, რომლებიც დაკავშირებულია ცხად ცოდნასთან.

ქსელის და მასთან დაკავშირებული მომხმარებლების, სერვისებისა და გარემოს მრავალგანზომილებიანი სუბიექტური და ობიექტური ცოდნა შეიძლება იქნეს დახასიათებული, აგებული, შესწავლილი და განახლებული AI ტექნოლოგიების დანერგვით. მიღებული ცოდნის საფუძველზე სტერეო (ანუ მოცულობითი) აღქმა,

გადაწყვეტილების მიღება და ქსელის დინამიკური კონფიგურაცია, შეიძლება განხორციელდეს ისე, რომ ქსელი საჭიროებისამებრ შეიცვალოს ნებისმიერი ახალი სერვისისთვის, რომელიც ჩვენ გვსურს. ასეთ ქსელს ინტელექტუალურ-ენდოგენურ ქსელს (IEN) ვუწოდებთ.

ამჟამად, საკმაოდ ბევრი AI მეთოდი გამოიყენება ქსელის კვლევაში, მაგრამ მათი უმეტესობა უბრალოდ იყენებს ML ალგორითმებს ქსელის კონკრეტული პრობლემების გადასაჭრელად. თუმცა, IEN არ არის AI მეთოდების მარტივი გამოყენების შედეგი კონკრეტული პრობლემების გადასაჭრელად, არამედ ის წარმოდგენს AI სისტემის დიზაინსა და კონსტრუირების პრაქტიკას, რომელიც კარგად ერგება ქსელურ სისტემებს, მათთან დაკავშირებულ ქსელურ არქიტექტურასა და სამუშაო გარემოს, ასეთი სისტემებისთვის, თანამედროვე AI პრინციპებისა და მეთოდოლოგიების, ასევე ქსელის სისტემის ბუნებრივი ატრიბუტებისა და მუშაობის მახასიათებლების ხელმძღვანელობით.

მომავალი 6G სერვისების ხარისხის მაღალი მოთხოვნების და მომხმარებლების გამოცდილების ხარისხის სათანადო მოთხოვნების დასაკმაყოფილებლად, ასევე ქსელის ავტომატური მუშაობისა და კონტროლის განსახორციელებლად, ეს თავი გვათავაზობს IEN-ის კონცეფციას. ამ თავის ძირითადი პუნქტები შემდეგნაირადაა შეჯამებული:

- შემოთავაზებულია ორფენიანი ჩაკეტილი მარყუჟის ქსელის ფუნქციონირებისა და მართვის ინოვაციური არქიტექტურა. შიდა მარყუჟი არის თვითგანვითარებადი ბირთვი, რომელიც იღებს ცოდნის გრაფს, როგორც ბირთვს ცოდნის შეგროვების, ანალიზისა და განახლებისთვის. გარე მარყუჟი არის ენდოგენურ ინტელექტუალურ ორიენტირებული ქსელის ოპერირებისა და მართვის დახურული კონტური, რომელიც ახორციელებს ქსელის ავტონომიურ მუშაობას და თვითკონტროლს.
- გაანალიზებულია IEN-ის აგების პრინციპები და ძირითადი ტექნოლოგიები. იგი მოიცავს ცოდნის გრაფს, დიდი მონაცემების, AI-ისა და სხვა ტექნოლოგიების გამოყენებას ცოდნის წარმოდგენისა და განახლების განსახორციელებლად, ქსელის სტერეო აღქმას, მოთხოვნილი რესურსების განაწილებას და ქსელის თვითგანვითარების მექანიზმს.
- წარმოდგენილია IEN-ის გამოყენების ტიპური შემთხვევა. ავტონომიური ქსელის გამტარუნარიანობის გაზრდის გამოყენების შემთხვევის აღწერით, IEN-ის მუშაობის მექანიზმი ინტუიციურად არის ახსნილი და გაანალიზებულია მისი უნარი დამოუკიდებლად გაუმკლავდეს ქსელის გამოწვევებს.

3.2. 6G ქსელების ინტელექტუალური მოთხოვნები

მომავალ 6G ეპოქაში უფრო პოპულარული გახდება ქსელური აპლიკაციები, რომლებიც წარმოდგენილია ჰოლოგრაფიული კომუნიკაციით, მრავალენსორული კომუნიკაციით, ინტელექტუალური ტრანსპორტით და ინტელექტუალური წარმოებით. თუმცა, ამ აპლიკაციების საჭიროებები ისეთ ასპექტებში, როგორცაა ვირტუალური და ფიზიკური ინტეგრაცია, იმერსია, კონტექსტუალიზაცია და პერსონალიზაცია, უკიდურესად მაღალ მოთხოვნებს უყენებს 6G ქსელების ინტელექტს.

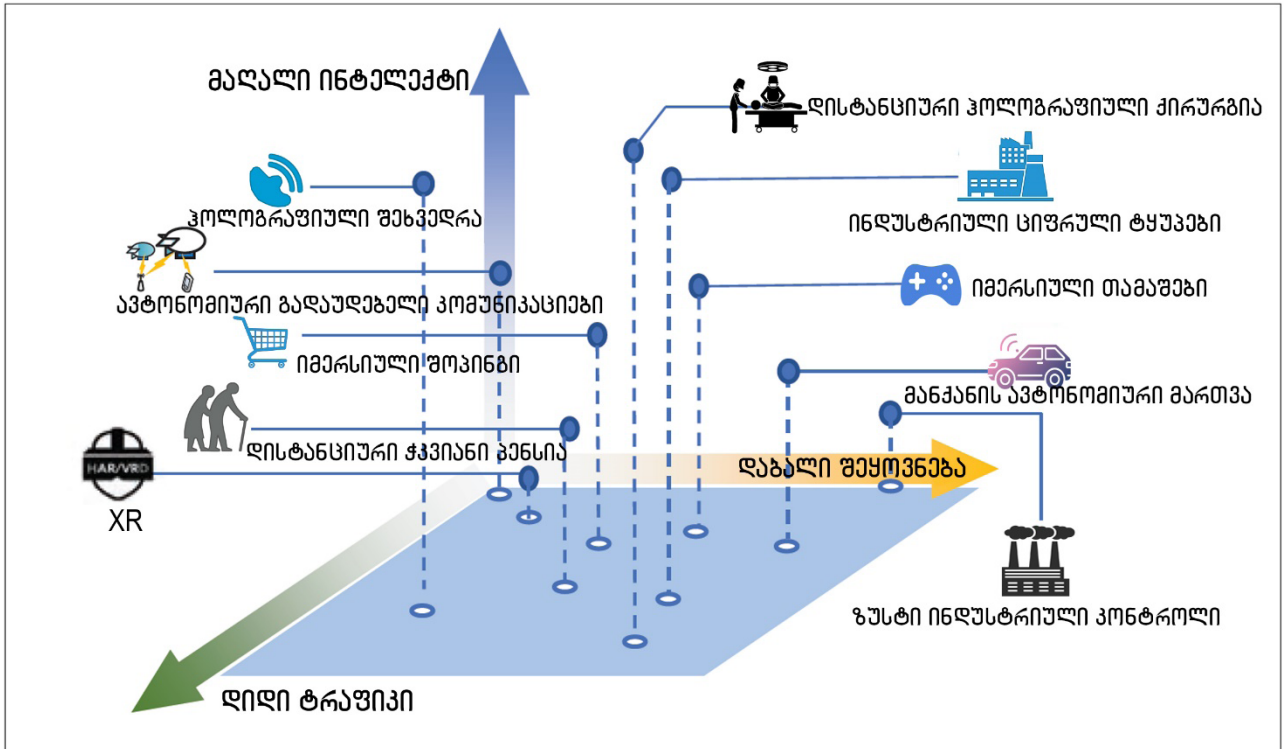
ჰოლოგრაფიული კომუნიკაციები გულისხმობს ჰოლოგრაფიული გამოსახულების ინტერაქტიულ გადაცემას წყაროდან სამიზნეზე. 6G ეპოქაში მედიასთან ურთიერთქმედების ფორმა დღევანდელი უბრალო მულტიმედიიდან გადაიქცევა მაღალი სიზუსტის მომატებული/ვირტუალური რეალობის (AR/VR) ურთიერთქმედებებამდე და ჰოლოგრაფიული ინფორმაციის ურთიერთქმედებამდე კი. ადამიანებს შეეძლება ისარგებლონ სრულად იმერსიული ჰოლოგრაფიული ინტერაქტიული გამოცდილებით ნებისმიერ დროსა და ნებისმიერ ადგილას. ჰოლოგრაფიულ სურათებზე დაფუძნებული დისპლეი, როგორც წესი, მოითხოვს მაღალ გამოთვლით სიმძლავრეს 3D სურათების სინთეზის, რენდერირების (ვიზუალიზაციის) და რეკონსტრუქციის დასასრულებლად. ჰოლოგრაფიული კომუნიკაცია მოითხოვს უკიდურესად მაღალ გამოთვლით სიმძლავრეს და უკიდურესად დაბალ შეყოვნებას, რაც მის მხარდასაჭერად საჭიროებს 3D მონაცემების მიმღები ტერმინალის მახლობლად გამოთვლების ჩატარებას. როგორც ადაპტირდება 6G ქსელები დინამიკურად ჰოლოგრაფიული კომუნიკაციის საჭიროებებთან და როგორ ეფექტიანად დაგეგმავს და დინამიკურად დაარეგულირებს ისეთ მრავალდომენიან ჰეტეროგენულ რესურსებს, როგორცაა ქსელი და გამოთვლითი სიმძლავრე, 6G ქსელის ინტელექტის წინაშე არსებული გამოწვევებია.

მრავალსენსორული კომუნიკაცია ნიშნავს, რომ ინფორმაცია უფრო მეტ სენსორულ გრძნობებს ატარებს. 5G ეპოქაში, ციფრული ვირტუალური აღქმის დანერგვით, ადამიანის ხედვა, სმენა, ყნოსვა, გემო, შეხება და სხვა გრძნობები სრულად იქნება მობილიზებული, ადამიანებს, მანქანებსა და სხვა ობიექტებს შორის მრავალსენსორული ურთიერთქმედების მისაღწევად. ამ საფუძველზე, ასევე უზრუნველყოფილი იქნება სხვადასხვა სახის მომსახურება, რომელიც მჭიდროდ არის დაკავშირებული ადამიანის ცხოვრების ისეთ საჭიროებებთან, როგორცაა დისტანციური დიაგნოსტიკა, დისტანციური ფსიქოლოგიური ჩარევა, დისტანციური ქირურგია და იმერსიული შოპინგი. ვინაიდან სხვადასხვა ტიპის სერვისებს განსხვავებული მოთხოვნილებები აქვთ სხვადასხვა გრძნობისთვის, მრავალსენსორული კომუნიკაცია 5G ქსელებისთვის აყენებს შემდეგ მოთხოვნებს: არა მხოლოდ მაღალსიჩქარიანი გადაცემა დაბალი შეყოვნებით, არამედ ინტელექტუალური აღქმისა და სერვისების ინტელექტუალური მიწოდების შესაძლებლობები სხვადასხვა პრიორიტეტით მრავალფეროვანი პრაქტიკული საჭიროებებისთვის.

მომავლის ინტელექტუალური ტრანსპორტი უაღრესად დინამიკური ინტელექტუალური სისტემაა, რომელიც ჩამოყალიბებულია აღქმის, კომუნიკაციის, გამოთვლითი ტექნოლოგიებისა და სატრანსპორტო ინფრასტრუქტურის შერწყმით. ავტონომიური მართვის ტექნოლოგიების განვითარებით და კოსმოსურ-საჰაერო-სახმელეთო-საზღვაო ინტეგრირებული საკომუნიკაციო ქსელების ამენებით ჩამოყალიბდება პეტეროგენული მრავალდონიანი ინტელექტუალური სატრანსპორტო სისტემა, რომელიც მოიცავს მთელ საზოგადოებას. იგი მოიცავს არა მხოლოდ ტრადიციულ სახმელეთო კომუნიკაციებს სატრანსპორტო საშუალებებსა და ინფრასტრუქტურას შორის, არამედ საჰაერო-სახმელეთო, კოსმოსურ-სახმელეთო და წყალქვეშა მანქანებთან დაკავშირებულ კომუნიკაციებსაც კი. ყველა ტიპის ავტომატიზებულმა მანქანამ უნდა დაგვემოს და გადაწყვიტოს საკუთარი ქმედებები რეალურ დროში ისეთ ინფორმაციასთან ერთად, როგორცაა გზებისა და მიმდებარე მანქანების მდგომარეობა. აღქმის, კომუნიკაციისა და გამოთვლითი ტექნოლოგიების ინტეგრირება სხვადასხვა ტიპის სატრანსპორტო საშუალებასთან, დიდი მონაცემების ეფექტიანი აღქმის, გადაცემისა და ანალიზის მისაღწევად, ასევე ტრანსპორტირების სწრაფი დაგეგმვა და ინტელექტუალური გადაწყვეტილების მიღება ახალი მოთხოვნები მომავალი 5G ქსელებისთვის.

ჭკვიანი წარმოება გულისხმობს გონივრულ და ინტენსიურ წარმოების პროცესს ინტელექტუალური წარმოების სისტემის მეშვეობით, ქსელის მიხედვით განაწილებული სამრეწველო ობიექტების გამოყენებით. ციფრულ ტყუპებზე დაფუძნებულ ინტელექტუალურ წარმოებას შეუძლია დამოუკიდებლად შეაგროვოს და გააანალიზოს წარმოების მონაცემები და განახორციელოს წარმოების პროცესის დასკვნის წარმოდგენა, პროგნოზირება, დამოუკიდებელი დაგეგმვა და გადაწყვეტილების მიღება ისეთი მიდგომების საშუალებით, როგორცაა ციფრული იმიტაცია და ყოვლისმომცველი კომპიუტერული სიმულაცია. ის ასევე უზრუნველყოფს ციფრული მულტიმედისა და AR ტექნოლოგიების გამოყენებას დიზაინისა და წარმოების პროცესის დემონსტრირებისთვის, რათა ხელი შეუწყოს ადამიანსა და მანქანას შორის მაღალი დონის ღრმა თანამშრომლობას. ამ პროცესში, 5G ქსელური ინტელექტის წინაშე მდგარი კრიტიკული გამოწვევებია ინდუსტრიული გარემოს სწრაფი მოცულობითი აღქმა, ინდუსტრიული აღჭურვილობის რეალურ დროში საიმედო კონტროლი და ციფრული და რეალური სივრცის თანამშრომლობის ინტელექტუალური მხარდაჭერა.

ამრიგად, ახალი 5G სერვისების გამოჩენა არა მხოლოდ დიდ მოთხოვნებს უყენებს ტრადიციულ KPI-ებს, როგორცაა შეყოვნება, სიჩქარე და საიმედოობა, არამედ უკიდურესად მოთხოვნადია ქსელის ინტელექტუალურობა. სხვადასხვა სერვისის დამოკიდებულება ქსელურ ინტელექტზე ნაჩვენებია ნახ. 3.1-ზე, საიდანაც ჩანს, რომ სერვისებს სხვადასხვა დონის დამოკიდებულება აქვთ ქსელის ინტელექტზე. ისეთ სერვისებს, როგორცაა დისტანციური ჰოლოგრაფიული ქირურგია, ჰოლოგრაფიული შეხვედრები და სამრეწველო ციფრული ტყუპები, დიდი მოთხოვნა აქვთ ქსელის ინტელექტზე, რადგან ქსელური ინტელექტი ხელს შეუწყობს ამ სერვისების სასურველი დონის შენარჩუნებას ექსტრემალური მოთხოვნებით სხვადასხვა განზომილებაში. ამის საპირისპიროდ, ისეთი სერვისები, როგორცაა დისტანციური ჭკვიანი პენსია (ანუ მოხერხებული გზით პენსიის მიღება) და ავტონომიური მართვა, აჩვენებს შედარებით დაბალ ინტელექტუალურ დამოკიდებულებას, რადგან მათ აქვთ უკიდურესი მოთხოვნები მხოლოდ შეზღუდულ განზომილებაში.



ნახ. 3.1. სხვადასხვა სერვისის დამოკიდებულება ქსელის ინტელექტზე: XR — გაფართოებული რეალობა, რომელიც მოიცავს AR-ს და VR-ს

3.3. ინტელექტუალური მოთხოვნები 5G-ის ექსპლუატაციისა და ტექნიკური მომსახურებისთვის

მომავალი 5G ქსელები უზრუნველყოფენ პერსონალიზებულ და განმსაზღვრელ ქსელურ კავშირებს და სერვისებს დიდი რაოდენობის დინამიკური კვანძების გამოყენებით, ისეთი შესაძლებლობებით, როგორცაა ავტონომიური განლაგება, ავტონომიური თანამშრომლობა, თვითოპტიმიზაცია და თვითგანკურნება. 5G-ისთან შედარებით, მოსალოდნელია, რომ 5G ქსელები განათავსებენ უფრო მეტ კვანძს უფრო მაღალი სიმკვრივით; დეცენტრალიზებული სტრუქტურა ქსელს უფრო შეკუმშულს ხდის; ქსელის ღიაობა (openness) იწვევს ქსელის ფუნქციების განვითარებას პროგრამულ უზრუნველყოფაზე და სერვისებზე ორიენტირებულად; ქსელისა და გამოთვლების ინტეგრაცია რესურსების დაგეგმვას უფრო აართულებს; და ასე შემდეგ. ყველა ეს მახასიათებელი აყენებს უფრო მაღალ ინტელექტუალურ მოთხოვნებს 5G ქსელების ექსპლუატაციისა და ტექნიკური მომსახურებისთვის (O&M).

მოთხოვნა O&M ცოდნაზე:

ქსელის ტრადიციული O&M ორიენტირებულია ქსელის ობიექტებზე. მისი მიზანია უზრუნველყოს ქსელური ობიექტების მაღალი წარმადობა და საიმედოობა, რაც, ძირითადად, კეთდება ხელით, პროფესიული ცოდნის მქონე ადმინისტრატორების მიერ. 5G-ში წამოყენებული მიზნის, „ქსელის ცვლილებები საჭიროებისამებრ, ნებისმიერი სერვისისთვის, რომელიც მომხმარებელს უნდა“, გამო, ქსელის O&M პარადიგმისთვის, რომელიც ეყრდნობა ხელით მუშაობის გამოცდილებას ან სტატიკური მომსახურების სტრატეგიას, რთულია ახალ ტენდენციასთან ადაპტირება. პარადიგმის ცვლილება მოითხოვს ქსელის O&M ცოდნის მნიშვნელოვან გაფართოებას, ანუ, არა მხოლოდ ქსელის O&M-ის პროფესიონალურ ცოდნას, არამედ აპლიკაციის მახასიათებლების, სერვისის განზრახვის, მომხმარებლის ქცევისა და პრეფერენციების, ქსელის მუშაობის გარემოს და ა. შ. ღრმა

გაგებას. ასევე საჭიროა მუდმივად სწავლისა და ახალი ცოდნის დაუფლების უნარი. დომენებში ცოდნის დიდი რაოდენობით დამუშავება, ინტეგრაცია, შენახვა და გამოყენება კრიტიკულ გამოწვევებს უქმნის ინტელექტუალური ქსელის O&M-ს.

ავტონომიური მართვის შემდგომი ინტელექტუალიზაციის მოთხოვნა:

4G ეპოქაში თვითორგანიზებული ქსელი (SON) გახდა მნიშვნელოვანი საშუალება ქსელის O&M-ის ეფექტიანობის გასაუმჯობესებლად. SON ქმნის ჩაკეტილ მარყუჟს მართვის სტრუქტურის აღქმის, დაგეგმვის, შესრულებისა და შეფასებისთვის და ახორციელებს ავტონომიურ მართვას ტიპური გამოყენების შემთხვევებისთვის, როგორცაა თვითკონფიგურაცია, თვითოპტიმიზაცია და თვითგანკურნება. მაგრამ SON-ის ავტონომიური მართვის სქემა ეყრდნობა კარგად შემუშავებულ ალგორითმებს ან მექანიზმებს. ბოლო წლებში, AI-ის სხვადასხვა მეთოდის გამოყენებამ, როგორცაა DL და RL, გააუმჯობესა მონაცემთა მასობრივი დამუშავების სიჩქარე და მიღებული გადაწყვეტილების სიზუსტე SON-ებში, რითაც გააუმჯობესა თვითორგანიზების მართვის შესაძლებლობები. თუმცა, კვლევების უმეტესობა, რომელიც იყენებს AI ტექნოლოგიებს SON-ზე, ჩვეულებრივ, ფოკუსირებულია მონაცემთა ბაზაზე ორიენტირებული ალგორითმის უზრუნველყოფაზე კონკრეტული გამოყენების შემთხვევისთვის. ამ ალგორითმებში გამოყენებული სხვადასხვა სასწავლო მოდელი, როგორც წესი, დამოკიდებულია კარგ დიზაინზე; მათ შეუძლიათ მოამზადონ და დაარეგულირონ საკუთარი პარამეტრები, მაგრამ ვერ გააუმჯობესებენ მოდელის სტრუქტურებს. 6G ქსელს უნდა ჰქონდეს თვითგანვითარების შესაძლებლობები, მათ შორის ქსელის ტოპოლოგიისა და ფუნქციების თვითგანვითარება და ა. შ., რაც უფრო მეტ მოთხოვნას აყენებს მართვის ინტელექტუალიზაციაზე.

მოთხოვნა ქსელის და გამოთვლითი რესურსების ინტელექტუალურ დაგეგმვაზე:

მომავალში, ახალი კომპიუტერული სერვისების მოთხოვნები მნიშვნელოვნად გაიზარდება და ინტეგრირებული გამოთვლითი ქსელის არქიტექტურა გახდება 6G-ის ახალი განვითარებადი ტენდენცია. ინტეგრირებულ არქიტექტურაში ტერმინალური კვანძები, პერიფერიული ღრუბლები და ძირითადი ღრუბლები ქმნიან მრავალდონიანი ქსელის გამოთვლით არქიტექტურას, რომელსაც შეუძლია უზრუნველყოს მდიდარი გამოთვლითი რესურსები სერვისებისთვის. მიუხედავად იმისა, რომ ქსელი ახორციელებს განაწილებული კვანძების ურთიერთდაკავშირებას, მას უნდა შეეძლოს ქსელისა და გამოთვლითი რესურსების ერთობლივი დაგეგმვა, რათა დინამიკურად დაგეგმოს სერვისის ნაკადი ოპტიმიზებულ გზაზე ოპტიმალური გამოთვლითი კვანძის დასამუშავებლად. შესაბამისად, გამოთვლითი კვანძები ყველა დონეზე დინამიკურად უნდა იყოს გამოყენებული, რათა უზასუხოს რეალურ დროში ცვალებად გამოთვლით ამოცანებს ქსელის ინტეგრაციის არქიტექტურაში და უზრუნველყოს ამოცანების დაგეგმვის შესაძლებლობა მოთხოვნისამებრ. ამავდროულად, ჰეტეროგენული ქსელისა და გამოთვლითი რესურსების ერთობლივი ინტელექტუალური დაგეგმვა არის ქსელის O&M ინტელექტუალიზაციის მნიშვნელოვანი გამოვლინება.

მომავალი 6G კომუნიკაციები განიხილავს ყველა ტიპის სცენარს და მომსახურებას. ამრიგად, ენდოგენური ინტელექტი 6G-ის ძირითადი კომპონენტი და 5G კომუნიკაციის ევოლუციის ხელშეწყობის გარდაუვალი ტენდენციაა. ქსელების ენდოგენური ინტელექტის კვლევა თანდათან ვითარდება. ლიტერატურაში შემოთავაზებულია რადიოწვდომის ქსელების ინტელექტუალური სისტემის არქიტექტურა და განხილულია შესაბამისი ფუნდამენტური თეორიები და ძირითადი ტექნოლოგიები; აღწერილია 6G IEN-ის მახასიათებლები და შემადგენლობა, რომელსაც მართავს ხელოვნური ინტელექტი, ასევე ძირითადი ტექნოლოგიები და პოტენციური აპლიკაციები.

ქსელის ენდოგენური ინტელექტის შესახებ არსებული კვლევები შეიძლება დაჯგუფდეს მონაცემებზე ორიენტირებულ მიდგომაზე და სერვისის ფუნქციონირების აპლიკაციაზე, ხოლო ავტონომიური ქსელის მართვის უფრო მაღალი დონე მონაცემთა მართვიდან გადადის ცოდნის მართვამდე. ჯერ კიდევ არსებობს შეზღუდვები IEN-ების არქიტექტურისა და მუშაობის მექანიზმის კვლევაში. ამიტომ, ეს თავი განმარტავს, თუ როგორ იყენებს IEN ცოდნის გრაფს და სხვა AI ტექნოლოგიებს, ქსელის მრავალგანზომილებიანი ცოდნის წარმოსაჩინად, მხარდასაჭერად და გამოსაყენებლად და წარმოგვიდგენს ინტუიციური გამოყენების შემთხვევას.

3.4. ინტელექტუალურ-ანდოგენური ქსელები 6G-ისტვის

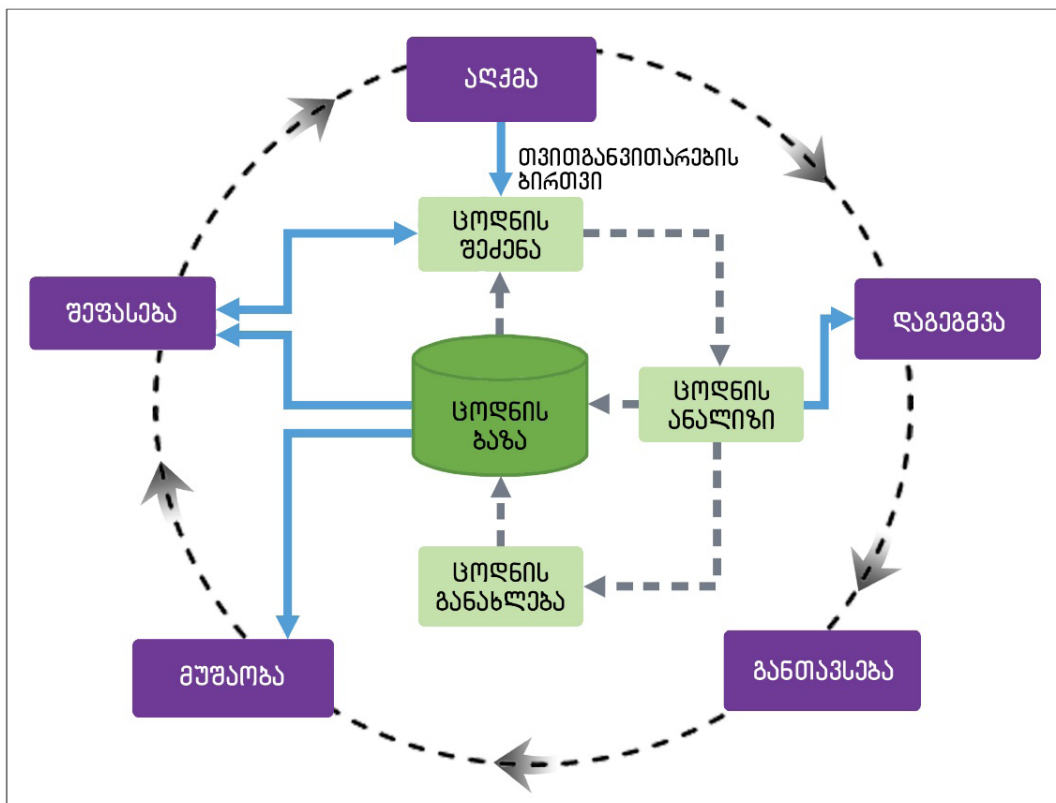
IEN ორიენტირებულია ცოდნაზე. იგი ნერგავს ქსელებში ხელოვნური ინტელექტის ტექნოლოგიას ქსელის მრავალგანზომილებიანი სუბიექტური და ობიექტური ცოდნის დასახასიათებლად, ასაგებად, სწავლის, გამოყენების და განახლებისთვის. ცოდნის გრაფისა და სტერეო აღქმის საფუძველზე, გადაწყვეტილების მიღება შესაძლებელია ქსელის დინამიკური კორექტირებისთვის, ასე რომ, ქსელი შეიძლება შეიცვალოს ისე, როგორც მას სჭირდება, ნებისმიერი ახალი სერვისისთვის, რომელიც ჩვენ გვსურს.

AI-ის თანამედროვე პრინციპებისა და მეთოდოლოგიების გათვალისწინებით, ქსელის სისტემის შინაგან ატრიბუტებთან და სამუშაო მახასიათებლებთან ერთად, IEN-ის განვითარება ნიშნავს ქსელის არქიტექტურისა და სამუშაო გარემოს ისეთი დიზაინის შექმნას, რომელიც ხელს შეუწყობს ქსელის ინტელექტის განვითარებას. IEN, რომელიც შედგება ორფენიანი ჩაკეტილი მარყუჟის სტრუქტურისგან და რომელსაც შეუძლია თვითგანვითარების რეალიზება, ნაჩვენებია ნახ. 3.2-ზე, სადაც შიდა ფენა წარმოადგენს თვითგანვითარებულ ბირთვის (SEC), რომელიც ასევე განიხილება როგორც „ცოდნის ტენი“, რომელიც იღებს ქსელის ცოდნის გრაფს, როგორც ბირთვს. თვითგანვითარების მოთხოვნილებების დასაკმაყოფილებლად, ქსელის ცოდნის გრაფში ცოდნის ფარგლები მოიცავს მომხმარებლის სუბიექტურ მოთხოვნებს და ობიექტურ მიზნებს, ქსელის რესურსების განაწილების სტრატეგიებს, ქსელურ გარემოს (მაგალითად, გეოგრაფიულ, კლიმატურ, საცხოვრებელ და ელექტრომაგნიტურ გარემოს), ქსელის მუშაობის ხარისხის შეფასების წესებს და ა. შ. ქსელის ცოდნა ინახება ქსელის ცოდნის გრაფში სამმაგი ფორმით (ერთეული 1, ურთიერთობა, ერთეული 2) ან (ერთეული, ატრიბუტი, ატრიბუტის მნიშვნელობა) ან N-მიმდევრობებით (დროის, სივრცის ან სხვა განზომილებების დამატებით ამ სამი მიმდევრობაზე). ის შეიძლება დინამიკურად განახლდეს რეალურ გარემოში ცვლილებების და თითოეული ქსელის ოპტიმიზაციის ეფექტიანობის შეფასების მიხედვით.

ცოდნის ეფექტიანი დაგროვებისთვის SEC ძირითადად შედგება სამი ეტაპისაგან.

ცოდნის შექმნა:

ეს არის ინტელექტუალური ანალიზი და ცოდნის შექმნა ქსელის რეალური მუშაობისა და შეფასების მონაცემებიდან. ცოდნის შექმნა შესაძლებელია საექსპერტო სისტემის ან ავტომატური საშუალებების გამოყენებით



ნახ. 3.2. IEN-ის ორფენიანი ჩაკეტილი მარყუჟი

სპეციალური ტექნიკის საშუალებით, რომელიც მოიცავს ლექსიკონის მოპოვებას, ერთეულების აღმოჩენას, ურთიერთობის მოპოვებას, მოვლენის მოპოვებას, წესების ამოღებას და ა. შ. მიღებული ქსელური ინფორმაცია და წესები შემდეგ გარდაიქმნება ცოდნის სტრუქტურირებულ ინფორმაციად და ინახება ქსელის ცოდნის გრაფში.

ცოდნის ანალიზი:

არსებული ცოდნის საფუძველზე, ძირითადი მონაცემები (მაგალითად, ინფორმაცია საბაზო სადგურისა და ფიჭის შესახებ) ქსელის არეალის შესახებ, სადაც პრობლემა გაჩენილი, შეიძლება მიღებულ იქნეს მომხმარებლის ინფორმაციის ან სხვა ინფორმაციის მიხედვით. შემდეგ, პრობლემურ ქსელთან დაკავშირებული ცოდნის მონაცემები მიიღება ქსელის ცოდნის გრაფიდან. ქსელის პრობლემა ავტონომიურად ფასდება ქსელის ცოდნის გრაფში „ქსელის პრობლემის დიაგნოსტიკის ცოდნის“ საფუძველზე. ქსელის ოპტიმიზაციის ფუნქციისთვის რეკომენდებულია შესაბამისი ოპტიმიზაციის ალგორითმები, როგორცაა ღრმა განმტკიცებული სწავლების (DRL) მოდელები და შესაბამისი ჰიპერპარამეტრები, ქსელის ოპტიმიზაციის ცოდნის საფუძველზე.

ცოდნის განახლება:

ცოდნის დამუშავების გზით ხდება ცოდნის ინტეგრაცია და განახლება, რაც ხელს უწყობს ცოდნის ეფექტიანი წრის ჩამოყალიბებას. კონკრეტულად, ცოდნის სხვადასხვა შემდგომი დამუშავება შეიძლება განხორციელდეს ისეთი ML მეთოდებით, როგორცაა ცოდნის შერწყმა, ცოდნის დასრულება და ცოდნის დასაბუთება, ცოდნის გრაფის მუდმივად გასაუმჯობესებლად.

ქსელის ცოდნის გრაფზე დაფუძნებული ცოდნის შიდა ფენის ჩაკეტილი მარყუჟი ახორციელებს ქსელის ცოდნის შექმნას, ანალიზს და განახლებას და შეუძლია განუწყვეტლივ დააგროვოს ქსელის O&M გამოცდილება. SEC უზრუნველყოფს ცოდნისა და ინტელექტის საფუძველს ქსელის თვითგანვითარებისთვის.

გარე ფენა შედგება მთელი რიგი ენდოგენური, ინტელექტზე ორიენტირებული ჩაკეტილი მარყუჟის ქსელის მუშაობისა და ტექნიკური ფუნქციებისგან, რომლებიც ახორციელებენ ქსელის ავტონომიურ მუშაობას და ექსპლუატაციას. გარე ფენა უზრუნველყოფს მონაცემებისა და გამოცდილების შეყვანას SEC-ში და შეუძლია განახორციელოს ქსელის დაგეგმვა, განლაგება, ექსპლუატაცია და კონფიგურაცია თვითგანვითარებადი ბირთვის მიერ მოწოდებული ცოდნის შესაბამისად, ასევე გააცნობიეროს ქსელის O&M თვითგანვითარების სტრატეგია. ეს ფენა მოიცავს ხუთ ეტაპს: აღქმას, დაგეგმვას, განლაგებას, მუშაობასა და შეფასებას. თითოეული ეტაპის ფუნქციები შემდეგია.

აღქმა:

მომხმარებლების, სერვისების, ქსელების, გარემოს და ა. შ. სტერეო აღქმა ხორციელდება იმისათვის, რომ გავიგოთ მომხმარებელთა მიერ ქსელის გამოყენების ჩვევები; ასევე გავიგოთ ქსელის მომხმარებლების სემანტიკური საჭიროებები და განვსაზღვროთ განსხვავებები ქსელებში, სერვისებსა და სამუშაო გარემოში, რათა შევქმნათ საინფორმაციო საფუძველი შემდგომი ეტაპებისთვის.

მომხმარებელთა ობიექტური განზრახვა გაანალიზებულია მომხმარებლების სუბიექტური საჩივრების შემადგენელი ინფორმაციის მიხედვით, რათა ქსელებმა შეძლოს მომხმარებლის განზრახვის ამოცნობა და გაგება, რამაც შეიძლება ქსელის ოპტიმიზაციის სქემა უფრო ყოვლისმომცველი გახადოს. ამ ეტაპზე მიღებული სხვადასხვა ინფორმაცია გამოყენებული იქნება SEC-ში ცოდნის შექმნის ფუნქციის ერთ-ერთ შემავალ მონაცემად.

დაგეგმვა:

აღქმული ინფორმაციის და ასევე SEC-ის ცოდნის მხარდაჭერით, ამ ეტაპზე ავტომატურად იქმნება ქსელის ტოპოლოგიის, ქსელის კონფიგურაციის, ფენების ჯვარედინი დომენური თანამშრომლობის და ქსელის მართვის გონივრული სქემები.

განლაგება:

ქსელის ფუნქციის ვირტუალიზაციის (NFV), პროგრამული უზრუნველყოფით განსაზღვრული ქსელის (SDN), მრავალჯერადი წვდომაზე დაფუძნებული პერიფერიული გამოთვლების, ქსელის ფენებად დაყოფის, ღრუბლოვანი გამოთვლების გამოყენებით და სხვა ტექნოლოგიებზე დაყრდნობით, IEN, დაგეგმილი სქემის მიხედვით, ავტომატურად წარმოქმნის სასურველი ქსელის ტოპოლოგიას, ავრცელებს ქსელის ფუნქციებს, ანაწილებს ქსელის რესურსებს, უზრუნველყოფს მართვისა და კონტროლის შესაძლებლობებს და ა. შ.

მუშაობა:

ქსელი ავტონომიურად მუშაობს ფუნქციური ლოგიკის მიხედვით. სერვისები მიწოდებულია ავტონომიურად, ბიზნეს ლოგიკის შესაბამისად. ქსელის მომსახურებისა და მართვის სისტემები მუშაობს დამოუკიდებლად, მართვისა და კონტროლის სტრატეგიების მიხედვით.

შეფასება:

SEC-ის მიერ განსაზღვრული შეფასების ინდიკატორებისა და შეფასების სქემების საფუძველზე ფასდება ქსელის მუშაობის სტატუსი და ქსელის მართვის სქემების ეფექტიანობა. შეფასების შედეგები შეიძლება გამოყენებულ იქნეს როგორც აღქმის ინფორმაციის ერთ-ერთი ტიპი, რათა უზრუნველყოს მონაცემთა საფუძველი ქსელის მუშაობის სქემისა და მართვის სქემის უწყვეტი განვითარებისთვის. ამავდროულად, შედეგები უკუკავშირით უნდა დაუბრუნდეს SEC-ს, რათა ხელი შეეწყოს SEC-ის გამოცდილების და ცოდნის დაგროვებასა და განახლებას.

თეორიული საკითხები და ძირითადი ტექნოლოგიები, რომლებიც უნდა განიხილებოდეს IEN-ის ასაგებად, მოიცავს შემდეგ ასპექტებს.

ცოდნის გრაფის მახასიათებლები და კონსტრუირების ტექნოლოგიები 6G ქსელებისთვის:

ეს ეხება 6G ქსელებისთვის მრავალგანზომილებიანი სუბიექტური და ობიექტური ცოდნის დახასიათების, აგების, შექმნის, მართვისა და განვითარების მექანიზმებს, მათ შორის მონაცემთა და ურთიერთქმედების მოდელების ინტელექტუალურ სტრუქტურირებას, პოტენციური მოდელების ავტომატურად მოპოვებას და დახვეწას, ცოდნის შერწყმას და დასაბუთებას, და ა. შ. ყოველივე აღნიშნული წარმოადგენს IEN-ების დანერგვის თეორიულ საფუძველს.

ქსელის თვითგანვითარების მექანიზმი:

ქსელის თვითგანვითარების მექანიზმს ახორციელებს შიდა SEC, რომელიც დაფუძნებულია ცოდნის გრაფზე და მართვის გარე ჩაკეტილ მარყუჟზე ქსელის მუშაობის მახასიათებლების შესაბამისად. მისი ძირითადი ტექნოლოგიები მოიცავს მიზნობრიობაზე ორიენტირებული ქსელის გამოცდილების მოპოვებას, რეორგანიზაციას და დედუქციის მეთოდებს, ქსელის კომპლექსური ევოლუციის დინამიკურ მოდელებს, მოქნილი ქსელის არქიტექტურას, რომელიც მხარს უჭერს დინამიკური რეორგანიზაციის ფუნქციებს და სერვისებს, და მტყუნების მიმართ მდგრად უსაფრთხოების მექანიზმებს ავტონომიური ინტერპრეტაციისთვის.

მოცულობითი აღქმის ტექნოლოგია პოლოგრადიული ქსელისთვის:

ეს მოიცავს ცოდნის გრაფზე დაფუძნებული მოცულობითი აღქმის ინფორმაციის აგების მეთოდებს, აგრეთვე ყოვლისმომცველი, ბუსტი და დროული მრავალგანზომილებიანი ინფორმაციის მოპოვებისა და ამოღების მექანიზმებს, რაც კარგ საინფორმაციო საფუძველს შექმნის ქსელისთვის გადაწყვეტილების მიღების პროცესში.

ქსელის რესურსების დაგეგმვის ელასტიკური მექანიზმი ცოდნის გრაფისა და მოცულობითი აღქმის საფუძველზე:

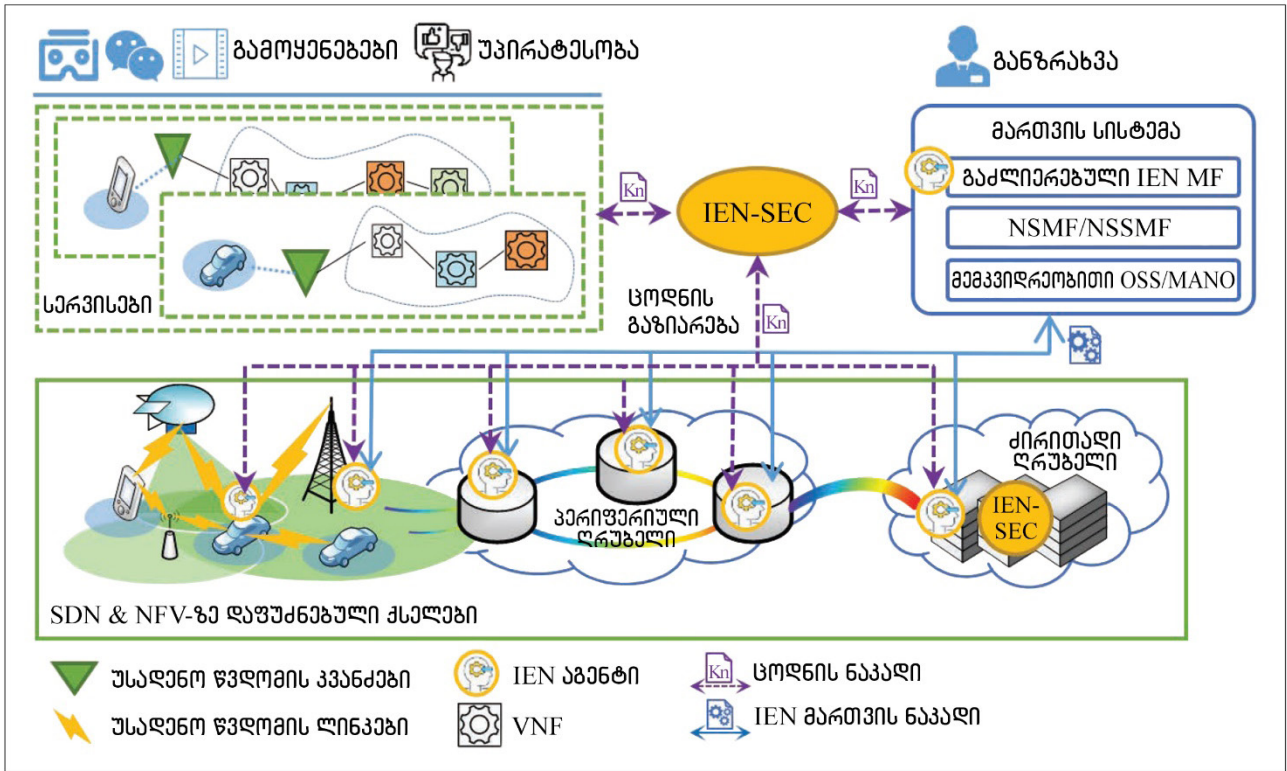
ეს ეხება სხვადასხვა ტიპის რესურსების ელასტიკურ დაგეგმვას, როგორცაა გამოთვლა, მარშრუტიზაცია, ქეშირება (caching) და ა. შ., ასევე რესურსების მძლავრ პროაქტიულ განაწილებას და ჯვარედინი ერთობლივი თანამშრომლობით ოპტიმიზაციას, რომლებიც ცოდნის გრაფსა და მრავალგანზომილებიან აღქმასთან დაკავშირებულ ინფორმაციაზეა დაფუძნებული.

IEN-ის არქიტექტურა შედგება შემდეგი ნაწილებისგან: IEN-SEC, ქსელის მართვის გაძლიერებული ფუნქციები, SDN და NFV მხარდაჭერი ქსელის ინფრასტრუქტურა და სერვისის ფუნქციების ჯაჭვზე დაფუძნებული ინტელექტუალური ქსელის სერვისები. ნაწილებს შორის ურთიერთობა ნაჩვენებია ნახ. 3.3-ზე, სადაც VNF აღნიშნავს ვირტუალიზებული ქსელის ფუნქციას.

IEN-ის თვითგანვითარების ბირთვი:

IEN-SEC ენდოგენური ინტელექტის წარმომქმნელი ბირთვია. მისი მთავარი ფუნქციაა ქსელის მუშაობის გამოცდილების ყველა ასპექტის დაგროვება და ქსელის O&M სტრატეგიების გაუმჯობესება და განახლება ორფენიანი ჩაკეტილი მარყუჟის ფორმით. იგი გავლენას მოახდენს ქსელის ინტელექტზე სამი ასპექტით:

- პირველი, ეს არის ქსელური სერვისების ინტელექტუალიზაცია. ის აღიქვამს მომხმარებელთა სერვისებს და გააანალიზებს მომხმარებელთა კმაყოფილებას სერვისებით. სერვისებისა და მომხმარებლის პრეფერენციების ბუსტი ანალიზის მეშვეობით, IEN მომხმარებლებს უფრო ჭკვიანურად მიაწვდის შესაბამის ქსელურ სერვისებს.
- მეორე, ეს არის ქსელის რესურსების მართვის ინტელექტუალიზაცია. ის აღიქვამს და შეაფასებს რესურსების



ნახ.3.3. IEN-ის არქიტექტურა

მარტვის არსებული სტრატეგიების გავლენას ქსელის მუშაობაზე, გააგრძელებს ამ სტრატეგიების გაუმჯობესებას შეფასების შედეგებზე დაყრდნობით და გაუმჯობესებულ სტრატეგიებს გაავრცელებს IEN-ის სხვადასხვა აგენტზე.

- მესამე, ეს არის ქსელური ინფრასტრუქტურის მარტვის ინტელექტუალიზაცია. ქსელური ინფრასტრუქტურის პარამეტრების კორექტირებასთან შედარებით, IEN-SEC-ს შეუძლია გამოიყენოს დაგროვილი სავსე გამოცდილება და ქსელის ცოდნის გრაფი, ქსელის ინფრასტრუქტურის განახლების, დამატებისა და სხვა გაუმჯობესების სტრატეგიების უზრუნველსაყოფად.

ქსელის მარტვის გაძლიერებული სისტემა:

ქსელის მარტვის გაძლიერებული სისტემა, ძირითადად, პასუხისმგებელია IEN-SEC-ის მიერ მოწოდებული გაუმჯობესებული სტრატეგიების ინტერპრეტაციაზე და განხორციელებაზე მარტვის კონკრეტული ფუნქციების განსახორციელებლად. იგი მოიცავს ქსელის მარტვის ტრადიციული სისტემის (NMS) ფუნქციებს (NMSF) და ასევე მარტვის ახალ ფუნქციებს, რომელთა შორისაა:

- NSMF (გამჭოლი სახის ფუნქციის დანაწილების მარტვისთვის) და NSSMF (ქსელის დაყოფისას მიღებული ქვექსელის (ქვედომენის) მარტვისთვის). ამ უკანასკნელის სახეობებია, ძირითადი ქსელის NSMF, სატრანსპორტო ქსელის NSSMF და ა. შ.
- NFV მენეჯმენტი და ექსპლუატაცია (MANO), რომელიც განკუთვნილია ქსელის ფუნქციების ჯაჭვის ექსპლუატაციისთვის, NFV-ის და ინფრასტრუქტურის მარტვისათვის.
- SDN კონტროლერი ბოლოების დამაკავშირებელი (გამჭოლი) მონაცემთა ნაკადის კონტროლისა და მარტვისთვის.
- IEN მენეჯმენტის გაძლიერებული ფუნქცია არის ახლად დანერგილი მენეჯმენტი IEN-ისთვის IEN აგენტების მარტვისათვის, რომლის ფუნქციები მოიცავს IEN აგენტების განსაზღვრას, აწყობას, განლაგებას და კონფიგურირებას, ასევე ცოდნის მონაცემთა ურთიერთქმედებას და სტრატეგიის სინქრონიზაციას IEN-SEC-სა და IEN აგენტებს შორის.

SDN/NFV–ით მხარდაჭერილი ქსელის ინფრასტრუქტურა:

აღნიშნული ინფრასტრუქტურა უზრუნველყოფს გამოთვლების, შენახვისა და საქსელო ფუნქციებს. ის ნერგავს ქსელის ფუნქციებს პროგრამული უზრუნველყოფის განსაზღვრისა და ვირტუალიზაციის საშუალებით. მას შეუძლია უპასუხოს ქსელის მართვისა და კონტროლის ბრძანებებს, რომლებიც გაცემულია ქსელის მართვის მოწინავე სისტემის მიერ და განახორციელოს ინფრასტრუქტურის რესურსების მოქნილი განლაგება და რეგულირება. ქსელის ინფრასტრუქტურაში, IEN-ის მხარდასაჭერად, დამატებულია IEN აგენტები, რომლებსაც შეუძლიათ მიიღონ სხვადასხვა სტრატეგია და ცოდნა, რომლებიც განაწილებულია IEN-SEC-ის მიერ და აქვთ გარკვეული ინტელექტუალური გამოთვლითი შესაძლებლობები გამოთვლითი ინტელექტის განაწილების განსახორციელებლად.

ინტელექტუალური ქსელის სერვისი, რომელიც დაფუძნებულია მომსახურების ფუნქციების ჯაჭვზე:

სერვისის ფუნქციების ჯაჭვზე დაფუძნებული ინტელექტუალური ქსელის სერვისი მომხმარებლებს სთავაზობს მოთხოვნილ სერვისებს. ერთი მხრივ, IEN ჰკვიანურად აწყობს ქსელის რესურსებს დანაწილებული ფორმით, მიკროსერვისის შესაძლებლობებზე დაყრდნობით, რაც უზრუნველყოფს QoS კონტროლს მომხმარებლებისთვის და სერვისებისთვის. მეორე მხრივ, IEN აღწერს და ზომავს მომხმარებლის პრეფერენციებს და აფასებს მომხმარებლის გამოცდილებას რეალურ დროში, მოცულობითი აღქმის შესაძლებლობებზე დაყრდნობით, რათა უზრუნველყოს მომხმარებელზე ორიენტირებული უკეთესი ქსელური სერვისები.

3.5. IEN გამოყენების შემთხვევა: გამტარუნარიანობის გაუმჯობესება საჰაერო საბაზო სადგურების მეშვეობით

IEN-ს შეუძლია გაიგოს მომხმარებლების სერვისის საჭიროებები და მაღალი დონის ქსელის მიზნები, ისწავლოს და დააგროვოს გამოცდილება ცოდნის ბაზის გასაუმჯობესებლად და დამოუკიდებლად და ჰკვიანურად აირჩიოს ქსელის კორექტირების სტრატეგიები, მათ შორის ქსელის პარამეტრების ოპტიმიზაცია, ქსელის რესურსების დაგეგმვა, ქსელში მტყუნებების გამოვლენა, თვითგანკურნება, და ა. შ. IEN აუმჯობესებს ორფენიან ჩაკეტილ მარყუჟს მართვის, კონტროლისა და სტრატეგიის მეშვეობით, რათა უზრუნველყოს ის პირობა, რომ ქსელის მუშაობის სტატუსი კვლავ დააკმაყოფილებს სერვისის საჭიროებებს. ქვემოთ წარმოდგენილია საჰაერო საბაზო სადგურზე დაფუძნებული ქსელის გამტარუნარიანობის გაუმჯობესების სცენარი, როგორც მაგალითი IEN-ის მუშაობის მექანიზმის წარმოსაჩენად.

აღვნიშნავთ, რომ IEN-SEC-მა უნდა დაასრულოს ცოდნის ბაზის კონსტრუქცია და იგი ძირითადად მოიცავს შემდეგ პროცესებს.

ცოდნის შექმნა:

აქ იგულისხმება რომ მიიღება ინფორმაცია საჰაერო საბაზო სადგურის ქსელური რესურსების შესახებ, გარემოს შესახებ, მომხმარებლის გამოცდილების შესახებ და ქსელის მუშაობის ხარისხის შესახებ. რესურსული ინფორმაცია მოიცავს ტერმინალებს, რომლებიც დაკავშირებულია საჰაერო საბაზო სადგურის ქსელთან, ქსელის ელემენტებთან, საკომუნიკაციო ლინკებთან და მათ ურთიერთკავშირთან. ქსელის ყველა სახის ინფორმაცია გარდაიქმნება სამეულებად ან N-მიმდევრობებად და ინახება ცოდნის გრაფში. ამ შემთხვევაში ქსელის გამტარუნარიანობის პრობლემების დიაგნოსტიკის წესები, საჰაერო ბაზის სადგურის განლაგება და ა. შ., შეიძლება შემუშავდეს გადაწყვეტილების ხეების სახით და შენახულ იქნეს ცოდნის გრაფში.

ცოდნის ანალიზი:

გამოვლენილი ქსელის გამტარუნარიანობის პრობლემისა და საჰაერო საბაზო სადგურის დაგეგმვის სტრატეგიის ცოდნის საფუძველზე, აღქმული ქსელისა და მომხმარებლის გამოცდილების ინფორმაციები გაანალიზებულია სცენის ცოდნის გრაფში, რათა დადგინდეს არის თუ არა ცოდნის განახლების საჭიროება.

ცოდნის განახლება:

ქსელის მუშაობის მდგომარეობის შეფასების შედეგების საფუძველზე, გარე ფენა იყენებს ML მეთოდებს ცოდნის ეფექტის შესაფასებლად. ამავდროულად, მსჯელობის ისეთი მეთოდები, როგორებიცაა გრაფის სტრუქტურა, წესების სწავლება და RL, გამოიყენება სხვადასხვა ცოდნას შორის პოტენციური კავშირების ავტომატურად აღმოსაჩენად და ცოდნის გრაფში არსებული ცოდნის მუდმივად განახლების მიზნით.

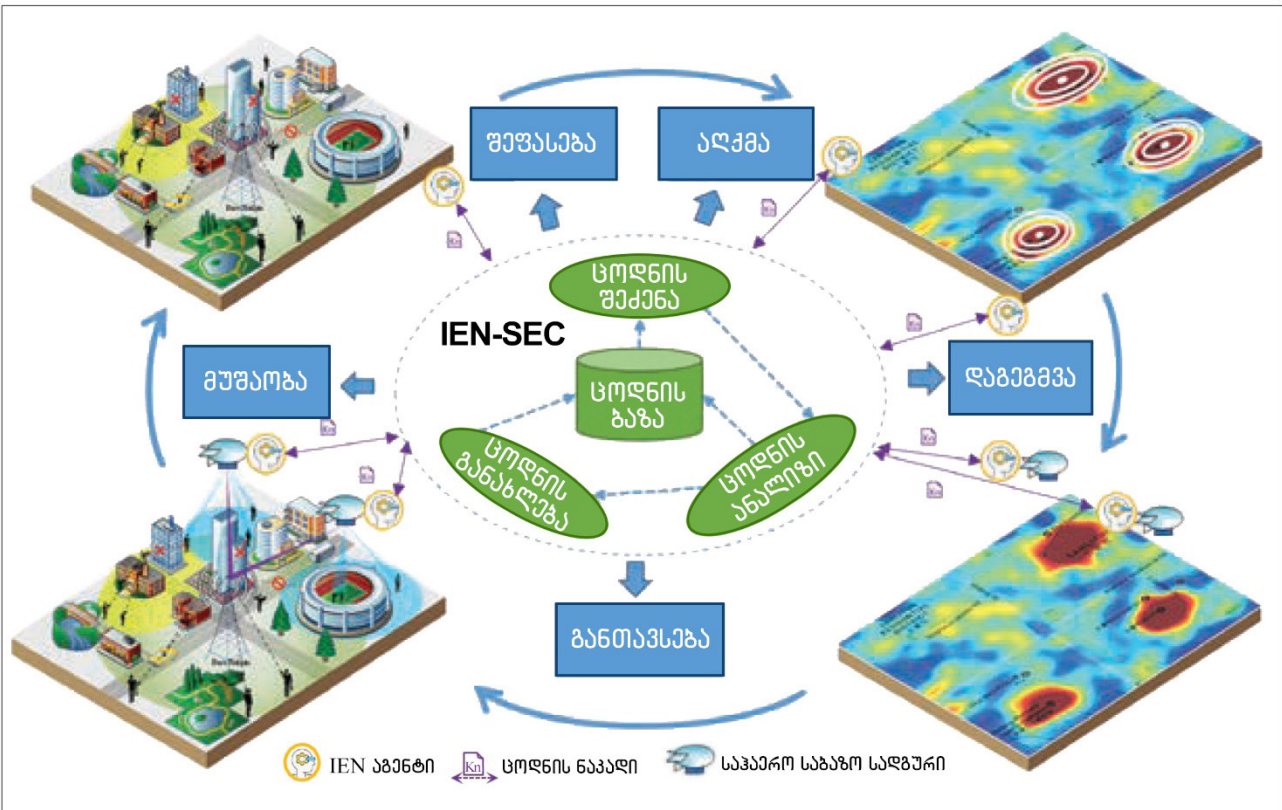
ამ გამოყენების შემთხვევაში, IEN-SEC დამოუკიდებლად წყვეტს თუ სად განათავსოს IEN აგენტები და რა ცოდნა გააგრძელოს IEN აგენტებს შორის, ქსელის ცოდნის გრაფიდან და ქსელის გარემოდან გამომდინარე. ის დამოუკიდებლად განსაზღვრავს გამტარუნარიანობის გაზრდის საჭიროებებს ქსელის სტერეო აღქმის მიხედვით და დამოუკიდებლად ახორციელებს საჭირო საბაზო სადგურების განლაგებასა და რესურსების დაგეგმვას ქსელის გამტარუნარიანობის გაზრდის მისაღწევად. როგორც ნახ. 3.4-ზე ნაჩვენებია, IEN-ის მოქმედი მექანიზმი შეიძლება აღიწეროს შემდეგნაირად.

ქსელის სტერეო აღქმა:

ქსელში IEN აგენტებს ქსელის და გარემოს მოცულობითი აღქმა სხვადასხვა მეთოდით შეუძლიათ, რომელთა შორისაა: გარემოს ცვლილების, ქსელის კონფიგურაციის პარამეტრების, ქსელის მუშაობის ინდიკატორების, ქსელის დატვირთვის ცვლილების, მომხმარებლის გამოცდილების აღქმა და ა. შ. IEN აგენტები ანალიზებენ ქსელის პრობლემებს მოცულობითი აღქმის ინფორმაციისა და შესაბამისი სტრატეგიების საფუძველზე ცოდნის ბაზიდან. როდესაც არსებობს გამტარუნარიანობის პრობლემა, ცხელი წერტილის (hotspot) იდენტიფიკაციის ალგორითმები შეიძლება გამოყენებულ იქნეს ტრაფიკის ცხელი წერტილების გამოსავლენად და საჭირო საბაზო სადგურის განლაგების ეტაპზე გადასასვლელად.

საჭირო საბაზო სადგურების განლაგება:

IEN აგენტი იყენებს ცხელი წერტილების ანალიზის ალგორითმებს ცხელი წერტილების არელებისა და მათი მხასიათებლების გასაანალიზებლად. ამის გათვალისწინებით, იგი აცალიბებს საჭირო საბაზო სადგურების განლაგების სქემას, ცოდნის ბაზაში არსებული საჭირო საბაზო სადგურების დაგეგმვის სტრატეგიის ცოდნის საფუძველზე. სქემა მოიცავს განლაგების ნომერს, მდებარეობას, სიმაღლესა და სხვის შესაბამის ორიენტაციას. IEN აგენტი უგზავნის ინფორმაციას შესაბამის საჭირო საბაზო სადგურებს, ხოლო საჭირო საბაზო სადგურები გადაადგილდებიან დანიშნულ ადგილებში განლაგების სქემის მიხედვით.



ნახ. 3.4. IEN-ის გამოყენების შემთხვევა, სადაც ნაჩვენებია ქსელის გამტარუნარიანობის ბრდა საჭირო საბაზო სადგურების მემეობით

გამტარუნარიანობის გაზრდის სქემის მუშაობა:

მას შემდეგ, რაც საჭირო საბაზო სადგური მიაღწევს დანიშნულ განლაგების ადგილს, ის აერთიანებს ინფორმაციას რეალური ქსელის გარემოსა და მომხმარებლის განაწილების შესახებ ამ ადგილას და მუდმივად ოპტიმიზაციას უკეთებს მის პოზიციას ოპტიმიზაციის ალგორითმების გამოყენებით. ასევე, ის მუდმივად ახდენს რესურსების განაწილების ოპტიმიზაციას, მომხმარებლის წვდომის შერჩევას, სხივის პარამეტრების განსაზღვრას და ა. შ., გამტარუნარიანობის გაუმჯობესების საუკეთესო ეფექტის მისაღწევად.

გამტარუნარიანობის გაუმჯობესების ეფექტის შეფასება:

IEN აგენტები აგრძელებენ ქსელისა და გარემოს აღქმას სხვადასხვა მეთოდით. ქსელის შეფასების ინდიკატორების, შეფასების სტანდარტებისა და ცოდნის ბაზაში არსებული შეფასების მეთოდების ცოდნის საფუძველზე, IEN აგენტები შემუშავებენ შეფასების გეგმას და აფასებენ შესაძლებლობების გაზრდის ეფექტს. შეფასების შედეგების მიხედვით, აგენტები ასევე აფასებენ ოპტიმიზაციის სქემის კორექტირების აუცილებლობას. ისინი განაგრძობენ იტერაციებს მანამ, სანამ აღარ იქნება საჭირო ოპტიმიზაციის სქემის ხელახალი კორექტირება. შემდეგ, ქსელის თითოეული ერთეულის შესაბამისი მდგომარეობის მონაცემები განახლდება და ინახება ცოდნის ბაზაში ისევე, როგორც პროცესების ისეთი შუალედური მონაცემები, როგორებიცაა მოცულობითი აღქმა, პრობლემის ანალიზი, სქემის შედგენა, სქემის მოქმედება და შეფასების შედეგი.

თითოეულ ეტაპზე, IEN-ის თითოეული აგენტის მიერ განხორციელებული სტრატეგია შეიძლება იყოს IEN-SEC-ის მიერ გავრცელებული იტერაციული სტრატეგია მის ადგილობრივ ცოდნასთან და გამოცდილებასთან ერთად. IEN აგენტს ასევე შეუძლია ატვირთოს ადგილობრივი ცოდნა ან გამოცდილება IEN-SEC-ში, რათა ხელი შეუწყოს სტრატეგიის განახლებას. IEN-SEC განახლებს ცოდნის ბაზას ფაქტობრივი სტრატეგიის განხორციელების ეფექტიანობის მიხედვით და განახლებულ სტრატეგიას მიაწვდის თითოეულ IEN აგენტს. თითოეული IEN აგენტის სამუშაო სტრატეგიების იტერაციული განახლებით, ქსელი შეიძლება ავტონომიურად განვითარდეს.

3.6. მომავალი სამუშაოების ანალიზი

მიუხედავად იმისა, რომ IEN პერსპექტიული გადაწყვეტაა, არსებობს მთელი რიგი პრობლემები, რომლებიც შემდგომ შესწავლას საჭიროებენ. ქვემოთ ჩვენ განვიხილავთ ზოგიერთ მათგანს.

1. ქსელის ცოდნის გრაფის ხარისხი პირდაპირ გავლენას ახდენს სტერეო აღქმის სიზუსტეზე, დინამიკურ კორექტირებასა და გადაწყვეტილების მიღებაზე. მიუხედავად იმისა, რომ მცირე მასშტაბის ცოდნის გრაფები ქსელის გამოყენების გარკვეული შემთხვევებისთვის შეიძლება ხელით აიგოს, სამომავლო სამუშაო უფრო რთულ გზას მოიცავს. პირველ რიგში, პრაქტიკული მიდგომა უნდა იყოს შემუშავებული მუდმივად აღმოჩენილი ცოდნის უწყვეტი ინტეგრაციისთვის არსებულ ცოდნის გრაფში. მეორეც, ცოდნის გრაფის განვითარების უკუკავშირის მექანიზმი უნდა გაუმჯობესდეს, რათა IEN-ს შეეძლოს გაიგოს, რომელი ქმედებები იწვევს ყველაზე ხშირად ინციდენტის მოგვარებას. გარდა ამისა, ინტელექტის სრული ენდოგენურობის მისაღწევად, აუცილებელია IEN-ში ძირითადი ელემენტებისა და ურთიერთობების ავტომატური აღმოჩენისა და ასახვის პრობლემის გადაჭრა.
2. მრავალგანზომილებიანი ინფორმაციის ყოვლისმომცველი, ზუსტი და დროული აღქმის, მოპოვებისა და პროგნოზირების მექანიზმები უზრუნველყოფს კარგ საინფორმაციო საფუძველს ქსელური გადაწყვეტილების მისაღებად. თუმცა, როგორ მივაღწიოთ კომპრომისს, ერთი მხრივ მოცულობითი აღქმის მასშტაბებსა და დროულობას, და მეორე მხრივ, მოცულობითი აღქმიდან მიღებული გლობალური ინფორმაციის ეფექტიანად გამოყენებას შორის, მომავალი კვლევის ერთ-ერთი მიმართულებაა.
3. ქსელის თვითგანვითარების ბირთვის ფორმირებისთვის აუცილებელია ქსელის კომპლექსური განვითარების დინამიკური მოდელის და განზრახვაზე ორიენტირებული ქსელის გამოცდილების ამოღების, რეკომბინაციისა და დედუქციის მეთოდების შესწავლა. ასევე, საჭიროა შესწავლილ იქნეს მოქნილი ქსელის არქიტექტურა, რომელიც მხარს უჭერს თვითგანვითარებას და მტყუნებების მიმართ მდგრად უსაფრთხოების მექანიზმს, რომელიც თავის მხრივ იყენებს თვითდედუქციას. ზემოთ წარმოდგენილი კვლევის მეშვეობით შეიძლება გაუმჯობესდეს ქსელის თვითგანვითარების ეფექტიანობა და საიმედოობა. ეს არის მნიშვნელოვანი წინაპირობა პრაქტიკული ხელმისაწვდომობისთვის.

