

ՀԱՅԱՍՏԱՆԻ ՀԱՆՐԱՊԵՏՈՒԹՅԱՆ ԳԻՏՈՒԹՅՈՒՆՆԵՐԻ ԱԶԳԱՅԻՆ ԱԿԱԴԵՄԻԱ
ԻՆՖՈՐՄԱՏԻԿԱՅԻ ԵՎ ԱՎՏՈՄԱՏԱՑՄԱՆ ՊՐՈԲԼԵՄՆԵՐԻ ԻՆՍՏԻՏՈՒՏ

ՀԱՅԿ ԱՆԴՐԱՆԻԿԻ ԳՐԻԳՈՐՅԱՆ

ԲԱՐՁՐ ԱՐՏԱԴՐՈՂԱԿԱՆՈՒԹՅԱՄԲ ՀԱՇՎԱՐԿՆԵՐ ԵՎ ՄԵԾ ՏՎՅԱԼՆԵՐ
ՊԱՐՈՒՆԱԿՈՂ ԽՆԴԻՐՆԵՐԻ ՎԵՐԼՈՒԾՈՒԹՅԱՆ ԵՎ ՎԻԶՈՒԱԼԻԶԱՑԻԱՅԻ
ՄԻՋԱՎԱՅՐԻ ՄՇԱԿՈՒՄ

ԱՏԵՆԱԽՈՍՈՒԹՅՈՒՆ

Ե.13.04 - «Հաշվողական մեքենաների, համալիրների, համակարգերի և ցանցերի
մաթեմատիկական և ծրագրային ապահովում» մասնագիտությամբ տեխնիկական
գիտությունների թեկնածուի գիտական աստիճանի հայցման համար

Գիտական ղեկավար՝
տեխնիկական գիտությունների թեկնածու Հրաչյա Ասցատրյան

Երևան-2018

Բովանդակություն

Ներածություն.....	4
ԳԼՈՒԽ 1 ԲԱՐՁՐ ԱՐՏԱԴՐՈՂԱԿԱՆՈՒԹՅԱՄԲ ՏՎՅԱԼՆԵՐԻ ՎԵՐԼՈՒԾՈՒԹՅՈՒՆ ԵՎ ՎԻՋՈՒԱԼԻԶԱՑԻԱ	10
1.1 Նախաբան.....	10
1.2 Բարձր արտադրողականությամբ հաշվողական համակարգեր և նրանց օգտագործումը.....	12
1.2.1 Զուգահեռ ծրագրավորման տեխնոլոգիաներ.....	14
1.2.2 Կլաստերային համակարգեր	18
1.2.3 Ամպային հաշվարկներ	22
1.3 Տվյալների վերլուծություն.....	25
1.3.1 Մեծ տվյալներ	27
1.3.2 Մեքենայական ուսուցում: Նեյրոնային ցանցեր	29
1.4 Տվյալների ինտերակտիվ վիզուալիզացիա	32
1.4.1 Վիզուալիզացիայի մեթոդներ և գործառույթներ.....	34
ԳԼՈՒԽ 2 ՏՎՅԱԼՆԵՐԻ ՎԵՐԼՈՒԾՈՒԹՅԱՆ ԵՎ ՎԻՋՈՒԱԼԻԶԱՑԻԱՅԻ ԵՆԹԱԿԱՌՈՒՑՎԱԾՔՆԵՐ ԵՎ ԳՈՅՈՒԹՅՈՒՆ ՈՒՆԵՑՈՂ ԼՈՒԾՈՒՄՆԵՐ	38
2.1 Նախաբան	38
2.2 Դեսկտոպ միջավայրում աշխատող ծրագրային գործիքներ	39
2.2.1 Հրամայական տողի կիրառմամբ ծրագրային գործիքներ	39
2.2.2 Օգտագործողի միջավայրով ծրագրային գործիքներ	42
2.3 Վեբ միջավայրում աշխատող ծրագրային գործիքներ	43
2.3.1 Վիզուալիզացիայի գրադարաններ.....	43
2.3.2 Օգտագործողի միջավայրով պատրաստի գործիքներ	47
2.4 Եզրակացություն	51
ԳԼՈՒԽ 3 ՏՎՅԱԼՆԵՐԻ ՎԵՐԼՈՒԾՈՒԹՅԱՆ ԵՎ ՎԻՋՈՒԱԼԻԶԱՑԻԱՅԻ ՀԱՄԱԼԻՐ ՀԱՄԱԿԱՐԳ.....	53
3.1 Նախաբան	53
3.2 Հաշվողական ենթակառուցվածք.....	55
3.3 Տվյալների կառավարում.....	57
3.4 Տվյալների վերլուծություն.....	58

3.5 Տվյալների վիզուալիզացիա	61
3.6 Տվյալների տարածում	62
3.7 Եզրակացություն	66
ԳԼՈՒԽ 4 ՄՇԱԿՎԱԾ ՀԱՄԱԼԻՐ ՀԱՄԱԿԱՐԳԻ ՓՈՐՁԱՐԿՈՒՄՆԵՐ	67
4.1 Ամպային ծառայություն ֆոտոնիկ դիսիպատիվ համակարգերի թվային հաշվարկների և վիզուալիզացիայի համար	67
4.1.1 Նախաբան	68
4.1.2 Կեռի ոչ գծային ռեզոնատորը չիրպացված արտաքին դաշտի ազդեցությամբ	69
4.1.3 Ամպային ծառայության ենթակառուցվածք	72
4.1.4 Տվյալների կառավարում	74
4.1.5 Օգտագործողի միջավայրի նկարագրություն	75
4.1.6 Կիրառություն	79
4.1.7 Եզրակացություն	82
4.2 Ամպային ծառայություն եղանակային տվյալների վերլուծության և վիզուալիզացիայի համար	83
4.2.1 Նախաբան	83
4.2.1 Ամպային ծառայության ենթակառուցվածքը	84
4.2.2 Տվյալների տարրեր, գործիքներ և մոդելներ	86
4.2.3 Եղանակի կանխատեսման WRF թվային մոդելը	89
4.2.4 Տվյալների կառավարում և վերլուծություն	96
4.2.5 Օգտագործողի միջավայր և վիզուալիզացիա	97
4.2.6 Կիրառություն	102
4.2.7 Եզրակացություն	104
Աշխատանքի հիմնական արդյունքները	106
Օգտագործված գրականության ցանկ	107

Ներածություն

Թեմայի արդիականությունը

Համակարգչային տեխնոլոգիաների զարգացման հետ զուգընթաց ավելանում են նաև գեներացվող տվյալները և նրանց տարաբնույթ բազմությունների (ստրուկտուրավորված և ոչ ստրուկտուրավորված) ծավալի և բարդ կառուցվածքի պատճառով անհրաժեշտություն է առաջանում կիրառել նոր մոտեցումներ՝ նրանց մշակման, պահպանման և վերլուծության համար:

Վերը նշված բազմությունները հանդիսանում են մեծ տվյալներ (Big Data), որը Դուգ Լանեյի կողմից սահմանվել է երեք հատկանիշների միջոցով (ծավալ, արագություն, բազմազանություն). «Մեծ տվյալները մեծ ծավալով, բարձր արագությամբ և/կամ մեծ բազմազանությամբ ինֆորմացիայի բազմություն է, որում որոշումների կայացման, իմաստների հայտնաբերման և գործընթացների օպտիմալացման համար պահանջվում են մշակման նոր ձևեր»: Ժամանակի ընթացքում ավելացել են մեծ տվյալները բնութագրող ևս երկու հատկանիշներ՝ ճշմարտացիություն և արժեք: Մեծ տվյալների մշակման, պահպանման, վերլուծության, վիզուալիզացիայի և տարածման համար անհրաժեշտ է կիրառել բարձր արտադրողականությամբ տվյալների վերլուծություն (HPDA – High Performance Data Analytics), որը միավորում է բարձր արտադրողականությամբ հաշվարկները (HPC – High Performance Computing) մեծ տվյալների հետ: HPDA խնդիրների լուծման համար առկա հայտնի գործիքները ապահովում են վերը նշված գործառույթների խմբից որոշակի ենթախմբերի իրականացում՝ կախված թիրախային խնդրից կամ յուրաքանչյուր գործառույթի համար գոյություն ունեցող ծրագրային լուծումները միմյանց հետ փոխկապակցելու բարդություններից: Օգտատերը գործառույթների ամբողջական ցիկլի իրականացման համար ստիպված է լինում խնդրի լուծման և արդյունքների ուսումնասիրության համար օգտագործել և փոխկապակցել մեկից ավելի ծրագրային միջավայր:

Այդ իսկ պատճառով խնդիր է առաջանում մշակել այնպիսի մեթոդներ և ծրագրային միջոցներ, որոնք հնարավորություն կտան միավորել բարձր արտադրողականությամբ հաշվարկների, տվյալների պահպանման, վերլուծության,

վիզուալիզացիայի և տարածման գործառույթները մեկ ընդհանուր համակարգում՝ օգտագործելով ժամանակակից ծրագրային լուծումները և ընդունված միջազգային հիմնօրինակները: Արդյունքում օգտատերը կկարողանա մեկ ծրագրային միջավայրի օգնությամբ իրականացնել խնդրի լուծման համար անհրաժեշտ ամբողջական գործընթացը, որը կնպաստի արդյունավետության բարձրացմանն ու ժամանակի խնայողությանը:

Աշխատանքի նպատակն ու խնդիրները

Աշխատանքի հիմնական նպատակն է բարձր արտադրողականությամբ հաշվարկներ և մեծ տվյալներ պարունակող խնդիրների լուծման համար կառուցել բաց կոդով, միջազգային հիմնօրինակներին համապատասխան, տվյալների մշակման, պահպանման, վերլուծությունների և վիզուալիզացիաներ տրամադրող, մետատվյալների տարածման հնարավորությամբ միասնական համակարգ: Այդ նպատակով դրվել են հետևյալ խնդիրները.

1. Մշակել տվյալների ստացման, ձևափոխման և վերլուծության ծրագրային մեթոդներ և գրադարաններ՝ տարբեր աղբյուրներից տվյալների հավաքագրման, ոչ ստրուկտուրավորված տվյալներից անհրաժեշտ գիտելիքի հայտնաբերման համար,
2. Մշակել այլընտրանքային պահոցում տվյալների պահպանման, նրանց հիման վրա մետատվյալների ստեղծման և տարածման ավտոմատ համակարգ՝ հաշվի առնելով միջազգային հիմնօրինակները,
3. Մշակել տվյալների փոխանցման և փոխակերպման ծրագրային մեթոդներ տրամաբանական տարբեր բաղադրիչների՝ հաշվողական ռեսուրսների, տվյալների պահպանման, վերլուծությունների, վիզուալիզացիաների և տարածման, միջև փոխհարաբերությունների ստեղծման համար:

Հետազոտման օբյեկտը

Աշխատանքում հիմնական հետազոտության օբյեկտներն են.

1. Բարձր արտադրողականությամբ համակարգերը և նրանց կիրառությունները,

2. Ստրուկտուրավորված և ոչ ստրուկտուրավորված տվյալների մշակման և պահպանման եղանակները,
3. Տվյալների վերլուծության միջոցները,
4. Մեծ ծավալով տվյալների մշակման համար գոյություն ունեցող մեթոդները, վերլուծության և ինտերակտիվ վիզուալիզացիայի ծրագրային միջավայրերը,
5. Մետատվյալների մշակման մեթոդները և լուծումները:

Հետազոտման մեթոդները

Աշխատանքի մեթոդաբանության և տեսական հիմքում ընկած են տվյալների վերլուծության, մշակման և վիզուալիզացիայի ոլորտների առաջատար մասնագետների աշխատանքները: Անհրաժեշտ ծրագրային միջոցների մշակման և կիրառման համար օգտագործվել են C/C++, Python, JAVA, JavaScript ծրագրավորման լեզուները և ազատ հասանելի NodeJs, Angular Js, Mondrian, Tensorflow ծրագրային միջավայրերը: Համակարգի նախագծման և մշակման ընթացքում հաշվի են առնվել տվյալ ոլորտներում աշխատող մասնագետների պահանջները, հնարավորինս հարմարավետ դարձնելու համար նրանց հետազոտությունների համար ստեղծված գործիքները: Մեթոդաբանության ցիկլը բաղկացած է հետազոտությունից, իրականացումից, թեստավորումից, արդյունքների վերլուծությունից, սխալների հայտնաբերումից և ուղղումներից:

Աշխատանքի գիտական նորույթը

- Մշակվել է միջազգային հիմնօրինակներին համապատասխան բարձր արտադրողականությամբ հաշվարկների իրականացման, տվյալների պահպանման, ձևափոխման, վերլուծության, ինտերակտիվ վիզուալիզացիայի և տարածման մեթոդներ պարունակող համալիր համակարգ՝ բարձր արտադրողականությամբ հաշվարկներ և մեծ տվյալներ պարունակող խնդիրների լուծման համար:
- Օգտագործելով առաջարկվող համալիր համակարգը, մշակվել են ամպային ծառայություններ.

1. Քվանտային ֆիզիկայում ֆոտոնիկ դիսիպատիվ համակարգերի (Photonic Dissipative System) թվային հաշվարկների և վիզուալիզացիայի համար,

2. Տարածաշրջանում եղանակի ուսումնասիրության և կանխատեսման համար օգտագործվող թվային մեթոդների արդյունքների ուսումնասիրության և նրանց ճշգրտության վերլուծությունների համար:

Ստացված արդյունքների կիրառական նշանակությունը

Մշակված ծրագրային միջավայրը կարելի է օգտագործել բարձր արտադրողականությամբ հաշվարկներ և մեծ տվյալներ պարունակող խնդիրներում՝ տվյալների պահպանման, վերլուծության, ինտերակտիվ վիզուալիզացիայի, տարածման և արդյունավետ հաշվարկների իրականացման համար, օգտագործելով ժամանակակից հաշվողական և ծրագրային միջոցները: Միջավայրը փորձարկվել է երկու տարբեր գիտական խնդիրների համար.

1. Քվանտային ֆիզիկայում ֆոտոնիկ դիսիպատիվ համակարգերում թվային հաշվարկների և վիզուալիզացիայի համար իրականացվել է բազմաօգտատեր ամպային ծառայություն, որը հնարավորություն է տալիս մեկ համակարգում՝ ընտրելով հետազոտության համար անհրաժեշտ մուտքային պարամետրեր, կատարել հաշվարկներ և հետազոտել արդյունքները տրամադրված վիզուալիզացիայի գործիքների միջոցով:
2. Տարածաշրջանում եղանակի ուսումնասիրության և կանխատեսման համար օգտագործվող թվային մեթոդների արդյունքների ուսումնասիրության և նրանց ճշգրտության վերլուծության համար մշակվել է ամպային ծառայություն, որը հնարավորություն է տալիս ոչ միայն ինտեգրել փաստացի առկա դիտարկման տվյալները, կանխատեսման թվային մոդելների արդյունքները, այլ նաև կատարել վերլուծություններ՝ օգտագործելով տրամադրված վիզուալիզացիայի տեսակները և վերլուծական պարամետրերը:

Ներդրումներ

Մշակված համակարգի կիրառությամբ կառուցված տարածաշրջանում եղանակի ուսումնասիրության և կանխատեսման համար նախատեսված ամպային ծառայությունը ներդրվել է «Հիդրոօդերևութաբանության և մթնոլորտային երևույթների վրա ակտիվ ներգործության ծառայություն» ՊՈԱԿ-ում:

Պաշտպանությանը ներկայացվում են հետևյալ դրույթները

- Բարձր արտադրողականությամբ հաշվարկներ և մեծ տվյալներ պարունակող խնդիրների լուծման համար մշակված՝ միջազգային հիմնօրինակներին համապատասխան, բարձր արտադրողականությամբ հաշվարկներ, տվյալների պահպանում, վերլուծություն և վիզուալիզացիա ապահովող, մետատվյալների տարածման հնարավորությամբ, համալիր համակարգը:
- Համակարգի կիրառության արդյունքում կառուցված հետևյալ ամպային ծառայությունները.
 1. Քվանտային ֆիզիկայում ֆոտոնիկ դիսիպատիվ համակարգերում թվային հաշվարկների և վիզուալիզացիայի համար,
 2. Տարածաշրջանում եղանակի ուսումնասիրության և կանխատեսման համար օգտագործվող թվային մեթոդների արդյունքների ուսումնասիրության և նրանց ճշգրտության վերլուծությունների համար:

Ստացված արդյունքների հավաստիությունը

Ստացված արդյունքների հավաստիությունը հիմնավորվում է մշակված ծրագրային համակարգի կիրառմամբ ստացված մի շարք փորձնական արդյունքներով:

Ստացված արդյունքների ապրոբացիան

Աշխատանքի հիմնական արդյունքները և դրույթները զեկուցվել և քննարկվել են Սոֆիայում կայացած «Two Years Avitohol: Advanced High Performance Computing Applications 2017» սեմինարում, Բելգրադում կայացած «VI-SEEM Climate» տարածաշրջանային վերապատրաստման ժամանակ, ինչպես նաև ԻԱՊԻ ընդհանուր սեմինարներում:

Հրապարակումներ

Ատենախոսության հիմնական արդյունքները հրապարակվել են 5 գիտական աշխատություններում, որոնց ցուցակը բերված է սեղմագրի վերջում:

Աշխատանքի կառուցվածքը և ծավալը

Ատենախոսությունը կազմված է ներածությունից, 4 գլխից և օգտագործված գրականության ցանկից: Ատենախոսությունը շարադրված է 112 էջերում և պարունակում է 68 գրականության հղում:

Աշխատանքի բովանդակությունը

Աշխատանքի առաջաբանում հիմնավորված է թեմայի արդիականությունը, սահմանված է աշխատանքի նպատակն ու հիմնական խնդիրները, նկարագրված են ուսումնասիրվող օբյեկտները և հետազոտման մեթոդները: Բերված են ստացված արդյունքները, դրանց գիտական նորույթը և կիրառական նշանակությունը:

Առաջին գլխում ներկայացված են աշխատանքի հիմքում ընկած տեսական և տեխնոլոգիական միջոցները: Այստեղ ներկայացված են բարձր արտադրողականությամբ համակարգերի դերը և կարևորությունը արդի խնդիրների լուծման համար, կլաստերային և ամպային համակարգերը իրենց տեսակներով և կիրառական ոլորտներով: Ներկայացված է տվյալների վիզուալիզացիայի դերը և նշանակությունը, մեծ տվյալների վերլուծությունների ժամանակ առաջացող բարդությունները:

Երկրորդ գլխում ներկայացված են տարբեր ոլորտներում կիրառություններ ունեցող տվյալների վերլուծության և վիզուալիզացիայի համար նախատեսված գրադարանները և ծրագրային միջավայրերը: Նկարագրված են նրանց տրամադրած հնարավորությունները և կատարվել են համեմատական վերլուծությունները:

Երրորդ գլխում ներկայացված է մշակված ծրագրային համալիր համակարգը, յուրաքանչյուր բաղադրիչի նշանակությունը, նրանց միջև գոյություն ունեցող կապերը և փոխազդեցությունների համար մշակված մեթոդներն ու ծրագրային միջոցները:

Չորրորդ գլխում ներկայացված են առաջարկվող համալիր համակարգի հենքի վրա մշակված ամպային ծառայությունները՝ քվանտային ֆիզիկայում ֆոտոնիկ դիսիպատիվ համակարգերի թվային հաշվարկների ու վիզուալիզացիայի և եղանակի կանխատեսման տվյալների վիզուալիզացիայի և նրանց համապատասխանելիության վերլուծության համար:

ԳԼՈՒԽ 1

ԲԱՐՁՐ ԱՐՏԱԴՐՈՂԱԿԱՆՈՒԹՅԱՄԲ ՏՎՅԱԼՆԵՐԻ ՎԵՐԼՈՒԾՈՒԹՅՈՒՆ ԵՎ ՎԻԶՈՒԱԼԻԶԱՑԻԱ

Այս գլխում ներկայացված է բարձր արտադրողականությամբ տվյալների վերլուծությունը ժամանակակից համակարգերում, նրանց դերն ու կարևորությունը: Ներկայացված են բարձր արտադրողականությամբ հաշվարկները, կլաստերային և ամպային համակարգերը, տվյալների վիզուալիզացիայի դերը գիտական տվյալներից ինֆորմացիայի ստացման և վերլուծություն համար:

1.1 Նախաբան

Տվյալների մասին գիտությունը գիտական մեթոդների, պրոցեսների, ալգորիթմների և համակարգերի միջոցառումներն է, որը նախատեսված է տվյալների տարբեր՝ ստրուկտուրավորված և ոչ ստրուկտուրավորված, ձևերից գիտելիքների և գաղափարների ներկայացման համար [1]: Ինտելեկտուալ հետազոտության գրեթե բոլոր բնագավառներում այն առաջարկում է նոր մոտեցում հայտնագործությունները կատարելու համար: Համատեղելով վիճակագրության, համակարգչային գիտության, կիրառական մաթեմատիկայի և վիզուալիզացիայի ասպարեզները, տվյալների մասին գիտությունը կարող է մեծ թվով տվյալները դարձնել նոր հասկացությունների և նոր գիտելիքների հիմք: Ընդհանուր առմամբ, տվյալների մասին գիտությունը ներառում է ավտոմատացված մեթոդների մշակումն ու կիրառումը մեծ ծավալով տվյալների վերլուծության և նրանցից գիտելիքների հայտնաբերման համար: Նմանատիպ ավտոմատացված մեթոդների շնորհիվ, ամենուրեք՝ գենետիկայից մինչև բարձր էներգիայի ֆիզիկա, տվյալների գիտությունը օգնում է ստեղծել գիտության նոր ճյուղեր և ազդել սոցիալական գիտությունների ու հումանիտար ոլորտների վրա: Ակնկալվում է, որ առաջիկա տարիներին այդ միտումը արագանալու է, քանի որ բջջային սենսորներից, բարդ գործիքներից, վեբ միջավայրերից և այլ աղբյուրներից ստացված տվյալները աճում են:

Տվյալների ծավալների աճը և նրանց բազմազանությունը առաջ է բերում նոր մոտեցումների կիրառման անհրաժեշտություն, քանի որ ավանդական մեթոդները

բավարար չեն: Մեծ քանակությամբ տվյալների վերլուծության համար օգտագործվում են բարձր արտադրողականությամբ համակարգեր, որոնք առաջարկում են նոր լուծումներ՝ առաջացած խնդիրների համար:

Բարձր արտադրողականությամբ տվյալների վերլուծությունը (HPDA) դա մեծ տվյալները վերլուծելու և նրանցից օրինաչափությունների և իմաստների հայտնաբերման համար բարձր արտադրողականությամբ հաշվարկների (HPC) օգտագործումն է [2]: Տվյալների ծավալների աճը պատճառ է դառնում բարձր արտադրողականությամբ տվյալների վերլուծության օգտագործման, որն առաջանում է գոյություն ունեցող և նոր գործոնների համակցությունից.