4. IEN აგებული იქნება AI-ის მრავალი გამოთვლითი ფუნქციის ბაზაზე, რომელიც მუშაობს IEN აგენტებზე. ადრე საკითხი, თუ სად უნდა განთავსებულიყო AI-ის ფუნქციები, მკაცრად არ განიხილებოდა, რადგან ქსელებში არ იყო ბევრი ასეთი აგენტი. თუმცა, IEN-ის შემთხვევაში, მიუხედავად იმისა, რომ AI-ის სხვადასხვა ფუნქციას აქვს განსხვავებული პრიორიტეტები და გამოთვლითი დროის შეზღუდვები, გამოთვლითი და ქსელური რესურსების ეფექტიანი გამოყენება ხდება მნიშვნელოვანი ამოცანა. ეს ფაქტორები IEN აგენტებისა და AI-ის ფუნქციების განლაგებას თვალსაჩინო და საინტერესო პრობლემად აქცევს.

3.7. მესამე თავის დასკვნა

ინტელექტუალურ-ენდოგენური ქსელები მნიშვნელოვანია 6G ქსელების სერვისის შესაძლებლობების გაუმჯობესებაში. ეს თავი გვთავაზობს IEN-ების კონცეფციას 6G ქსელის ინტელექტის მოთხოვნების გაცნობით, ასევე მისი ფუნქციონირების და ექსპლუატაციის მიმართულებით. ამ თავში ასევე წარმოდგენილია IEN-ების არქიტექტურა, ძირითადი პრინციპები, ტექნოლოგიები და განხილულია ასეთი ქსელის ზოგადი სამუშაო პროცესი, საჭირო საბაზო სადგურების საშუალებით გამტარუნარიანობის გაზრდის ინტუიციური გამოყენების შემთხვევაში.

თავი 4 ღრმა სწავლებაზე ორიენტირებული და ბუფარზე დაფუძნებული კოოპერატიული ქსელები 5G/6G-ისთვის

4.1. შესავალი

5G/6G ხელვა და BACN-ის საფუძვლები:

5G და 6G ქსელები უფრო მეტია, ვიდრე 5G ქსელის უბრალო გაუმჯობესება ან გაფართოება; ისინი წარმოგვიდგენენ პარადიგმის გასაოცარ ძვრებს. მობილური ტრაფიკის ზრდა, პირველ რიგში, გამოწვეულია ახალი აპლიკაციების სწრაფი გავრცელებით მობილურ მოწყობილობებზე, მათ შორისაა IoT, ელექტრონული ჯანდაცვა, „სატრანსპორტო საშუალება – ყველაფერთან“ (V2X) და M2M კომუნიკაციები, ასევე ვირტუალური/ მომატებული რეალობა. ეს აპლიკაციები მოითხოვენ ქსელის გამტარუნარიანობის მნიშვნელოვან ზრდას, ქსელში მკაცრ შეყვანებას და ქსელის გაუმჯობესებულ საიმედოობას. შედეგად, გარდა საკომუნიკაციო კვანძების ურთიერთდაკავშირებისა, 6G ხელს შეუწყობს საყოველთაოდ კავშირის მიღწევას, უზრუნველყოფს მომსახურების მაღალ ხარისხს (QoS) და ინტელექტუალურ შესაძლებლობებს.

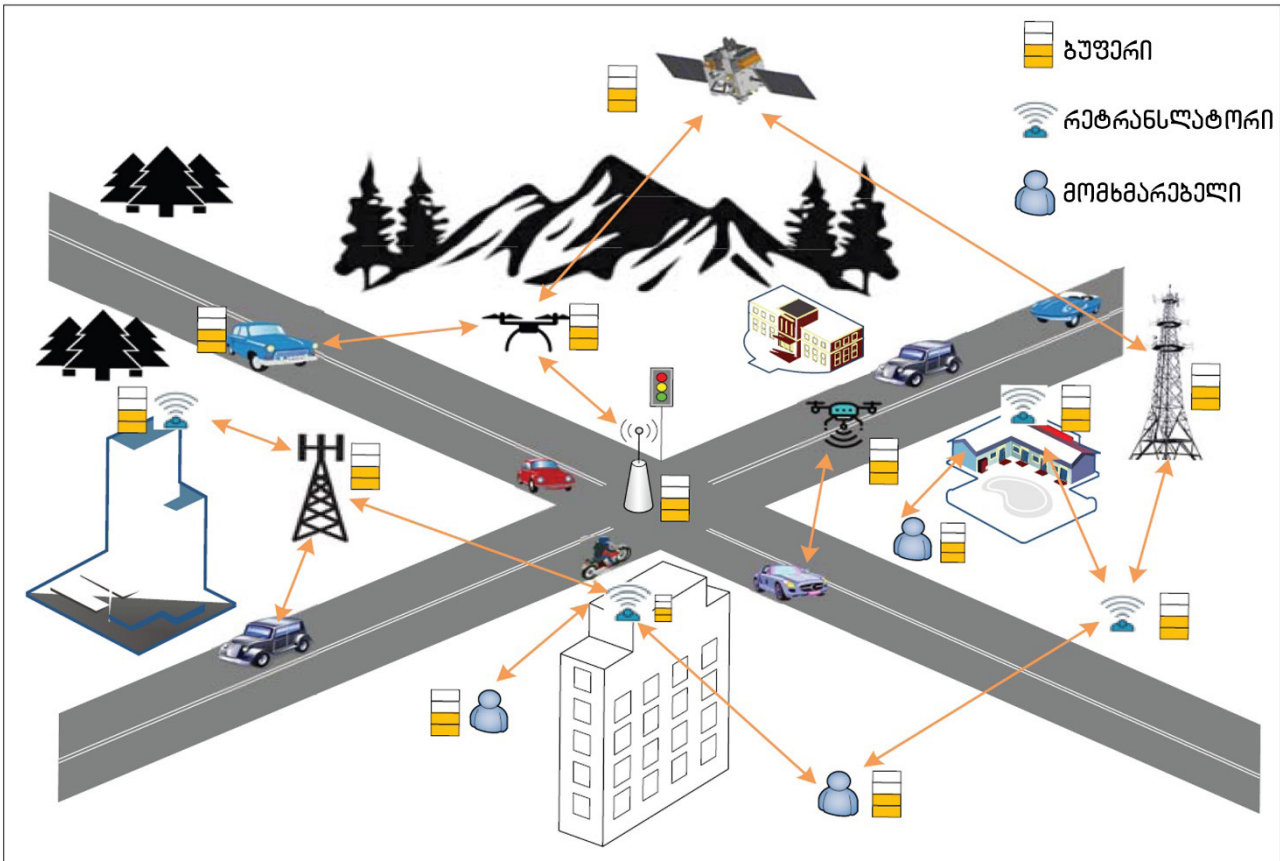
QoS მოთხოვნების დასაკმაყოფილებლად და რაც შეიძლება მეტი მოწყობილობის მომსახურებისთვის დაინერგა კოოპერატიული ქსელი. კოოპერატიული გადაცემის მნიშვნელოვანი მახასიათებლებია უსადენო ქსელის დაფარვის გაფართოება, გადაცემის საიმედოობის გაზრდა და მომხმარებლის წვდომის გაფართოება. გარდა ამისა, რეტრანსლატორები და წვდომის წერტილები, ჩვეულებრივ, გამოიყენება ტერესტრიალური და არატერესტრიალური ქსელების (NTN) დაფარვის გაუმჯობესებისათვის. სისტემის საიმედოობის, სპექტრის გამოყენებისა და დაფარვის შემდგომი გაზრდის მიზნით, ბუფერული (ანუ ბუფერის დახმარებით განხორციელებული) სარელეო გადაცემა (რეტრანსლაცია) გამოკვლეულია, როგორც პერსპექტიული ტექნიკა. ბუფერების გამოყენება ბუფერზე დაფუძნებულ კოოპერატიულ ქსელებში (BACN) საშუალებას აძლევს რეტრანსლატორს, დროებით შეინახოს მონაცემთა პაკეტები და გადასცეს ისინი დროის არამიმდევრობით ორ ინტერვალში, რაც უზრუნველყოფს სისტემის დაგეგმვის უფრო მეტ მოქნილობას და საშუალებას იძლევა მოხდეს ქსელის რესურსების ოპტიმიზაცია.

ტიპური გამოყენებები:

IoT სისტემები აკავშირებს არა მხოლოდ ადამიანებს, არამედ უამრავ სენსორს სახლებში, სამედიცინო ან სხვა დაწესებულებებში. ბუფერულ რეტრანსლატორს შეუძლია წინასწარ შეინახოს გაზიარებული ინფორმაცია სპეციალურ მოწყობილობაზე, რაც დაგვეხმარება ორ დისტანციურ IoT კვანძს შორის მონაცემთა გადაცემის პროცესში. V2X არის IoT-ის ტიპი, რომელიც აკავშირებს მანქანებს მიმდებარე ობიექტებთან, როგორცაა გზისპირა ინფრასტრუქტურა, მოწყობილობები, ფეხით მოსიარულეები, ქსელები და ახლომდებარე მანქანები. ის გახდა ინტელექტუალური სატრანსპორტო სისტემის საფუძველი და ცდილობს მართვის პროცესი უფრო უსაფრთხო და ინტელექტუალური გახადოს.

მობილურ სატრანსპორტო საშუალებებს შეუძლიათ გადაიტანოს გადაუდებელი ინფორმაცია ცენტრალურ კვანძში, რომელიც შეიძლება ინახებოდეს ბუფერებში და შემდეგ ეს ინფორმაცია გადასცეს ახალ სატრანსპორტო საშუალებებს, რომლებიც შედიან გზაჯვარედინზე, როგორც ნაჩვენებია ნახ. 4.1-ზე. თუმცა, V2X კომუნიკაციას აქვს მაღალი სივრცით-დროითი დინამიკა, რომელიც მოითხოვს ბუფერების განახლებას და იქ არსებული ინფორმაციის ხშირად შეცვლას სატრანსპორტო საშუალებების სხვადასხვა ადგილზე და დროით ინტერვალში შეჭრის გამო.

NTN-ები უზრუნველყოფენ ფართო დაფარვისა და მობილური კომუნიკაციის სერვისებს არასახმელეთო პლატფორმების ისეთი საშუალებით, როგორცაა თანამგზავრები, მაღალ სიმაღლეზე განთავსებული აეროსტატები ან უპილოტო საფრენი აპარატები (UAV). თანამგზავრული კომუნიკაციებისთვის, ლიტერატურაში გამოკვლეულია კავშირი QoS მეტრიკებს შორის (რომელთა შორისაა ბუფერის ზომა, შეყვანება და გადაცემის სიჩქარე, და ქსელის კონფიგურაცია) ახალი ტანდემური რიგების მოდელის აგებით, თანამგზავრთა მონაცემების სარელეო ქსელებში მონაცემთა ჩამოსვლის, გადაგზავნის და ჩამოტვირთვის პროცესების იმიტირებით. ასევე გამოიკვლიეს UAV-ებზე დაფუძნებული მობილური სარელეო გადაცემის სისტემა, რომელსაც უპილოტო საფრენი აპარატი ეხმარებოდა ბუფერის სასრული ზომით და რომელიც ითვალისწინებს შეყვანების შემღვდევებს. ბუფერზე



ნახ. 4.1. 5G/6G ქსელების კომუნიკაციის სცენარები ბუფერული რეტრანსლაციისთვის

დაფუძნებულ UAV-ებს შეუძლიათ დინამიკურად დაარეგულიროს NTN ტოპოლოგია მომსახურების მოთხოვნების შესაბამისად და საჭიროების შემთხვევაში გადაიტანოს მონაცემები დაბრკოლებების გარშემო. მეორე მხრივ, ბუფერულ UAV-ებს შეუძლიათ უზრუნველყოს მონაცემთა გადაცემა საუკეთესო შესაძლებლობებით, თუნდაც მკაცრი ამინდის პირობებში, რადგან UAV-ებს შეუძლიათ მონაცემთა პაკეტების შენახვა ბუფერებში, როდესაც არხის მდგომარეობა ცუდია და მონაცემთა პაკეტები გადასცენ მომხმარებლებს არხის გაუმჯობესების შემდეგ. ამრიგად, სპეციალური (ad-hoc) ქსელები, მათ შორის თანამგზავრები და UAV-ები, შეიძლება გამოყენებულ იქნეს სამამშველო ოპერაციებისთვის მიწისძვრის, წყალდიდობის ან შორეულ ადგილებში.

BACN-ებისთვის რეტრანსლატორის შერჩევის სქემები:

ტრადიციული ბუფერული რეტრანსლატორის სქემების შერჩევის პროცედურები, როგორც წესი, ევრისტიკულად არის შექმნილი, როგორც უსადენო არხის, ასევე ბუფერული მდგომარეობის გათვალისწინებით. მაგალითად, კლასიკური მაქსიმალური ლინკის (max-link) მეთოდი ირჩევს ლინკს ყველა ხელმისაწვდომ ლინკს შორის არხის საუკეთესო მდგომარეობით თითოეულ დროით ინტერვალში, რაც აღწევს განცალკევების მაქსიმალურ მნიშვნელობას. თუმცა, ვინაიდან max-link სქემა ეხება მხოლოდ ცარიელ ან სრულ ბუფერულ მდგომარეობას, მას აქვს პაკეტის საშუალო შეყოვნების დიდი მნიშვნელობა, რომელიც იზრდება რეტრანსლატორების რაოდენობასთან და ბუფერების ზომებთან ერთად. ამ პრობლემის გადასაჭრელად, ბუფერის მდგომარეობაზე დაფუძნებული რეტრანსლატორების სქემების შერჩევის პროცედურა იყო შემოთავაზებული, რომელიც ითვალისწინებს როგორც არხის ხარისხს, ასევე ბუფერულ მდგომარეობებს. გარდა ამისა, მაღალი საიმედოობის და დაბალი შეყოვნების მისაღწევად, შემოთავაზებული იყო გადახრის (დევიაციის) მნიშვნელობებზე დაფუძნებული რეტრანსლატორის სქემის შერჩევის პროცედურა, სადაც გადახრის მნიშვნელობების განაწილება დამოკიდებულია მყისიერ და წინასწარ განსაზღვრულ ბუფერულ მდგომარეობებს შორის განსხვავებაზე.

BACN ქსელებისთვის რეტრანსლატორის შერჩევის ხელსაყრელი სქემის შემუშავებისას საჭიროა სხვადასხვა ასპექტის გათვალისწინება, რომელთა შორისაა კავშირის არხის მდგომარეობა, ბუფერული მდგომარეობები, შეყოვნების მახასიათებლები და ა. შ. თუმცა, ეს ასპექტები ცვალებადია დროის მიხედვით და იცვლება გარემოს ცვლილებებთან ერთად. ამიტომ შეყოვნების მოთხოვნები ხშირად ეწინააღმდეგება რეტრანსლატორის

შერჩევას მტყუნების მოთხოვნებს. მაგალითად, არხის მდგომარეობით ლინკის არჩევამ შეიძლება არ დააკმაყოფილოს ბუფერული მდგომარეობის მოთხოვნა ან გამოიწვიოს ხანგრძლივი შეყოვნება. ამ დროში ცვალებადი პრობლემების მოგვარება ჯერ კიდევ ღია პრობლემაა, განსაკუთრებით მაშინ, როდესაც ბუფერის ზომა ან რეტრანსლატორების რაოდენობა მნიშვნელოვანია. ამრიგად, ევრისტიკულად შემუშავებული ტრადიციული ბუფერული რეტრანსლატორის შერჩევის სქემების მასხაიათებლები შორს არის ოპტიმალურისგან, რაც აიძულებს მკვლევარებს შეისწავლონ DL-ის მეთოდები ბუფერზე დაფუძნებულ კოოპერატიულ ქსელებში რეტრანსლატორების შერჩევის პრობლემის გადასაჭრელად, რომლებსაც აქვთ უპირატესობები მაღალგანზომილებიანი და დროში ცვალებადი პრობლემების გადაჭრაში. DL-ის უპირატესობა ის არის, რომ იგი წარმოადგენს მონაცემთა ბაზაზე ორიენტირებულ ინსტრუმენტს, რომელიც გამორიცხავს მათემატიკური მოდელის წინასწარი შემუშავების აუცილებლობას და საშუალებას იძლევა განხორციელდეს ოპტიმიზაცია მონაცემთა სასწავლო ნაკრების საფუძველზე. მაგალითად, ლიტერატურაში შემოთავაზებული იყო რეტრანსლატორების შერჩევის ორი მეთოდი DRL-ის გამოყენებით, რათა გადაჭრილიყო გამტარუნარიანობის მაქსიმიზაციის პრობლემა შეყოვნებაზე მკაცრი შეზღუდვების პირობებში. აღნიშნული მიდგომა განსხვავდება ბუფერზე დაფუძნებული არსებული სქემებისგან, რომლებიც, თავის მხრივ, როგორც წესი, ითვალისწინებენ პაკეტის საშუალო შეყოვნებას და აღწევენ შესაბამის გამტარუნარიანობას შეყოვნებაზე არსებული შეზღუდვების გათვალისწინებით. თუმცა, DL-ის ალგორითმების ჩართვა BACN-ში მრავალი მიზნებითა და შეზღუდვებით, როგორცაა უსაფრთხო გადაცემები, შეყოვნებების შემცველი გადაცემები, ჰეტეროგენული ქსელები და არასრულყოფილი CSI, არ არის მარტივი ამოცანა. გარდა ამისა, იმის გათვალისწინებით, რომ DL-ის თანამედროვე მეთოდების უმეტესობა ძირითადად ფოკუსირებულია კლასიფიკაციასთან დაკავშირებულ პრობლემებსა და ბუნებრივი ენის დამუშავების აპლიკაციებზე, კომპლექსური მიზნებისა და შეზღუდვების ინტეგრაცია DL მეთოდებში ზოგადად მნიშვნელოვან გამოწვევებს ქმნის.

4.2. ტექნიკური გამოწვევები BACN ქსელებისთვის

არსებულ სისტემებში BACN-ების ხშირი გამოყენების მიუხედავად, არსებობს რამდენიმე კრიტიკული გამოწვევა ბუფერის გამოყენებისას მომავალ უსადენო ქსელებში. ეს პარაგრაფი განიხილავს ზოგიერთ ამ გამოწვევას.

დაბალი შეყოვნება:

მომავალ 5G ქსელებში შეყოვნებისადმი მგრძობიარე უსადენო მულტიმედია აპლიკაციების სწრაფად მზარდი მოთხოვნების დასაკმაყოფილებლად, BACN-მა უნდა გაითვალისწინოს პაკეტის შეყოვნება. კონკრეტული შეყოვნების შეზღუდვის გათვალისწინებით, მულტიმედია უსადენო სერვისების QoS-ის უზრუნველყოფის დიზაინის საკითხებმა მნიშვნელოვანი გვლევითი ყურადღება მიიპყრო. მიყუჩების (ფედინგის) შემცველი უსადენო არხების დროში ცვალებადი ბუნების გამო, დეტერმინირებული შეყოვნებით განპირობებული QoS შეზღუდვები აღარ შეიძლება გამოყენებულ იქნეს მულტიმედია უსადენო სერვისების რიგების ქცევის დასახასიათებლად. ამ მიზნით, შემოთავაზებულია პაკეტის საშუალო და მკაცრი შეყოვნება, რათა მხარი დაუჭიროს შეფერხებისადმი მგრძობიარე უსადენო კომუნიკაციას 5G მულტიმედია მობილურ უსადენო ქსელებში. ბუფერულ კოოპერატიულ კომუნიკაციაზე არსებული საშუალებები ძირითადად ითვალისწინებს პაკეტის საშუალო შეყოვნებას. თუმცა, შეყოვნების მკაცრი შეზღუდვა კონკრეტულ 5G სცენარებში აუცილებელია, რადგან პაკეტის საშუალო შეყოვნება ვერ უზრუნველყოფს მონაცემთა სხვადასხვა პაკეტის შეყოვნების „სამართლიანობას“. მიუხედავად ამისა, შეყოვნების მკაცრ შეზღუდვებს შეუძლია უზრუნველყოს მონაცემთა სხვა პაკეტების შეყოვნების „სამართლიანობა“. ამიტომ, უფრო დაბალი შეყოვნებისკენ სწრაფვა არის ზოგადი ტენდენცია 5G-ში, იმისგან დამოუკიდებლად, არის ეს საშუალო შეყოვნება თუ მკაცრი შეყოვნება.

მაღალი საიმედოობა:

ასევე გამოკვლეულია შეყოვნებასა და საიმედოობას შორის ურთიერთკომპრომისული დაბალანსება. უსადენო სისტემების მზარდი სირთულისა და კვანძების რაოდენობის ფონზე, შეყოვნებისა და საიმედოობის დაბალანსება კრიტიკულ გამოწვევას წარმოადგენს. ქსელში IoT მოწყობილობების მოსალოდნელი დიდი რაოდენობა აიძულებს 5G არქიტექტურას შეიმუშაოს საიმედო გადაცემის სისტემა, რომელიც დაფუძნებულია ბუფერულ კოოპერატიულ კომუნიკაციაზე. სამომავლო ქსელი უნდა აკმაყოფილებდეს პაკეტის/მონაცემთა გადაცემის მაღალ სიჩქარეს და მონაცემთა საიმედო კომუნიკაციას რეალურ დროში ფუნქციონირებად აპლიკაციებსა (როგორცაა V2X,

ციფრული ინტელექტუალური მედიცინა და ა. შ.) და განაწილებულ პერიფერიულ მოწყობილობებს შორის. კონკრეტულად, მონაცემთა საიმედოობის საჭიროება იქნება 99.9999 პროცენტამდე. მასობრივი ულტრა საიმედო დაბალი შეყოვნების კომუნიკაციის (mURLLC) სერვისები საჭიროა 5G-ში და BACN-ის წინაშე დგას საიმედოობის გამოწვევები შეყოვნების მოთხოვნების დაკმაყოფილებისას.

გადაცემა არასრულყოფილი CSI-ით:

BACN-ზე არსებული საშუალების უმეტესობაში განხილულია გადაცემა მყისიერი სრულყოფილი CSI-ის საფუძველზე. თუმცა, მყისიერი სრულყოფილი CSI-ის მიღება რთული ან შეუძლებელია ფართომასშტაბიან ქსელებში უკუკავშირის შეყოვნებისა და არხის შეფასების შეცდომების გამო. არასრულყოფილი CSI-ის განხილვისას, რეტრანსლაციონების შერჩევის არსებული ტრადიციული სქემების უმეტესობისთვის არსებობს მაღალი ალბათობა იმისა, რომ შერჩეულ იქნეს არაკვალიფიციური ლინკები გადაცემისთვის და, შესაბამისად, სისტემის მუშაობა გაუარესდება ისეთი ფაქტორებით, როგორცაა მტყუნების ალბათობა, შემცირებული გამტარუნარიანობა და ხანგრძლივი შეყოვნებები. აქედან გამომდინარე, რეტრანსლაციონების შერჩევის ახალი სქემების შემუშავება არასრულყოფილი CSI-ით აუცილებელია, რაც ასევე რთული ამოცანაა, რადგან გადაწყვეტილების მიღებმა უნდა შეაფასოს CSI-ის არასრულყოფილების გავლენა სისტემის მუშაობის მახასიათებლებზე.

უსაფრთხო გადაცემა ფიზიკურ ფენაში:

მონაცემთა დიდი მოცულობის გადაცემა, მასობრივი საკომუნიკაციო კვანძების არსებობა, მონაცემთა დროებითი შენახვა ბუფერებში და პოტენციური მოსმენების არსებობა აჩენს პრობლემებს მონაცემთა უსაფრთხოებასთან დაკავშირებით. კერძოდ, უსადენო საკომუნიკაციო გარემოში გავრცელების ხასიათიდან გამომდინარე, ინფორმაციის გაუთხრევა გადაცემის დროს ძალზე სავარაუდოა. ამიტომ, ბუფერზე დაფუძნებული რეტრანსლაციის შესაბამისი ტექნიკა და პროტოკოლები უნდა შემუშავდეს მომხმარებლის კონფიდენციალურობისა და მონაცემთა უსაფრთხოების დასაცავად. თუმცა, მომხმენი მოწყობილობების არსებობის გამო, ბუფერის დახმარებით შემუშავებული რეტრანსლაციის სქემის დაპროექტება ძალიან რთულია, რადგან შეიძლება საჭირო გახდეს დამატებითი კომპლექსური ფაქტორების ერთობლივი განხილვა, როგორცაა CSI-ის მოსმენა, საიდუმლოების გამტარუნარიანობა, ხელოვნური ხმაური და ა. შ. ამ შემთხვევაში, ხელსაყრელი ბუფერული რეტრანსლაციის სქემის შემუშავება პრობლემატურია და ანალიზი ხდება ძალიან რთული. უფრო მეტიც, საკომუნიკაციო სისტემაში მოსმენის კვანძების რაოდენობის ზრდასთან ერთად, ევრისტიკულად შემუშავებული ბუფერული რეტრანსლაციის სქემების დახმარებით ადვილი არ არის უსაფრთხო გადაცემის უზრუნველყოფა.

ჰეტეროგენული ქსელები:

ჰეტეროგენული ქსელი არის ქსელის ტიპი, რომელიც შედგება ისეთი სხვადასხვა ტიპის კვანძებისა და მოწყობილობებისგან, როგორცაა მანქანები და ზოგიერთი IoT კვანძი, რომელსაც შეუძლია სხვადასხვა სერვისისა და აპლიკაციის მხარდაჭერა. სხვადასხვა ქსელის ინტეგრაცია უზრუნველყოფს უწყვეტ კავშირს, ასევე გაზრდილ დაფარვასა და მოცულობას, რაც მომხმარებლებს საშუალებას აძლევს მიიღოს წვდომა სერვისებზე მათი მდებარეობის ან კონკრეტული მოწყობილობის მიუხედავად. ამიტომ აუცილებელია ქსელის ჰეტეროგენული მართვის შესწავლა სხვადასხვა კვანძის დასაკავშირებლად. აქ ერთ-ერთი მთავარი გამოწვევაა სხვადასხვა ტიპის კვანძებისა და მოწყობილობების ინტეგრაციის მართვა უწყვეტი კავშირის, გაზრდილი დაფარვისა და სიმძლავრის უზრუნველსაყოფად. გარდა ამისა, ჰეტეროგენულ ქსელში მოწყობილობების დიდი რაოდენობა იწვევს ინტერფერენციას, რაც გავლენას ახდენს ქსელის მუშაობაზე. მაშასადამე, ბუფერის დახმარებით რეტრანსლაცია არის სფერო, რომელიც მოითხოვს მეტ გამოკვლევას ჰეტეროგენულ ქსელებში მონაცემთა გადაცემის ეფექტიანობის გასაუმჯობესებლად.

აღნიშნავთ, რომ BACN ასევე ქმნის დამატებით საინჟინრო გამოწვევებს, რომლებიც დაკავშირებულია არხის მოდელირებას, შეფასებასა და არხის რეალისტური მონაცემების გამომუშავებასთან. მოთხოვნილ სცენარებში, როგორცაა V2X, NTN და დრონით კომუნიკაციები, არხები დინამიკურობის მაღალ დონეს ავლენენ, რაც ართულებს არხის ზუსტი ინფორმაციის მიღებას. აქედან გამომდინარე, არსებობს საიმედო მეთოდების საჭიროება, რომლებსაც შეუძლიათ ეფექტურად წარმოქმნას და ზუსტად განსაზღვროს არხის მონაცემები. ერთ-ერთი პოტენციური მიდგომა სხივების მიკვლევისა და ხელოვნური ინტელექტის ალგორითმების შერწყმაა, რაც შესაძლებელს გახდის სუპერგარჩევადობის მოდელირების მეთოდის შემუშავებას, რომელიც რეალურ დროში სიმულაციის მოთხოვნებს აკმაყოფილებს.

4.3. DL-ის პოტენციური ტექნოლოგიები

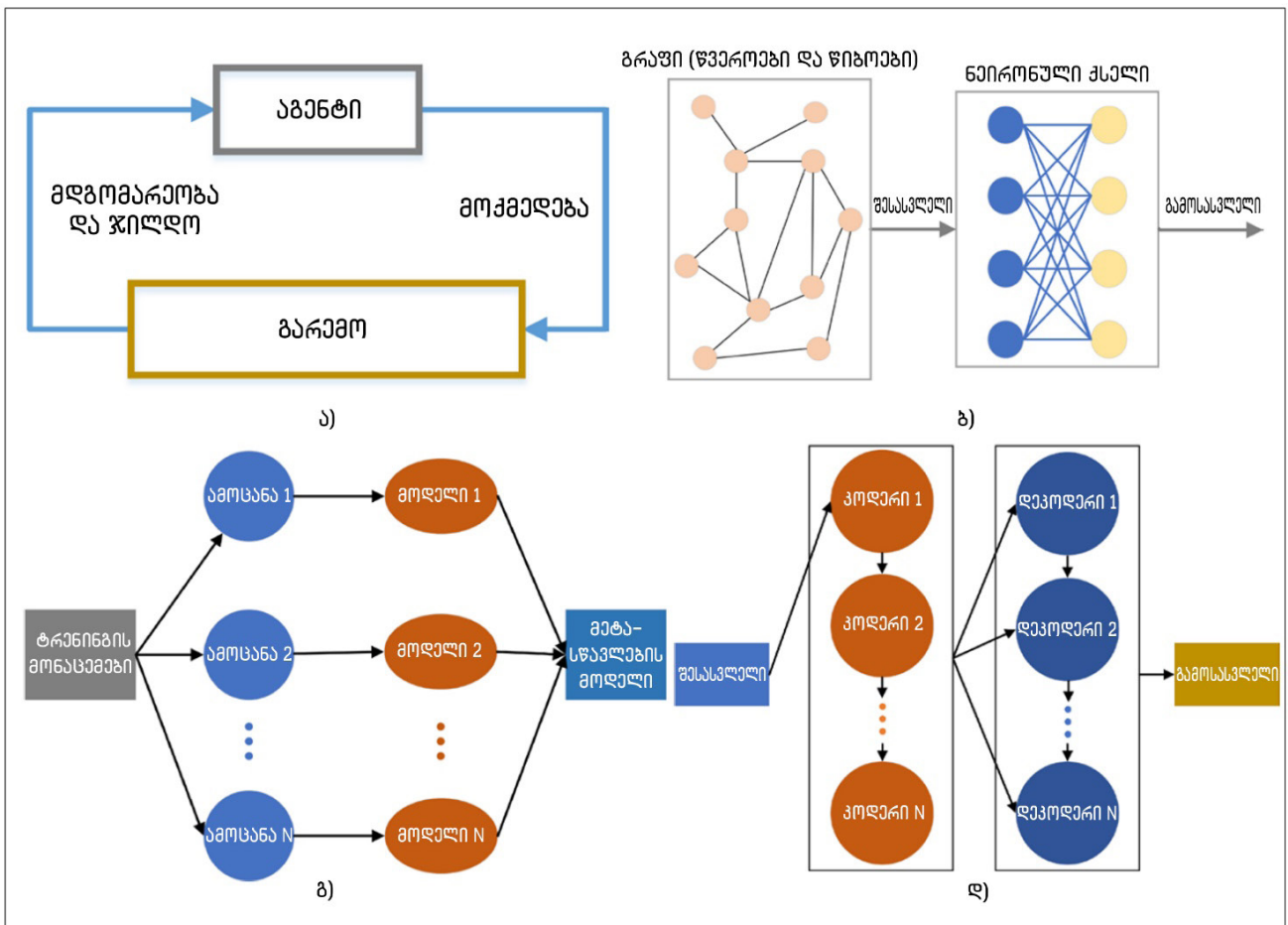
ამ პარაგრაფში წარმოდგენილია DL-ის მიმდინარე ტექნოლოგიები, რომლებსაც აქვთ BACN-ში გამოყენების პოტენციალი.

ღრმა განმტკიცებელი სწავლება:

უსადენო კომუნიკაციების ოპტიმიზაციის პრობლემების ფართოდ გამოყენებული გადაწყვეტა არის DRL. DRL ალგორითმები ეფუძნება მარკოვის გადაწყვეტილების პროცესების (MDP) იდეას, რომელიც წარმოქმნის მათემატიკურ ჩარჩოს. ეს უკანასკნელი იყენებს მდგომარეობებს, მოქმედებებს, გადასვლებს და ჯილდოებს გარემოს აღსაწერად, როგორც ნაჩვენებია ნახ. 4.2ა-ბე. MDP-ებში აგენტი გამოიყენება გარემოსთან ურთიერთობისთვის, ოპტიმალური გადაწყვეტილებების შესასწავლად. ამ პროცესს ეწოდება განმტკიცებელი სწავლება, რადგან აგენტი სწავლობს გარემოდან მიღებული ჯილდოებიდან. ტრადიციულ ოპტიმიზაციის მეთოდებთან და DL-ის ალგორითმებთან შედარებით, DRL უფრო მოქნილია რთული პრობლემების გადასაჭრელად დინამიკურ გარემოში, როგორცაა BACN.

გრაფებზე დაფუძნებული ნეირონული ქსელები:

მეორე მხრივ, გრაფზე დაფუძნებული ნეირონული ქსელები (GNN) არის DL-ის ტიპი, რომელიც სპეციალურადაა შექმნილი გრაფებზე განთავსებული სტრუქტურირებული მონაცემების დასამუშავებლად. უსადენო კომუნიკაციებში, გრაფის კვანძები წარმოადგენენ მოწყობილობებს ქსელში, ხოლო წიბოები — ურთიერთობებს მოწყობილობებს შორის, როგორც ნაჩვენებია ნახ. 4.2ბ-გე. მატრიცა ასახავს ურთიერთობებს სხვადასხვა მოწყობილობას შორის BACN-ებში და ნეირონული ქსელი გამოიყენება შესაბამისი ინფორმაციის დასამუშავებლად მოდულების გენერირებისთვის BACN-ებში მოწყობილობებს შორის ურთიერთობების აღსაწერად. უფრო მეტიც, მატრიცასა და ნეირონულ ქსელებს GNN-ში შეუძლიათ დიდი რაოდენობით მონაცემების დამუშავება, რაც ძალიან მნიშვნელოვანია მომავალი ფართომასშტაბიანი BACN-ებისთვის.



ნახ. 4.2. DL-ის ალგორითმების სტრუქტურა: ა) RL სტრუქტურა; ბ) გრაფული ნეირონული ქსელის სტრუქტურა; გ) მეტასწავლების სტრუქტურა; დ) ტრანსფორმული სწავლების სტრუქტურა.

მეტასწავლება:

ბოლო წლებში მეტასწავლება გახდა ფართოდ გამოყენებული ინსტრუმენტი მანქანური სწავლების სფეროში. ML-ის ტრადიციულ ალგორითმებში სწავლების მოდელი დატრენინგებულია კონკრეტული მონაცემთა ნაკრებისთვის კონკრეტული პრობლემის გადასაჭრელად. თუმცა, მიღებული გამოცდილება შესაძლოა აღარ იყოს შესაფერისი ახალი ამოცანებისთვის. ამრიგად, შემოთავაზებულია მეტასწავლება, რათა ვისწავლოთ თუ როგორ უნდა შევცვალოთ სწავლების მოდელი წინა ამოცანების საფუძველზე, რათა გაუმჯობესდეს მისი ადაპტირება ახალ ამოცანებთან. BACN-ებში მეტასწავლებას შეუძლია არსებული სწავლების მოდელის შეცვლა ახალი ამოცანებისთვის უფრო სწრაფად, ვიდრე ახლის დატრენინგება. აქედან გამომდინარე, მეტასწავლებას აქვს პოტენციური გააუმჯობესოს ML-ზე დაფუძნებული სისტემების ეფექტიანობა და მოქნილობა, რაც მას მნიშვნელოვანს გახდის მომავალ ფართომასშტაბიან BACN-ებში.

ტრანსფორმერი:

ტრანსფორმერი კიდევ ერთი მნიშვნელოვანი ალგორითმია DL-ის მიმდინარე სამუშაოებში. ტრანსფორმერის მთავარი იდეა თვითყურადღებაა, რაც საშუალებას აძლევს მას ერთდროულად განიხილოს მიმდევრობის ყველა პოზიცია და ისწავლოს კონტექსტური ურთიერთობები ნიშნებს შორის თანმიმდევრობით. ტრანსფორმერის მოდელი გამოიყენება ტრენინგის მონაცემთა ნაკრების ფარგლებში გრძელვადიანი დამოკიდებულებების შესასწავლად. თანმიმდევრული მოდელირების ეს შესაძლებლობა საშუალებას აძლევს მას უფრო ეფექტიანად გამოიტანოს ურთიერთობები წინა და მომდევნო მდგომარეობებს შორის BACN-ებში. ამრიგად, მას შეუძლია გადაწყვეტილების მიღების პროცესების ოპტიმიზაცია და გადაწყვეტილების სტაბილურობის გაუმჯობესება.

ცენტრალიზებული და განაწილებული სწავლება:

ცენტრალიზებული სწავლება არის DL-ის გამოყენების ერთ-ერთი სტრატეგია უსადენო კომუნიკაციებში. კერძოდ, ქსელის სხვა მოწყობილობების მონაცემები გროვდება საკონტროლო კვანძში და გამოიყენება გადაწყვეტილების მიღებისთვის სწავლების მოდელის ტრენინგისთვის. შემდეგ საკონტროლო კვანძში დატრენინგებული მოდელი გამოიყენება გადაწყვეტილების მისაღებად და მთელი ქსელისთვის მოქმედებების განსაზღვრისთვის. ეს მიდგომა საშუალებას იძლევა გამოიყენოს დიდი რაოდენობით მონაცემები უაღრესად ბუსტი მოდელის მოსამზადებლად, ამასთან ტრენინგის დროს ქსელში განთავსებულ მოწყობილობებს შორის დიდი რაოდენობით მონაცემების გადაცემას თავიდან ავიცილებთ.

მეორე მხრივ, განაწილებული სწავლებისას, შესაძლებელია ქსელში არსებული ყველა მოწყობილობის გამოთვლითი რესურსების გამოყენება და შესაბამისად თითოეულ მოწყობილობაზე საჭირო გამოთვლითი რესურსები მცირდება. გარდა ამისა, განაწილებული სწავლება საშუალებას იძლევა შემუშავდეს ოპტიმიზაციის გადაწყვეტილებები ფართომასშტაბიან ქსელებში, რადგან თითოეულ მოწყობილობას შეუძლია გადაწყვეტილების მიღება ადგილობრივად, რათა შემცირდეს ქსელში გადაცემული მონაცემების რაოდენობა. დაბოლოს, განაწილებულ სწავლებას შეუძლია გააუმჯობესოს კონფიდენციალურობა უსადენო ქსელისთვის, რადგან მგრძობიარე მონაცემები არ საჭიროებს მოწყობილობებს შორის გადაცემას.

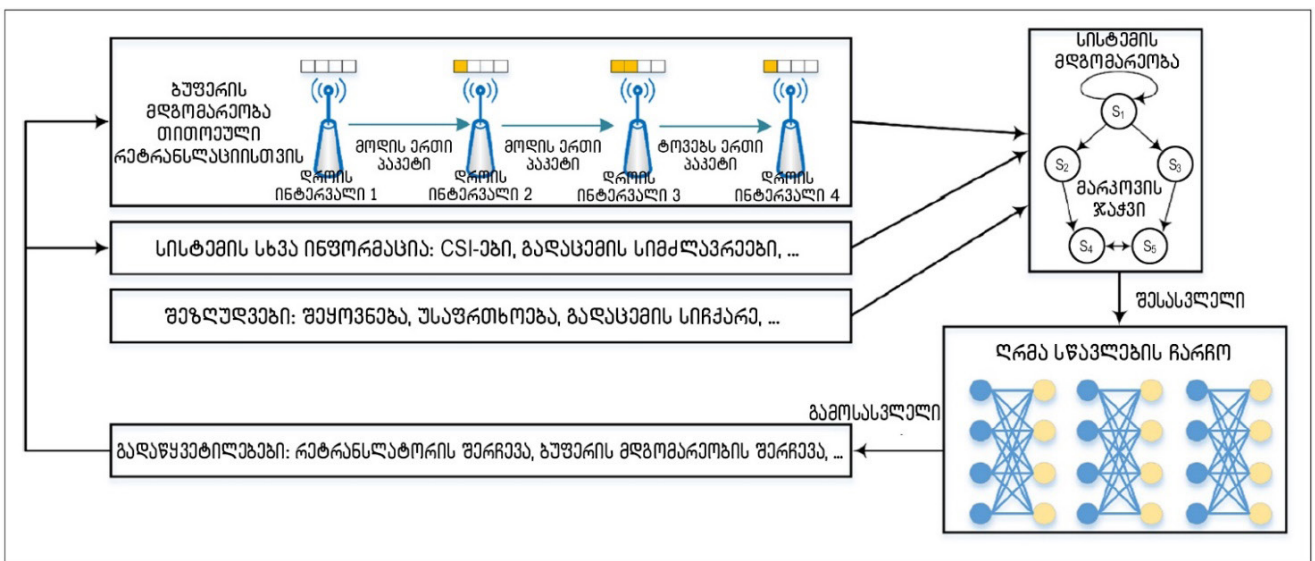
4.4. DL-ით გაძლიერებული BACN-ებისთვის შემოთავაზებული სტრუქტურები

DL-ის მთავარი იდეა არის ხელოვნური ნეირონების მრავალი ფენის გამოყენება მონაცემთა წარმოდგენის შესასწავლად. უსადენო ქსელებში DL-ის გამოყენების ერთ-ერთი მთავარი უპირატესობა ის არის, რომ მას შეუძლია შეისწავლოს ცვლადებს შორის ურთიერთობა მონაცემებში, რაც საშუალებას იძლევა მიღებულ იქნეს ბუსტი გადაწყვეტილებები ქსელის ოპტიმიზაციის კუთხით. გარდა ამისა, DL-ის ალგორითმების დატრენინგება შესაძლებელია ძველი (ისტორიული) მონაცემების გამოყენებით, რაც მას საშუალებას აძლევს მუდმივად მოერგოს უსადენო ქსელის ცვლილებებს. უფრო მეტიც, ბუფერზე დაფუძნებული რეტრანსლატორის შერჩევის პროცესი შეიძლება მოდელირებული იყოს როგორც MDP და შემდეგ DRL ალგორითმები შეიძლება გამოყენებულ იქნეს MDP ელემენტებიდან ამოხსნის შესასწავლად. DRL ფოკუსირებულია სიტუაციების შესატყვის ქმედებებთან, რათა მაქსიმალურად გაზარდოს ჯილდო. ამრიგად, DRL ალგორითმები შეიძლება გამოყენებულ იქნეს გამტარუნარიანობასა და შეყვანებას შორის ურთიერთობის ოპტიმიზაციისთვის ბუფერული რეტრანსლაციის სისტემებში. უფრო მეტიც, NTN-ებში DRL-ის კონვერგენციის მახასიათებლის გასაუმჯობესებლად, მეტასწავლება არის პერსპექტიული ინსტრუმენტი DRL-ში დატრენინგებული მოდელის დასარეგულირებლად.

მეორე მხრივ, მომავალი უსადენო ქსელი შეიძლება იყოს ფართომასშტაბიანი ჰეტეროგენული სისტემა, რომელიც გამოიწვევს ოპტიმიზაციის უფრო რთულ პრობლემას. GNN არის პერსპექტიული ინსტრუმენტი ამ პრობლემის გადასაჭრელად გრაფის სტრუქტურის მატრიცად წარმოდგენით და ამ მატრიცის დამუშავებით ნეირონული ქსელის გამოყენებით. თუმცა, ძნელია ტრენინგის მონაცემთა ნაკრების მოპოვება GNN-ისთვის NTN-ებში. ამრიგად, GNN-ის გამოყენება DRL-თან კომბინაციაში იძლევა რესურსების ოპტიმიზაციას DRL გადაწყვეტილებების საფუძველზე, სადაც GNN ეხმარება ქსელის ფუნქციების ამოღებას DRL ტრენინგისთვის.

როგორც ნახ. 4.3-ზეა ნაჩვენები, DL-ზე დაფუძნებული BACN-ები იყენებენ DL-ის ალგორითმების შესაძლებლობებს BACN-ებში რესურსის ოპტიმიზაციისთვის. სისტემის მდგომარეობის შესახებ ინფორმაციას, რომელიც მოიცავს ბუფერულ მდგომარეობას, CSI-ის და ოპტიმიზაციის შეზღუდვებს, აქვს პოტენციური გადავიდეს მრავალ მდგომარეობაში დროის შემდგომ პერიოდში, რითაც ჩამოყალიბდება მარკოვის ჯაჭვი. გარდა ამისა, სისტემის მდგომარეობა განიხილება, როგორც DL სტრუქტურის შემავალი მონაცემები, რომელიც შექმნილია როგორც ოპტიმიზატორი BACN-ებში. ზემოთ ნახსენები DL ალგორითმები გაერთიანებულია და გამოიყენება მისი სტრუქტურის შესაქმნელად, რათა მოამზადოს ნეირონული ქსელის მოდელი BACN-ებში გადაწყვეტილების მისაღებად. წინასწარ დატრენინგებული მოდელი ეხმარება DL სტრუქტურას გამომავალი შედეგების (ანუ BACN-ებისთვის გადაწყვეტილების ცვლადების) ოპტიმიზაციაში, როგორცაა ბუფერული რეტრანსლატორის შერჩევა და მარშრუტიზაციის სტრატეგია.

უფრო მეტიც, არსებული ბუფერული რეტრანსლატორების შერჩევის სქემების უმეტესობა ვარაუდობს, რომ არსებობს ცენტრალური საკონტროლო კვანძი, რომელიც მოიპოვებს ყველა საჭირო ინფორმაციას ქსელში და მიიღებს გლობალურ გადაწყვეტილებებს. თუმცა, კომუნიკაციის ღირებულების გათვალისწინებით, სარეულო კვანძებმა შეიძლება მოიპოვოს საჭირო ინფორმაცია მხოლოდ მებობელი კვანძებიდან, ან ტექნოლოგიური შეზღუდვები ცენტრალურ კვანძს არ აძლევს საშუალებას მოიპოვოს ყველა ინფორმაცია და გამოთვალოს გადაწყვეტილება. ამიტომ, ამ პრობლემის გადასაჭრელად დაინერგა DRL ალგორითმი, რომელსაც ეწოდება მრავალაგენტური პროქსიმალური პოლიტიკის ოპტიმიზაცია (MAPPO). ამ შემთხვევაში, თითოეული აგენტი მუშაობს შესაბამის საკომუნიკაციო კვანძზე და ყველა აგენტს შეუძლია ისწავლოს, რათა მიიღოს გადაწყვეტილება, მათი ურთიერთქმედებისა და ადგილობრივი გარემოს მეშვეობით. შემდეგ ყველა აგენტი თანამშრომლობს განაწილებული გზით, რათა გააუმჯობესოს ჯამური გამტარუნარიანობა შეყოვნების შეზღუდვით. უფრო მეტიც, GNN გამოიყენება ქსელის გრაფების მოდელირებისთვის შემოთავაზებულ ქსელში სხვადასხვა მოწყობილობას შორის ურთიერთობების დასაფიქსირებლად, რათა დაეხმაროს DRL გადაწყვეტილებების ოპტიმიზაციას. შემოთავაზებულ NTN-ში მოძველებული CSI-ის და MAPPO-ში მაღალი კონვერგენციის სირთულის გათვალისწინებით, მეტასწავლება ჩაშენებულია განხილულ MAPPO ალგორითმში დატრენინგებული მოდელის კორექტირებისთვის. კონკრეტულად, მეტასწავლებას შეუძლია შეისწავლოს ურთიერთობა ადრე დატრენინგებულ მოდელსა და



ნახ. 4.3. DL სტრუქტურა BACN-ებისთვის

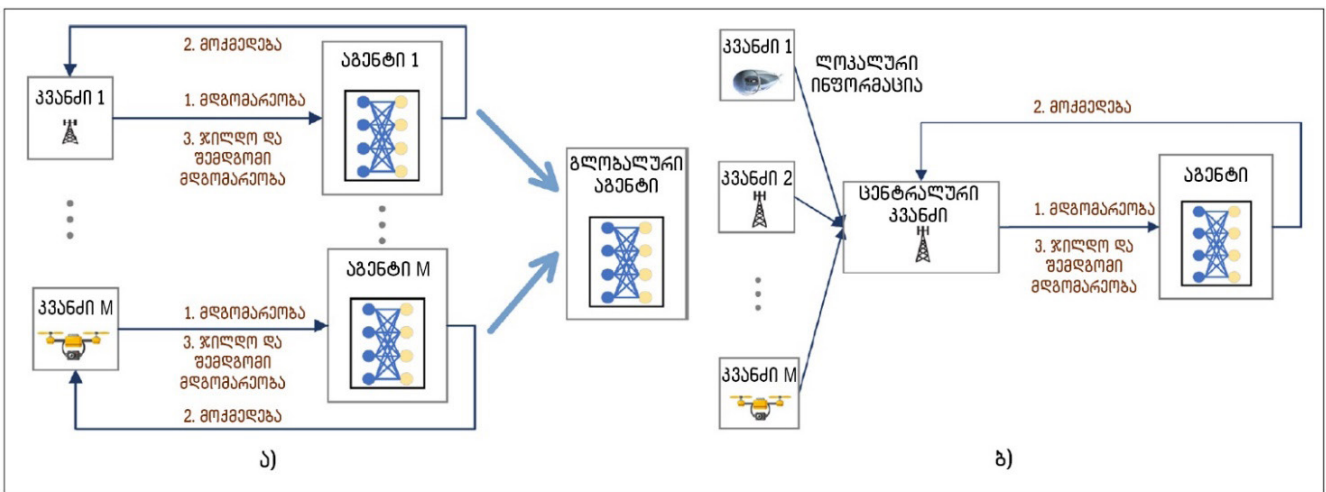
მათ პროგნოზებს შორის და შემდეგ შექმნას მეტამოდელი გადაწყვეტილების მოდელში ჰიპერპარამეტრების დასარეგულირებლად.

DL-ზე დაფუძნებული ოპტიმიზაციის ეფექტიანობის უზრუნველსაყოფად ბუფერზე დაფუძნებულ NTN-ებში, ჩვენ განვიხილავთ დედამიწის დაბალი ორბიტის თანამგზავს, როგორც წყაროს, მრავალ მაღალი სიმაღლის პლატფორმას (HAP), UAV-ებს და საბაზო სადგურებს, როგორც სარელეო კვანძებს, და მიწისპირა მომხმარებელს, როგორც დანიშნულების წერტილს. თითოეული სარელეო კვანძი აღჭურვილია ბუფერით და მუშაობს ნახევრად დუბლექსურ რეჟიმში. ამრიგად, სარელეო კვანძს არ შეუძლია ერთდროულად გადასცეს და მიიღოს სიგნალი მრავალნახტომიან ქსელში. ივარაუდება, რომ ყველა არხი განიცდის რაისის მიყუჩებას. არასანდო მომხმარებელი განიხილება როგორც მომხმენი, რომელსაც შეუძლია ბუფერზე დაფუძნებული NTN-ის სიგნალების მოსმენა, რაც ქმნის უსაფრთხო გადაცემის პრობლემას ფიზიკურ ფენაში. გარდა ამისა, მკაცრი შეყოვნების შეზღუდვები გათვალისწინებულია გამტარუნარიანობის მახასიათებელში, რაც ქმნის დაბალი შეყოვნების პრობლემას. უფრო მეტიც, მოძველებული CSI შემოთავაზებულ ქსელში ქმნის CSI-ის არასრულყოფილების პრობლემას უსადენო გადაცემაში. ამ შეზღუდვებისა და გამოწვევების ფონზე, ძალიან მნიშვნელოვანია შემოთავაზებული DL სტრუქტურის გამოყენება სისტემის გამტარუნარიანობის მაქსიმალურად გაზრდისა და მონაცემთა საიმედო გადაცემის უზრუნველსაყოფად.

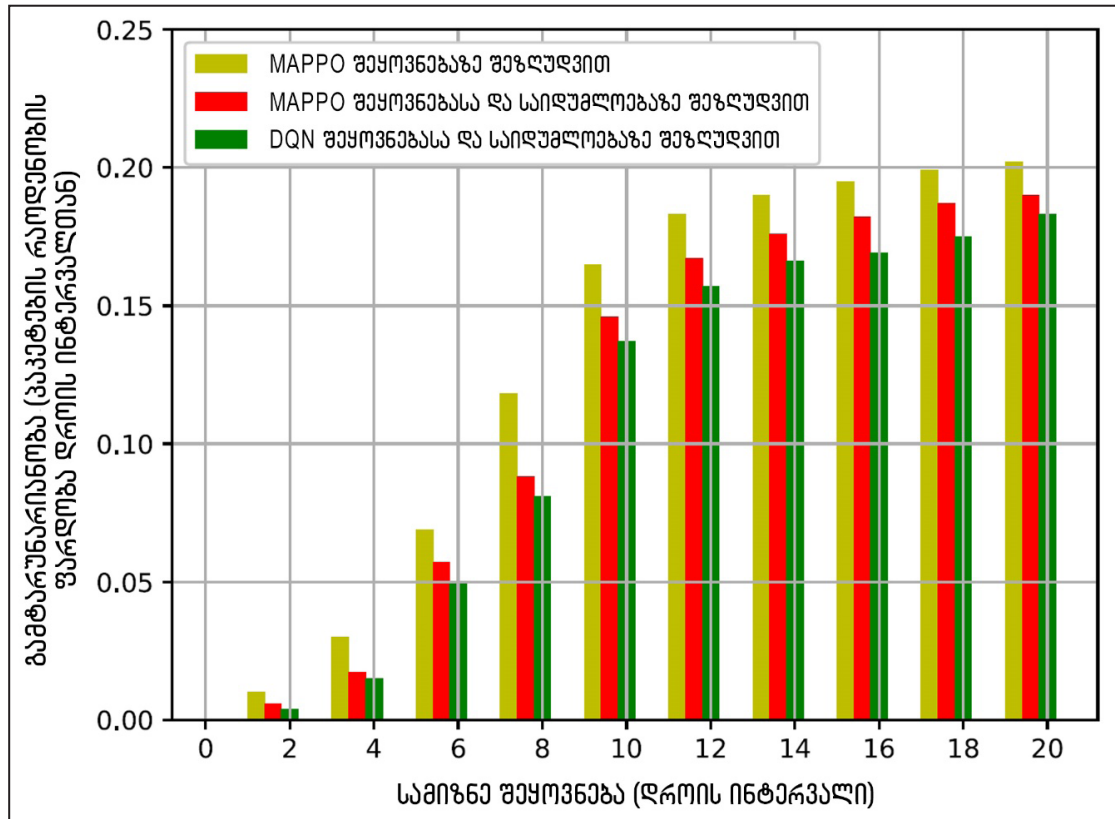
იმის გათვალისწინებით, რომ ცენტრალიზებული მართვა შეუძლებელია დეცენტრალიზებულ NTN-ში, ჩვენ ვიყენებთ MAPPO-ს, როგორც ნაჩვენებია ნახ. 4.4ა-ზე, ბუფერული რეტრანსლაციის შემცველ NTN-ში, რათა მოვახდინოთ მრავალნახტომიანი მარშრუტიზაციის სტრატეგიის ოპტიმიზაცია, MDP გარემოსთან ურთიერთქმედებით დროის 5000 ინტერვალზე.

ნახ. 4.5 აჩვენებს, რომ მიუხედავად იმისა, რომ მოსმენის არსებობა ამცირებს გამტარუნარიანობას, MAPPO-ზე დაფუძნებულ შემოთავაზებულ ალგორითმს მაინც შეუძლია მოახდინოს გადაწყვეტის ოპტიმიზაცია და, შესაბამისად, გამტარუნარიანობის მაქსიმიზაცია შეყოვნებისა და საიდუმლოების შეზღუდვებით. გარდა ამისა, იმის გათვალისწინებით, რომ თითოეულ აგენტს შეუძლია მიიღოს ცოდნა მხოლოდ მიმდებარე კვანძებიდან, ყველა აგენტი სწავლობს ადგილობრივ გარემოს და თანამშრომლობს ერთობლივი გადაწყვეტის შესაქმნელად. უფრო მეტიც, ღრმა Q-სწავლების ქსელთან (DQN) შედარებით, MAPPO-ს შეუძლია ჰქონდეს უკეთესი კონვერგენცია (კრებადობა) ნდობის რეგიონის განახლების გაუმჯობესებული სტრატეგიის გამოყენებით.

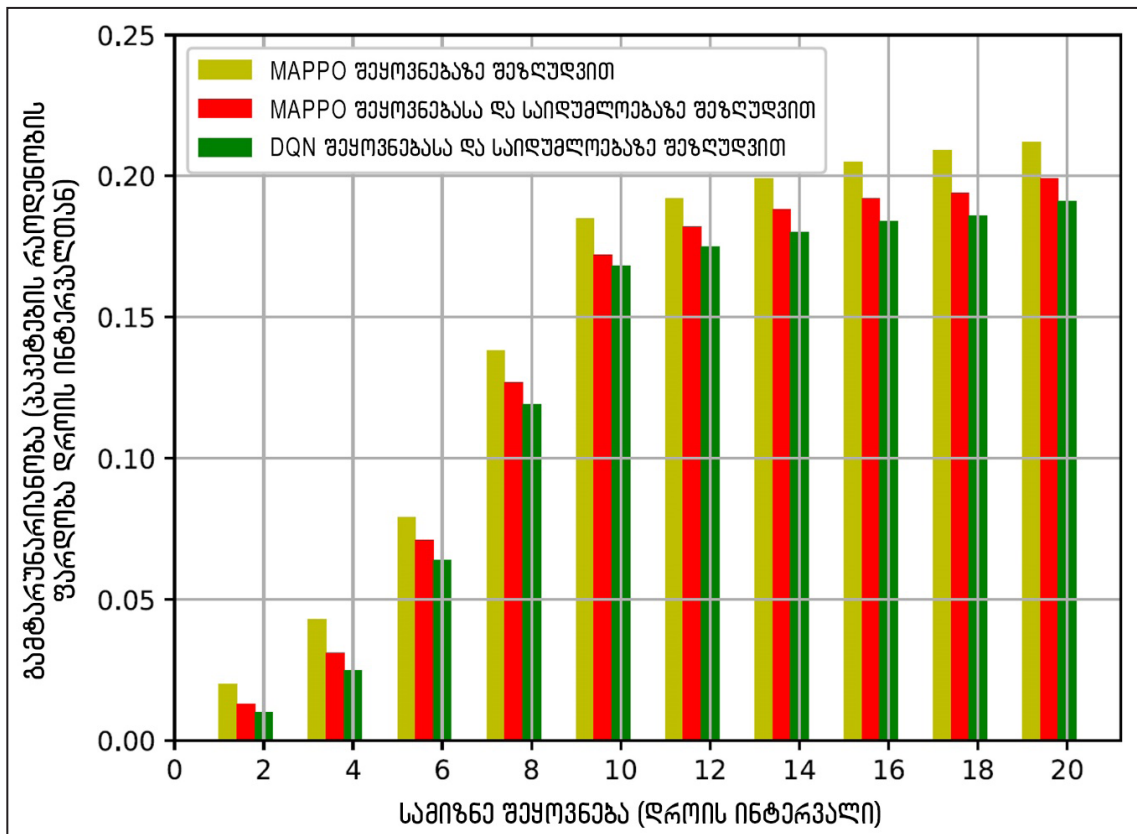
ნახ. 4.6 გვიჩვენებს DRL ალგორითმების მახასიათებლებს ცენტრალიზებულ ბუფერზე დაფუძნებულ NTN-ში. ცენტრალიზებულ სცენარში, ჩვენ განვიხილავთ ცენტრალურ კონტროლერს, რომელიც შეიძლება იყოს ნებისმიერი კვანძი და რომელიც სტრატეგიულად განლაგებულია გლობალური ინფორმაციის ოპტიმალურად მისაღებად პრაქტიკულ ქსელში. შემოთავაზებული ცენტრალიზებული DRL ალგორითმის სტრუქტურა ნაჩვენებია ნახ. 4.4ბ-ზე. შედეგები აჩვენებს, რომ სამიზნე შეყოვნებას დიდი გავლენა აქვს გამტარუნარიანობაზე NTN-ის შეყოვნების შეზღუდვის გამო. ბუფერული ტექნოლოგია აუმჯობესებს გადაცემის საიმედოობას სიგნალისა



ნახ. 4.4. დეცენტრალიზებული DRL-ისა და ცენტრალიზებული DRL-ის სტრუქტურა NTN-ში ბუფერული რეტრანსლაციით: ა) დეცენტრალიზებული DRL NTN-ში; ბ) ცენტრალიზებული DRL NTN-ში.



ნახ. 4.5. გამეტარუნარიანობის დამოკიდებულება სამიჯნე შეყოვნებაზე დეცენტრალიზებულ NTN-ში ბუფერული რეტრანსლაციით



ნახ. 4.6. გამეტარუნარიანობის დამოკიდებულება სამიჯნე შეყოვნებაზე ცენტრალიზებულ NTN-ში ბუფერზე დაფუძნებული რეტრანსლაციით

და ხმაურის დაბალი თანაფარდობის (SNR) არხებისთვის, მაგრამ ბუფერის გამოყენება ჩვეულებრივ იწვევს შეყოვნების ხარჯებს. შემოთავაზებულ MAPPO-ზე დაფუძნებულ ალგორითმს შეუძლია ურთიერთქმედება გარემოსთან, რათა დადგინდეს შეფერხების შეზღუდვის გავლენა და მოხდეს გადაწყვეტილებების პოვნის შესწავლა სისტემის თითოეული მდგომარეობისთვის. ამრიგად, MAPPO-ზე დაფუძნებულ ალგორითმს შეუძლია გადაწყვეტის ოპტიმიზაცია სხვადასხვა სამიზნე შეყოვნებით. გარდა ამისა, მოსმენის არსებობის გათვალისწინებით, MAPPO-ზე დაფუძნებული ალგორითმი უნდა ითვალისწინებდეს საიდუმლოების დონეს გადაცემისთვის რეტრანსლატორების არჩევისას. ამრიგად, მახასიათებელი ოდნავ უარესდება, რადგან აგენტმა შეიძლება თავი აარიდოს სარელეო კვანძების შერჩევას მომსმენის ახლომახლო. გარდა ამისა, ცენტრალიზებულ მიდგომას უკეთესი მახასიათებლები აქვს, ვიდრე განაწილებულს, შეყოვნებაზე სხვადასხვა შეზღუდვის პირობებში. მიზეზი არის ის, რომ ცენტრალიზებულ უსადენო საკომუნიკაციო გარემოში, ცენტრალურ კონტროლერს აქვს წვდომა ყველა ინფორმაციაზე, რათა მიადგინოს ოპტიმიზაციის გაუმჯობესებულ შედეგებს ყოვლისმომცველი მონაცემების შესწავლის გზით. ამის საპირისპიროდ, განაწილებულ გარემოში, რთულია ინფორმაციის გადაცემა ყველა კვანძიდან ცენტრალურ კვანძში ოპტიმალური გლობალური კონტროლისთვის. ამრიგად, თითოეულმა კვანძმა უნდა მოახდინოს სტრატეგიის ოპტიმიზაცია მისი განსჯის და მიმდებარე კვანძებიდან მიღებული ინფორმაციის საფუძველზე, რაც გამოიწვევს არაოპტიმალურ მახასიათებლებს ცენტრალიზებულ სისტემებთან შედარებით.

4.5. მომავალი კვლევითი მიმართულებები

მიუხედავად იმისა, რომ BACN 6G კომუნიკაციაში საინტერესო სფეროა კვლევითი საზოგადოებისთვის, კვლევა რეტრანსლატორების შერჩევისა და ბუფერული მდგომარეობის მართვის შესახებ BACN-ში ჯერ კიდევ ადრეულ ეტაპზეა. აქედან გამომდინარე, ჩვენ გამოვყოფთ 6G ქსელში ბუფერული რეტრანსლატორების რეალიზაციასთან დაკავშირებულ ძირითად კვლევით მიმართულებებს.

DL-ის გამოყენების პერსპექტივა BACN-ში:

მომავალში DL მნიშვნელოვან როლს ითამაშებს BACN-ში. თუმცა, იმის გათვალისწინებით, რომ BACN იქნება უფრო რთული (რაც განპირობებულია დიდი რაოდენობის სარელეო კვანძებით, სარელეო კვანძების გამოჩენით და გაქრობით და ენერჯის მოხმარებით რეტრანსლატორებში) და დაკავშირებული იქნება სხვა ისეთ ქსელებთან, როგორცაა თანამგზავრული ქსელები მომავალ 6G კომუნიკაციებში, არსებულ DL მეთოდებს შეიძლება შეექმნათ პრობლემები ტრენინგის მონაცემთა ნაკრების გენერირების მაღალი სირთულის და სისტემის გადაწყვეტილების დიდი სივრცის გამო. ამრიგად, DL-ის ზოგიერთი სხვა ახალი ტექნოლოგია შეიძლება დაინერგოს BACN-ში, რათა გადაჭრას ეს გამოწვევები მომავალი 6G ქსელებისთვის. მაგალითად, ორკასკადიანი კორელაციური ქსელი შეიძლება გამოყენებულ იქნეს ტრენინგის მონაცემთა ნაკრების შესაქმნელად, ხოლო ვოლპერტინგერის (Wolpertinger) არქიტექტურა შექმნილია MDP-ებში მოქმედებების შეფასების ღირებულების შესამცირებლად. იმის გათვალისწინებით, რომ BACN სცენარები შეიძლება შეიცვალოს რეალურ საკომუნიკაციო სისტემებში, მეტაევრისტიკული ალგორითმი შეიძლება გამოყენებულ იქნეს BACN-ისთვის ადაპტიური გადაწყვეტილების სისტემის შესაქმნელად. მას შეუძლია DL-ზე დაფუძნებული გადაწყვეტილების სისტემების შემდგომი ოპტიმიზაცია, რაც შეიძლება გამოყენებულ იქნეს კონვერგენციის მახასიათებლის გასაუმჯობესებლად სხვადასხვა BACN სცენარისთვის.

რეკონფიგურირებად ინტელექტუალურ ზედაპირებთან თავსებადობა:

ბუფერზე დაფუძნებული რეტრანსლატორი ასევე შეუძლია დაუკავშირდეს RIS-ს უსადენო გარემოს გასაუმჯობესებლად. თანდაყოლილი პრობლემების გარდა, არსებობს რამდენიმე ახალი გამოწვევა. უპირველეს ყოვლისა, ისეთი რთული არხების მუშაობის ოპტიმიზაცია, როგორცაა ჰიბრიდულ უსადენო ქსელებში და ასევე RIS-ის სინთეზი ბუფერულ რეტრანსლატორებთან წარმოადგენს კრიტიკულ საკითხს, რომელიც უნდა გადაიჭრას. მეორეც, RIS-ისთვის პასიური სხივის ფორმირების ელემენტების დაპროექტება შეიცავს ამპლიტუდისა და ფაზური წანაცვლების დისკრეტულ დონეებს, სადაც ამომწურავი ძეხნა არაპრაქტიკულია; ზოგიერთი არსებული ნაშრომი ასუსტებს შეზღუდვებს და იღებს სავარაუდო მნიშვნელობებს. თუმცა, ამან შეიძლება გამოიწვიოს მახასიათებლების

გაუარესება დაქვანტვის შეცდომების გამო. მახასიათებლების შემდგომი გაუმჯობესების მიზნით, ალტერნატიული ევრისტიკული ოპტიმიზაციის ტექნიკა შეიძლება გამოყენებულ იქნეს დისკრეტული ამპლიტუდის/ფაზური წანაცვლების მნიშვნელობების იტერაციების გზით ოპტიმიზაციისთვის. შემდეგ ოპტიმიზებული შედეგები შეიძლება გამოყენებულ იქნეს როგორც ტრენინგის მონაცემთა ნაკრები DL-ზე დაფუძნებული მეთოდებისთვის. გარდა ამისა, ევრისტიკული ალტერნატიული ოპტიმიზაციისა და DRL მეთოდების გამოთვლითი სირთულის გათვალისწინებით პრაქტიკული RIS კოფიციენტებისთვის, ორკასკადიანი კორელაციური ქსელი შეიძლება დაინერგოს კონვერგენციის დაჩქარებისა და მახასიათებლების გასაუმჯობესებლად.

უსაფრთხოებისა და კონფიდენციალურობის საკითხები:

რაც შეეხება ბუფერულ რეტრანსლაციას, არსებული სამუშაოები ძირითადად ფოკუსირებულია მონაცემთა ეფექტიან გადაცემაზე. თუმცა, ზოგიერთმა მომხმარებელმა შეიძლება უარი თქვას სხვა მომხმარებლების დახმარებაზე ინფორმაციის გადაცემაში კონფიდენციალურობის გამო, სენსიტიური ინფორმაციის გაზიარებისას, რამაც შეიძლება გამოიწვიოს უსაფრთხოების დარღვევა ან ტექნიკური სირთულეები. კონფიდენციალურობამ და უსაფრთხოებამ სულ უფრო მზარდი ყურადღება მიიპყრო 5G/6G-ში, სადაც მომხმარებლებს შეიძლება დასჭირდეთ მათი მონაცემების ადგილობრივად შენახვა და მოსმენის თავიდან აცილება. მეორე მხრივ, ხელოვნური ინტელექტის მოდელში ტრენინგის მონაცემები ასევე დაკავშირებული უნდა იყოს უსაფრთხოების საკითხებთან. ამიტომ, განაწილებული სწავლების არქიტექტურა სრულად უნდა ითვალისწინებდეს მონაცემთა განაწილებას, გამოთვლის შესაძლებლობებს და ქსელს. ამრიგად, მანქანურ სწავლებაზე ორიენტირებული ბუფერზე დაფუძნებული რეტრანსლატორის შერჩევის ახალ მეთოდებს, როგორცაა ფედერირებული სწავლება, შეუძლია უზრუნველყოს სასურველი კონფიდენციალურობა და უსაფრთხოება. უფრო მეტიც, DL ალგორითმები შეიძლება გამოყენებულ იქნეს მონაცემთა დასაშიფრად, რაც ართულებს თავდამსხმელებისთვის BACN-ში სენსიტიური ინფორმაციის აღმოჩენას და მასზე წვდომას.

ბუფერისა და ქეშ-მეხსიერების შერჩევის ჰიბრიდული ტექნოლოგია:

უსადენო ქეშირება არის განვითარებადი ტექნოლოგია, რომელიც საშუალებას იძლევა საჭირო მონაცემები ქეშ-მეხსიერებებში იქნეს შენახული, რაც, თავის მხრივ, საშუალებას იძლევა განხორციელდეს მონაცემთა პირდაპირი გადაცემა ქეშ-მეხსიერიდან დისტანციური ღრუბლის ნაცვლად. გადაცემის ეფექტიანობის შემდგომი გაუმჯობესების მიზნით, ბუფერზე დაფუძნებული შერჩევის სქემები შეიძლება გაერთიანდეს ქეშირების ტექნოლოგიასთან. DL ალგორითმები შეიძლება გამოყენებულ იქნეს ჰიბრიდულ ბუფერზე დაფუძნებულ და ქეშირების სქემების დიზაინთან და ოპტიმიზაციასთან დაკავშირებული ისეთი გამოწვევების მოსაგვარებლად, რომლებსაც რთული და მაღალგანზომილებიანი მონაცემების დამუშავების უნარი და მონაცემთა შიგნით რთული შაბლონებისა და ურთიერთობების აღმოჩენის პოტენციალი აქვთ. DL ალგორითმების გამოყენებით, ჰიბრიდული ბუფერისა და ქეშირების შერჩევის ტექნოლოგიები შეიძლება შემუშავდეს 6G ქსელებში კონტენტის მიწოდების ეფექტიანად სამართავად.

BACN მომავალ 6G SAGIN სისტემებში:

ვინაიდან კოსმოსურ-საჰაერო-სახმელეთო ინტეგრირებული ქსელები (SAGIN) შეიცავს უამრავ დინამიკურ კვანძს მათი ფართო კავშირის დიაპაზონში, ბუფერზე დაფუძნებული რეტრანსლატორების შერჩევა SAGIN-ში იმედისმომცემია. პირველ რიგში, სამომავლო მიმართულებაა დინამიკური მოწყობილობების კომპლექსურ ბუნებასთან გამკლავება, როგორცაა UAV-ები და თანამგზავრები SAGIN-ში. ზოგიერთ არსებულ ნაშრომში გამოკვლეულია ბუფერული რეტრანსლაცია UAV-ების ქსელებში. თუმცა, UAV-ის სწრაფი მობილურობის საკითხი გადაუჭრელი რჩება. მეორე მხრივ, კომუნიკაციის შეზღუდვები სხვადასხვა საკომუნიკაციო ფენას შორის და მაღალი ზედნადები ხარჯები, რომლებიც გამოწვეულია SAGIN-ის უაღრესად დინამიკური გარემოთი, კრიტიკული გამოწვევებია. გარდა ამისა, რთულია არხის შესახებ ინფორმაციის მიღება დინამიკურ გარემოში. ამრიგად, DL-ზე ორიენტირებულ მეთოდებს, როგორცაა ინვერსიული RL ბუფერული რეტრანსლაციისთვის, შეუძლია ეფექტიანად მართოს რესურსების და არხების მოპოვება მოქნილი SAGIN-ისთვის.

4.6. მეოთხე თავის დასკვნა

ამ თავში წარმოდგენილია ბუფერზე დაფუძნებული კოოპერატიული სისტემების რეტრანსლატორების შერჩევის არსებული სქემების გაუმჯობესების პოტენციური ახალი მოთხოვნების დასაკმაყოფილებლად. იდენტიფიცირებულია BACN-ების გამოყენებასთან დაკავშირებული გამოწვევები და ნაჩვენებია DL-ზე დაფუძნებული გადაწყვეტილებები, რომლებიც გვპირდებიან ამ პრობლემების მოგვარებას. ულტრა საიმედო და დაბალი შეყოვნების მოთხოვნებით უსაფრთხო კომუნიკაციის მისაღწევად, არასრულყოფილი CSI-სა და ჰეტეროგენული ქსელების გათვალისწინებით, მნიშვნელოვანია BACN-ის შინაგანი მახასიათებლების მკაფიო ფორმით გაგება და შესწავლა. ჩვეულებრივი ბუფერული რეტრანსლატორების შერჩევის მეთოდებისგან განსხვავებით, ამ თავში შემოთავაზებულია DL-ზე ორიენტირებული მიდგომები, რომლებიც გამოიყენება როგორც ცენტრალიზებულ, ასევე დეცენტრალიზებულ სცენარებში. კომპიუტერულ სიმულაციურ შედეგებზე დაყრდნობით DRL-ის ეფექტიანობა ნაჩვენებია ცენტრალიზებულ ბუფერზე დაფუძნებულ NTN-ში, ისევე როგორც წარმოდგენილია მრავალაგენტიანი DRL ალგორითმის ეფექტიანობა დეცენტრალიზებულ NTN-ში. შემოთავაზებულ DL-ზე ორიენტირებულ მეთოდებს აქვთ განზოგადების პოტენციური სხვადასხვა BACN-ში და გვთავაზობენ ღირებულ შეხედულებებს მომავალი კვლევითი მიმართულებებისთვის.

თავი 5 AI-ზე დაფუძნებული ქსელის ფენაზე დაყოფა 6G ქსელისთვის

5.1. შესავალი

ჩინა თავებში უკვე აღვნიშნეთ, რომ არსებულ უსადენო ქსელთან შედარებით (5G-ის ჩათვლით), 6G შემოგვთავაზებს უფრო მეტს, ვიდრე ეს არის უბრალოდ ისეთი KPI მოთხოვნების გაუმჯობესება, როგორებიცაა მონაცემთა გადაცემის გაზრდილი სიჩქარე, ქსელის უფრო მაღალი გამტარუნარიანობა და დაბალი შეყოვნება. 6G ქსელს ექნება შემდეგი უნიკალური მახასიათებლები: პირველი, კოსმოსური ქსელები, როგორცაა დაფუძნებული დედამიწის დაბალი ორბიტის (LEO) თანამგზავრებზე, საჰაერო ქსელები (მაგალითად, UAV-ებით), სახმელეთო ქსელები ფიჭური საბაზო სადგურებით (BS) შევლენ გლობალური დაფარვის კოსმოსურ-საჰაერო-სახმელეთო ინტეგრირებულ ქსელში (SAGIN) მცისიერად, ანუ მოთხოვნისთანავე, მრავალი სერვისის უზრუნველსაყოფად. მეორე, რესურსების ვირტუალიზაციას ქსელის ფენაზე დაყოფის ტექნიკის გამოყენებით და საბოლოო მომხმარებლის ვირტუალიზაციას ციფრული ტყუპების ტექნიკის გამოყენებით, შეუძლია ხელი შეუწყოს ქსელის გაუმჯობესებულ ვირტუალიზაციას მისი მოქნილი მენეჯმენტის უზრუნველსაყოფად (განმარტავთ, რომ საბოლოო მომხმარებლის ვირტუალიზაცია არის მისი ვირტუალიზაცია ქსელის ექსპლუატაციაში და მენეჯმენტში მომხმარებელთა ქცევისა და სტატუსის დახასიათებით (მაგალითად, სერვისის მოთხოვნების და QoS-ის დაკმაყოფილება), რაც შეიძლება მიღწეული იყოს ციფრული ტყუპების კონცეფციების გამოყენებით). მესამე, ინტელექტი აღწევს ქსელების ყველა კუთხეში, დაწყებული საბოლოო მომხმარებლებიდან ქსელის პერიფერიის გავლით დისტანციურ ღრუბლამდე, რაც უზრუნველყოფს საყოველთაო ინტელექტს. ქსელის რიგი კვანძები აღჭურვილია ჩაშენებული AI ფუნქციებით, რითაც არა მხოლოდ ხელს უწყობს ქსელის ინტელექტუალურ მართვას, არამედ AI სერვისების განვითარებასაც (ამის მაგალითია, ღრმა ნეირონულ ქსელზე დაფუძნებული აპლიკაციები). აქედან გამომდინარე, მოსალოდნელია, რომ 6G ქსელები შექმნიან უსადენო ქსელის ახალ ეკოსისტემას, რომელიც მოიტანს სოციალურ და ეკონომიკურ სარგებელს.

6G ქსელები მხარს დაუჭერს სერვისების მრავალფეროვნებას სხვადასხვა QoS მოთხოვნით, ისეთებით როგორცაა მრავალსენსორული გაფართოებული რეალობა (XR) და ჰოლოგრამის ვიდეო ნაკადი. 5G ქსელებში დამკვიდრებული დივერსიფიცირებული სერვისების მხარდასაჭერად, ქსელის ფენაზე დაყოფა არის პოტენციური მიდგომა მრავალი ლოგიკურად იზოლირებული ვირტუალური ქსელის (ანუ ფენების) ასაგებად სხვადასხვა სერვისისთვის საერთო ფიზიკური ქსელის თავზე. სხვადასხვა სერვისის QoS მოთხოვნები შეიძლება გარანტირებული იყოს აღნიშნული ფენების მართვის ეკონომიური სტრატეგიებით, დაწყებული მომზადების, დაგეგმვისა და ექსპლუატაციის ფაზებიდან, ქსელის ფენაზე დაყოფის სასიცოცხლო ციკლში.

ქსელის დაყოფის სქემების შემუშავება მრავალი გამოწვევის წინაშე დგას 6G ქსელებში მათი უნიკალური მახასიათებლების გამო. პირველი, SAGIN-ში კოსმოსური, საჰაერო და სახმელეთო ქსელის სეგმენტების ფენების მართვა მოითხოვს ჰეტეროგენული ქსელის სეგმენტების გონივრულ კოორდინაციას. უფრო მეტიც, 6G ქსელებს სჭირდებათ სხვადასხვა სახის ახალი სერვისის მხარდაჭერა, მათი განსხვავებული და მკაცრი QoS მოთხოვნების დაკმაყოფილების დროს, რაც კიდევ უფრო ართულებს ფენების მართვას. აქედან გამომდინარე, გადამწყვეტი მნიშვნელობა აქვს 6G ქსელებში ფენების მართვის ინტელექტუალური გადაწყვეტილებების შემუშავებას. მეორე, მძლავრი გამოთვლითი შესაძლებლობებითა და მოწინავე ხელოვნური ინტელექტის ტექნოლოგიებით, ყველგან გავრცელებული ინტელექტი ხელს უწყობს ხელოვნური ინტელექტის მრავალ სერვისის ახალი QoS მოთხოვნებით, როგორცაა მონაცემთა ხარისხი, დასკვნის სიზუსტე და ტრენინგის შეყოვნება. ამიტომ, აუცილებელია პერსონალურად მორგებული ქსელის ფენების შექმნა 6G ქსელებში ახალი AI სერვისების მხარდასაჭერად.

ამ თავში ჩვენ წარმოვადგინეთ AI-ზე დაფუძნებული ქსელის დაყოფის არქიტექტურას 6G სისტემებისთვის, რათა ხელი შეუწყოს ქსელის ინტელექტუალურ მართვას და AI სერვისების მხარდაჭერას. AI-ზე დაფუძნებული ნიშნავს, რომ, როგორც ჩაშენებული კომპონენტი ქსელის დაყოფის არქიტექტურაში, AI არსებობს არა მხოლოდ

SDN-ის კონტროლერში ქსელის ფუნქციების მართვისთვის, არამედ ქსელის ფუნქციებში, როგორც სერვისები საბოლოო მომხმარებლებისთვის. მაშასადამე, AI-ისა და ქსელის დაყოფის სინერჯია შემოთავაზებულ არქიტექტურაში ორმაგია. ერთი მხრივ, AI ტექნოლოგიის გამოყენება შესაძლებელია ქსელის ფუნქციების მართვისთვის, კერძოდ, ხდება AI-ის გამოყენება ფუნქციების დაყოფისთვის. წარმოდგენილია ქსელის ფუნქციების დაყოფის სასიცოცხლო ციკლი, მომზადების, დაგეგმვისა და ექსპლუატაციის ფაზების ჩათვლით და თითოეული ფაზისთვის AI-ზე დაფუძნებული გადაწყვეტილებების მითითებით. გარდა ამისა, წარმოდგენილია საბოლოო მომხმარებლებს, წვდომის წერტილებსა და SDN კონტროლერს შორის ინფორმაციის გაცვლის დეტალური პროცედურა. მეორე მხრივ, ქსელის დაყოფა შეიძლება გამოყენებულ იქნეს პერსონალურად მორგებული ქსელის ფუნქციების შესაქმნელად სხვადასხვა AI სერვისისთვის, კერძოდ, განხილულია ფუნქციების დაყოფის გამოყენება AI-ისთვის. განხილულია ისეთი პოტენციური მიდგომები, როგორცაა AI მაგალითების შემუშავება და რესურსების ეფექტური მართვა AI სერვისებისთვის.

ამ თავის დარჩენილი ნაწილი ორგანიზებულია შემდეგნაირად. განიხილება 6G ქსელების მოსალოდნელი მახასიათებლები და შემდეგ შემოთავაზებულია AI-ზე დაფუძნებული ქსელის დაყოფის არქიტექტურა. წარმოდგენილია, შესაბამისად, AI ფუნქციების დაყოფისთვის და ფუნქციების დაყოფა AI-ისთვის. შესწავლილია კონკრეტული გამოყენების შემთხვევა. გამოკვეთილია მომავალი კვლევითი მიმართულებები, რასაც მოჰყვება შესაბამისი დასკვნა.

5.2. AI-ზე დაფუძნებული ქსელის ფუნქციების დაყოფა

ქსელის ფუნქციების დაყოფა განვითარებადი ტექნოლოგიაა, რომელიც ხელს უწყობს დივერსიფიცირებულ აპლიკაციებს ხარჯთაღრიცხვითი გზით. ქსელის დაყოფის კონცეფცია წარმოიშვა 1980-იანი წლების ბოლოს. დღესდღეობით, ქსელის დაყოფა ძირითადი ტექნოლოგიაა 5G ქსელებში, რომელსაც მხარს უჭერს NFV და SDN ტექნიკა. კერძოდ, NFV იძლევა ვირტუალიზებულ რესურსებს და ქსელის ფუნქციებს, რესურსების მოქნილი მართვისთვის, ხოლო SDN ხელს უწყობს ქსელის ცენტრალიზებულ მართვას ქსელის ოპტიმიზაციისთვის. 5G ქსელებში, ქსელის დაყოფა განსაზღვრულია 3GPP-ის მე-15 გამოშვებაში. უფრო მეტიც, მომავალ 6G ქსელებში, ქსელის დაყოფა გააგრძელებს განვითარებას და უფრო მნიშვნელოვან როლს შეასრულებს.

ქსელის დაყოფის ძირითადი იდეა არის მრავალი ლოგიკურად იზოლირებული ქსელის შექმნა საერთო ფიზიკური ინფრასტრუქტურის თავზე, რომელსაც შეუძლია მიაღწიოს ქსელის მოქნილ და ადაპტირებულ მართვას. მისი სარგებელი სამმაგია:

1. მრავალჯერადი ქირავნობა: მრავალ ვირტუალურ ქსელს შეუძლია საერთო ფიზიკური ინფრასტრუქტურის გაზიარება, რაც ამცირებს კაპიტალურ ხარჯებს ქსელის განლაგებისთვის.

2. სერვისის იზოლაცია: მრავალი ფუნქციონირება სხვადასხვა სერვისისთვის რესურსების ჭკვიანური მენეჯმენტის საშუალებით, რაც ეფექტიანად უზრუნველყოფს მომსახურების დონის შეთანხმებებს სხვადასხვა ფუნქციისთვის.

3. მოქნილობა: ქსელის ფუნქციების დაყოფას შეუძლია ქსელის მოქნილი მენეჯმენტის მხარდაჭერა, რადგან მოთხოვნისამებრ შეიძლება შეიქმნას, შეიცვალოს ან წაიშალოს ფუნქციები.

5G-დან 6G-მდე, ზოგადად მოსალოდნელია, რომ KPI მოთხოვნები გაიზარდება მინიმუმ ათჯერ. ბოლო ნაშრომების მიხედვით, 6G ქსელების KPI მოთხოვნები მოიცავს 1 ტბიტი/წმ მონაცემთა გადაცემის პიკურ სიჩქარეს, 20-100 გბიტი/წმ მომხმარებლის გამოცდილებაზე დაფუძნებულ მონაცემთა გადაცემის სიჩქარეს, 0.1 მწმ-მდე შეყოვნებას, 10 მილიონ მოწყობილობას კვადრატულ კილომეტრზე და პრაქტიკულად 100 პროცენტთან დაფარვას. ასეთი KPI მოთხოვნები საჭიროებს რამდენიმე კანდიდატ ტექნოლოგიას, როგორცაა ტერაჰერცული (ტჰც) კომუნიკაციები და AI.

3GPP სამუშაო ჯგუფი განიხილავს 5G-ის კანდიდატ ტექნოლოგიებს 2026 წლის ბოლოსთვის, ხოლო 6G სტანდარტის პირველი დოკუმენტაცია, სავარაუდოდ, 2030 წლისთვის იქნება.

5G ქსელებისგან განსხვავებით, 6G ქსელებს აქვთ რამდენიმე გამორჩეული მახასიათებელი:

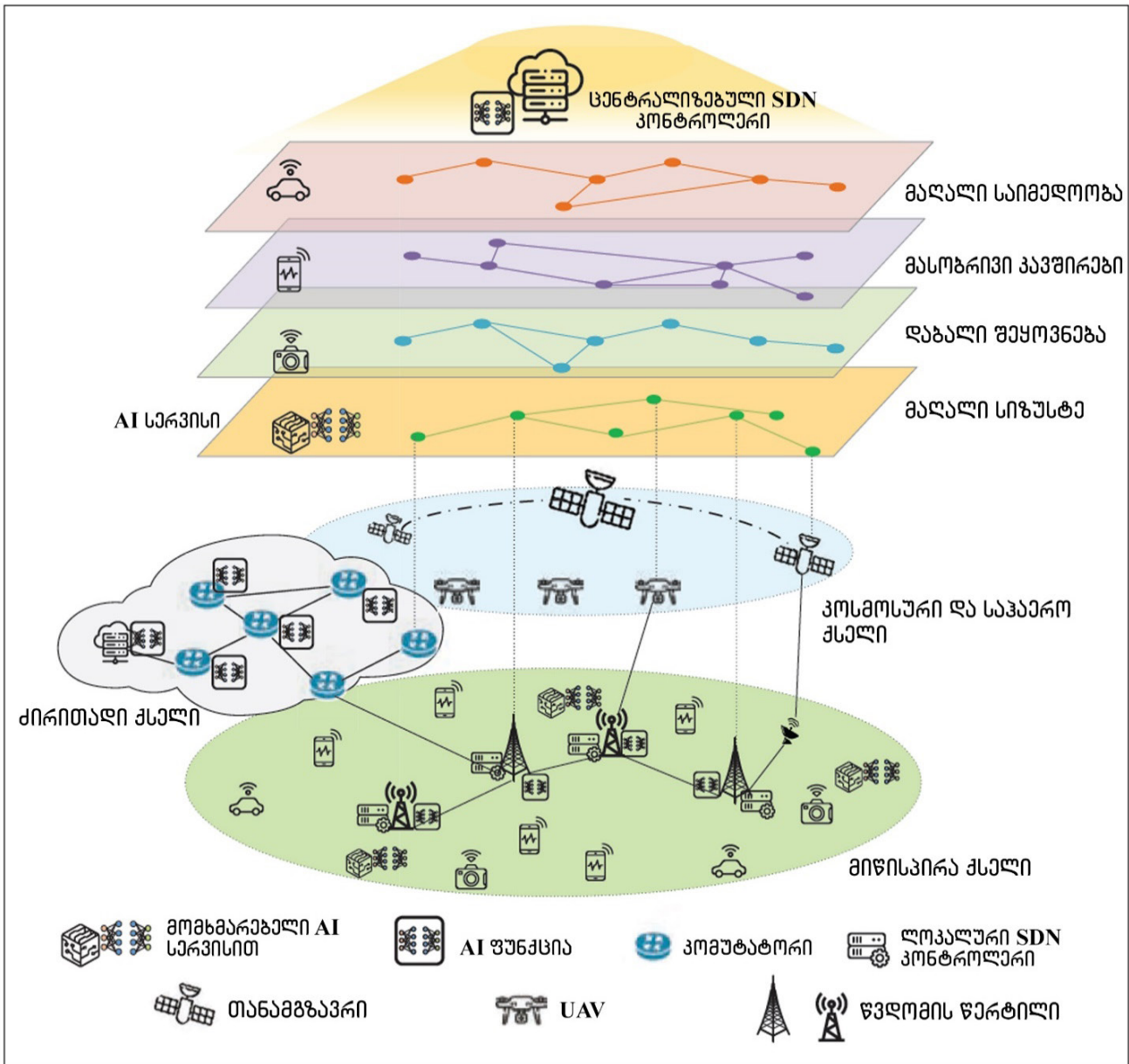
ა) SAGIN: მიუხედავად იმისა, რომ ამჟამინდელი მიწისპირა ქსელები უზრუნველყოფენ კარგ დაფარვას მჭიდროდ დასახლებულ რაიონებში, 6G-მ უნდა უზრუნველყოს უნივერსალური დაფარვა, მათ შორის სოფლად, შორეულ და იშვიათად დასახლებულ რეგიონებში. ამ მიზნის მისაღწევად, 6G გამოიყენებს სიმაღლის განზომილებას. კოსმოსური, საჰაერო და სახმელეთო ქსელის სეგმენტები ინტეგრირებულია SAGIN-ში, რომელსაც შეუძლია უზრუნველყოს გლობალური დაფარვა, ხელი შეუწყოს მოთხოვნილ მომსახურებას და მხარი დაუჭიროს მაღალსიჩქარიან, დაბალი შეყოვნების სერვისებს.

ბ) სერვისების მრავალფეროვნება: ბევრ სერვისს აქვს მკაცრი QoS მოთხოვნები სხვადასხვა ასპექტში. მობილური VR-ის და ჰოლოგრამის ვიდეო ნაკადის აპლიკაციები საჭიროებენ მონაცემთა გადაცემის მაღალ სიჩქარეს; მაგალითად, მობილური VR-ის მონაცემთა გადაცემის სიჩქარე აპლიკაციაში 5 გბიტ/წმ-მდეა. სხვა აპლიკაციებმა შეიძლება მოითხოვოს ულტრა მაღალი საიმედოობა, როგორცაა ავტონომიური მართვა, სამრეწველო კონტროლის სისტემები და რობოტების/UAV-ების გუნდები; მაგალითად, ავტონომიური მართვის საჭირო საიმედოობა 99.999 პროცენტამდეა.

გ) საყოველთაო ინტელექტი: ქეშირების შესაძლებლობით, დიდი რაოდენობით მონაცემების შენახვა შესაძლებელია ქსელში. გარდა ამისა, AI ტექნიკის, პერიფერიული გამოთვლებისა და მოწყობილობების გამოთვლების განვითარებით, ინტელექტი დისტანციური ღრუბლიდან გადადის ქსელის პერიფერიაზე და საბოლოო მომხმარებლებამდე. ამ გზით, AI ინტეგრირებული იქნება 6G ქსელებში ქსელის ჰიერარქიული მართვისთვის, ქსელში არსებული უზარმაზარი რაოდენობის მონაცემების უშუალო შესწავლით. უფრო მეტიც, საყოველთაო ინტელექტი ხელს შეუწყობს AI-ის მთელი რიგი სერვისების განვითარებას, რომლებშიც AI უზრუნველყოფილია როგორც სერვისი.

ეს ფუნქციები აჩენს ახალ გამოწვევებს 6G ქსელებისთვის ქსელის დაყოფის სქემების შემუშავებაში. პირველი, SAGIN არა მხოლოდ ზრდის ინტეგრირებული ქსელის სეგმენტების რაოდენობას, არამედ შემოაქვს დამატებითი დინამიკა ქსელის რესურსების ხელმისაწვდომობაზე თანამგზავრების მობილურობისა და UAV-ების მანევრირების გამო. ქსელის დაყოფის სქემები უნდა მოიცავდეს ფართომასშტაბიან SAGIN-ს, რესურსის დინამიკური ხელმისაწვდომობის გათვალისწინებით. უფრო მეტიც, მრავალფეროვანი სერვისების მხარდაჭერა მკაცრი QoS მოთხოვნებით კიდევ უფრო ართულებს ქსელის დაყოფის სქემის დიზაინს. მეორე, საყოველთაო ინტელექტი ხელს უწყობს მრავალ განვითარებად AI სერვისს, რომელიც გავრცელებული იქნება 6G ქსელებში. ჩვეულებრივი სერვისებისგან განსხვავებით, AI სერვისების მხარდაჭერა მოითხოვს მრავალ ნაბიჯს, მათ შორის მაღალი ხარისხის მონაცემთა ნიმუშების შეგროვებას, დამაკმაყოფილებელი AI მოდელების მომზადებას და მოდელის შემუშავებას დაბალი შეყოვნებით, რომელიც უნდა აკმაყოფილებდეს სხვადასხვა QoS მოთხოვნას. როგორც დავაკმაყოფილოთ ასეთი მრავალფეროვანი QoS მოთხოვნები AI სერვისებისთვის, კვლავ გამოწვევად რჩება.

ზემოაღნიშნული პრობლემების გადასაჭრელად, ნახ. 5.1-ზე წარმოდგენილია AI-ზე დაფუძნებული ქსელის დაყოფის არქიტექტურა 6G სისტემებისთვის. ეს არქიტექტურა მიზნად ისახავს SAGIN-ისა და საყოველთაო ინტელექტის ინტეგრირებას და სხვადასხვა სერვისის მხარდაჭერას მკაცრი QoS მოთხოვნებით. 5G ქსელების ქსელის დაყოფასთან შედარებით, შემოთავაზებულ არქიტექტურას ორი ახალი მახასიათებელი აქვს. პირველი მიდგომით, AI ინტეგრირებულია SDN კონტროლერებში, რათა განხორციელდეს ინტელექტუალური ქსელის დაყოფა ისე, რომ ქსელის მრავალი ფენა მკაცრი QoS მოთხოვნებით შეიძლება ეფექტიანად და ეკონომიურად იმართებოდეს AI ტექნიკის საშუალებით. ამ მიდგომას ეწოდება AI-ის გამოყენება ფენებზე დაყოფისთვის. მეორე მიდგომის მიხედვით, ახალი AI სერვისები მხარდაჭერილია ქსელის დაყოფით. გარდა ჩვეულებრივი სერვისების ქსელის ფენებისა, AI სერვისებისთვის ახალი ქსელის ფენები აგებულია საერთო ფიზიკური ინფრასტრუქტურის თავზე. ასეთ მიდგომას მოიხსენიებენ, როგორც ფენებზე დაყოფას AI-ისთვის.



ნახ. 5.1. AI-ზე დაფუძნებული ქსელის დაყოფის არქიტექტურა 6G ქსელებისთვის

შემოთავაზებულ არქიტექტურაში განლაგებულია SDN კონტროლერის ორი ტიპი. ერთი არის ცენტრალიზებული SDN კონტროლერი, რომელიც მდებარეობს ღრუბელში და მართავს ქსელის ფენებს; ხოლო მეორე არის ადგილობრივი SDN კონტროლერი, რომელიც მდებარეობს წვდომის წერტილებში და შექმნილია ქსელის თითოეულ ფენაზე საბოლოო მომხმარებლებისთვის რესურსების დასაგეგმად. ქვემოთ მოყვანილი SDN კონტროლერი ეხება ცენტრალიზებულ SDN კონტროლერს, თუ სხვა რამ არ არის მითითებული. შემდეგში ჩვენ წარმოვადგენთ AI-ის გამოყენების ძირითად იდეებს, შესაბამისად, AI-ის ფენებად დაყოფისთვის და ფენებად დაყოფას AI-ისთვის.

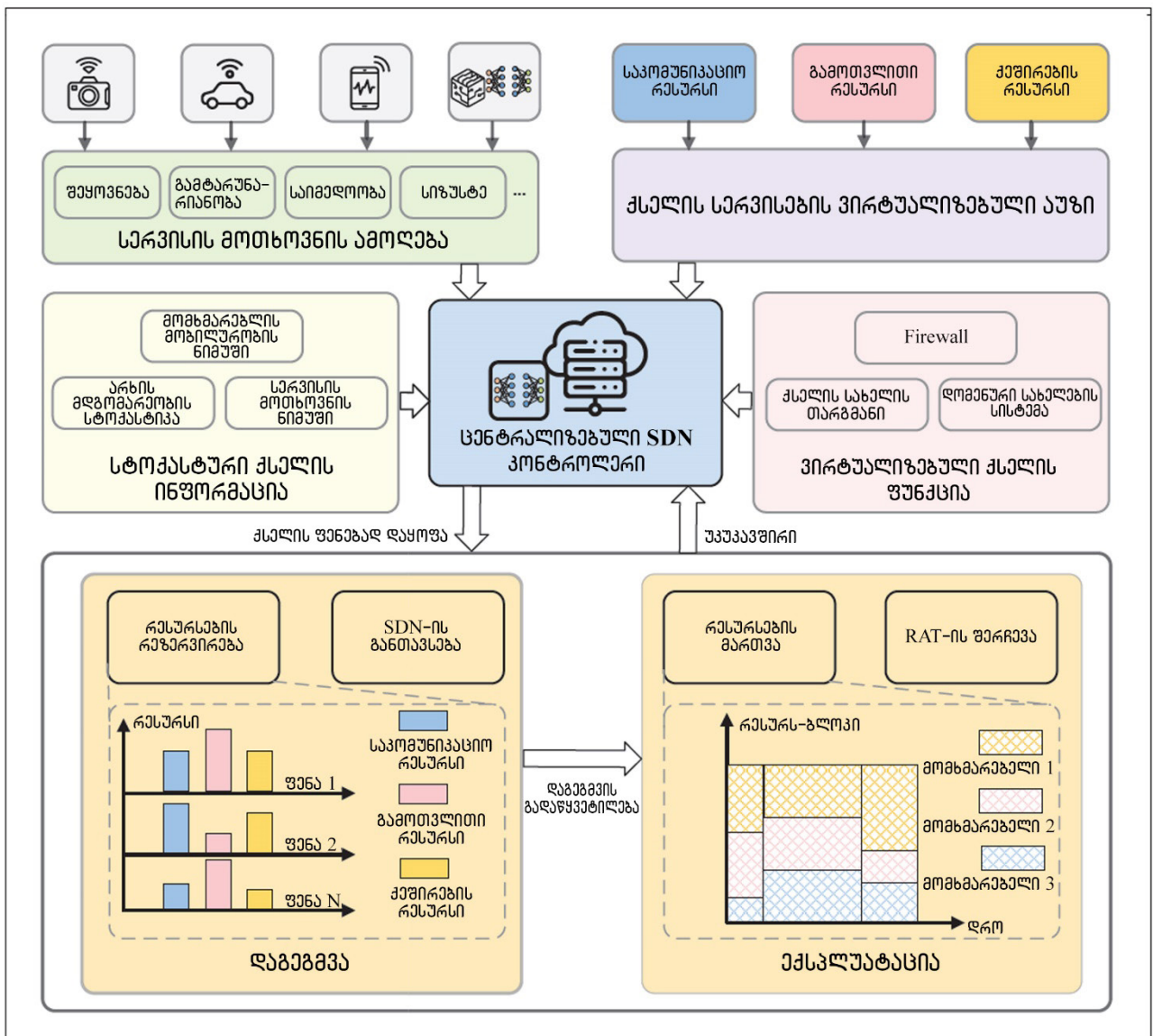
5.3. AI ფენებად დაყოფისთვის

ამ პარაგრაფში ჩვენ სამ ფაზად წარმოვადგენთ ქსელის ფენებად დაყოფის სასიცოცხლო ციკლს და შემდეგ გამოვიკვლევთ AI-ის პოტენციურ გადაწყვეტილებებს თითოეული ფაზისთვის. ამის შემდეგ განიხილება AI-ში ინფორმაციის გაცვლის შესაბამისი პროცედურა, რომელიც გამოიყენება დაყოფისთვის.

როგორც ნაჩვენებია ნახ. 5.2-ზე ქსელის ფენებად დაყოფის სასიცოცხლო ციკლი შედგება სამი ეტაპისგან: მომზადება, დაგეგმვა და ექსპლუატაცია. აღვნიშნავთ, რომ ცენტრალიზებული SDN კონტროლერი პასუხისმგებელია მომზადებისა და დაგეგმვის ფაზებზე, ხოლო ექსპლუატაციის ფაზას კოორდინაციას უწევენ ადგილობრივი SDN კონტროლერები.

მომზადების ფაზა:

ამ ფაზაში ხდება ქსელის ფენების აგება და კონფიგურაცია, რომელიც ეფუძნება სერვისის მოთხოვნებს, მონაცემთა ტრაფიკს, მომხმარებლის ინფორმაციას და ვირტუალური ქსელის რესურსების ხელმისაწვდომობას. მიზნის მისაღწევად, SDN კონტროლერი ასრულებს შემდეგ დავალებებს:



ნახ. 5.2. ქსელის ფენებად დაყოფის სასიცოცხლო ციკლი, რომელიც შედგება სამი ეტაპისგან: მომზადება, დაგეგმვა და ექსპლუატაცია

- სერვისის მოთხოვნის ამოღება: ამ ამოცანის მიზანია სერვისების კლასიფიკაცია მათი იმ QoS მოთხოვნების მიღებით, როგორცაა მომსახურების შეყოვნება, სერვისის პრიორიტეტი, გამტარუნარიანობა და საიმედოობა. 3GPP-ის აქვს სტანდარტიზებული სპეციფიკური სერვისის/ფენის ტიპის ისეთი მნიშვნელობები კლასიფიცირებული სერვისებისთვის, როგორცაა გაუმჯობესებული მობილური ფართობოლოვანი ქსელი, ულტრა საიმედო დაბალი შეყოვნების კომუნიკაციები და მანქანური ტიპის მასობრივი საკომუნიკაციო სერვისები.
- ქსელის რესურსებისა და ფუნქციების ვირტუალიზაცია: ქსელის რესურსები, როგორცაა კომუნიკაცია, გამოთვლები და ქეშირების რესურსები, გაერთიანებულია ვირტუალიზებულ რესურს-ბლოკებში ვირტუალიზაციის მოწინავე ტექნიკის საშუალებით. ანალოგიურად, ქსელის ფუნქციები, როგორცაა firewall, ქსელის სახელის გარჩევადობა, დომენის სახელების სისტემა, განცალკევებულია გამოყოფილი აპარატურის ქსელის ფუნქციებიდან VNF-ებად. ვირტუალიზაციის საშუალებით SDN კონტროლერს შეუძლია მოქნილად მართოს ქსელის რესურსები და ფუნქციები.

ამ ამოცანების დასრულების შემდეგ, SDN კონტროლერს შეუძლია შექმნას ქსელის ფენები თითოეული მიღებული მოთხოვნისთვის.

დაგეგმვის ფაზა:

ეს ფაზა მიზნად ისახავს ქსელის რესურსების რეზერვირებას ფენებისთვის, რათა უზრუნველყოს სერვისი. დაგეგმვის ფაზა ოპერირებს გრძელვადიან პერიოდში. დრო დაყოფილია რამდენიმე დაგეგმვის პერიოდად (ფანჯრებად) თითოეული ფენისთვის. თითოეული დაგეგმვის ფანჯრის ხანგრძლივობა დამოკიდებულია სერვისის მოთხოვნასა და ქსელის დინამიკაზე, რომელიც მერყეობს რამდენიმე წუთიდან რამდენიმე საათამდე. მიზნის მისაღწევად, დაგეგმვის ფაზაში სრულდება შემდეგი ორი ნაბიჯი:

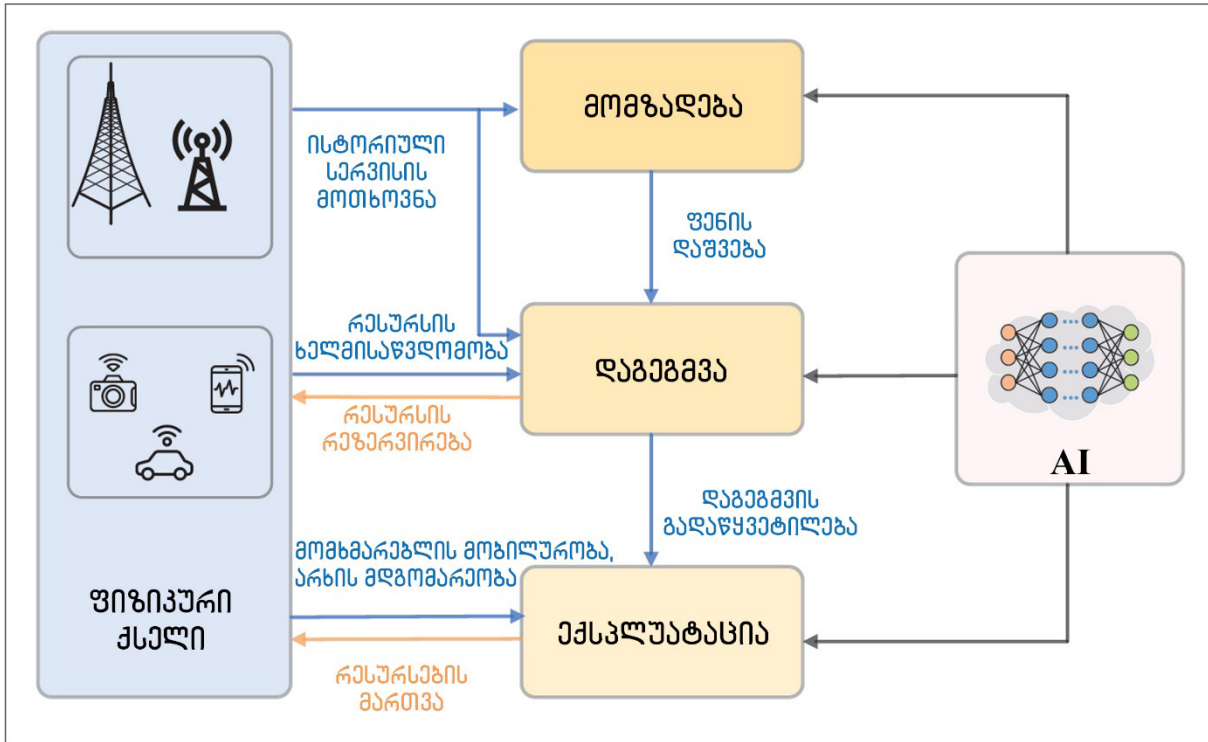
- ინფორმაციის შეგროვება სერვისისა და ქსელის შესახებ: SDN კონტროლერის გლობალური მართვის ფუნქციების წყალობით, შესაძლებელია ქსელის შესახებ ვრცელი ინფორმაციის შეგროვება ძირითადი ფიზიკური ქსელებიდან, როგორცაა სერვისის მოთხოვნები, არხის სტოქასტური პირობები და მომხმარებლის მობილურობის ნიმუშები. შეგროვებული ინფორმაცია გამოიყენება რესურსების რეზერვირების (დაჯავშნის) შესახებ შემდგომი გადაწყვეტილების მისაღებად.
- რესურსების რეზერვირება: ყოველი დაგეგმვის ფანჯრის დასაწყისში, SDN კონტროლერი არეგულირებს ქსელის რეზერვირებული რესურსების რაოდენობას თითოეული ფენისთვის, ფენის მახასიათებლის მონიტორინგზე დაყრდნობით. თითოეული ფენისთვის რეზერვირებული ვირტუალიზებული ქსელის რესურსები აისახება ფიზიკურ ქსელში. ყოველი დაგეგმვის ფანჯრის ბოლოს, სისტემის გარკვეული ინფორმაცია უბრუნდება SDN კონტროლერს, როგორცაა რესურსების გამოყენება, სისტემის მუშაობის მახასიათებელი და სერვისის დონის შეთანხმების დაკმაყოფილება. უკუკავშირის ინფორმაციაზე დაყრდნობით, SDN კონტროლერს შეუძლია რესურსების რეზერვირების გადაწყვეტილებების კორექტირება, რათა მოერგოს ქსელის დინამიკურ გარემოს და უზრუნველყოს QoS მოთხოვნები.

ექსპლუატაციის ფაზა:

ამ ფაზის განმავლობაში დაგეგმილია ფენების მომსახურება ხელმომწერი საბოლოო მომხმარებლებისთვის რეზერვირებული რესურსების გამოყენებით. ექსპლუატაციის ფაზას სჭირდება ბევრად უფრო მოკლე დრო (მაგალითად, 100 მწმ), ვიდრე დაგეგმვის ფაზას. კონკრეტულად, ცენტრალიზებული SDN კონტროლერის კოორდინაციით, ადგილობრივი SDN კონტროლერები ანაწილებენ ქსელის რესურსებს საბოლოო მომხმარებლების თითოეულ ფენაზე, რეალურ დროში მათი მონაცემთა ტრაფიკის მიხედვით. ამ ფაზაში მიღებული გადაწყვეტილებები მოიცავს რადიოწვდომის ტექნოლოგიის (RAT) შერჩევას, მომხმარებლის ასოციაციის განსაზღვრას კონკრეტულ რადიოწვდომის წერტილებთან, სათანადო პროტოკოლისა და მასთან დაკავშირებული პარამეტრების არჩევას და რესურსების მართვას საბოლოო მომხმარებლებს შორის.

AI-ის როლი ქსელის ფენებად დაყოფაში:

მიუხედავად იმისა, რომ ქსელის დაყოფამ შეიძლება ხელი შეუწყოს სერვისების უზრუნველყოფას, ქსელის მრავალი ფენის მართვა მნიშვნელოვან ხარჯებს იწვევს ქსელის მართვის პროცესში, განსაკუთრებით 6G ქსელებში. როგორც ნაჩვენებია ნახ. 5.3-ზე, AI-ზე დაფუძნებული ქსელის ფენებად დაყოფა არის პოტენციური გადაწყვეტა, რომელშიც AI თამაშობს სხვადასხვა როლს ქსელის დაყოფის სხვადასხვა ფაზაში.



ნახ. 5.3. განხილული AI-ზე დაფუძნებული ქსელის ფენებად დაყოფის გადაწყვეტა, სადაც AI თამაშობს განსხვავებულ როლს მომზადების, დაგეგმვისა და ექსპლუატაციის ფაზებში

მომზადების ფაზაში AI-ის სჭირდება შეასრულოს ორი დავალება:

1. სერვისის მოთხოვნის პროგნოზირება: ისტორიულ მონაცემებზე დაყრდნობით, სერვისებზე მოთხოვნის პროგნოზირება შესაძლებელია AI-ის ისეთი ტექნიკის საშუალებით, როგორცაა რეკურენტული ნეირონული ქსელები. კვლევებმა აჩვენა, რომ სერვისზე მოთხოვნილება და რესურსების გამოყენება შესაძლებელია ზუსტად იყოს წინასწარ განსაზღვრული. პროგნოზირების შედეგები შეიძლება გამოყენებულ იქნეს გადაწყვეტილების მისაღებად დაგეგმვის ფაზაში.
2. ფენის დაშვება: SDN კონტროლერი საშუალებას აძლევს ფენას, მაქსიმალურად გამოიყენოს ქსელის რესურსები მათი ხელმისაწვდომობისა და სერვისის მოთხოვნების საფუძველზე. ვინაიდან ფენის დაშვების გადაწყვეტილება ორობითია, ეს პრობლემა განიხილება, როგორც მთელი რიცხვის ოპტიმიზაციის პრობლემა. ფართომასშტაბიან ქსელებში, რესურსების ხელმისაწვდომობის კომპლექსური განაწილებით, ჩვეულებრივი ოპტიმიზაციის გადაწყვეტილებები ხდება რთული და ძნელად ასახსნელი, ხოლო AI-ზე დაფუძნებულ გადაწყვეტილებებს აქვს კარგი პოტენციალი.

დაგეგმვის ფაზაში AI-ის შეუძლია შეასრულოს ორი დავალება:

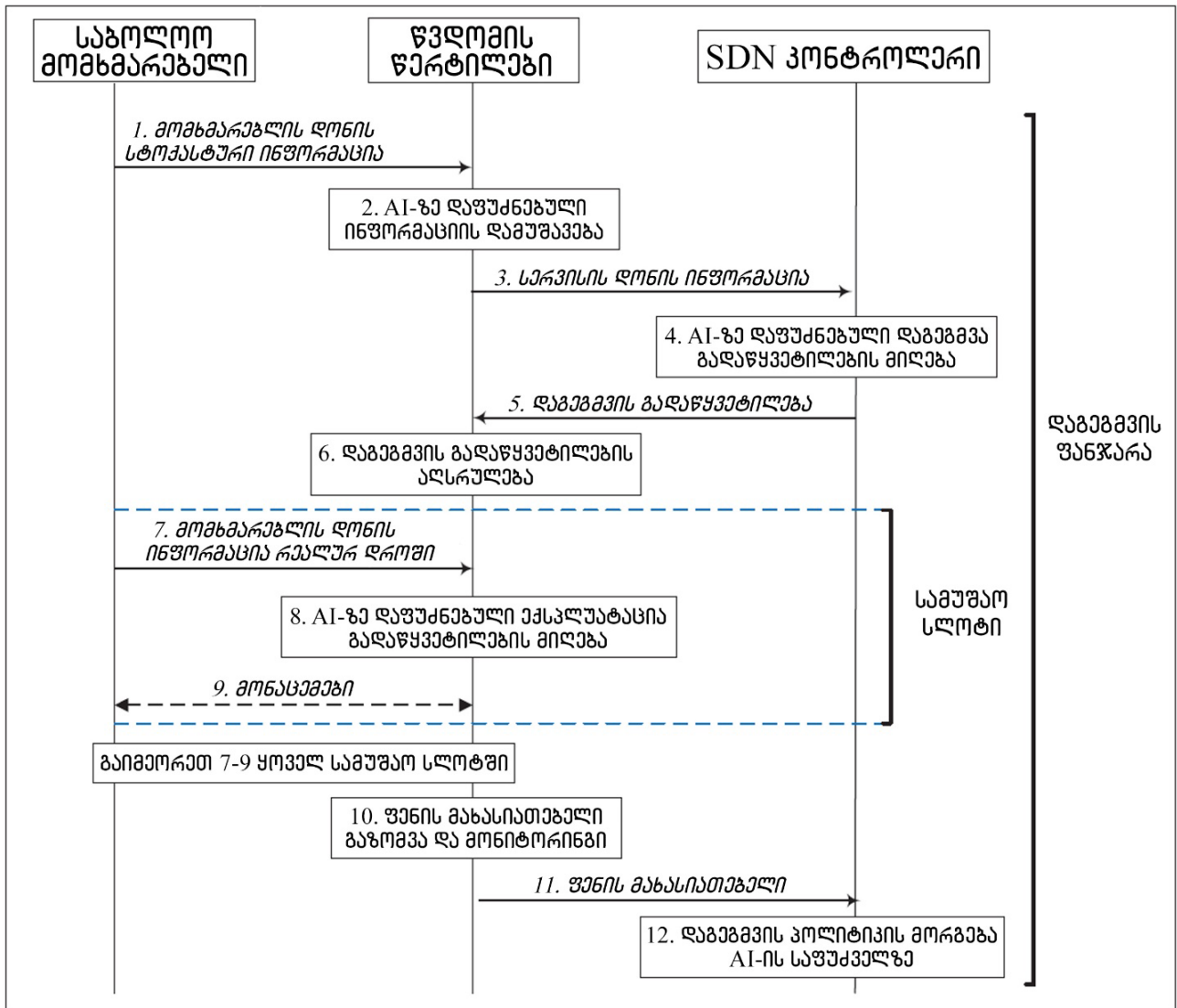
1. VNF-ის განთავსება: SDN კონტროლერი ავრცელებს VNF-ებს ქსელში სერვისების მხარდასაჭერად. VNF-ებისთვის გამოყოფილი რესურსები დინამიკურად უნდა დარეგულირდეს მომსახურების დროში ცვალებად მოთხოვნილებებზე დაყრდნობით, რათა გარანტირებული იყოს მომსახურების შეყვანების მოთხოვნები. DL მეთოდები შეიძლება მისადაგებულ იქნეს დინამიკურ ქსელურ გარემოზე რესურსების გამოყენების გასაუმჯობესებლად.
2. რესურსების რეზერვირება: SDN კონტროლერი ინახავს რესურსებს სხვადასხვა ფენისთვის, მათი სერვისის მოთხოვნებიდან გამომდინარე. ვინაიდან მონაცემთა ტრაფიკის დატვირთვა დროში ცვალებადია, რესურსების რეზერვირება უნდა იყოს ადაპტირებული დინამიკურ მოთხოვნებთან რეალურ დროში, რომელთა მოგვარებაც შესაძლებელია ისეთი RL მეთოდებით, როგორცაა ღრმა დეტერმინირებული პოლიტიკის გრადიენტი (DDPG).

ექსპლუატაციის ფაზაში AI-ის დახმარებით უნდა შესრულდეს ორი დავალება:

1. რესურსების მართვა: ფენის რეზერვირებული რესურსები გამოიყოფა საბოლოო მომხმარებლებისთვის. გადაწყვეტილებები მიიღება რეალურ დროში მომხმარებლის მობილურობის, სერვისის მოთხოვნების და ა. შ. მიხედვით. რესურსების ეფექტიანად გამოყენების მიზნით, RL მეთოდები შეიძლება გამოყენებულ იქნეს რესურსების დინამიკური მართვისთვის.
2. RAT-ის შერჩევა: სისტემის სარგებლობის მაქსიმალურად გასაზრდელად, თითოეული საბოლოო მომხმარებლისთვის ოპტიმალური RAT შერჩევა რამდენიმე კანდიდატ RAT-ს შორის. მომხმარებლის მობილურობის გამო, მომხმარებლის მიერ აღქმული RAT სერვისის შესრულება სტოქასტურია. ამ პრობლემის მოგვარება შესაძლებელია RL-ში განსაზღვრული მრავალხელიანი ბანდიტის (multi-arm bandit) მეთოდებით (მაგალითად, კონტექსტური ბანდიტის მეთოდით).

AI-ის ბაზაზე ფენებად დაყოფის პროცედურა გულისხმობს ინფორმაციის გაცვლას საბოლოო მომხმარებლებს, წვდომის წერტილებსა და SDN კონტროლერს შორის. პროცედურა ილუსტრირებულია ნახ. 5.4-ზე და მოიცავს შემდეგი ნაბიჯებს:

1. წვდომის წერტილები აგროვებს მომხმარებლის დონის იმ სტოქასტურ ინფორმაციას, როგორებიცაა საბოლოო მომხმარებლების სერვისის მოთხოვნის შაბლონები, მობილურობის ნიმუშები და სტოქასტური არხის პირობები.



ნახ. 5.4. ინფორმაციის გაცვლის პროცედურა AI-ის გამოყენებისას ქსელის ფენებად დაყოფისთვის

2. წვდომის წერტილები გარდაქმნის მომხმარებლის დონის ინფორმაციას სასურველ სერვისის დონის ინფორმაციად. მაგალითად, მომხმარებლის სიმკვრივის შესახებ ინფორმაციის მიღება შესაძლებელია მომხმარებლის შესახებ ინფორმაციის დამუშავებით და AI ტექნიკა შეიძლება გამოყენებულ იქნეს ასეთი მონაცემების აბსტრაქციის, შერწყმისა და ანალიზისთვის.
3. დამუშავებული სერვისის დონის ინფორმაცია მიეწოდება SDN კონტროლერს.
4. SDN კონტროლერი აწარმოებს AI-ზე დაფუძნებულ დაგეგმვის ალგორითმებს, რათა გადაწყვეტილებები მიიღოს სერვისის დონის შეგროვებული ინფორმაციის საფუძველზე.
5. დაგეგმვის განსაზღვრული გადაწყვეტილებები იგზავნება უკან, წვდომის ყველა წერტილში.
6. წვდომის წერტილები აღასრულებენ მიღებული დაგეგმვის გადაწყვეტილებებს (მაგალითად, ქსელის რესურსების რეზერვირებას შესაბამისი ფუნქციისთვის).
7. სერვისის საბოლოო მომხმარებლები აცნობებენ რეალურ დროში მისაღებ თავიანთ ისეთ ინფორმაციას მათთან დაკავშირებულ წვდომის წერტილებს, როგორცაა რეალურ დროში სერვისის მოთხოვნები, არხის პირობები და ამოცანების მონაცემთა ბოძები.
8. წვდომის წერტილები იყენებს AI-ზე დაფუძნებულ ოპერაციულ ალგორითმს საბოლოო მომხმარებლებისთვის რესურსების გამოსაყოფად, რეალურ დროში მომხმარებლის დონის ინფორმაციის საფუძველზე.
9. საბოლოო მომხმარებლების სერვისის მოთხოვნები მხარდაჭერილია გამოყოფილი ქსელის რესურსებით. მაგალითად, გამოთვლითი ამოცანები შეიძლება გადაიტვიტოს წვდომის წერტილებში საკომუნიკაციო რესურსის გამოყენებით და შემდეგ დამუშავდეს გამოთვლითი რესურსების გამოყენებით. დაგეგმვის ფანჯარაში თითოეული ოპერაციული ინტერვალისთვის მეორდება ნაბიჯები ნომრებით 7-9.
10. წვდომის წერტილები მონიტორინგს უწევს ქსელში ფუნქციის მახასიათებელს, განხორციელებული დაგეგმვის გადაწყვეტილებების გათვალისწინებით, საბოლოო მომხმარებლების კმაყოფილების მაჩვენებლის გაზომვით, დაგეგმვის ფანჯარაში ყველა ოპერაციული ინტერვალის მიხედვით.
11. წვდომის წერტილები აცნობებენ ქსელის მახასიათებლებს SDN კონტროლერს.
12. SDN კონტროლერი იღებს გადაწყვეტილებას დაგეგმვის შესახებ შემდეგი დაგეგმვის ფანჯრისთვის და არეგულირებს დაგეგმვის პოლიტიკას უკუკავშირის ინფორმაციის საფუძველზე.

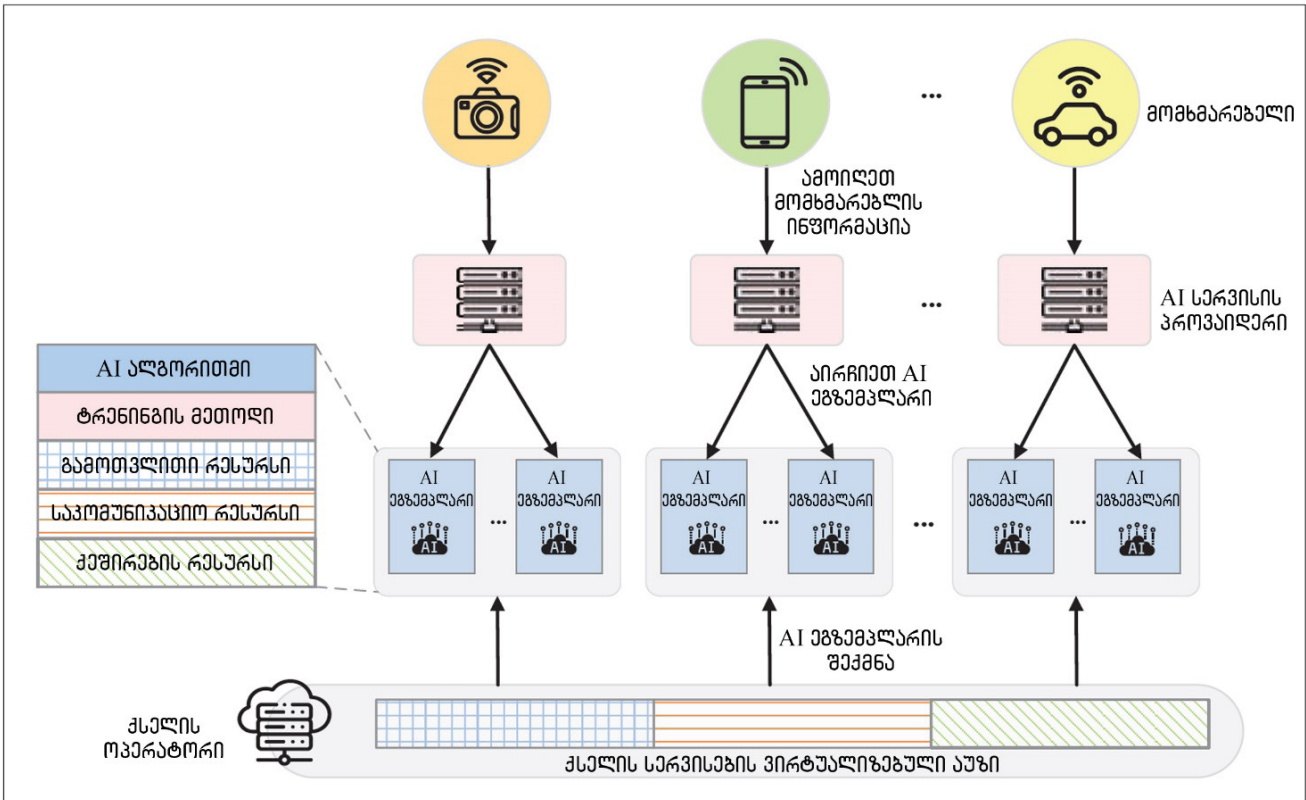
აღვნიშნავთ, რომ წარმოდგენილ პროცედურაში 1-6 და 10-12 ნაბიჯები სრულდება ქსელის დაგეგმვის ფაზაში, ხოლო ნაბიჯები 7-9 მისი ექსპლუატაციის ფაზაში.

5.4. ფუნჯად დაყოფა AI-ისთვის

ფუნჯად დაყოფა AI-ისთვის იყენებს ქსელის დაყოფას AI სერვისების მხარდასაჭერად QoS მოთხოვნების დაკმაყოფილებისას. პოტენციური გადაწყვეტილებები მოიცავს AI-ის ეგზემპლარების შექმნას, შერჩევას და რესურსების ეფექტურ მართვას AI სერვისის სასიცოცხლო ციკლში.

AI სერვისების მხარდაჭერის განხორციელების სხვადასხვა ვარიანტი არსებობს. AI სერვისის დანერგვა შესაძლებელია სხვადასხვა სახის ალგორითმების, ტრენინგის მეთოდებისა და ქსელის რესურსების განაწილების მეშვეობით. მაგალითად, ობიექტების აღმოჩენის სერვისები შეიძლება განხორციელდეს ResNet32, Inception-v3, AlexNet ან VGG16 ალგორითმების მეშვეობით. აქედან გამომდინარე, AI სერვისის მხარდაჭერის მთავარი გამოწვევაა ქსელში განხორციელების შესაბამისი ვარიანტის განსაზღვრა.

ამ პრობლემის გადასაჭრელად, ჩვენ წარმოვადგინებთ AI ეგზემპლარის კონცეფციას, როგორც ნაჩვენებია ნახ. 5.5-ზე.



ნახ. 5.5. AI ეგზემპლარების მართვის კონცეპტუალური სტრუქტურა AI სერვისებისთვის

AI სერვისის AI ეგზემპლარი წარმოადგენს AI სერვისის განხორციელების ვარიანტს. ძირითადი იდეა არის შესაძლო განხორციელების მრავალი ვარიანტის პოვნა და შემდეგ შესაბამისი არჩევა ქსელის გარემოზე დაყრდნობით. AI ეგზემპლარის მართვის კონცეპტუალური ჩარჩო-პროცედურა შედგება ორი ეტაპისგან:

1. **AI ეგზემპლარის შემუშავება:** ქსელის ოპერატორი აყალიბებს მრავალკანდიდატიან AI ეგზემპლარებს თითოეული AI სერვისისთვის, არსებული ქსელის რესურსებისა და სერვისის მოთხოვნების საფუძველზე. AI ეგზემპლარი შეიძლება შეიცავდეს:
 - AI ალგორითმს, რომელიც განსაზღვრავს განხორციელების ალგორითმს და შესაბამისი ნეირონული ქსელის არქიტექტურას;
 - AI ალგორითმის ტრენინგის მეთოდს (მაგალითად, ცენტრალიზებულ ან განაწილებულ ტრენინგს);
 - ქსელის რესურსების საჭირო რაოდენობას.
2. **AI ეგზემპლარის არჩევა:** ამ ეტაპზე, AI სერვისის პროვაიდერი ირჩევს შესაბამის AI ეგზემპლარს კანდიდატი AI ეგზემპლარებიდან მომხმარებლის სერვისის უპირატესობის საფუძველზე (მაგალითად, კონფიდენციალურობის დაცვის უპირატესობით). თუ არჩეულია AI ეგზემპლარი, AI სერვისი შესრულდება AI ალგორითმის და შესაბამისი საჭირო რაოდენობის ქსელის რესურსების გამოყენებით AI ეგზემპლარის მიერ. შედეგად, AI ეგზემპლარის იდეა უზრუნველყოფს მოქნილობას AI სერვისის მართვისთვის.

მოთხოვნის შემთხვევაში ქსელის ოპერატორს შეუძლია დინამიკურად მართოს AI სერვისები კანდიდატი AI ეგზემპლარების შექმნის, შეცვლისა და წაშლის გზით. გასათვალისწინებელია, რომ AI ეგზემპლარები არ შემოიფარგლება მხოლოდ ნეირონული ქსელებით და ისინი შეიძლება შეიცავდეს სხვა ტიპის AI ალგორითმებს, როგორცაა დამხმარე ვექტორული მანქანები და გადაწყვეტილების ხეები.

AI სერვისების გაშვება მოიცავს სამ ეტაპს: მონაცემთა შეგროვებას, მოდელის ტრენინგს და მოდელის დასკვნას (ანუ, ეს არის AI სერვისის მთლიანი სასიცოცხლო ციკლი). კერძოდ, მონაცემთა შეგროვების პროცედურა ხორციელდება საკომუნიკაციო ლინკების საშუალებით და შეგროვებული მონაცემები შეიძლება ინახებოდეს ქსელის პერიფერიულ სერვერებში. შეგროვებულ მონაცემებზე დაყრდნობით, AI მოდელის ტრენინგი შესაძლებელია

მოდელის მომზადების ეტაპზე. მოდელის ტრენინგი შეიძლება განხორციელდეს როგორც ცენტრალიზებული, ისე განაწილებული გზით. მაგალითად, მრავალ მოწყობილობას შეუძლია ერთობლივად იმუშაოს გლობალური მოდელის მოსამზადებლად FL-ის საშუალებით. შემდეგი არის ის რომ, კარგად დატრენინგებული AI მოდელები განლაგებულია კონკრეტული გამოთვლითი ამოცანების შესასრულებლად; ამას ეწოდება მოდელის დასკვნა. მოდელის დასკვნა შეიძლება განხორციელდეს რამდენიმე გზით. მაგალითად, პერიფერიული მოწყობილობების ერთობლივი დასკვნის მიდგომებით შეიძლება განაწილდეს და დამუშავდეს გამოთვლითი ამოცანები ქსელის სხვადასხვა კვანძში, დასკვნის დაბალი შეყოვნების მისაღწევად.