- Հզոր HPC համակարգերը ունակ են լուծելու մեծ ծավալով, ավելի բարձր թույլատրությամբ և ավելի շատ տվյալների տարրեր պարունակող խնդիրներ:
- Ավելի մեծ, ավելի բարդ գիտական գործիքների և սենսորային ցանցերի լայնորեն կիրառումը, «խելացի» էլեկտրական ցանցերից մինչև խոշոր Հադրոն կոլայդեր:
- Որոշակի առարկաների վերափոխումը տվյալների վրա հիմնված գիտությունների մեջ: Կենսաբանությունը հայտնի օրինակ է, բայց այս վերափոխումը նույնիսկ տարածվում է հումանիտար առարկաների վրա, ինչպիսիք են՝ հնագիտությունը և լեզվաբանությունը:
- Ստոխաստիկ մոդելավորման (ֆինանսական ծառայությունների) աճը, պարամետրային մոդելավորում (արտադրություն) և այլ խնդիրների լուծման իտերացիոն մեթոդներ, որոնց կուտակային արդյունքները մեծ տվյալների ծավալ են կազմում:
- Նոր, ավելի առաջադեմ վերլուծական մեթոդների և գործիքների առկայություն. MapReduce / Hadoop, գրաֆիկական և սեմանտիկական վերլուծություն, գիտելիքների բացահայտման ալգորիթմներ և այլն:
- Առաջնային վերլուծության իրականացման պահանջը իրական ժամանակում, որը անհրաժեշտություն է առաջացնում կոմերցիոն ընկերություններում HPC համակարգերի կիրառման:

Ներկայումս HPC-ն ունի լայն կիրառություն գիտության և այլ ոլորտներում, քանի որ այն օգտագործում է միասնական հաշվողական ուժը այնպիսի ձևով, որն ապահովում է ավելի բարձր ցուցանիշներ, քան գիտական, ճարտարագիտական կամ բիզնեսի խոշոր խնդիրների լուծման համար կարող են ապահովել սովորական համակարգիչները կամ համակարգչային կայանները:

1.2 Բարձր արտադրողականությամբ հաշվողական համակարգեր և նրանց օգտագործումը

Համակարգիչների զարգացման պատմությունը ցույց է տալիս, որ երբ գոյություն ունեցող տեխնոլոգիաների օգնությամբ արդեն լիովին հնարավոր է լինում լուծել հայտնի խնդիրները, ապա առաջանում են նոր խնդիրներ, որոնք պահանջում են օգտագործել նոր տեխնոլոգիաներ, մոտեցումներ և հզորություններ: Այդպիսի խնդիրներ գոյություն ունեն ինչպես քվանտային ֆիզիկայի, այնպես էլ կլիմայագիտության, գենային ինժեներիայի և այլ բնագավառներում: Սակայն մեկ համակարգչի օգնությամբ անհնար է լուծել նման խնդիրները ողջամիտ ժամանակահատվածում: Որպես այդպիսի խնդիր կարելի է դիտարկել 2007թ. European Organization for Nuclear Research (CERN) [3] լաբորատորիայի կողմից գործարկված արագացուցիչը, որի նպատակն է միջուկային ֆիզիկայի բնագավառում գտնել նոր տարրական մասնիկներ: Գործարկված արագացուցիչից տարեկան արտածվելու է մոտ 5000 տերաբայթ ինֆորմացիա, որի մշակման համար անհրաժեշտ կլինեն տարբեր երկրներում գտնվող ավելի քան 25000 համակարգիչ:

Այսպիսի խնդիրների առկայության պարագայում ակնհայտ է դառնում, որ հարկավոր է կիրառել խնդիրների լուծման համար նոր մոտեցումներ ու տեխնոլոգիաներ: Սկզբում փորձ արվեց համախմբել մի քանի համակարգիչ այնպես, որ դրանք աշխատեն միևնույն խնդրի տարբեր հատվածների վրա միաժամանակ: Սակայն համակարգիչների համախմբումը միասնական մեկ համակարգի մեջ առաջ բերեց նոր խնդիրներ, քանի որ շատ կարևոր էր, որ առաջադրված խնդիրը տրոհվի մասերի այնպես, որ նրանց միաժամանակյա իրականացումը լինի հնարավոր: Արդյունքում խնդիրը մասերի տրոհման եղանակով լուծելու համար առաջացան հատուկ օպերացիոն

համակարգեր, ծրագրավորման լեզուներ, ծրագրային և ապարատային միջոցներ: Ահա այսպես ստեղծվեցին բարձր արտադրողականությամբ հաշվողական համակարգերը (գերհամակարգիչ, կլաստեր), որոնք իրենցից ներկայացնում են բարդ համալիրներ և, որոնց պարամետրերը որոշվում են ոչ միայն առանձին բաղադրիչների բնութագրիչներով (պրոցեսոր, հիշողություն և այլն), այլ նաև իրենց ճարտարապետությամբ:

Բարձր արտադրողականությամբ հաշվարկները, որոնք հայտնի են որպես HPC, սահմանվում են որպես հաշվարկային և տվյալների ինտենսիվ խնդիրների՝ ներառյալ հաշվարկների, մոդելավորման և մատուցման, լուծման համար համակցված հաշվարկային հզորության օգտագործումը, որ ստանդարտ համակարգերը չեն կարողանում լուծել՝ սահմանափակ հնարավոր ռեսուրսների պատճառով (չափից շատ տվյալներ են պահանջվում) [4]: HPC-ն այս սահմանափակումները հաղթահարելու մոտեցումն է, որն օգտագործում է մասնագիտացված, բարձրակարգ սարքավորումներ՝ հաշվարկային հզորություն կուտակելու համար: Տվյալները և գործողությունները համապատասխան ձևով մի քանի մասերի բաշխելու համար անհրաժեշտ է իրականացնել զուգահեռացում:

Բարձր արտադրողականությամբ հաշվարկների պատմությունը սկիզբ է առնում 1960-ականներից. դիսկրետ գերմանիումով տրանզիստորներ օգտագործող երկրորդ սերնդի Atlas մեքենան Մանչեստրի համալսարանում, IBM-ի կողմից ստեղծված առաջին տրանզիստորներով IBM 7030 Stretch սուպերհամակարգիչը և Seymour Cray-ի կողմից Control Data Corporation (CDC)-ում մշակված համակարգիչների սերիան [5]:

HPC-ները հիմնականում օգտագործվում են երկու պատճառով: Առաջին պատճառը դա կենտրոնական հաշվողական միավորների և հանգույցների թվի ավելացման շնորհիվ հասանելի է դառնում ավելի մեծ հաշվողական հզորություն: Այն հնարավորություն է տալիս մոդելները հաշվարկել ավելի արագ, քանի որ միավոր ժամանակում կատարվում է ավելի շատ գործողություն: Զուգահեռ համակարգում կատարման ժամանակի հարաբերությունը հաջորդական համակարգի կատարման ժամանակին կոչվում է արագացում: Արագացման վերին սահմանը կախված է, թե ինչ չափով է ծրագրային կոդը զուգահեռացված: Ըստ Ամդահալի օրենքի [6].

$$S_{max}(\varphi) = \lim_{P \rightarrow \infty} S(P) := 1/(1 - \varphi),$$

որտեղ S -ը արագացումն է, φ -ն կողի զուգահեռացված հատվածի հարաբերությունն է ընդհանուր կողին, իսկ P -ն պրոցեսների քանակը:

HPC-ների օգտագործման երկրորդ պատճառը դա այն է, որ կլաստերային համակարգի օգտագործման դեպքում հասանելի հիշողության ծավալը գծային ձևով կախված է լրացուցիչ հանգույցների ավելացման հետ: Այդ պատճառով հաշվողական միավորների աճի դեպքում ավելի մեծ մոդելներ կարող են հաշվարկվել: Սա կոչվում է լայնածավալ արագացում: Այս մոտեցման կիրառումը հնարավորություն է տալիս շրջանցել Ամդահալի օրենքը, որը դիտարկվում է ֆիքսված չափի խնդիրների համար: Հաշվողական հզորության կրկնապատկումը և հիշողությունը հնարավորություն են տալիս երկու անգամ ավելի մեծ խնդիրը հաշվարկել միևնույն ժամանակում: Գուստաֆսոն-Բարսիսի օրենքը բացատրում է այս երևույթը [6].

$$S(P) = \Phi \cdot P - (1 - \Phi),$$

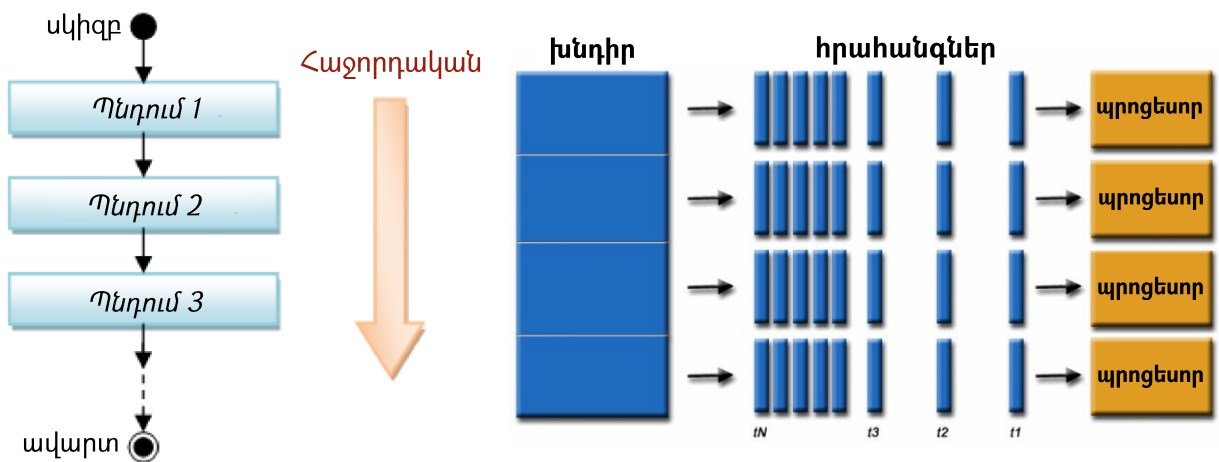
որտեղ $S(P)$ -ն լայնածավալ արագացումն է, Φ -ն կողի զուգահեռացված հատվածի հարաբերությունն է ընդհանուր կողին, իսկ P -ն պրոցեսների քանակը:

1.2.1 Զուգահեռ ծրագրավորման տեխնոլոգիաներ

Սովորաբար համակարգչային ծրագրերը ստեղծվել են հաջորդական հաշվարկների համար: Խնդիրը լուծելու համար ալգորիթմը կառուցվում և իրականացվում էր որպես հրահանգների հաջորդական հոսք, որոնք կատարվում էին համակարգչի CPU-ի վրա: Միաժամանակ կարող էր իրականացվել միայն մեկ հրահանգ և ամեն մի հրահանգի ավարտից հետո սկսվում էր մյուսը:

Մյուս կողմից, զուգահեռ հաշվողականությունը խնդիրը լուծելու համար օգտագործում է բազմաթիվ մշակման տարրեր՝ զուգահեռ աշխատանքի սկզբունքով: Դա իրականացնելու համար խնդիրը բաժանվում է իրարից անկախ մասերի այնպես, որ ամեն մշակող տարր կարողանա իրականացնել իր մասը զուգահեռ: Մշակման տարրերը կարող են լինել բազմազան և ներառել այնպիսի ռեսուրսներ, ինչպիսիք են բազմակի

պրոցեսորներով համակարգիչը, մի քանի ցանցային համակարգիչներ, մասնագիտացված սարքավորումներ կամ նշվածներից ցանկացած կոմբինացիա (նկ. 1.1):



Նկար 1.1. Զուգահեռ և հաջորդական ծրագրավորման սխեմաներ:

Զուգահեռ համակարգիչները կարելի է պայմանականորեն դասակարգել ըստ զուգահեռացում տրամադրող ապարատային մակարդակի: Ստորև ներկայացված են միջոցներ, որոնք թույլ են տալիս որոշ համակարգերում իրականացնել զուգահեռացում:

Ընդհանուր հիշողությամբ հաշվողական համակարգեր - OpenMP

OpenMP-ն (Open Multi-Processing) API (Application Program Interface) է, որը հնարավորություն է տալիս բազմապլատֆորմ ընդհանուր հիշողությամբ բազմամշակման ծրագրավորում C/C++ և Fortran լեզուներով պրոցեսորային ճարտարապետություններում և օպերացիոն համակարգերում, ինչպիսիք են՝ Linux, Mac OS X, Windows, Solaris և այլն [7]: Այն ընդունված է որպես բարձր մակարդակի ընդհանուր հիշողությամբ զուգահեռ ծրագրավորման հիմնօրինակ: Այն հիմնված է կոմպիլատորի հրահանգների, գրադարանային պրոցեսորների և միջավայրի փոփոխականների համադրության վրա:

OpenMP-ի օգտակար հատկություններից մեկն այն է, որ այն թույլ է տալիս օգտագործել նույն ելակետային կոդը և՛ OpenMP-ի, և՛ սովորական կոմպիլատորների հետ: Այս հատկությունը հասանելի է շնորհիվ նրան, որ OpenMP-ի հրահանգներն ու

հրամանները անտեսանելի են սովորական կոմպիլատորների համար: Պրոցեսը հանդիսանում է համակարգչային կատարվող ծրագրի նմուշ: Այն իր մեջ պարունակում է ծրագրային կոդը և ընթացիկ գործունեությունը: Իրականացման թրեդը հանդիսանում է մշակման ամենափոքր միավորը, որը նախատեսվում է օպերացիոն համակարգի կողմից:

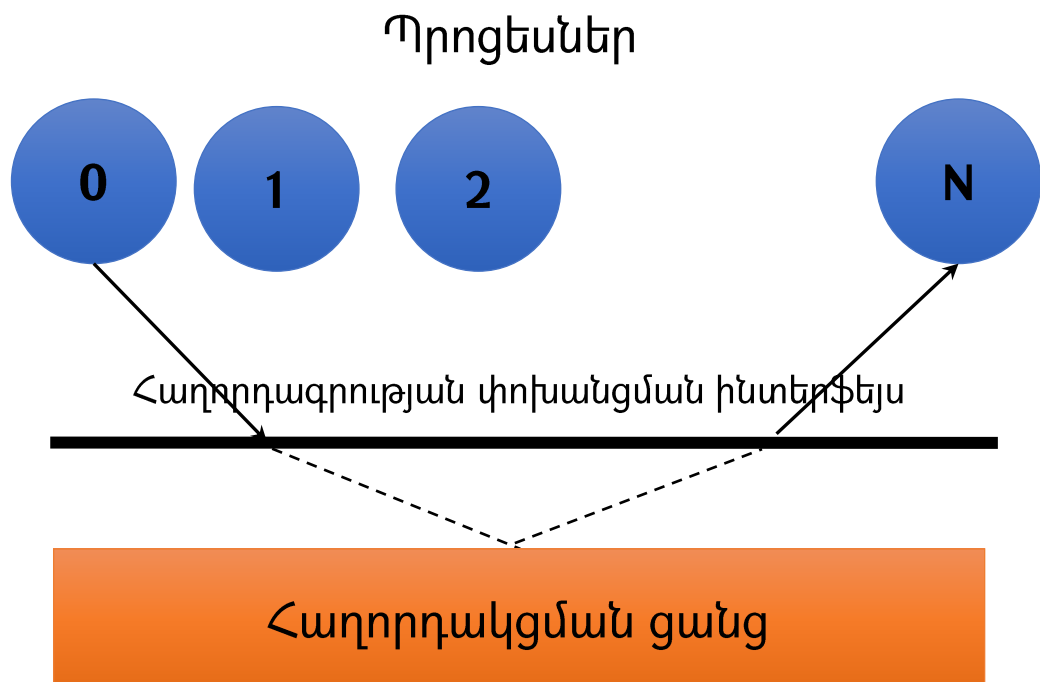
Ծրագրային կոդի OpenMP-ով զուգահեռացնելու հիմնական տեխնիկան կոմպիլատորի հրահանգներն օգտագործելն է: Հրահանգները ավելացվում են կոդի աղբյուրում որպես ցուցիչ կոմպիլատորին, որ այդ բլոկը պետք է իրականացվի զուգահեռ որոշ հրահանգավորումներով, որոնք տեղեկացնում են, թե ինչպես պետք է իրականացվի զուգահեռացումը:

OpenMP-ն հիմնված է ընդհանուր հիշողությամբ ծրագրավորման պարադիգմի վրա, որում առկա է բազմակի թրեդների հնարավորությունը: Ընդհանուր հիշողությամբ պրոցեսները բաղկացած են բազմակի թրեդներից, որոնք ծրագրի աշխատանքի ընթացքում կարող են ստեղծվել և ոչնչացվել: Բազմակի համագործակցող թրեդներ կարող են աշխատել միաժամանակ: OpenMP-ով կառուցված ծրագրերը բաղկացած են զուգահեռ բլոկներից և երկու այդպիսի բլոկներում միայն մեկն է որպես գլխավոր թրեդ: Նոր թրեդները աշխատում են միաժամանակ գլխավոր թրեդի հետ և զուգահեռ բլոկի վերջում ոչնչանում են:

Տարաբաշխված հիշողությամբ հաշվողական համակարգեր - MPI

Հաղորդագրությունների փոխանցման ինտերֆեյսը (MPI) բազմաթիվ զուգահեռ համակարգիչների վրա աշխատող ստանդարտացված և շարժական հաղորդագրություն-փոխանցում մի համակարգ է՝ նախագծված կրթության և արդյունաբերության ոլորտի մի խումբ հետազոտողների կողմից (նկ. 1.2) [8]: Ստանդարտով որոշվում է գրադարանի ենթածրագրերի միջուկի ձևաչափը և իմաստաբանությունը՝ օգտակար օգտագործողների լայն զանգվածի համար, ովքեր գրում են հաղորդագրություն փոխանցելու շարժական ծրագրեր տարբեր համակարգչային ծրագրավորման լեզուներով, ինչպիսիք են՝ Fortran, C, C++ և Java. Կան մի քանի լավ փորձարկված և արդյունավետ MPI-ի իրականացումներ, այդ թվում նաև

անվճար և համընդհանուր հասանելի, ինչը նպաստեց արդյունաբերության համար զուգահեռ ծրագրային ապահովման զարգացմանը և բերեց մեծաքանակ շարժական, գործնական, էֆեկտիվ, ճկուն և ընդլայնելի զուգահեռ ծրագրերի ստեղծման:



Նկար 1.2. MPI-ի կառուցվածքը:

MPI-ն ապահովում են զուգահեռ սարքավորումների մատակարարները՝ հստակ որոշված ֆունկցիաների հիմնական հավաքածուներով, որոնք կարող են արդյունավետ իրագործվել: Որպես արդյունք, սարքավորումների մատակարարները այդ հավաքածուներով կարող են կառուցել ցածր մակարդակի ստանդարտ պրոցեդուրա՝ ստեղծելու համար ավելի բարձր մակարդակի ենթածրագրեր բաշխված հիշողության հաղորդակցման միջավայրի համար: MPI-ն ապահովում է հիմնական օգտագործողների համար հասարակ, բայց բավականին հզոր ինտերֆեյս, թույլ տալու համար ծրագրավորողներին օգտագործել բարձր արտադրողականության հաղորդագրություն փոխանցելու գործողություններ:

Որպես փորձ՝ ստեղծելու հաղորդագրություն փոխանցելու համար «իրական» ստանդարտ, հետազոտողները MPI-ի մեջ ներառել են որոշ համակարգերի առավել օգտակար հատկանիշներ, այլ ոչ թե որպես ստանդարտ վերցրել են մեկ համակարգ: Հատկանիշները ընտրվել են IBM, Intel, nCUBE, PVM, Express, P4 և PARMACS համակարգերից:

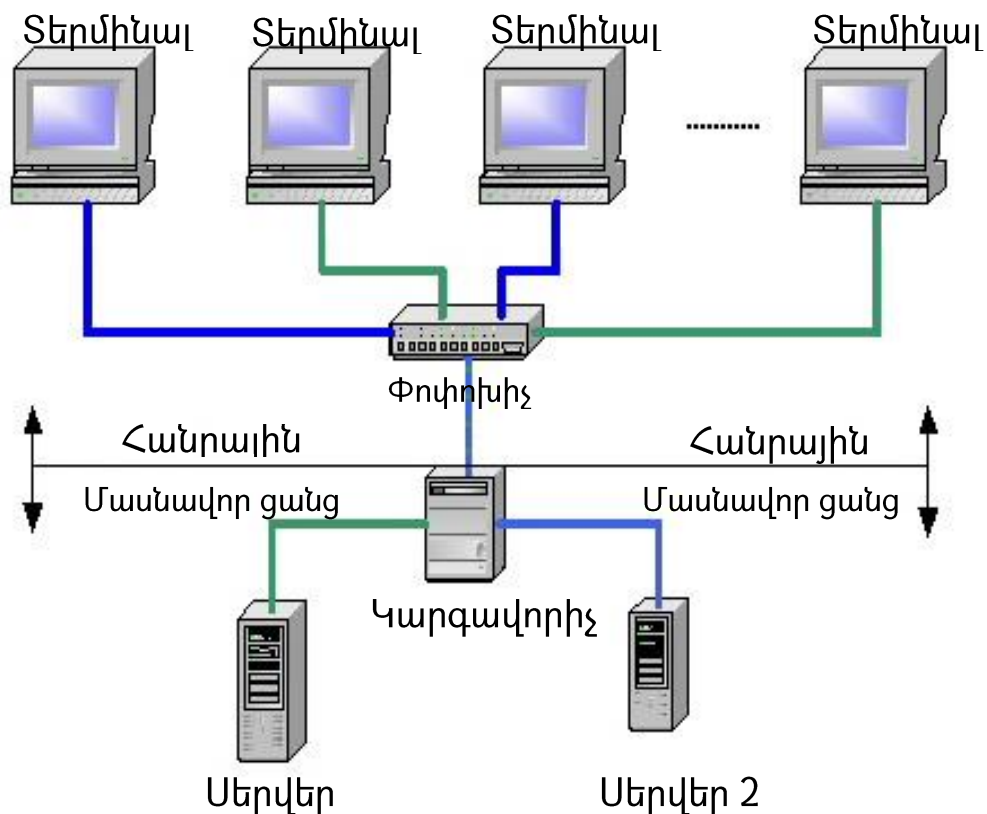
MPI-ը զուգահեռ համակարգիչների ծրագրերում օգտագործվող ծրագրավորման լեզվից անկախ հաղորդագրությունների արձանագրություն է: Այն ապահովում է և՛ կետ առ կետ, և՛ կոլեկտիվ կապերը: MPI-ի նպատակն է՝ բարձր արտադրողականությունը, մասշտաբայնեցումը և տեղափոխելիությունը: Այսօր MPI-ը հանդիսանում է գերիշխող մոդել բարձր արտադրողականությամբ հաշվարկներում:

Ի սկզբանե MPI-ը նախագծված էր բաշխված հիշողությամբ ճարտարապետությունների համար, որը ժամանակի ընթացքում դառնում է ավելի ու ավելի հանրահայտ: Քանի որ ճարտարապետության միտումները փոխվել են, ընդհանուր հիշողությամբ SMP-ները միավորվել են ցանցերով՝ ստեղծելով հիբրիդային բաշխված հիշողություն/ընդհանուր հիշողության համակարգեր: MPI-ի իրականացնողները հարմարեցրել են իրենց գրադարանները, որպեսզի ապահովեն վերը նշված հիշողության ճարտարապետությունները: Նրանք նաև մշակել են ճանապարհներ՝ ապահովելու համար տարբեր միջկապեր և արձանագրություններ:

1.2.2 Կլաստերային համակարգեր

Կլաստերը բաղկացած է մի շարք իրար միացված համակարգիչներից, որոնք աշխատում են միասին և ավելի հաճախ դիտարկվում են որպես մեկ համակարգ (նկ. 1.3) [9]: Կլաստերում յուրաքանչյուր հանգույց կարգավորված է նույն գործողությունը կատարելու համար և ղեկավարվում է ծրագրային ապահովման միջոցով: Կլաստերի բաղադրիչները սովորաբար միացվում են լոկալ ցանցերի միջոցով, որտեղ յուրաքանչյուր հանգույց կամ համակարգիչ գործարկում է օպերացիոն համակարգից ստացած իր հրահանգը: Մեծամասամբ բոլոր հանգույցները օգտագործում են միևնույն ապարատային ապահովումը և միևնույն օպերացիոն համակարգը, այնուամենայնիվ, որոշ կարգաբերումներում հնարավոր է օգտագործվեն տարբեր ապարատային ապահովումներ և/կամ օպերացիոն համակարգեր յուրաքանչյուր համակարգչի վրա: Կլաստերները ստեղծվել են, որպեսզի բարելավեն մեկ համակարգչի արագագործությունը և հուսալիությունը, միևնույն ժամանակ, որպես կանոն, ավելի ցածր գին են ունենում, քան նույն արագագործությունն ու հուսալիությունն ունեցող հզոր համակարգիչը: Դրանք առաջացել են մի շարք հաշվարկային տենդենցների

համապատասխանության արդյունքում, ներառյալ՝ էժանագին միկրոպրոցեսորների, գերարագ ցանցերի, բարձր արտադրողականություն ունեցող համակարգերի համար ծրագրային ապահովման հասանելիությունը: Կլաստերները լայն կիրառություն ունեն աշխարհում, սկսած փոքր բիզնեսի կլաստերներից մինչև գերարագ սուպերհամակարգիչները, ինչպիսին է՝ IBM Sequoia: Կլաստերների լայն կիրառությունը, այնուամենայնիվ, կարող է սահմանափակ լինել կախված նրանից, որ ծրագրային ապահովումը պետք է մշակված և ներդրված լինի որպես մեկ խնդիր: Հետևաբար, հնարավոր չէ պատահական հաշվողական առաջադրանքներ կատարելու նպատակով օգտագործել կլաստերային համակարգերը:



Նկար 1.3. Կլաստերի կառուցվածքը:

Կլաստերային համակարգերը կարելի է կարգաբերել տարբեր նպատակներով, սկսած հիմնական բիզնես նպատակներից՝ վեբ ծառայությունների սպասարկման, մինչև գիտական գերմեծ հաշվարկներ: Ամեն դեպքում, կլաստերները կարելի է օգտագործել

նրանց բարձր հուսալիության պատճառով: Ստորև բերված են կլաստերներին յուրահատուկ երկու հատկությունները.

- «Բեռնման հավասարակշռում» հատկությամբ կլաստերներում հավասարակշռվում են հանգույցների հաշվողական ծանրաբեռնվածությունները միմյանց միջև և, այդպիսով, ավելանում է կլաստերի արտադրողականությունը: Օրինակ, վեբ սերվերը կարող է տարբեր հարցումները նշանակել տարբեր հանգույցների, այդպիսով, օպտիմիզացնելով հարցմանն արձագանքելու ընդհանուր ժամանակը:
- «Բարձր հուսալիությամբ կլաստերները» բարձրացնում են կլաստերային հուսալիության մոտեցումը: Դրանք գործում են ունենալով կրճատված հանգույցներ, որոնք հետո օգտագործվում են, երբ համակարգի կոմպոնենտները շարքից դուրս են գալիս:

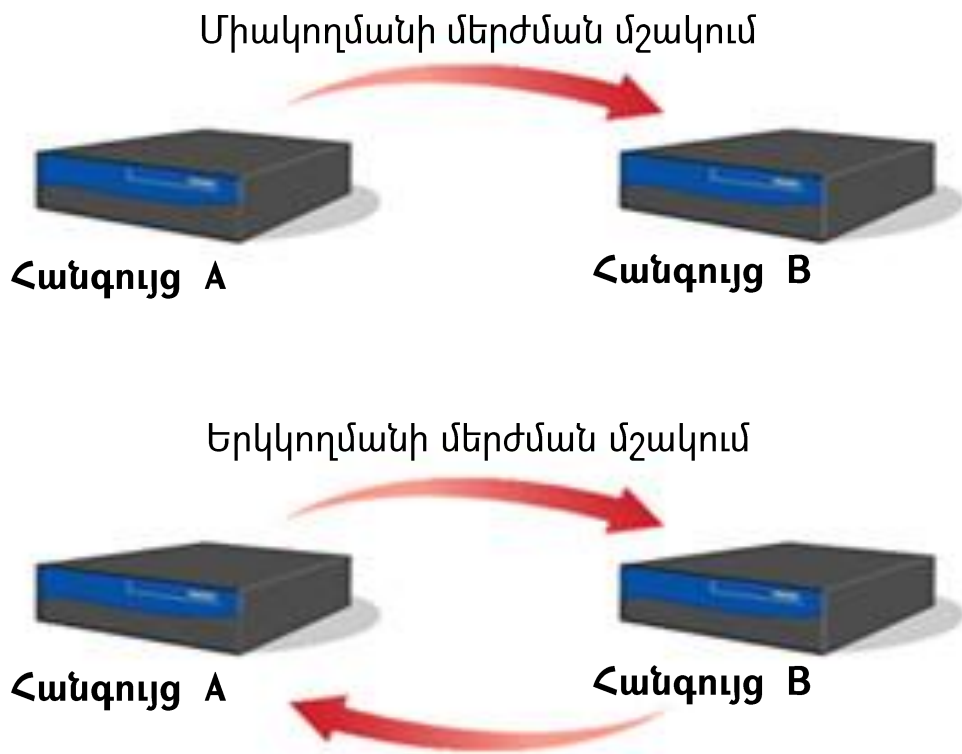
Կլաստերները հիմնականում դասակարգվում են ըստ բարձր հասանելիության, ծանրաբեռնվածության հավասարակշռության և բարձր արտադրողականության:

Բարձր հասանելիության կլաստերը հիմնականում օգտագործվում է բարելավելու համար կլաստերի կողմից տրամադրված ծառայության հասանելիությունը (նկ. 1.4): Նրանք գործում են ունենալով լրացուցիչ հանգույցներ, որոնք նախատեսված են փոխարինելու այն հանգույցներին, որոնք կարող են շարքից դուրս գալ աշխատանքի ընթացքում: Հաճախ օգտագործվում է կրիտիկական տվյալների բազաների, ցանցային ֆայլերի փոխանակման և բիզնես ծրագրերի համար:

Ծանրաբեռնվածության հավասարակշռության կլաստերի դեպքում մի քանի համակարգիչներ կապվում են, որպեսզի կիսեն հաշվողական ծանրաբեռնվածությունը: Տրամաբանորեն նրանք մի քանի համակարգիչ են, բայց գործում են որպես միասնական վիրտուալ համակարգիչ: Օգտագործողի կողմից կատարված հարցումը բաշխվում է բոլոր հանգույցներով մեկ կամ մի քանի հավասարակշռողներով:

Բարձր արտադրողականությամբ կլաստերներում հաշվողական առաջադրանքը բաժանվում է տարբեր հանգույցների միջև, որի շնորհիվ բարձրանում է արտադրողականությունը: Հիմնականում օգտագործվում է գիտական հաշվարկներում:

Այս կլաստերի հայտնի իրականացումներից են Linux օպերացիոն համակարգի հետ աշխատող հանգույցները և զուգահեռացում իրականացնող ծրագրերը:



Նկար 1.4. Բարձր հասանելիության կլաստեր:

Կլաստերի այլ կարևոր բաղադրիչներից է տեղային պլանավորող ծրագիրը, որը ղեկավարում է կլաստերի հանգույցների ծանրաբեռնվածության հավասարակշռությունը, և զուգահեռ ծրագրային միջավայրը, ինչպիսիք են՝ MPI-ը կամ PVM-ը, որոնք թույլ են տալիս կլաստերային ծրագրերին լինել շարժական կլաստերների լայն դասի համար:

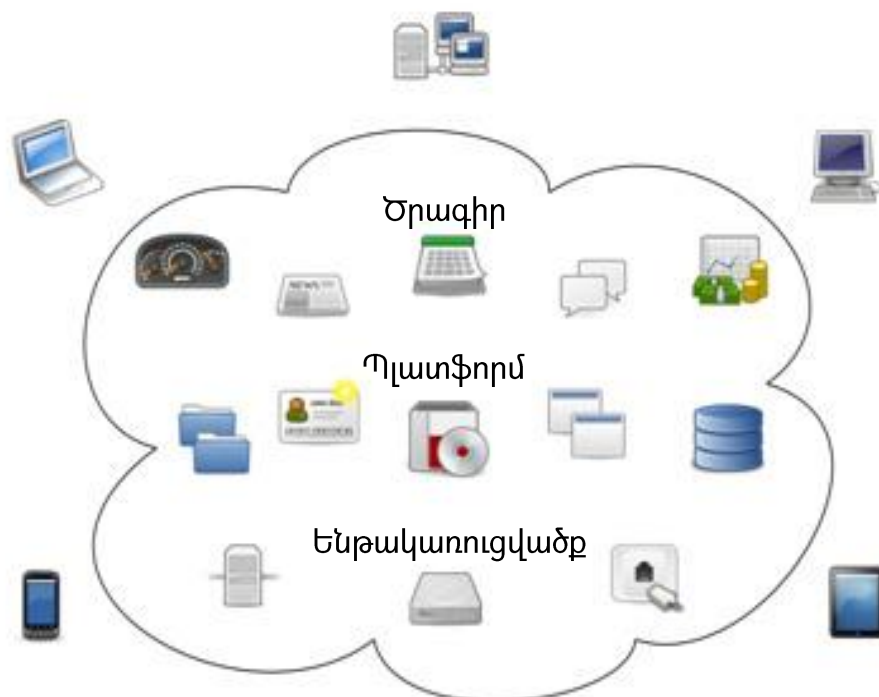
Կլաստերային համակարգերի առավելություններն են.

- Կառավարելիություն – շատ ջանքեր և ծախսեր են պահանջվում կառավարելու համար մեծ թվով բաղադրիչներ: Սակայն կլաստերի օգնությամբ բազմաթիվ բաղադրիչներ աշխատում են համակցված, որպես մեկ միավոր, որը կառավարումը դարձնում է ավելի դյուրին:
- Միասնական համակարգ – օգտագործողին թվում է, թե նա աշխատում է մեկ համակարգի հետ, սակայն իրականում նա աշխատում է մի քանի բաղադրիչների հետ, որոնց մասին նա կարող է չանհանգստանալ:

- Բարձր մատչելիություն – քանի որ բաղադրիչները հանդիսանում են մեկը մյուսի պատճեն, ապա տեխնիկական խնդիրների պատճառով շարքից դուրս եկած բաղադրիչը փոխարինվում է մեկ ուրիշով, որի շնորհիվ օգտագործողի աշխատանքը չի խափանվում:

1.2.3 Ամպային հաշվարկներ

Ամպային հաշվողական համակարգը նոր համակարգչային պարադիգմ է, որն ըստ պահանջի ապահովում է դինամիկ ընդլայնվող և վիրտուալ ռեսուրսներ (նկ. 1.5) [10]:



Նկար 1.5. Ամպային համակարգեր:

Այս համակարգերի հիմնական սկզբունքն է առաջարկել հաշվողական համակարգերի, պահեստավորումների և ծրագրերի ծառայություններ: Այս ծառայություններից օգտվելու համար անհրաժեշտ է միայն ինտերնետ: Օրինակ, Բույան սահմանել է այս համակարգը հետևյալ ձևով. «Ամպը զուգահեռ և բաշխված հաշվողական համակարգ է՝ բաղկացած փոխկապակցված և վիրտուալ համակարգիչներից, որոնք ներկայացվում են որպես մեկ կամ ավելի միասնական և

դինամիկ հաշվողական ռեսուրսներ՝ համաձայն ծառայություն մատուցողի և օգտագործողի միջև կնքված ծառայության մատուցման համաձայնագրի»:

Նկար 1.5-ում ներկայացված է, թե ինչպես կարելի է միանալ ամպային ծառայություններին օգտագործելով ցանկացած սարք, որն ունի ինտերնետի հնարավորություն: Ամպային ենթակառուցվածքը ներառում է ընդլայնվող ռեսուրսներ պահեստում, ցանց, հաշվարկներ և վիրտուալ ենթակառուցվածքներ: Ամպային հաշվողական համակարգերը կրճատում են գործառնական և կապիտալ ծախսերը և թույլ են տալիս SS գերատեսչություններին ավելի կենտրոնանալ ծրագրերի ռազմավարության վրա: Այն տրամադրում է ծառայություններ ենթակառուցվածքի, հարթակի և ծրագրային ապահովման մակարդակներում: Այն տրամադրում է նաև այնպիսի առանձնահատկություններ, ինչպիսիք են՝ արագությունը, ռեսուրսների ընդլայնումը, զուգահեռ մշակումը, միայն օգտագործած ռեսուրսի դիմաց վճարումը, ցանկացած ժամանակ տեխնոլոգիայի փոփոխությունը, ծառայության շուրջօրյա հասանելիությունը, հուսալիությունը, անվտանգությունը և այլն:

Ամպային համակարգերն ունեն հինգ հիմնական հատկանիշները՝ արագ առաձգականություն, չափված ծառայություններ, ինքնասպասարկում անհրաժեշտության դեպքում, ռեսուրսների միաձուլում և ցանցային մուտք:

Ամպային համակարգերը առաջարկում են իրենց առավելությունները ծառայությունների երեք տեսակների միջոցով.

- Ենթակառուցվածքը որպես ծառայություն (IaaS),
- Հարթակը որպես ծառայություն (PaaS),
- Ծրագրային ապահովումը որպես ծառայություն (SaaS):

Այն նաև տրամադրում է իր ծառայությունները չորս տեղակայման մոդելների միջոցով, որոնք են՝ հասարակական ամպ, մասնավոր ամպ, հիբրիդային ամպ և համայնքային ամպ:

- Հասարակական ամպը տեղակայման մոդել է, որում ենթակառուցվածքը հասանելի է ընդհանուր հասարակությանը և այն օգտագործում է վճարման pay-as-you-go մոդելը: Ամպային ռեսուրսները հասանելի են ինտերնետի միջոցով, և

տրամադրողը պատասխանատվություն է կրում ապահովելու մասշտաբայնեցման տնտեսությունն ու համատեղ ենթակառուցվածքի կառավարումը: Այս մոդելում հաճախորդները կարող են ընտրել անվտանգության մակարդակը և բանակցել սպասարկման մակարդակի համար: Այս մոդելի ամպերից է Amazon Web Services EC2-ը:

- Մասնավոր ամպը տեղակայման այլ մոդել է ամպային ծառայությունների համար: Այս մոդելում ամպը տեղյակ չէ երրորդ կողմից և չի կիսում ռեսուրսները: Ամպային ռեսուրսները կարող են տեղակայված լինել հաճախորդի կազմակերպության տարածքում: Այս մոդելը չի ազդում օգտագործողի անվտանգության և համապատասխանության պահանջների վրա, սակայն այն կազմակերպելու համար պահանջվում է կատարել SS ենթակառուցվածքային ներդրումների կապիտալ ծախսեր: Այսպիսի ամպային միջավայրում հասարակությունը չունի մուտք մասնավոր ամպային միջավայր:
- Հիբրիդային ամպը տեղակայման մոդել է, որը խմբավորում է տարբեր տեսակի ամպեր, օրինակ՝ մասնավոր և հասարակական: Այս մոդելում համախմբված ամպերը պահպանում են իրենց ինքնությունը, բայց ստանդարտեցված տեխնոլոգիայով աշխատում են միասին: Ամպային այս մոդելում ընդհանուր հանրությունը ամպին մուտք չունի, բայց կազմակերպությունը օգտագործում է ենթակառուցվածք և՛ հասարակական, և՛ մասնավոր ամպերում:
- Համայնքային ամպը տեղակայման չորրորդ մոդելն է, որը կարող է օգտագործվել ամպային ծառայություններ տրամադրելու համար: Այս մոդելում ամպային բազմաթիվ կազմակերպություններ կամ հաստատություններ են, որոնք ունեն ընդհանուր հետաքրքրություններ կամ մտահոգություններ, ինչպիսիք են՝ համապատասխանության նկատառումները, անվտանգության պահանջները և բաժնետոմսերի ենթակառուցվածքները: Այս տեսակի ամպերը կարող են կառավարվել կազմակերպությունների կամ երրորդ կողմի կողմից և կարող են գտնվել տարածքից ներս կամ դուրս: Հանրությունները և համայնք կազմող կազմակերպությունները կարող են օգտվել համայնքային ամպերի տրամադրած ծառայություններից:

Ամպային հաշվողական համակարգի առավելություններն են.

- Ընդհանուր ռեսուրսներ - Ամպային հաշվողական համակարգը ընդհանրացնում է ռեսուրսները, որպեսզի իր ծառայությունները ապահովի բազմաթիվ օգտագործողների: Այդ իսկ պատճառով, այն հեշտությամբ կարող է տրամադրել ծրագիր, որը հնարավորություն է տալիս ըստ պահանջի մեծացնել կամ փոքրացնել ռեսուրսները:
- Pay-As-You-Go - Օգտագործողը վճարում է միայն այն ռեսուրսների համար, որոնք օգտագործում է: Հետագայում նա կարող է ավելացնել կամ պակասեցնել ռեսուրսները ըստ պահանջի:
- Ավելի լավ ապարատային կառավարում - Քանի որ բոլոր համակարգիչները օգտագործում են միևնույն սարքավորումները, ամպային ծառայությունը հեշտությամբ կարողանում է կառավարել այն:

1.3 Տվյալների վերլուծություն

Տվյալների վերլուծությունը գործընթաց է, որն օգտագործվում է տվյալները ստուգելու, մաքրելու, վերափոխելու և ձևափոխելու համար՝ տվյալ իրավիճակում որոշակի եզրակացության հանգելու համար: Այն սովորաբար լինում է երկու տեսակի՝ որակական և քանակական: Տվյալների տեսակը թելադրում է վերլուծության մեթոդը: Որակական հետազոտությունում վերլուծվում են տեքստի կամ անհատական բառերի նման ոչ թվային տվյալները: Քանակական վերլուծությունը, մյուս կողմից, կենտրոնանում է տվյալների չափման վրա և կարող է օգտագործել վիճակագրությունը՝ օգնելու համար արդյունքների և եզրակացությունների բացահայտմանը: Ստացված արդյունքները թվային են և որոշ դեպքերում վերլուծության երկու ձևերը կիրառվում են միասին: Օրինակ, քանակական վերլուծությունը կարող է օգնել որակյալ եզրակացությունների հաստատմանը:

Ստորև ներկայացված են տվյալների վերլուծության որոշ առավելություններ.

- Տվյալների վերլուծությունը օգնում է տարբեր աղբյուրներից ստացված տվյալների հայտնաբերումների կառուցվածքը ստանալուն:

- Տվյալների վերլուծությունը շատ օգտակար է մեծ խնդիրները փոքր մասերի բաժանելու համար:
- Տվյալների վերլուծությունը գործում է որպես ֆիլտր, երբ խոսքը վերաբերում է հսկայական տվյալների հավաքածուների իմաստալից հասկացությունների ձեռքբերմանը:
- Տվյալների վերլուծությունը վիճակագրական մշակման միջոցով օգնում է մարդկային կողմնորոշմանը չշեղվել հետազոտության արդյունքներից:

Տվյալների վերլուծության քննարկման ժամանակ կարևոր է նշել, որ անհրաժեշտ է ընտրել այդ տվյալների վերլուծության համար համապատասխան մեթոդաբանությունը, հակառակ դեպքում՝ տվյալները չեն կարող հավաքագրվել և վերլուծվել:

Ստրուկտուրավորված կոչվում են այն տվյալները, որոնք կարգավորված և կազմակերպված են որոշակի ձևով՝ նրանց նկատմամբ որոշ գործողությունների կիրառման հնարավորություն ապահովելու նպատակով: Ոչ ստրուկտուրավորված տվյալների վերլուծությունը վերաբերում է տվյալների օբյեկտների վերլուծման այն գործընթացին, որը չի հետևում կանխորոշված տվյալների մոդելին/ճարտարապետությանը և/կամ անկազմակերպ է: Սովորաբար այն ներառում է վերլուծություն յուրաքանչյուր տվյալների օբյեկտի համար, որը տվյալների բազայում պահված չէ: Դրանք ներառում են փաստաթղթեր, մեդիա ֆայլեր, պատկերներ և այլն: Այս տեսակի տվյալների վերլուծությունը սովորաբար կատարվում է ինֆորմացիայի, թաքնված մտքերի, տվյալների խմբերի միջև փոխհարաբերությունների հայտնաբերման համար: Տվյալների պեղումը ներկայացնում է հետաքրքիր և օգտակար ձևանմուշների, ինչպես նաև մեծածավալ տվյալներում փոխհարաբերությունների հայտնաբերման գործընթաց: Այն մեծ չափի թվային հավաքածուների վերլուծության համար համախմբում է վիճակագրական և արհեստական բանականության (օրինակ, նեյրոնային ցանցեր և մեքենայական ուսուցում) գործիքները տվյալների բազաների կառավարման հետ: Այն ներառում է խնդիրների 6 ընդհանուր դաս [11].

- Անոմալիաների հայտնաբերում – Անսովոր տվյալների հայտնաբերում, որոնք կարող են հետաքրքիր լինել, կամ տվյալների սխալներ, որոնք պահանջում են հետագա վերանայում:
- Ասոցիացիայի կանոնների ուսուցում – Փոփոխականների միջև փոխհարաբերությունների որոնում:
- Կլաստերացում – Առանց տվյալներում հայտնի կառուցվածքների օգտագործմամբ բացահայտում է տվյալներում այնպիսի խմբեր և կառուցվածքներ, որոնք ինչ-որ ձևով «նման» են:
- Դասակարգում – Նոր տվյալներում կիրառելու համար հայտնի կառուցվածքների ընդհանրացում:
- Ռեգրեսիա – Նվազագույն սխալանքով տվյալների մոդելավորման համար գործառնությունների հայտնաբերում՝ տվյալների կամ տվյալների հավաքածուների միջև հարաբերությունները գնահատելու համար:
- Ընդհանրացում – Տվյալների հավաքածուների ավելի ամփոփ ներկայացում՝ ներառյալ վիզուալիզացիա և զեկույցների ձևավորում:

1.3.1 Մեծ տվյալներ

Մեծ տվյալները նոր մարտահրավեր է էլեկտրոնային արդյունաբերության համար: Վերջին տարիների ընթացքում «Մեծ Տվյալներ» տերմինը օգտագործվել է տարբեր խոշոր «խաղացողների» կողմից՝ տարբեր հատկանիշներով տվյալների բնութագրման համար [12]: META Group-ի (այնուհետև Gartner-ի) վերլուծաբան Դուգ Լանեյի կողմից մեծ տվյալները սահմանվել է երեք հատկանիշների միջոցով (ծավալ, արագություն, բազմազանություն). «Մեծ տվյալները մեծ ծավալով, բարձր արագությամբ և/կամ մեծ բազմազանությամբ ինֆորմացիայի բազմություն է, որում որոշումների կայացման, իմաստների հայտնաբերման և գործընթացների օպտիմալացման համար պահանջվում են մշակման նոր ձևեր»: Ժամանակի ընթացքում ավելացել են մեծ տվյալները բնութագրող ևս երկու հատկանիշներ՝ ճշմարտացիություն և արժեք:

Մեծ տվյալները համատեղում են տվյալների կառավարման մարտահրավերները տվյալների հետ աշխատելու համար՝ չափերի և բարդության նոր մասշտաբով: Այս խնդիրներից շատերը նոր չեն, սակայն նոր են 5V-ի հետ կապված մեծ տվյալների առանձնահատկությունների հետ կապված խնդիրները [13] [14] (նկ. 1.6):

Ծավալ (Volume)	Վերաբերում է յուրաքանչյուր վայրկյան գեներացվող մեծ քանակի տվյալների և ղեկավարվում է մեծ տվյալների համակարգի բաշխվածության շնորհիվ:
Բազմազանություն (Variety)	Վերաբերում է գեներացվող տվյալների տարբեր տիպերին, որոնք կարող են վերլուծվել մեծ տվյալների տեխնոլոգիաների շնորհիվ:
Արագություն (Velocity)	Վերաբերում է նոր տվյալների ստեղծման և նրանց ղեկավարելու արագությանը:
Ճշմարտացիություն (Veracity)	Վերաբերում է տվյալների որակին և հուսալիությանը: Գոյություն ունեն տվյալների տարբեր աղբյուրներ, և կախված տվյալների բնույթից, արդյունքների կարևորությունը տարբեր է:
Արժեք (Value)	Վերաբերում է մեծ տվյալների հիմնական նպատակին՝ ստանալ արժեք տվյալներից, գեներացնել ինֆորմացիայից գիտելիք:

Նկար 1.6. Մեծ Տվյալներ: 5 V-եր:

Կազմակերպությունների, տեխնոլոգիաների և այլ սոցիալական ցանցերի զարգացումը հանգեցրել է մեծ ծավալի ինֆորմացիայի առաջացման, որոնք չեն կարող կառավարվել ավանդական տվյալների բազաների միջոցով: Մեկ միավոր տվյալների հավաքածուի ծավալը կարող է հասնել տերաբայթերից մինչև էքսաբայթերի: Ահռելի քանակությամբ տվյալներից նշանակալի ինֆորմացիայի հայտնաբերումը բավականին բարդ խնդիր է, սակայն գոյություն ունեն տիրույթներ և տնտեսական ոլորտներ [15], ինչպիսիք են՝ ֆիզիկական գիտությունները, բժշկական կրթությունը, կառավարությունը, բանկային և ֆինանսական ծառայությունները, որոնք մեծ տվյալներ օգտագործելով ստանում են առավելություններ: Մեծ տվյալների հիմնական աղբյուրներից են հանդիսանում կազմակերպությունների ներսում գեներացված տվյալները, ինչպիսիք են՝ էլեկտրոնային նամակները, փաստաթղթերը, ստրուկտուրավորված և ոչ

ստրուկտուրավորված տվյալները, կազմակերպությունից դուրս առկա հանրային տվյալները: Մեծ տվյալները ստրուկտուրավորված և ոչ ստրուկտուրավորված տվյալներից բաղկացած հսկայական հավաքածու է, և չնայած նրան, որ ստրուկտուրավորված տվյալների համար գոյություն ունեն բազմաթիվ ալգորիթմներ և պրոցեսորաներ, ոչ ստրուկտուրավորված տվյալների համար անհրաժեշտ է լինում կիրառել մեծ տվյալների վերլուծություններ:

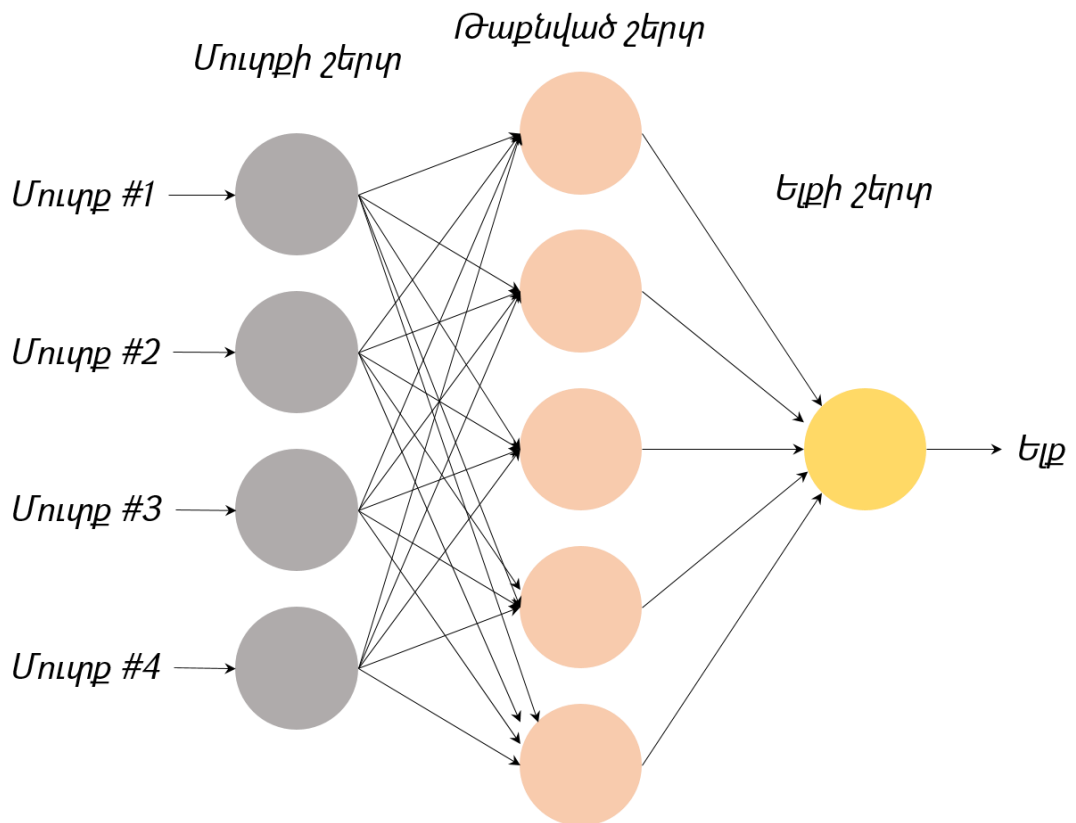
1.3.2 Մեքենայական ուսուցում: Նեյրոնային ցանցեր

Մեքենայական ուսուցումն արհեստական բանականության ճյուղ է, որտեղ համակարգը ստեղծվում է այն տվյալների հիման վրա, որոնք կարող են օգտագործվել կանխատեսումների, ստանդարտ ճանաչման և մուտքագրման տվյալների դասակարգման համար [16]: Մեքենայական ուսուցումը նաև սերտ կապված է վիճակագրության և օպտիմիզացիայի հետ: Այն կիրառելի է այն բնագավառներում, որտեղ բացահայտ ծրագրավորված ալգորիթմները անօգուտ են: Տվյալների դասակարգման ենթաբաժնում տարբեր մեքենաների ուսումնառության մոտեցումների շարքում հայտնաբերվել են նեյրոնային ցանցերի մեթոդներ, որոնք օգտագործվում են որպես վիճակագրական մեթոդների այլընտրանքային տարբերակ:

Տվյալների վերլուծության ոլորտում մեքենայական ուսուցումն այն մեթոդն է, որն օգտագործվում է բարդ մոդելների և ալգորիթմների մշակման համար, որոնք ենթարկվում են կանխատեսման: Այս վերլուծական մոդելները թույլ են տալիս հետազոտողներին, գիտնականներին, ճարտարագետներին և վերլուծաբաններին «արտադրել հուսալի, կրկնվող որոշումներ և արդյունքներ» և պատմական տվյալների հարաբերություններից ու միտումներից սովորելու միջոցով բացահայտել «թաքնված պատկերացումները»:

Արհեստական նեյրոնային ցանցը (ANN – Artificial Neural Network) տեղեկատվության մշակման պարադիգմ է, որը նման է կենսաբանական նեյրոնային համակարգերին, ինչպիսիք են ուղեղը, պրոցեսային տեղեկատվությունը [17]: Արհեստական նեյրոնային ցանցը մաթեմատիկական մոդել է և օգտագործվում է մուտքային տվյալների և արդյունքների միջև բարդ հարաբերությունների մոդելավորման

կամ տվյալների օրինակների հայտնաբերման համար: Այս պարադիգմի հիմնական տարրը տեղեկատվության մշակման համակարգի նորագույն կառուցվածքն է: Այն բաղկացած է միանգամայն փոխկապակցված վերամշակող տարրերի (նեյրոնների) մեծ թվից, որոնք խնդիրների լուծման համար աշխատում են միասնական: ANN-ները, ինչպես մարդիկ, սովորում են օրինակով: ANN-ն կարգավորվում է կոնկրետ խնդրի համար, օրինակ, օրինակների ճանաչում կամ տվյալների դասակարգում՝ ուսումնական գործընթացի միջոցով: Ինչպես կենսաբանական համակարգերում, այնպես էլ ANN-ում ուսուցումը ներառում է նեյրոնների միջև գոյություն ունեցող սինապտիկ կապերի ճշգրտում:



Նկար 1.7. Պարզ նեյրոնային ցանց:

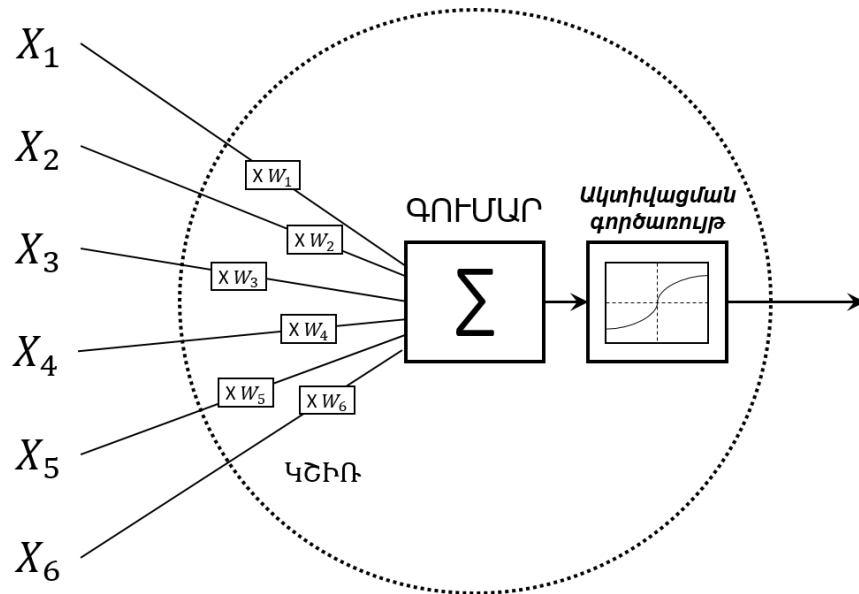
Նկար 1.7-ում ներկայացված է պարզ նեյրոնային ցանցի ճարտարապետությունը: Այն կազմված է մուտքի, ելքի և մեկ կամ մի քանի թաքնված շերտերից: Մուտքի շերտի յուրաքանչյուր հանգույց կապված է թաքնված շերտի հանգույցի հետ և, որի հանգույցներն էլ կապված են ելքի շերտի հանգույցների հետ: Սովորաբար յուրաքանչյուր կապի հետ ասոցացվում է կշիռ: Մուտքի շերտը ներկայացնում է ցանցին տրված

չմշակված ինֆորմացիան և այս շերտում արժեքները երբեք չեն փոխվում: Ցանցին տրված յուրաքանչյուր մուտքային տվյալներ կրկնօրինակվում է և ուղարկվում է թաքնված շերտում գտնվող հանգույցներին, որոնք օգտագործելով որոշակի կշիռներ փոփոխում են այդ տվյալները: Այնուհետև այդ նոր տվյալները փոխանցվում են ելքի շերտ, որտեղ նույնպես կատարվում են փոփոխություններ՝ օգտագործելով թաքնված և ելքի շերտերի միջև կապերի կշիռները, և արտադրում է արդյունքը, որը մշակվում է ակտիվացման գործառույթի միջոցով:

Յուրաքանչյուր շերտում հանգույցների թվի ընտրությունը կախված է խնդրից, որը ներյոնային ցանցը փորձում է լուծել: Մուտքի և ելքի հանգույցների քանակը կախված է ուսուցման տվյալների հավաքածուներից: Թաքնված շերտի հանգույցների օպտիմալ քանակի ընտրությունը հանդիսանում է բարդ խնդիր, քանի որ շատ հանգույցների դեպքում ալգորիթմում հնարավոր հաշվարկների քանակը աճում է, իսկ քիչ հանգույցները կարող են հանգեցնել ցանցի սխալ ուսուցման:

Ներյոնային ցանցի կառավարման եղանակը հանգույցների միջև կշիռների կարգավորումն ու ճշգրտումն է: Նախնական կշիռները սովորաբար սահմանվում են որոշակի պատահական թվերով, և այնուհետ ճշգրտվում են ներյոնային ցանցի ուսուցման ընթացքում: Ուսուցման յուրաքանչյուր քայլից հետո կշիռները թարմացվում են նոր արժեքներով: Եթե թարմացումից հետո ներյոնային ցանցի արդյունքը ավելի լավն է, քան նախորդ կշիռների խմբի դեպքում, ապա նոր արժեքները պահպանվում են և շարունակվում է քայլերի կատարումը: Կշիռների որոշման ժամանակ կարևոր նպատակը պետք է հանդիսանա գտնել այնպիսի կշիռների հավաքածու, որոնց դեպքում ստացված արդյունքի սխալը կլինի նվազագույնը:

Ներյոնային ցանցի թաքնված շերտի հանգույցներում ակտիվացման գործառույթները օգտագործվում են ոչ գծայնության ապահովման համար: Ակտիվացման գործառույթը կարող է լինել, օրինակ՝ գծային, սահմանային կամ սիգմոիդ գործառույթներից որևէ մեկը: Նկար 1.8-ում նկարագրված է ներյոնային ցանցում թաքնված շերտի հանգույցների գործառույթը:



Նկար 1.8. Նեյրոնային ցանցում թաքնված շերտի հանգույցի գործառույթ:

Հանգույցում մուտքային տվյալները բազմապատկվում են համապատասխան կողի կշիռով և մեկ թիվ ստանալու համար բոլոր արժեքները գումարվում են և ստացված արժեքը օգտագործվում է ակտիվացման գործառույթում:

1.4 Տվյալների ինտերակտիվ վիզուալիզացիա

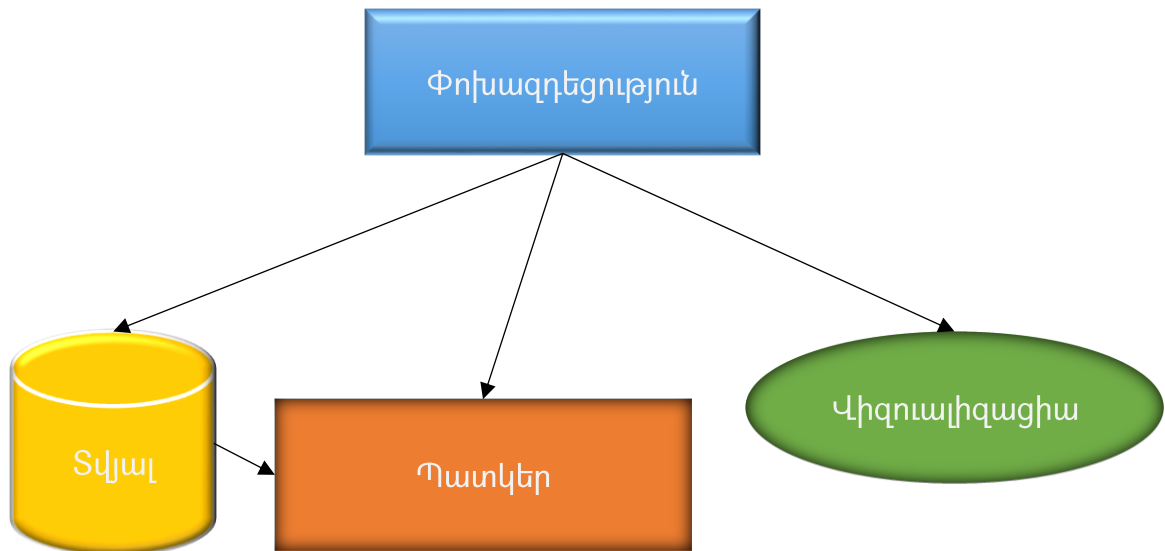
Գիտական տվյալների վիզուալիզացիան հանդիսանում է գիտության զարգացման և ընկալման առանցքային մասը: Այն հանդիսանում է նաև համակարգչային գրաֆիկայի ենթախումբ և համակարգչային գիտության ճյուղ: Գիտական տվյալների վիզուալիզացիայի նպատակն է գրաֆիկական պատկերացում տալ գիտնականներին, որպեսզի ավելի դյուրին լինի հասկանալ, նկարագրել և ստանալ անհրաժեշտ ինֆորմացիան իրենց ունեցած տվյալներից [18]:

Տվյալների վիզուալիզացիայի հիմնական նպատակն է՝ հստակ, արդյունավետ ձևով հաղորդել ինֆորմացիան վիճակագրական և ինֆորմացիոն գրաֆիկների ու ինֆորմացիոն պատկերների միջոցով: Թվային ինֆորմացիան կարող է կոդավորվել՝ օգտագործելով կետեր, գծեր, պատկերներ, որպեսզի հնարավոր լինի ավելի պատկերավոր ձևով իրականացնել այդ տվյալների վիզուալիզացիան [19]: Այն հնարավորություն է տալիս բարդ տվյալները դարձնել ավելի հասկանալի, հասանելի և օգտագործելի:

Ինտերակտիվ վիզուալիզացիայի ոլորտը ուսումնասիրում է, թե ինչպես են մարդիկ համագործակցում համակարգիչների հետ՝ ստեղծելու համար ինֆորմացիայի վիզուալ ներկայացում և, թե ինչպես այդ գործընթացը կարելի է դարձնել ավելի արդյունավետ: Վիզուալիզացիայի տեխնոլոգիաները հնարավորություն են տալիս օգտագործողին ընկալել մեծ քանակությանը տվյալներից կարևոր բաղադրիչները, գտնել անհրաժեշտ բաժինները, որոնք կնպաստեն ճշգրիտ որոշումների կայացմանը: Այնուամենայնիվ, առանց ինտերակտիվ բաղադրիչի վիզուալիզացիան հաճախ դիտվում է որպես աշխատանքի արդյունքի վերջնական կետ կամ որպես դիտարկումների համագործակցության եղանակ: Օգտագործողներին անհրաժեշտ է մանիպուլացնել և ուսումնասիրել տվյալները: Վիզուալիզացիաների հետ մարդկային ընկալման ու փոխազդեցության ձևերը և վիզուալիզացիայի համակարգի օգտակարությունը կարող են մեծապես ազդել տվյալ տեղեկատվության հասկացության վրա: Բացի այդ, որոշ ոլորտներում դեռ պարզ չէ, թե որոնք են, ինչպիսիք են հետաքրքրությունների առանձնահատկությունները և ինչպես սահմանել դրանք, որպեսզի հնարավոր լինի դրանց հայտնաբերել: Արդյունքում օգտագործողների փոխազդեցությունը դառնում է ավելի կարևոր:

Օգտագործողի տվյալները ուսումնասիրելու և ավելի լավ հասկանալու կարողությունները մեծացնելու համար, լայն հաշվարկների փոխազդեցության և վիզուալիզացիայի հնարավորությունների վրա հիմնված փորձերի արդյունքները պետք է օպտիմալացված լինեն այնպես, որ տվյալների հասանելիությունը և առընչվող գործառնությունները լինեն ակնհայտ [20]: Վիզուալիզացիայի և փոխգործակցության տեխնոլոգիաների արդյունավետ ինտեգրումը գիտական հաշվարկների ոլորտում դարձել է շատ կարևոր:

Ինտերակտիվ վիզուալիզացիան թույլ է տալիս ավելի սերտ կապ հաստատել հետազոտողների, նրանց մոդելների և ուսումնասիրվող տվյալների միջև: Ինտերակտիվ վիզուալիզացիան ոչ միայն տրամադրում է տվյալները դիտարկելու գործիք, այլ նաև թույլ է տալիս՝ օգտագործելով փոխազդեցության հնարավորությունը, շրջել տվյալների միջով և փոխանակվել այդ գաղափարներով: Նկար 1.9-ում ցուցադրված է այս գործընթացի բազային մոդելը:



Նկար 1.9. Ինտերակտիվ վիզուալիզացիայի գործընթացի բազային մոդելը:

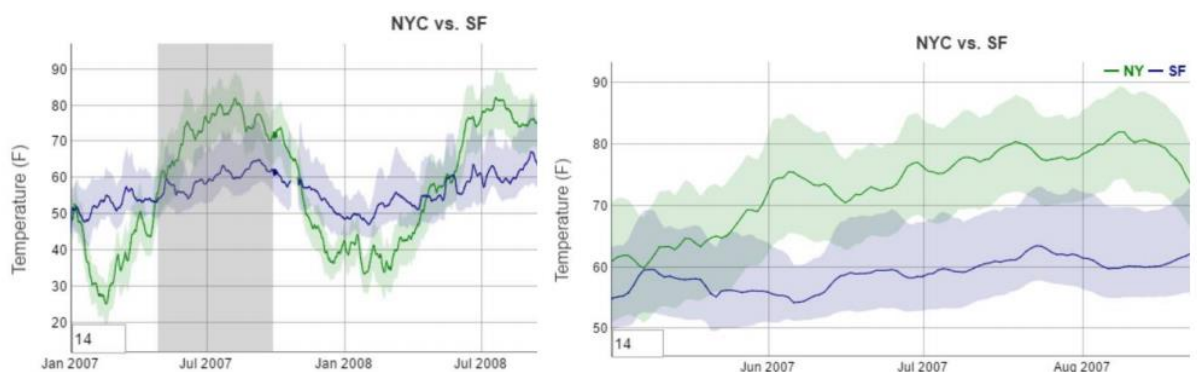
1.4.1 Վիզուալիզացիայի մեթոդներ և գործառնություններ

Տվյալների վիզուալ ներկայացումը մեծ տվյալների մշակման և հետազոտման կարևոր քայլերից է: Վիզուալիզացիայի գործիքների առանցքային առավելությունների թվում հիմնական ուշադրությունը սևեռվում է որոշումների կայացման բարելավմանը, սակայն ոչ պակաս կարևոր է նաև հաշվի առնել տվյալների հնարավորինս արդյունավետ վերլուծությունը և օգտագործողների միջև համագործակցության բարելավումը [21]: Գոյություն ունեն տվյալների վիզուալ ներկայացման ընդհանուր մեթոդներ՝ ներառյալ աղյուսակներ, հիստոգրամներ, պատմական ցուցանակներ կամ վահանակներ, ինչպես ցույց է տրված նկար 1.10-ում: Արտացոլման յուրաքանչյուր մեթոդ ունի կոնկրետ նպատակ, այդ իսկ պատճառով, անհրաժեշտ է որոշել, թե որ տեսակի տեղեկատվությունն է ցուցադրման համար ցանկալի, ապա ընտրել առավել համապատասխան գրաֆիկական ներկայացումը: Օրինակ, եթե անհրաժեշտ է ցուցադրել թվային հատվածներ, հարմար է օգտագործել շրջաձև դիագրամներ, իսկ համեմատություններ ցույց տալու համար ավելի նպատակահարմար է օգտագործել հիստոգրամներ:



Նկար 1.10. Վիզուալիզացիայի որոշ ներկայացումներ¹:

Առավել տարածված մեխանիզմները հիմնված են ինֆորմացիայի ներկայացման ստատիկ մեթոդների վրա, սակայն գոյություն ունեն ինտերակտիվ գործիքներ, որոնք թույլ են տալիս բազմակի պարամետրերի փոփոխման միջոցով ձևափոխել վիզուալիզացիոն արտապատկերումները: Այս տեսակի վիզուալիզացիաների շրջանակներում գոյություն ունեն որոշակի կարևոր գործառույթներ, որոնք պետք է հաշվի առնել [22].

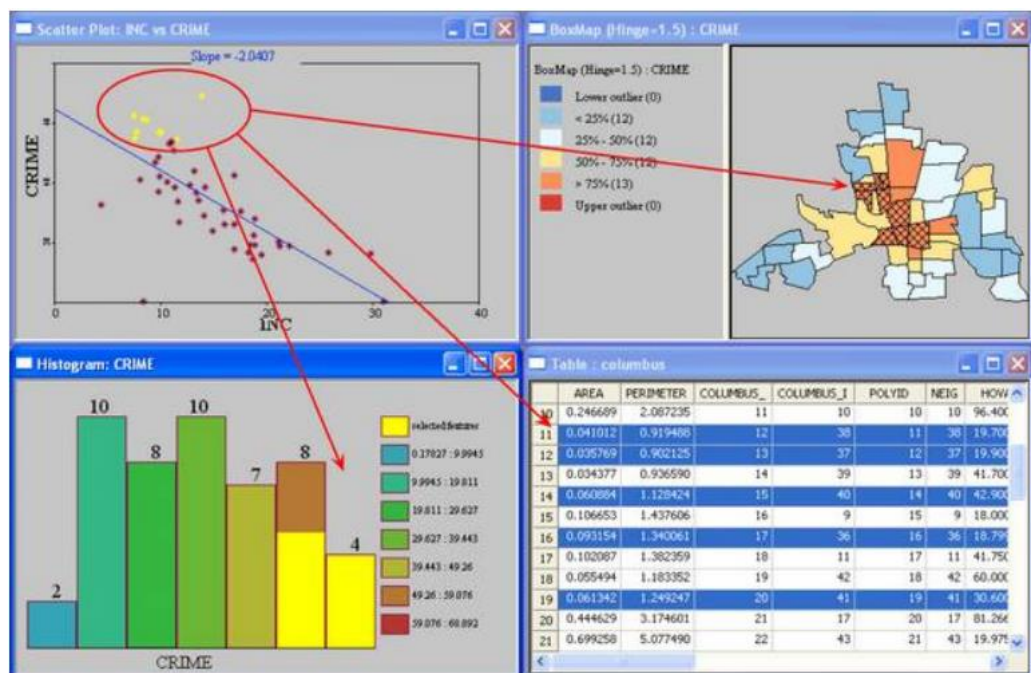


Նկար 1.11. Ընտրության գործառույթ²:

¹ <http://www.butleranalytics.com/big-data-visualization-tools/>

² <https://torbjornzetterlund.com/how-to-visualize-your-data-on-your-website/>

- Ընտրություն. Այս գործառույթը հնարավորություն է տալիս նշելու վիզուալիզացիայի տիրույթներ՝ մոտեցնելու և ավելի մանրամասն ինֆորմացիա տեսնելու համար: Նկար 1.11-ում ներկայացված են երկու պատկերներ, որոնցից ձախում կարելի է տեսնել նշված տիրույթը՝ վերցված մոխրագույն ուղղանկյան մեջ, որի ընդարձակ և ավելի մանրամասն ինֆորմացիա պարունակող տարբերակը ցուցադրված է աջ կողմի նկարում:
- Հղում և վրձին. Այս գործառույթը հնարավորություն է տալիս միմյանց հետ կապել միևնույն տվյալի տարբեր ներկայացումները, այսպիսով, ապահովելով ինֆորմացիայի միևնույն կտորի համար մի քանի դիտանկյուններ: Նկար 1.12-ը ցույց է տալիս 4 տարբեր պատկերներ, որտեղ հիմնական ինֆորմացիան պարունակող առաջին պատկերում նշելով հետաքրքրող հատվածը, մնացած երեքը թարմացվում և ցուցադրում են այդ նշված հատվածի մասին ինֆորմացիան:



Նկար 1.12. Հղում և վրձին գործառույթ³:

- Ֆիլտրում. Այս գործառույթը թույլ է տալիս զտել հետաքրքրող ինֆորմացիան ընդհանուրից, որի շնորհիվ նվազում է յուրաքանչյուր պահին ցուցադրվող

³ <https://geospatialro.wordpress.com/tag/open-source/>

ինֆորմացիայի քանակը, հնարավորություն տալով ավելի մանրամասն ուսումնասիրել հետաքրքրող հատվածը:

- Վերակազմավորում. Այս գործառույթը հնարավորություն է տալիս օգտագործելով տարածական և 3D վիզուալիզացիաներ՝ տրամադրել ինֆորմացիայի ավելի մեծ համատեքստ: Պտտելու և պատկերով շրջելու հնարավորությունը թույլ է տալիս տեսնել այն ինֆորմացիաները, որոնք կարող են թաքնված լինել, այսպիսով՝ բացառելով ինֆորմացիայի կորուստները:

Ստորև ներկայացված են որոշ հատկանիշներ, որոնք գերադասելի են, որ ներառված լինել վիզուալիզացիայի գործիքներում.

- Տվյալի ներմուծում տարբեր աղբյուրներից,
- Տեքստային նշումներով վիզուալ ներկայացման համալրում,
- Տեղեկատվության և գիտելիքի ստացում ներկայացման ձևից,
- Տվյալների մեծ ծավալների ներկայացում,
- Տվյալների ինտելեկտուալ վերլուծության ինտեգրում,
- Համատեղ աշխատանք,
- Հեշտ և հարմարավետ օգտագործում,
- Արդյունավետություն:

Հիմնվելով այս հատկությունների վրա, գրաֆիկական ներկայացման մասշտաբայնեցումը և դինամիկան կարելի է դիտարկել որպես տվյալների վիզուալիզացիայի հիմնական մարտահրավերներից երկուսը: Այնուամենայնիվ, կան մի քանի ֆունկցիոնալ հնարավորություններ, որոնք վիզուալիզացիայի գործիքները փորձում են կիրառել ներկայումս տվյալների ներկայացման համար ոչ միայն ինֆորմացիա ցուցադրելու, այլ նաև նրանից գիտելիք ստանալու համար:

ԳԼՈՒԽ 2

ՏՎՅԱԼՆԵՐԻ ՎԵՐԼՈՒԾՈՒԹՅԱՆ ԵՎ ՎԻԶՈՒԱԼԻԶԱՑԻԱՅԻ ԵՆԹԱԿԱՌՈՒՑՎԱԾՔՆԵՐ ԵՎ ԳՈՅՈՒԹՅՈՒՆ ՈՒՆԵՑՈՂ ԼՈՒԾՈՒՄՆԵՐ

Այս գլխում ներկայացված են առկա և կիրառելի ծրագրային գործիքներից որոշները, որոնք հնարավորություն են տալիս կատարել տվյալների վերլուծություններ և վիզուալիզացիաներ տարբեր խնդիրների լուծման համար: Ներկայացված են նրանց համեմատական վերլուծությունները:

2.1 Նախաբան

Գոյություն ունեն տվյալների վերլուծության և վիզուալիզացիայի բազմաթիվ գործիքներ տարբեր նպատակներով կիրառությունների համար, որոնք առաջարկում են հատկությունների լայն շրջանակ՝ հնարավորություն տալով լուծելու մի շարք տարբեր խնդիրներ: Որպես արդյունք, տարբեր կազմակերպություններ, հետազոտողներ և այլ օգտատերեր շատ հաճախ հանդիպում են բարդությունների, երբ փորձում են գտնել իրենց խնդիրների պահանջներին բավարարող համապատասխան գործիք: Հիմնական խնդիրը, որն առաջանում է, դա այն է, որ չկա մի ընդհանուր ճանապարհ ընտրելու ճիշտ գործիք և չկա այդ գործիքի գնահատման մեկ ընդհանուր մեթոդ: Այդ իսկ պատճառով, կոնկրետ խնդրի համար ճիշտ ծրագրային միջավայրի ընտրությունը լինում է բարդ, երբ հաշվի չեն առնվում այդ գործիքների տեսակների դասակարգումները:

Դասակարգումը հնարավորություն է տալիս խմբավորել գործիքները՝ օգտագործելով չափանիշների ստանդարտ բազմությունը՝ դիտարկելով խնդրի յուրահատկություններին համապատասխան գործիքի ընտրությունը: Տվյալների վերլուծության և վիզուալիզացիայի գործիքները կարելի է դասակարգել ըստ իրենց տրամադրած ֆունկցիոնալության, կառուցվածքային ձևերի և աշխատանքային միջավայրերի: Ըստ աշխատանքային միջավայրի կարելի է բաժանել երկու խմբի՝ դեսկտոպ և վեբ, իսկ ըստ կառուցվածքի՝ առանց օգտագործողի միջավայրի և օգտագործողի միջավայրով: Այսպիսով, դիտարկենք այս դասակարգումների հիման վրա գոյություն ունեցող որոշ վիզուալիզացիայի գործիքներ:

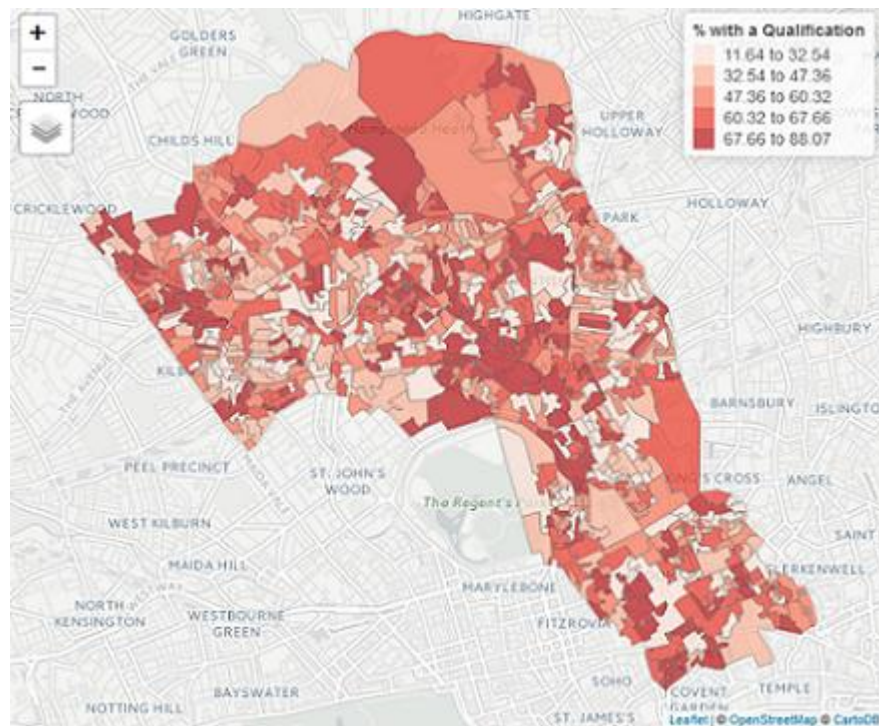
2.2 Դեսկտոպ միջավայրում աշխատող ծրագրային գործիքներ