AI სერვისის მახასიათებელი დამოკიდებულია AI მომსახურების სასიცოცხლო ციკლის სამივე ეტაპზე. მაგალითად, მოდელის დასკვნის სიზუსტე დამოკიდებულია ისეთ მრავალ ფაქტორზე, როგორცაა შეგროვებული მონაცემების ხარისხი, ტრენინგის გამეორებების რაოდენობა და მოდელის დასკვნის მისაღებად გამოყენებული მიდგომა. იმავდროულად, სამივე ეტაპი მოიხმარს მრავალგანზომილებიანი ქსელის რესურსებს. შედეგად, AI სერვისების მუშაობის ოპტიმიზაციის მიზნით, ქსელის რესურსები ერთობლივად უნდა გამოიყოს ამ სამი ეტაპისთვის. რეზერვირებული ქსელის რესურსები AI-ის ნაწილებში შემდგომში უნდა გადანაწილდეს ამ სამი ეტაპზე, რათა დააკმაყოფილოს მათი შესაბამისი QoS მოთხოვნები.

5.5. კარძო შემთხვევის შესწავლა

ამ პარაგრაფში წარმოდგენილია AI-ის დახმარებით რესურსების დაჯავშნის პრაქტიკული მაგალითი, რომელიც გვიჩვენებს თუ როგორ შეიძლება შემცირდეს სისტემის საერთო ღირებულება გრძელვადიან პერსპექტივაში.

ჩვენ განვიხილავთ ინტეგრირებულ საჰაერო-სახმელეთო ქსელს ავტონომიური მართვის სერვისების უზრუნველსაყოფად ავტომანქანებისთვის, რომლებიც გადიან ავტომაგისტრალის სეგმენტს. საავტომობილო გზის განხილული მონაკვეთისთვის, რომლის სიგრძეა 2 კმ, ორი BS თანაბრად განლაგებულია მარშრუტის გასწვრივ 1 კმ განცალკევების მანძილით, ხოლო ცენტრში მოთავსებულია ერთი UAV, რომელიც მოძრაობს 100 მ სიმაღლეზე. ამოცანის მონაცემთა ზომა და გამოთვლის ინტენსივობა დაყენებულია 0.6 მბიტზე და 6×10^8 ციკლზე, შესაბამისად. სერვისის შეფერხება ხასიათდება მასობრივი მომსახურების (რიგების) თეორიით, ვინაიდან დავალების ჩამოსვლა განიხილება როგორც პუასონის პროცესი სიჩქარით $\lambda = 1$ პაკეტი/წმ. სერვისის შეყოვნების მოთხოვნის უზრუნველსაყოფად, აგებულია ქსელის ფენა, რომელშიც რეზერვირებულია სპექტრი და გამოთვლითი რესურსები.

რესურსების დაჯავშნის გადაწყვეტილებები მიიღება სისტემის მთლიანი ღირებულების მინიმიზაციის მიზნით, მანქანის დინამიკის გათვალისწინებით. სისტემის მთლიანი ღირებულება მოცემულია ფორმულით

$$C = \sum_{t=1}^T (\omega_r C_r^t + \omega_s C_s^t + \omega_d C_d^t),$$

რომელიც არის სამი ღირებულების კომპონენტის შეწონილი ჯამი დაგეგმვის ყველა T ფანჯარაში.

1. რესურსის რეზერვირების ღირებულება C_r^t ითვალისწინებს დაჯავშნილი სპექტრისა და გამოთვლითი რესურსების რაოდენობას BS-ებზე დაგეგმვის t ფანჯარაში. სპექტრის რესურსი გამოყოფილია 5 მჰც ქვეგადამტანის ერთეულზე, ხოლო გამოთვლითი რესურსი გამოყოფილია ვირტუალური მანქანის (VM) ეგზემპლარზე დამუშავების სიჩქარით 10×10^9 ციკლი/წმ.
2. ფენის რეკონფიგურაციის ღირებულება C_s^t ითვალისწინებს სხვაობას რესურსის რეზერვირების ორ თანმიმდევრულ გადაწყვეტილებას შორის.
3. შეყოვნების მოთხოვნის დარღვევისთვის ჯარიმა C_d^t ეხება შემთხვევას, როდესაც მომსახურების შეყოვნება გადააჭარბებს შეყოვნების მოთხოვნას.

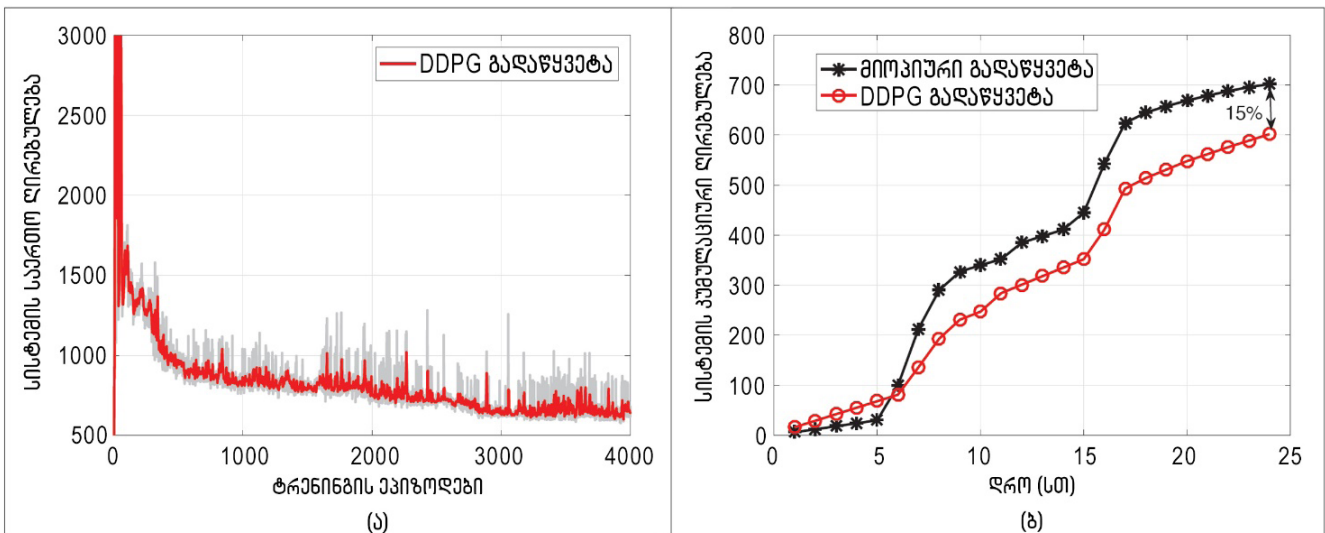
შეწონილი პარამეტრები დაყენებულია როგორც $\omega_r = 1$, $\omega_s = 20$ და $\omega_d = 200$. დაგეგმვის ფანჯრის ზომა დაყენებულია ერთ საათზე.

ჩვენ წარმოვადგენთ DDPG-ზე დაფუძნებულ გადაწყვეტას სისტემის მთლიანი ღირებულების შესამცირებლად. აღვნიშნავთ, რომ DDPG ალგორითმი აერთიანებს პოლიტიკაზე დაფუძნებული და ღირებულებაზე დაფუძნებული მეთოდების ძლიერ მხარეებს ორი ნეირონული ქსელის ჩართვის გზით. ესენია, აქტორის (Actor) ქსელი, რომელიც განსაზღვრავს ოპტიმალურ მოქმედებებს არსებული მდგომარეობის გათვალისწინებით, და კრიტიკოსის (Critic)

ქსელი, რომელიც აფასებს მდგომარეობის-მოქმედების მნიშვნელობის ფუნქციას. ამ გადაწყვეტაში ორივე ქსელი არის სრულად დაკავშირებული ნეირონული ქსელი ოთხი ფენით, ნეირონების რაოდენობა ორ ფარულ ფენაში არის 128 და 64, შესაბამისად, და მათი სწავლების კოეფიციენტი დაყენებულია, შესაბამისად, როგორც 2×10^{-4} და 2×10^{-3} . მახასიათებლების შედარებისთვის, ჩვენ ვიღებთ ოპტიმიზაციაზე დაფუძნებულ გადაწყვეტას, სახელად „რესურსების მიოპიური დაჯავშნა“, რომელშიც ქსელის რესურსები რეზერვირებულია იმისათვის, რომ მინიმუმამდე დაიყვანოს რესურსის რეზერვირების ღირებულება დაგეგმვის თითოეულ ფანჯარაში და შეყოვნების მოთხოვნის დაკმაყოფილების დროს.

ჩვენ ვაფასებთ შემოთავაზებული DDPG-ზე დაფუძნებული გადაწყვეტის მახასიათებელს, რომელიც ეფუძნება რეალურ სიტუაციაში ავტომაგისტრალზე მანქანების მოძრაობის ნაკადის კვალს, და რომელიც შეგროვებულია ალბერტას (კანადა) სატრანსპორტო კომპანიის მიერ. როგორც ნაჩვენებია ნახ. 5.6-ბ-ზე, თავდაპირველად ჩვენ წარმოვადგენთ DDPG-ზე დაფუძნებული შემოთავაზებული გადაწყვეტის კონვერგენციის (კრებადობის) მახასიათებლებს. ხუთი წერტილის მიხედვით მოძრავი საშუალო მნიშვნელობა გამოიყენება სიმულაციის დაუმუშავებელ წერტილებზე, რათა ხაზი გაესვას კონვერგენციის ტენდენციას (იხ. წითელი მრუდი). ჩანს, რომ DDPG-ზე დაფუძნებული რესურსების რეზერვირების გადაწყვეტა 4000 ტრენინგის ეპიზოდის შემდეგ განიცდის კონვერგენციას.

შემდეგ, ნახ. 5.6-ბ-ზე, წარმოდგენილია სისტემის კუმულაციური ღირებულება ერთი დღის განმავლობაში. შეიძლება აღინიშნოს, რომ DDPG-ზე დაფუძნებულ გადაწყვეტას შეუძლია შეამციროს სისტემის მთლიანი ღირებულება ერთი დღის განმავლობაში დაახლოებით 15 პროცენტით მიოპიურ გადაწყვეტასთან შედარებით. მიზეზი არის ის, რომ შემოთავაზებულ DDPG-ზე დაფუძნებულ გადაწყვეტას შეუძლია მინიმუმამდე დაიყვანოს სისტემის მთლიანი გრძელვადიანი ღირებულება, ხოლო მიოპიური გადაწყვეტა მინიმუმამდე ამცირებს სისტემის მოკლევადიან ღირებულებას, რაც იწვევს ფენის რეკონფიგურაციის აკრძალულ ხარჯებს ქსელის რესურსების რეზერვირების ხშირი კორექტირების გამო, მაღალი დინამიკის მქონე სატრანსპორტო ქსელებში. სიმულაციის შედეგები აჩვენებს, რომ შემოთავაზებული AI-ზე დაფუძნებულ რესურსების რეზერვირების გადაწყვეტილებას შეუძლია სისტემის დაბალი ღირებულების უზრუნველყოფა.



ნახ. 5.6. DDPG-ზე დაფუძნებული რესურსების რეზერვირების შემოთავაზებული გადაწყვეტის მახასიათებლის შეფასება:
 ა) კონვერგენციის მახასიათებელი; ბ) სისტემის კუმულაციური ღირებულება ერთი დღის განმავლობაში.

5.6. მომავალი კვლევითი მიმართულებები

ქვემოთ ჩვენ განვიხილავთ კვლევის ზოგიერთ ღია საკითხს, რომლებიც დაკავშირებულია AI-ზე დაფუძნებულ ქსელის ფენებზე დაყოფასთან.

ქსელის დაგეგმვისა და ექსპლუატაციის ერთობლივი დიზაინი:

დაგეგმვა და ექსპლუატაცია ხორციელდება და შერწყმულია ორ განსხვავებულ ვადაში. კონკრეტულად, დაგეგმვის ფაზა ხორციელდება გრძელვადიან მასშტაბში (მაგალითად, წუთების დონეზე) სხვადასხვა ფენისთვის, რესურსების დარეზერვების მიზნით და სერვისის მოთხოვნილებებზე დაყრდნობით, ხოლო ექსპლუატაციის ფაზა ხორციელდება მოკლე დროში (მაგალითად, ქვემოთ დონეზე) რეზერვირებული რესურსების გამოყოფისთვის მოთხოვნილ მომხმარებლებზე ყოველი ფენის ფარგლებში. ქსელის ფენებზე დაყოფის ოპტიმალური მახასიათებლის მისაღწევად საჭიროა დაგეგმვისა და ექსპლუატაციის ერთობლივი ოპტიმიზაცია.

მონაცემთა მართვის სტრუქტურა:

AI-ზე დაფუძნებული ქსელის დაყოფის ქვაკუთხედი არის უამრავი მონაცემი, რომელიც შეიძლება გამოყენებულ იქნეს AI-ის მოდელის ტრენინგისთვის. 5G ქსელებში მონაცემები ფართოდ არის გავრცელებული. შეზღუდული საკომუნიკაციო რესურსების გამო აქ დიდი რაოდენობით მონაცემების შეგროვების ხარჯების უგულვებელყოფა არ შეიძლება და, გარდა ამისა, შეგროვებული მონაცემები უნდა დამუშავდეს ქსელის მართვისთვის ღირებული ინფორმაციის მოსაპოვებლად. მაგალითად, ინდივიდუალური მომხმარებლების ქცევის უხვი ისტორიული მონაცემები შეიძლება გაანალიზდეს, რათა წინასწარ განისაზღვროს სივრცით-დროითი სერვისების მოთხოვნის განაწილება. აქედან გამომდინარე, აუცილებელია მონაცემთა მართვის სტრუქტურის ჩამოყალიბება მონაცემთა შეგროვებისა და ანალიზისთვის.

პროგნოზზე დაფუძნებული ქსელის დაყოფა:

მოწინავე AI ტექნოლოგიების განვითარებით, უკვე შესაძლებელია ქსელში მონაცემთა ტრაფიკის პროგნოზირება. თუ როგორ გამოვიყენოთ ეფექტიანად პროგნოზირების სიმძლავრე ქსელის ფენებზე დაყოფისთვის, წარმოადგენს საინტერესო თემას. ვინაიდან პროგნოზი არასრულყოფილია, პროგნოზირების შეცდომამ შეიძლება გააუარესოს ქსელის დაყოფის მახასიათებელი. შესაბამისად მნიშვნელოვანი საკითხია, თუ როგორ შევაფასოთ პროგნოზირების შეცდომების გავლენა სისტემის მუშაობაზე და შევიმუშავოთ შესაბამისი გადაწყვეტილებები.

5.7. მიხუთი თავის დასკვნა

ამ თავში ჩვენ წარმოვადგინეთ AI-ზე დაფუძნებული ქსელის ფენებზე დაყოფის არქიტექტურა, რათა ხელი შეეწყოს ქსელის ინტელექტუალურ მართვას და 5G ქსელებში AI სერვისების მხარდაჭერას. არქიტექტურა მიზნად ისახავს AI-ისა და ქსელის დაყოფის სინერგიას. AI უზრუნველყოფს ქსელის მართვის სირთულის შემცირებას, დინამიკურ ქსელურ გარემოსთან ადაპტაციის დროს, AI-ის შესაძლებლობების გამოყენებით ქსელის დაყოფის პროცესში. დაყოფა AI-ისთვის არის ქსელის მორგებული ფენების შექმნა, სხვადასხვა განვითარებად AI სერვისთან უკეთესი ურთიერთობისთვის. AI-ზე დაფუძნებული ქსელის დაყოფის არქიტექტურის განვითარების ტემპების დასაჩქარებლად საჭიროა ფართო კვლევითი ძალისხმევა, რაც ასახულია მომავალ კვლევით მიმართულებებში.

თავი 6 განაწილებული მანქანური სწავლების გამოყენების ეფექტიანობა 5G ქსელებისთვის

6.1. შესავალი

ბოლო სამი ათწლეულის განმავლობაში ჩვენ მოწმენი გავხდით სატელეკომუნიკაციო ინდუსტრიის ევოლუციისა 2G-დან 5G-მდე, სადაც თითოეული თაობა უზრუნველყოფს მომხმარებლის უფრო მდიდარ გამოცდილებას მის წინა თაობასთან შედარებით, რაც მხარდაჭერილია რადიონტერფეისისა და ძირითადი ქსელის მზარდი წინსვლებით. მიუხედავად იმისა, რომ კომერციული 5G არაავტონომიური ქსელები პირველად 2019 წელს ამოქმედდა, სამართლიანობისთვის უნდა ითქვას, რომ მათ ჯერ კიდევ არ აქვთ რეალიზებული თავიანთი სრული პოტენციური აპლიკაციების ფართო სპექტრში, დაწყებული eMBB ქსელიდან, mMTC, IoT და URLLC ტექნოლოგიების ჩათვლით. კომერციული მობილური ქსელების თაობების გზის გადახედვისას, აღმოვაჩინეთ, რომ თითოეულ თაობას ესაჭიროება დაახლოებით ათი წლის ფოკუსირებული კვლევა ამ ტექნოლოგიაზე დაფუძნებული წარმატებული ფართომასშტაბიანი განლაგებისთვის. შესაბამისად, გონივრული იქნება ველოდოთ პირველი 5G ქსელის გაშვებას 2030 წლისთვის. ამიტომ, დროა ვიკითხოთ: როგორი იქნება 5G?

პასუხი იმაზე, თუ როგორი უნდა იყოს 5G ასეთია: – „იქნება მრავალგანზომილებიანი“, და 5G-ის ჭეშმარიტად პერსპექტიული ხედვა უნდა ითვალისწინებდეს არა მხოლოდ ფუნქციურ ტექნოლოგიებსა და სერვისებს, არამედ უნდა გამოიყენოს განვითარებადი ღირებულების ჯაჭვებისა და მომავალი ბიზნეს-ვერტიკალების სრული სპექტრი. პირველი განლაგებული 5G ქსელი შეიძლება იყოს გარკვეულწილად 5G ქსელების ევოლუცია, მაგრამ ასევე ერთდროულად გააერთიანებს რადიკალურად „გამრღვევ“ ტექნოლოგიებს, გამოყენების შემთხვევების რეალიზაციისთვის, რომლებიც ამჟამად არ არის 5G-ის ნაწილი და რომელიც შეიძლება სრულად არ იყოს რეალიზებადი არსებული ტექნოლოგიებით.

იმისათვის, რომ 5G გახდეს საყოველთაოდ უსადენო ქსელი, მნიშვნელოვანია, რომ მან საშუალება მისცეს მობილური ქსელის ოპერატორებს უზრუნველყონ მომსახურების ფართო სპექტრი არა მხოლოდ საბოლოო მომხმარებლებისთვის (მომხმარებლის მოწყობილობები, UE-ები), არამედ უფრო დიდი საწარმოო სერვისებისთვისაც, რომლებიც შეიძლება კლასიფიცირდეს როგორც „ბიზნესი ბიზნესისთვის“ (B2B) სერვისების კატეგორია, რომელიც მოიცავს ლოჯისტიკას, წარმოებას, ტრანსპორტირებას, ჯანდაცვას, საბანკო და სამთავრობო სექტორს და ა. შ. მიუხედავად იმისა, რომ 5G ცდილობდა მომსახურებოდა ასეთი B2B ვერტიკალების ქვეჯგუფს mMTC და URLLC-ის საშუალებით, მნიშვნელოვანი ძალისხმევაა საჭირო 5G-ში KPI-ების განსაზღვრისთვის სხვადასხვა ვერტიკალისთვის B2B კატეგორიაში. მუდმივად დაკავშირებულ საბოლოო მომხმარებელს, რომელიც შეიძლება იყოს მომხმარებელი „ბიზნესი მომხმარებლისთვის“ (B2C) სერვისში ან B2B-ში, უნდა ჰქონდეს წვდომა ყველა ხელმისაწვდომ ციფრულ სერვისზე 5G ფართო ქსელის სხვადასხვა, კარგად განსაზღვრული შესაძლებლობის გამოყენებით. ამ მიზნით, მოცემულ პარაგრაფში, ჩვენ წარმოვადგენთ 5G-ის გამოყენების შემთხვევების მაგალითებს. ეს სია არ არის ამომწურავი, მაგრამ მოიცავს ვერტიკალების ერთობლიობას, რომლებიც შეიძლება განხორციელდეს 5G-ში არსებული KPI სპეციფიკაციების საფუძველზე და რომლებიც ამჟამად ნაწილობრივ ხელმისაწვდომია 5G სტანდარტებში.

წარმოება:

მძიმე ინდუსტრია, განსაკუთრებით წარმოების სექტორი, ეყრდნობა მაღალი სიზუსტის აღჭურვილობას, რომელსაც, ხშირად, ერთმანეთთან თანამშრომლობით უწევს ფუნქციონირება, მაგალითად, რობოტოტექნიკა საავტომობილო ინდუსტრიაში ან მათი სათადარიგო ნაწილების წარმოებაში. 5G-ის ხედვით განსაზღვრულ ინდუსტრიაში, მოსალოდნელია, რომ მრავალჯერადი RAT ტექნოლოგია, დროისადმი მგრძობიარე ქსელთან ერთად, უზრუნველყოფენ მაღალ საიმედოობასა და დაბალ შეყოვნებას, რაც ხელს შეუწყობს სხვადასხვა ინდუსტრიულ აპლიკაციას. ამის წინასწარი ვერსია გამოიყენება როგორც ინდუსტრია 4.0-ის ხედვის ნაწილი; თუმცა, 5G NR-ის სტანდარტები, რომლებიც მიმართულია ინდუსტრია 4.0-ში, საკმარისი არ არის შემდეგი თაობის ინდუსტრიული კუთხით მოწინავე გამოყენების შემთხვევების გასააქტიურებლად. მაგალითად, კრიტიკული ინფრასტრუქტურის რეალურ დროში ავტომატიზებული მონიტორინგი, როგორცაა ელექტროგადამცემი ქსელები

და ენერგომომარაგების ხაზები, მოითხოვს URLLC ბმულებს, რომლებიც მოიცავს რობოტებისა და ადამიანების ორმაგ მობილურობას.

ჯანმრთელობის დაცვა:

COVID-19 პანდემიამ აჩვენა ჯანდაცვის სერვისების პირდაპირ მიწოდების მნიშვნელობა პაციენტებისთვის საკუთარ სახლში და ასეთი სერვისების მხარდაჭერის აუცილებლობა 5G-ში. სახლში მოვლაზე გადასვლა ასევე ამცირებს საავადმყოფოების საოპერაციო და ადმინისტრაციულ ხარჯებს და უკეთესად ემსახურება მაღალი რისკის მქონე პაციენტებს მობილურობის პოტენციური დარღვევებით. პაციენტის დისტანციური მონიტორინგი შეიძლება გაუმჯობესდეს AR/VR-ის გამოყენებით. მონაცემთა გადაცემის მაღალი სიჩქარე და უკიდურესად მაღალი საიმედოობის ლინკები მილიწამზე ნაკლები შეყოვნების მოთხოვნებით საჭირო იქნება რობოტის დახმარებით ტელეოპერაციისთვის. კიდევ ერთი ძირითადი ტექნოლოგია, რომელიც ტელემედიცინას შესაძლებელს გახდის, იქნება რეალურ დროში ტაქტილური უკუკავშირი, რომელიც ჯერ არ არის მიღწეული. თუმცა, 5G-ის, რომელიც, სავარაუდოდ, უზრუნველყოფს უფრო მაღალსიხშირული სპექტრის გამოყენებას, პერიფერიული მოწყობილობების განაწილებული სწავლების ტექნიკასთან ერთად, შეუძლია პოტენციურად უზრუნველყოს ტელემედიცინაში საჭირო მკაცრი KPI-ები.

საზოგადოებრივი სერვისები და უსაფრთხოება:

საზოგადოებრივი უსაფრთხოების (PS) ოპერაციები გადამწყვეტია იმისთვის, რომ მოქალაქეებს მიაწოდოს მნიშვნელოვანი და ზოგიერთ შემთხვევაში სიცოცხლის გადამრჩენი ინფორმაცია სახელმწიფო უწყებებიდან. პირველი მოპასუხეები ყველაზე მნიშვნელოვანი კომპონენტია ბრძანების სრულ ჯაჭვში და ტრადიციულად ეყრდნობოდნენ არასაიმედო ხმოვან საკომუნიკაციო ლინკებს. თუმცა, განახლებული შესაძლებლობები, როგორცაა რეალურ დროში მაღალი გარჩევადობის ვიდეო ნაკადი სპეციალიზებული კამერებიდან და რეალურ დროში წვდომა სენსორების მონაცემებზე, მათ შორის თერმული სენსორების ჩათვლით, გაზრდის პირველი რეაგირების უნარს, უფრო ეფექტიანად შეამსუბუქოს კრიზისები. D2D კომუნიკაციები, მათ შორის რობოტების დისტანციური მართვა ისეთი აპლიკაციებისთვის, როგორცაა ბომბების განადგურება და რობოტების ფუნქციონირება ინციდენტის ადგილებში, ასევე საჭიროებს KPI-ებს, რომლებსაც ამჟამად არ აკმაყოფილებს 5G. 5G-ში მათ დასაკმაყოფილებლად, ყურადღება უნდა გამახვილდეს გაუმჯობესებულ დაფარვაზე და მკვრივ გარემოში დიდი რაოდენობის კავშირების მხარდაჭერის შესაძლებლობაზე. PS ქსელის ინფრასტრუქტურა 5G-ში ასევე უნდა მოიცავდეს KPI-ებს, რომლებიც უზრუნველყოფენ ბატარეით მომუშავე საბოლოო მომხმარებლის მოწყობილობების ეფექტიან გამოყენებას PS შეტყობინებების მიღებისას.

6.2. კომპრომისი კომუნიკაციასა და გამოთვლებს შორის

ადრეული სამეცნიერო ნაშრომები, რომლებმაც გადამწყვეტი როლი ითამაშეს 5G-ის იდეის ჩამოყალიბებაში, ხაზს უსვამდნენ პროგრამულ უზრუნველყოფაზე ორიენტირებულ მიდგომაზე გადასვლის აუცილებლობას, დაწყებული ქსელის ბირთვიდან რადიოინტერფეისამდე. SDN, რომელიც აღნიშნავს გადასვლას ტრადიციული, აპარატურაზე ორიენტირებული მიდგომიდან, NFV-თან ერთად, კვლავაც იქნება მამოძრავებელი ძალა 5G ქსელებში, როგორც ეს იყო 5G-ში. პარალელურად, მობილური პერიფერიული გამოთვლების (MEC) პარადიგმა საჭიროებს RAN-ის და ძირითადი ქსელის ფუნქციების რესტრუქტურირებისას BS-ის ზოგიერთი ფუნქციის ღრუბელში გადატანით და ძირითადი ქსელის ზოგიერთი ფუნქციის ქვემოთ გადმოტვირთვით. ეს ქმნის მკაფიოდ იდენტიფიცირებულ „პერიფერიულ“ ან „ნისლის“ დომენს BS-სა და საბოლოო მოწყობილობას შორის.

მიუხედავად იმისა, რომ ღრუბლოვანი გამოთვლები უზრუნველყოფდა უფრო მდიდარ და რთულ აპლიკაციებს საბოლოო მომხმარებლებისთვის დისტანციური ღრუბლოვანი სერვერის სიმძლავრის გამოყენებით, 5G და პოტენციურად 5G გამოყენების შემთხვევებისთვის განსაზღვრული უკიდურესად მგრძობიარე შეყოვნების მოთხოვნები მოითხოვს ალტერნატიულ მიდგომას. თანამედროვე უსადენო ქსელებში ტრაფიკის რთული განაწილების გამო, ქსელის არქიტექტურა სულ უფრო ჰეტეროგენული ხდება.

არსებობს მრავალი ტიპის ქსელის წვდომის კვანძი, რომელიც უზრუნველყოფს საიმედო და შეუფერხებელ კავშირს მობილური მომხმარებლებისთვის, როგორცაა, მაგალითად, მაკრო BS, პატარა ფიჭის BS და WiFi წვდომის წერტილი (AP). ქსელის წვდომის ეს კვანძები მხარს უჭერენ პერიფერიულ გამოთვლებს ქსელის

პერიფერიებზე გადაცემის დაბალი შეყოვნებით. ქსელის წვდომის კვანძების განსხვავებული მახასიათებლების გამო, როგორცაა დაფარვის უნარი და გადაცემის სიმძლავრე, ჰეტეროგენული MEC ქსელების თანაარსებობის დიზაინმა მზარდი ყურადღება მიიპყრო. ამასთან, ქსელის წვდომის მრავალ კვანძს შორის კოორპერატიული გამოთვლითი გადმოტვირთვა კარგად უნდა იყოს შემუშავებული.

6.3. განაწილებული სწავლება 6G-ისთვის

ფიჭური უსადენო ქსელების დისპერსიული და ზოგჯერ მწირი ბუნების გამო, რომელიც შედგება შესაძლო ჰეტეროგენული საბოლოო მოწყობილობებისგან, განაწილებული ML-ის (DML) პარადიგმა გახდა სასიცოცხლოდ მნიშვნელოვანი ML-ის მიდგომების გამოსაყენებლად, ზოგადად, უსადენო ქსელის პრობლემების გადაჭრის პროცესში. ფაქტორები, რომლებიც DML-ს შესაფერისს ხდის უსადენო ქსელებში გამოსაყენებლად, მრავალია; მაგალითად:

- ვინაიდან მობილური და IoT მოწყობილობები გამოთვლების კუთხით უფრო უნარიანი ხდებიან და აქვთ უფრო გაზრდილი შენახვითი ტევადობა, ისინი ასევე გამოიმუშავენ ლოკალური მომხმარებლის მონაცემების ექსპონენციალურად მზარდ მოცულობას და კონტექსტურ სენსორულ მონაცემებს, რომლებიც წარმოიქმნება სხვადასხვა აპლიკაციიდან.
- სერვერის კვანძებში საბოლოო მოწყობილობებიდან დიდი რაოდენობით მონაცემების გაგზავნაზე შეზღუდვების გამო და მომხმარებლის მონაცემების კონფიდენციალურობის შესუსტების გამო, ლოკალური მონაცემების გაგზავნა სერვერის კვანძში (აგრეგატორში) ტრენინგის ყოველ რაუნდში არაოპტიმალურია.

ამიტომ, საბოლოო მოწყობილობებისთვის სასარგებლოა ლოკალურად წარმოქმნილი მონაცემების გენერირება და შენახვა მოწყობილობაზე და მხოლოდ ლოკალური ტრენინგიდან მიღებული მოდელის პარამეტრების განახლებების გადატანა ცენტრალურ სერვერზე, რომელიც შეიძლება გამოყენებულ იქნეს გლობალური ML მოდელის განახლებისთვის. ამას მოიხსენიებენ, როგორც „პარამეტრული სერვერის“ არქიტექტურას, რომელიც შეიძლება დაიყოს ცენტრალიზებულ, მრავალ კვანძზე განაწილებულ ML მიდგომად. FL არის პარამეტრული სერვერის არქიტექტურის ერთ-ერთი ყველაზე პოპულარული ვარიანტი და წარმოადგენს უსადენო საკომუნიკაციო სისტემებში განაწილებული ML კვლევის დიდ უმრავლესობას. ლიტერატურიდან ცნობილია სხვა დეცენტრალიზებული DML-ის მიდგომები. თუმცა, ისინი ფართოდ არ გამოიყენება უსადენო ქსელებში პრაქტიკული გამტარუნარიანობისა და შეყოვნების შეზღუდვების გამო და, შესაბამისად, ჩვენი ყურადღება ამ თავში გამახვილებული იქნება FL არქიტექტურაზე და მასთან დაკავშირებული ალგორითმებზე, რომლებიც შეიძლება გამოყენებულ იქნეს უსადენო ქსელების სისტემებში.

FL განსაკუთრებით კარგად შეეფერება ფართომასშტაბიან უსადენო ისეთ ქსელებს, როგორცაა ფიჭური სისტემები. კერძოდ, FL განიხილავს მომხმარებლის კონფიდენციალურობის არაერთგვაროვან საკითხებს, რომლებიც კარგად არ არის შესაწავლილი უფრო ტრადიციულ DML არქიტექტურებში, რაც შეიძლება მოიცავდეს ლოკალური მომხმარებლის მონაცემების გაზიარებას ცენტრალურ სერვერთან ან ერთმანეთთან. გარდა ამისა, რადგან FL მოითხოვს მხოლოდ პარამეტრის განახლებების გაზიარებას მონაწილე მოწყობილობებიდან აგრეგატორთან და არა თავად ლოკალურ მონაცემებთან, FL ამცირებს კომუნიკაციის საერთო ხარჯებს და შეუძლია უფრო ეფექტიანად მოაგვაროს უსადენო არხის გაურკვევლობა, რითაც აუმჯობესებს საიმედოობას.

თავდაპირველად განვიხილოთ FL-ის ძირითადი პრინციპები.

მხოლოდ ღრუბელზე დაფუძნებული ML-ის ტრადიციული მიდგომა ატვირთავს საბოლოო მოწყობილობებზე მიღებულ მონაცემებს დისტანციურ ღრუბლოვან სერვერზე ცენტრალიზებული ტრენინგისთვის, რათა მოამზადოს საერთო მოდელი მომავალი დასკვნისთვის. თუმცა, ღრუბელში ტრენინგის დრო შეიძლება არაპრაქტიკული იყოს მგრძობიარე მონაცემების დიდი მოცულობის გამო, რომელიც უნდა იქნეს გამოყენებული ტრენინგის პროცესში, და ნაწილობრივ იმის გამო, რომ ML მოდელის ტრენინგის გამოთვლითი სირთულე მნიშვნელოვნად დიდია. იმავდროულად, იმის გამო, რომ ღრუბლოვანი სერვერი შეიძლება ფიზიკურად იყოს დაშორებული საბოლოო მოწყობილობებიდან, ამ მოწყობილობებს შეიძლება დისფუნქციურს ხდიდეს კომუნიკაციის დიდი შეყოვნებები. ამ პრობლემის გადასაჭრელად, MEC-ის მიერ ხელშეწყობილი FL შეიძლება იყოს პერსპექტიული მიდგომა ცენტრალიზებული სასწავლო პარადიგმიდან უფრო პრაქტიკულ დეცენტრალიზებულ ტრენინგზე გადასვლისთვის. FL იძლევა ML მოდელის აგრეგაციას სხვადასხვა საბოლოო მოწყობილობაზე, რომლებიც გადიან ტრენინგს

მათი ლოკალური მონაცემების ნაკრების გამოყენებით და ერთობლივად სწავლობენ გლობალურ მოდელს. კონკრეტულად, ყოველი რაუნდის დასაწყისში, სერვერი აგზავნის მიმდინარე გლობალურ მოდელს თითოეულ მონაწილე საბოლოო მოწყობილობაზე. საბოლოო მოწყობილობები (კლიენტები) შემდეგ ატრენინგებენ თავიანთ ინდივიდუალურ ლოკალურ მოდელებს მათი შეზღუდული მონაცემების ნაკრების საფუძველზე და გადააქვთ მოდელის პარამეტრების განახლებები ტრენინგის ყოველი რაუნდის ბოლოს სერვერზე განთავსებულ აგრეგატორზე. ეს შეიძლება განმეორდეს ტრენინგის იმდენი რაუნდისთვის, რამდენიც საჭიროა გლობალური მოდელისთვის სასურველი სიზუსტის მისაღწევად. FL DML-ის სხვა სქემებისგან გარკვეული უნიკალური ფაქტორებით განსხვავდება. პირველი, ვარაუდი, რომ მონაცემთა ნიმუშები, რომლებიც განთავსდება სხვადასხვა საბოლოო მოწყობილობაზე, არის დამოუკიდებელი და თანაბრად განაწილებული შემთხვევითი ცვლადების რეალიზაცია, შეიძლება არ იყოს მართებული FL-ში, რადგან ერთი მომხმარებლის საბოლოო მოწყობილობის ლოკალურ მონაცემთა ნაკრები შეიძლება არ ასახავდეს ყველა მომხმარებლის განაწილებას. მეორე, FL-ში გენერირებული ლოკალურ მონაცემთა ნაკრები შეიძლება მნიშვნელოვნად განსხვავდებოდეს ზომით, რაც იწვევს დაუბალანსებელ განაწილებას. მონაცემთა ზომეში ეს დისბალანსი უპირველეს ყოვლისა განპირობებულია სხვადასხვა ტიპის IoT მოწყობილობით (მაგალითად, სმარტფონი ან მანქანა) და აპლიკაციის სხვადასხვა სცენარით (მაგალითად, რუკების აპლიკაცია სმარტფონზე შეიძლება უფრო მეტ მონაცემს გამოიმუშავებდეს ქალაქის აქტიური მომხმარებლისთვის, ვიდრე სოფლის ნაკლებად აქტიური მომხმარებლისთვის). მესამე, FL-ის პირობებში, გაზომილი მონაცემების ნიმუშების ჯამური რაოდენობა, რომელიც ხელს უწყობს გლობალური მოდელის შესწავლას პერიფერიულ სერვერზე, გაცილებით მეტია, ვიდრე ის ხელმისაწვდომია თითოეული მომხმარებლის ლოკალური ტრენინგისთვის. დაბოლოს, FL მოსწავლეთა უმეტესობა არის მობილური მოწყობილობები (სმარტფონები, ტარებადი მოწყობილობები, დრონები, მანქანები და ა. შ.), რომელთაც, შესაძლოა, აქვთ FL პერიფერიულ სერვერთან არასაიმედო უსადენო კავშირი. ეს გულისხმობს, რომ აგრეგატორს შეიძლება მოუწიოს ხაზგარეშე ან ნელი დაკავშირების მქონე მოსწავლეების მხარდაჭერა, განსაკუთრებით ფიჭური აპლინკის სცენარში. ამ განმასხვავებელი ფაქტორების კონტექსტში, FL იძლევა აშკარა უპირატესობებს უსადენო აპლიკაციებში, რომლებიც შეიძლება არ იყოს ხელმისაწვდომი სხვა დეცენტრალიზებულ ML მიდგომებში.

6.4. კერძო შემთხვევის შესწავლა: FL-ის გამოყენება სპექტრზე დინამიკური წვდომისთვის

ამ პარაგრაფში განვიხილავთ სპექტრზე დინამიკურ წვდომას (DSA), როგორც სპეციალურ აპლიკაციას, რომელშიც FL შეიძლება გამოყენებულ იქნეს საუკეთესო მახასიათებლების მისაღწევად. პირველ რიგში, ჩვენ წარმოვადგენთ რამდენიმე წინასწარ ინფორმაციას DSA-ს შესახებ. სპექტრული რესურსების ეფექტიანად გამოსაყენებლად სპექტრის მართვის ორი ტიპის მექანიზმი შეიძლება იქნეს წარმოდგენილი: სტატიკური და დინამიკური. სპექტრის სტატიკური გაზიარება ავგუფებს და აწესრიგებს სპექტრის ყველა რესურსს, რათა იგივე ნაწილი დაუბრუნდეს სერვისის პროვაიდერებს (SP). ლიცენზირებული SP-ები შესაბამისად გეგმავენ ამ სპექტრულ რესურსებს თავიანთი აბონენტებისთვის. მეორე მხრივ, DSA-ში სპექტრული რესურსები დინამიკურად ნაწილდება როგორც ლიცენზირებულ SP-ებზე, ასევე არალიცენზირებულ SP-ებზე, შესაბამისად, QoS-ის გარანტიით და მის გარეშე. ეს უზრუნველყოფს ეფექტიან გზას ხელმისაწვდომი რადიორესურსების გამოსაყენებლად და სპექტრის დეფიციტის აღმოსაფხვრელად არალიცენზირებული SP-ებისთვის ახალი სპექტრული რესურსების დამატების გარეშე. ქვემოთ, ლიცენზირებული და არალიცენზირებული მომხმარებლები, შესაბამისად, მოიხსენიებიან, როგორც პირველადი მომხმარებლები (PrU) და მეორადი მომხმარებლები (SeU).

ML მეთოდები ადრე გამოიყენებოდა DSA აპლიკაციებში სპექტრული რესურსების უფრო ეფექტიანად გადანაწილებისთვის. მაგალითად, DRL დაინერგა როგორც ბუნებრივი ინსტრუმენტი DSA-ისთვის და გაზიარებისთვის. კერძოდ, DRL აგენტი ახორციელებს მოქმედებას გარემოზე დაკვირვების საფუძველზე, იღებს ჯილდოს გარემოსგან, განხორციელებული მოქმედების მიხედვით და შემდეგ გადადის ახალ მდგომარეობაში. DRL აგენტის მიზანია მოიძიოს პოლიტიკა, რომელიც აუქმობეს კუმულაციურ ჯილდოს. ლიტერატურაში განხილული DRL სტრუქტურა არის მრავალაგენტიანი განმტკიცებული სწავლება (MARL). აქ, სისტემაში ჩართულია მრავალი აგენტი, რითაც ეს გარდაიქმნება ოპტიმიზაციის პრობლემად, რომელიც მოიცავს ყველა ჩართული აგენტის

პოლიტიკას. თითოეული აგენტის ინდივიდუალური ქმედებები, ჯილდოები და მდგომარეობა გავლენას ახდენს ყველა სხვა აგენტის ქმედებებზე.

MARL საშუალებას აძლევს ამ აგენტებს დაუკავშირდნენ სერვერს და დაამუშაონ მათი განაწილებული ამოცანები პარალელურად, როგორც კი აგენტები მიიღებენ მათ. მაგალითად, აგენტებს შეუძლიათ თავიანთი გამოცდილება გაუზიარონ ერთმანეთს, რათა გაზარდონ სწავლების კონვერგენცია. გარდა ამისა, MARL საშუალებას იძლევა სისტემაში მოხდეს ახალი აგენტების დამატება და არააქტიური აგენტების ჩანაცვლება. თუმცა, MARL იყენებს ჭარბ გამოთვლით დროს მისი ექსპონენციალური სირთულის გამო, რაც პრობლემის განზომილების ფუნქციაა. მასზე ასევე გავლენას ახდენს გარემოს არასტაციონარული მდგომარეობა, ისევე როგორც კომპრომისი კვლევასა და ექსპლუატაციას შორის. ამ პრობლემების გადასაჭრელად, ლიტერატურაში შემოთავაზებულია სხვადასხვა გადაწყვეტილება, მათ შორის ღრმა Q-ქსელი, მისი ვერსია რეზერვუარული გამოთვლებით, რომელიც ცნობილია როგორც ღრმა ექო მდგომარეობის Q-ქსელი (DEQN). ჩვენ აღვწერთ, თუ როგორ შეიძლება გამოვიყენოთ MARL-ზე დაფუძნებული FL კონფიგურაცია DSA პრობლემის გადასაჭრელად.

მოკლე აღწერა:

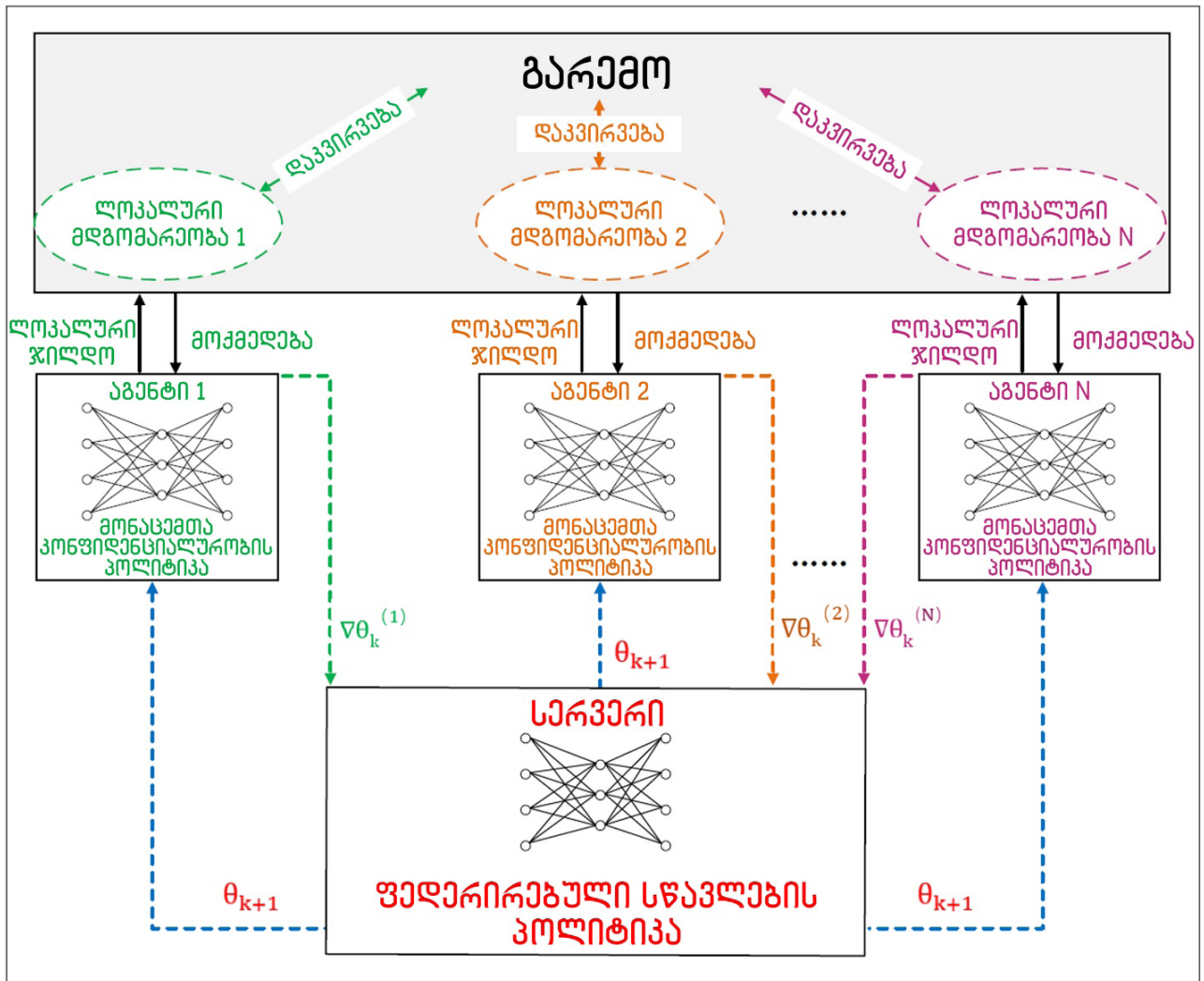
არსებული MARL ალგორითმები ვარაუდობენ, რომ ერთობლივი ჯილდო მიიღება ყველა აგენტის მიერ, ან რომ თითოეული აგენტი იღებს ინდივიდუალურ ჯილდოს, მაგრამ უზიარებს მას სხვა აგენტებს. თუმცა, ეს ვარაუდი შეიძლება არ იყოს პრაქტიკული გარკვეულ რეალურ აპლიკაციებში, რადგან აგენტებმა შეიძლება არ გააზიარონ თავიანთი დაკვირვებები და მიღებული ჯილდოები მონაცემთა კონფიდენციალურობისა და უსაფრთხოების პრობლემების გამო. MARL-ზე დაფუძნებულ FL სისტემაში (რომელსაც განვიხილავთ), აგენტები არ უზიარებენ თავიანთ ლოკალურ დაკვირვებებსა და ჯილდოებს სხვა აგენტებს, მაგრამ ანახლებენ თავიანთ პოლიტიკას, რათა მაქსიმალურად გაზარდონ ინდივიდუალური გრძელვადიანი ლოკალური ჯილდოები. ამ სისტემის მიზანია ერთობლივი გრძელვადიანი ჯილდოს ოპტიმიზაცია, რომელიც გამოიხატება როგორც თითოეული აგენტის გრძელვადიანი ლოკალური ჯილდოს ჯამი.

FL-ის არქიტექტურა MARL სქემაში გამოსახულია ნახ. 6.1-ზე, სადაც Θ_k წარმოადგენს მოდელის პარამეტრებს კომუნიკაციის k რაუნდში და $\nabla\Theta_k^{(n)}$ წარმოადგენს მოდელის პარამეტრების განახლებებს k რაუნდში, რომელიც გაგზავნილია n მომხმარებლის მიერ სერვერზე. ჩვენ ვირჩევთ სიგნალის ინტერფერენცია-პლუს-ხმაურთან თანაფარდობას (SINR), როგორც ჩვენი ხარისხის მეტრიკას, რომელიც ითვალისწინებს ყველა ფაქტორს (რომელიც გავლენას ახდენს SeU-ზე ქსელში), როგორცაა მიმღების თერმული ხმაური, BS-ის გადაცემის სიმძლავრე და ინტერფერენცია პოტენციურად ერთდროულად გადამცემ BS-მომხმარებლების წყვილებს შორის. მომხმარებლის გამტარუნარიანობა გამოიყენება როგორც თითოეული მომხმარებლის ან აგენტის ლოკალური ჯილდოს ფუნქცია. ვინაიდან MARL საშუალებას აძლევს მომხმარებელს ჰქონდეს ინტერაქცია გარემოსთან და დაატრენინგოს საერთო მოდელი გრძელვადიანი ჯილდოს მაქსიმიზაციისთვის, ის კარგად ემთხვევა FL იდეას, რომ გამოიყენოს საერთო გლობალური მოდელი ყველა მომხმარებლისთვის. ეს ასევე იძლევა მაგალითს იმის შესახებ, თუ როგორ შეიძლება FL-ის გამოყენება დიდი რაოდენობის SeU-ების განსათავსებლად საჯარო სერვისებში და უსაფრთხოების გამოყენების შემთხვევაში 6G-ისთვის, რომელიც ზემოთ იყო აღწერილი.

სპექტრზე წვდომის პოლიტიკა:

სპექტრული რესურსების ეფექტიანი გამოყენებისთვის ჩვენ ვამოდელირებთ სპექტრზე წვდომის სტრატეგიას შემდეგნაირად. არსებობს N SeU და M არხი ($N > M$), ამიტომ თითოეულ SeU-ს შეუძლია კონკრეტულ დროს მხოლოდ ერთ არხზე წვდომა. არალიცენზირებული მომხმარებლებიდან ინტერფერენციების თავიდან ასაცილებლად, SeU-ს არ აქვს უფლება გადასცეს კონკრეტულ არხში, როდესაც PRU იკავებს ამ არხს. თუმცა, ერთმა SeU-მ შეიძლება ხელი შეუშალოს მეორე SeU-ს. PRU არხზე წვდომის აქტივობა მოდელირებულია, როგორც მარკოვის პროცესი. SeU-ებს შორის ინტერფერენციების თავიდან აცილების მიზნით, ჩვენ ვიყენებთ ადრე აღწერილ სპექტრზე წვდომის სტრუქტურას SeU-ებისთვის.

განხილულ MARL სისტემაში ჩვენ ვიყენებთ დეცენტრალიზებული პოლიტიკის გრადიენტის მეთოდს ერთობლივი ჯილდოს ოპტიმიზაციისთვის. ინიციალიზებული პოლიტიკის ქსელი პირველად ნაწილდება ყველა აგენტზე. თითოეული აგენტის პოლიტიკის ქსელი, რომელიც დანერგილია როგორც ნეირონული ქსელი, განახლებულია გარემოზე საკუთარი დაკვირვების საფუძველზე. თითოეულ საკომუნიკაციო ციკლში აგენტი ასუფთავებს თავის ბუფერს, აკვირდება გარემოს, ასრულებს მოქმედებას მისი პოლიტიკის საფუძველზე და იღებს ჯილდოს გარემოსგან. ზემოაღნიშნული ნაბიჯების გამეორების შემდეგ საკმარისი რაოდენობის იტერაციებისთვის,



ნახ. 6.1. MARL-ის გამოყენება ფედერირებული სწავლების სტრუქტურაში

თითოეული აგენტი სწავლობს განახლებულ ლოკალურ მოდელს. განახლებული ლოკალური მოდელები გადადის ცენტრალურ სერვერზე და გროვდება იქ გლობალური მოდელის განახლებისთვის.

მომხმარებლის მონაწილეობა:

FL ტრენინგის პროცესის თითოეულ საკომუნიკაციო ციკლში მომხმარებელთა შესაბამისი რაოდენობის შერჩევა ვადამწყვეტია კონვერგენციის დაჩქარებისთვის. ამ მიზნით, ჩვენ განვიხილავთ მომხმარებლის ნაწილობრივი მონაწილეობის შესაძლებლობას FL ალგორითმში აგრეგაციის ყოველი რაუნდის განმავლობაში. ჩვენ ვვარაუდობთ, რომ მონაწილე მომხმარებლების მოცემული რაოდენობისთვის, ალბათობა იმისა, რომ კონკრეტული მომხმარებელი შეირჩევა ტრენინგის კონკრეტულ რაუნდში მონაწილეობის მისაღებად, ერთნაირია. ცენტრალურ სერვერზე, მხოლოდ მონაწილე მოწყობილობების ლოკალური მოდელის წონებია აგრეგირებული და მხოლოდ მონაწილე მოწყობილობები თითოეულ აგრეგაციის ციკლში იღებენ მოდელის განახლებას შემოთავაზებული სტრუქტურის მიხედვით. ამრიგად, RL აგენტმა, რომელიც განლაგებულია თითოეულ SeU-ში, არ არის საჭირო იცოდეს სხვა SeU-ების მონაცემთა ნიმუშები ან იყოს მათზე დამოკიდებული.

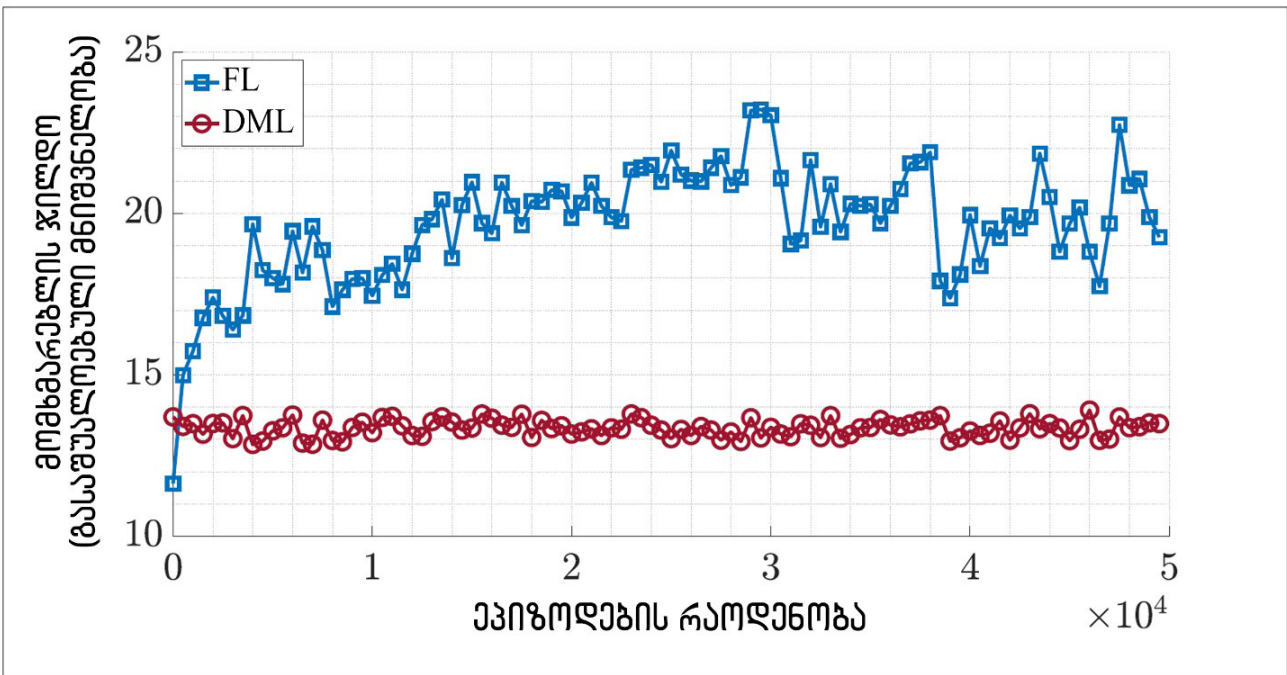
კომპიუტერული სიმულაციის შედეგები:

ჩვენი სიმულაციის პარამეტრებში შემთხვევითი წესით ვათავსებთ რვა BS-UE წყვილს $400 \text{ მ} \times 400 \text{ მ}$ ფართობზე და ვაკონფიგურირებთ ოთხ განსხვავებულ სიხშირულ ზოლს, როგორც ხელმისაწვდომ საკომუნიკაციო არხებს. ჩვენ განვიხილავთ მხოლოდ დაუნლინკს, ანუ კომუნიკაციას რვა BS-იდან მათ შესაბამის UE-ებთან. ეს რვა BS მოქმედებს როგორც SeU გადამცემი (SeU Tx-ები), ხოლო რვა UE მოქმედებს როგორც SeU მიმღები (SeU Rx-ები). ალბათობა იმისა, რომ ოთხი არხიდან თითოეული დაიკავოს PrU-მ, დაყენებულია 20 პროცენტზე. ყოველ

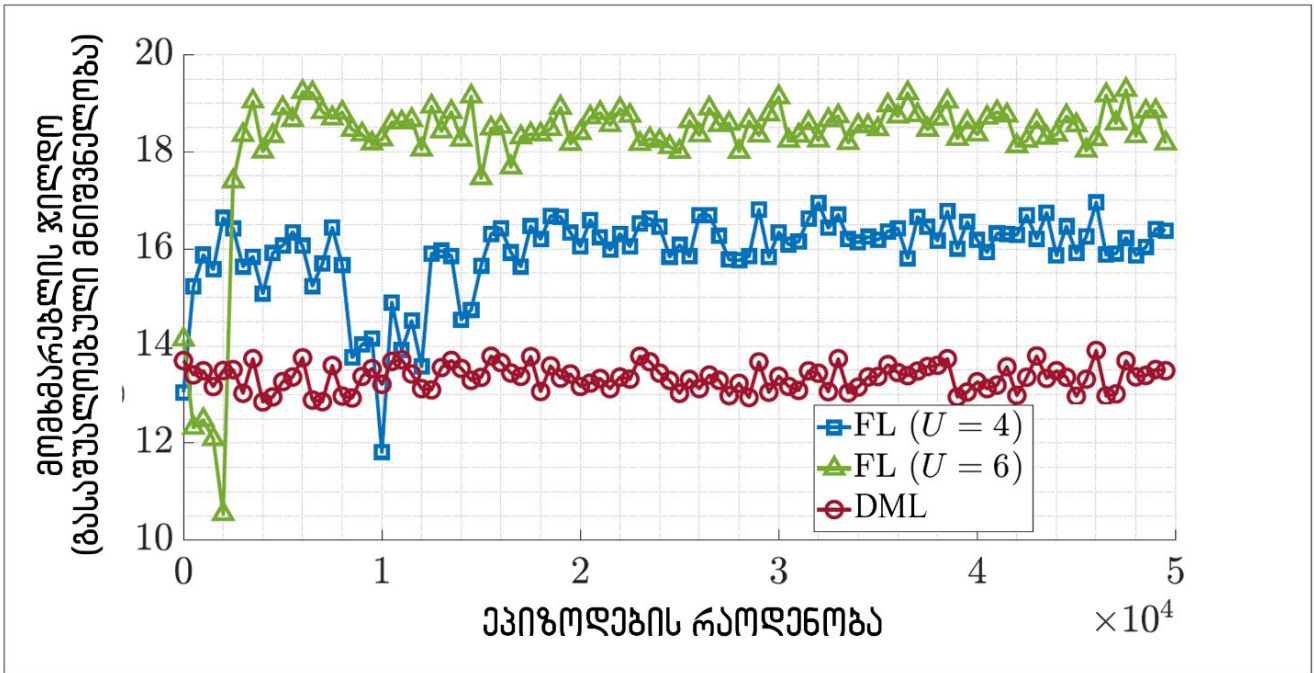
ტ საფეხურზე, პოლიტიკა ასახავს აგენტის დაკვირვებას მის მოქმედებაზე ისე, რომ აგენტი ირჩევს წვდომას ოთხი არხიდან ერთ-ერთზე ან რჩება უმოქმედო, რის შემდეგაც ის მიიღებს ჯილდოს. დაკვირვება შედგება გასაშუალოებული ისტორიული გამტარუნარიანობისგან წინა დროით საფეხურამდე და გამტარუნარიანობისგან წინა დროით საფეხურზე ყველა არხის მიხედვით. ჩვენ შევადარებთ ჩვენს MARL-ზე დაფუძნებული FL-ის ტრადიციულ განაწილებული სწავლების მიდგომას, სადაც ორივე იყენებს RL აგენტს, რომელიც იყენებს ორ-ფენიან წინაკავშირიან (feedforward) ნეირონულ ქსელს თავის ბირთვად. ტრადიციული DML-ის დაყენებისას, თითოეული SeU Tx იღებს მოდელს ცენტრალური სერვერიდან და იწყებს მისი ლოკალური მოდელის განახლებას ისე, რომ მისი ლოკალური ჯილდო მაქსიმალური იყოს, ყოველგვარი შემდგომი კომუნიკაციის გარეშე ცენტრალურ სერვერთან ან სხვა SeU Tx-ებთან.

მეორე მხრივ, MARL-ით მხარდაჭერილ FL-ს შეუძლია უზრუნველყოს უკეთესი საერთო მახასიათებელი სისტემურ დონეზე, რადგან ის იძლევა არაპირდაპირ თანამშრომლობას SeU Tx-ებს შორის ცენტრალურ სერვერზე მათი ინდივიდუალური ლოკალური მოდელის პერიოდული აგრეგაციის გზით. ტრადიციული DML-ის შემთხვევაში მომხმარებლები ეჯიბრებიან ერთმანეთს შებლუდული რესურსებისთვის, თანამშრომლობის გარეშე. მომხმარებლებს შეუძლიათ განახორციელონ ხუთი ქმედებიდან ერთი, კერძოდ, [0, 1, 2, 3, 4], სადაც 0 მიუთითებს, რომ არ არის არხზე წვდომა, ხოლო 1-დან 4-მდე ნებისმიერი ინდექსი წარმოადგენს ოთხი ხელმისაწვდომი არხიდან ერთ-ერთზე წვდომას. თითოეული აგენტის ლოკალური მოდელი დატრენინგებულია 50 დროითი საფეხურის შემდეგ, რაც განისაზღვრება, როგორც ერთი ეპიზოდი. გლობალური მოდელი FL-ში გროვდება ყოველი მეოთხე ეპიზოდის შემდეგ. მოდელის ეს პერიოდული აგრეგაცია და განახლება ხელს უწყობს ჯილდოების გრძელვადიან ჯამს, რომელიც დაკავშირებულია მომხმარებლის გამტარუნარიანობასთან. სხვა სიტყვებით რომ ვთქვათ, მოდელის აგრეგაცია საშუალებას აძლევს მომხმარებლებს „მოიხედონ“ სხვა მომხმარებლების გარემოში და უზრუნველყოფს მომხმარებელთა თანამშრომლობას.

ნახ. 6.2 გვიჩვენებს, რომ MARL-ით მხარდაჭერილი FL იწვევს უფრო მაღალ (ლოკალურ) ჯილდოს, რომელიც საშუალოდ არის მომხმარებლებში და, შესაბამისად, უფრო მაღალ ერთობლივ ჯილდოს DML-თან შედარებით, რაც ასევე გულისხმობს მომხმარებლის მთლიან გამტარუნარიანობას, თუმცა FL მოითხოვს SeU-ებს შორის კომუნიკაციის უფრო მეტ რაუნდს, Tx-ებსა და ცენტრალურ სერვერს შორის, ჩვეულებრივ DML-თან შედარებით. ეს მიუთითებს გარკვეულ კომპრომისზე კომუნიკაციის ზედნადებ ხარჯებსა და მისალწევ გამტარუნარიანობას შორის DML/FL ტექნიკის გამოყენებისას.



ნახ. 6.2. მომხმარებლის ჯილდოს გასაშუალოებული მნიშვნელობები FL-ში და ჩვეულებრივ DML-ში



ნახ. 6.3. მომხმარებლის ჯილდოს გასაშუალოებული მნიშვნელობების დამოკიდებულება მონაწილე მომხმარებელთა რაოდენობაზე

ნახ. 6.3 განიხილავს FL სტრუქტურას სხვადასხვა რაოდენობის მონაწილე მომხმარებლისთვის $U (< N)$ თითოეული აგრეგაციის რაუნდში, ანუ ყოველ მეოთხე ეპიზოდში. ეს აჩვენებს, რომ რაც უფრო მეტია მონაწილე მომხმარებელთა რაოდენობა, მით უფრო მაღალია მომხმარებლის საშუალო ჯილდო, და შესაბამისად, სისტემის გამტარუნარიანობა. იმავდროულად, ეს ნაწილობრივი მონაწილეობის მექანიზმი საშუალებას აძლევს სტრუქტურას იყოს უფრო მოქნილი მომხმარებლების არჩევისას უკეთესი არხის პირობებით, ენერჯის უფრო მოქნილი შეზღუდვით და საკმარისი გამოთვლითი რესურსებით. ჩვენს სიმულაციაში გამოყენებული პარამეტრების სრული ნაკრები შეჯამებულია ცხრილში 6.1. გამოკვლეული ეპიზოდების ჯამური რაოდენობა MARL ალგორითმში დაყენებულია 50000-ზე.

6.5. 6G-ში DML-ის გამოყენების მთავარი გამოწვევები

რთული ბალანსი დისტანციურ ღრუბლოვან სერვერსა და პერიფერიულ კვანძს შორის (რომელიც საბოლოო მომხმარებლებს მაღალი QoS-ით უზრუნველყოფს) მოითხოვს გამოთვლების ძალიან დიდ რაოდენობას, ამასთან უკიდურესად დაბალი შეყოვნების მიღწევას, რაც მთავარ გამოწვევად რჩება 6G გამოყენების უმეტეს შემთხვევაში, სადაც DML-ს და განსაკუთრებით FL-ს შეუძლია პოტენციურად ფუნქციონირება. ქვემოთ ჩვენ წარმოვადგენთ მოკლე მიმოხილვას სხვა დაკავშირებული ღია გამოწვევების შესახებ, რომელთა წინაშე შეიძლება აღმოჩნდეს DML და, კერძოდ, FL სტრატეგიები 6G-ისთვის.

პოტენციურად დიდი რაოდენობის ჰეტეროგენული მოწყობილობების შემცველი განაწილებული უსადენო ქსელის ერთ-ერთი ყველაზე გამორჩეული თვისება არის ამ მოწყობილობების მობილურობის შესაძლებლობა იმდენად, რამდენადაც საკმარისად მაღალი სიჩქარით მობილურობამ, თუნდაც მოწყობილობების ქვეჯგუფისთვის, შეიძლება წარმოადგინოს ტრენინგისა და ტესტირების მონაცემების განაწილებები, რომლებიც მნიშვნელოვნად განსხვავდება. მიუხედავად იმისა, რომ ისეთი მიდგომები, როგორცაა დომენის ადაპტაცია, შეიძლება გამოყენებულ იქნეს დასკვნის მიღების გასაუმჯობესებლად ასეთი ტრენინგი-ტესტის შესაბამისობის არსებობის შემთხვევაში, მისი დიდი მასშტაბით დანერგვა მოწყობილობაზე მისაღები გამოთვლითი სირთულით კვლავ ღია პრობლემაა. ეს ეხება არა მხოლოდ FL-ს, არამედ ნებისმიერ DML მიდგომას, რომელშიც სტატისტიკურად განსხვავებული ტრენინგის და ტესტის მონაცემთა ნაკრები არის შესაძლებელი.

FL სასწავლო პროცესში მონაწილე მოწყობილობების მაქსიმალური შესაძლო რაოდენობის შენარჩუნება ყოველთვის პრობლემაა, განსაკუთრებით არასაიმედო უსადენო გარემოში. გარდა ამისა, ენერჯის დაზოგვის მიზნით, ბატარეაზე მომუშავე IoT მოწყობილობები აერთიანებს DML-ის შესაფერის სტრატეგიებს (მაგალითად, გარკვეული ტრენინგის ტურებზე უარის თქმას). მიუხედავად იმისა, რომ ლოკალური მომხმარებლის მონაცემების კონფიდენციალურობის უზრუნველყოფა FL-ის გამოჩენული მახასიათებელია, თავდამსხმელებს შეუძლიათ მაინც შეაგროვონ სისტემის კრიტიკული ინფორმაცია მოდელის ცვლილებებიდან. მიუხედავად იმისა, რომ უფრო ახალი ისეთი მეთოდები, როგორებიცაა უსაფრთხო მრავალმხრივი გამოთვლა, დიფერენციალური კონფიდენციალურობა და უსაფრთხო აგრეგაცია, მიზნად ისახავს გააუმჯობესოს FL კონფიდენციალურობა, ეს მიდგომები ზოგადად სწირავენ დასკვნის მახასიათებლებს სათანადო კონფიდენციალურობას. ამ ხარჯების გაგება და დაბალანსება წარმოადგენს მნიშვნელოვან გამოწვევას კერძო FL სისტემების დანერგვისას, როგორც თეორიულად, ასევე პრაქტიკულად.

პარამეტრი	მნიშვნელობა
არხის ბატარეის ზოლი	10 მჰს
გზაში დანაკარგების მოდელი	$41 + 22.7 \log_{10}(d[\text{მ}])$ დბ
მცირემასშტაბიანი უაღიანი	რაიისის განაწილება
პირდაპირი ხელშის გზის კოეფიციენტი	5
ხმაურის სპექტრული სიმკვრივე	- 174 (დბმ/ჰს)
გამოკვლეული ეპიზოდების ჯამური რაოდენობა	50000
დროითი ბიჯები თითო ეპიზოდზე	50
სწავლების ლოკალური კოეფიციენტი	0.01
შემცირების მაჩვენებელი	0.9

ცხრილი 6.1. კომპიუტერული სიმულაციის პარამეტრები

მიუხედავად იმისა, რომ სინქრონული FL მოდელი უზრუნველყოფს კონვერგენციის უკეთეს გარანტიებს, ის მგრძობიარე სტრაგლერის (Straggler) ეფექტის მიმართ. ასინქრონული FL მოდელი უფრო შესაფერისია პრაქტიკაში, განსაკუთრებით მაშინ, როდესაც საბოლოო მოწყობილობები განსხვავდება ტექნოლოგიით, ქსელის კავშირის საიმედოობით და ბატარეის სიმძლავრის თვალსაზრისით, რაც იწვევს სისტემის პარამეტრებში მნიშვნელოვან ჰეტეროგენულობას მთელ ქსელში. საჭიროა თეორიული გამოკვლევა პოპულარული ალგორითმების კონვერგენციის საზღვრებზე, როგორიცაა სტოქასტური გრადიენტული დაღმართი (SGD), რომელიც შეიძლება იყოს შესაფერისი სხვადასხვა აპლიკაციისთვის. არსებული კვლევების უმეტესობა, რომელიც ანალიზებს FL-ს, განიხილავდა დამოუკიდებელი და თანაბარი განაწილების დაშვებას ლოკალური

მომხმარებლის მონაცემებისთვის. ზოგიერთ ნაშრომში შესწავლილი იყო სხვა განაწილებების შემთხვევებიც, ასინქრონული კომუნიკაციის შემცირების მეთოდებით კონფიდენციალურობის პირობებში. თუმცა, თეორიული და გამოცენებითი ანალიზის ხარისხი, რომელიც FL-ის მონაცემებზე დაფუძნებული ასეთი ტიპის განაწილებებზეა მორგებული, ჯერ კიდევ შესასწავლია.

6.6. მუშაობის ტიპის დასკვნა

ამ თავში ასახულია 5G ქსელების პერსპექტიული ხედვა, რომელიც ხაზს უსვამს გამოცენების სპეციფიკურ შემთხვევებს, რომლებიც აფართოებს ან განაახლებს 5G NR-ში დანერგილ ტექნოლოგიებს. უსადენო ქსელების თანდაყოლილი შეზღუდვებისა და მახასიათებლების სპეციფიკაციების გამო, რომელიც განსაკუთრებით მკაცრი იქნება 5G-ში, DML-მა, როგორც პარადიგმამ შეიძლება მნიშვნელოვანი როლი შეასრულოს ახალი აპლიკაციების ჩართვაში. როგორც კონკრეტული მაგალითი, ჩვენ ვიყენებთ FL-ს MARL-ით სპექტრზე დინამიკური წვდომის პრობლემის გადასაჭრელად და სიმულაციების საშუალებით პერსპექტიული შედეგების დემონსტრირებისთვის. MARL-ით მხარდაჭერილი FL კარგია 5G გამოცენების შემთხვევებისთვის, რომელიც ეყრდნობა განაწილებული მომხმარებლების დიდი რაოდენობის თანამშრომლობას. ასევე წარმოდგენილია გამოწვევების შესაბამისი მაგალითები და პოტენციური სამომავლო მიმართულებები DML და FL მიდგომების გამოცენებისთვის 5G ქსელებში.

თავი 7 ფედერირებული ღრმა განმტკიცებული სწავლება 6G ქსელებში ღია RAN-ის დაყოფისთვის

7.1. შესავალი

მოსალოდნელია, რომ 6G ქსელები ხელს შეუწყობს მონაცემთა ტრაფიკის ფეთქებად ზრდას, რომელიც გენერირებულია დიდი რაოდენობის დაკავშირებული მოწყობილობების მიერ. ეს მოწყობილობები უზრუნველყოფს მოწინავე სერვისებს სხვადასხვა ვერტიკალურ ინდუსტრიაში მომსახურების ხარისხის დივერსიფიცირებული მოთხოვნებით. აქედან გამომდინარე, ტრაფიკის მოცულობა შემდეგი თაობის ქსელებში ექსპონენციალურად გაიზრდება, რაც მოითხოვს უზარმაზარ რაოდენობის მონაცემების გადატანას. ეს გაზრდის გადაცემის სიჩქარეზე მოთხოვნას, რათა უზრუნველყოს კავშირი, რომელიც საჭიროა ამ მონაცემების ეფექტიანად და სწრაფად გადასაცემად; შესაბამისად, ეს ქსელები დაექვემდებარება უფრო მკაცრ მოთხოვნებს, ვიდრე არსებული 5G ქსელები. ყველა ამ სერვისის მხარდასაჭერად, მობილური ქსელის ოპერატორებს (MNO) მოეთხოვებათ უზრუნველყონ ადეკვატური ქსელური ინფრასტრუქტურა. ამის მისაღწევად მაღალი ტექნიკური ეფექტიანობით და ხარჯთეფექტური გზით, MNO-ები ეყრდნობიან რამდენიმე უახლეს ტექნოლოგიას, როგორცაა, მაგალითად, ქსელის ფენებად დაყოფა. ეს საშუალებას იძლევა შეიქმნას რამდენიმე ლოგიკურად დამოუკიდებელი ქსელი, ე. წ. ქსელის ფენები, რომლებიც მოქმედებენ საერთო ფიზიკურ ინფრასტრუქტურაში. კერძოდ, RAN-ის დაყოფა შედგება RAN რესურსების დაყოფისგან, რათა შეიქმნას სხვადასხვა RAN ფენა, თითოეული მორგებული და განაწილებული კონკრეტული სერვისის მოთხოვნების დასაკმაყოფილებლად. MNO ფლობს RAN-ის ფიზიკურ რესურსებს, მათ შორის რადიორესურსებს, და იჯარით აძლევს მათ მობილური ქსელის ვირტუალურ ოპერატორებს (MVNO), რათა განათავსონ RAN ფენები მათი შეთავაზებული სერვისების საფუძველზე.

მომხმარებლებისთვის რადიორესურსების განაწილება არის ძალიან რთული ოპერაცია MVNO-სთვის. ეს ძირითადად გამოწვეულია რადიორესურსების სიმცირით და მათი სერვისების მიმართ არაერთგვაროვანი QoS მოთხოვნებით. ამ გამოწვევების დასაძლევად, სხვადასხვა ძალისხმევა დაეთმო RAN-ის დაყოფის პრობლემის ფორმულირებას ოპტიმიზაციის ტექნიკის გამოყენებით და მის გადაჭრას ევრისტიკული გზით. თუმცა, ეს მიდგომები შეიძლება არ ემთხვეოდეს ნულოვანი შეხების (zero touch) ქსელის პერსპექტივებს, რადგან მათ არ შეუძლიათ სწრაფად მოერგოს RAN გარემოში დინამიკურ და მუდმივ ცვლილებებს. ML ტექნიკას, კონკრეტულად DRL ალგორითმებს, შეუძლიათ ამ პრობლემის მოგვარება RAN-ის დაყოფის ოპერაციების მენეჯმენტში მეტი ავტომატიზაციის შემოტანით. სტოქასტური RAN-ის გარემო ფაქტორები, როგორცაა მომხმარებელთა სიმკვრივე, მომხმარებლის მოთხოვნები და უსადენო არხის გადაცემის პირობები დიდ გავლენას ახდენს DRL მოდელის სიზუსტეზე, რაც ამცირებს მომხმარებლებისთვის რადიორესურსების განაწილების ეფექტიანობას. მართლაც, როდესაც MVNO აშენებს რესურსების განაწილების თავის DRL მოდელს ტრენინგის მონაცემთა ნაკრების გამოყენებით, რომელიც დაკავშირებულია მხოლოდ მისი მომხმარებლების ქცევასთან და მის გარემომცველ გარემოსთან, DRL მოდელის სიზუსტე შეიძლება შეზღუდული იყოს. დივერსიფიცირებული მონაცემთა ნაკრებიდან სარგებლობისთვის, MVNO-ებს შეუძლიათ ითანამშრომლონ თავიანთი მონაცემების ერთმანეთთან გაზიარებით, რათა უზრუნველყონ მრავალფეროვანი და მაღალი ხარისხის მონაცემთა ბაზა. თუმცა, MVNO-ები ხშირად კონკურენტი სუბიექტები არიან და ნაკლებად სავარაუდოა, რომ სურთ თავიანთი მონაცემების გაზიარება კონფიდენციალურობასთან დაკავშირებული პრობლემების გამო. FL გაჩნდა როგორც მონაცემთა გაზიარების პერსპექტიული ალტერნატივა, რადგან ის ხელს უწყობს კონფიდენციალურობის შემანარჩუნებელ ერთობლივ ტრენინგს მოდელის განახლების გაზიარების გზით.

FL არის კოოპერატიული სწავლების მიდგომა, რომელშიც მრავალი თანამშრომელი, ჩვენს შემთხვევაში MVNO, ატრენინგებს ML მოდელს მათი პირად მონაცემთა ნაკრების გამოყენებით და შემდეგ აგზავნის თავის მომზადებულ მოდელს აგრეგაციის ერთეულში გლობალური მოდელის შესაქმნელად. აგრეგაციის ერთეული

უბრუნებს ამ მოდელს ყველა კოლაბორატორს გამოსაყენებლად ან შემდგომი ტრენინგისთვის. ამრიგად, FL საშუალებას აძლევს MVNO-ებს შექმნას ML რესურსების განაწილების რობასტული (მტკიცე) მოდელი მონაცემთა კონფიდენციალურობის შენარჩუნებით, რადგან მხოლოდ დატრენინგებული მოდელია გაზიარებული. შესაბამისი გამოცდილება საშუალებას მისცემს RAN-ის ფენებად დაყოფის მოდელს ისწავლოს სხვადასხვა სცენარიდან, რაც მას უფრო ადაპტირებულს ხდის გარემოს ცვლილებებთან. MVNO-ებში მომხმარებლების დაუბალანსებელი, არადამოუკიდებელი და არაერთგვაროვანი განაწილების გამო (non-i.i.d) და, ამასთან, მათი განსხვავებული რაოდენობისა და მოთხოვნების შესაბამისად, FL ხდება მიმზიდველი გადაწყვეტა რობასტული მოდელების შესაქმნელად.

FL-ზე დაფუძნებული RAN-ის ფენებად დაყოფის მექანიზმის ავტომატიზაციისა და მოქნილობის მაღალი დონის დასაწინააღმდეგებლად გაჩნდა O-RAN, როგორც რევოლუციური გადაწყვეტა (აღვნიშნავთ, რომ O-RAN აღიანსება შექმნა O-RAN არქიტექტურა, როგორც საფუძველი ვირტუალური RAN-ისთვის ღია აპარატურულ უბრუნველყოფაზე და ღრუბელზე). O-RAN ეყრდნობა ღიაობასა და ინტელექტს, რათა უბრუნველყოს მოქნილი და ეფექტიანი RAN სერვისები. ღიაობა გულისხმობს RAN-ის დაყოფას რამდენიმე დამოუკიდებელ ფუნქციად, შემდეგ კი ღია, სტანდარტიზებული და თავსებადი ინტერფეისების გამოყენებას ამ RAN ფუნქციებს შორის კომუნიკაციის უბრუნველსაყოფად. ამრიგად, შემცირდება საფირმო ალჭურვილობის გამოყენების აუცილებლობა, რაც გამოიწვევს უფრო მეტ კონკურენციას მომწოდებლებს შორის და შემდგომ ინოვაციებს. ასევე, O-RAN ხელს უწყობს მეტი ინტელექტისა და ავტომატიზაციის გამოყენებას RAN ოპერაციების მართვაში RAN ინტელექტუალური კონტროლერების (RIC) მეშვეობით. RIC-ები იყენებენ ML-ის შესაძლებლობებს RAN-ის რესურსებისა და კომპონენტების მართვისთვის ისე, რომ ის გახდეს ნულოვანი შეხების (zero touch) ქსელის სისტემა.

ქვემოთ წარმოდგენილი სამუშაოს მიზანია შექმნას DRL-ზე დაფუძნებული რადიორესურსების განაწილების საიმედო მოდელი MVNO-ებისთვის მათი ჰეტეროგენული და მრავალფეროვანი მონაცემთა ნაკრების გამოყენებით მონაცემთა კონფიდენციალურობისა და უსაფრთხოების შენარჩუნებით. შესაბამისად, ჩვენ განვიხილავთ FL-ზე დაფუძნებულ რადიორესურსების განაწილების კოპერატიულ მექანიზმს MVNO-ებისთვის. ამ მექანიზმში, თითოეული MVNO ატრენინგებს DRL რადიორესურსების განაწილების მოდელს მისი მომხმარებლების მოთხოვნების შესაბამისად და აგზავნის მომზადებულ მოდელებს RIC-ში აგრეგაციისთვის. შემდეგ RIC უკან უგზავნის გლობალურ DRL მოდელს თითოეულ MVNO-ს, რათა განაახლოს მისი ლოკალური DRL მოდელები. RAN-ის ფენებად დაყოფის შემოთავაზებული მიდგომა შეესაბამება ნულოვანი შეხების ქსელის ჩარჩოს, რადგან ის მიზნად ისახავს მომხმარებლებისთვის რადიორესურსების განაწილების ავტონომიურ მართვას, განაწილების პოლიტიკის ადაპტირებით გარემოს შეზღუდვებთან და მომხმარებელთა მოთხოვნებთან. მართლაც, DRL-ის გამოყენება იძლევა RAN-ის ფენებად დაყოფის ოპერაციებში ჩაკეტილი მარყუჟის კონტროლის პროცესის განხორციელების საშუალებას. გარდა ამისა, ამ თავში შემოთავაზებული RAN-ის ფენებად დაყოფის მექანიზმი დანერგულია O-RAN არქიტექტურაში, რომელიც ემხრობა მეტი ინტელექტისა და ავტომატიზაციის გამოყენებას RAN-ის ოპერაციების მონიტორინგში.

ამ თავის ძირითადი შედეგები შეიძლება შეჯამდეს შემდეგნაირად:

- ჩვენ განვიხილავთ DRL-ზე დაფუძნებულ RAN-ის ფენებად დაყოფის ძალისხმევას და დაკავშირებულ გამოწვევებს მრავალ MVNO-იან გარემოში.
- ჩვენ ვქმნით ფედერირებულ DRL (FDRL) მექანიზმს O-RAN არქიტექტურაზე, MVNO-ების რადიორესურსების განაწილების ფუნქციონირების გასაუმჯობესებლად.
- ფართო და დეტალური კომპიუტერული სიმულაციების საშუალებით ნაჩვენებია, რომ შემოთავაზებული RAN-ის ფენებად დაყოფის მექანიზმი საშუალებას იძლევა უკეთ განაწილდეს საჭირო რადიორესურსები მომხმარებლის QoS მოთხოვნების დასაკმაყოფილებლად შეყვანებისა და მონაცემთა გადაცემის სიჩქარის თვალსაზრისით.

7.2. DRL, როგორც RAN-ის ფენებად დაყოფის საშუალება

არსებულ ლიტერატურაში DRL ფართოდ გამოიყენება როგორც ეფექტიანი ინსტრუმენტი RAN-ის ფენებად დაყოფისთვის. მაგალითად, DRL გამოიყენება RAN-ის ფენების კონფიგურაციის პარამეტრების დასადგენად. იგი განსაზღვრავს RAN-ის თითოეული ფენისთვის გადაცემის სიძლიერეს, სიხშირის გატარების ზოლს (Bandwidth) (ამ უკანასკნელს ქვემოთ ყველგან მოვიხსენიებთ, როგორც გატარების ზოლს), გაზიარებულ გატარების ზოლს ფენებს შორის და ნუმეროლოგიას. ლიტერატურაში ასევე შემუშავებულია DRL-ზე დაფუძნებულ RAN-ის ფენებად დაყოფის მექანიზმი, რომელსაც იყენებენ BS-ები რადიორესურსების გასანაწილებლად მათთან ასოცირებულ მომხმარებლებთან. როდესაც BS-ის რესურსები არ არის საკმარისი, ეს მექანიზმი მოიცავს სხვა BS-ებისგან დამატებითი რესურსების მოთხოვნას მისი მომხმარებლების QoS მოთხოვნების დასაკმაყოფილებლად.

ასევე ახლახან იქნა წარმოდგენილი DRL-ზე დაფუძნებული კვლევები RAN-ის ფენებად დაყოფის ოპერაციის გასაუმჯობესებლად მრავალ MVNO-იან სცენარში. მაგალითად, შემოთავაზებულია ორდონიანი რესურსების განაწილების სქემა DRL-ის გამოყენებით. MNO-დან მიღებულ რესურსებზე დაყრდნობით, თითოეული MVNO თავის მომხმარებლებისთვის ანაწილებს საჭირო რადიორესურსს მათი QoS-ის დასაკმაყოფილებლად. MVNO იყენებს DRL ალგორითმს თავისი მომხმარებლებისთვის რესურსების გამოსაყოფად მათი კმაყოფილების მაჩვენებლის მაქსიმალურად გაზრდის მიზნით, რაც განისაზღვრება ქსელიდან წარმატებით მიღებული პაკეტების რაოდენობით. ასევე, წარმოდგენილია DRL-ზე დაფუძნებული რესურსების განაწილების ალგორითმი MVNO-ების სარგებლიანობის მაქსიმალურად გაზრდის მიზნით, რომლის მიხედვითაც უნდა მოხდეს RAN ფენების QoS მოთხოვნების დაკმაყოფილებაც. MVNO-ს სარგებლიანობა განისაზღვრება, როგორც მოგება ფენაზე რესურსის განაწილებიდან მინუს MNO-დან რესურსის იჯარით აღების ღირებულება. RAN ფენების QoS მოთხოვნები განისაზღვრება მონაცემთა გადაცემის სიჩქარისა და შეყოვნების თვალსაზრისით.

ყველა ამ მიდგომაში, DRL გამოიყენებოდა, როგორც RAN-ის ფენებად დაყოფის ორგანიზების საშუალება. თუმცა, DRL მოდელის განზოგადების შესაძლებლობა RAN-ის ფენებად დაყოფის სხვადასხვა გარემოზე შეიძლება იყოს შეზღუდული მათი ძლიერი დამოკიდებულების გამო იმ გარემოს მონაცემებზე, რომელშიც ისინი გადიან ტრენინგს. შემოთავაზებული FDRL-ზე დაფუძნებული RAN-ის ფენებად დაყოფის მექანიზმი გადალახავს რობასტულობის პრობლემას, რომელიც შეიძლება წარმოიშვას მხოლოდ DRL-ზე დაფუძნებულ მიდგომებში. FDRL საშუალებას აძლევს MVNO-ებს გამოიყენონ ტრენინგის უფრო მრავალფეროვანი საშუალებები DRL-ისთვის, და ამასთან ერთად თავიდან აიცილონ მონაცემების გაზიარება თავიანთ RAN გარემოში.

ბევრმა მიდგომამ აჩვენა, რომ DRL უზრუნველყოფს RAN-ის ეფექტიან დაყოფას. მიუხედავად ამისა, ზოგიერთი საკითხი გადაუჭრელი რჩება და იმსახურებს შემდგომ შესწავლას.

რობასტულობა:

პოლიტიკის შემსწავლელი (Policy-learning) მეთოდების მთავარი ნაკლი, როგორცაა DRL, არის მათი დამოკიდებულება მონაცემებზე. რობასტული DRL მოდელის ტრენინგი დიდ მონაცემებს მოითხოვს. გასათვალისწინებელია, რომ RAN-ის გარემოს აქვს მაღალი გაურკვევლობა, რომელიც დაკავშირებულია მომხმარებლის სხვადასხვა მოთხოვნასთან, უსადენო გადაცემის დინამიკურ პირობებთან და ტრაფიკის მუდმივ რყევებთან. შესაბამისად, ასეთ გარემოში ტრენინგი შეიძლება დაზარალებული იყოს მონაცემთა სიმცირის გამო. ამრიგად, საჭიროა ჰეტეროგენული სცენარების უზრუნველყოფა DRL-ზე დაფუძნებული RAN-ის ფენებად დაყოფის რობასტული მოდელის შესაქმნელად.

SLA-ის დაკმაყოფილება:

RAN-ის ფენებად დაყოფისას, სერვისის დონის შეთანხმება (SLA) წარმოადგენს მინიმალურ QoS მოთხოვნებს, რომლებიც გარანტირებული უნდა იყოს მიწოდებულ სერვისზე. 5G/6G ქსელებში, RAN-ის ფენებად დაყოფა სავარაუდოდ ასოცირდება სერვისის დონის უფრო დეტალურ შეთანხმებებთან. გარდა ამისა, რესურსები RAN გარემოში შეზღუდულია, რაც იწვევს რესურსების დეფიციტის პრობლემებს. ამიტომ, ქსელის ოპერატორებმა უნდა შეიმუშაონ DRL ალგორითმები, რომლებიც ოპტიმიზაციას გაუკეთებენ რესურსების განაწილებას SLA-ების დასაკმაყოფილებლად და ამით უზრუნველყოფენ სერვისის მაქსიმალურ დონეს.

RAN-ის ფენების იზოლაცია:

DRL-ზე დაფუძნებული RAN-ის ფენებად დაყოფის მექანიზმებს შეუძლია უზრუნველყოს RAN-ის ფენების მოქნილი მართვა. მათ შეუძლიათ სწრაფად შექმნან ან გადააკეთონ RAN-ის ფენები მომხმარებლების

მოთხოვნების დასაკმაყოფილებლად. თუმცა, RAN-ის ფენების სწრაფი კონფიგურირების მიღწევა შესაძლებელია მხოლოდ მათი იზოლაციის ხარჯზე. ამიტომ, DRL მექანიზმებმა უნდა განიხილონ ურთიერთკომპრომისი სწრაფ მართვასა და RAN-ის ფენების ბუსტ იზოლაციას შორის.

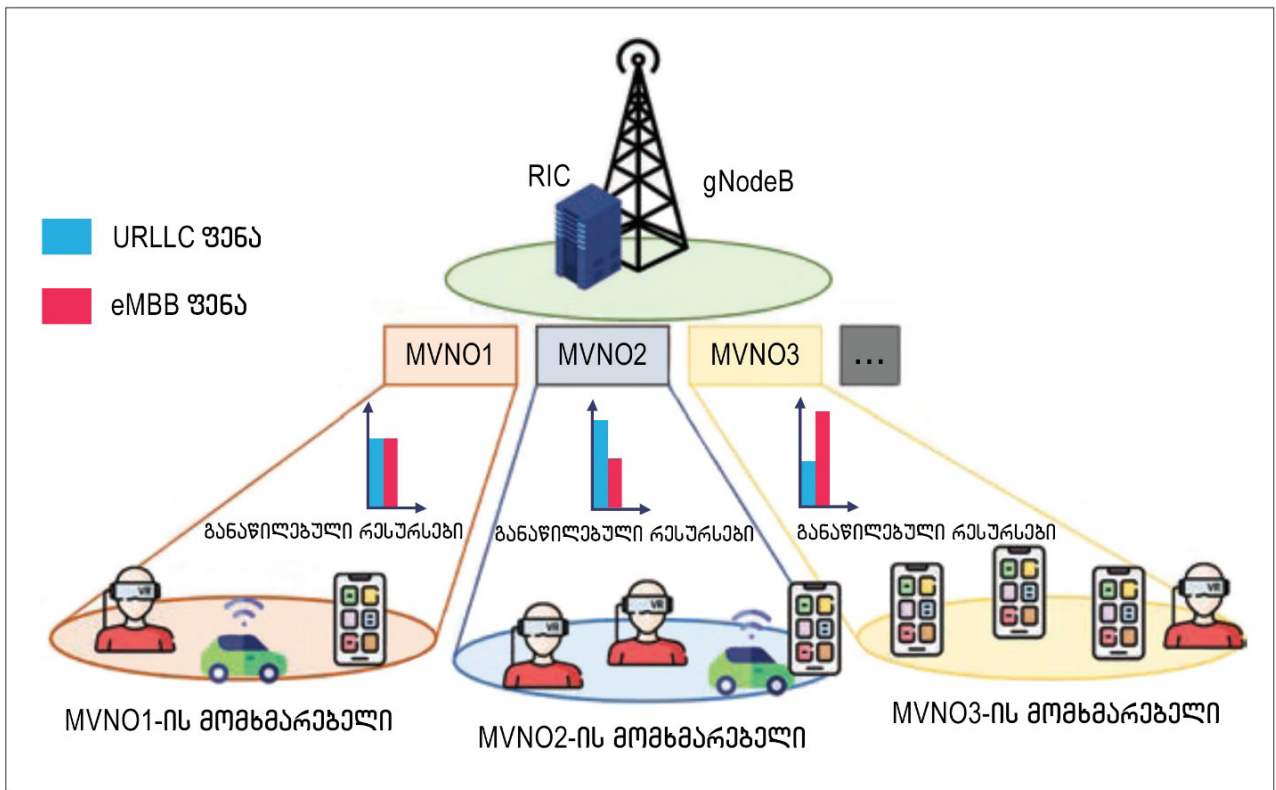
კონფიდენციალურობა და კონკურენცია:

მიუხედავად იმისა, რომ სასურველია MVNO-ების წახალისება თანამშრომლობისთვის, რათა შექმნან ფენებად დაყოფის რობასტული მოდელები, ჩნდება მრავალი პრობლემა კონფიდენციალურობის კუთხით. მართლაც, MVNO-ებს არ აქვთ სტიმული, გაუზიარონ კონკურენტებს მათი მომხმარებლების QoS-თან დაკავშირებული მონაცემები. ამის ნაცვლად, საჭიროა ახალი მეთოდები, რომლებიც ხელს შეუწყობენ როგორც თანამშრომლობას, ასევე კონფიდენციალურობას, რათა უზრუნველყოს მაღალი SLA შესაბამისობა.

შემდეგ პარაგრაფებში ჩვენ შევისწავლით FDRL მექანიზმს რადიორესურსების გასანაწილებლად მრავალ MVNO-იან გარემოში. ჩვენ გამოვიყენებთ FL-ს, რობასტულ და კერძო DRL-ზე დაფუძნებული RAN-ის ფენებად დაყოფის მექანიზმის შესაქმნელად, რომელიც აკმაყოფილებს MVNO-ების მომხმარებლებს SLA-ის თვალსაზრისით. შემოთავაზებული მექანიზმის რობასტულობა მიიღწევა მას შემდეგ, რაც MVNO-ები ითანამშრომლობენ გლობალური DRL მოდელის მოსამზადებლად. MVNO-ები იზიარებენ ლოკალურად დატრენინგებულ მოდელებს მათი სცენარების საფუძველზე, რათა შექმნან RAN-ის ფენებად დაყოფის რობასტული გლობალური მოდელი.

7.3. FDRL -ზე დაფუძნებული კოოპერატიული RAN-ის ფენებად დაყოფა MVNO-ებს შორის

სისტემის მოდელი წარმოდგენილია ნახ. 7.1-ზე. ჩვენ განვიხილავთ RIC-ით მხარდაჭერილ RAN არქიტექტურას ერთი BS-ით (ეს არის gNodeB), რომელსაც ფლობს MNO. BS მუშაობს გატარების მთლიან B ზოლში. MNO პასუხისმგებელია MVNO-ების ნაკრების მომსახურებაზე თითოეულ მათგანზე გაზიარებული გატარების ზოლის ნაწილის იჯარით აღებით. თითოეულ MVNO-ს ჰყავს მომხმარებლების ნაკრები, რომლებიც ატვირთავენ თავიანთ პაკეტებს ქსელში. ჩვენ განვიხილავთ მომხმარებელთა ორ ტიპს, კერძოდ, eMBB-ს და URLLC-ს.



ნახ. 7.1. სისტემის მოდელის წარმოდგენა

ვვარაუდობთ, რომ MVNO-ებისთვის გატარების ზოლის განაწილება უკვე შესრულებულია MNO-ს მიერ. MVNO თავის თითოეულ მომხმარებელს ანიჭებს MNO-ს მიერ წინასწარ განაწილებული გატარების ზოლის ნაწილს, რათა დააკმაყოფილოს მისი QoS მოთხოვნები პაკეტების ატვირთვისას. მართლაც, თითოეული MVNO-ს მიზანია ამ ზოლის ეფექტიანად განაწილება, რათა უზრუნველყოს მონაცემთა გადაცემის მაღალი სიჩქარე eMBB მომხმარებლებისთვის და გადაცემის დაბალი შეყოვნება URLLC მომხმარებლებისთვის. ამ თავში ჩვენ წარმოვადგენთ FDRL-ზე დაფუძნებულ RAN-ის ფენებად დაყოფის მექანიზმს MVNO-ებისთვის, რომლებიც თანამშრომლობენ თავიანთი მომხმარებლებისთვის საკომუნიკაციო რესურსების განაწილების მახასიათებლის გასაუმჯობესებლად.

RAN-ის გარემო უადრესად დინამიკურია, სადაც მოძრაობის რყევები მუდმივია და უსადენო გადაცემის პირობები ხშირად იცვლება. ამრიგად, MVNO-სთვის რთულია ასეთ გარემოში რადიორესურსების ფენებად დაყოფის ოპერაციის შესრულება. ამ პრობლემის დასაძლევად შეიძლება გამოყენებულ იქნეს ML-ის ისეთი ტექნიკა, როგორცაა DRL. DRL ალგორითმზე დაფუძნებული MVNO-ების RAN ფენებად დაყოფის პრობლემის გადასაჭრელად, ჩვენ გვჭირდება მისი მოდელირება მარკოვის გადაწყვეტილების პროცესით (MDP).

MDP-ის ფორმულირება:

თითოეული MVNO-სთვის, გატარების ზოლის განაწილების პრობლემა მოდელირებულია როგორც ერთაგენტიანი MDP შემდეგნაირად:

- მდგომარეობათა სივრცე: აგენტის მიერ დაკვირვებული მდგომარეობა მოიცავს მისი მომხმარებლების ტიპს და მათ არხში გაძლიერებას (channel gain). მომხმარებლების ტიპი (მაგალითად, eMBB და URLLC) აუცილებელია თითოეული მომხმარებლის მოთხოვნების დასადგენად. არხში გაძლიერება მომხმარებელსა და gNodeB-ს შორის გამოითვლება gNodeB-ის მიერ შეგროვებული არხის მდგომარეობის ინფორმაციის გამოყენებით. ვინაიდან ყოველი MVNO შეიძლება ემსახურობდეს მომხმარებლის სხვადასხვა რაოდენობას დროის თითოეულ მომენტში, ის აჩენს პრობლემას, რომელიც დაკავშირებულია დატრენინგებული მოდელის შესასვლელი და გამოსასვლელი მონაცემების ზომასთან. ამ პრობლემის დასაძლევად, ჩვენ დავაყენეთ შესასვლელი და გამოსასვლელი მონაცემების ზომა მომხმარებელთა მაქსიმალური რაოდენობის მიხედვით, რომელსაც შეუძლია MVNO-ს მოემსახუროს ერთდროულად. იმ შემთხვევაში, თუ დაკვირვების შედეგად განსაზღვრული მომხმარებლების რაოდენობა ნაკლებია მომხმარებლის ამ მაქსიმალურ ზღვრულ ოდენობაზე, ჩვენ ვიყენებთ ნულოვან შევსებას. ეს საშუალებას გვაძლევს მოვახდინოთ თითოეული MVNO-ს მომხმარებელთა განსხვავებული რაოდენობის ადაპტირება და დატრენინგებული მოდელის გაერთიანება.
- სამოქმედო სივრცე: აგენტმა უნდა გადაწყვიტოს, თუ რა ოდენობის გატარების ზოლი უნდა გაუნაწილდეს მის თითოეულ მომხმარებელს. მომხმარებლისთვის შერჩეული ქმედება არის რეალური მნიშვნელობა 0-სა და ზღვრულ შორის, რომელიც განსაზღვრავს გატარების ზოლის იმ მაქსიმალურ მნიშვნელობას, რომელიც შეიძლება ჰქონდეს მომხმარებელს.
- ჯილდოს ფუნქცია: ამ თავში წარმოდგენილი სამუშაოს მთავარი მიზანია უზრუნველყოს მონაცემთა გადაცემის მაღალი სიჩქარე eMBB მომხმარებლებისთვის და გადაცემის დაბალი შეყოვნება URLLC მომხმარებლებისთვის. ჯილდო დამოკიდებულია თითოეული მომხმარებლისთვის უზრუნველყოფილ QoS-ზე, მონაცემთა გადაცემის სიჩქარისა და შეყოვნების თვალსაზრისით. შესაბამისად, აგენტის მიერ მიღებული ჯილდო არის eMBB მომხმარებლების მიერ მიღწეული მონაცემთა გადაცემის სიჩქარეების მთლიანი ჯამი და URLLC მომხმარებლების მიერ განცდილი შეყოვნების შებრუნებული მნიშვნელობა გადაცემისას. თუ eMBB მომხმარებლის მონაცემთა გადაცემის მიღწეული სიჩქარე ნაკლებია მინიმალურ ზღვრულზე, ან თუ URLLC მომხმარებლის მიერ განცდილი შეყოვნება აღემატება ზღვრულის მაქსიმალურ მნიშვნელობას, ეს ქმედება დაჯარიმდება -0.1 -ით. ეს ჯარიმები დაემატება აგენტის მიერ მიღებულ ჯილდოს. უფრო მეტიც, თუ გატარების ზოლის ნაწილების წილადების ჯამი 1-ზე მეტია, გლობალური ქმედება ჩაითვლება არასწორად და ის დაჯარიმდება -0.05 -ით. გარდა ამისა, იმ შემთხვევის თავიდან აცილების მიზნით, როდესაც გატარების ზოლის ნაწილი ენიჭება მომხმარებელს, რომელიც არ არსებობს, ჩვენ ამ მოქმედებას ვუკავშირებთ დაჯარიმებას, რომელიც უარყოფითი მნიშვნელობის ტოლია და რომელსაც ვამატებთ ჯილდოს (-0.1).

DRL იძლევა საშუალებას შეიქმნას აგენტები, რომლებიც სწავლობენ მრავალგანზომილებიან მდგომარეობებს, სადაც პოლიტიკა წარმოდგენილია ღრმა ნეირონული ქსელის სახით. DRL პირველად დაინერგა DQN-ის მეშვეობით და სწრაფად იქნა ატაცებული კვლევითი საზოგადოების მიერ მრავალი პრაქტიკული გადაწყვეტილების პრობლემის გადასაჭრელად. თუმცა, DQN წარმოადგენს პოლიტიკის გარეთ სწავლებას (Off-policy learning) და შეიძლება კარგად არ მუშაობდეს მაღალი გაურკვევლობის მქონე გარემოში, როგორცაა უსადენო ქსელები. აღსანიშნავია, რომ ღირებულებებზე დაფუძნებული RL ალგორითმები, როგორცაა Q-learning, ჯერ ოპტიმიზაციას უკეთებენ მნიშვნელობის ფუნქციას და შემდეგ იღებენ ოპტიმალურ გადაწყვეტილებებს, ხოლო პოლიტიკაზე დაფუძნებული მეთოდები პირდაპირ ოპტიმიზაციას უკეთებენ ჯილდოებზე დაფუძნებულ ობიექტურ ფუნქციას, რაც მათ შესაფერისს ხდის დიდი ან უსასრულო სამოქმედო სივრცეებისთვის. თუმცა, პოლიტიკაზე დაფუძნებულ RL-ს შეიძლება ჰქონდეს ხმაურიანი და არასტაბილური გრადიენტები. შედეგად, ჩვენ ვთავაზობთ აქტორზე და კრიტიკოსზე დაფუძნებული (actor-critic-based) ალგორითმის გამოყენებას. სინამდვილეში, „აქტორი-კრიტიკოსის“ მიდგომები აერთიანებს ძლიერ მხარეებს როგორც ღირებულებებზე დაფუძნებული, ასევე პოლიტიკაზე დაფუძნებული RL ალგორითმებიდან. გარდა ამისა, ვინაიდან წილადის მნიშვნელობები უწყვეტია, ჩვენ ვიყენებთ ღრმა DDPG ალგორითმს, რომელიც ერთდროულად სწავლობს Q-ფუნქციას და პოლიტიკას და ასრულებს მოქმედებებს უწყვეტი სივრციდან.

DDPG ალგორითმი:

იგი იყენებს ოთხ ნეირონულ ქსელს: აქტორის ქსელს, კრიტიკოსის ქსელს, აქტორის სამიზნე ქსელსა და კრიტიკოსის სამიზნე ქსელს. გარემოს მოცემულ დაკვირვებაში მყოფი მდგომარეობისთვის აქტორი ირჩევს მოქმედებას და კრიტიკოსი იყენებს მდგომარეობა-მოქმედების Q ფუნქციას მის შესაფასებლად. ტრენინგის ნომუშებს შორის კორელაციის შესამცირებლად, DDPG იყენებს გამოცდილების აღდგენის მესხიერებას, რომელიც ასტაბილურებს მის ქცევას. გამოცდილება განისაზღვრება გარემოს ამჟამინდელი მდგომარეობით, არჩეული მოქმედებით, მიღებული ჯილდოთი და გარემოს შემდეგი მდგომარეობით. აგენტი ინახავს თავის გამოცდილებას და შემდეგ არჩევს მათგან შემთხვევით მინი პაკეტებს, რათა მოამზადოს თავისი DQN. DDPG-ის საძიებო პოლიტიკა ხორციელდება ტრენინგის პროცესში მოქმედებებზე ხმაურის დამატებით. დამატებული ხმაური DDPG აგენტს საშუალებას აძლევს ეფექტიანად გამოიკვილოს თავისი გარემო. ჩვენ გამოვიყენეთ ორნშტეინ-უჰლენბეკის (Ornstein-Uhlenbeck) პროცესი ხმაურის მნიშვნელობების შესაქმნელად.

მომხმარებელთა განაწილების უთანასწორობა სხვადასხვა გეოგრაფიულ არეალში (მაგალითად, გარეუბნებში და ქალაქის ცენტრებში), არაადეკვატურს ხდის ერთი და იმავე მოდელის გამოყენებას ყველა დაფარულ ზონაში. უფრო მეტიც, თითოეული MVNO-ს მიერ შეგროვებული მონაცემების რაოდენობა გარკვეულ რეგიონებში (მაგალითად, სოფლად) საკმაოდ შეზღუდულია. ვინაიდან თითოეული MVNO-სთვის სასარგებლოა გატარების ზოლის განაწილების მოდელის გაუმჯობესება, FL-მა შექმნა შესაძლებლობა, მრავალი MVNO-სთვის გამოიყენონ მონაცემები მომხმარებელთა უფრო ფართო ერთობლიობიდან და თავიდან აიცილონ მისი გაზიარება. ამ შემთხვევაში, თითოეული MVNO ატრენინგებს გატარების ზოლის განაწილების მოდელს მომხმარებელთა ურთიერთქმედების გამოყენებით. შემდეგ MVNO ატვირთავს თავის ადგილობრივად დატრენინგებულ მოდელს მიმდინარე რაუნდისთვის არარეალურ დროში (non-RT) RIC-ში. non-RT RIC ახორციელებს მოდელის აგრეგაციას შეწონილი ჯამის გამოყენებით, თითოეული MVNO-ს მომხმარებელთა რაოდენობის მიხედვით.

FDR-ზე დაფუძნებული მექანიზმი შედგება რამდენიმე საკომუნიკაციო რაუნდისგან. თითოეული საკომუნიკაციო რაუნდი შედგება ტრენინგის რამდენიმე ლოკალური ეპიზოდისგან, რის შემდეგაც MVNO განახლებს ადგილობრივ მოდელს და აგზავნის განახლებებს non-RT RIC-ში. ეპიზოდის ყველა ეტაპზე, თითოეული MVNO აკვირდება ადგილობრივ გარემოს, ირჩევს მოქმედებას და იღებს შესაბამის ჯილდოს. ყოველი გადასვლა ორ მდგომარეობას შორის (ანუ გამოცდილება) ინახება გამეორების ბუფერში. როდესაც გამოცდილების წინასწარ განსაზღვრული რაოდენობა შენახულია, MVNO-ები იღებენ შემთხვევით მინი პაკეტებს გამეორების ბუფერიდან და განახლებენ სხვადასხვა DDPG ქსელს. აქტორის ქსელი განახლებულია პოლიტიკის გრადიენტით, ხოლო კრიტიკოსის ქსელი განახლებულია დანაკარგების ფუნქციის მინიმიზაციის გზით. შემდგომში განახლება სამიზნე ქსელებიც. ლოკალური ეპიზოდების წინასწარ განსაზღვრული რაოდენობის დასასრულს, თითოეული MVNO აგზავნის ადგილობრივ განახლებულ მოდელს non-RT RIC-ში აგრეგაციისთვის. ეს უკანასკნელი აგროვებს ყველა ლოკალურ განახლებას MVNO-ებიდან და ქმნის გლობალურ მოდელს შეწონილი ჯამის გამოყენებით.

O-RAN ალიანსის მეორე სამუშაო ჯგუფმა განსაზღვრა სამი განლაგების სცენარი ML მოდელის სწავლებისა და ML მოდელის დასკვნისთვის. ქვემოთ ჩვენ ვირჩევთ სცენარს, როდესაც ოფლაინ (offline) ტრენინგის პროცესს მასპინძლობს non-RT RIC და ონლაინ (online) დასკვნის პროცესს მასპინძლობს უახლოესი (ახლომდებარე) RT RIC. ეს არჩევანი დაკავშირებულია შემოთავაზებულ FDRL-ზე დაფუძნებულ RAN-ის ფენებად დაყოფის მექანიზმთან. მართლაც, ჩვენ გამოვიყენეთ DDPG ალგორითმი, რომელიც მოითხოვს უამრავ მონაცემს RAN დაყოფის მოდელების ოფლაინ ტრენინგის შესასრულებლად.

ამ ტიპის ოფლაინ ტრენინგი მოითხოვს რესურსებით სავსე ბლოკს მაღალი გამოთვლითი და შენახვის შესაძლებლობებით. ამიტომ, non-RT RIC, როგორც ჩანს, უფრო შესაფერისია RAN-ის დაყოფის მოდელების ოფლაინ ტრენინგის მასპინძლობისთვის. უახლოესი RT RIC არჩეულ იქნა მოდელის დასკვნის მასპინძლად, რადგან:

1. მომხმარებლისთვის რადიორესურსების განაწილების ოპერაცია უნდა განხორციელდეს მოკლე დროში.
2. RAN-ის ფენებად დაყოფის ოპერაციის შესრულების დროს საჭირო მონაცემები ხელმისაწვდომია E2 ინტერფეისის მეშვეობით, რომელიც აკავშირებს უახლოეს RT RIC-ს RAN-ის კვანძებთან.

7.4. მასპინძლების რისხვითი შედეგები

ჩვენ განვიხილავთ RIC-ით მხარდაჭერილ RAN არქიტექტურას ერთი BS-ით. სიმულირებული მომხმარებლები შემთხვევითი წესით არიან მიმოფანტული 500 მ × 500 მ ფართობზე BS-ის გარშემო და მათ ემსახურება სამი MVNO.

MVNO-ები ერთობლივად ატრენინგებენ DDPG მოდელს. მოდელის ოთხ ქსელს აქვს ორი ფარული სრულად დაკავშირებული ფენა, შესაბამისად 400 და 300 ნეირონით. მიგვაჩნია, რომ თითოეულ MVNO-ს ჰყავს მომხმარებელთა მაქსიმალური რაოდენობა 5-ის ტოლია. ვინაიდან მომხმარებლების მაქსიმალური რაოდენობაა 5, შესასვლელი ფენის ზომაა 10, ხოლო გამოსასვლელი ფენის ზომაა 5. ჩვენ გამოვიყენეთ ReLU, როგორც აქტივაციის ფუნქცია, რადგან ის თავიდან აიცილებს გრადიენტების გაქრობას უკუმიმართულებით გავრცელებისას, მით უმეტეს, რომ სამოქმედო სივრცე შეზღუდულია მცირე მნიშვნელობებით. ეს განპირობებულია ერთ-ერთი პირობით, რომელიც გვეუბნება, რომ გამოყოფილი წილი უნდა იყოს 0.3-ზე ნაკლები, რათა თავიდან იქნეს აცილებული კატარების ზოლის გაფლანგვა მომხმარებელთა მცირე რაოდენობაზე. ჩვენ გამოვიყენეთ ადამ (Adam) ოპტიმიზატორი ორი განსხვავებული სწავლების სიჩქარით აქტორისა და კრიტიკოსისთვის.

შემოთავაზებული არქიტექტურა რეალიზებული იქნა პროგრამირების ენა Python-ზე. DDPG ალგორითმის განსახორციელებლად გამოყენებული იქნა PyTorch პლატფორმა. კომპიუტერული სიმულაციები ჩატარდა 2.6 გიგაჰერციან ლეპტოპზე Intel i7 პროცესორით, 16 გიგაბაიტი ოპერატიული მეხსიერებით და NVIDIA GeForce RTX 2070 გრაფიკული ბარათით.

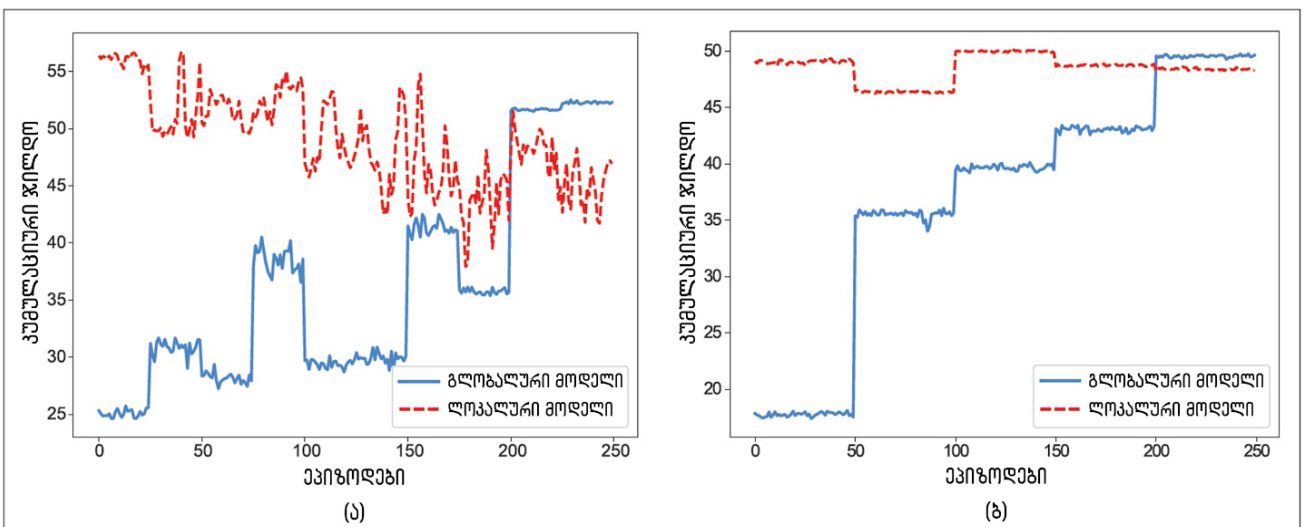
FDRL მექანიზმისთვის ტრენინგი ტარდება სულ ხუთ საკომუნიკაციო რაუნდში. თითოეულ რაუნდში მოდელის დატრენინგება ხდება თითოეული MVNO-ს მიერ 500 ეპიზოდისთვის, სანამ მოდელი გაიგზავნება RIC-ში აგრეგაციისთვის. თითოეული ეპიზოდი შედგება 50 ნაბიჯისგან, სადაც არხში გაძლიერების მნიშვნელობები გადაიტვირთება ყოველ ნაბიჯზე, ხოლო მომხმარებლების მდებარეობების გადატვირთვა ხდება ყოველ 25 ეპიზოდში. მომხმარებლის მოთხოვნების შესაბამისად non-i.i.d განაწილებების გენერირებისთვის, ვაყენებთ URLLC და eMBB მომხმარებლების განსხვავებული ალბათობას თითოეული MVNO-სთვის. URLLC მომხმარებელთა ალბათობების ნაკრები არის 25, 50 და 75 პროცენტი პირველი, მეორე და მესამე MVNO-სთვის, შესაბამისად. ჩვენი შემოთავაზებული გადაწყვეტის შემდგომი შემოწმების მიზნით, ჩვენ შევქმენით მომხმარებლების არათანაბარი განაწილება. კონკრეტულად, ჩვენ განვიხილეთ შემთხვევა, როდესაც MVNO-ებს ნომრებით 1, 2 და 3 ჰყავს 5, 4 და 3 მომხმარებელი, შესაბამისად. ამ შემთხვევაში, თითოეულ MVNO-ზე განაწილებული სიხშირული ზოლის კატარების წილი პროპორციულია მისი მომხმარებელთა რაოდენობისა.

ქვემოთ ვსწავლობთ ორ სცენარს: non-i.i.d-ის მომხმარებელთა თანაბარი რაოდენობით და non-i.i.d-ის მომხმარებელთა არათანაბარი რაოდენობით. FDRL-ის მუშაობის შესაფასებლად, ჩვენ შევადარებთ გლობალურ მოდელს (ანუ FDRL-ით დატრენინგებულ ჩვენს მოდელს) ლოკალურ მოდელებთან, რომლებიც დატრენინგებული იყო თითოეული MVNO-ს მიერ თანამშრომლობის გარეშე. ჩვენ ვერ ვაჩვენებთ ტრენინგის ჯილდოს ევოლუციას; შემდეგ ჩვენ ვართულებთ ტესტირებას მიღებული მოდელებისთვის განაწილების სხვადასხვა წანაცვლების სცენარში, რათა შევადგასოთ მათი რობასტულობა.

პირველი განხილული სცენარი არის non-i.i.d მომხმარებელთა თანაბარი რაოდენობით. მომხმარებელთა საერთო რაოდენობაა 15, თითოეული MVNO ემსახურება 5 მომხმარებელს. ნახ. 7.2ა გვიჩვენებს ლოკალური მოდელების საშუალო ჯილდოსა და გლობალური მოდელის საშუალო ჯილდოს ევოლუციას ხუთ ექსპერიმენტში. მიუხედავად იმისა, რომ გლობალური მოდელი უმჯობესდება გაზიარებული გამოცდილებით, თანაც ლოკალური მოდელების საშუალო მაჩვენებელსაც კი აჭარბებს შემდგომ რაუნდებში, ლოკალურ მოდელებს აქვთ გაუარესებული მახასიათებელი ტრენინგის განმავლობაში, სავარაუდოდ გადაჭარბებული ოვერფიტინგის (overfitting) გამო (განვმარტავთ, რომ ოვერფიტინგი არის ML-ის არასასურველი ქცევა, რომელიც ხდება მაშინ, როდესაც ML-ის მოდელი იძლევა ზუსტ პროგნოზებს ტრენინგის მონაცემებისთვის, მაგრამ არა ახალი მონაცემებისთვის). ფაქტობრივად, ვინაიდან შემდგომ რაუნდებში ორნმტინ-უპლენბეკის ხმაურით გამოწვეული კვლევა მცირდება, ადგილობრივი მოდელები მომხმარებლებს უნაწილებენ ნაკლებ გატარების ზოლს, რაც ამცირებს ჯილდოს მნიშვნელობებს. ამის საპირისპიროდ, გლობალური მოდელი უფრო ნელა სწავლობს განზოგადებას, მაგრამ მთლიანობაში უფრო რობასტულ ტრენინგს აღწევს საერთო გამოცდილების გამოყენებით.

მეორე განხილული სცენარი არის non-i.i.d მომხმარებელთა არათანაბარი რაოდენობით. მომხმარებელთა საერთო რაოდენობაა 12, საიდანაც 5, 4 და 3 მომხმარებელს ემსახურება, შესაბამისად, პირველი, მეორე და მესამე MVNO. სურათი 7.2ბ გვიჩვენებს ლოკალური მოდელების საშუალო ჯილდოს და გლობალური მოდელის საშუალო ჯილდოს ევოლუციას ხუთ ექსპერიმენტში. ჩვენი პირველი დაკვირვება არის ის, რომ ორივე მოდელისთვის კუმულაციური ჯილდო ნაკლებია, ვიდრე ეს მიიღწევა მომხმარებელთა თანაბარი რაოდენობის შემთხვევაში. აღნიშნული ეფექტი ძირითადად გამოწვეულია ჯარიმით, რომელიც დაკავშირებულია არარსებული მომხმარებლებისთვის გატარების ზოლის განაწილებასთან. გარდა ამისა, წინა ექსპერიმენტების მსგავსად, გლობალური მოდელი ნელ-ნელა უმჯობესდება საკომუნიკაციო რაუნდების განმავლობაში, ხოლო ლოკალური მოდელები არ უმჯობესდებიან.

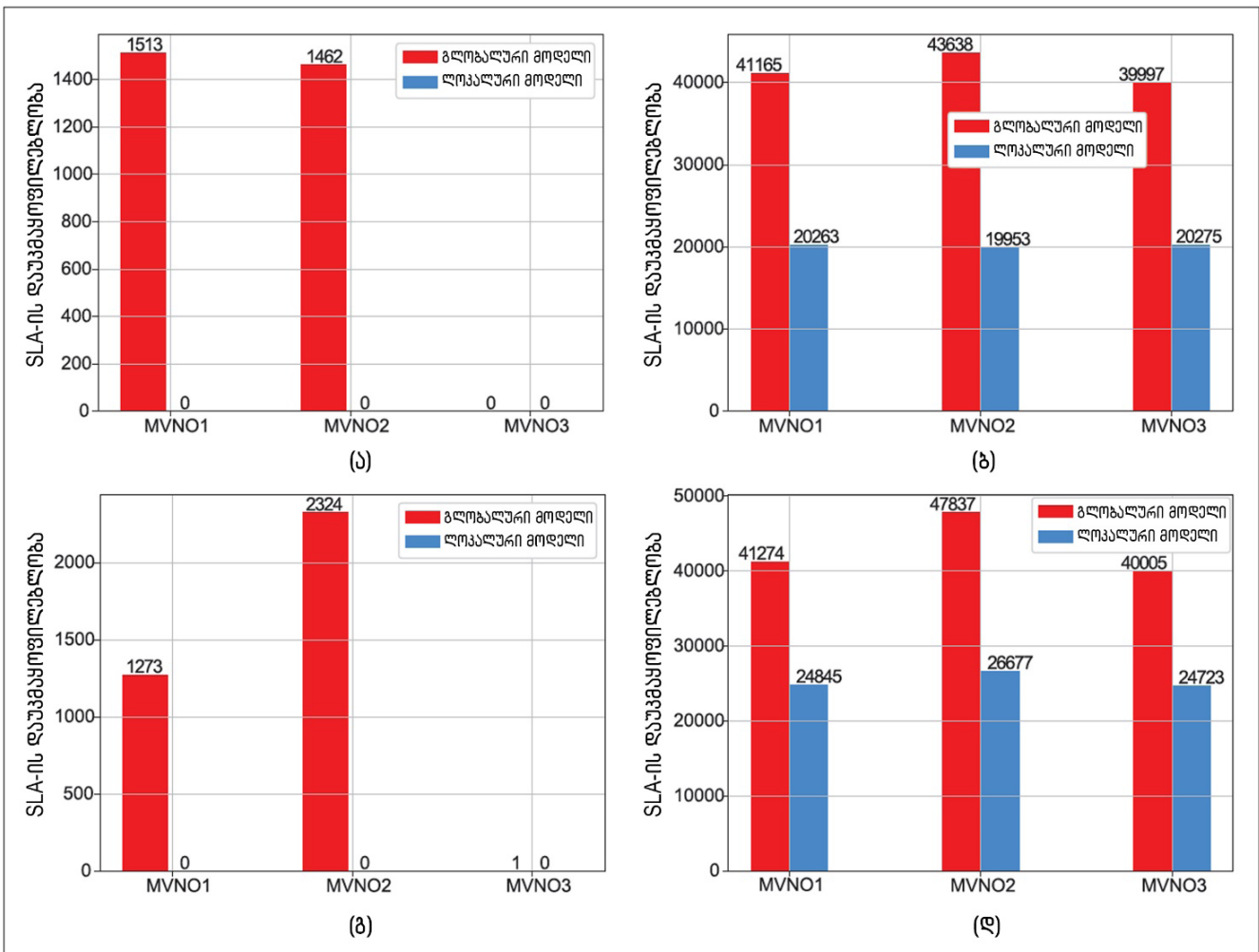
შემოთავაზებული FDRL მექანიზმის მუშაობის შესაფასებლად, ჩვენ შევადარეთ გლობალური მოდელის არასწორი მოქმედებების რაოდენობა თითოეულ ლოკალურ მოდელთან. განვმარტავთ, რომ ქმედება ითვლება არასწორად, თუ ის არ აკმაყოფილებს მომხმარებლის SLA მოთხოვნებს. ჩვენ ვიყენებთ მიღებულ ლოკალურ და გლობალურ მოდელებს და ვამოწმებთ მათ სხვადასხვა გარემოში თითოეული MVNO-ს ძირითადი მომხმარებლის ტიპების განაწილების ცვალებადობით და შემდეგ თითოეული MVNO-ს მიერ მომსახურება გაწეული მომხმარებლების რაოდენობის ცვალებადობით. შემოთავაზებული FDRL მექანიზმის მუშაობის შესაფასებლად, ჩვენ შევადარეთ გლობალური მოდელის არასწორი მოქმედებების რაოდენობა თითოეულ ლოკალური მოდელის ანალოგიურ სიდიდეებს. ჩავთვალეთ, რომ ქმედება ითვლება არასწორად, თუ ის არ აკმაყოფილებს მომხმარებლის SLA მოთხოვნებს. ჩვენ ვიღებთ მიღებულ ლოკალურ და გლობალურ მოდელებს და ვამოწმებთ მათ სხვადასხვა გარემოში, ყოველი MVNO-ს ძირითადი მომხმარებლის ტიპების განაწილების ცვალებადობით და შემდეგ თითოეული MVNO-ს მიერ მომსახურებული მომხმარებლების რაოდენობის ცვალებადობით.



ნახ. 7.2. ტრენინგის მახასიათებელი: ა) non-i.i.d და მომხმარებელთა თანაბარი განაწილება; ბ) non-i.i.d და მომხმარებელთა არათანაბარი განაწილება.

მომხმარებლის ტიპების განაწილების ცვალებადობა:

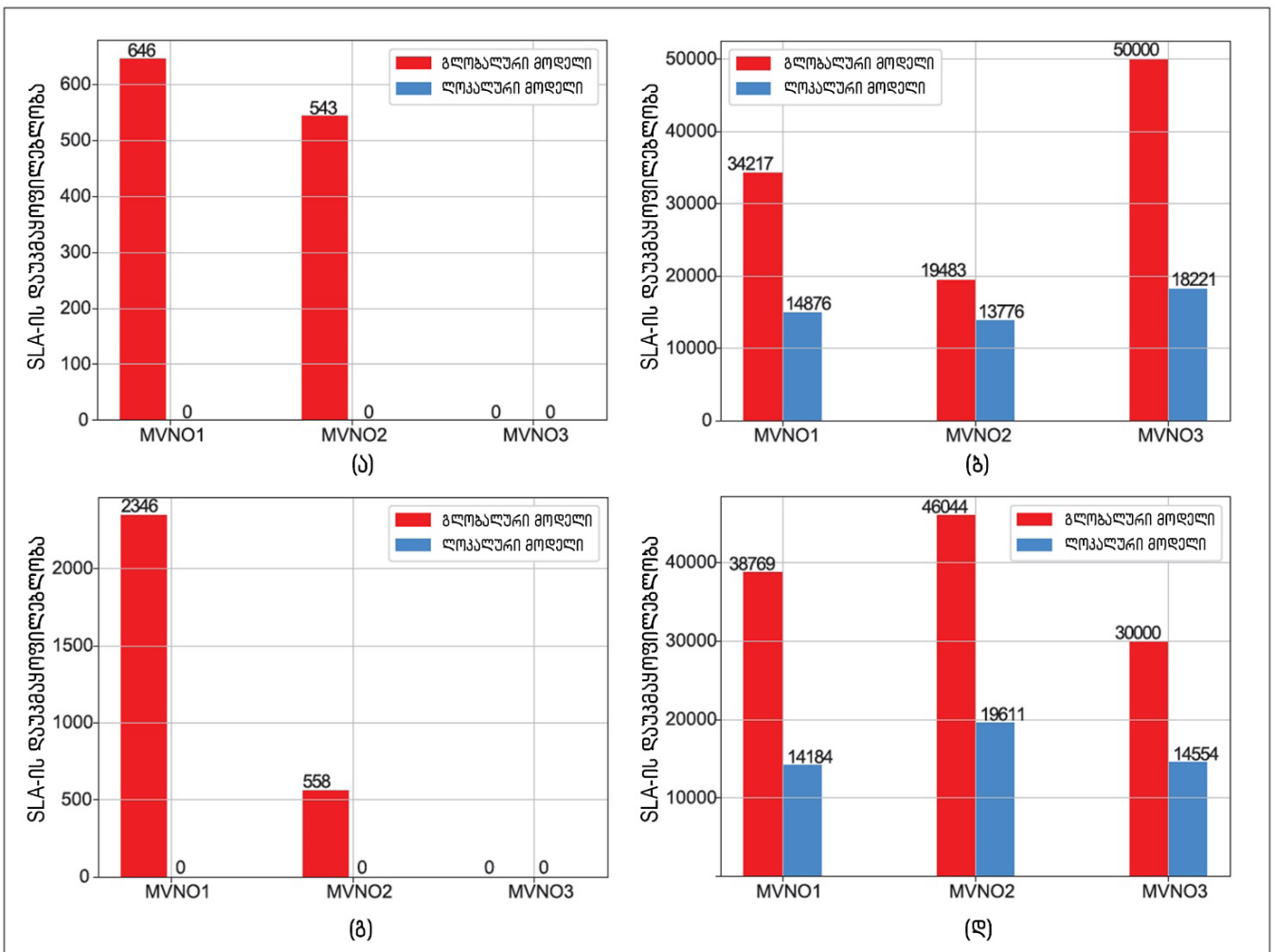
პირველი განხილული სცენარი არის non-i.i.d მომხმარებელთა თანაბარი რაოდენობით. მოდელები დატრენინგებული არიან 15 მომხმარებელით, სადაც თითო MVNO 5 მომხმარებელს ემსახურება. მოდელების რობასტულობის შესაფასებლად, ჩვენ ვცვლიდით მომხმარებლების ძირითად განაწილებებს თითოეული MVNO-სთვის. URLLC მომხმარებლების ალბათობა დატრენინგებულ მოდელებში არის 25, 50 და 75 % პირველი, მეორე და მესამე MVNO-სთვის, შესაბამისად. პირველ ექსპერიმენტში ჩვენ შევცვალეთ URLLC მომხმარებელთა ალბათობები ტესტირების ფაზაში 75, 25 და 50 %-მდე პირველი, მეორე და მესამე MVNO-სთვის, შესაბამისად. მეორე ექსპერიმენტში ჩვენ შევცვალეთ ეს ალბათობები 50, 75 და 25 %-მდე. ნახ. 7.3 გვიჩვენებს კუმულაციურ რაოდენობას, რამდენჯერაც მომხმარებლების SLA მოთხოვნები არ იყო დაკმაყოფილებული MVNO-ების ლოკალური მოდელებით ან გლობალური მოდელით, ერთსა და იმავე გარემოზე დაკვირვებისას (სულ 20000 დაკვირვებისთვის). ჩვენ შევამჩნიეთ, რომ მთლიანობაში, გლობალური მოდელის ქმედებები ნაკლებად მიდრეკილია დაარღვიოს SLA მოთხოვნები eMBB და URLLC მომხმარებლებისთვის, ინდივიდუალურად დატრენინგებულ მოდელებთან შედარებით. გარდა ამისა, რამდენადაც ჩვენ უფრო დიდი წონა მივაკუთვნეთ URLLC მომხმარებლებს, გლობალური მოდელი პრიორიტეტს ანიჭებს ამ ტიპის მომხმარებლებს და ნაკლებად სავარაუდოა, რომ დაარღვიოს მათი საჭირო შეყოვნება.