Այս տեսակի գործիքները օգտագործելու համար անհրաժեշտ է ունենալ համակարգիչ, բեռնել համապատասխան ծրագրային փաթեթը և նոր միայն կարողանալ օգտագործել: Դեսկտոպ միջավայրում աշխատող ծրագրային միջոցները հնարավորություն են տալիս օգտագործել լոկալ գրաֆիկական սարքավորումների հաշվողական հզորությունները, սակայն միևնույն ժամանակ իրենց բնույթով առաջացնում են բարդություններ, քանի որ պահանջում են համապատասխան օպերացիոն համակարգ և անհրաժեշտություն է առաջանում համակարգչում տեղադրել լրացուցիչ ծրագրային փաթեթներ: Հրամանների տողով աշխատող գործիքները շատ հարմար են ներդրված կիրառությունների դեպքերում, երբ հնարավոր է մեկ այլ համակարգից օգտագործել տրամադրած միջոցները, սակայն հետազոտողից պահանջվում է լրացուցիչ գիտելիքների իմացություն:

2.2.1 Հրամայական տողի կիրառմամբ ծրագրային գործիքներ

R

R-ը վերլուծական հաշվարկների ու գրաֆիկների լեզու և միջավայր է (նկ. 2.1) [23]: Այն GNU-ի նախագիծ է և նման է S միջավայրին, որը մշակվել է Bell Laboratories-ի կողմից: Այն շատ ընդարձակելի է և տրամադրում է վերլուծությունների (գծային և ոչ գծային, դասական վերլուծական թեստեր, ժամանակային հաջորդականությունների վերլուծություններ, դասակարգում, կլաստերինգ և այլն) և գրաֆիկական տեխնիկաների/հմտությունների լայն շրջանակ: R-ի առավելություններից մեկն այն է, որ թույլ է տալիս կառուցել հրատարակչությունների համար որակով արտապատկերումներ՝ ներառյալ մաթեմատիկական սիմվոլներ և բանաձևեր:



Նկար 2.1. R գրադարանի կողմից տրամադրված վիզուալիզացիայի օրինակ⁴:

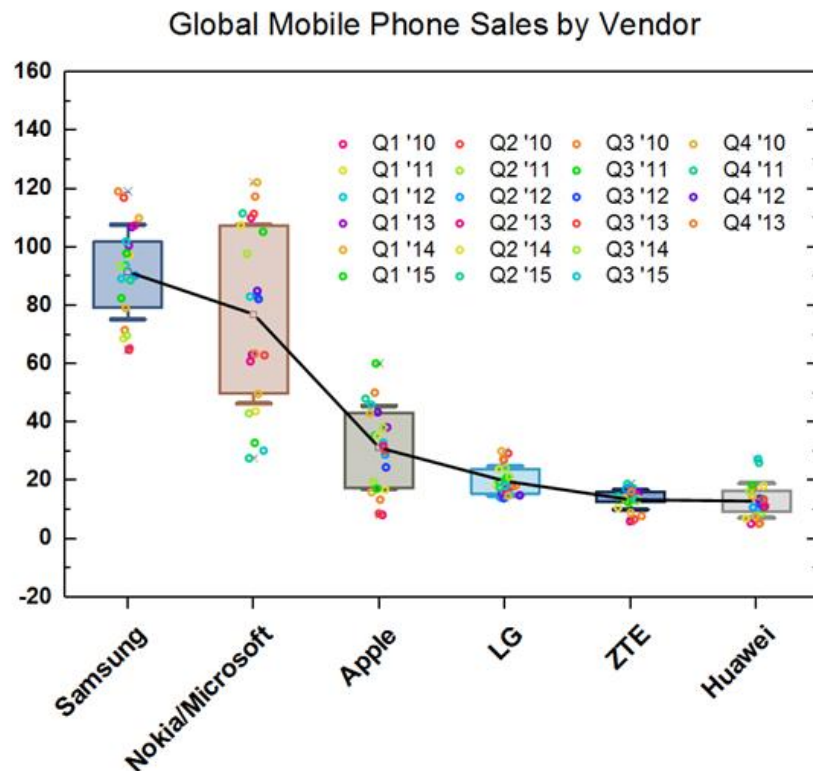
R-ը հնարավորություն է տալիս օգտագործել տարբեր տեսակի ֆայլեր՝ CSV, SAS, TXT, JSON և SPSS, կամ միանգամից օգտվել գործիքներից, ինչպիսիք են՝ Microsoft Excel, Microsoft Access, Oracle, MySQL և SQLite: Գոյություն ունեն որոշ գրաֆիկական ինտերֆեյսներ, խմբագիրներ և IDE-ներ, որոնք հեշտացնում են աշխատանքը: Այն աշխատում է բոլոր հիմնական օպերացիոն համակարգերում: Այն կառավարում է մեծ ծավալով տվյալներ և, պրոցեսների ավտոմատացման միջոցով բեռնումը դառնում է ավելի դյուրին: Անհրաժեշտ են մասնագիտացված մշակողներ, քանի որ սա հանդիսանում է բարդ լեզու և համարվում է ոչ դյուրին ուսուցման համար:

Origin

Origin-ը տվյալների վերլուծությունների և գրաֆիկական ներկայացման ծրագրային ապահովում է՝ նախատեսված կոմերցիոն արդյունաբերություններում գիտնականների

⁴ <http://blog.revolutionanalytics.com/2017/05/cdrc-spatial-course.html>

և ինժեներների, ակադեմիկոսների և կառավարական լաբորատորիաների համար (նկ. 2.2) [24]:



Նկար 2.2. Origin-ի կողմից տրամադրված վիզուալիզացիայի օրինակ⁵:

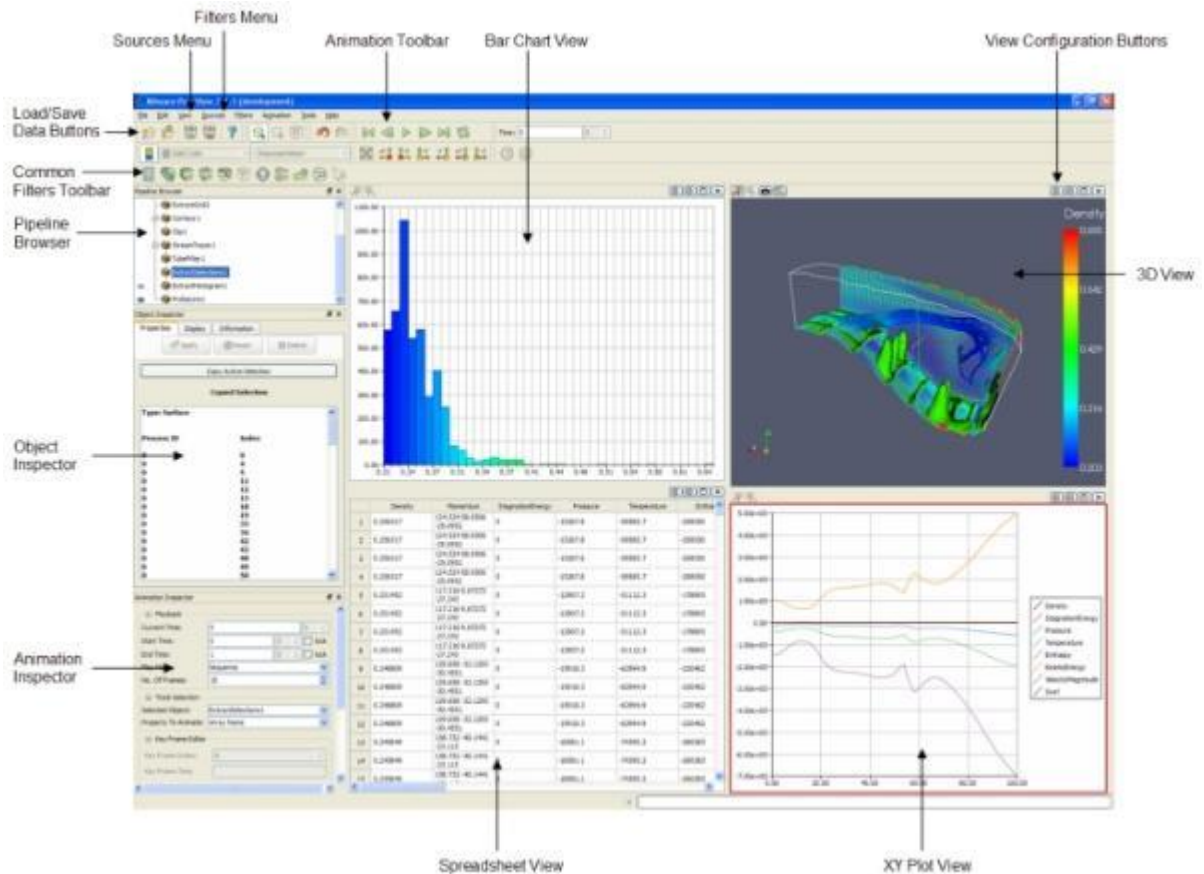
Այն առաջարկում է հեշտ օգտագործելի միջավայր սկսնակների համար, հնարավորություն է տալիս կատարել առաջադեմ փոփոխություններ՝ միջավայրի հետ ավելի լավ ծանոթանալու զուգընթաց: Origin-ի գրաֆիկներն ու վերլուծությունների արդյունքները թարմացվում են տվյալների պարամետրերի փոփոխության հետ մեկտեղ, հնարավորություն է տալիս ստեղծել կադապարներ կրկնվող խնդիրների դեպքում օգտագործելու համար և օգտագործողի միջավայրից՝ առանց ծրագրավորման, իրականացնել խմբակային գործողություններ: C ծրագրավորման լեզվի, ներդրված Python-ի կամ R-ի միջոցով հնարավոր է կապել այլ ծրագրերի հետ, ինչպիսիք են՝ MATLAB, LabVIEW, Microsoft© Excel կամ Origin-ում ստեղծված ռեժիմները:

⁵ <https://www.originlab.com/2017>

2.2.2 Օգտագործողի միջավայրով ծրագրային գործիքներ

ParaView

ParaView-ն տվյալների վերլուծության և վիզուալիզացիայի բաց, բազմապլատֆորմային գործիք է՝ մշակված Սանդիայի ազգային լաբորատորիայի, Kitware ընկերության և Լոս-Ալամոսի ազգային լաբորատորիայի կողմից (նկ. 2.3) [25]:



Նկար 2.3. ParaView գործիքի օգտագործողի միջավայր⁶:

Այն բաց կոդով, բազմաֆունկցիոնալ տվյալների վերլուծության և վիզուալիզացիայի համակարգ է: Օգտատերերը կարող են արագ ձևավորել պատկերներ, վերլուծել իրենց տվյալները՝ օգտագործելով որակական և քանակական մեթոդներ: Տվյալների հետազոտությունները կարող են կատարվել ինտերակտիվորեն 3D ձևաչափով կամ տրամադրված ծրագրային միջավայրով՝ օգտագործելով ParaView-ի խմբաքանակի մշակման հնարավորությունները: ParaView-ն մշակվել է վերլուծելու համար չափազանց մեծ տվյալների հավաքածուները՝ օգտագործելով տարածական հիշողության

⁶ <http://www.bu.edu/tech/support/research/training-consulting/online-tutorials/paraview/>

հաշվողական ռեսուրսները: Այն կարող է գործարկվել սուպերհամակարգիչների համար՝ վերլուծելու համար պետամասշտաբ չափի տվյալներ, ինչպես նաև նոութբուքների՝ փոքր տվյալների վերլուծության համար: Համակարգը աշխատում է տվյալների ֆորմատների մեծ մասի հետ և հնարավորություն է տալիս բեռնել արդեն պատրաստի վիզուալիզացիոն վիճակներ, և ստեղծել նաև նոր գրաֆիկական վիզուալիզացիաներ հենց գործիքի միջոցով:

ParaView-ն նաև տրամադրում է առցանց միջավայր, որը դեռ ամբողջովին ավարտուն վիճակում չէ, և չի պարունակում բոլոր հնարավորությունները, ինչպես դեսկտոպ տարբերակում, և առաջացնում է տեխնիկական բարդություններ որոշ օպերացիոն համակարգերի տարբերակների համար:

2.3 Վեբ միջավայրում աշխատող ծրագրային գործիքներ

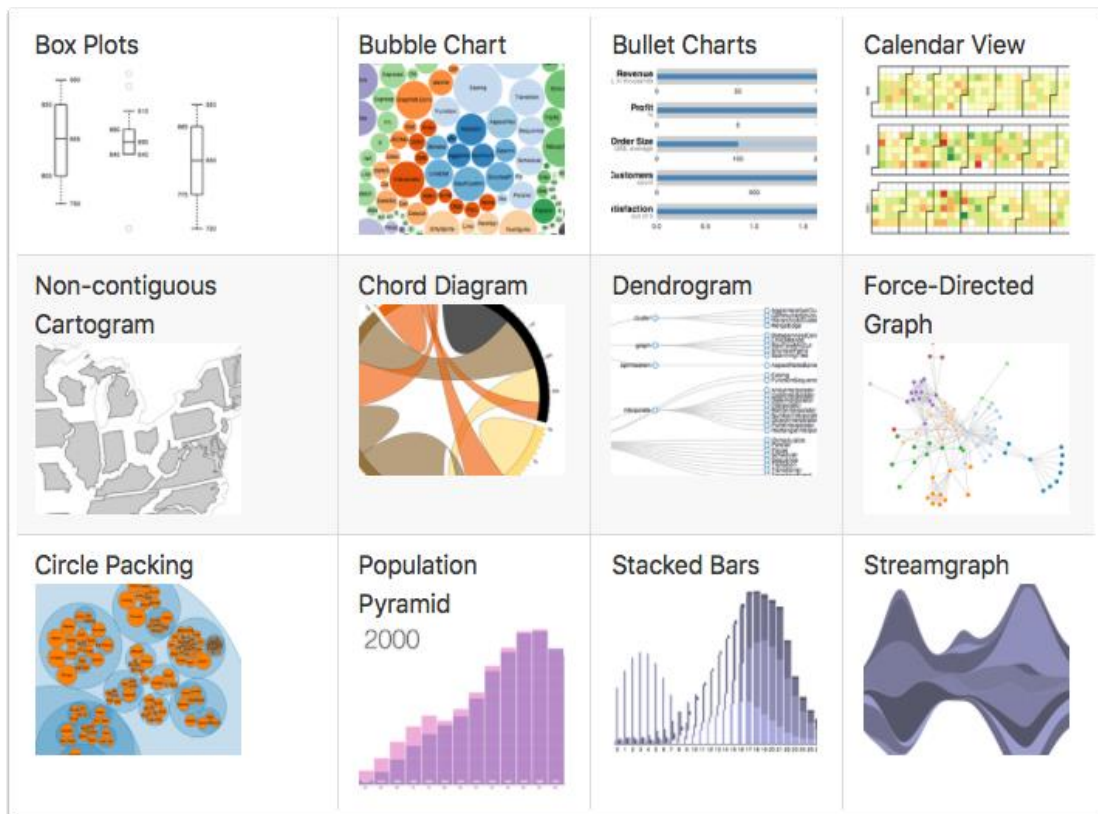
Ժամանակակից իրականությունում ինտերնետի լայնորեն կիրառության և հանրահայտության հետ մեկտեղ ստեղծվել են վիզուալիզացիայի նոր մոտեցումներ՝ հիմնված վեբ միջավայրի վրա, որոնք կարող են օգնել տվյալների տարբեր աղբյուրների միջև կապի ստեղծմանը և ծրագրային ապահովման ծախսերի նվազեցմանը, հնարավորություն տալով ներկայացումը դարձնել արդիական և թարմացված: Այս գործիքները շատ հանրահայտ են բիզնես վերլուծության ոլորտում, սակայն ոչ այդքան դեռևս կիրառելի գիտական ոլորտներում:

2.3.1 Վիզուալիզացիայի գրադարաններ

Այս գործիքների օգտագործումը պահանջում է տեխնիկական հմտություններ տվյալների հետ աշխատելու և ճիշտ վիզուալիզացիա ստեղծելու համար: Ստորև ներկայացված են այս խմբին պատկանող ամենահայտնի գործիքներից որոշները.

D3js

D3js-ը (Data Driven Document) հանդիսանում է JavaScript գրադարան մեծ տվյալների ինտերակտիվ վիզուալիզացիայի համար, որն ունի բոլոր հնարավորությունները իրական ժամանակում կիրառության համար (Նկ. 2.4) [26]:



Նկար 2.4. *D3.js* գրադարանի կողմից տրամադրված վիզուալիզացիայի օրինակներ⁷:

Սա գործիք չէ և, հետևաբար, օգտագործողը պետք է ունենա JavaScript ծրագրավորման լեզվի բավարար գիտելիքներ, որպեսզի կարողանա ներկայացնել տվյալները մարդու համար հասկանալի տեսքով: Այս գրադարանը մշակում է տվյալները՝ օգտագործելով SVG և HTML5 ձևաչափերը, այդ իսկ պատճառով, դիտարկիչների հին տարբերակները, ինչպիսիք են՝ IE7 և IE8-ը, չեն կարող բավարարել D3.js-ի պայմաններին:

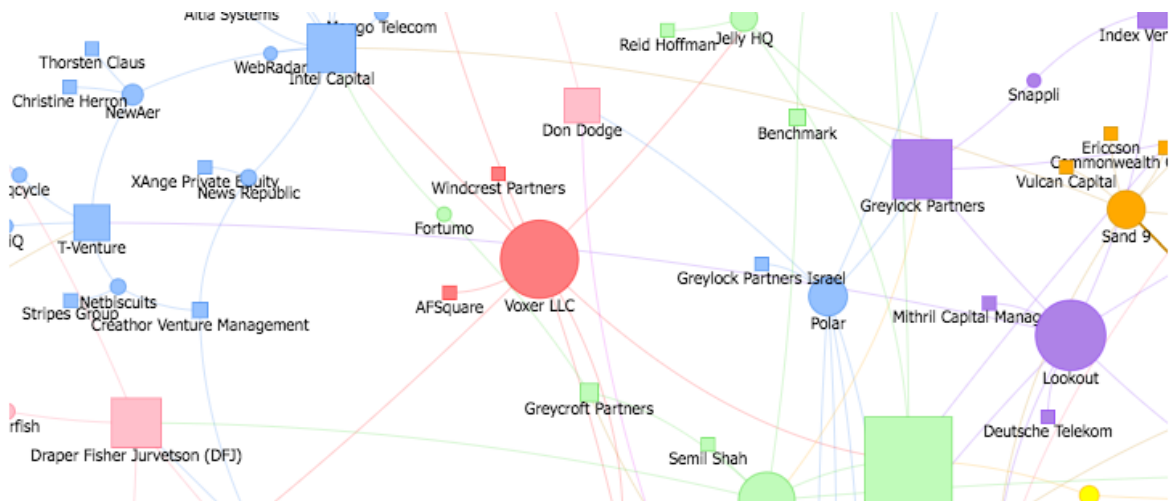
Տարբեր անհամաչափ աղբյուրներից հավաքագրված մեծ ծավալի տվյալների հավաքածուները իրական ժամանակում կապվում են DOM-ի [27] հետ՝ հնարավորություն տալով չափազանց արդյունավետ ճանապարհով կառուցելու ինտերակտիվ անիմացիաներ՝ 2D և 3D ձևաչափերով: D3-ի ճարտարապետությունը թույլ է տալիս օգտագործողներին այն օգտագործել տարբեր համակարգերում որպես հավելումներ:

⁷ <https://spin.atomicobject.com/2013/08/12/d3-javascript-data-visualization/>

Անհրաժեշտ են վիզուալիզացիաների ստեղծման համար մասնագիտացված մշակողներ: Մշակման ժամանակը երկար է և, վիզուալիզացիայի համար անհրաժեշտ էլեմենտների աճի հետ զուգընթաց, սահմանափակվում և նվազում է արդյունավետությունը:

VisJs

VisJS-ը դինամիկ, դիտարկիչների վրա հիմնված վիզուալիզացիայի գրադարան է: Այն մշակված է մեծ ծավալով դինամիկ տվյալների մշակման և տվյալները ինտերակտիվ ձևով մանիպուլացնելու հնարավորություն ընձեռելու համար: Գրադարանը բաղկացած է DataSet, Timeline, Network, Graph2d and Graph3d բաղադրիչներից (նկ. 2.5) [28]:



Նկար 2.5. Vis/s գրադարանի կողմից տրամադրված վիզուալիզացիայի օրինակ⁸:

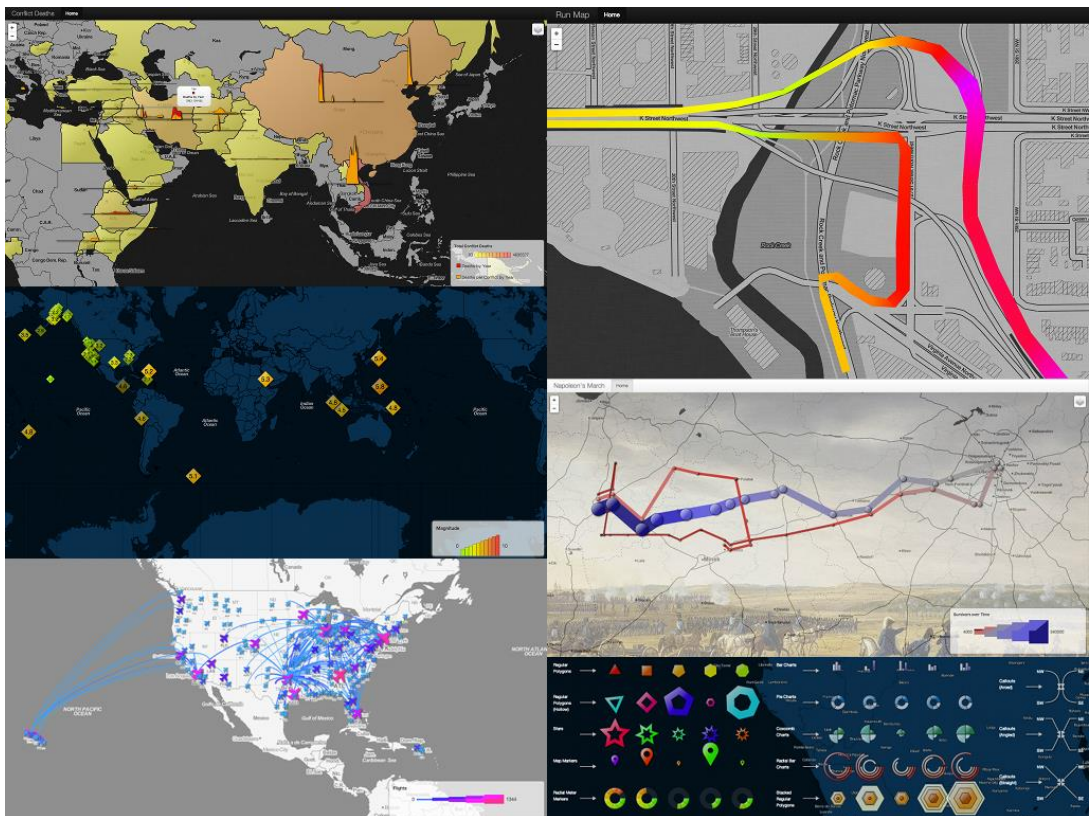
Տվյալների բեռնման համար անհրաժեշտ է ստեղծել սերվերային հատված յուրաքանչյուր լուծման համար: Այն թույլ է տալիս կապել տարբեր տեսակի տվյալներ DOM-ի հետ և ստեղծել խնդրի պայմաններին համապատասխան վիզուալիզացիաներ: Վերջիններս ստեղծելու համար անհրաժեշտ են մասնագիտացված մշակողներ: Մշակման ժամանակը երկար է և, վիզուալիզացիայի համար անհրաժեշտ էլեմենտների աճի հետ զուգընթաց, սահմանափակվում և նվազում է արդյունավետությունը:

⁸ <https://constructive.co/insights/6-best-interactive-data-visualization-tools-2016-part-2/>

Leaflet

Leaflet-ը հանդիսանում է JavaScript գրադարան, որը նախատեսված է ինտերակտիվ քարտեզներ ստեղծելու համար (նկ. 2.6) [29]:

Այն արդյունավետ ձևով աշխատում է բոլոր հիմնական դեսկտոպ ու բջջային պլատֆորմների համար և հնարավորություն ունի ընդլայնվելու՝ օգտագործելով հավելումներ: Ունի գեղեցիկ, հեշտ օգտագործելի և բավականին լավ API-ի թղթաբանություն և պարզ ու ընթեռնելի աղբյուրի կոդ: Հնարավորություն է տալիս օգտագործել տվյալներ տարբեր աղբյուրներից, սակայն տվյալների բեռնման համար անհրաժեշտ է մշակել համապատասխան կոդ: Քարտեզներում տվյալներ ցուցադրելու համար օգտագործվում են GeoJSON ֆայլեր: Այս գրադարանը օգտագործելու համար անհրաժեշտ է ունենալ որոշակի GIS գիտելիքներ, իսկ համակարգը կարգավորելու համար պահանջվում է մասնագետի օգնությունը:



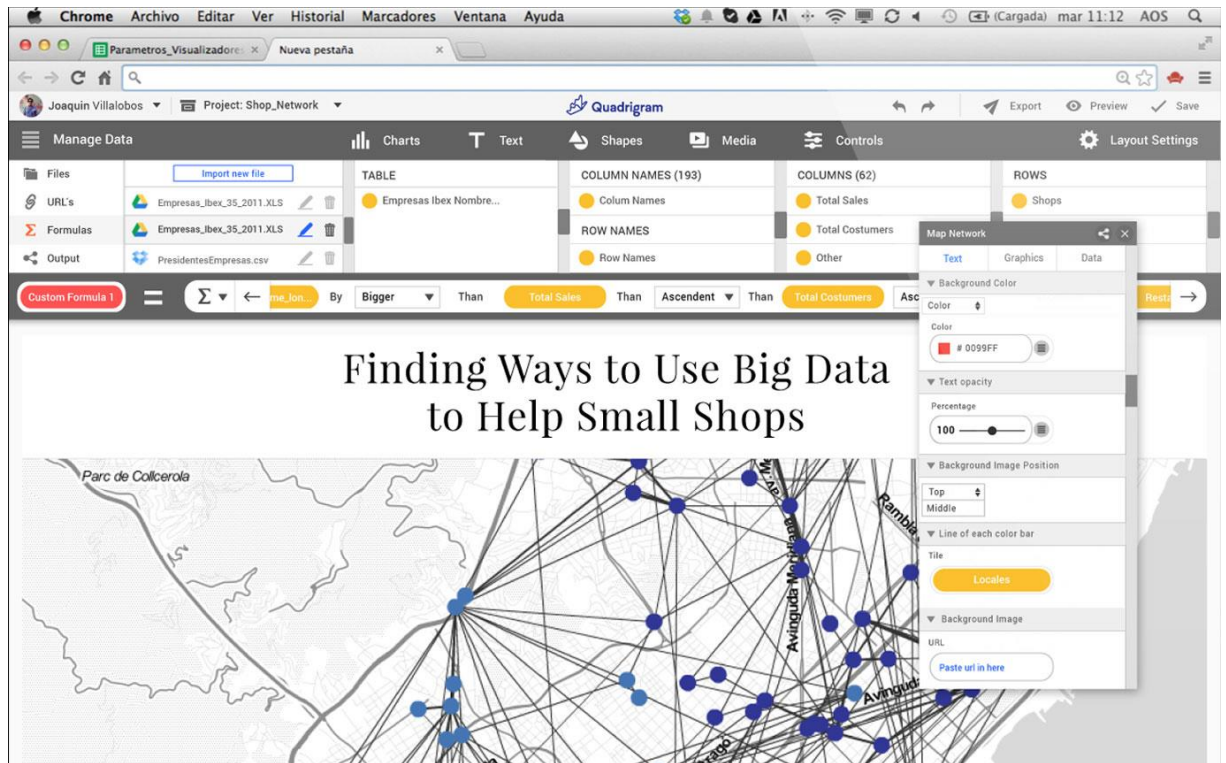
Նկար 2.6. Leaflet գրադարանի կողմից տրամադրված վիզուալիզացիայի օրինակներ⁹:

⁹ <http://dmitrynazarov.blogspot.com/2014/03/>

2.3.2 Օգտագործողի միջավայրով պատրաստի գործիքներ

Quadrigram

Quadrigram-ը վիզուալ ծրագրավորման միջավայր է, որը հնարավորություն է տալիս կառուցել և տարածել տվյալների ինտերակտիվ վիզուալիզացիաները, առանց որևէ ծրագրավորման հմտությունների (նկ. 2.7) [30]:



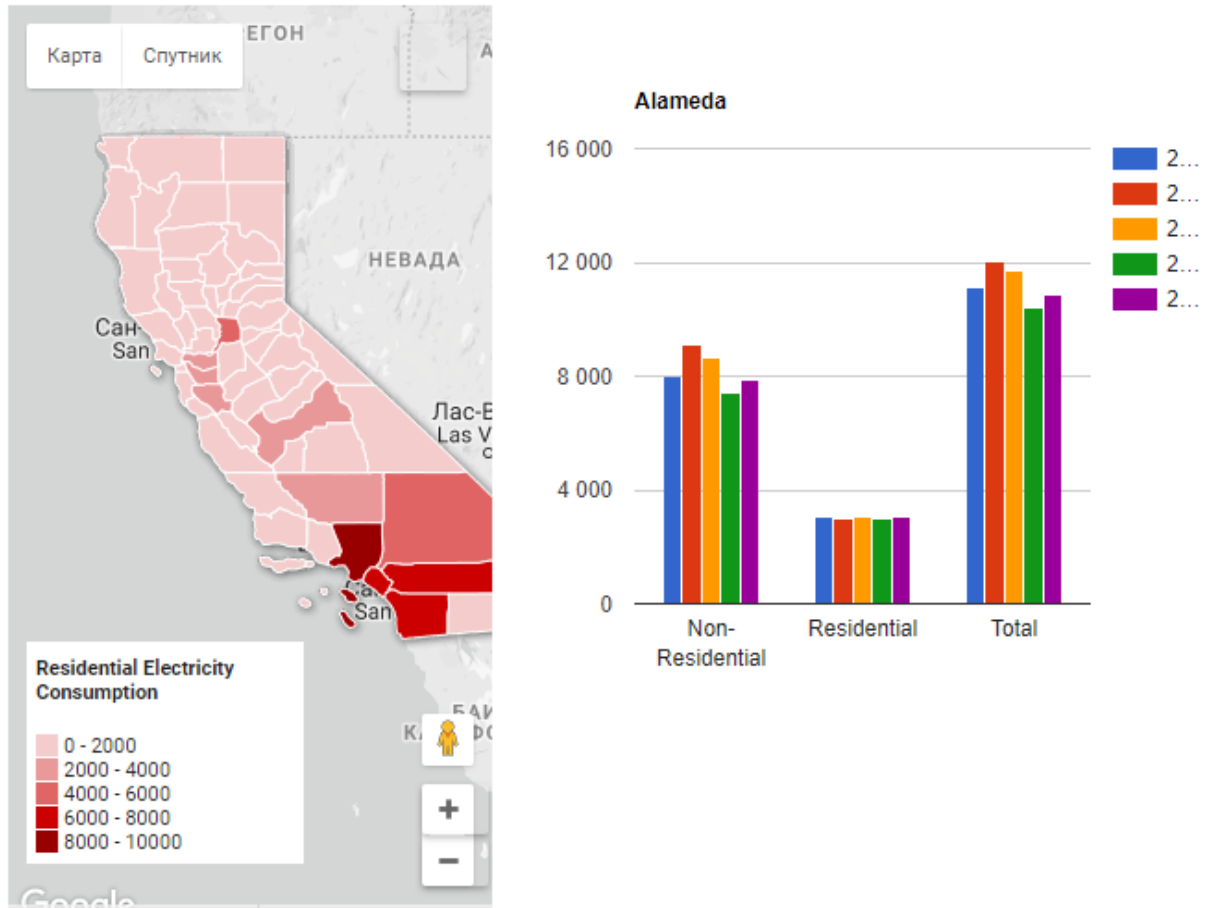
Նկար 2.7. *Quadrigram* գործիքի օգտագործողի միջավայր¹⁰:

Համակարգը պարունակում է ավելի քան 50 տեսակի ինտերակտիվ վիզուալիզացիաներ: Որպես տվյալների պահոց օգտագործվում է Google Drive ծրագրային համակարգը: Տվյալների մուտքի համար անհրաժեշտ է բեռնել տվյալները JSON [31] ձևաչափով: Quadrigram-ում գոյություն ունի առանց նոր գրաֆիկների դասերի ստեղծման հնարավորության, առկա վիզուալիզացիաների տեսակների նկատմամբ հնարավորությունները սահմանափակ են: Վիզուալիզացիա ստեղծելու համար անհրաժեշտ է տիրապետել հատուկ ծրագրավորման լեզվի:

¹⁰ <http://www.topexpert.com.ua/it/servisy-vizualizacii-i-infografiki/>

Google Fusion Tables

Google Fusion Tables-ը վեր միջավայրի ծրագիր է, որը նախատեսված է աղյուսակային տվյալները վիզուալիզացնելու և տարածելու համար (նկ. 2.8) [32]:



Նկար 2.8. Google Fusion Tables գործիքի տրամադրած վիզուալիզացիայի օրինակ¹¹:

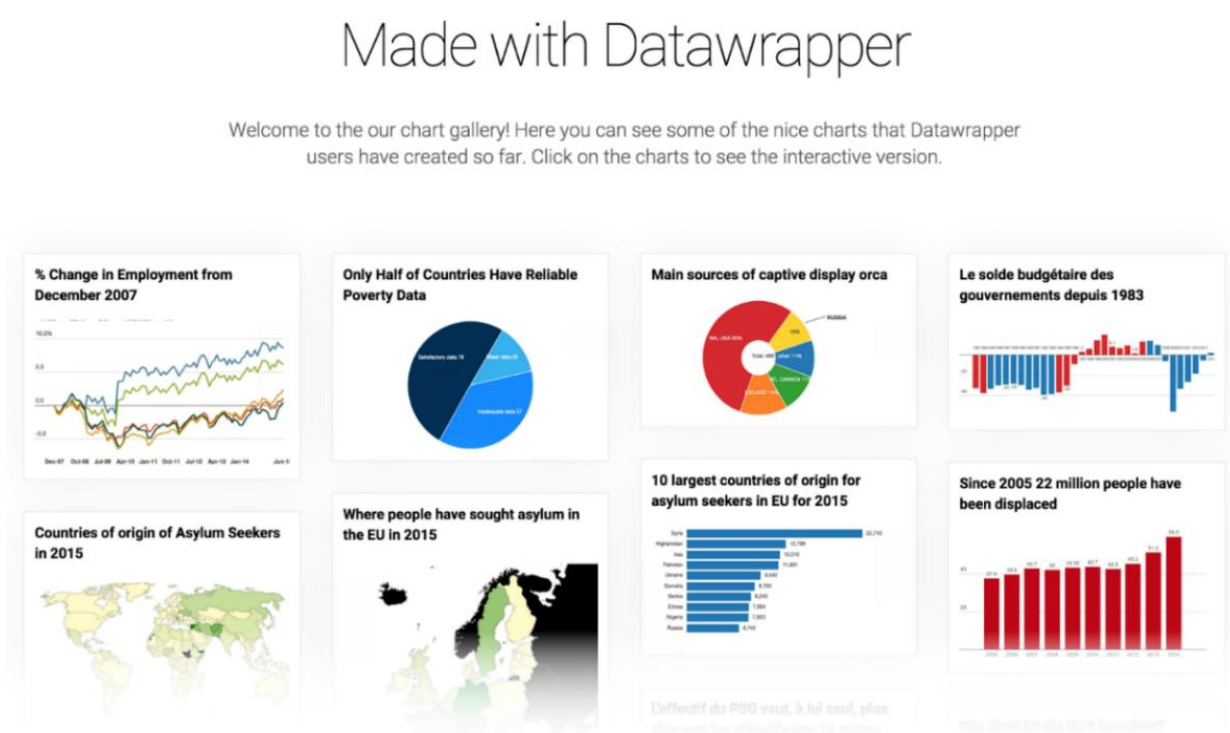
Օգտատերը կարող է բեռնել CSV, KML, ODS, XLS կամ Google Spreadsheet ձևաչափերով տվյալներ Fusion Tables աղյուսակում, օգտագործել այդ տվյալները միաժամանակ այլ օգտատիրոջ հետ իրական ժամանակում, տարածել այն Google Search համակարգում, ստեղծել քարտեզի և գծագրերի վիզուալիզացիաներ մասնավոր օգտագործման համար կամ օգտագործել որպես ներդրված մոդուլ այլ վեբ համակարգերում, անհրաժեշտ չափանիշներով կատարել ֆիլտրում: Հասանելի վիզուալիզացիաների տեսակները սահմանափակ են օգտագործողի համար և առկա է

¹¹ <http://qaru.site/questions/1978479/in-fusion-tables-can-a-gradient-be-set-via-the-google-maps-javascript-based-api>

ոչ բարդ վիզուալ ծրագրավորման լեզու, որն անհրաժեշտ է իմանալ վիզուալիզացիաներ ստեղծելու համար: Ավելի բարդ վիզուալիզացիաներ ստեղծելու համար հնարավորություն է տալիս օգտագործելու Google գրադարանները կամ JavaScript ծրագրավորման լեզուն: Անհրաժեշտ է նախօրոք պատրաստել տվյալները և ձևափոխել տվյալ գործիքի համար համապատասխան ֆորմատի:

Datawrapper

Datawrapper-ը առցանց գործիք է, որը նախատեսված է ինտերակտիվ գծապատկերներ ստեղծելու համար (նկ. 2.9) [33]: Տվյալների բեռնման համար անհրաժեշտ է այն ձևափոխել CSV ֆորմատի, այնուհետև այդ տվյալները լրացվում են համապատասխան դաշտերում և Datawrapper-ը գեներացնում է վիզուալիզացիաների տարբեր տեսակներ: Հիմնականում օգտագործվում են նորություններ տրամադրող համակարգերում՝ ներդրված տարբերակով:

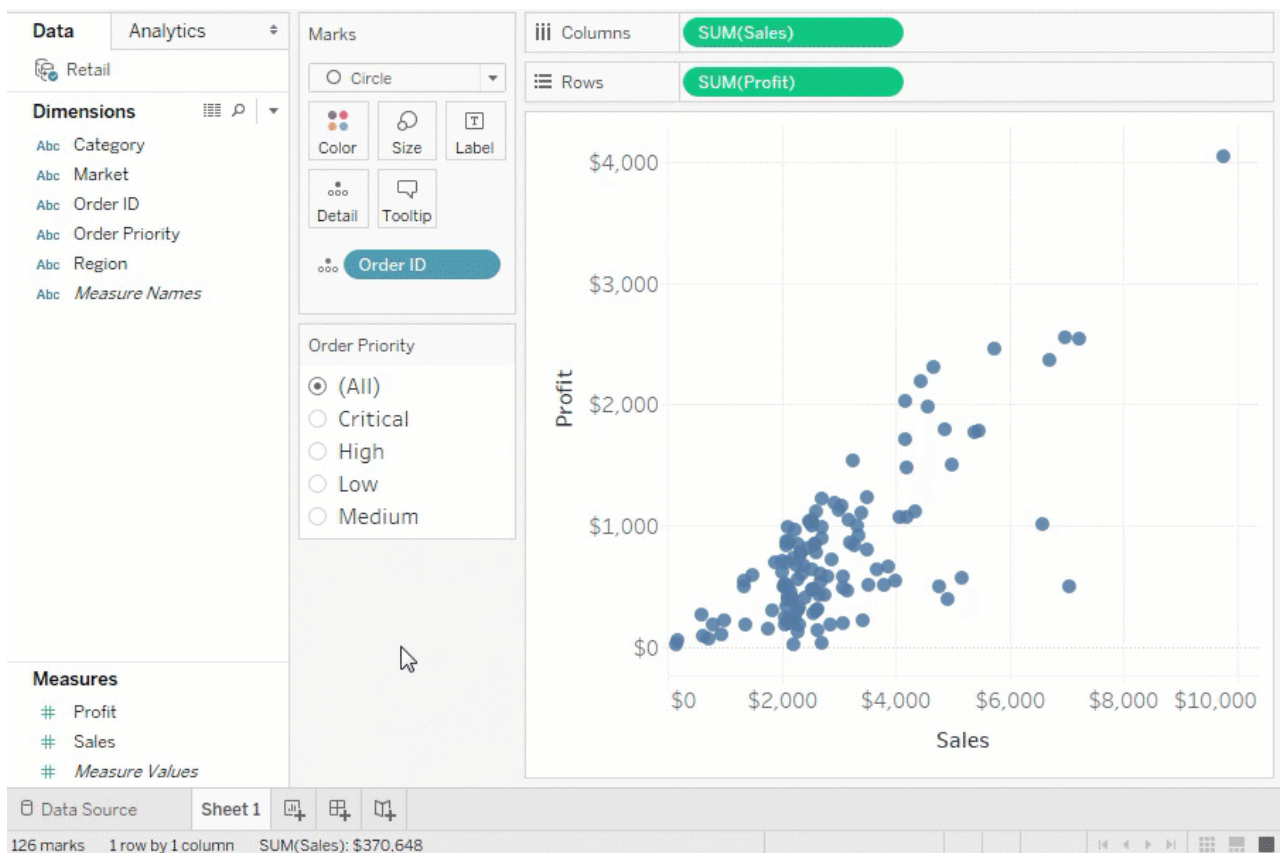


Նկար 2.9. *Datawrapper* գործիքի կողմից տրամադրված վիզուալիզացիայի օրինակներ¹²:

¹² <https://www.justinmind.com/blog/8-data-visualization-tools-for-ux-designers-visualize-insights-prototype-better-products/>