ნახ. 7.3. შეფასება მომხმარებელთა სხვადასხვა რაოდენობის პირობებში: ა) URLLC (4, 3, 5); ბ) eMBB (4, 3, 5); გ) URLLC (3, 5, 4); დ) eMBB (3, 5, 4).

მომხმარებლების რაოდენობის ცვალებადობა:

მეორე განხილული სცენარი არის non-i.i.d მომხმარებელთა არათანაბარი რაოდენობით. მოდელები პირველად ატრენინგებენ ჯამში 12 მომხმარებელს, სადაც 5, 4 და 3 მომხმარებელს ემსახურება, შესაბამისად, პირველი, მეორე და მესამე MVNO. ჩვენ შევცვადეთ შევცვლასებინა მოდელების რობასტულობა მომხმარებელთა რაოდენობის ცვლილების შემთხვევაში. პირველი, ჩვენ შევცვალეთ მომხმარებლების რაოდენობა ტესტირების დროს 4, 3 და 5-მდე პირველი, მეორე და მესამე MVNO-ებისთვის, შესაბამისად. მეორე ექსპერიმენტში ჩვენ შევცვალეთ ეს რიცხვები 3, 5 და 4-მდე. ნახ. 7.4 გვიჩვენებს რამდენჯერ არ იყო დაკმაყოფილებული მომხმარებლის SLA-ები MVNO-ების ლოკალური და გლობალური მოდელით. დაკვირვება ხორციელდებოდა ერთსა და იმავე გარემოში და ჩატარდა 20000 დაკვირვება. წინა ექსპერიმენტების მსგავსად, გლობალური მოდელის ქმედებები ნაკლებად სავარაუდოა რომ არღვევდეს SLA მოთხოვნებს eMBB და URLLC მომხმარებლებისთვის, ინდივიდუალურად დატრენინგებულ მოდელებთან შედარებით. უფრო მეტიც, მესამე MVNO-ს, რომელიც ძირითადად დატრენინგებულია URLLC მომხმარებლებთან, აქვს მაღალი კმაყოფილების მაჩვენებელი ამ ტიპისთვის, მაგრამ ცუდად მუშაობს eMBB მომხმარებლებისთვის. ზოგადად, QoS-ის გაუმჯობესება ორივე ტიპის მომხმარებლისთვის გლობალური მოდელის გამოყენებით ამართლებს MVNO-ების თანამშრომლობას.



ნახ. 7.4. შეფასება მომხმარებელთა სხვადასხვა რაოდენობის პირობებში: ა) URLLC (4, 3, 5); ბ) eMBB (4, 3, 5); გ) URLLC (3, 5, 4); დ) eMBB (3, 5, 4).

7.5. მეზვიდა თავის დასკვნა და სამომავლო სამუშაოები

ამ თავში, ჩვენ განვიხილეთ FDRL RAN-ის ფენებად დაყოფის მექანიზმი მრავალ MVNO-იან გარემოში, რათა გადანიჭილდეს რადიორესურსები URLLC და eMBB მომხმარებლებისთვის. ჩვენ ვიყენებთ FL-ს, რათა შეიმუშაოს RAN-ის ფენებად დაყოფის მექანიზმი, რომელშიც MVNO-ები თანამშრომლობენ თავიანთი მომხმარებლებისთვის რადიორესურსების განაწილების პოლიტიკის გასაუმჯობესებლად. ჩვენ ვაყალიბებთ რესურსების განაწილების პრობლემას თითოეული MVNO-ს მომხმარებლებისთვის, როგორც ერთი აგენტის MDP-ის და ვიყენებთ DDPG ალგორითმს მის გადასაჭრელად. შემდეგ FL გამოიყენება MVNO-ების რადიორესურსების განაწილების მოდელების გასაუმჯობესებლად. შემოთავაზებული FDRL მექანიზმი რობასტულია URLLC და eMBB მომხმარებლების სხვადასხვა მოთხოვნების დასაკმაყოფილებლად სხვადასხვა სცენარში და უფრო მდგრადია გარემოს ცვლილებების მიმართ.

მიუხედავად იმისა, რომ შემოთავაზებული FDRL-ზე დაფუძნებული RAN-ის ფენებად დაყოფის მექანიზმი იძლევა რადიორესურსების განაწილების მოდელების უფრო მეტ რობასტულობას, სხვადასხვა საინტერესო კვლევითი მიმართულება იმსახურებს ყურადღებას. მიზანშეწონილია შევისწავლოთ, თუ როგორ გამოვაყვლინოთ და დავაკვირდეთ ნებისმიერ დეგრადაციას DRL მოდელების მახასიათებლებში, რათა პროაქტიულად დაიწყოს გადამზადების პროცესი. კიდევ ერთი საინტერესო საკითხი, რომელიც უნდა გამოვიკვლიოთ, არის ის, თუ როგორ გავაძლიეროთ კონფიდენციალურობა იმის მიღმა, რასაც ფედერირებული სწავლება თავისთავად უწყობს ხელს. მართლაც, მიუხედავად იმისა, რომ FL გარკვეულწილად იძლევა მონაცემთა კონფიდენციალურობის გარანტიას, მოწინააღმდეგეებს მაინც შეუძლიათ მიიღონ მონაცემები მოდელის განახლებიდან. ამის შესამცირებლად, კონფიდენციალურობის დაცვის სხვა ტექნიკის ინტეგრირება, როგორცაა, მაგალითად, დიფერენციალური კონფიდენციალურობა, შეიძლება შემდგომში იყოს შესწავლილი MNVO-ების შემთხვევაში, რომლებიც თანამშრომლობენ FDRL სქემებში.

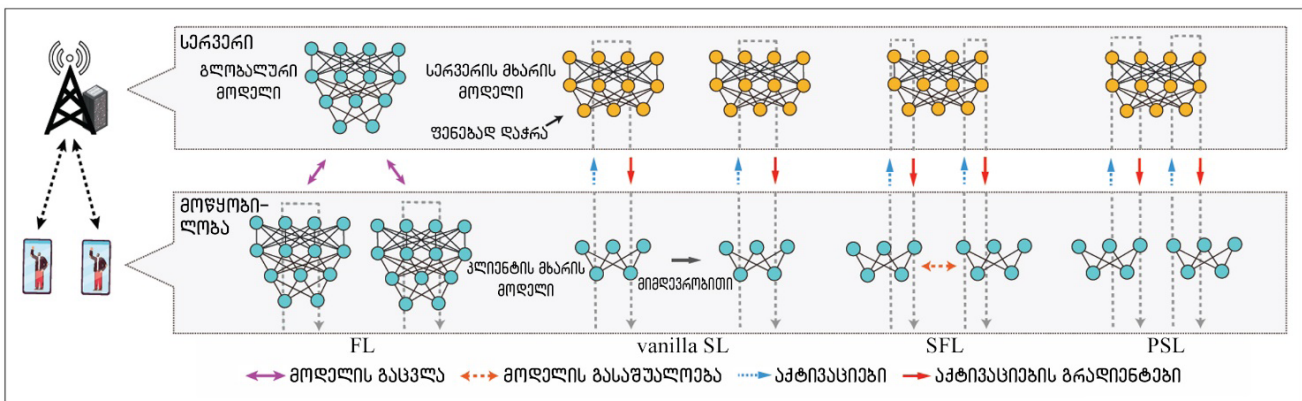
თავი 8 ბაზოფილი სწავლება 5G პერიფერიულ ქსელებში

8.1. შესავალი

დრუბელზე დაფუძნებული მოდელის ტრადიციული ტრენინგი, რომელიც ახდენს ყველა მონაცემის ცენტრალიზებას დამუშავებისთვის, აღარ არის საკმარისი იმისათვის, რომ დააკმაყოფილოს მზარდი მოთხოვნები მონაცემთა ტრაფიკზე, საყოველთაოდ გამოთვლითი საჭიროებები, მკაცრი შეყოვნება და IoT-ის განვითარებადი აპლიკაციებისთვის პერსონალიზაციის მოთხოვნები. ამ გამოწვევების დასაძლევად, პერიფერიული სწავლება (EL) გაჩნდა, როგორც საინტერესო კვლევითი მიმართულება, რომელიც იყენებს მრავალწვდომიანი MEC-ის შესაძლებლობებს ML-ისა და ლოკალური ტრენინგის მხარდასაჭერად, რითაც მიიღწევა ბექჰაულის (backhaul) გატარების ბოლის შემცირებული ხარჯები, ულტრა დაბალი შეყოვნება და კონტექსტური ცნობიერება. მაგალითად, როგორც EL-ის ქვედომენი, ფედერირებული პერიფერიული სწავლება (FEEL) მიიპყრო მნიშვნელოვანი კვლევითი და ინდუსტრიული ინტერესი ბოლო რამდენიმე წლის განმავლობაში მისი კონფიდენციალურობის კუთხით გაუმჯობესებული ბუნების გამო, რაც უკვე განხილულია 3GPP-ის მე-18 გამოშვებაში 5G სტანდარტიზაციისთვის.

თუმცა, FL-ს ასევე აქვს თავისი შეზღუდვები. FL-ის ძირითადი იდეაა გამოიყენოს ლოკალური მოდელის ტრენინგი და გლობალური მოდელის აგრეგაცია თანამშრომლობითი სწავლებისთვის მომხმარებელთა დაუმუშავებელ მონაცემებზე წვდომის გარეშე, როგორც ეს ილუსტრირებულია ნახ. 8.1-ზე. სამწუხაროდ, FL შეიძლება განუხორციელებელი იყოს რესურსებით შეზღუდული IoT მოწყობილობების დიდი რაოდენობის გამო, ვინაიდან მთელი მოდელი დატრენინგებულია საბოლოო მოწყობილობებზე. ამ დაბრკოლების დასაძლევად, გაცოფილი სწავლება (SL) გამოჩნდა, როგორც მოდელის ტრენინგის პერსპექტიული სქემა. მოდელის გაცოფით და მისი ნაწილის პერიფერიულ სერვერზე განთავსებით, SL საშუალებას აძლევს სერვერს გაუმკლავდეს DNN-ის ძირითად დატვირთვას მოდელის გაცოფის საფუძველზე, და ამავე დროს შეინარჩუნოს რამდენიმე ადრეული ფენა და დაუმუშავებელი მონაცემები ლოკალურ მოწყობილობებზე კონფიდენციალურობის უზრუნველყოფისთვის. ეს მიდგომა მნიშვნელოვნად ამცირებს გამოთვლების, შენახვისა და მეხსიერების მოთხოვნებს მოდელის ტრენინგისთვის, რაც ML-ს უფრო ხელმისაწვდომს ხდის რესურსებით შეზღუდული მოწყობილობებისთვის.

ჩვენ გვჯერა, რომ, როგორც FL-ის ალტერნატივა/დამატება, SL მიდგომა შეიძლება გახდეს ერთ-ერთი გაბატონებული AI ტექნოლოგია 5G-ის პერიფერიაზე. 5G იქნება სენსორების, გამოთვლითი მოწყობილობების და ML-ის ქსელი, რათა მიაღწიოს მაღალ მახასიათებლებს ყველგან, არსებული AI ამოცანების გადასაჭრელად, მონაცემთა მფლობელობისა და კონფიდენციალურობის საკითხების გათვალისწინებით. SL-ის ხელსაყრელი მახასიათებლები სრულყოფილად შეესაბამება 5G ხედვას. ერთი მხრივ, როგორც უკვე ავხსენით, SL საშუალებას იძლევა გადმოიტვირთოს ტრენინგის დატვირთვა მონაცემთა კონფიდენციალურობის გაზრდისას. ამას უდიდესი მნიშვნელობა აქვს ფართოდ გავრცელებული მობილური და IoT მოწყობილობებისთვის შეზღუდული აპარატურული



ნახ. 8.1. FL და თანამედროვე SL სტრუქტურების ილუსტრაცია, სადაც მოდელის გასაშუალებული შეფასება SFL-ში შესრულებულია მიმწოდებელ სერვერზე, რომელიც მდებარეობს კლიენტებთან ახლოს

უბრუნველყოფით, როგორცაა მობილური ტელეფონები და ჭკვიანი კამერები, რომლებსაც შეიძლება გაუჭირდეთ ინტენსიური ლოკალური გამოთვლითი ტრენინგის მხარდაჭერა, როგორც ამას FL მოითხოვს. ასევე რთულია დაუშუშავებელი მონაცემების აღდგენა SL-ში აქტივაციების გაცვლის და სერვერის მიერ კლიენტის მხარის მოდელის შესახებ ცოდნის ნაკლებობის გამო. მეორე მხრივ, SL აღწევს რესურსების უკეთეს გამოყენებას ქსელის პერიფერიაზე განაწილებული გამოთვლითი და მეხსიერების რესურსების გამოყენებით. ამჟამად DNN-ები, როგორც წესი, შეიცავს მილიონობით ან თუნდაც მილიარდობით პარამეტრს, რაც მათ პერიფერიულ ტრენინგს რთულს ხდის. მაგალითად, დიდი ენის მოდელს 7B LLaMA-ს, რომელიც არის LLaMA-ს ყველაზე პატარა ვერსია და რომელიც შესაფერისია მოწყობილობაზე განთავსებისთვის, ჯერ კიდევ აქვს 7 მილიარდი პარამეტრი. 6G-ის ეპოქაში, პერიფერიული გამოთვლითი რესურსები გახდება საყოველთაოდ, სადაც სერვერები შეიძლება განთავსებული იყოს ქსელის ბირთვში, მაკრო BS-ებში, მცირე BS-ებში, პიკო BS-ებში და თვით ავტონომიურ მანქანებსა და მობილურ ტელეფონებშიც კი. მოდელის გაყოფის ჩართვით, SL-ს 6G პერიფერიაზე შეუძლია ხელი შეუწყოს მოქნილი გამოთვლითი დატვირთვის გაზიარებას საბოლოო მოწყობილობებსა და მრავალ სერვერს შორის ერთობლივი ტრენინგისთვის. ეს საშუალებას იძლევა საუკეთესოდ გამოვიყენოთ განაწილებული გამოთვლითი რესურსები, რესურსებით ინტენსიური ტრენინგის შესასრულებლად.

მობილური პერიფერიული ქსელების და SL-ის ინტეგრაცია წარმოადგენს უნიკალურ ტექნიკურ გამოწვევებს და საინტერესო კვლევით შესაძლებლობებს. დღემდე ეს სფერო შედარებით შეუსწავლელია. პერიფერიაზე ეფექტიანი SL-ის უბრუნველსაყოფად, დიზაინის ძირითადი გამოწვევები წარმოიქმნება მაღალგანზომილებიანი მახასიათებლების ხშირი გადაცემის და უკან გავრცელების გრადიენტების აუცილებლობის გამო, გატარების ბოლის მიხედვით შეზღუდულ უსადენო არხებში, და ღრმა მოდელის ტრენინგის აუცილებლობის გამო რესურსებით შეზღუდულ პერიფერიულ ქსელებში. ამ პრობლემების სწორად მოგვარება შესაძლებელია მხოლოდ SL-ის და საკომუნიკაციო-გამოთვლითი რესურსების მართვის ჰოლისტიკური დიზაინით, რაც ძალიან რთულია. ამ თავის ძირითადი ამოცანაა განისაზღვროს, თუ როგორ შეიძლება ეფექტიანად მოხდეს გაყოფილი პერიფერიული სწავლების (SEL) მხარდაჭერა რესურსებით შეზღუდულ უსადენო ქსელებში.

მერვე თავი მიზნად ისახავს საფუძვლიანად შეისწავლოს SL-ის განთავსება მობილურ პერიფერიულ ქსელებში. ამ მიზნით, თავდაპირველად წარმოვადგენთ SL-ისთვის მორგებულ 6G არქიტექტურულ დიზაინს. შემდეგ, განვსაზღვრავთ პოტენციურ მიმართულებებს ეფექტიანი SL დიზაინისთვის, რომელიც იყენებს მოდელისა და აქტივაციის შეკუმშვას და უკან გავრცელებას SL-ის რესურსის მოხმარების შესამცირებლად. დაბოლოს, ჩვენ განვიხილავთ ინოვაციურ გადაწყვეტილებებს SEL-თან დაკავშირებული რესურსების მართვის პრობლემებისთვის (როგორც ერთპერიფერიული, ისე მრავალპერიფერიული სცენარების მიხედვით), როგორცაა რესურსების დინამიკური განაწილება უმოქმედო რესურსების გამოყენების ოპტიმიზაციისთვის, იერარქიულ ან მრავალნახტომიან (multi-hop) SL-ს დიდი მოდელის ერთობლივი ტრენინგისთვის და მოდელის განთავსების/მიგრაციის სტრატეგიებს მომხმარებელთა განაწილებისა და მობილურობის დასაკმაყოფილებლად. ამგვარად, ეს თავი გვთავაზობს SEL-ის ყოვლისმომცველ მიმოხილვას და ხაზს უსვამს კვლევით შესაძლებლობებს. მნიშვნელოვანია აღინიშნოს, რომ მოდელის გაყოფის მსგავსების მიუხედავად, SEL-ის პრობლემები მნიშვნელოვნად განსხვავდება პერიფერიების გაყოფის დასკვნისგან (inference), რომელიც აღწერილია ზოგიერთ არსებულ ნაშრომში, რადგან SEL-ის მიზანია მოდელის ტრენინგი და დახვეწა-დაყენება და არა პროგნოზირება პერიფერიაზე.

ეს თავი ორგანიზებულია შემდეგნაირად. ვიწყებთ არსებული SL მიდგომების შესწავლით. შემდეგ ჩვენი დისკუსია შემდეგ გადავა სინერგიაზე SL-სა და 6G პერიფერიას შორის. კერძოდ, ჩვენ განვიხილავთ 6G არქიტექტურულ დიზაინს SL-ისთვის, რასაც მოჰყვება სიღრმისეული დისკუსია ინოვაციური რესურსების მიხედვით ეფექტიან SL-ის სტრუქტურის დიზაინზე. შემდეგ წარმოვადგენთ რესურსების მართვის სტრატეგიებს SL-ისთვის ერთფიჭიანი და მრავალფენიანი სცენარებით. საბოლოოდ, განვსაზღვრავთ ღია პრობლემებს SEL-ისთვის და ამ თავის დასკვნებს.

8.2. FL-ის და SL-ის შედარება

სტანდარტული ML მიდგომები, როგორც წესი იყენებენ ოპტიმიზაციის ალგორითმებს, როგორცაა SGD დისტანციურ ღრუბლოვანი ცენტრში ტრენინგის ცენტრალიზებული მონაცემებით, რაც იწვევს კონფიდენციალურობის სერიოზულ გაფონვას. ამ პრობლემის შესამსუბუქებლად, 2016 წელს Google-მა შემოგვთავაზა FL, რაც საშუალებას აძლევს მობილურ ტელეფონებს ერთობლივად ისწავლონ გაზიარებული მოდელი, ხოლო ტრენინგის ყველა მონაცემი ლოკალურად შეინახონ. კონკრეტულად, მოწყობილობები მხოლოდ ატრენინგებენ და ატვირთავენ მოდელის განახლებებს სერვერზე აგრეგაციისთვის და ჩამოტვირთავენ აგრეგირებულ მოდელს სერვერიდან, როგორც ნაჩვენებია ნახ. 8.1-ზე. მოდელის უზარმაზარი ზომისა და მოდელის განმეორებითი ატვირთვა/ჩამოტვირთვის გამო, FL იწვევს კომუნიკაციის მნიშვნელოვან შეყვანებას და დიდ ტვირთად აწვება სატელეკომუნიკაციო ინფრასტრუქტურას. ამ მიზეზით, FL-მა გამოიწვია მნიშვნელოვანი ინტერესი სატელეკომუნიკაციო ინდუსტრიის მხრიდან, რომელიც მიზნად ისახავს განახორციელოს კომუნიკაციის კუთხით FL-ის ეფექტიანი განთავსება მობილურ პერიფერიაზე, რაც გამოიწვევს ახალი სფეროს გაჩენას, რომელიც ცნობილია როგორც FEEL.

მოდელის მზარდი ზომისა და რესურსებით შეზღუდული პერიფერიული მოწყობილობების გათვალისწინებით, FL შეიძლება არ იყოს შესაფერისი სხვადასხვა ინტელექტუალური აპლიკაციისთვის, მოწყობილობებზე მოდელის სრული ტრენინგის საჭიროების გამო. დაინერგა რა 2018 წელს, SL გაჩნდა, როგორც კონფიდენციალურობის გამაძლიერებელი თანამშრომლობითი სასწავლო სტრუქტურა, რომელიც ითვალისწინებს რესურსების შეზღუდვებს მონაცემთა კონფიდენციალურობის შენარჩუნებისას. იდეა არის მოდელის უბრალოდ გაყოფა ორ ან მეტ ნაწილად და ამ ნაწილების განთავსება კლიენტისა და სერვერის მხარეს, შესაბამისად, ეს საშუალებას აძლევს სერვერს გაიზიაროს ტრენინგის დატვირთვა.

SL-ის რამდენიმე ვარიანტი არსებობს. vanilla SL, რომელიც არის ორიგინალური ფორმა, მუშაობს თანმიმდევრულად, რომელიც ატრენინგებს მოდელს დროის კონკრეტულ მონაკვეთში ერთი კლიენტისთვის. თუმცა, vanilla SL-ის თანმიმდევრული ტრენინგის პროცესი იწვევს ტრენინგის გადაჭარბებულ შეყვანებას. უფრო მეტიც, non-i.i.d პარამეტრების შემთხვევაში, თანმიმდევრულმა ტრენინგმა შეიძლება გამოიწვიოს ტრენინგის ცუდი მახასიათებელი, რადგან მოდელს აქვს ტენდენცია, კარგად მოერგოს მხოლოდ ბოლო კლიენტის მონაცემების განაწილებას.

ამ საკითხების გადასაჭრელად, გაყოფილი ფედერირებული სწავლება (SFL) და პარალელური გაყოფილი სწავლება (PSL) შემუშავდა კლიენტის მხარის მოდელის ტრენინგის დაპარალელუებისათვის, რაც კლიენტებს აძლევს უფლებას, თავიანთი ქვემოდელი ერთდროულად მოამზადონ, როგორც ეს ილუსტრირებულია ნახ. 8.1-ზე. განსხვავება SFL-სა და PSL-ს შორის მდგომარეობს სინქრონიზაციის მოთხოვნებში: FL-ის სულისკვეთების შესაბამისად, პირველი მოითხოვს კლიენტის მხარის მოდელების პერიოდულ გასაშუალოებას, ხოლო მეორე არ საჭიროებს ამას. შედეგად, SFL იწვევს კომუნიკაციის ხარჯების გაზრდას კლიენტის მხარის მოდელის გადაცემის გამო.

უფრო მეტიც, SFL მოითხოვს ცალკეულ და შეთანხმებულ სერვერებს, შესაბამისად, სერვერის მხარის მოდელის ტრენინგისთვის და კლიენტის მხარის მოდელის გასაშუალოებისთვის. წინააღმდეგ შემთხვევაში, სერვერის მხარის მოდელის გამოსასვლელით და კლიენტის მხარის მოდელის პარამეტრებით, მანე (მოწინააღმდეგე) სერვერს შეუძლია ადვილად აღადგინოს შეყვანილი დაუმუშავებელი მონაცემები. პირიქით, PSL გამოირიცხავს კლიენტის მხრიდან მოდელის სინქრონიზაციის აუცილებლობას, რაც გადალახავს ზემოთ მოცემულ შეზღუდვებს. თუმცა, ის ბუნებრივად იწვევს კლიენტის მხარის მოდელის განსხვავებულ პარამეტრებს მოწყობილობებში, რამაც შეიძლება უარყოფითად იმოქმედოს ტრენინგის კონვერგენციაზე SFL-თან შედარებით, როგორც ნაჩვენებია ლიტერატურაში.

ბუნებრივი კითხვაა, როგორ გავაკეთოთ არჩევანი SL-სა და FL-ს შორის, როდესაც ორივე მათგანი კონფიდენციალურობის გამაძლიერებელი თანამშრომლობითი სასწავლო სტრუქტურებია? გადაწყვეტი ფაქტორი არის საბოლოო მოწყობილობების გამოთვლითი შესაძლებლობები. როგორც უკვე აღვნიშნეთ, SL არის ბუნებრივი გადაწყვეტა, როდესაც საბოლოო მოწყობილობა შეზღუდულია რესურსებით და ნაკლებად სავარაუდოა, რომ შეძლებს დიდი მოდელის დატრენინგებას გამოთვლითი და მეხსიერების შეზღუდვების გამო.

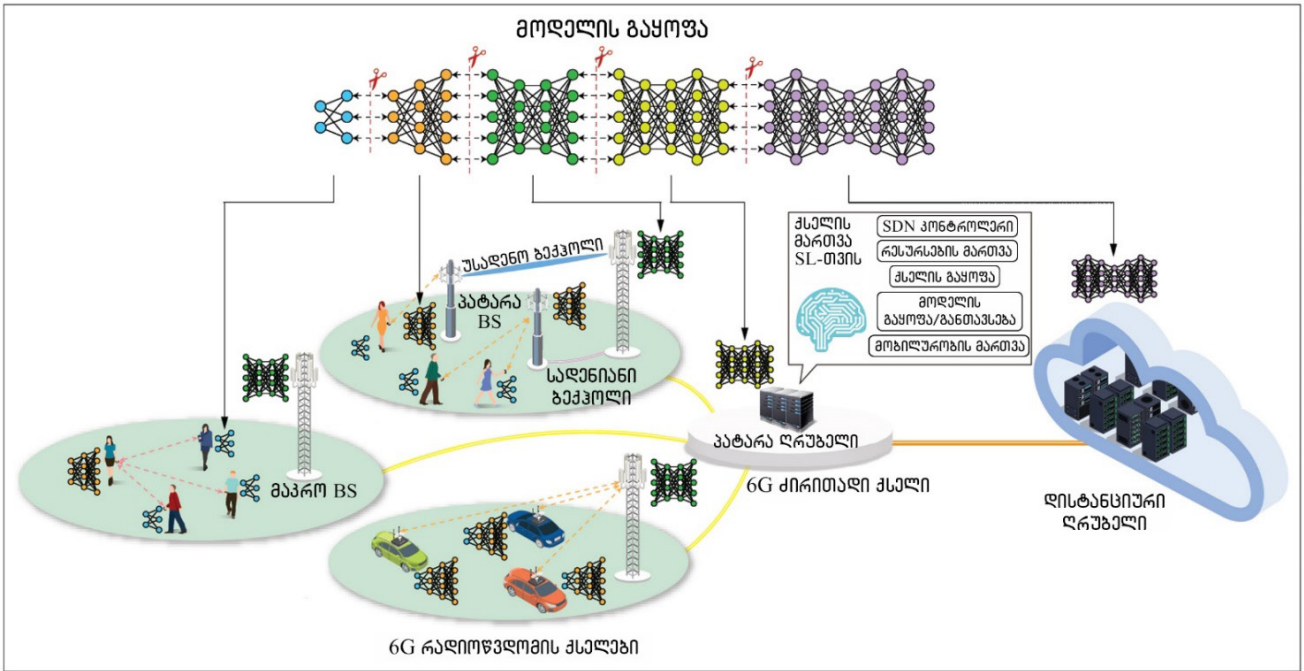
კომუნიკაციის პერსპექტივიდან, როდესაც კლიენტების ტრენინგის მონაცემთა ნაკრები დიდია, FL შეიძლება იყოს სასურველი, რადგან SL მოითხოვს მონაცემთა დიდ მოცულობებს, რომლებიც პროპორციულია მონაცემთა ნაკრების ზომისა. ამის საპირისპიროდ, ვინაიდან FL მოითხოვს მოდელის გადაცემას, SL ხდება უფრო ეფექტიანი კომუნიკაციის თვალსაზრისით, როდესაც მოდელის მონაცემთა ზომა უფრო დიდია, ვიდრე სმეშირებული მონაცემების ზომა (მაგალითად. ResNet-152-ისთვის, როდესაც პაკეტის ზომაა 32, 37-ე ფენის სმეშირებული მონაცემთა მოცულობა დაახლოებით 0.49 მბაიტია, თუმცა მოდელის ზომაა 230 მეგაბაიტი). (განვმარტავთ, მონაცემთა სმეშირება (Data smashing) არის ახალი პრინციპი მონაცემთა ნებისმიერი ნაკადების წყაროებს შორის მსგავსების შესაფასებლად, არც დომენის ცოდნის და არც სწავლების გამოყენებით. ეს არის პროცესი, რომლის დროსაც ნაკადები და ანტი-ნაკადები შეიძლება ალგორითმულად შეეჯახონ ერთმანეთს, რათა გამოავლინონ განსხვავებები, რომელთა აღმოჩენა ძნელია ტრადიციული მეთოდების გამოყენებით). მიგვაჩნია, რომ SL შეიძლება იყოს FL-ის მნიშვნელოვანი ალტერნატივა რესურსებით შეზღუდულ პერიფერიულ ქსელებში.

8.3. არჩიტექტურული დიზაინი გაყოფილი პერიფერიული სწავლებისთვის

SEL მოითხოვს კომუნიკაციებისა და ტრენინგის ჰოლისტიკური დიზაინის შექმნას, რადგან არსებობს ფუნდამენტური კომპრომისი სასწავლო ქვემოდელის გამოთვლის ღირებულებასა და კოლაბორაციულ მოწყობილობებს შორის სმეშირებული მონაცემების (ანუ შუალედური აქტივაციები/უკან გავრცელებული გრადიენტები) გადაცემის საფასურს შორის. კარგად შემუშავებული დამხმარე არჩიტექტურა აუცილებელია რესურსებით შეზღუდული უსადენო ქსელის პირობებში ტრენინგის კონვერგენციის ოპტიმიზაციისთვის. აღვნიშნავთ, რომ 6G იძლევა კომუნიკაციებისა და გამოთვლების ნამდვილ კონვერგენციას, რაც უზრუნველყოფს MEC-ისა და SL-ის შემდგომი ინტეგრაციის შესაძლებლობებს. ამის გათვალისწინებით, ჩვენ წარმოვადგენთ 6G ქსელის პოტენციურ არჩიტექტურას, რომელიც მორგებულია SL-ზე.

რესურსების შეზღუდვის გამო, ერთმა გამომთვლელმა მოწყობილობამ/სერვერმა შეიძლება ვერ შეძლოს დიდი AI მოდელის დატრენინგება/განთავსება. ამიტომ SL სერვისების ეფექტიანი უზრუნველყოფის გასააქტიურებლად, საჭიროა იერარქიული სისტემა, რომელიც ამუშავებს სერვისებს სხვადასხვა გამოთვლითი დატვირთვით, შეყოვნების შეზღუდვით და პერსონალიზაციის მოთხოვნებით. 6G SL სისტემა მოიცავს მონაცემთა წყაროებს, როგორცაა ჯგვიანი კამერები, მობილური ტელეფონები და ავტონომიური მანქანები, ჰეტეროგენულ ფიჭურ საბაზო სადგურებს, როგორცაა პატარა და მაკრო საბაზო სადგურები, პერიფერიული სერვერები (ღრუბლები), რომლებიც მდებარეობს ფიჭების აგრეგაციის ადგილებში და დისტანციურ ღრუბელს. შედეგად, მრავალმხრივი AI მოდელები, მათ შორის სრული ან ნაწილობრივი მოდელები, ნაწილდება სისტემის მრავალ დონეზე ერთობლივი ტრენინგისთვის. ზოგადად, უფრო დიდი მოდელი (ქვემოდელი) შეიძლება განთავსდეს უფრო მძლავრ კვანძში, რომელიც უფრო შორს არის მონაცემთა წყაროებიდან, ხოლო უფრო მცირე მოდელი (ქვემოდელი) შეიძლება ინახებოდეს რესურსებით შეზღუდულ პერიფერიულ მოწყობილობებზე/სერვერებზე. მაღალი დონის გამოთვლით კვანძებს ასევე შეუძლიათ შეინახონ მოდელები (ქვემოდელები) უფრო ზოგადი წარმოდგენით მომხმარებლების მიერ ფართომასშტაბიანი გამოყენებისთვის, ხოლო ქვედა დონის პერიფერიული კვანძები ინახავენ მოდელებს (ქვემოდელებს), რომლებიც უკეთესად მოერგება ლოკალურ გარემოს. კონფიგურაციიდან გამომდინარე, სმეშირებული მონაცემების გაცვლა შესაძლებელია ამ მოწყობილობებს/სერვერებს შორის, ანუ D2D, ასევე სატრანსპორტო საშუალება — სატრანსპორტო საშუალებასთან (V2V) ლინკების, უსადენო ბექჰოლის ან სადენიანი ბექჰოლის მეშვეობით, როგორც ილუსტრირებულია ნახ. 8.2-ზე.

SL-ში, მოდელის ტრენინგი მრავალ გამოთვლით მოწყობილობაზე/სერვერზე მოითხოვს რესურსების გონივრულ კოორდინაციას. გამჭოლი QoS გარანტიების უზრუნველსაყოფად, სავარაუდოდ, აუცილებელია ცენტრალიზებული მართვა. საბედნიეროდ, ცენტრალიზებული ინტელექტი კარგად არის მორგებული 6G არჩიტექტურასთან. ამ მიზნის მისაღწევად, 6G პერიფერიას შეუძლია დაწეროს SDN, რათა ხელი შეუწყოს მოდელის გადაცემას, სმეშირებული მონაცემების მარშრუტიზაციას და გამოთვლითი რესურსების განაწილებას. ქსელის ლინკის სტატუსისა და გამოთვლითი/შენახვის/მეხსიერების რესურსის ხელმისაწვდომობის მონიტორინგით, ცენტრალური კონტროლერი პროაქტიულად ყოფს მოდელებს, მართავს გამოთვლით და ქსელის რესურსებს, აკონფიგურირებს მონაცემთა მარშრუტიზაციას და ახორციელებს მოდელის განთავსებას/მიგრაციას. გამჭოლი



ნახ. 8.2. SL-ის სრული არქიტექტურა 6G-ში

(ბოლოების დამაკავშირებელი) ქსელის ფენებად დაყოფა (დაჭრა) შეიძლება გამოყენებულ იქნეს სხვადასხვა SL ამოცანისთვის, დიფერენცირებული QoS უზრუნველყოფის მისაღწევად. მაგალითად, SL ამოცანები ავტონომიური მართვისთვის ან რობოტის კონტროლისთვის მოითხოვს ულტრა დაბალ შეყოვნებას, ხოლო SL ამოცანები შემდეგი სიტყვის პროგნოზირების მოდელის ტრენინგისთვის შეიძლება არ იყოს დროისადმი მგრძობიარე. დაბოლოს, 6G პერიფერიას ასევე უნდა ჰქონდეს მობილურობის მართვის კომპონენტები, რათა უზრუნველყოს უწყვეტი სერვისი და მოდელის მიგრაცია, რაც საშუალებას მისცემს კონტექსტზე დაფუძნებულ მოდელებს მიჰყვეს მომხმარებლებს მათი გადაადგილებისას. ამ მიდგომების დანერგვით, 6G სისტემას შეუძლია გამოიყენოს ქსელის მასშტაბით განაწილებული რესურსები და დააკმაყოფილოს სხვადასხვა SL აპლიკაციის QoS მოთხოვნები.

8.4. რესურსების კუთხით ეფექტიანი გაყოფილი სწავლების სტრუქტურები

პერსპექტიული უპირატესობების მიუხედავად, შეზღუდული სპექტრი და გამოთვლითი რესურსები ქსელის პერიფერიაზე მნიშვნელოვან დაბრკოლებებს უქმნის SL-ის ეფექტიან განხორციელებას. ხშირად მიზანშეწონილია ტრენინგის სიზუსტის გაცვლა შემცირებულ შეყოვნებაზე ქსელის და გამოთვლითი რესურსების შეზღუდვების პირობებში. ქვემოთ ჩვენ წარმოვადგენთ ინოვაციურ SL სტრუქტურებს, რომლებიც ამცირებს რესურსების მოთხოვნებს სხვადასხვა ასპექტში.

უპირველეს ყოვლისა, გადამწყვეტი მნიშვნელობა აქვს კომუნიკაციის ზედნადები ხარჯების შემცირებას, რომელიც დაკავშირებულია სემირებულ მონაცემთა გაცვლასთან მოწყობილობებსა და პერიფერიულ სერვერს შორის გაყოფილი ფენის მეშვეობით. ამ პრობლემის შესამსუბუქებლად, ერთ-ერთი პერსპექტიული მიმართულებაა ავტოკოდერის დანერგვა, რომელიც ატრენინგებს კოდერს მონაცემების შეკუმშვისთვის და შემდეგ დეკოდერს მონაცემების აღდგენისთვის. ლიტერატურაში ციკლური კონვოლუცია გამოიყენება მრავალი მახასიათებლის შეკუმშვისთვის ერთ შეკუმშულ ფუნქციაში, რომელიც დეკოდირდება სერვერის მხარეს ციკლური კორელაციის მეშვეობით. მიუხედავად იმისა, რომ ეს პროცესი იწვევს ხმაურს, სწავლების მახასიათებელზე გავლენა მცირეა.

როგორც ნეირონული ქსელები, ისე ავტოკოდერები საჭიროებენ დამატებითი გამოთვლებისა და ტრენინგის

ხარჯებს და რთულია თეორიულად მათი გაგება. ასე რომ, სხვა კვლევითი მიმართულება არის სემპირებული მონაცემების პირდაპირ შეკუმშვა. ამ მიზნით, ფუნქციის შეკუმშვა შესწავლილი იქნა გაყოფილი დასკვნის სახით, გასხვლის (pruning) აქტივაციებით. თუმცა, მისი გავლენა SL-ზე იმსახურებს შემდგომ შესწავლას. ასევე, მნიშვნელოვანია თეორიულად დახასიათდეს SL-ის კონვერგენციის ზღვარი შეკუმშვის კოეფიციენტების თვალსაზრისით, რომლის საფუძველზეც შეიძლება შემუშავდეს SL სქემა საგულდაგულოდ შემუშავებული მახასიათებლების შეკუმშვის კოეფიციენტით, რათა მიაღწიოს ოპტიმალურ ბალანსს ტრენინგის სიზუსტესა და შეყოვნებას შორის.

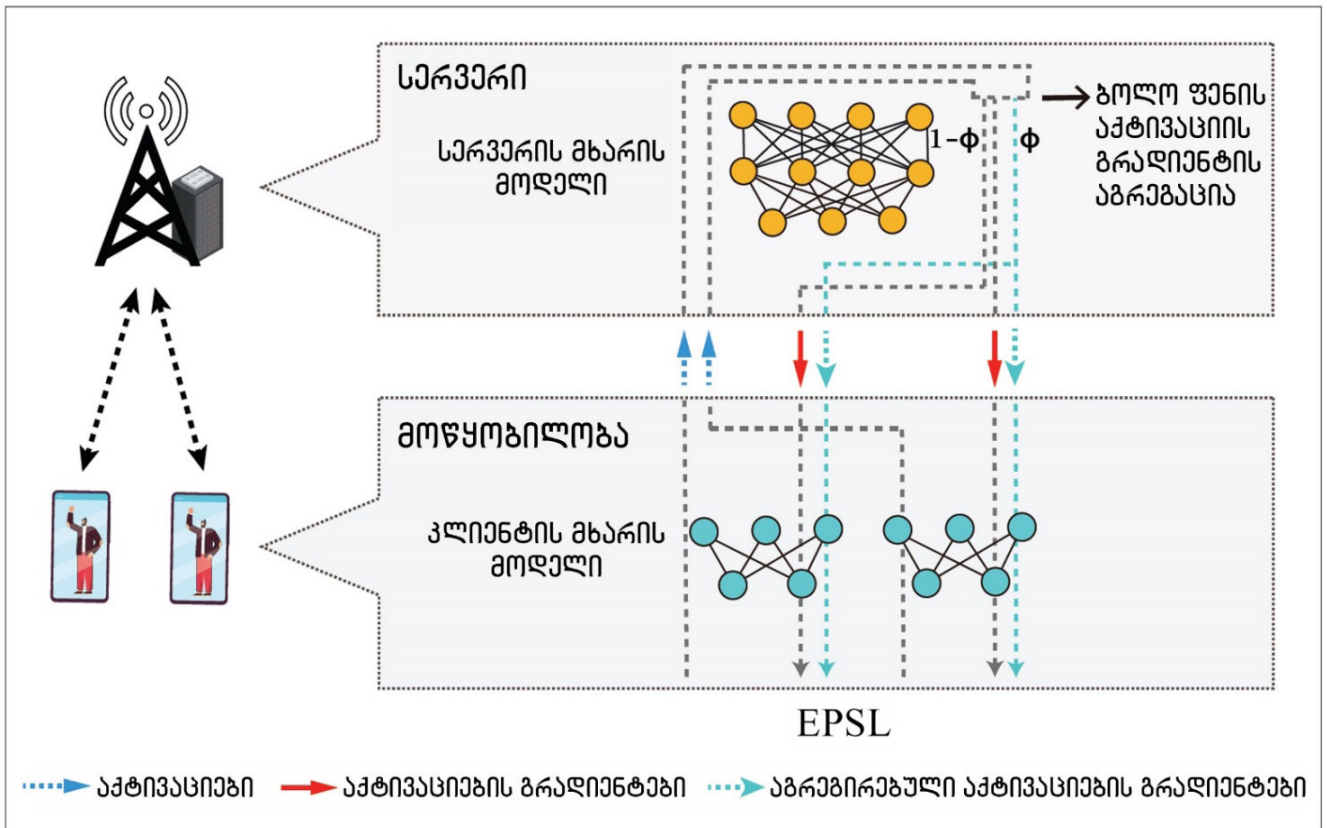
მეორე პრობლემა არის ის, თუ როგორ შევამციროთ გამოთვლითი დატვირთვა, განსაკუთრებით მოწყობილობებზე. მიუხედავად იმისა, რომ სერვერიამუშავეებს SL-მისამუშაო დატვირთვის უმეტეს ნაწილს, დარჩენილი გამოთვლითი დატვირთვა, როგორცაა რამდენიმე ადრეული ფენა, რომელიც აუცილებელია გადაუმუშავებელი მონაცემების დასამალად, შეიძლება მაინც იყოს ძალიან მოთხოვნადი რესურსებით შეზღუდული მობილური/IoT მოწყობილობებისთვის. მიზანშეწონილი გამოსავალი არის ტრენინგი შეკუმშულ მოდელზე, რათა შემცირდეს გამოთვლითი და მესხიერების ხარჯები. არსებობს შეკუმშვის სხვადასხვა პოპულარული მეთოდი:

- მოდელის დაქვანტვა ამცირებს როგორც წონების, ისე აქტივაციების ბიტების სიგრძეს (bit widths), ანუ რაოდენობას (მაგალითად, სრული სიზუსტიდან 8 ბიტამდე), რათა შეამციროს ტრენინგის შეყოვნება და მესხიერების მოთხოვნილება.
- მოდელის გასხვლა პირდაპირ ამცირებს DNN-ის პარამეტრების რაოდენობას.

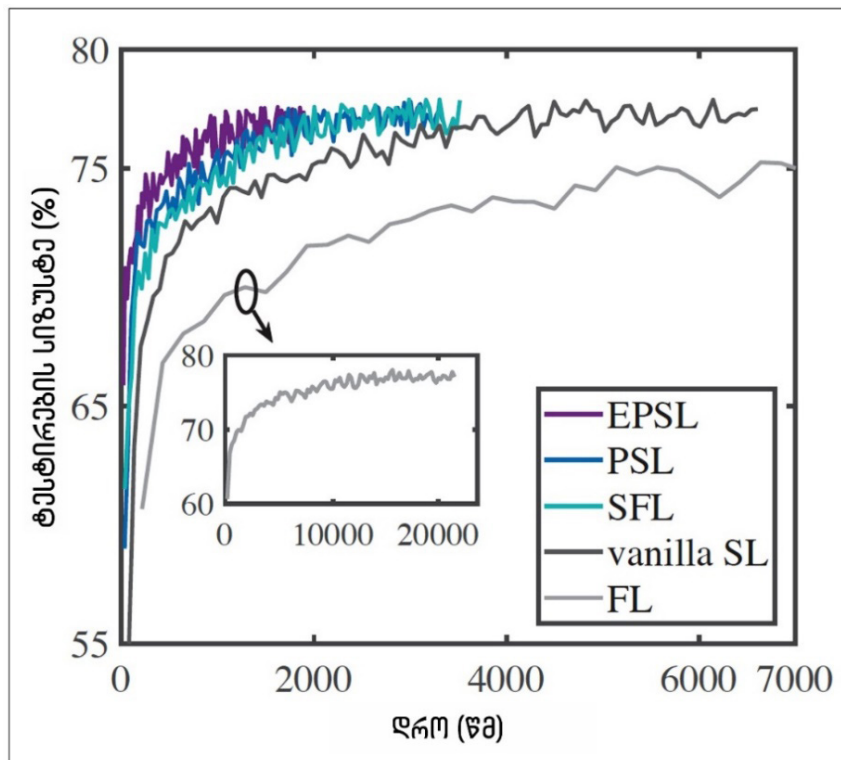
ჩვენი მიზანია შევამციროთ გამოთვლითი ხარჯები უშუალოდ შეკუმშულ წონებზე ტრენინგის გზით, ასევე შევამციროთ კომუნიკაციის ოვერჰედები დაჭრილ ფენაზე აქტივაციების ბიტების სიგრძის შემცირებით. SL-ში მომგებიანია მოდელის შეკუმშვის სქემის შემუშავება, რომელიც საშუალებას აძლევს კლიენტებს და პერიფერიულ სერვერს გამოიყენონ სხვადასხვა შეკუმშვის კოეფიციენტი მათი გამოთვლითი შესაძლებლობების საფუძველზე. მაგალითად, ვინაიდან სერვერი შეიძლება იყოს ბევრად უფრო ძლიერი ვიდრე კლიენტი, კლიენტებს შეუძლიათ ტრენინგი გაიარონ დაბალი სიზუსტის ქვემოდელზე, ხოლო სერვერს შეუძლია გაუმკლავდეს მაღალი სიზუსტის ქვემოდელს. ამ სქემის გავლენის თეორიული ანალიზი საინტერესო თემაა მომავალი კვლევისთვის, რომელსაც შეუძლია მოგვაწოდოს ინსტრუქცია, თუ როგორ უნდა განვითაროთ SL შეკუმშვით, რესურსებით შეზღუდულ სიტუაციებში.

ბოლო ამოცანა, რომელიც უნდა გადავწყვიტოთ, არის სერვერის მხრიდან გამოთვლითი დატვირთვის შემცირება. მიუხედავად იმისა, რომ პერიფერიული სერვერი ზოგადად უფრო მძლავრია, ვიდრე პერიფერიული მოწყობილობა, ის ასევე შეიძლება გახდეს სუსტი წერტილი PSL/SFL-ში, რადგან სერვერი შეიძლება მოემსახუროს კლიენტების დიდ რაოდენობას და ხშირად აიღოს ტრენინგის დატვირთვის უმეტესი ნაწილი. ამ საკითხის მოსაგვარებლად, განხილულია ეფექტიანი პარალელური გაყოფილი სწავლება (EPSL), რათა შემცირდეს უკან გავრცელებული გრადიენტების განზომილება მათი ბოლო შრებზე აგრეგაციის გზით, როგორც ეს ნაჩვენებია ნახ. 8.3-ზე.

ლიტერატურაში აღწერილ არსებულ თანამედროვე SL-ებთან შედარებით, როგორცაა SFL და PSL, ამ მეთოდს შეუძლია შეამციროს გამოთვლითი და კომუნიკაციის ხარჯები უკან გავრცელების დროს $O(M)$ -დან (M კლიენტების რაოდენობა) $O(1)$ -მდე. აგრეგაციის კოეფიციენტი $\phi \in [0, 1]$ წარმოადგენს აქტივაციის გრადიენტების ნაწილს, რომელიც აგრეგირებულია ბოლო ფენაზე. გასათვალისწინებელია ის, რომ EPSL-ს ასევე შეუძლია აკონტროლოს აგრეგაციის კოეფიციენტი ϕ უკან გავრცელების პროცესში, რათა დაამყაროს ბალანსი კომუნიკაციების/გამოთვლითი ხარჯების შემცირებასა და სწავლების სიზუსტეს შორის, სადაც $\phi = 0$ ამცირებს EPSL-ს PSL-მდე. EPSL-ის უპირატესობა სხვა SL მიდგომებთან მიმართებაში ნაჩვენებია ნახ. 8.4-ზე, სადაც უკან გავრცელების გრადიენტები მცირდება სწავლების სიზუსტეზე შესამჩნევი ზემოქმედების გარეშე (ანუ 0.46 პროცენტიანი გაუარესებით, როდესაც მოდელი კონვერგირდება).



ნახ. 8.3. შემოთავაზებული EPSSL სქემა

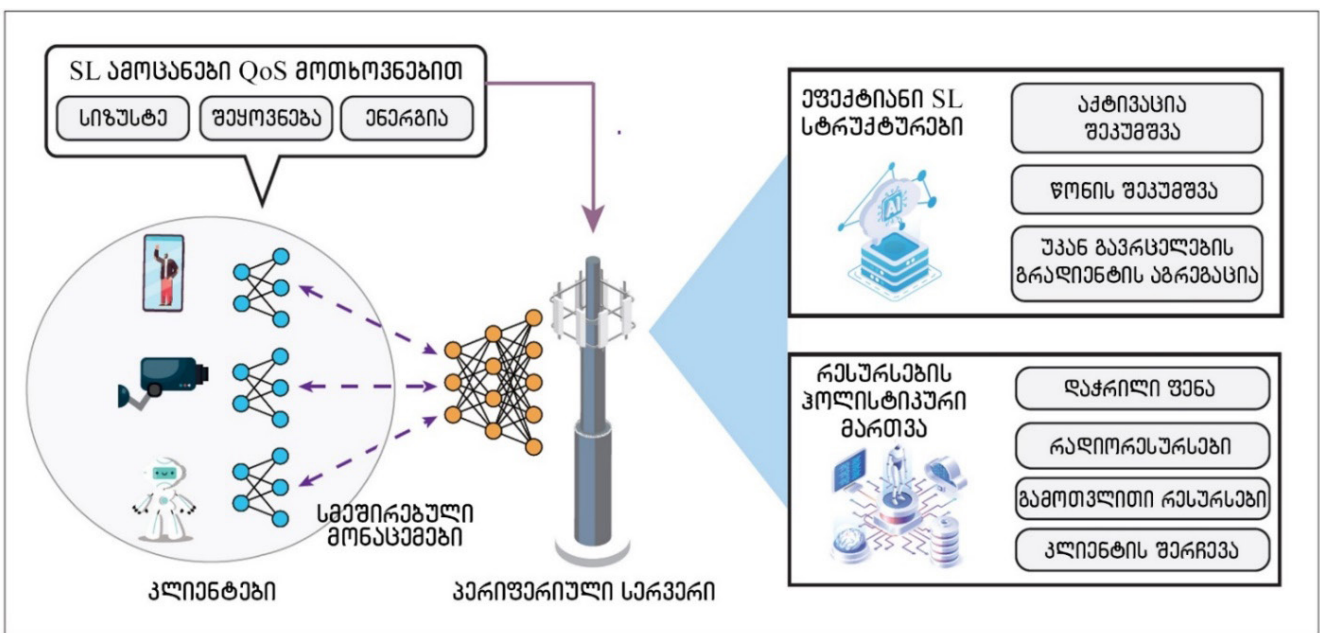


ნახ. 8.4. ResNet-18-ის ტესტირების სიზუსტე HAM10000 მონაცემთა ნაკრებისთვის, შემდეგი ტექნოლოგიებისთვის: FL, vanilla SL, SFL, PSL და EPSSL. მონაცემთა ნიმუშები ნაწილდება 5 კლიენტზე დამოუკიდებლად და თანაბრად განაწილებული შემთხვევითი პარამეტრების მიხედვით, სადაც საერთო ხელმისაწვდომი გატარების ბოლი არის 70 მპკ, თითოეული კლიენტის გამოთვლითი შესაძლებლობა თანაბრად ნაწილდება სეგმენტზე $[0.1, 0.5] \times 10^9$ ციკლი/წმ-ში, ხოლო სერვერის გამოთვლითი შესაძლებლობა არის 7×10^9 ციკლი/წმ.

8.5. რესურსების მართვა SL-ისთვის: ერთი ფიჭის პერსპექტივა

SPL-ის დროს ტრენინგის შეყოვნებას განსაზღვრავს ყველაზე ნელი კლიენტი, რომელიც ასევე ცნობილია როგორც სტრაგლერი. ამ პრობლემის შესამსუბუქებლად, არხები და სერვერის გამოთვლითი რესურსები გონივრულად უნდა გამოიყოს სტრაგლერებისთვის, რათა მოხდეს ტრენინგის პროცესის ოპტიმიზაცია. მიუხედავად იმისა, რომ სტრაგლერის ეფექტი ასევე არსებობს FL-ში, SPL მოიცავს მოდელის გაყოფას და სმეშირებულ მონაცემთა გაცვლას, რაც დიზაინს მნიშვნელოვნად განასხვავებს FL-ის მიდგომებისგან. ამ საჭიროებების გათვალისწინებით, ჩვენ განვიხილავთ ქსელის რესურსების განაწილების პრობლემებს, რომლებიც მორგებულია SEL-ზე ერთი ფიჭაზე, როგორც ნაჩვენებია ნახ. 8.5-ზე.

ქსელის რესურსების განაწილება მჭიდროდ არის დაკავშირებული მოდელის გაყოფასთან SL-ში, რაც განასხვავებს მას FL-ისგან. გაყოფილი ფენა მნიშვნელოვან გავლენას ახდენს ტრენინგის შეყოვნებაზე, რამაც შეიძლება გამოიწვიოს ტრენინგის ცვალებადი დატვირთვა მოწყობილობებსა და პერიფერიულ სერვერებს შორის და სხვადასხვა საკომუნიკაციო ზედნადები ხარჯები ფენის გამოსასვლელი მონაცემების ზომის გამო. კერძოდ, მოდელის „უფრო ღრმა“ ფენაზე გაყოფისას, მეტი გამოთვლითი დატვირთვა რჩება კლიენტის მხარეს, ხოლო კომუნიკაციის ზედნადები შეიძლება პოტენციურად შემცირდეს, რადგან ფენის გამოსასვლელი მონაცემების ზომა ხშირად მცირდება, როდესაც ის გადადის უფრო ღრმად, მაგალითად, როგორც ეს ხდება უმეტეს CNN-ებში. შესაბამისად, მოდელის გაყოფისა და რესურსების განაწილების ერთობლივი ოპტიმიზაცია აუცილებელია გამოთვლით და საკომუნიკაციო რესურსებს შორის კარგი ბალანსის მისაღწევად. მეორე მხრივ, რადგან SL სერვერი მხარს უჭერს პარალელურ ტრენინგს მრავალი კლიენტისთვის, უფრო მეტი გამოთვლითი და არხის რესურსების გამოყოფა სტრაგლერისთვის აუცილებელია მისი შეზღუდული ლოკალური გამოთვლითი და კომუნიკაციის შესაძლებლობების კომპენსაციისთვის. ამ მიმართულებით, ლიტერატურაში განხილულია კლასტერზე დაფუძნებული SL, რომელშიც კლიენტები ერთდროულად ატრენინგებენ მოდელს თითოეულ კლასტერში, რომელიც SFL-ზეა დაფუძნებული. შემდგომში, მოდელი გადის ტრენინგს სხვადასხვა ჯგუფში, ტრადიციული SL მეთოდის საფუძველზე. ეს მიდგომა სტოქასტურად ოპტიმიზაციას უკეთებს დაჭრილი ფენის შერჩევას, მოწყობილობების კლასტერირებას და რადიოსპექტრის განაწილებას, სადაც დაჭრილი ფენის არჩევის გადაწყვეტილება მიიღება უფრო დიდ დროში, ხოლო მოწყობილობის კლასტერირება და რადიოსპექტრის განაწილების გადაწყვეტილებები მიიღება უფრო მცირე დროში. შემდეგი ნაბიჯის გადადგმისას, სასიცოცხლოდ მნიშვნელოვანია PSL/SFL-ისთვის რესურსების მოთხოვნის დაგეგმვის სქემების შემუშავება. არსებული გადაწყვეტილებები გამოყოფს ფიქსირებულ რესურსებს თითოეული კლიენტისთვის ტრენინგის რაუნდის განმავლობაში. თუმცა, სტატიკური რესურსის ეს დანაწევრება რესურსებს უმოქმედოდ ტოვებს SL-ში დროის მნიშვნელოვანი ნაწილის განმავლობაში. მაგალითად,



ნახ. 8.5. SL-ის ილუსტრაცია ერთფიჭიანი სისტემებისთვის

როდესაც კლიენტი ახორციელებს პირდაპირ გავრცელებას, ამ დროს არ არის გადასაცემი მონაცემები, რის გამოც მინიჭებული არხები და სერვერის გამოთვლითი რესურსები უმოქმედოა. FL-ისგან განსხვავებით, რომელსაც აქვს გაცვლილი მონაცემების ფიქსირებული ზომა (ანუ მოდელის ზომა), SL-ს აქვს მოქნილობა, მართოს გამოთვლითი და საკომუნიკაციო ოვერჰედები სერიული დამუშავების გზით. ამიტომ, PSL/SFL-ისთვის რესურსების მოთხოვნის დაგეგმვის შესწავლა, რომელიც დინამიკურად ანაწილებს არხებსა და გამოთვლით რესურსებს კლიენტებზე, რომლებსაც სჭირდებათ შეყვანების მინიმუმამდე შემცირება, შემდგომი კვლევის საკითხია.

რესურსების შეზღუდვის გამო, ტრენინგისთვის ყველა აქტიური კლიენტის შერჩევა შეიძლება არაპრაქტიკული იყოს. კლიენტის ნაწილობრივი მონაწილეობის გათვალისწინებით, მისი შერჩევა გადამწყვეტ როლს თამაშობს SEL-ში. 6G პერიფერია მოითხოვს კლიენტის შერჩევის ერთიან სტრუქტურას ორი ფაქტორის გათვალისწინებით: შერჩეული კლიენტების რაოდენობა (ან ტრენინგის მონაცემების ნიმუშები) და მონაცემთა მრავალფეროვნება. ერთი მხრივ, ზოგიერთი სამეცნიერო ნაშრომი DL-ისთვის მიზნად ისახავს რაც შეიძლება მეტი კლიენტის შერჩევას რესურსების ჰეტეროგენურობით ვადის მოთხოვნების შესაბამისად. ეს დასტურდება იმით, რომ ტრენინგში უფრო მეტი მონაწილის (ან ეკვივალენტური მონაცემების მეტი ნიმუშის) ჩართვა ზოგადად იწვევს კონვერგენციის უფრო სწრაფ სიჩქარეს. მეორე მხრივ, კლიენტების რაოდენობის მაქსიმიზაციამ შეიძლება გამოიწვიოს მიკერძობული მოდელი, რადგან ამ შემთხვევაში კლიენტის მოწყობილობები არხის ცუდი პირობებით (მაგალითად, ფიჭის პერიფერიაზე) და შეზღუდული გამოთვლითი შესაძლებლობებით, სავარაუდოდ, გამოირიცხება. ამიტომ, ასევე აუცილებელია კლიენტების შერჩევა მათი მონაცემთა განაწილების მიხედვით. FL-ისგან განსხვავებით, SL-ს შეუძლია კლიენტების შერჩევა სემშირებული მონაცემების საფუძველზე, რაც არსებითად ორიგინალური მონაცემების მაღალი დონის მახასიათებლებს წარმოადგენს. პერსპექტიული იდეაა შევარჩიოთ კლიენტების ნაკრები სემშირებული მონაცემებით, რომლებიც უკეთ წარმოადგენენ სემშირებულ მონაცემთა გლობალურ განაწილებას. ამ სტრატეგიის ეფექტიანობა მოითხოვს შემდგომ დადასტურებას. გასათვალისწინებელია, რომ მონაცემთა მრავალფეროვნების უზრუნველყოფა შეიძლება ეწინააღმდეგებოდეს მოცემულ ვადაში მეტი კლიენტის შერჩევის მიზანს. კლიენტის შერჩევის ერთიანი სისტემა, სავარაუდოდ, დააბალანსებს არჩეულ კლიენტთა რაოდენობას და მონაცემთა მრავალფეროვნებას SL-ისთვის სწავლების ეფექტიანობის ოპტიმიზაციის მიზნით.

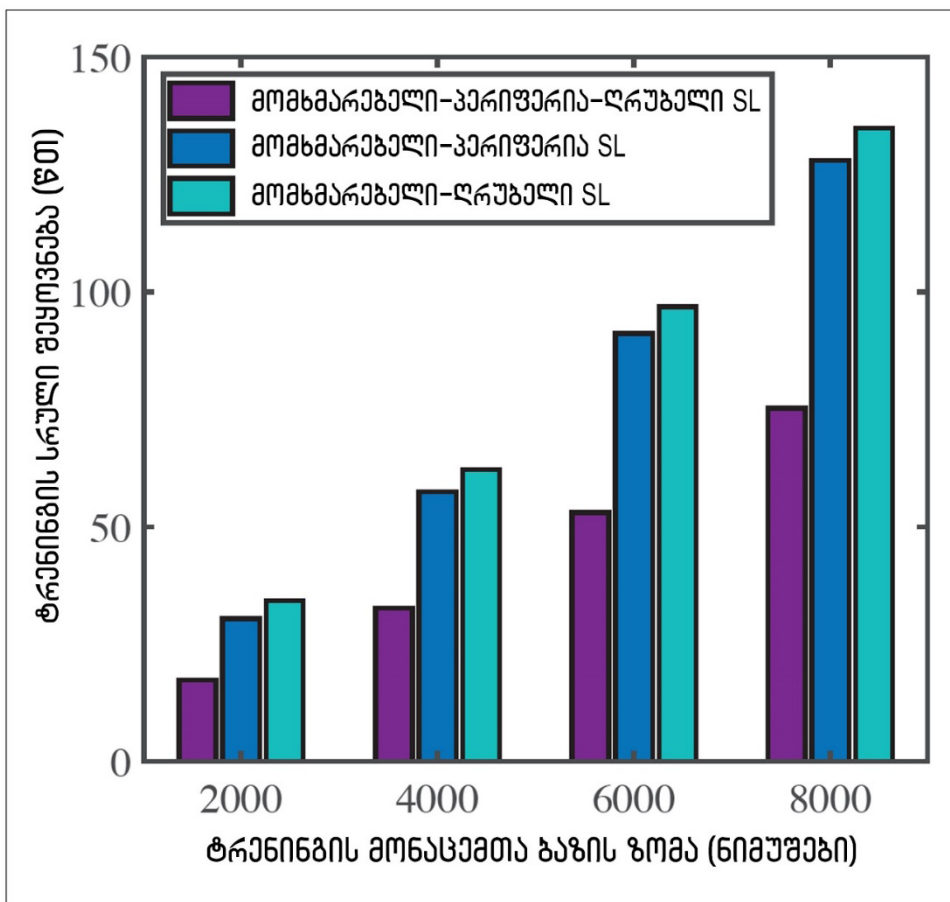
8.6. რესურსების მართვა SL-ისთვის: ქსელური ინჟინერიის პერსპექტივა

ხელოვნური ინტელექტის მოდელის მზარდი ზომა წარმოადგენს არსებით გამოწვევას EL-ისთვის. მრავალპერიფერიული გაყოფილი სწავლების საფუძველზე, ჩვენ შეგვიძლია განვათავსოთ დიდი მოდელები 6G პერიფერიაზე, ხოლო გამოთვლითი და მეხსიერების შეზღუდვები გადავლახოთ სამუშაო დატვირთვის გაზიარებით განაწილებულ პერიფერიულ სერვერებზე. გარდა ამისა, მოდელის განთავსება და მიგრაცია სავარაუდოდ გახდება SEL-ის ძირითადი კომპონენტები. ეს პარაგრაფი ეძღვნება იმ ასპექტებს, რომლებიც განიხილავს SEL-ს ქსელის პერსპექტივიდან.

პრაქტიკული 6G სისტემები აღჭურვილია იერარქიული გამოთვლითი არქიტექტურით, სხვადასხვა დონის ღრუბლოვანი/პერიფერიული სერვერებით, როგორც ილუსტრირებულია ნახ. 8.2-ზე. ეფექტიანი სწავლების გასაადვილებლად, აუცილებელია რესურსების მრავალსაფეხურიანი კოორდინაცია. მნიშვნელოვანია აღინიშნოს, რომ უფრო გავრცელებულ ორდონიან SL-თან შედარებით, მრავალდონიანი სერვერის თანამშრომლობა უზრუნველყოფს უფრო მეტ მოქნილობას კომუნიკაციასა და გამოთვლებს შორის დაბალანსებული ურთიერთობის მისაღწევად. მრავალდონიანი SL-ის ეფექტიანობის საჩვენებლად, განვიხილოთ სამსაფეხურიანი არქიტექტურა: მომხმარებელი-პერიფერია-ღრუბელი. ამ შემთხვევაში, კომუნიკაციის სუსტი წერტილი და შეყვანება ხშირად დევს პერიფერია-ღრუბლის ლინკზე. იმავდროულად, როგორც უკვე აღვნიშნეთ, ფენის ზომა მცირდება, რადგან ის უფრო ღრმად პროგრესირებს ბევრ პრაქტიკულ მოდელში, როგორცაა CNN. ამ დაკვირვების საფუძველზე, ზოგიერთი ფენის მინიჭება საბოლოო მომხმარებლებისთვის და ზოგიერთი სხვა ფენის მინიჭება პერიფერიულ სერვერზე საშუალებას იძლევა გვექონდეს ღრმა და უფრო „ვიწრო“ გაყოფილი ფენა პერიფერიასა და ღრუბელს შორის, რითაც მცირდება კომუნიკაციის ხარჯები. ამის საპირისპიროდ, ორსაფეხურიანი მომხმარებლის ღრუბლოვანი არქიტექტურა შეიძლება მოიცავდეს კომუნიკაციის გადაჭარბებულ შეყვანებას ღრუბელთან დიდი მოცულობის სემშირებული მონაცემების გაცვლის გამო უფრო „ფართო“ ადრეული ფენისთვის (რადგან მომხმარებელს შეუძლია მხოლოდ რამდენიმე ფენის რეალიზაცია). მეორე მხრივ, ორსაფეხურიანი არქიტექტურის

გამოყენება მომხმარებლის პერიფერიაზე, თუმცა გამორიცხავს ღრუბელი-პერიფერია გადაცემების საჭიროებას, მას არ გააჩნია ადეკვატური გამოთვლითი სიმძლავრე პერიფერიაზე. ხუთკლიენტისანი PSL-ის განხილვით, ნახ. 8.6 გვიჩვენებს ღრუბელი-პერიფერია-მომხმარებელი SL არქიტექტურის იერარქიულ უპირატესობას აღნიშნულ ორსაფეხურიან სისტემებთან შედარებით. ამ მიზნით, იერარქიული SL-ის შესწავლა პოტენციურად მეტი დონეებით ფართომასშტაბიანი მომხმარებლებისთვის არის პერსპექტიული კვლევითი მიმართულება.

ჩვენ განვიხილავთ წარმოდგენილ სცენარებს საერთო უჯრედოვან ქსელში (mesh network). ზემოხსენებული იერარქიული SL არის მრავალნახტომიანი SL-ის ტიპი, მაგრამ შემოიფარგლება მხოლოდ „ვერტიკალური“ პარადიგმით, რომელიც შედგება მხოლოდ სხვადასხვა დონის სერვერებისგან. უფრო ზოგადი გაგებით, მრავალრიცხოვან პატარა/მაკრო საბაზო სადგურებს შეუძლიათ შექმნან პერიფერიული სერვერების უჯრედოვანი ქსელი მრავალნახტომიანი SL-ისთვის. ძირითადი მოტივაცია არის სამუშაო დატვირთვის უკეთ გაზიარება მრავალ სერვერზე, რათა გაუმკლავდეს მოდელის გამოთვლითი კუთხით ინტენსიურ ტრენინგს. მუშაობის ოპტიმიზაციისთვის აუცილებელია მოდელის გაყოფისა და მონაცემთა მარშრუტიზაციის ერთობლივი სისტემის დიზაინის შესწავლა მრავალნახტომიან გამოთვლით ქსელებში, გატარების ბოლის, გამოთვლების და მესხიერების შეზღუდვების გათვალისწინებით. 5G მობილურ ქსელებში, ცენტრალიზებულ სმეშირებულ მონაცემთა მარშრუტირება შეიძლება განხორციელდეს ქვემოდელების გაყოფის/განთავსების და გამოთვლითი/ (გატარების ბოლის) რესურსის შეზღუდვების გათვალისწინებით ცენტრალიზებული გზით SDN-ით. ეს მიდგომა უფრო ეფექტიანი იქნება, ვიდრე განაწილებული მარშრუტიზაცია ცენტრალური კონტროლერის მიერ მიღებული გლობალური ცოდნის გამო.



ნახ. 8.6. მომხმარებელი-პერიფერია-ღრუბელი, მომხმარებელი-პერიფერია და მომხმარებელი-ღრუბელი არქიტექტურების ტრენინგის სრული შეყოვნება HAM10000-ზე სამიზნე სიზუსტის მისაღწევად, როგორც ტრენინგის მონაცემთა ნაკრების ზომის ფუნქცია. აქ მონაცემთა ნიმუშები ნაწილდება 5 კლიენტზე დამოუკიდებელად და თანაბრად განაწილებული შემთხვევითი პარამეტრების მიხედვით, სადაც ღრუბლის გამოთვლითი შესაძლებლობა არის 20×10^9 ციკლი/წმ, პერიფერია-ღრუბელი ლინკის გამტარუნარიანობა დაყენებულია მომხმარებლის პერიფერიული ლინკის გამტარუნარიანობის 1/20-ზე და სხვა ძირითადი პარამეტრები შეესაბამება ნახ. 8.4-ის პარამეტრებს.

სტატიკური პერიფერიული სერვერების გარდა, მრავალნახტომიანი SL ასევე შეიძლება განხორციელდეს მობილურ სპეციალიზებულ (ad hoc) ქსელებში 6G პარადიგმის ფარგლებში. მაგალითად, სატრანსპორტო შემადგენლობას შეუძლია განახორციელოს SL ჯგუფში მანქანებს შორის მოდელის გაყოფით და გაზიარებით, რომელიც ეფუძნება V2V კომუნიკაციებს. D2D კომუნიკაციის საფუძველზე, სმარტფონის მომხმარებლებს ასევე შეუძლიათ დიდი მოდელის მომზადება მისი მრავალ ნაწილად დაყოფით. ყველა ამ სცენარს შეუძლია გამოიყენოს დისპერსიული რესურსები ქსელის პერიფერიაზე, რადგან ერთმა მოწყობილობამ ან პერიფერიულმა სერვერმა შეიძლება ვერ შეძლოს ინდივიდუალურად გაუმკლავდეს გამოთვლითი ინტენსიური ტრენინგის ამოცანებს.

6G ქსელის პერიფერია ამუშავებს შემნახველი მოწყობილობების განაწილებულ რესურსებს, რომლებიც შეიძლება გამოყენებულ იქნეს ზოგადი AI მოდელების განთავსებისა და მიგრაციისთვის SL ოპერაციების გასაადვილებლად. SL-ს და დასკვნას შეუძლია გამოიყენოს „მოდელის ნაწილობრივი განთავსება“ ქეშირების მუშაობის გასაუმჯობესებლად, იმის გამო, რომ მომხმარებლებს და სერვერებს შეუძლიათ ნეირონული ქსელების ნაწილის რეალიზაცია. აქედან გამომდინარე, არსებობს კომპრომისი პერიფერიაზე შენახვისა და კომუნიკაციისა და გამოთვლით ხარჯებს შორის. მიუხედავად იმისა, რომ მოდელის უფრო დიდი ნაწილის განთავსება პერიფერიულ კვანძში იკავებს მეტ ადგილს, ეს პოტენციურად ამცირებს კომუნიკაციის ხარჯებს სხვა კვანძებთან მონაცემთა გაცვლისთვის. აქედან გამომდინარე, გადამწყვეტი მნიშვნელობა აქვს მოდელის გაყოფისა და მოდელის განთავსების ერთობლივად შემუშავებას სერვისის განთავსების/მიგრაციისთვის, გატარების ბოლის, გამოთვლითი და მეხსიერების შეზღუდვების გათვალისწინებით. გარდა ამისა, მოდელის განთავსება/მიგრაცია SL-ში შეიძლება ითვალისწინებდეს კლიენტების გეოგრაფიული მონაცემების დროში ცვალებად განაწილებას. არსებობს ზოგადი (ნაწილობრივი) მოდელები, რომლებიც ერგება მომხმარებელთა/მომსახურების ფართო სპექტრს (მაგალითად, ავტონომიური მართვა), მაგრამ არ გააჩნია უმაღლესი ამოცანის შესრულების მახასიათებლები, ასევე დახვეწილი (ნაწილობრივი) მოდელები, რომლებიც სპეციალიზებულია გარკვეულ ამოცანებზე (მაგალითად, ავტონომიური მართვა ქალაქისთვის წვიმიან დღეებში). მოდელების ტრენინგისას/დასკვნისას, ამ (ნაწილობრივი) მოდელების სათანადო განთავსება იძლევა მონაცემთა/მოდელის რეალურ დროში გადაცემას მონაცემთა წყაროებსა და გამოთვლით სერვერებს შორის. ეს ფაქტორები საჭიროებს სერვისის განთავსების/მიგრაციის პრობლემების ხელახლა გადახედვას პერიფერიული გამოთვლითი პარადიგმის ქვეშ.

8.7. ღია პრობლემები და მომავალი კვლევითი მიმართულებები

მიუხედავად იმისა, რომ ჩვენ გამოვყავით ზოგიერთი კვლევითი პრობლემა და შესაბამისი გადაწყვეტილებები, ჯერ კიდევ არსებობს რამდენიმე აქტუალური კვლევითი მიმართულება. ჩვენ განვიხილავთ ამ ღია პრობლემებს შემდეგნაირად.

კონვერგენციის ანალიზი PSL-ისთვის:

კონვერგენციის ანალიზი გადამწყვეტ როლს თამაშობს SEL-ისთვის რესურსების ოპტიმიზაციაში, რადგან ის გვეხმარება რესურსების გამოყოფაში ტრენინგის დასაჩქარებლად. არსებითად, PSL შეიძლება ჩაითვალოს SFL-ის განსაკუთრებულ შემთხვევად, სადაც კლიენტის მხარის მოდელები არასოდეს იქნებიან აგრეგირებული. ზოგადად, SFL-ის კონვერგენცია ჯერ კიდევ საჭიროებს დამატებით გაგებას, განსაკუთრებით იმაზე, თუ როგორ იმოქმედებს კლიენტის მხრიდან მოდელის აგრეგაცია მოდელის კონვერგენციაზე. უკიდურეს შემთხვევაში, PSL გამორიცხავს კლიენტის მხრიდან მოდელების აგრეგაციის აუცილებლობას, რის შედეგადაც წარმოიქმნება ერთი და იგივე სერვერის მოდელი და კლიენტის მხარის განსხვავებული მოდელები ყველა მოწყობილობაზე. მიუხედავად იმისა, რომ ემპირიულმა ექსპერიმენტებმა აჩვენა, რომ მისი უარყოფითი გავლენა სწავლების მახასიათებელზე მცირეა, რამდენადაც ჩვენთვის არის ცნობილი, ჯერ არ არსებობს თეორიული ანალიზი, რომელიც აჩვენებს PSL-ის კონვერგენციას, რაც მოითხოვს შემდგომ კვლევით ძალისხმევას.

ასინქრონული SEL:

მიმდინარე SFL/PSL სტრუქტურაში, პერიფერიული სერვერი განაახლებს მოდელს მხოლოდ ყველა კლიენტისთვის ტრენინგის შესრულებისას. თუმცა, როდესაც პერიფერიული მოწყობილობა მოითხოვს ტრენინგის შეფერხების ან გადაცემის შეყოვნების გაცილებით მეტ დროს არხის მკაცრი პირობების გამო, სხვა მონაწილეები უნდა იყვნენ ლოდინის რეჟიმში. ასინქრონული PSL საშუალებას აძლევს სერვერს განაახლოს სერვერის მხარის მოდელი, სანამ ის დაასრულებს ტრენინგს ერთი ან გარკვეული რაოდენობის კლიენტებისთვის, რითაც

გაზრდის რესურსების გამოყენებას. თუმცა, ასინქრონული FL-ის მსგავსად, ეს პროცესი პოტენციურად აფერხებს მოდელის კონვერგენციას, რადგან სტრაგეგიები ნაკლებად იქნება წარმოდგენილი მოდელის განახლებებში ნაკლები მონაწილეობის გამო. შესაბამისად, გადამწყვეტი მნიშვნელობა აქვს „მოდელის დაბერების“ მართვას ასინქრონულ SFL-სა და PSL-ში შესაბამისი მოდელის აგრეგაციის სიხშირის შერჩევით, რომელიც უნდა მოერგოს რესურსების ჰეტეროგენურობას უსადენო პერიფერიაზე.

SEL ეტიკეტების კონფიდენციალურობის შენარჩუნებით:

ჩვეულებრივ SL-ში, ეტიკეტები უნდა განთავსდეს სერვერის მხარეს. თუმცა, მონაცემთა ეტიკეტი ზოგჯერ შეიცავს კერძო სენსიტიურ ინფორმაციას (მაგალითად, დაავადებას, რომელიც შეიძლება ჰქონდეს პაციენტს), რომელიც უნდა იყოს დაცული პერიფერიული სერვერიდან. ამის დასაძლევად, ლიტერატურაში შემოთავაზებულია P-ფორმის SL, სადაც პირველი და ბოლო ფენა მოთავსებულია კლიენტის მხარეს, რაც საშუალებას აძლევს გამომავალ ფენებსა და მათ შესაბამის ეტიკეტებს დარჩეს ლოკალური. თუმცა, ეს პარადიგმა აწესებს დამატებით საკომუნიკაციო ხარჯებს დამატებითი გაყოფის წერტილის არსებობის გამო, რაც მოითხოვს ორი გაყოფილი ფენის ფრთხილად შერჩევას, ასევე უსადენო ქსელებზე დამატებით მონაცემთა გადაცემის ეფექტიან მართვას.

8.8. მერვე თავის დასკვნა

NG-ის ეპოქაში, ჩვენ მოველით, რომ გაყოფილმა სწავლებამ შეიძლება მნიშვნელოვნად შეამციროს რესურსების მოთხოვნა მოწყობილობაზე მოდელის ტრენინგისთვის, რაც საშუალებას მისცემს ML-ს სწრაფად გაფართოვდეს მასობრივ IoT მოწყობილობებზე. ამ თავში მიმოხილულია SL-ის ბოლოდროინდელი მიღწევები და აღწერილია მისი უწყვეტი ინტეგრაცია NG პერიფერიასთან, გამომდინარე როგორც სწავლების, ასევე კომუნიკაციის პერსპექტივიდან. ამ სფეროში, რომელიც ძირითადად შეუსწავლელია, არსებობს კვლევითი შესაძლებლობების მდიდარი ნაკრები, როგორცაა უფრო ეფექტური და ეფექტიანი SL სტრუქტურების შემუშავება და რესურსების განაწილების სტრატეგიები, რომლებიც მორგებულია SL-ზე. ვიმედოვნებთ, რომ ამ თავში წარმოდგენილი შედეგები მიიპყრობენ კვლევითი ორგანიზაციების, ხელოვნური ინტელექტის სექტორების, სატელეკომუნიკაციო ინდუსტრიებისა და სტანდარტიზაციის ორგანიზაციის ყურადღებას, რაც საბოლოოდ გადააქცევს SEL-ს სიცოცხლისუნარიან რეალობად მომავალ NG ეპოქაში.

თავი 9 AI-ზე ორიენტირებული რესურსების უზრუნველყოფა 6G იმერსიული სერვისებისთვის

9.1. შესავალი

დღესდღეობით, 6G-ის მიღწევებით, იმერსიული სერვისები, როგორებიცაა ვირტუალური რეალობა (VR) და მომატებული რეალობა (AR), იწყებენ უზარმაზარ ინტერესს მოზიდვას. 6G-ის უმაღლესი დონის მახასიათებლები, როგორცაა ულტრა დაბალი შეყოვნება, ულტრა მაღალი საიმედოობა, ულტრა მაღალი ენერგოეფექტიანობა და სისტემის ულტრა მაღალი სიმძლავრე, ქმნის იმერსიული სერვისის ქვაკუთხედს. კერძოდ, კიბერსივრცის ყველა არსებითი ასპექტი გაადვილებულია იმერსიულ სერვისებში ინტეგრაციისთვის, რაც უზრუნველყოფს 6G სერვისის მომხმარებლებისთვის სამგანზომილებიან (3D) ვირტუალურ სამყაროსა და ფიზიკურ სამყაროს შორის ღრმა ურთიერთქმედების შესაძლებლობას.

კერძოდ, იმერსიული სერვისები მოიცავს სხვადასხვა ტარებად მოწყობილობას, როგორებიცაა თავზე დასამაგრებელი დისპლეები (HMD), ჭკვიანი საათები და ჭკვიანი სათვალეები, რათა უზრუნველყოს ვირტუალური სცენების მაღალი ხარისხის აღქმა SU-ებისთვის. იმერსიული სერვისების პოპულარობა და სატარებელი ხელსაწყოების გამრავლება სერიოზულ წინააღმდეგობრივ პრობლემებს ქმნის რესურსების მაღალ მოთხოვნებსა და მათ შეზღუდვებს შორის. იმერსიული სერვისების ბაზრის გაჩენა წინააღმდეგობების მოგვარების იმედს იძლევა. იმერსიული სერვისების ბაზარი შედგება SU-ებისა და SP-ებისგან, სადაც SU-ებს შეუძლიათ საფასური გადაიხადონ რესურსების გამოყენებისთვის, მაგალითად გატარების ზოლის, ცენტრალური პროცესორების (CPU) და გრაფიკული დამუშავების ბლოკების (GPU), SP-ების, ღრუბლოვანი ცენტრებისა და პერიფერიული სერვერების ხელმისაწვდომობისთვის. ასეთი სერვისების ბაზრის კეთილდღეობა შესაძლებელს ხდის იმერსიულ სერვისების გავრცელებას. ბლოკჩეინის დახმარებით, ასევე უზრუნველყოფილია რესურსების უსაფრთხო გაქირავება და წახალისება, რაც კიდევ უფრო აღძრავს გაძლიერებულ ინტერესს იმერსიული სერვისების ბაზრის მიმართ.

მიუხედავად იმისა, რომ ზემოაღნიშნულ ტექნოლოგიებს ახალი სარგებელი მოაქვთ იმერსიული სერვისებისთვის, რჩება მნიშვნელოვანი გამოწვევები რესურსების მიწოდებასთან დაკავშირებით:

- იერარქიასთან შესაფერისი სტრუქტურის არარსებობა იწვევს 6G რესურსების გამოყენების ნაკლებ ეფექტიანობას. იმერსიული სერვისების ბაზრის აყვავებასთან ერთად, აქ შემოდის SP-ების დიდი რაოდენობა. ბაზრის იერარქიული მახასიათებლები სულ უფრო მკაფიო ხდება, დაწყებული ღრუბლოვანი მონაცემთა ცენტრებიდან დამთავრებული სერვერებით და თუნდაც ინდივიდებით. არსებობს მთელი რიგი კვლევები, რომლებიც ფოკუსირებულია ლატენცურ ანალიზზე, წამახალისებელი მექანიზმის დიზაინსა და ბლოკჩეინზე დაფუძნებულ სავაჭრო მექანიზმზე. იერარქიასთან დაკავშირებული სტრუქტურის არარსებობა იწვევს რესურსების ცუდ და არასრულ გამოყენებას.
- სერვისებსა და რესურსებს შორის დროითი დამოკიდებულების უგულებელყოფა იწვევს 6G რესურსების ნაკლებ ეფექტიანობას. წინა კვლევები რესურსებით უზრუნველყოფის პრობლემას განიხილავს, როგორც ერთჯერად ამოცანას. მიუხედავად იმისა, რომ ოპტიმიზებული პრობლემის გადაწყვეტა შედარებით მარტივია, გლობალური ოპტიმალური მახასიათებელი, როგორცაა რესურსების ეფექტიანობა, არადაამაკმაყოფილებელია.
- 6G ჰეტეროგენული რესურსების სტიმულირება არასაკმარისია. როგორც უკვე აღვნიშნეთ, ჰეტეროგენული 6G რესურსები ასოცირდება და დაკავშირებულია ისეთ მახასიათებლებთან, როგორებიცაა გატარების ზოლი, CPU და GPU. სასწრაფოდ საჭიროა წამახალისებელი მექანიზმი, რომელიც მოიცავს ჰეტეროგენულ 6G რესურსებს.

ზემოაღნიშნული პრობლემების გადასაჭრელად, ამ თავში, AI ტექნოლოგიის გამოყენებით, ჩვენ განვიხილავთ AI-ზე ორიენტირებული რესურსების უზრუნველყოფის საკითხებს 6G იმერსიული სერვისებისთვის, ანუ Almers-6G. აქ რესურსებით უზრუნველყოფა და ჰეტეროგენული აუქციონი განიხილება სხვადასხვა იერარქიულ რეგიონში. დიდ რეგიონში რესურსების უზრუნველყოფა მიზნად ისახავს ჰეტეროგენული რესურსების გამოყოფას 6G SU-ების აღქმის გამოცდილების ხარისხის დასაკმაყოფილებლად. მცირე რეგიონში, ჰეტეროგენული აუქციონი

ფოკუსირებულია რესურსების შესაბამისობასა და ფასების განსაზღვრაზე. ის მიზნად ისახავს ოპტიმალური კომპრომისის მიღწევას იმერსიული სერვისების შეყოვნებასა და ენერგოეფექტიანობას შორის. გარდა ამისა, შემოთავაზებული პრობლემის უშუალო გადაჭრის გამოთვლითი სირთულის შესამცირებლად, გამოყოფილია ოპტიმიზაციის ცვლადები ეფექტური ალგორითმის შესაქმნელად. ამ თავის შედეგების რეზიუმირება შეიძლება შემდეგნაირად:

- წარმოდგენილია AI-ზე ორიენტირებული რესურსების უზრუნველყოფის მიდგომა 6G იმერსიული სერვისებისთვის, სახელად Almers-6G. Almers-6G იყენებს რეგიონთაშორის მექანიზმს, ის იძლევა რესურსების დროთა შორის უზრუნველყოფას, რათა დააკმაყოფილოს SU-ების ინტერაქტიული გამოცდილება ჰეტეროგენული აუქციონის მექანიზმით.
- დიდ რეგიონში, SU-ების ასოციაციისა და რესურსების განაწილების ერთობლივი ოპტიმიზაციის გზით, ჩამოყალიბებულია რესურსების უზრუნველყოფის პრობლემა, რომელიც მინიმუმამდე ამცირებს დროის მიხედვით გასაშუალოებულ შეყოვნებას ენერჯის შეზღუდვების პირობებში. ზემოაღნიშნული პრობლემის შესასწავლად წარმოდგენილია ლიაპუნოვის ოპტიმიზაციის პრობლემა და პრობლემის გადასაჭრელად გამოიყენება კონტექსტურ-იმერსიული სწავლების ალგორითმი.
- პატარა რეგიონში წარმოდგენილია BDDA-ის მექანიზმი ჰეტეროგენული რესურსებისთვის, სამსახურებრივი ფუნქციების შემუშავებისას, როგორც SU-ის და SP-ის სავაჭრო სტრატეგიები BDDA-ში. გარდა ამისა, შემუშავებულია BDDA სმარტ-კონტრაქტი და ის დანერგულია კონსორციუმის ბლოკჩეინზე მომავალი პოტენციური 6G იმერსიული სერვისებისთვის.

9.2. Almers-6G-ის სტრუქტურა

Almers-6G ძირითადად ორიენტირებულია 6G იმერსიული სერვისების თავსებადობაზე სხვადასხვა დაინტერესებულ მხარეს შორის. კერძოდ, Almers-6G-ში არის ოთხი ტიპის სუბიექტი: 6G სერვისების მომხმარებლები, რომლებიც მოქმედებენ როგორც მყიდველები; 6G სერვისების პროვაიდერები, რომლებიც მოქმედებენ როგორც გამყიდველები; ბლოკჩეინის საშუალებით მოქმედი აუქციონერები; და სერვისების რესურსების რეგიონი. უნდა აღინიშნოს, რომ გამოთვლითი რესურსებით 6G სერვისის ყველა პროვაიდერს შეუძლია შეითავსოს ბლოკჩეინის კვანძების ფუნქცია, რადგან მათ აქვთ უზარმაზარი გამოთვლითი რესურსები. ანუ მათ შეუძლიათ მონაწილეობა მიიღონ რესურსების უზრუნველყოფაში Almers-6G-ში, რაც უფრო მეტად შეესაბამება საჯარო ბლოკჩეინის ბუნებას. ამრიგად, Almers-6G იყენებს საჯარო ბლოკჩეინს და სამუშაოს დადასტურების კონსენსუსს მისი უსაფრთხოების უზრუნველსაყოფად. იმავდროულად, ბლოკჩეინის მიერ შექმნილი პოტენციური წახალისების მექანიზმი უფრო მეტ წევრს აიძულებს მონაწილეობა მიიღონ Almers-6G-ში და დაეხმარება მათ მიიღონ მეტი კეთილდღეობა, რითაც შეუძლიათ მიაღწიონ მრავალ გამარჯვებას.

6G სერვისების მომხმარებელი:

SU აღნიშნავს ობიექტს, რომელიც გამოიწერს 6G იმერსიულ სერვისებს, როგორცაა AR, VR, ვირტუალური შეხვედრები, ვირტუალური თამაშები და ა. შ. 6G სერვისებისთვის გადახდის შემდეგ, SU-ებს აქვთ წვდომა Almers-6G-ზე ნებისმიერ დროს და ნებისმიერ ადგილას ჭკვიანი მოწყობილობებით, როგორცაა HMD, ტარებადი მოწყობილობები და ტვინი-კომპიუტერის ინტერფეისიცი კი. მაღალი ხარისხის იმერსიული სერვისების ძიებაში, SU-ებს უფრო მეტად აწუხებთ ურთიერთქმედების ხარისხი და ფასი, უფრო ზუსტად კი შეყოვნება. SU-ებისთვის, რომლებიც სარგებლობენ მაღალსიჩქარიანი იმერსიული სერვისებით, ურთიერთქმედების შეყოვნება დიდ გავლენას ახდენს ფუნქციონირებაზე. გარდა ამისა, 6G ტექნოლოგიით მხარდაჭერილი საყოველთაოდ გავრცელებული და უსაფრთხოება აბონენტებს საშუალებას აძლევს მიაღწიონ უწყვეტ და უპრობლემო მუშაობას.

6G სერვისების პროვაიდერი:

SP აღნიშნავს ობიექტს, რომელიც უზრუნველყოფს 6G იმერსიულ სერვისებს 6G რესურსების გაყიდვით; ეს შეიძლება იყოს გატარების ზოლი, CPU, GPU და ა. შ. მაგალითად, ზოგიერთი ტარებადი მოწყობილობა შექმნილია ადამიანების სასიცოცხლო ნიშნების მონაცემების შესამოწმებლად, შესაგროვებლად და გასაზომად. ზოგიერთი მოწყობილობა გლობალური პოზიციონირების სისტემით (GPS) გამოიყენება ფიზიკური ობიექტების მოძრაობისა და პოზიციის შესახებ ინფორმაციის შესაგროვებლად. გარდა ამისა, მოწინავე 6G ტექნოლოგია, რომელიც

უზრუნველყოფილია იმერსიული სერვისების ბაზრით, ამ ჭკვიან მოწყობილობებს უფრო ინტერაქტიულს ხდის და ამავე დროს გამოირჩევა უფრო ძლიერი კომუნიკაციისა და გამოთვლითი შესაძლებლობებით.

6G სერვისების რესურსების რეგიონი:

6G სერვისების რესურსების რეგიონი ეხება ობიექტს, რომელიც შედგება მრავალი SP-ისგან, როგორებიცაა ზოგიერთი მაღალტექნოლოგიური სტარტაპი, საზოგადოებები და ორგანიზაციები. 6G სერვისების რესურსების რეგიონში უკიდურესად მასობრივი SP-ები განლაგებულია ფიზიკური სამყაროს ინფორმაციის აღქმის, შეგროვებისა და დამუშავებისთვის, რაც უზრუნველყოფს SU-ებისთვის 6G იმერსიულ ვირტუალურ სამყაროს. რესურსების თითოეულ რეგიონს ჰყავს მენეჯერი, რომელსაც ირჩევენ SP-ები და რომელიც პასუხისმგებელია Almers-6G-ში ბლოკჩეინის ფუნქციონირებაზე. 6G სერვისების რესურსების ყველა რეგიონი ქმნის დიდ რეგიონს, ხოლო SP-ები რესურსების თითოეულ რეგიონში ქმნიან პატარა რეგიონს.

ბლოკჩეინზე დაფუძნებული აუქციონერი:

SU-ები და SP-ები გვთავაზობენ და აფასებენ 6G იმერსიულ სერვისებს წარმოდგენილი BDDA მექანიზმით. აღსანიშნავია, რომ ისინი ერთმანეთს ეჯიბრებიან ფასების სტრატეგიების დამოუკიდებლად განსაზღვრისას. BDDA-ის სამართლიანობის, თავსებადობის, კონფიდენციალურობისა და უსაფრთხოების გასაუმჯობესებლად, ჩვენ განვიხილავთ ბლოკჩეინს, როგორც აუქციონერს. იგი ადარებს SU-ს და SP-ს, BDDA-ში აუქციონის გამარჯვებულების დასადგენად. გარდა ამისა, ბლოკჩეინი აღრიცხავს ქსელის ინფორმაციას, როგორც ტრანზაქციებს, რაც კიდევ უფრო აუმჯობესებს აუქციონის ეფექტიანობას.

Almers-6G-ის შემოთავაზებული სტრუქტურა ნაჩვენებია ნახ. 9.1-ზე, რომელიც მოიცავს ფიზიკური სამყაროს, ურთიერთქმედების, ვირტუალური სამყაროსა და ბლოკჩეინის ფენებს.

ფიზიკური სამყაროს ფენა:

როგორც ნაჩვენებია ნახ. 9.1-ის ქვედა ფენაში, ფიზიკური სამყაროს ფენა მოიცავს ფიზიკურ ობიექტებს და აქტივობებს. კონკრეტულად, ფიზიკური ობიექტები, როგორც წესი, მოიცავს სერვისის აბონენტებს, პროვაიდერებს და რესურსების რეგიონებს. SP-ები რესურსების რეგიონებში განათავსებენ ფიზიკურ ინფრასტრუქტურას, რათა უპასუხონ აბონენტების 6G იმერსიულ მოთხოვნებს. გარდა ამისა, ამ ობიექტების ურთიერთქმედება წარმოშობს სხვადასხვა სახის ინფორმაციას ფიზიკური აქტივობის შესახებ, ანუ ფიზიკური სამყაროს შესახებ მონაცემების აღქმას, რომელიც მუდმივად გროვდება მოწინავე სენსორული ტექნოლოგიების საშუალებით. გარდა ამისა, ფიზიკური აქტივობის ინფორმაცია წარმოქმნის ფიზიკური სამყაროს სარკეს ციფრული ტყუპით და ვიზუალიზდება ვირტუალურ სამყაროში.

ურთიერთქმედების ფენა:

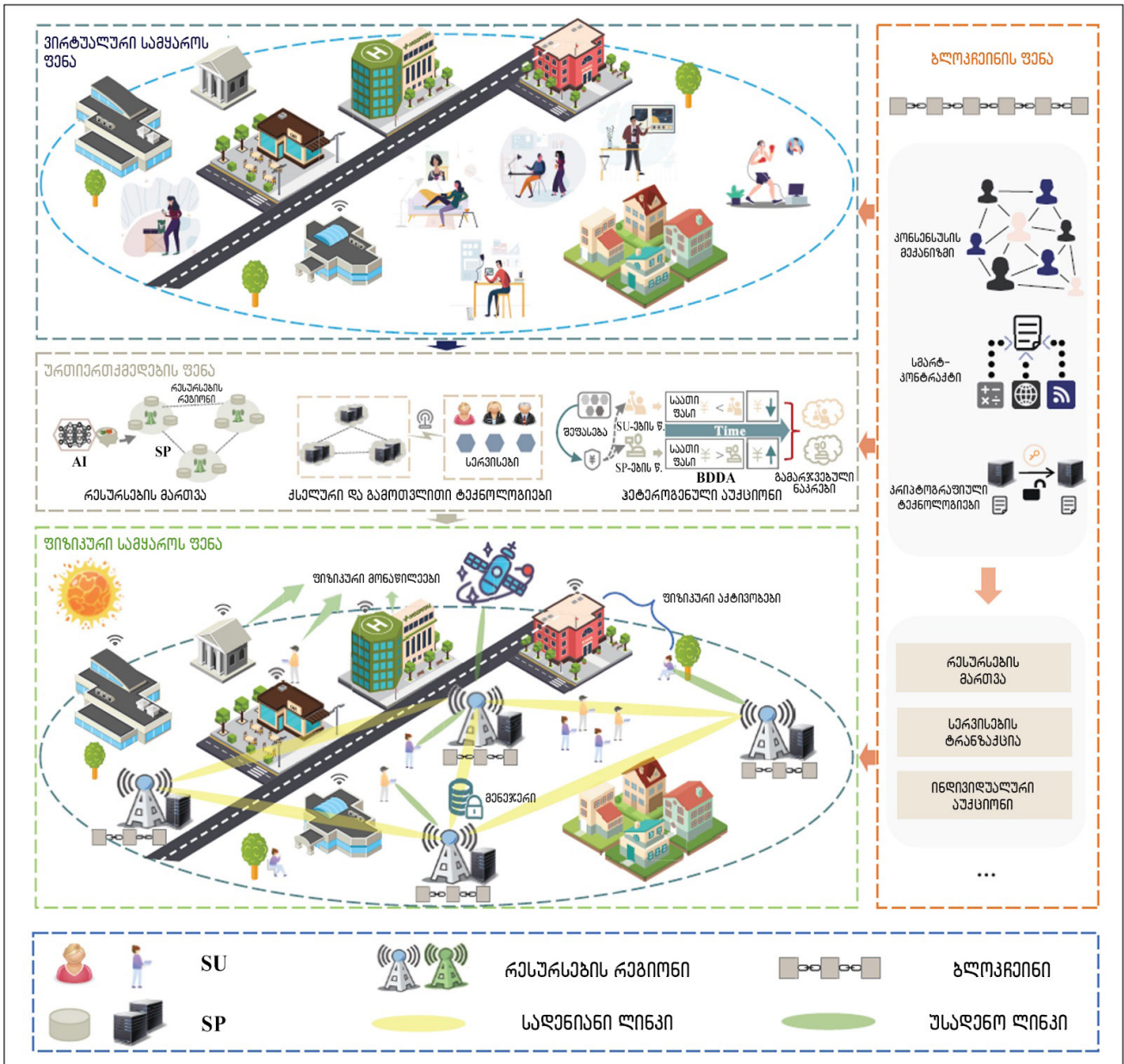
ურთიერთქმედების ფენა არის მნიშვნელოვანი კავშირი ფიზიკურ სამყაროსა და ვირტუალურ სამყაროს შორის. ურთიერთქმედების ფენის მეშვეობით, ფიზიკური სამყაროდან გენერირებული ფიზიკური აქტივობის ინფორმაცია გადაიცემა ვირტუალურ სამყაროში, შემდეგ კი მიიღება უკუკავშირი ვირტუალური სამყაროდან. ამ ინტერაქტიულ პროცესში აუცილებელია უწყვეტი და სრული დაფარვის კავშირის მხარდაჭერა ისეთი ტექნოლოგიების საშუალებით, როგორცაა რესურსების უზრუნველყოფა და ქსელებისა და გამოთვლების ჯვარედინი ინტეგრაცია.

ვირტუალური სამყაროს ფენა:

Almers-6G მიზნად ისახავს შექმნას ზოგადი ვირტუალური სამყარო, რომელშიც ფიზიკურ ობიექტებს შორის აქტივობები სინქრონიზებულია ფიზიკურ სამყაროსთან. სხვა სიტყვებით რომ ვთქვათ, ფიზიკური სამყარო შეიძლება ზუსტად განმეორდეს ვირტუალურ სამყაროში ვირტუალურ ტყუპებთან დინამიკური ურთიერთქმედებით. ვირტუალურ სამყაროში SU-ები აღიქვამენ ვირტუალურ ობიექტებს და იღებენ იმერსიულ სერვისებს. ამჟამად ვირტუალურ სამყაროში იმერსიულმა სერვისებმა დივერსიფიცირებული მიმართულებები მიიღო. მაგალითად, ვირტუალური თამაშები, ვირტუალური განათლება, შეხვედრები და სხვა იმერსიული სერვისები შეიძლება იყოს წარმოდგენილი ნახ. 9.1-ის ზედა ფენაში.

ბლოკჩეინის ფენა:

როგორც საკვანძო ტექნოლოგია Almers-6G-ის საგრძნობლად გასაუმჯობესებლად, ბლოკჩეინის ფენა უზრუნველყოფს უსაფრთხო სერვისის ტრანზაქციებს მესამე მხარის გარეშე. Almers-6G-ში ბლოკჩეინი მუშაობს 6G სერვისების რესურსების რეგიონის მენეჯერებზე. როგორც ნაჩვენებია ნახ. 9.1-ის მარჯვენა ფენაში, ბლოკჩეინი ხასიათდება დეცენტრალიზებით, უცვლელობითა და გამჭვირვალობით. ბლოკჩეინის

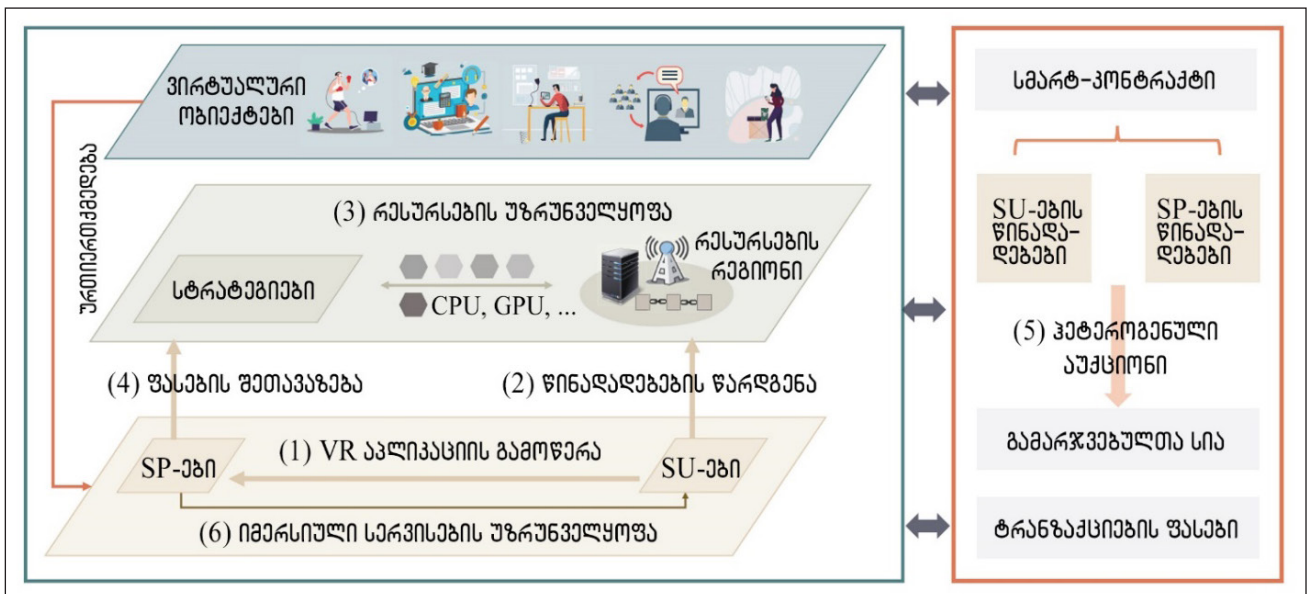


ნახ. 9.1. Almers-6G-ის სტრუქტურა

ფენა უზრუნველყოფს ტრანზაქციების უსაფრთხო შენახვას, რესურსების მენეჯმენტს და მორგებულ აუქციონს უცვლელი წიგნების (ლედჯერების), ჰვანი კონტრაქტებისა და დაშიფრის ტექნოლოგიების გამოყენებით და ა. შ. უფრო კონკრეტულად რომ ვთქვათ, ყველა სერვისის ტრანზაქცია შეიძლება ჩაიწეროს ბლოკჩეინის ლედჯერში, სმარტ-კონტრაქტებით, რომლებიც მოქმედებენ როგორც აგენტები, ფასების და 6G ჰეტეროგენული რესურსების განაწილების ოპტიმიზაციისთვის, როგორცაა გატარების ბოლი, CPU და GPU.

Almers-6G-ის უკეთეს გასაგებად, ქვემოთ მოცემულია გამოყენების შემთხვევის კონკრეტული მაგალითი (ნახ. 9.2). VR აპლიკაციები, როგორცაა ვირტუალური თამაშები და ვირტუალური კონცერტები, დღესდღეობით სარგებლობს უდიდესი პოპულარობით და SU-ებს სთავაზობს იმერსიულ გამოცდილებას ულტრა დაბალი შეყოვნებით და ულტრა საიმელო მოთხოვნებით.

ნახ. 9.2-ის მიხედვით, როდესაც SU გამოიწერს VR აპლიკაციას პირველ ეტაპზე, SU წარუდგენს წინადადებებს Almers-6G-ს, რათა ისარგებლოს ფიზიკურ-ვირტუალური სამყაროს ურთიერთქმედებით მეორე ეტაპზე. ურთიერთქმედების ფენამ უნდა გადაწყვიტოს რამდენი ჰეტეროგენული 6G რესურსი (ანუ კომუნიკაცია და გამოთვლა) უნდა იყოს უზრუნველყოფილი თითოეული სერვისის რესურსების რეგიონიდან, რათა მიაღწიოს



ნახ. 9.2. Almers-6G-ის გამოყენების შემთხვევის მაგალითი

SU-ის კმაყოფილებას მესამე ეტაპზე. რესურსების უზრუნველყოფის სტრატეგიების მოპოვების შემდეგ, მრავალი SP სერვისების რესურსების ყველა რეგიონში გვერდობს ფასებს მეოთხე ეტაპზე აბონენტების მომსახურების მოთხოვნების დასაკმაყოფილებლად. გარდა ამისა, აუქციონის სმარტ-კონტრაქტი ჰეტეროგენული რესურსების უზრუნველყოფისთვის შესრუდება, რათა განისაზღვროს გამარჯვებული შესატყვისობების სია SU-ებსა და SP-ებს შორის, ტრანზაქციის ფასების მისაღებად მეხუთე ეტაპზე. საბოლოოდ, SU იხდის და აღწევს ვირტუალურ სამყაროს და შემდეგ ურთიერთქმედებს სხვა ვირტუალურ ობიექტებთან მეექვსე ეტაპზე.

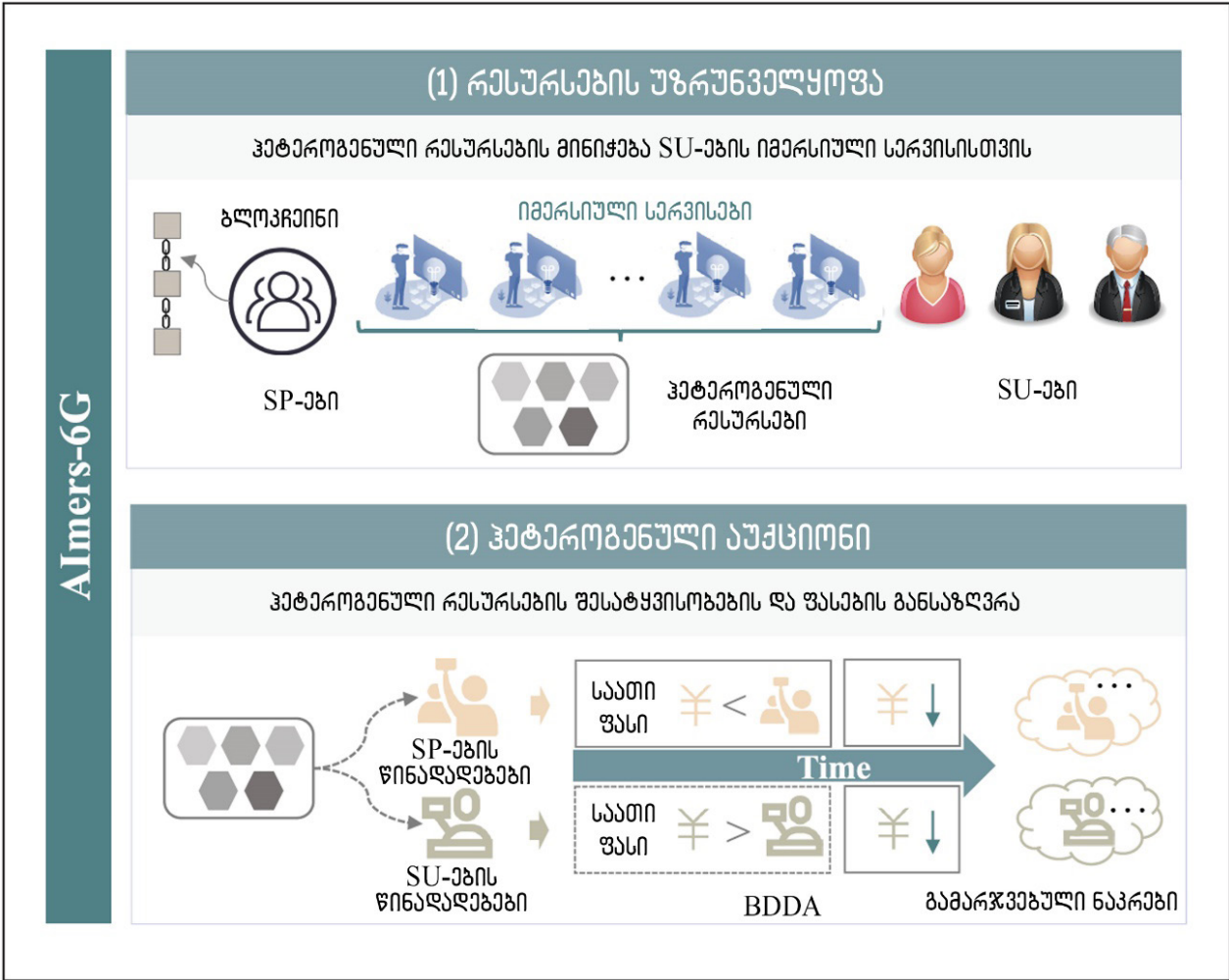
9.3. სისტემის მოდელი

ამ თავში განვიხილავთ ბლოკჩეინზე დაფუძნებულ 6G იმერსიული სერვისების ბაზარს. საკმარისი გამოთვლითი და საკომუნიკაციო რესურსების გამო, SP-ები მოქმედებენ როგორც ბლოკჩეინის კვანძები, რომლებიც ერთდროულად ასრულებენ როგორც გადმოტვირთვის ამოცანებს, ასევე ბლოკჩეინის მომსახურებას. Almers-6G-ში, შეყვანებისადმი მგრძობიარე SU ამოცანები იტვირთება 6G სერვისების რესურსების რეგიონებში. ამ ამოცანების დამუშავების შემდეგ, გამოთვლითი ჩანაწერები იტვირთება ბლოკჩეინში, რათა ჩაიწეროს განაწილებულ ლედაჯერში. შემდეგ ეს ჩანაწერები გადამოწმდება და იფუთება ახალ ბლოკში. კონსენსუსის მიღწევის შემდეგ, ახალი ბლოკი მიმაგრდება ბლოკჩეინზე.

როგორც ნახვენებია ნახ. 9.3-ზე, ეს პრობლემა შეიძლება დაიყოს ორ ნაწილად: რესურსების უზრუნველყოფა და ჰეტეროგენული აუქციონი. რესურსების უზრუნველყოფისას, 6G სერვისების რესურსების რეგიონები გამოყოფენ გამოთვლით და გატარების ბოლის რესურსებს ამოცანების რაოდენობისა და ზომის მიხედვით, შეყვანებისა და ენერჯის მოხმარების გათვალისწინებით. დავალების ჩამოყალიბების სტრატეგია და ამ პრობლემის გადაჭრის შედეგად მიღებული რესურსების სპეციფიკური განაწილება განსაზღვრავს შემდეგ საფეხურს. ჰეტეროგენულ აუქციონში ვიყენებთ BDDA მექანიზმს, რათა განვსაზღვროთ ოპტიმალური ფასები და განაწილების წესები წინა საფეხურიდან მიღებული განაწილების სტრატეგიის საფუძველზე.

საკომუნიკაციო მოდელი:

დავუშვათ, რომ ყველა SP იყენებს სიხშირულ დაყოფას მრავალჯერადი წვდომით (FDMA), სადაც თითოეული რესურს-ბლოკი ემსახურება ერთ მომხმარებელს, განსხვავებული გატარების ბოლით. არსებობს რამდენიმე ფაქტორი, რომელიც გავლენას ახდენს დაუნლინკის გატარების ბოლზე და ესენია: გამოყოფილი გატარების ბოლი, ხმაურის სიმძლავრის სპექტრული სიმკვრივე, არხის გაძლიერება, გეოგრაფიული მდგომარეობა და ა. შ. ზოგადად, ვერტიკალურ გეოგრაფიულ მანძილს აქვს უმნიშვნელო ეფექტი ჰორიზონტალურ მანძილთან შედარებით. ითვლება, რომ თითოეული ორთოგონალური რესურს-ბლოკი, რომელიც გამოყოფილია იმერსიული



ნახ. 9.3. Almers-6G-ის რეალიზაციის პრობლემის დაყოფა

სერვისისთვის, დაკავებულია მაქსიმუმ ერთი SU-ით. დაუნლინკის გადაცემის სიჩქარეზე გავლენას არ ახდენს სხვა სერვისების რესურსების რეგიონები ან SU-ები, რომლებიც მონაწილეობენ ავტორიზაციაში უსადენო ქსელის პერიფერიაზე. მაშასადამე, გადაცემის ეს დრო შეიძლება გამოითვალოს ამოცანის ზომით და დაუნლიკში გადაცემის ბიტური სიჩქარით. ენერჯის მოხმარება შეიძლება განისაზღვროს გადაცემის სიმძლავრით და გადაცემის დროით.

გამოთვლითი მოდელი:

როდესაც SU-ების ამოცანა გადმოტვირთულია სერვისების რესურსების რეგიონში, ამ რეგიონმა უნდა გამოეყოს გამოთვლითი რესურსების ნაწილი ამოცანის შესასრულებლად. დავალების დამუშავების დრო დამოკიდებულია CPU-ში ციკლების რაოდენობაზე, რომელიც საჭიროა ერთბიტიანი შესასვლელი მონაცემების დასამუშავებლად და CPU ციკლების სიხშირეზე, რომელიც საჭიროა ამოცანის დასამუშავებლად (ციკლი/წმ). CPU-ის ენერჯის გამოთვლა არის შესრულების სიხშირის სუპერწრფივი ფუნქცია, რომელიც გამოითვლება დავალების დამუშავების დროით, CPU-ის ციკლის სიხშირით და CPU-ის გადართვის ეფექტური ტევადობის საფუძველზე.

ბლოკჩეინის მოდელი:

იმერსიული 6G სერვისებით სარგებლობისთვის, SU-ები ატვირთავენ თავიანთ გამოთვლით ამოცანებს სერვისების რესურსების რეგიონებში, სადაც კონფიდენციალურობის საკითხები გადაწყვეტია. SU-ების პირადი ინფორმაცია შეიძლება დაზიანდეს არასანდო სერვისების რესურსების რეგიონების მიერ. ამიტომ, Almers-6G-ში ჩვენ ვიყენებთ სამუშაოს მტკიცებულებებზე (proof-of-stake) დაფუძნებულ ბლოკჩეინს, რათა დავრწმუნდეთ, რომ სერვისების რესურსების რეგიონები პატიოსანი და სანდოა. მიუხედავად ამისა, სხვა ტიპის მაინინგ (mining) მექანიზმები, როგორცაა ფსონის დადასტურება, ასევე შეიძლება იქნეს გამოყენებული ჩვენს მოდელში. ანუ, სერვისების რესურსების რეგიონმა, რომელმაც ახლახან დაასრულა ამოცანის დამუშავება, ჯერ უნდა გადაჭრას

ჰეშირების (hashing) პრობლემა, შემდეგ კი სერვისების რესურსის რეგიონი უფლებამოსილია შეფუთოს ეს ჩანაწერი ბლოკში. კონსენსუსის მიღწევის შემდეგ, ახალი ბლოკი მიჯაჭვულია (დამატებულია) ბლოკჩეინში. ამოცანების დამუშავებაში ბლოკჩეინის დანერგვით, უსაფრთხოება და კონფიდენციალურობა შეიძლება მნიშვნელოვნად გაუმჯობესდეს.

დავალების დასრულების შემდეგ, სერვისების რესურსების რეგიონმა უნდა უზრუნველყოს გამოთვლითი რესურსების ნაწილი ჰეშირების პრობლემის მოსაგვარებლად და, ასევე, ამოცანების დამუშავების ჩანაწერების ბლოკში შეფუთვა. ჰეშირების პრობლემის ზომის გათვალისწინებით, ჩვენ ვვარაუდობთ, რომ გამოყოფილი გამოთვლითი რესურსები განსაზღვრავს ამოცანების ამოხსნასა და შეფუთვაზე დახარჯულ დროს, დამუშავების სიძვერევეს (CPU ციკლებში/ბიტზე) და ამოცანის ზომას. ეს ამოცანა შეიცავს ჰეშირების პრობლემის ზომას, ჩანაწერის გადმოტვირთვას და დაშიფრულ მონაცემებს. გამოთვლითი მოდელის მსგავსად, ენერჯის მოხმარება გამოითვლება CPU-ის გადართვის ეფექტური ტევადობით, გამოყოფილი გამოთვლითი რესურსებით და მოხმარებული დროით.

9.4. რესურსების უზრუნველყოფა

ჩვენ ვაყალიბებთ სერვისის შეყოვნების მინიმიზაციის პრობლემას, როგორც სტოქასტური ოპტიმიზაციის ერთობლივ პრობლემას, რომელიც მოიცავს გადმოტვირთვის გადაწყვეტილებებს, გატარების ზოლის დანაწევრებას და გამოთვლითი რესურსების განაწილებას. სერვისების რესურსების რეგიონები მინიმუმამდე ამცირებენ SU-ების ლოდინის დროს NG გატარების ზოლის და გამოთვლითი რესურსების რეგულირებით, რამდენიმე შეზღუდვის ქვეშ, მათ შორის იმით რომ, თითოეული SU დაკავშირებულია სერვისების რესურსების ერთ რეგიონთან, გატარების ზოლის და გამოთვლითი რესურსების ორი კრიტიკული მნიშვნელობით. მომსახურების რესურსების რეგიონების მიერ გამოყოფილი გატარების ზოლების ჯამი არ შეიძლება აღემატებოდეს მათ საკუთრებაში არსებულ მთლიან გატარების ზოლს და იგივე სამართლიანი უნდა იყოს გამოთვლითი რესურსებისთვის. გარდა ამისა, გრძელვადიანი საშუალო ჯამური ენერჯის მოხმარება არ უნდა აღემატებოდეს ზედა ზღვარს.

ჩვენ ვიყენებთ ლიაპუნოვის თეორიას ზემოაღნიშნული სტოქასტური ოპტიმიზაციის პრობლემის გადასაჭრელად. კონკრეტულად, ჩვენ ვქმნით ვირტუალურ ენერჯეტიკულ რიგებს ენერჯიაზე შემოღებული შეზღუდვის წარმოსადგენად. სერვისების რესურსების თითოეული რეგიონისთვის, ყოველ SU-ის აქვს თავისი ვირტუალური ენერჯის რიგი. დროითი ინტერვალის (სლოტის) დასაწყისში, განახლებება ენერჯის რიგების მნიშვნელობები, რომლებიც გამოითვლება რიგის სიგრძით წინა დროით ინტერვალზე, პლუს დროის იმავე მონაკვეთში გენერირებული ენერჯის მთლიანი მოხმარება და გამოკლებული დროის მიხედვით გასაშუალოებული ენერჯის ზედა ზღვარი (რიგის სიგრძე არასოდეს შემცირდება 0-ის ქვემოთ). როდესაც რიგის მნიშვნელობები ერთმანეთს ემთხვევა, ენერჯის მოხმარების შეზღუდვა დაკმაყოფილებულია.

ლიაპუნოვის ოპტიმიზაციიდან გამომდინარე, არსებობს სიდიდის „დრიფტი პლუს ჯარიმა“ ზედა ზღვარი. ჩვენ ვიღებთ ზღვრის ზედა მნიშვნელობას და ახალ ობიექტურ ფუნქციას ფორმულის დამტკიცების მეშვეობით. ეს ობიექტური ფუნქცია შეიცავს შეზღუდვებს ენერჯის მოხმარებაზე და ვირტუალური ენერჯის რიგების ზომაზე. ამ პრობლემის გადაჭრით, გრძელვადიანი შეზღუდვები შეიძლება გარდაიქმნას რიგის სტაბილურობის შეზღუდვებად, და ამავე დროს ხდება სერვისის შეყოვნების ოპტიმიზაცია.

ისეთი ტრადიციული მიდგომები, როგორებიცაა ევრისტიკა და ორობითი ცვლადები, იძლევა სუბოპტიმალურ, და არაოპტიმალურ შედეგებს და შესაბამისი მახასიათებლების გაუარესებას. მიმდინარე სიტუაციის კონტექსტის სრულად გათვალისწინება და იმერსიული გადაწყვეტილებების მიღება არის პერსპექტიული ალტერნატივა, რომელსაც ამ თავში ეწოდება კონტექსტურ-იმერსიული სწავლება. ზემოაღნიშნული ლიაპუნოვის ოპტიმიზაციის დროითი მახასიათებლების გათვალისწინებით, ინტუციურ LSTM-ზე დაფუძნებული აქტორ-კრიტიკოსის ალგორითმი გამოიყენება კონტექსტურ-იმერსიულ სწავლებაში. ის გამორიცხავს გამოცდილებიდან სწავლების გამოთვლით სირთულეს რეალურ დროში, ტრენინგის მონაცემების ნიმუშების ხელით ეტიკეტირების საჭიროების გარეშე.

ნაწილობრივი სტრატეგიების გენერაცია (აქტორი):

ყოველი დროითი ინტერვალის დასაწყისში მონაწილეები იყენებენ LSTM ქსელს და მოქმედების გენერატორს, რათა შექმნან კონტექსტით გაცნობიერებული, იმერსიული სერვისის დაგეგმვის გადაწყვეტილებები. ჩვენ განვიხილავთ არხის გაძლიერების კოეფიციენტს თითოეული რეგიონის მომსახურე რესურსისთვის და SU-ის

ენერგეტიკული რიგის სიგრძეს, როგორც შემავალ მონაცემებს. ჩვენ ვიღებთ ალბათობას, რომ SU-მ განუსაზღვროს დავალებები სერვისების რესურსების სხვადასხვა რეგიონს სიგმოიდის (sigmoid) დაყენებით, როგორც აქტივაციის ფუნქცია. ამის შემდეგ, ალბათობები რაოდენობრივად ფასდება სერვისის დაგეგმვის სტრატეგიების მიხედვით. კონკრეტულად, ჩვენ ვაწარმოებთ სერვისის დაგეგმვის პოლიტიკას მაქსიმალური ალბათობის მიხედვით. ამავდროულად, ჩვენ ვქმნით დამატებით დაგეგმვის პოლიტიკას ერთი SU-ისთვის ყველაზე მაღალი ალბათობის არჩევით და სხვების შემთხვევით არჩევით.

რესურსების უზრუნველყოფის ოპტიმიზაცია (კრიტიკოსი):

კონტექსტური დისპეტჩერიზაციის სტრატეგიებიდან გამომდინარე, გადაწყვეტილება გამოთვლითი რესურსების ზომაზე და გატარების ზოლზე შეიძლება ჩამოყალიბდეს არაწრფივი პროგრამირების პრობლემად. რესურსების საუკეთესო უზრუნველყოფის მისაღწევად, ამ პრობლემას ვავაგრებთ ოპტიმიზაციის ხელსაწყო გამოყენებით, ანუ გურობის (Gurobi) ოპტიმიზაციის საშუალებით. გურობი არის სრულფასოვანი ამომხსნელი, ასევე კომერციული ოპტიმიზაციის პროგრამული ინსტრუმენტი. ეს ამომხსნელი იყენებს განშტოებისა და შემოსაზღვრის (branch-and-bound) მეთოდს ოპტიმიზაციის პრობლემის გადასაჭრელად. ამომხსნელი არ დაასრულებს პროცესს მანამ, სანამ ქვედა და ზედა საზღვრებს შორის განსხვავება არ იქნება $1e-10$ -ზე ნაკლები (განსხვავების აბსოლუტური ზღვარი) ან $1e-4$ -ზე ნაკლები (განსხვავების ფარდობითი ზღვარი), გამრავლებული ზედა ზღვრის აბსოლუტურ მნიშვნელობაზე. აქედან გამომდინარე, გადაწყვეტის გლობალური ოპტიმალურობა დამოკიდებულია გონივრულ სიზუსტეზე. ამის შემდეგ, თითოეული დისპეტჩერიზაციის სტრატეგიისა და შესაბამისი რესურსის უზრუნველყოფისთვის, ჩვენ ვიანგარიშებთ მათ მიზნობრივ ფუნქციას და ვირჩევთ მინიმალურს.

რიგის და ქსელის განახლება:

ზემოთ მოყვანილ კრიტიკოსის მოდულში მიღებული ოპტიმალური დისპეტჩერიზაციის სტრატეგიისა და რესურსების განაწილების მიხედვით, ჩვენ ვიღებთ ყველა SU-ის ენერჯის მოხმარებას და ვაახლებთ ენერჯის რიგებს. განახლებული ენერგეტიკული რიგები გამოყენებული იქნება შემავალი მონაცემების სახით მომდევნო დროით ინტერვალში. კონტექსტურ-იმერსიული სწავლება იყენებს რიგის სიგრძეებს, არხის გაძლიერების კოეფიციენტებს და არჩეულ განაწილების სტრატეგიებს, როგორც ეტიკეტირებულ შესასვლელ-გამოსასვლელ ნიმუშებს მოქმედი ნეირონული ქსელების პოლიტიკის განახლებისთვის. ჩვენ ვიყენებთ SGD ალგორითმს საშუალო ჯვარდინი ენტროპიის კარგვის ფუნქციის შესამცირებლად და კონტექსტური მოდულის პარამეტრების გასაახლებლად.

9.5. ჰიბრიდული აუქციონი

NG იმერსიული სერვისების ბაზარი არის ორმხრივი ბაზარი — SU-ებით როგორც მყიდველებით და SP-ებით როგორც გამყიდველებით. უსადენო აპლიკაციებს, როგორცაა ვირტუალური კონცერტები, ვირტუალური შეხვედრები, ვირტუალური ტრენინგი და ა. შ. აქვთ უკიდურესად მკაცრი მოთხოვნები შეყვანებისა და გატარების ზოლის მიმართ. უწყვეტი ვაჭრობის ბაზრებისგან განსხვავებით, NG იმერსიული სერვისების ბაზარი უნდა იყოს ერთგვარი „გამოძახების ბაზარი“ (call market), რომელშიც SP და SU მყისიერად ირთებიან და ასრულებენ იმერსიული სერვისებით ვაჭრობას.

SU-ების იმერსიული სერვისის საჭიროებების დასაკმაყოფილებლად, სერვისების რესურსების რეგიონებმა ერთობლივად უნდა განიხილონ გატარების ზოლის და გამოთვლითი რესურსების გამოყენების შესაძლებლობები. მაგალითად, გატარების ზოლის რესურსების ღირებულება დამოკიდებულია სიხშირული სპექტრის დიაპაზონებზე, რომლებიც იცვლება დროში ცვალებადი მოთხოვნების მიხედვით. გამოთვლითი რესურსების ღირებულება დამოკიდებულია სერვისების რესურსების რეგიონების CPU სიჩქარეებზე (სიხშირეებზე) და CPU ჩიპების გამოთვლითი ენერგოეფექტიანობის კოეფიციენტზე. ამრიგად, სერვისების რესურსების რეგიონებს აქვთ განსხვავებული შეფასება მოხმარებლის სხვადასხვა მოთხოვნისთვის. აუქციონში მონაწილეობის მიღება და საკუთარი წინადადებების წარმოდგენა შეუძლია ერთი და იმავე სერვისების რესურსების რეგიონის ყველა SP-ს.

SU-ებისთვის შემოთავაზებული შესყიდვის ფასი მჭიდრო კავშირშია მომსახურების ხარისხთან, რომლის მიღებაც სურთ. ისინი აფასებენ თავიანთი ალქმის გამოცდილების ხარისხს მყისიერი სტრუქტურული მსგავსების (SSIM) და ვიდეო მრავალმეთოდური შეფასების შერწყმის (VMAF) მეშვეობით. VMAF არის Netflix-ის მიერ

შემუშავებული აღქმის ვიდეო ხარისხის შეფასების ალგორითმი, რომელიც ცდილობს ასახოს მაყურებლის აღქმა ნაკადის ხარისხის შესახებ. SSIM ადარებს პიქსელების ინტენსივობის ადგილობრივ ნიმუშებს, რომლებიც ნორმალიზებულია სიკაშკაშისა და კონტრასტის მიხედვით. მათ მნიშვნელობებზე გავლენას ახდენს ბრუნვის საშუალო სიჩქარე, ისევე როგორც ბიტური სიჩქარე. წინა სამუშაოების მიხედვით ადამიანის ჯანმრთელობაზე გემოქმედებით განპირობებული შეზღუდვისა და ძალადობრივი ონლაინ თამაშების გამო, ბრუნვის საშუალო სიჩქარე წამში 360°-დან 1800°-მდეა. ამასთან, ბიტური სიჩქარის ზომა არის ამ პარამეტრის საჭირო მნიშვნელობა, რომელიც საჭიროა მისი აპლიკაციების მხარდასაჭერად. ბრუნვის საშუალო სიჩქარისა და მოსალოდნელი ბიტური სიჩქარის მიხედვით, SU-ებს შეუძლიათ საკუთარი შეფასების პოვნა SSIM-ის და VMAF-ის გამოთვლით.

6G იმერსიული სერვისების ბაზარზე მრავალი SU-ისა და SP-ის არსებობის გათვალისწინებით, აუქციონისთვის ვიყენებთ ორმხრივ გამოძახების ბაზარს (double-side call market). BDDA-ის მიზნებია SP-სა და SU-ის შედარება და შესყიდვისა და გაყიდვის ფასების განსაზღვრა. ნახ. 9.4 გვიჩვენებს გამოყენებული საათობრივი აუქციონის (clock auction) წესებს, მათ შორის ბლოკჩეინზე დაფუძნებულ აუქციონერს, SU-ის მხარეს და SP-ის მხარეს.

სანამ შემდეგ მსჯელობებზე გადავიდოდეთ მოვიყვანოთ ერთი განმარტება:

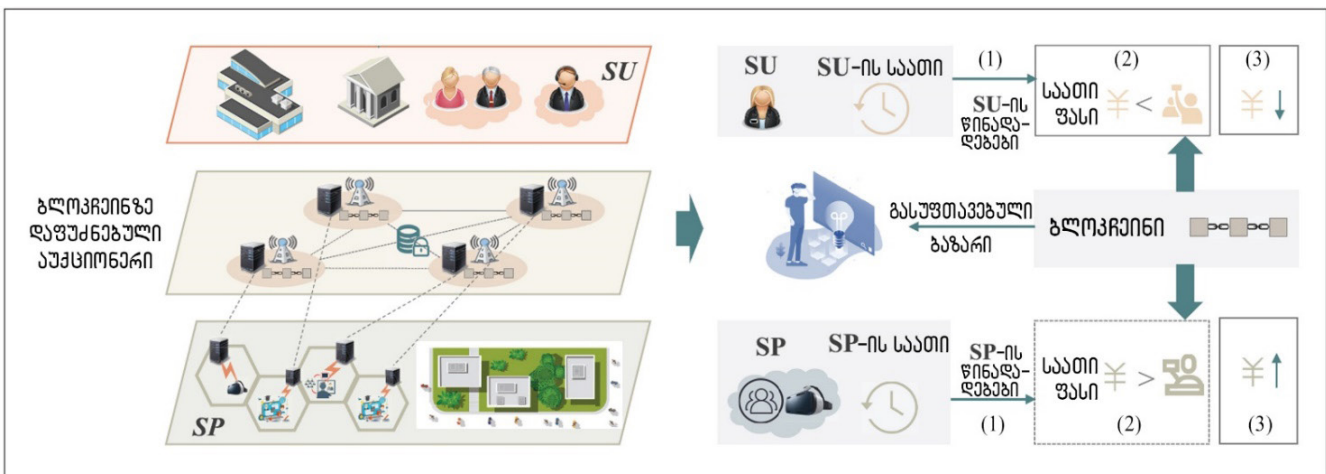
საათობრივი აუქციონები არის ვაჭრობის მექანიზმის ტიპი, რომელიც გამოიყენება სხვადასხვა კონტექსტში, მაგალითად, სატელეკომუნიკაციო სექტორის განაწილებაში. საათობრივ აუქციონზე ფასების მნიშვნელობები (მათ უწოდებენ საათის ფასებს, clock prices) იწყება მაღლიდან და იკლებს მომდევნო რაუნდებში, სანამ პრეტენდენტები არ გამოაცხადებენ თავიანთ მზადყოფნას მიიღონ მიმდინარე ფასი.

ბლოკჩეინზე დაფუძნებული აუქციონერის მხარე:

აუქციონერი ერთდროულად ინახავს პოლანდიურ აუქციონთან დაკავშირებულ ორი საათის ფასს SU და SP მხარისთვის. SU-ის საათის ფასი აჩვენებს შესყიდვის ფასს და ის იხსნება წინა გამოცდილების საფუძველზე დაყენებული უმაღლესი ფასით. მოცემული სერვისების რესურსების რეგიონისთვის, SP-ის საათის ფასი აჩვენებს გასაყიდ ფასს. გარდა ამისა, აუქციონერმა ასევე უნდა აირჩიოს ბიჯების ზომა საათის ფასის დასარეგულირებლად. აუქციონერი ჩვეულებრივ იწყება მყიდველის რაუნდით.

SU-ის მხარე:

აუქციონი ჩვეულებრივ იწყება SU-ის მხრიდან. ქვემოთ მოცემულია კონკრეტული ბიჯები, რომლებიც შეესაბამება ნახ. 9.4-ის ზედა მარჯვენა კუთხეში არსებულ ელემენტებს. პირველი, აუქციონერი ატყობინებს SU-ის საათის ფასს, თუ SU აკმაყოფილებს შეთავაზების კრიტერიუმებს, ანუ მოსალოდნელი ფასი უფრო მაღალია ვიდრე საათის ფასი. SU შეირჩევა გამარჯვებულად SU-ის მხრიდან და წარადგენს შესყიდვის წინადადებას მყიდველის მიმდინარე საათის ფასად. მეორე, ჩვენ შევადარებთ SU-ის წინადადებებს და საათის ფასებს. გასათვალისწინებელია, რომ რამდენიმე SU შეიძლება იყოს კვალიფიცირებული, მაგრამ შეირჩევა მხოლოდ ერთი. მესამე, თუ SU-ის საათის ფასი ძალიან მაღალია და არ არის არჩეული შესაბამისი SU, SU-ის საათის ფასი სათანადოდ შემცირდება მანამ, სანამ არ მოიძებნება შესაბამისი SU. მაშინ გამარჯვებული ამოღებულ იქნება ამჟამინდელი SU-ების ნაკრებიდან.



ნახ. 9.4. ჰეტეროგენული აუქციონის მექანიზმი

SP-ის მხარე:

უფლებამოსილი SU-ის არჩევის შემდეგ, აუქციონის ჩატარების ჯერი დგება SP-ისთვის. ანალოგიურად, კონკრეტული ბიჯები შეესაბამება ნახ. 9.4-ის ქვედა მარჯვენა მხარეს არსებულ პუნქტებს: პირველი, ის SP-ები, რომლებსაც ჯერ არ უვაჭრიათ, წარადგენენ თავიანთ წინადადებებს აუქციონერისგან SU-ის საათის ფასის მიღების შემდეგ. მეორე, ჩვენ შევადარებთ SP-ის წინადადებებს და საათის ფასებს. თუ SP აკმაყოფილებს ტრანზაქციაში მონაწილეობის პირობებს, ანუ მოსალოდნელი გასაყიდი ფასი საათის ფასზე დაბალია, ეს SP შეირჩევა გამარჯვებულად SP მხარეს და მიიღებს მის გასაყიდ წინადადებას, როგორც მიმდინარე SP-ის საათის ფასს. მესამე, SU-ის მხარის მსგავსად, შეირჩევა მხოლოდ ერთი SP. თუ ბაზარზე არ არის SP-ები, რომლებსაც სურთ ვაჭრობა, SP-ის საათის ამჟამინდელი ფასი უნდა გაიზარდოს. შემდეგ დგება SP-ის ჯერი. აუქციონი მთავრდება, როდესაც SP-ის საათის ფასი უფრო მაღალია, ვიდრე SU-ის საათის ფასი.

გამარჯვებული მყიდველი და გამყიდველი შესატყვისია (დაკავშირებულია) და ამოღებულია აუქციონიდან, ხოლო ტრანზაქციის ფასი დგინდება SU-ის საათის ფასისა და SP-ის საათის ფასის გადაკვეთის საერთო წერტილის საფუძველზე. ჩვეულებრივ SU-ის საათის და SP-ის საათის საწყის ფასებს ადგენს აუქციონერი თავიდანვე. სოციალური კეთილდღეობა განისაზღვრება, როგორც წარმატებული SU-ებისა და SP-ების წინადადებების ჯამი. ჩვენს ალგორითმში, რესურსების უზრუნველყოფის პრობლემის გადაჭრისას, სერვისების რესურსების რეგიონი იღებს ყველა იმ SU-ის შეფასებას, რომელსაც ის გაუწევს მომსახურებას. აუქციონის დაწყებამდე SP-ის საათის საწყისი ფასი დგინდება, როგორც გასაშუალოებული შეფასება.

აუქციონის ბაზარზე საკრედიტო კრიზისის შემთხვევაში შემუშავებულია ჰკვიანი კონტრაქტი, რომელიც უზრუნველყოფს ავტომატურ ვაჭრობას. კონტრაქტის შემსრულებელი რეგისტრირებულია როგორც სავაჭრო სისტემის აუქციონერი. მყიდველები და გამყიდველები არეგისტრირებენ თავიანთ ვინაობას და წინადადებებს სარეგისტრაციო მოდულის მეშვეობით. გამყიდველებს მოეთხოვებათ უზრუნველყონ როგორც ხელმისაწვდომი გატარების ზოლის, ასევე გამოთვლითი რესურსები. როდესაც ვაჭრობა იწყება, აუქციონერი იღებს შეთავაზებებს ყველა მყიდველისგან და ადგენს გამყიდველისა და მყიდველის საათის ფასებს. მყიდველები და გამყიდველები კავშირდებიან სავაჭრო მოდულში დაყენებული ფუნქციის მიხედვით. საათის არასათანადო ფასის შემთხვევაში, რომელიც ხელს უშლის აუქციონის ჩატარებას, აუქციონერი ასრულებს საათის ფასის შეცვლის ფუნქციას (შეცვლილი ფასი სულაც არ არის ერთნაირი ყოველ ჯერზე). მას შემდეგ რაც განხილული ორი საათის ფასი გადაიკვეთება, ტრანზაქციის მოდული შეწყდება. გადაცემის მოდულში ყველა გამარჯვებული გამყიდველი გადასცემს ადრე გამოთვლილ გატარების ზოლს და გამოთვლით რესურსებს.

9.6. კომპიუტერული სიმულაციის შედეგები

მოცემულ პარამეტრებში ჩვენ ვაფასებთ წარმოდგენილი სქემის მახასიათებლებს. განიხილება სცენარი ათი SU-ით და სერვისების რესურსების სამი რეგიონით. SU-ების ამოცანების ზომა ეფუძნება რეალურად არსებულ პერიფერიაზე ფუნქციონირებადი ღრუბლოვანი კომპანიის მონაცემთა ნაკრებს. ჩვენი მონაცემები ასახავს 10 მილიონზე მეტი მომხმარებლის მოთხოვნას 3.51 გბაიტზე მეტი მოცულობის მონაცემთა ჩანაწერით. დროითი ინტერვალი დაყენებულია 1000-ზე, გადაცემის სიმძლავრე მერყეობს 10-დან 20 დბმ-მდე, არხის გაძლიერების კოეფიციენტი მიღებულია გზაში დანაკარგების (pathloss) მოდელიდან და იღებს საბოლოო მნიშვნელობას 0-დან 1-მდე. ასევე, ზოგადად დამუშავების სიმკვრივე და ბლოკისთვის დამუშავების სიმკვრივე დაყენებულია, შესაბამისად, სიდიდეებზე 279.62 ციკლი/ბიტი და 737.5 ციკლი/ბიტი. CPU-ის გადართვის ეფექტური ტევადობა დაყენებულია 10^{-27} -ზე. მაქსიმალური გატარების ზოლი სამი სერვისის რესურსების რეგიონისთვის დაყენებულია, როგორც 0.1 მპც, 0.15 მპც და 0.2 მპც, შესაბამისად, და მაქსიმალური გამოთვლითი რესურსები სამი სერვისის რესურსების რეგიონისთვის არის 5 გპც, 4 გპც და 3 გპც, შესაბამისად. SU-ების ბრუნვის საშუალო სიჩქარე მერყეობს წამში 360°-დან 1800°-მდე. SU აპლიკაციების მხარდასაჭერად საჭირო ბიტური სიჩქარე მერყეობს 15 მბიტი/წმ-დან 30 მბიტი/წმ-მდე.

ჩვენს შეფასებაში საბაზისო მეთოდები მოიცავს:

- რიგის პრიორიტეტის (QP) მეთოდს, რომელიც ითვალისწინებს მხოლოდ ენერჯის რიგის სიგრძეს. ყოველ დროით ინტერვალში, SU-ის ამოცანები გადაიტვირთება სერვისების რესურსების რეგიონებში უმოკლესი რიგის სიგრძით.

- შემთხვევით მეთოდს, რომლის მიხედვითაც SU-ის ამოცანები გადაიტვირთება სერვისების რესურსების შემთხვევით არჩეულ რეგიონებში.

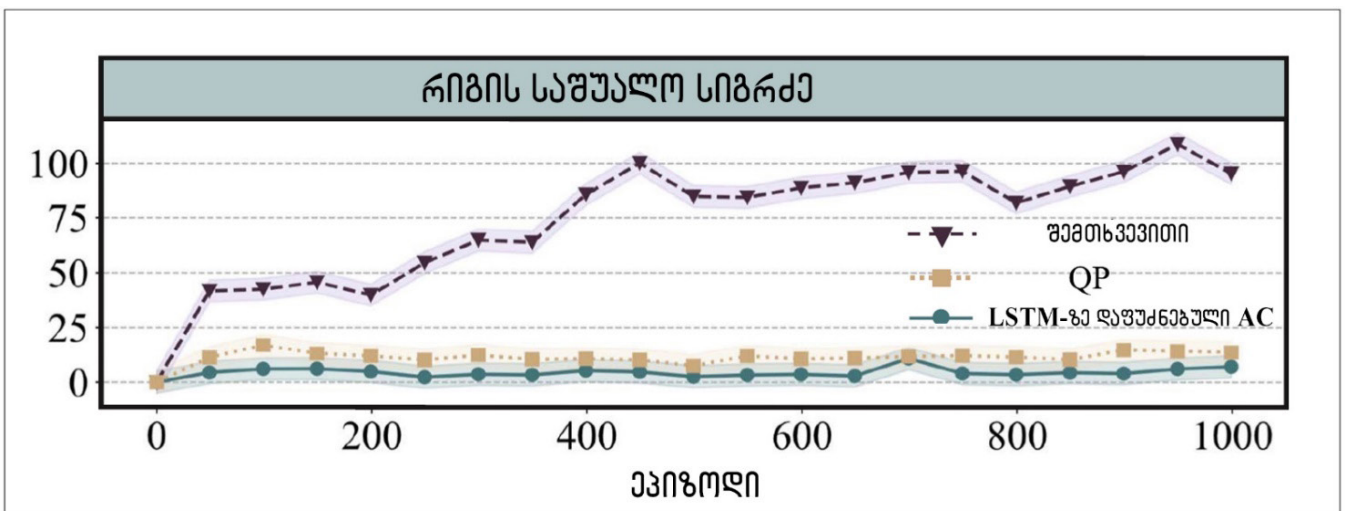
გარდა ამისა, ვაჭრობის ძირითადი მეთოდები მოიცავს:

- ორმაგი აუქციონის მეთოდს (DA), სადაც მყიდველისა და გამყიდველის ფასები ჩამოთვლილია თანმიმდევრობით, ამასთან მყიდველები შეესაბამებიან პრინციპს „მალიდან დაბლა“ და გამყიდველები შეესაბამებიან პრინციპს „დაბლიდან მაღლა“.
- ორმაგი ჰოლანდიური აუქციონის მეთოდს (DDA), სადაც გამყიდველის საათის ფასი დადგინდება აუქციონერის გამოცდილებიდან გამომდინარე, და არა მყიდველის სავარაუდო საშუალო ფასის საფუძველზე. აქ ჩვენ დავაყენეთ საწყისი საათის ფასები როგორც მყიდველებისთვის, ისევე გამყიდველებისთვის ფიქსირებულ მნიშვნელობაზე.

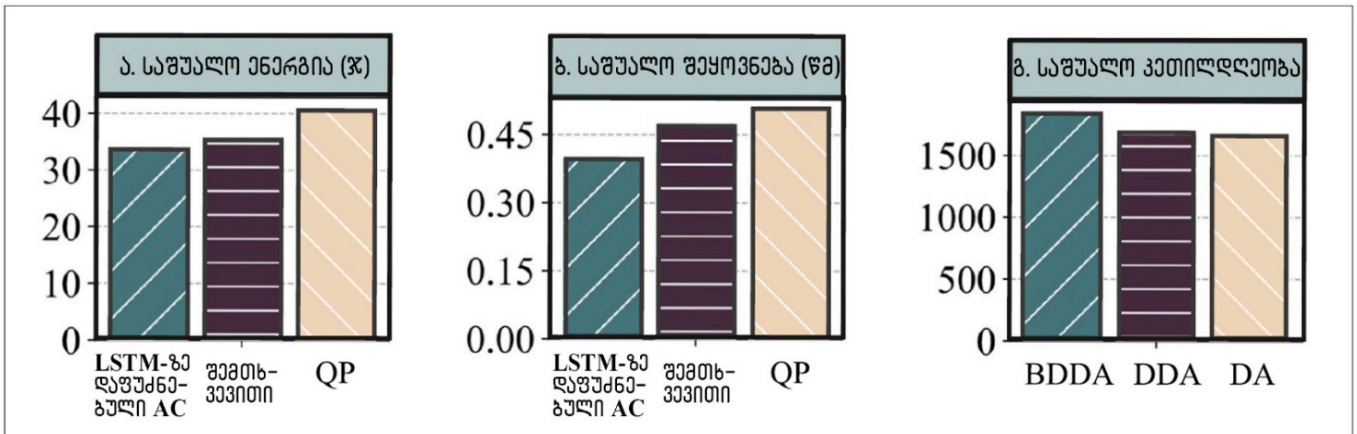
ნახ. 9.5 გვიჩვენებს კონვერგენციისა და რიგის საშუალო სიგრძის მახასიათებლებს. გამომდინარე იქიდან, რომ ჩვენი შემოთავაზებული სქემა ითვალისწინებს რიგის დროით ინფორმაციას, SU-ების მეტი მოთხოვნა შეიძლება გაიგზავნოს ოპტიმალურ რესურს-რეგიონებში, რითაც მნიშვნელოვნად შემცირდება ენერჯის რიგის სიგრძე. რიგის მოკლე სიგრძე საშუალებას იძლევა სწრაფად უპასუხოთ SU-ების მოთხოვნებს, რაც კიდევ უფრო აუმჯობესებს იმერსიული სერვისის ხარისხს.

ნახ. 9.6ა და 9.6ბ გვიჩვენებს, რომ ჩვენი შემოთავაზებული სქემა აღწევს სერვისის საშუალო ენერჯის ყველაზე დაბალ მოხმარებას და შეყოვნებას, რომელიც აუმჯობესებს QP-ის და შემთხვევით მეთოდს, მთლიანი შეყოვნებით 16.20 %-ით და 36.56 %-ით, ხოლო სერვისში ენერჯის მოხმარებით 4.14 %-ით და 17.77 %-ით. მიზეზი ის არის, რომ QP ითვალისწინებს მხოლოდ რიგის სიგრძეს, და არა დროით მახასიათებლებს. მიუხედავად იმისა, რომ QP იძლევა გარანტიას, რომ საშუალო სიგრძე სწრაფად არ გაიზრდება, ის არ იძლევა გარანტიას, რომ ვაკეთდება ოპტიმალური არჩევანი, ან რომ სერვისის მიერ გენერირებული ენერჯის მოხმარება იქნება მინიმალური. შემთხვევითი მეთოდი იწვევს რიგის არასტაბილურობას, როგორც ეს ნაჩვენებია ნახ. 9.5-ზე მოყვანილ კონვერგენციის მახასიათებელში.

ნახ. 9.6გ გვიჩვენებს შემოთავაზებული აუქციონის სქემის სოციალურ კეთილდღეობას, რომელიც განსაზღვრება როგორც გამარჯვებული მყიდველებისა და გამყიდველების ჯამი. ჩვენი მექანიზმი აჭარბებს DA-ის და DDA-ის შესაბამისად 11.02 %-ით და 10.53 %-ით. მიზეზი არის ის, რომ ბიჯების ზომის სათანადო დაყენებით, გამყიდველსაც და მყიდველსაც შეეძლება აუქციონზე შესაბამისი ფასის მიღება. ასევე, გამყიდველის საათის ფასად საშუალო ფასის გამოყენებამ შეიძლება უფრო დიდი სარგებელი მოუტანოს გამყიდველებს. DA-ის არ შეუძლია შეცვალოს ორივე მხარის შეთავაზებები, ამიტომ გამყიდველის დაბალი ფასი იწვევს უფრო მცირე სოციალურ კეთილდღეობას. DDA-მ ვერ დაასახელა გამყიდველის საათის ფასის სწორი მნიშვნელობა, რაც



ნახ. 9.5. კონვერგენციის მაჩვენებელი და რიგის საშუალო სიგრძის მახასიათებელი



ნახ. 9.6. საშუალო ენერჯიის, საშუალო შეყვანების და საშუალო კეთილდღეობის მასხაითებლების შედარება

იმას ნიშნავს, რომ მაღალი ფასი იწვევს აუქციონების სწრაფად დასრულებას, ხოლო დაბალი ფასი გამყიდველს ნაკლებად მომგებიანს ხდის. გარკვეულ დროით ინტერვალებში, DA-ის და DDA-ის სოციალური კეთილდღეობა შეიძლება იყოს უფრო მაღალი, ვიდრე ამ თავში შემოთავაზებული სქემაში, თუმცა, 1000 დროითი ინტერვალის ჯამის გასაშუალოებით, ჩვენ ვიღებთ უკეთეს სოციალურ კეთილდღეობას. საერთო ჯამში, განხილული მეთოდი აღწევს ენერჯიის ნაკლებ მოხმარებას, დაბალ შეყვანებას და უკეთეს სოციალურ კეთილდღეობას, რაც უკეთესი არჩევანია იმერსიული მომსახურების ბაზრისთვის.

9.7. მეცხრე თავის დასკვნა და სამომავლო კვლევითი საგუშაო

ამ თავში განხილული იყო Almers-6G მიდგომა 6G იმერსიული სერვისებისთვის გატარების ზოლისა და გამოთვლითი რესურსების განაწილების, ასევე ბაზრის აუქციონის პრობლემის ერთობლივად გადასაჭრელად. ასევე, წარმოდგენილია კონტექსტურ-იმერსიული სწავლების მეთოდი გადმოტვირთვის გადაწყვეტილების მისაღებად და კომუნიკაციისა და გამოთვლითი რესურსების გამოყოფისთვის განაწილებულ დიდ რეგიონებში. აღნიშნულის საფუძველზე, ჩვენ განვიხილავთ ბლოკჩეინზე დაფუძნებულ ორმაგ პოლანდიურ აუქციონს, რათა მაქსიმალურად გავზარდოთ სოციალური კეთილდღეობა პატარა რეგიონში. კომპიუტერული სიმულაციის შედეგები აჩვენებს, რომ შემოთავაზებული სქემა აღწევს ოპტიმალურ შეყვანებას, ენერჯიის მოხმარებას და სოციალურ კეთილდღეობას. ვფიქრობთ, რომ სამომავლო სამუშაოებში უნდა მოხდეს რესურსების განაწილების გადაწყვეტილების ოპტიმიზაციისა და პრაქტიკული გამოყენების საკითხების გამოკვლევა რეალური სისტემის ასაშენებლად.

ბოლოთქმა

ნავარაუდებია, რომ IMT-2030-ით უზრუნველყოფილი აპლიკაციები და სერვისები დააკავშირებს ადამიანებს, მანქანებსა და სხვა ობიექტებს. ადამიანი-მანქანის ინტერფეისების, ინტერაქტიული და მაღალი გარჩევადობის ვიდეო სისტემების, როგორცაა გაფართოებული რეალობის დისპლეები, ტაქტილური სენსორებისა და აქტივატორების და მრავალსენსორული ინტერფეისების სფეროებში ტექნოლოგიური მიღწევების გამოყენებით, მოსალოდნელია, რომ IMT-2030 შესთავაზებს ადამიანებს იმერსიულ გამოცდილებას, რომელიც წარმოიქმნება ვირტუალურად ან ხორციელდება დისტანციურად. მეორე მხრივ, მანქანები უნდა იყოს ინტელექტუალური, ავტონომიური, საიმედო და ზუსტი, მანქანური აღქმის მიღწევების, მანქანების ურთიერთქმედების შესაძლებლობის ფარგლებში და ხელოვნური ინტელექტის (AI) ეფექტიანი მართვის დემონსტრირებით. ფიზიკურ სამყაროში ადამიანები და მანქანები მუდმივად ურთიერთობენ ერთმანეთთან, მუშაობენ ციფრულ სამყაროსთან, რომელიც აფართოებს რეალურ სამყაროს დიდი რაოდენობით მოწინავე სენსორებისა და ხელოვნური ინტელექტის გამოყენებით. ასეთი ციფრული სამყარო არა მხოლოდ აკოპირებს რეალურ სამყაროს, არამედ გავლენას ახდენს მასზე, აძლევს ადამიანებს ვირტუალურ გამოცდილებას და მანქანებს გამოთვლით და მართვით შესაძლებლობებს.

მომავალი გამოყენების შემთხვევები მხარს დაუჭერს განაწილებულ გამოთვლით და AI აპლიკაციებს. ტიპური გამოყენების შემთხვევები მოიცავს IMT-2030-ის დახმარებით ავტომატიზებულ მართვას, ავტონომიურ თანამშრომლობას მოწყობილობებს შორის ჯანდაცვის აპლიკაციებისთვის, რთული და ხანგრძლივი გამოთვლითი ოპერაციების გადატვირთვას პერიფერიულ მოწყობილობებსა და ქსელებში, ციფრული ტყუპების შექმნას და სხვა. ეს სცენარები მოითხოვს ტრაფიკის მაღალი გამტარუნარიანობისა და მონაცემთა გადაცემის გაზრდილი სიჩქარის მხარდაჭერას, ასევე ულტრა დაბალ შეყოვნებას და ულტრა მაღალ საიმედოობას. კონკრეტული გამოყენების შემთხვევიდან გამომდინარე, კომუნიკაციის ასპექტების გარდა, მოსალოდნელია, რომ გამოყენების სცენარები მოიცავს ახალ შესაძლებლობებს, რომლებიც დაკავშირებულია AI-ისა და გამოთვლითი ფუნქციების ინტეგრაციასთან IMT-2030-ში, მათ შორის სხვადასხვა წყაროდან მონაცემთა შეგროვებას, მომზადებას და დამუშავებას, განაწილებული AI მოდელის ტრენინგს, მოდელის გაზიარებას და განაწილებას, და გამოთვლითი რესურსების ორკესტრირებას და გააერთიანებას. AI-ისთან დაკავშირებული მოქმედი შესაძლებლობები ეხება IMT-2030-ის ფარგლებში სპეციფიკური ფუნქციების უზრუნველყოფის შესაძლებლობას AI-ზე ორიენტირებული აპლიკაციების მხარდასაჭერად. ეს ფუნქციები მოიცავს მონაცემთა განაწილებულ დამუშავებას, განაწილებულ სწავლებას, AI გამოთვლებს, AI მოდელის რეალიზაციას, AI მოდელის დასკვნას და ა. შ.

ეს წიგნი შეისწავლის ახალ ტექნოლოგიებსა და კვლევით პერსპექტივებს AI-ით მხარდაჭერილ 6G მობილური საკომუნიკაციო სისტემებისთვის, ასევე ძირითად ტექნიკურ გამოწვევებს, რომელიც წარმოდგენილია უახლეს პუბლიკაციებზე და სამეცნიერო მიღწევებზე დაყრდნობით. წიგნში მოცემული კვლევები დაგეგმარება მრავალი სატელეკომუნიკაციო პრობლემის გადაჭრის პროცესში, როგორცაა, მაგალითად, განზოგადებული/გაუმჯობესებული ინდექსი მოდულაციის (IM) სქემების შემუშავება უფრო მაღალი სპექტრული ან/და ენერგოეფექტიანობით, ასევე გადამცემი/მიმღები სისტემის დაბალი სირთულით და შეცდომების უკეთესი მახასიათებლებით. წიგნის შედეგებზე დაყრდნობით შესაძლებელი გახდება IM ტექნიკის ინტეგრირება თანამედროვე MIMO სისტემებში, რომლებიც გამოყენებული იქნება 6G უსადენო ქსელებში და IM ტექნოლოგიების პოტენციალის გამოკვლევა პრაქტიკული განხორციელების სცენარების მიხედვით.

ვფიქრობთ, რომ მასალა სასარგებლო იქნება და დახმარებას გაუწევს საინფორმაციო ტექნოლოგიებისა და კომუნიკაციების დარგში მომუშავე სპეციალისტებს და აკადემიურ პერსონალს, ბაკალავრიატის მაღალი კურსის სტუდენტებს, მაგისტრანტებს, დოქტორანტებს და მათ უდავოდ ახლოს მიიყვანს ამ მიმართულებით არსებულ სამეცნიერო-ტექნოლოგიურ მიღწევებთან და პრობლემატიკასთან.

ბევრი მცდელობის მიუხედავად, გვერდი ვერაფრით ავუარეთ უამრავ გადაუჭრელ ტერმინოლოგიურ პრობლემას, რომელთა დაძლევა ჩვენთვის შეუძლებელი იყო; ამასთან დაკავშირებით მკითხველს ბოლოში გვინდა მოუხადოთ.

შეუძლებელია მაღლიერებით არ მოვიხსენიოთ კავკასიის უნივერსიტეტისა და საქართველოს უნივერსიტეტის მხარდაჭერა. ასევე, მადლობა გვინდა გადავუხადოთ აკადემიურ დოქტორ მამუკა ჩხაიძეს, წიგნის მომზადებისას გაწეული დახმარებისთვის.

სერგო შავგულიძე, ნოდარ უდრელიძე
თბილისი, ივნისი, 2024 წელი

ლიტერატურა

პირველ თავში გამოყენებული ლიტერატურა

- A.-A. A. Boulogeorgos *et al.*, "Machine learning: A catalyst for THz wireless networks," *Frontiers Commun. Netw.*, vol. 2, Sep. 2021, p. 37.
- M. M. Butt, A. Pantelidou, and I. Z. Kovács, "ML-assisted UE positioning: Performance analysis and 5G architecture enhancements," *IEEE Open J. Veh. Technol.*, vol. 2, Sep. 2021, pp. 377-388.
- J. Hoydis *et al.*, "Toward a 6G AI-native air interface," *IEEE Commun. Mag.*, vol. 59, no. 5, 2021, pp. 76-81.
- B. Mao *et al.*, "Optimizing computation offloading in satellite-UAV-served 6G IoT: A deep learning approach," *IEEE Netw.*, vol. 35, no. 4, 2021, pp. 102-108.
- T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, 2017, pp. 563-575.
- M. K. Shehzad, L. Rose, and M. Assaad, "ML-based massive MIMO channel prediction: Does it work on real-world data?" *IEEE Wireless Commun. Lett.*, vol. 11, no. 4, 2022, pp. 811-815.
- R. Shafin *et al.*, "Artificial intelligence-enabled cellular networks: A critical path to beyond-5G and 6G," *IEEE Wireless Commun.*, vol. 27, no. 2, 2020, pp. 212-217.
- M. K. Shehzad *et al.*, "ML-based massive MIMO channel prediction: Does it work on real-world data?" *IEEE Wireless Commun.* Vol. 11, no. 4, 2022, pp. 811-815.
- M. K. Shehzad *et al.*, "Artificial intelligence for 6G networks: Technology advancement and standardization," *IEEE Vehicular Technology Magazine*, vol. 17, no. 3, 2022, pp. 16 - 25.
- F. Tariq *et al.*, "A speculative study on 6G," *IEEE Wireless Commun.*, vol. 27, no. 4, 2020, pp. 118-125.
- R. Zhong *et al.*, "AI empowered RIS-assisted NOMA networks: Deep learning or reinforcement learning?" *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, 2022, pp. 182-196.
- "IMT traffic estimates for the years 2020 to 2030," International Telecommunication Union, Geneva, Switzerland, Rep. ITU M.2370 0, 2015.
- 3GPP TR 23.791: "Study of enablers for network automation for 5G (Release 16)," 2019.
- 3GPP TR 38.331 "NR; Radio Resource Control (RRC); Protocol specification (Release 15)," 2021.
- 3GPP TR 37.317: Study on enhancement for data collection for NR and ENDC (Release 17), 2022.
- 3GPP TR 22.874 "5G System (5GS); Study on traffic characteristics and performance requirements for AI/ML model transfer (Release 18), 2021.

მეორე თავში გამოყენებული ლიტერატურა

- J. Bao *et al.*, "Towards a Theory of Semantic Communication," *Proc. IEEE Network Science Wksp.*, West Point, USA, 2011, pp. 110-117.
- N. Farsad *et al.*, "Deep Learning for Joint Source channel Coding of Text," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 2326-2330, Calgary, Canada, Apr. 2018.
- H.-T. He *et al.*, "Model-Driven Deep Learning for Physical Layer Communications," *IEEE Wireless Commun.*, vol. 26, no. 5, 2019, pp. 77-83.
- T. Hospedales *et al.*, "Meta-Learning in Neural Networks: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, 2022, pp. 5149-5169.
- P. Kairouz *et al.*, "Advances and Open Problem in Federated Learning," *Foundations and Trends in Machine Learning*, vol. 14, no. 1-2, 2021, pp. 1-210.

- R. J. McEliece et al., "The Capacity of the Hopfield Associative Memory," *IEEE Trans. Inf. The.*, vol. 33, no. 4, 1987, pp. 461-82.
- H. B. McMahan et al., "Communication-Efficient Learning of Deep Networks From Decentralized Data," *Proc. Mach. Learn. Res.*, vol. 54, 2017, pp. 1273-82.
- Z.-J. Qin, G. Y. Li, and H. Ye, "Federated Learning and Wireless Communications," *IEEE Wireless Commun.*, vol. 28, no. 5, 2021, pp. 134-40.
- C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, The University of Illinois Press, 1949.
- N. C. Thompson et al., "The Computing Limits of Deep Learning," <https://arxiv.org/abs/2007.05558>, July 2020.
- W. Tong and G. Ye Li, "Nine Challenges in Artificial Intelligence and Wireless Communications For 6G," *IEEE Wireless Commun.*, vol. 29, no. 4, 2022, pp. 140-145.
- H. Ye et al., "Deep Learning-Based End-to-End Wireless Communication Systems With GAN as Unknown Channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, 2020, pp. 3133-43.
- H. Ye, L. Liang, and G. Y. Li, "Decentralized Learning With Unreliable Communications," *IEEE J. Select. Topics Signal Proc.*, Apr. 2022, pp. 487-500.
- H. Q. Xie et al., "Deep Learning Enabled Semantic Communication Systems," *IEEE Trans. Signal Process.*, vol. 69, May 2021, pp. 2663-75.
- X.-Y. Wei et al., "Accretionary Learning With Deep Neural Networks With Applications," *IEEE Trans. Cognit. Commun. and Networking (Early Access)*, Dec, 2023.

შესაბუთო თავში გამოყენებული ლიტერატურა

- N. E. Bezai et al., "Future Cities and Autonomous Vehicles: Analysis of the Barriers to Full Adoption," *Energy and Built Environment*, 2020.
- A. Clemm et al., "Toward Truly Immersive Holographic-Type Communication: Challenges and Solutions," *IEEE Commun. Mag.*, vol. 58, no. 1, 2020, pp. 93-99.
- X. Huang, "Intelligent Remote Monitoring and Manufacturing System of Production Line Based on Industrial Internet of Things," *Comp. Commun.*, Elsevier, vol. 150, 2020, pp. 421-428.
- Y. Lu, X. Xu, and L. Wang, "Smart Manufacturing Process and System Automation: A Critical Review of the Standards and Envisioned Scenarios," *J. Manuf. Sys.*, Elsevier, vol. 56, 2020, pp. 312-325.
- G. Manogaran et al., "AI-Assisted Service Virtualization and Flow Management Framework for 6G-Enabled Cloud-Software-Defined Network Based IoT," *IEEE IoT J.*, 2021.
- P. Mugen, S. Yaohua, and W. Wenbo, "Intelligent-Concise Radio Access Networks in 6G: Architecture, Techniques and Insight," *J. Beijing Univ. Posts and Telecommun.*, vol. 43, no. 3, 2020.
- S. Nayak and R. Patgiri, "6G Communication Technology: A Vision on Intelligent Healthcare," *Health Informatics: A Computational Perspective in Healthcare*, ser. *Studies in Computational Intelligence*, R. Patgiri, A. Biswas, and P. Roy, Eds., 2021, pp. 1-18.
- A. Rasheed, O. San, and T. Kvamsdal, "Digital Twin: Values, Challenges and Enablers from a Modeling Perspective," *IEEE Access*, vol. 8, 2020, pp. 21,980-22,012.
- S. Saharan, S. Bawa, and N. Kumar, "Dynamic Pricing Techniques for Intelligent Transportation System in Smart Cities: A Systematic Review," *Comp. Commun.*, Elsevier, vol. 150, 2020, pp. 603-625.
- H. Sami et al., "AI-Based Resource Provisioning of IoE Services in 6G: A Deep Reinforcement Learning Approach," *IEEE Trans. Network and Service Management*, vol. 18, no. 3, 2021, pp. 3527-3540.
- A. H. Sodhro et al., "Toward 6G Architecture for Energy-Efficient Communication in IoT-Enabled Smart Automation Systems," *IEEE IoT J.*, vol. 8, no. 7, 2021, pp. 5141-5148.

-
- H. Tataria et al., "6G Wireless Systems: Vision, Requirements, Challenges, Insights, and Opportunities," *Proc. IEEE*, vol. 109, no. 7, 2021, pp. 1166-1199.
- H. Yang et al., "Artificial-Intelligence-Enabled Intelligent 6G Networks," *IEEE Network*, vol. 34, no. 6, 2020, pp. 272-280.
- X. You et al., "Towards 6G Wireless Communication Networks: Vision, Enabling Technologies, and New Paradigm Shifts," *Science China Info. Sciences*, vol. 64, no. 110301, 2021, pp. 1-74.
- T. Zhang et al., "Artificial Intelligence Driven 6G Networks: Endogenous Intelligence," *Telecommun. Science*, vol. 36, no. 9, 2020, pp. 14-22.
- F. Zhou et al., "Intelligence-Endogenous Networks: Innovative Network Paradigm for 6G," *IEEE Wireless Commun.*, vol. 29, no. 5, Feb. 2022, pp. 40-47.

მეოთხე თავში გამოყენებული ლიტერატურა

- M. Z. Alam et al., "Low-Delay Path Selection for Cluster-Based Buffer-Aided Vehicular Communications," *IEEE Trans. Vehicular Technology*, vol. 69, no. 9, 2020, pp. 9356-9363.
- D. He et al., "Physics and AI-Based Digital Twin of Multi-Spectrum Propagation Characteristics for Communication and Sensing in 6G and Beyond," *IEEE JSAC*, vol. 41, no. 11, 2023, pp. 3461-3473.
- J. He et al., "Jamming and Link Selection for Joint Secrecy/Delay Guarantees in Buffer-Aided Relay System," *IEEE Trans. Commun.*, vol. 70, no. 8, 2022, pp. 5451-5468.
- C. Huang, G. Chen, and Y. Gong, "Delay-Constrained Buffer-Aided Relay Selection in the Internet of Things With Decision-Assisted Reinforcement Learning," *IEEE Internet of Things J.*, vol. 8, no. 12, 2021, pp. 10198-10208.
- C. Huang et al., "Buffer-Aided Relay Selection for Cooperative Hybrid NOMA/OMA Networks With Asynchronous Deep Reinforcement Learning," *IEEE JSAC*, vol. 39, no. 8, 2021, pp. 2514-2525.
- I. Krikidis, T. Charalambous, and J. S. Thompson, "Buffer-Aided Relay Selection for Cooperative Diversity Systems Without Delay Constraints," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, 2012, pp. 1957-1967.
- J.-H. Lee et al., "Throughput Maximization of Mixed FSO/RF UAV-Aided Mobile Relaying With a Buffer," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, 2021, pp. 683-694.
- S. Luo and K. C. Teh, "Buffer State Based Relay Selection for Buffer-Aided Cooperative Relaying Systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, 2015, pp. 5430-5439.
- N. Nomikos et al., "Low-Complexity Buffer-Aided Link Selection With Outdated CSI and Feedback Errors," *IEEE Trans. Commun.*, vol. 66, no. 8, 2018, pp. 3694-3706.
- J. Vanschoren, "Meta-Learning," *Automated Machine Learning: Methods, Systems, Challenges*, 2019, pp. 35-61.
- P. Xu et al., "A Novel Link Selection in Coordinated Direct and Buffer-Aided Relay Transmission," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, 2023, pp. 3296-3309.
- P. Xu et al. "Deep Learning Driven Buffer-Aided Cooperative Networks for B5G/6G: Challenges, Solutions, and Future Opportunities," *IEEE Wireless Commun.*, vol. 31, no. 4, 2024, pp. 215-222.
- Z. Wei et al., "Toward Multifunctional 6G Wireless Networks: Integrating Sensing, Communication, and Security," *IEEE Commun. Mag.*, vol. 60, no. 4, 2022, pp. 65-71.
- J. Zhou et al., "Graph Neural Networks: A Review of Methods and Applications," *AI Open*, vol. 1, 2020, pp. 57-81.
- Y. Zhu et al., "Modeling and Performance Analysis for Satellite Data Relay Networks Using Two-Dimensional Markov-Modulated Process," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, 2020, pp. 3894-3907.

N. Zlatanov et al., "Buffer-Aided Cooperative Communications: Opportunities and Challenges," *IEEE Commun. Mag.*, vol. 52, no. 4, 2014, pp. 146-153.

მეხუთე თავში გამოყენებული ლიტერატურა

- C. Campolo et al., "5G Network Slicing for Vehicle-to-Everything Services," *IEEE Wireless Commun.*, vol. 24, no. 6, 2017, pp. 38-45.
- C. Gutterman et al., "RAN Resource Usage Prediction for a 5G Slice Broker," *Proc. ACM MobiHoc*, Catania, Italy, 2019.
- I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
- A. Kalokylos, "A Survey and an Analysis of Network Slicing in 5G Networks," *IEEE Commun. Stds. Mag.*, vol. 2, no. 1, 2018, pp. 60-65.
- M. Li et al., "Slicing-Based Artificial Intelligence Service Provisioning on the Network Edge: Balancing AI Service Performance and Resource Consumption of Data Management," *IEEE Vehic. Tech. Mag.*, vol. 16, no. 4, 2021, pp. 16-26.
- S.-C. Lin et al., "The Architectural Implications of Autonomous Driving: Constraints and Acceleration," *Proc. ASPLOS*, 2018, pp. 751-766.
- R. Minerva, G. M. Lee, and N. Crespi, "Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models," *Proc. IEEE*, vol. 108, no. 10, 2020, pp. 1785-1824.
- X. Shen et al., "Holistic Network Virtualization and Pervasive Network Intelligence for 6G," *IEEE Commun. Surveys & Tutorials*, vol. 24, no. 1, First quarter, 2022, pp. 1-30.
- X. Shen et al., "AI-Assisted Network-Slicing Based Next-Generation Wireless Networks," *IEEE Open J. Vehic. Tech.*, vol. 1, no. 1, 2020, pp. 45-66.
- W. Wu et al., "Dynamic RAN Slicing for Service-Oriented Vehicular Networks via Constrained Learning," *IEEE JSAC*, vol. 39, no. 7, 2021, pp. 2076-2089.
- W. Wu et al. "AI-Native Network Slicing for 6G Networks," *IEEE Wireless Commun.*, vol. 29, no. 1, Feb. 2024, pp. 96-103.
- X. You et al., "Towards 6G Wireless Communication Networks: Vision, Enabling Technologies, and New Paradigm Shifts," *Sci. China Info. Sci.*, vol. 64, no. 1, 2021, pp. 1-74.
- N. Zhang et al., "Software Defined Space-Air-Ground Integrated Vehicular Networks: Challenges and Solutions," *IEEE Commun. Mag.*, vol. 55, no. 7, 2017, pp. 101-109.
- C. Zhou et al., "Deep Reinforcement Learning for Delay-Oriented IoT Task Scheduling in Space-Air-Ground Integrated Network," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, 2021, pp. 911-925.
- W. Zhuang et al., "SDN/NFV-Empowered Future IoV with Enhanced Communication, Computing, and Caching," *Proc. IEEE*, vol. 108, no. 2, 2020, pp. 274-291.

მეექვსე თავში გამოყენებული ლიტერატურა

- K. Bonawitz et al., "Practical Secure Aggregation for Privacy-Preserving Machine Learning," *Proc. 2017 ACM SIGSAC Conf. Computer and Commun. Security*, 2017, pp. 1175-1191.
- K. Cao et al., "An Overview on Edge Computing Research," *IEEE Access*, vol. 8, 2020, pp. 85,714-85,728.
- H.-H. Chang, L. Liu, and Y. Yi, "Deep Echo State Q-network (DEQN) and Its Application in Dynamic Spectrum Sharing for 5G and Beyond," *IEEE Trans. Neural Networks and Learning Systems*, 2020, pp. 1-11.
- Y. Chen et al., "Asynchronous Online Federated Learning for Edge Devices with Non-IID Data," *Proc. 2020 IEEE Int'l. Conf. Big Data*, 2020, pp. 15-24.

-
- M. J. Haber, B. Chappell, and C. Hills, "Cloud Computing," *Cloud Attack Vectors*, Springer, 2022, pp. 9-25.
- S. Jere and Y. Yi, "Edge Intelligence for Beyond-5G through Federated Learning," *Proc. 2021 IEEE/ACM Symp. Edge Computing*, 2021, pp. 345-349.
- T. Li et al., "Federated Learning: Challenges, Methods, and Future Directions," *IEEE Signal Processing Mag.*, vol. 37, no. 3, 2020, pp. 50-60.
- Y. Li et al., "Privacy-Preserving Federated Learning Framework Based on Chained Secure Multiparty Computing," *IEEE IoT J.*, vol. 8, no. 8, 2020, pp. 6178-6186.
- Z. Li et al., "SoteriaFL: A Unified Framework for Private Federated Learning with Communication Compression," *Advances in Neural Information Processing Systems*, 2022.
- N. Mohammadi et al., "Differential Privacy Meets Federated Learning under Communication Constraints," *IEEE IoT J.*, 2021.
- S. Rajendran et al., "MapReduce-Based Big Data Classification Model Using Feature Subset Selection and Hyperparameter Tuned Deep Belief Network," *Scientific Reports*, vol. 11, no. 1, 2021, pp. 1-10.
- R. Shafin et al., "Artificial Intelligence-Enabled Cellular Networks: A Critical Path to Beyond-5G and 6G," *IEEE Wireless Commun.*, vol. 27, no. 2, 2020, pp. 212-217.
- H. Wang et al., "Optimizing Federated Learning on Non-IID Data with Reinforcement Learning," *Proc. IEEE INFOCOM 2020*, 2020, pp. 1698-1707.
- K. Zhang, Z. Yang, and T. Basar, "Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms," *Handbook of Reinforcement Learning and Control*, 2021, pp. 321-384.
- L. Zhang et al., "DQN-Based Mobile Edge Computing for Smart Internet of Vehicle," *EURASIP J. Advances in Signal Processing*, vol. 2022, no. 1, 2022, pp. 1-16.
- S. Jere et al. "Distributed Learning Meets 6G: A Communication and Computing Perspective," *IEEE Wireless Commun.*, vol. 28, no. 1, 2023, pp. 112-117.

მეშვიდე თავში გამოყენებული ლიტერატურა

- A.Abouaomar et al., "Federated Deep Reinforcement Learning for Open RAN Slicing in 6G Networks," *IEEE Comun. Mag.*, vol. 61, no. 2, 2023, pp. 126-132.
- G. Chen et al., "Two Tier Slicing Resource Allocation Algorithm Based on Deep Reinforcement Learning and Joint Bidding in Wireless Access Networks," *Sensors*, vol. 22, no. 9, 2022, p. 3495.
- A. Filali et al., "Multi-Access Edge Computing: A Survey," *IEEE Access*, vol. 8, 2020, pp. 197017-197046.
- A. Filali et al., "Dynamic SDN-Based Radio Access Network Slicing with Deep Reinforcement Learning for URLLC and eMBB Services," *IEEE Trans. Net. Sci. Eng.*, 2022.
- W. Jiang et al., "The Road Towards 6G: A Comprehensive Survey," *IEEE Open J. Commun. Society*, vol. 2, 2021, pp. 334-366.
- P. Kairouz et al., "Advances and Open Problems in Federated Learning," *Foundation and Trends in Machine Learning*, vol. 14, no. 1-2, 2021, pp. 1-210.
- M. K. Motaleb et al., "Resource Allocation in an Open RAN System Using Network Slicing," *IEEE Trans. Net. Service Management*, 2022.
- O-RAN Alliance, "O-RAN: Towards an Open and Smart RAN," *White Paper*, Oct. 2018.
- O-RAN Alliance, "O-RAN WG 2 AI/ML Workflow Description and Requirements," *Tech. Spec. v. 1.02.02*, 2020.
- D. A. Ravi et al., "RAN Slicing in Multi-MVNO Environment under Dynamic Channel Conditions," *IEEE IoT J.*, vol. 9, no. 6, 2021, pp. 4748-4757.

- M. Setayesh, S. Bahrami, and V. W. Wong, "Resource Slicing for eMBB and URLLC Services in Radio Access Network Using Hierarchical Deep Learning," *IEEE Trans. Wireless Commun.*, vol. 21, no. 11, 2022, pp. 8,950-8,966.
- H. Song et al., "A Deep Reinforcement Learning Framework for Spectrum Management in Dynamic Spectrum Access," *IEEE IoT J.*, vol. 8, no. 14, 2021, pp. 11208-11218.
- Z. Wang et al., "Utility Optimization for Resource Allocation in Edge Network Slicing Using DRL," *IEEE GLOBECOM 2020*, 2020.
- T. Wang et al., "LinkSlice: Fine-grained Network Slice Enforcement Based on Deep Reinforcement Learning," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 8, 2022, 2378-2394.

მერვე თავში გამოყენებული ლიტერატურა

- X. Chen et al., "Federated Learning over Multihop Wireless Networks With In-Network Aggregation," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 6, 2022, pp. 4622-4634.
- L. Deng et al., "Model Compression and Hardware Acceleration for Neural Networks: A Comprehensive Survey," *Proc. IEEE*, vol. 108, no. 4, 2020, pp. 485-532.
- C.-Y. Hsieh, Y.-C. Chuang, and A.-Y. Wu, "C3-SL: Circular Convolution-Based Batch-Wise Compression for Communication-Efficient Split Learning," *Proc. MLSP*, Aug. 2022.
- Huawei, *6G: The Next Horizon: From Connected People and Things to Connected Intelligence*, Cambridge, U.K.: Cambridge Univ. Press, 2021.
- M. Kim, A. DeRieux, and W. Saad, "A Bargaining Game for Personalized, Energy Efficient Split Learning over Wireless Networks," *Proc. WCNC*, Mar. 2023.
- Z. Lin et al., "Efficient Parallel Split Learning over Resource-Constrained Wireless Edge Networks," *IEEE Trans. Mobile Comput.*, Jan. 2024.
- Z. Lin et al. "Split Learning in 6G Edge Networks," *IEEE Wireless Commun.*, vol. 31, no. 4, 2024, pp. 170-176.
- C. Thapa et al., "Splitfed: When Federated Learning Meets Split Learning," *Proc. AAAI*, Feb. 2022.
- B. McMahan et al., "Communication-Efficient Learning of Deep Networks From Decentralized Data," *Proc. AISTATS*, Apr. 2017.
- J. Shao and J. Zhang, "Communication-Computation Trade Off in Resource-Constrained Edge Inference," *IEEE Commun. Mag.*, vol. 58, no. 12, 2020, pp. 20-26.
- S. Wang et al., "HiveMind: Towards Cellular Native Machine Learning Model Splitting," *IEEE JSAC*, vol. 40, no. 2, 2022, pp. 626-640.
- W. Wu et al., "Split Learning over Wireless Networks: Parallel Design and Resource Management," *IEEE JSAC*, vol. 41, no. 4, 2023, pp. 1051-1066.
- 3GPP, "Study on Traffic Characteristics and Performance Requirements for AI/ML Model Transfer in 5GS," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 22.874, 2021, version 18.2.0., Dec. 2021.

მეცხრე თავში გამოყენებული ლიტერატურა

- B. Bixby, "The Gurobi Optimizer," *Transp. Research Part B*, vol. 41, no. 2, 2007, pp. 159-178.
- M. Chen et al., "Creativebioman: A Brain- and Body-Wear able, Computing-Based, Creative Gaming System," *IEEE Sys. Man. and Cybernetics Maga.*, vol. 6, no. 1, 2020, pp. 14-22.

-
- S. Chen et al., "Vision, Requirements, and Technology Trend of 6G: How to Tackle the Challenges of System Coverage, Capacity, User Data-Rate and Movement Speed," *IEEE Wireless Commun.*, vol. 27, no. 2, 2020, pp. 218-228.
- Y. Hao et al., "Data-Driven Resource Management in a 5G Wearable Network Using Network Slicing Technology," *IEEE Sensors J.*, vol. 19, no. 19, 2019, pp. 8379-8386.
- Y. Hao et al., "Deep Reinforcement Learning for Edge Service Placement in Softwarized Industrial Cyber-Physical System," *IEEE Trans. Industrial Infor.*, vol. 17, no. 8, 2021, pp. 5552-5561.
- J. Kang et al., "Blockchain-Based Federated Learning for Industrial Metaverses: Incentive Scheme With Optimal Aol," *Proc. Conf. Blockchain*, 2022, pp. 71-78.
- V. Kelkkanen, M. Fiedler, and D. Lindero, "Bitrate Requirements of Non-Panoramic VR Remote Rendering," 2020, pp. 3624-3631.
- W. Y. B. Lim et al., "Dynamic Edge Association and Resource Allocation in Self-Organizing Hierarchical Federated Learning Networks," *IEEE JSAC*, vol. 39, no. 12, 2021, pp. 3640-3653.
- X. Lin et al., "Blockchain-Based On-Demand Computing Resource Trading in IoV-Assisted Smart City," *IEEE Trans. Emerging Topics in Comp.*, vol. 9, no. 3, 2021, pp. 1373-1385.
- A. I. Middy et al., "Auction-Based Resource Allocation Mechanism in Federated Cloud Environment: TARA," *IEEE Trans. Services Comp.*, vol. 15, no. 1, 2022, pp. 470-483.
- C. Qiu et al., "AI-mers-6G: AI-Driven Region-Temporal Resource Provisioning for 6G Immersive Services," *IEEE Wirel. Commun.*, vol. 30, no. 3, June 2023, pp. 196-203.
- X. Ren et al., "AI-Bazaar: A Cloud-Edge Computing Power Trading Framework for Ubiquitous AI Services," *IEEE Trans. on Cloud Comp.*, 2022, pp. 1-13.
- F. Tang et al., "The Roadmap of Communication and Networking in 6G for the Metaverse," *IEEE Wireless Commun.*, 2022, pp. 1-15.
- W. Xiao et al., "Dynamic Request Redirection and Resource Provisioning for Cloud-Based Video Services Under Heterogeneous Environment," *IEEE Trans. Parallel Distributed Syst.*, vol. 27, no. 7, 2016, pp. 1954-1967.
- M. Xu et al., "Wireless Edge-Empowered Metaverse: A Learning-Based Incentive Mechanism for Virtual Reality," *Proc. ICC, IEEE*, 2022, pp. 5220-5225.
- P. Yang et al., "6G Wireless Communications: Vision and Potential Techniques," *IEEE Net.*, vol. 33, no. 4, 2019, pp. 70-75.

აბრევიატურები და აკრონიმები

- 3GPP — Third Generation Partnership Project — მე-3 თაობის პარტნიორობის პროექტი
- AC — Actor-Critic — აქტორ-კრიტიკოსის (ალგორითმი)
- AF — Application Function — აპლიკაციის ფუნქცია
- AI — Artificial Intelligence — ხელოვნური ინტელექტი
- ANN — Artificial Neural Network — ხელოვნური ნეირონული ქსელი
- AP — Access Point — წვდომის წერტილი
- AR — Augmented Reality — მომატებული რეალობა
- B2B — Business-to-Business — ბიზნესი ბიზნესისთვის
- B2C — Business-to-Consumer — ბიზნესი მომხმარებლისთვის
- B5G — Beyond 5G — 5G-ის შემდგომი
- BACN — Buffer-Aided Cooperative Network — ბუფერზე დაფუძნებული კოოპერატიული ქსელი
- BDDA — Blockchain-based Double Dutch Auction — ბლოკჩეინზე დაფუძნებული ორმაგი პოლანდიური აუქციონი
- BLEU — Bilingual Evaluation Understudy — ორენოვანი შეფასების შესწავლა
- BS — Base Station — საბაზო სადგური
- CNN — Convolutional Neural Network — კონვოლუციური ნეირონული ქსელი
- CPU — Central Processing Unit — ცენტრალური პროცესორი
- CSI — Channel State Information — არხის მდგომარეობის შესახებ ინფორმაცია
- D — Dimension — განზომილება
- D2D — Device-to-Device — მოწყობილობებს შორის (კომუნიკაცია)
- DA — Double Auction — ორმაგი აუქციონი
- DDA — Double Dutch Auction — ორმაგი პოლანდიური აუქციონი
- DDPG — Deep Deterministic Policy Gradient — ღრმა დეტერმინირებული პოლიტიკის გრადიენტი
- DeepSC — Deep Semantic Communication — ღრმა სემანტიკური კომუნიკაცია
- DEQN — Deep Echo State Q-Network — ღრმა ექო მდგომარეობის Q-ქსელი
- DL — Deep Learning — ღრმა სწავლება
- DML — Distributed ML — განაწილებული ML
- DMRS — Demodulation Reference Signal — დემოდულაციის საცნობარო სიგნალი
- DNN — Deep Neural Network — ღრმა ნეირონული ქსელი
- DQN — Deep Q-Learning Network — ღრმა Q-სწავლების ქსელი
- DRL — Deep Reinforcement Learning — ღრმა განმტკიცებელი სწავლება
- DSA — Dynamic Spectrum Access — სპექტრზე დინამიკური წვდომა
- EL — Edge Learning — პერიფერიული სწავლება
- ELU — Exponential LU — ექსპონენციალური LU
- eMBB — enhanced Mobile Broadband — გაძლიერებული მობილური ფართობოლოვანი (ქსელი)
- EPSL — Efficient Parallel Split Learning — ეფექტიანი პარალელური გაყოფილი სწავლება
- FDMA — Frequency Division Multiple Access — სიხშირული დაყოფა მრავალჯერადი წვდომით
- FDRL — Federated DRL — ფედერირებული DRL
- FEEL — Federated Edge Learning — ფედერირებული პერიფერიული სწავლება
- FL — Federated Learning — ფედერირებული სწავლება
- G — Generation — თაობა
- GAN — Generative Adversarial Network — გენერაციულ-შეჯიბრებითი ქსელი
- gNB — Next Generation Node B — შემდგომი თაობის კვანძი B
- GNN — Graph Neural Network — გრაფზე დაფუძნებული ნეირონული ქსელი
- GPU — Graphics Processing Unit — გრაფიკული დამუშავების ბლოკი
- HAP — High Altitude Platform — მაღალი სიმაღლის პლატფორმა
- HMD — Head-Mounted Display — თავზე დასამაგრებელი დისპლეი

HO — Handover — ჰენდოვერი

IEN — Intelligence-Endogenous Network — ინტელექტუალურ-ენდოგენური ქსელი

IM — Index Modulation — ინდექს ოდულაცია

IMT — International Mobile Telecommunications — საერთაშორისო მობილური ტელეკომუნიკაციები

IoT — Internet of Things — საგნების ინტერნეტი

JSCC — Joint Source-Channel Coding — წყაროსა და არხისა ერთობლივი კოდირება

KPI — Key Performance Indicator — მახასიათებლების ძირითადი ინდიკატორი

LeakyReLU — ReLU-ის გაჟონილი ვერსია

LEO — Low Earth Orbit — დედამიწის დაბალი ორბიტა

LoS — Line of Sight — პირდაპირი ხედვის ხაზი

LSTM — Long Short-Term Memory — გრძელი მოკლევადიანი მეხსიერება

LTE — Long-Term Evolution — გრძელვადიანი ევოლუცია

M2M — Machine to Machine — მანქანა მანქანასთან

MANO — Management And Operation — მენეჯმენტი და ექსპლუატაცია

MAPPO — Multi-Agent Proximal Policy Optimization — მრავალაგენტიანი პროქსიმალური პოლიტიკის ოპტიმიზაცია

MARL — Multi-Agent Reinforcement Learning — მრავალაგენტიანი განმტკიცებელი სწავლება

MBB — Mobile Broadband — მობილური ფართობლოვანი (ქსელი)

MDP — Markov Decision Process — მარკოვის გადაწყვეტილების პროცესი

MEC — Mobile Edge Computing — მობილური პერიფერიული გამოთვლები

MF — Management Function — მართვის ფუნქცია

MIMO — Multiple Input Multiple Output — მრავალშესასვლელიანი და მრავალგამოსასვლელიანი (სისტემა)

ML — Machine Learning — მანქანური სწავლება

mMTC — massive Machine-Type Communication — მასობრივი მანქანური ტიპის კომუნიკაცია

MNO — Mobile Network Operator — მობილური ქსელის ოპერატორი

MSE — Mean Squared Error — საშუალო კვადრატული შეცდომა

mURLLC — massive Ultra-Reliable Low Latency Communication — მასობრივი ულტრა საიმედო დაბალი შეყოვნების კომუნიკაცია

MVNO — Mobile Virtual Network Operator — მობილური ქსელის ვირტუალური ოპერატორი

NF — Network Function — ქსელის ფუნქცია

NFV — Network Function Virtualization — ქსელის ფუნქციის ვირტუალიზაცია

NMS — Network Management System — ქსელის მართვის სისტემა

NMSF — NMS Function — NMS ფუნქცია

NN — Neural Network — ნეირონული ქსელი

NOMA — Non-Orthogonal Multiple Access — არაორთოგონალური მრავალჯერადი წვდომა

non-i.i.d — not Independent and Identically Distribution — არადამოუკიდებელი და არაერთგვაროვანი განაწილება

non-RT — non-Real-Time — არარეალური დრო

NR — New Radio — ახალი რადიო

NSMF — Network Slicing Management Function — ქსელის დაყოფის მართვის ფუნქცია

NSSMF — Network Slice Sub-domain Management Function — ქსელის ფენების ქვე-დომენის მართვის ფუნქცია

NTN — Non-Terrestrial Network — არატერესტრიალური ქსელი

NWDAF — Network Data Analytics Function — ქსელის მონაცემთა ანალიზის ფუნქცია

O&M — Operation and Maintenance — ექსპლუატაცია და ტექნიკური მომსახურება

OPEX — Operational Expenditure — საოპერაციო დანახარჯები

OptSFC — Optimizer for Security Functions — უსაფრთხოების ფუნქციების ოპტიმიზატორი

O-RAN — Open RAN — ღია RAN

PCA — Principal Component Analysis — ძირითადი კომპონენტების ანალიზი

PHY — Physical — ფიზიკური (ფენა)
PMI — Precoding Matrix Index — წინასწარი კოდირების მატრიცის ინდექსი
PReLU — Parametric ReLU — პარამეტრული ReLU
PrU — Primary User — პირველადი მომხმარებელი
PS — Public Security — საზოგადოებრივი უსაფრთხოება
PSL — Parallel Split Learning — პარალელური გაყოფილი სწავლება
QoS — Quality of Service — მომსახურების ხარისხი
QP — Queue Priority — რიგის პრიორიტეტი
RAN — Radio Access Network — რადიოწვდომის ქსელი
RAT — Radio Access Technology — რადიოწვდომის ტექნოლოგია
ReLU — Rectified Linear Unit — გასწორებული წრფივი ერთეული
RF — Radio Frequency — რადიოსიხშირე
RIC — RAN Intelligent Controller — RAN-ის ინტელექტუალური კონტროლერი
RIS — Reconfigurable Intelligent Surface — რეკონფიგურირებადი ინტელექტუალური ზედაპირი
RL — Reinforcement Learning — განმტკიცებელი სწავლება
RNN — Recurrent Neural Network — რეკურენტული ნეირონული ქსელი
RRM — Radio Resource Management — რადიორესურსების მართვა
RSRP — Reference Signal Receive Power — საცნობარო სიგნალის მიღებული სიმძლავრე
RT — Real-Time — რეალური დრო
SAGIN — Space-Air-Ground Integrated Network — კოსმოსურ-საჰაერო-სახმელეთო ინტეგრირებული ქსელი
SDN — Software Defined Networking — პროგრამული უზრუნველყოფით განსაზღვრული ქსელი
SEC — Self-Evolving Core — თვითგანვითარებული ბირთვი
SEL — Split Edge Learning — გაყოფილი პერიფერიული სწავლება
SELU — Scaled ELU — მასშტაბირებული ELU
SeU — Secondary User — მეორადი მომხმარებელი
SFL — Split Federated Learning — გაყოფილი ფედერირებული სწავლება
SGD — Stochastic Gradient Descent — სტოქასტური გრადიენტული დაღმართი
SINR — Signal-to-Interference-plus-Noise Ratio — სიგნალის ინტერფერენცია-პლუს-ხმაურთან თანაფარდობა
SL — Split Learning — გაყოფილი სწავლება
SLA — Service Level Agreement — სერვისის დონის შეთანხმება
SNR — Signal-to-Noise Ratio — სიგნალისა და ხმაურის თანაფარდობა
SON — Self-Organizing Network — თვითორგანიზებული ქსელი
SP — Service Provider — სერვისის პროვაიდერი
SSIM — Structural Similarity — სტრუქტურული მსგავსება
SU — Service User — სერვისის მომხმარებელი
SuL — Supervised Learning — ზედამხედველობითი სწავლება
SVM — Support Vector Machine — დამხმარე ვექტორული მანქანა
UAV — Unmanned Aerial Vehicle — უპილოტო საფრენი აპარატი
UE — User Equipment — მომხმარებლის მოწყობილობა
UL — Unsupervised Learning — ზედამხედველობის გარეშე სწავლება
URLLC — Ultra-Reliable Low-Latency Communication — ულტრა საიმედო დაბალი შეყოვნების კომუნიკაცია
V2V — Vehicle-to-Vehicle — სატრანსპორტო საშუალება-სატრანსპორტო საშუალებასთან
V2X — Vehicle-to-Everything — სატრანსპორტო საშუალება-ყველაფერთან
VM — Virtual Machine — ვირტუალური მანქანა
VMAF — Video Multi-method Assessment Fusion — ვიდეო მრავალმეთოდური შეფასების შერწყმა
VNF — Virtual Network Function — ვირტუალიზებული ქსელის ფუნქცია
VR — Virtual Reality — ვირტუალური რეალობა
XR — eXtended Reality — გაფართოებული რეალობა
WGAN — Wasserstein GAN — ვასერშტეინის GAN