Tableau Online

Tableau Online-ը հանդիսանում է ամպային համակարգի վրա հիմնված տվյալների ինտերակտիվ վիզուալիզացիայի միջավայր (նկ. 2.10) [34]: Այն տրամադրում է բազմաթիվ գրավիչ և ֆունկցիոնալ վիզուալիզացիաներ: Այս գործիքը վերլուծության գործիք չէ, սակայն հիմնականում օգտագործվում է ներքին վահանակներում, որը հնարավորություն է տալիս հեշտությամբ կատարել տվյալների վերլուծություն: Դյուրին է օգտագործման և կառավարման համար՝ ստանալու համար ավելի հարմարավետ վիզուալիզացիաներ: Սակայն եթե անհրաժեշտ է ստանալ ավելի բարդ վիզուալիզացիաներ, անհրաժեշտ է տիրապետել R ծրագրավորման լեզվին: Վիզուալիզացիաների որոշ տեսակներ, օրինակ՝ գործիքների վահանակները, ստացվում են Tableau սերվերի միջոցով, իսկ ֆորմատավորման ֆունկցիաները՝ միայն Tableau-ի կողմից տրամադրվածներն են: Հնարավոր չէ ստեղծել նոր վիզուալիզացիաներ, եթե տրամադրված չեն:



Նկար 2.10. Tableau գործիքի օգտագործողի միջավայր¹³:

¹³ https://data-visualization.cioreview.com/vendor/2017/tableau_software

2.4 Եզրակացություն

Այսպիսով, կարող ենք տեսնել, որ տվյալների վերլուծության և վիզուալիզացիայի ոլորտը հագեցած է ծրագրային միջավայրերով, որոնք բավարարում են տարբեր գիտական խնդիրների պահանջներին: Սակայն հաշվի առնելով այդ գործիքների տրամադրած հնարավորությունները, աշխատանքային կառուցվածքը և առանցքային կետերը, նրանց համեմատական վերլուծությունները, հասկանալի է դառնում, որ դեռ գոյություն չունի այնպիսի գործիք, որում հնարավորինս հաշվի առնված լինեն որոշակի կարևոր բաղադրիչներ:

Հետազոտողի տեսանկյունից, դեսկտոպ միջավայրերում աշխատող ծրագրային միջոցները իրենց բնույթով այդքան էլ հարմար չեն, քանի որ պահանջում են համապատասխան օպերացիոն համակարգ և կիրառությունների արդյունքները մնում են այդ համակարգում: Կան համակարգեր, որոնք լուծում են այդ հարցը՝ տրամադրելով առցանց պահոցներ, սակայն դա հաճախ բարդություն է առաջացնում, քանի որ փոխելով աշխատանքային միջավայրը և գործիքը, անհրաժեշտ է նորից իրականացնել այդ ծրագրային միջոցի ներդնումը: Գիտական խնդիրների լուծման համար ոչ կոմերցիոն բնույթ ունեցող գործիքները շատ հաճախ պահանջում են հատուկ օպերացիոն համակարգերի տարբերակներ և ունենում են որոշակի ծրագրային թերություններ, որոնք կիրառության համար բարդություններ են առաջացնում:

Վեբ միջավայրում աշխատող ծրագրային միջավայրերը հնարավորություն են տալիս օգտագործել տրամադրվող միջոցները անկախ ծրագրային միջավայրից: Ինչպես նշվեց, այս միջավայրում գոյություն ունեն երկու դասի գործիքներ՝ գրադարաններ և պատրաստի օգտագործողի միջավայրով: Գրադարանները նախատեսված են որպես վիզուալիզացիա տրամադրող միջոց այլ համակարգերում կիրառությունների համար, այսինքն՝ իրենց որպես առանձին ծրագրային միջոց հնարավոր չէ օգտագործել: Իսկ օգտագործողի միջավայրով միջոցներից որոշները պահանջում են տվյալների բեռնման համար օգտագործել սահմանված ձևաչափ, սակայն ոչ բոլոր խնդիրներում է, որ վիզուալիզացիայի համար անհրաժեշտ արդյունքը ստացվում է նման ձևաչափերում: Շատ հաճախ անհրաժեշտություն է լինում կիրառելու միջանկյալ միջոցներ, նման ձևաչափերի ձևափոխելու համար: Որոշ ծրագրեր անվճար

են տրամադրում միայն բազային հնարավորություններ, իսկ լրացուցիչ և ավելի ճկուն միջոցներ տրամադրում են միայն որոշակի գումարով:

Ներկայացված միջոցները ունենում են իրենց բավարար կիրառությունները որոշ գիտական խնդիրների պահանջները բավարարելու համար, սակայն կան այնպիսիք, որտեղ մեկ գործիքի կիրառումը դեռ բավարար չէ, քանի որ մեկ համակարգում ներդրված չեն անհրաժեշտ բաղադրիչները: Օրինակ, եղանակի ուսումնասիրության խնդրում կարելի է գտնել բազմաթիվ ծրագրային միջոցներ, որոնք հնարավորություն կտան ստանալ գեղեցիկ և ինֆորմացիոն քարտեզներ, սակայն մինչև վիզուալիզացիայի փուլին հասնելը անհրաժեշտ է հավաքագրել այդ տվյալները, վերլուծել, բերել համապատասխան ձևաչափի: Այս տեսակի խնդիրների դեպքում, հետազոտողի համար շատ արդյունավետ կլինի ունենալ այս գործառույթները մեկ գործիքում, որպեսզի խնդիրներ չառաջանան օգտագործման ձևի և գիտելիքների չիմացության:

Հաշվի առնելով վերը նշված համեմատական վերլուծությունները, գոյություն ունեցող ծրագրային միջավայրերի հնարավորությունները և գիտական խնդիրներով զբաղվող հետազոտողների պահանջները, կառուցվել է համալիր համակարգ, որի կիրառությամբ հնարավոր կլինի շրջանցել առաջացած բարդությունները և կենտրոնանալով կոնկրետ խնդրի լուծման համար անհրաժեշտ պահանջներին, հնարավորություն տալով տվյալների պահպանումը, հաշվարկները, վերլուծությունները և վիզուալիզացիաները իրականացնել մեկ համակարգում:

ԳԼՈՒԽ 3

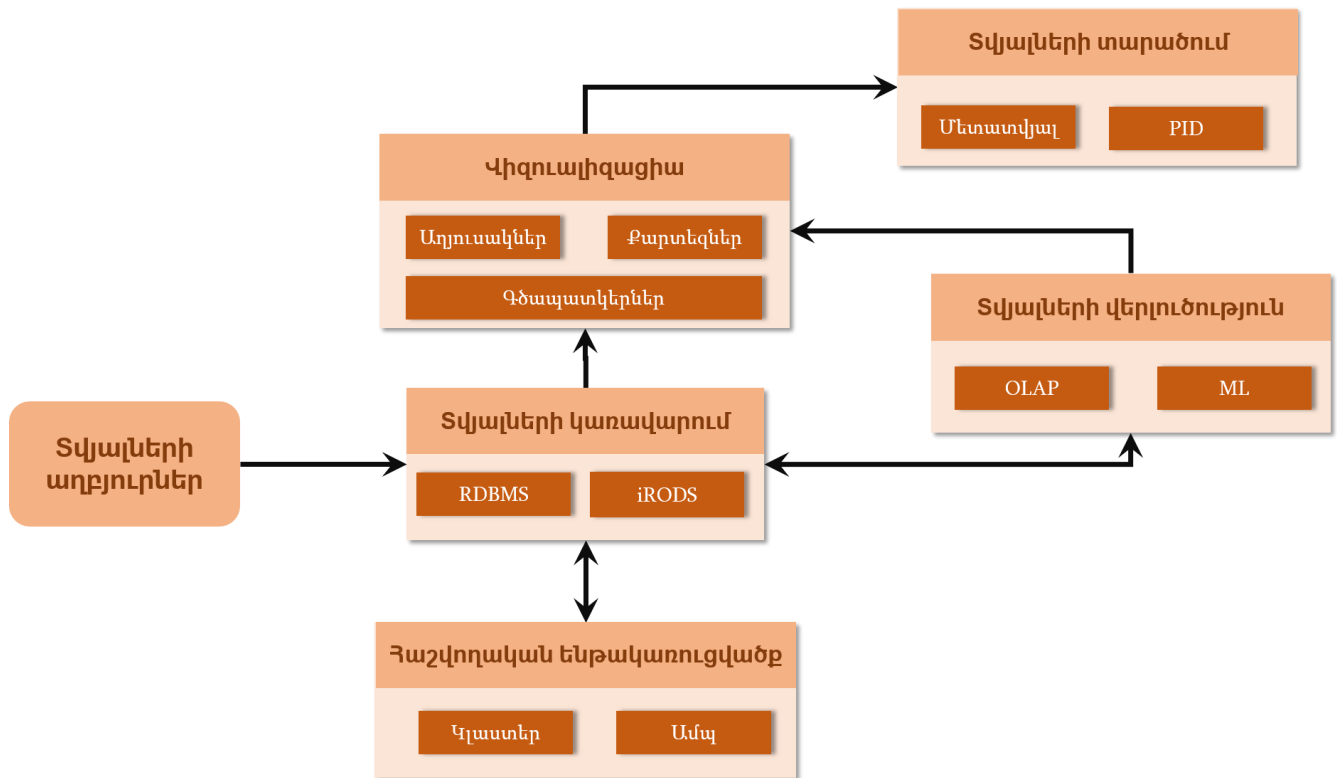
ՏՎՅԱԼՆԵՐԻ ՎԵՐԼՈՒԾՈՒԹՅԱՆ ԵՎ ՎԻԶՈՒԱԼԻԶԱՑԻԱՅԻ ՀԱՄԱԼԻՐ ՀԱՄԱԿԱՐԳ

Այս գլխում ներկայացված է տվյալների վերլուծության և վիզուալիզացիայի համար մշակված ծրագրային համալիր համակարգը (Data Analytics and Visualization System - DAVS), որը հնարավորինս բավարար չափով լուծում է երկրորդ գլխում ներկայացված խնդիրները և, որի կիրառության միջոցով կարելի է կառուցել տվյալների վերլուծությունների և ինտերակտիվ վիզուալիզացիաների համակարգեր վեբ միջավայրում և հնարավոր կլինի լուծել այնպիսի խնդիրներ, որոնք պարունակում են մեծ տվյալներ և բարձր արտադրողականությամբ հաշվարկներ՝ հաշվի առնելով գիտական ոլորտի կամ առաջադրված խնդրի պահանջները:

3.1 Նախաբան

Հաշվի առնելով նախորդ գլխում կատարված ուսումնասիրությունները, գոյություն ունեցող ծրագրային գործիքները, նրանց համեմատական վերլուծությունները մշակվել է գիտական խնդիրներում տվյալների վերլուծությունների և ինտերակտիվ վիզուալիզացիայի համակարգ, որը լուծում է բարձր արտադրողականությամբ հաշվարկների, մեծ ծավալով տվյալների մշակման և պահպանման խնդիրները՝ հնարավորություն տալով տվյալների վերլուծության և վիզուալիզացիայի գործիքների միջոցով ավելի արդյունավետ դարձնել տվյալների օգտագործումը և միջազգային հիմնօրինակներին համապատասխան ձևով տարածել ստացված արդյունքները: Կախված պահանջներից ու տվյալների տեսակներից և վիզուալիզացիայի նախընտրությունից, մշակված համակարգի ճկունությունը հնարավորություն է տալիս հարմարեցնել միջավայրը առաջադրված խնդրի պայմաններին:

Ստորև ներկայացված է ծրագրային համալիր համակարգի մոդելի բաղադրիչները (նկ. 3.1).



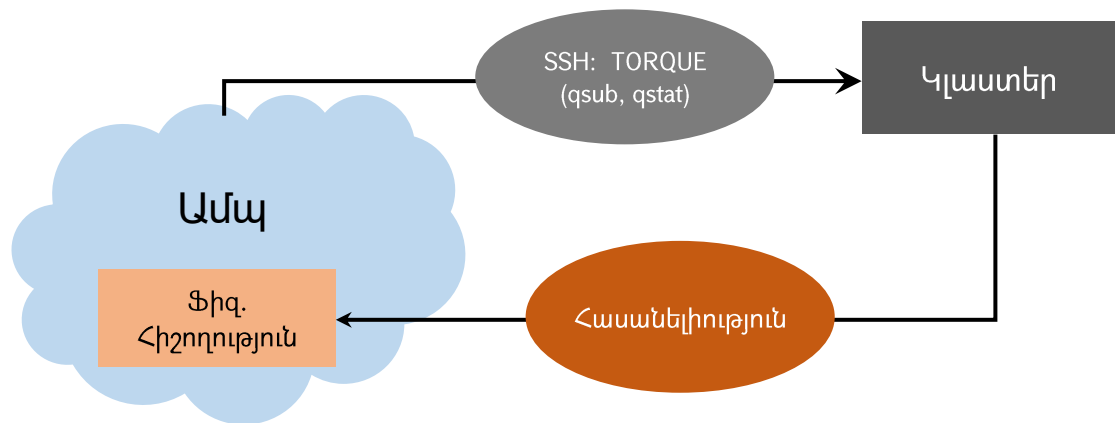
Նկար 3.1. Մշակված համալիր համակարգի մոդելի բաղադրիչները:

Տվյալների աղբյուրներից ստացված տվյալները պահպանվում են տվյալների կառավարման բաղադրիչում: Այնուհետև, կախված խնդրի դրվածքից, հաշվարկներ իրականացնելու անհրաժեշտության դեպքում հաշվողական ենթակառուցվածք բաղադրիչը, ստանալով անհրաժեշտ տվյալները տվյալների կառավարման բաղադրիչից, կատարում է համակարգում ներդրված խնդրի հաշվարկը և ավարտից հետո ստացված արդյունքները գրանցում է այդ բաղադրիչում: Եթե խնդիրը պարունակում է վերլուծական գործառույթ, ապա տվյալների հիման վրա կատարվում է վերլուծություն, և արդյունքը պահպանվում է տվյալների կառավարման բաղադրիչում կամ անմիջապես կատարվում է այդ տվյալների վիզուալիզացիա: Ստացված արդյունքների համար, անհրաժեշտության դեպքում, կարելի է ստեղծել մետատվյալներ և մշտական նույնացուցիչների միջոցով տարածել այն:

Այժմ դիտարկենք յուրաքանչյուր բաղադրիչ առանձին-առանձին՝ նշելով նրանց ընտրության պատճառներն ու համակարգի առավելությունները:

3.2 Հաշվողական ենթակառուցվածք

Հաշվողական ենթակառուցվածք բաղադրիչը հիմք է հանդիսանում մյուս բոլոր բաղադրիչների համար և իրականացնում է հաշվողական և ղեկավարման գործառույթ: Բարձր արտադրողականությամբ հաշվարկների իրականացման համար օգտագործվում է կլաստերային համակարգ, իսկ սերվերին, տվյալների պահուցներին, ծրագրերին, ծառայություններին և հաշվողական ռեսուրսներին ցանցային մուտքի հասանելիության ապահովման համար՝ վիրտուալ ռեսուրսներ տրամադրող ամպային տեխնոլոգիա: Վերջինիս կիրառությունը հնարավորություն է տալիս մշակելու ծառայություններ վեբ միջավայրում, որոնցից օգտագործողը կարող է օգտվել ցանկացած վայրից՝ առանց ներբեռնելու ծրագրային միջոցը:



Նկար 3.2 Կլաստերային և ամպային համակարգերի միջև փոխհարաբերությունների մոդել:

Այս երկու համակարգերը միմյանցից անկախ են և բարդություն է առաջանում նրանց միջև փոխհարաբերությունների ստեղծման: Այդ խնդիրը լուծելու համար մշակվել են ծրագրային միջոցներ, որոնց օգնությամբ ամպային համակարգից հաշվարկի համար անհրաժեշտ հարցումը կլաստերային համակարգում իրականացնելու համար օգտագործվում են Secure Shell (SSH) ցանցային արձանագրությունը և TORQUE ռեսուրսների կառավարիչի հրամանները (qsub, qstat) (նկ. 3.2):

Սկզբում ամպային միջավայրում համապատասխան առկա ֆիզիկական հիշողությունը կլաստերի համար դարձվում է հասանելի, քանի որ այնտեղ կատարվող հաշվարկները օգտագործում են հասանելի տիրույթում գտնվող տվյալները: Ամպային

միջավայրից յուրաքանչյուր նոր հաշվարկի հարցման համար ավելացվում է նոր տող համակարգի տվյալների բազայում: Նախապես պատրաստված ծրագրի միջոցով, օգտագործելով հեռահար SSH (Secure Shell) միացումը, օգտագործվում է “qsub” հրամանը այդ հաշվարկը կլաստերի հերթում ավելացնելու և իրականացնելու համար: “qsub” հրամանը որպես պատասխան տրամադրում է այդ հաշվարկի հերթում գտնվելու նույնացուցիչը, որը հետագայում օգտագործելով “qstat” հրամանում, ստուգվում է այդ աշխատանքի ընթացիկ վիճակը: Հաշվարկի ավարտից հետո ստեղծվում են երկու տեքստային ֆայլեր՝ error և out, որտեղ համապատասխան սխալի դեպքում գրվում է առաջացման պատճառը, իսկ out ֆայլում ծրագրի աշխատանքի ընթացքում գեներացված նշումները: Այսպիսով, որոշակի ժամանակային պարբերությամբ կանչելով “qstat” հրամանը կարելի է հայտնաբերել հաշվարկի ավարտը և կատարել համապատասխան գործողությունները՝ սխալի դեպքում արտածել պատճառները, իսկ հաջող կատարման դեպքում օգտագործել ստացված արդյունքները:

Տվյալ բաղադրիչը հնարավորություն է տալիս խնդրի համար մշակված միջավայրին տրամադրել բարձր արտադրողականությամբ հաշվարկներ, որն ավելի արդյունավետ կդարձնի խնդրի համար անհրաժեշտ տվյալների մշակումը, մեծ հաշվողականություն պարունակող ծրագրային հաշվարկների իրականացումը և ոչ ստրուկտուրավորված տվյալներից անհրաժեշտ ինֆորմացիայի հայտնաբերումը:

Փորձերի համար օգտագործվել է Centos 6.9 օպերացիոն համակարգով Linux կլաստեր 1 վերահսկիչ հանգույցով (1 միջուկ, 2GB RAM) և 8 աշխատող հանգույցներով (2 միջուկներ 8 GB RAM), CPU-ի մոդելը՝ Intel(R) Xeon(R) CPU E5-2620 v2@ 2.10GHz: Բոլոր հանգույցները փոխկապակցված են 1GB արագությամբ ինտերֆեյսով:

Ամպային հաշվարկների համար օգտագործվել է OpenStack համակարգը՝ մեկ վերահսկիչ և մեկ ցանցային հանգույցներ (Intel(R) Core(TM) i5-6500 CPU @ 3.20GHz 4 միջուկներ, 16GB RAM), 4 հաշվողական հանգույցներ (Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz 40 միջուկներ, 265GB RAM յուրաքանչյուրում): Բոլոր հանգույցները փոխկապակցված են 1GB արագությամբ ենթակառուցվածքով: Օգտագործվել է Centos 7 օպերացիոն համակարգը:

3.3 Տվյալների կառավարում

Տվյալների հետ աշխատելու համար անհրաժեշտ է ունենալ պահոց, որը հնարավորություն կտա կառավարելու այդ տվյալները: Տվյալները լինում են տարբեր տեսակների և տարբեր ձևաչափերի և, շատ հաճախ անհրաժեշտություն է առաջանում վերափոխել այն մեկ ձևաչափից մյուսին՝ ավելի ընթեռնելի և ծրագրային միջոցներին հասանելի դարձնելու համար: Այս նպատակով այս բաղադրիչը հնարավորություն է տալիս օգտագործել միջանկյալ ծրագրային միջոցներ այդ տեսակի բարդությունները հաղթահարելու համար (նկ. 3.3):



Նկար 3.3. Տվյալների կառավարման բաղադրիչի աշխատանքային մոդել:

Համակարգում ոչ ստրուկտուրավորված տվյալների պահպանման համար օգտագործվում է կանոններով ուղղորդվող ինտեգրված տվյալների համակարգը (iRODS - The integrated Rule-Oriented Data System) [35]: iRODS-ում պահվող տվյալների հետ աշխատելու համար տրամադրված է iRODS Client, որը ֆայլերը ներկայացնում է որպես տվյալի օբյեկտներ (Data Objects) հավաքագրված հավաքածուներում (Collections): Տվյալների արդյունավետ հայտնաբերման համար չափազանց օգտակար է յուրաքանչյուր տվյալի համար ունենալ մետատվյալ, հատկապես երբ աշխատում ենք մեծ ծավալի տվյալների բազմությունների հետ: iRODS-ում տվյալի օբյեկտը բացի համակարգի կողմից տրամադրված մետատվյալներից (ֆայլի անվանում, չափ, ստեղծման ամսաթիվ) կարող է ունենալ նաև օգտագործողի կողմից սահմանված մետատվյալներ: Այս հարուստ մետատվյալները թույլ են տալիս բացահայտել տվյալները այնպիսի հատկանիշներով, ինչպիսիք են հեղինակների անունները, բանալի բառերը և պարունակության տեսակը:

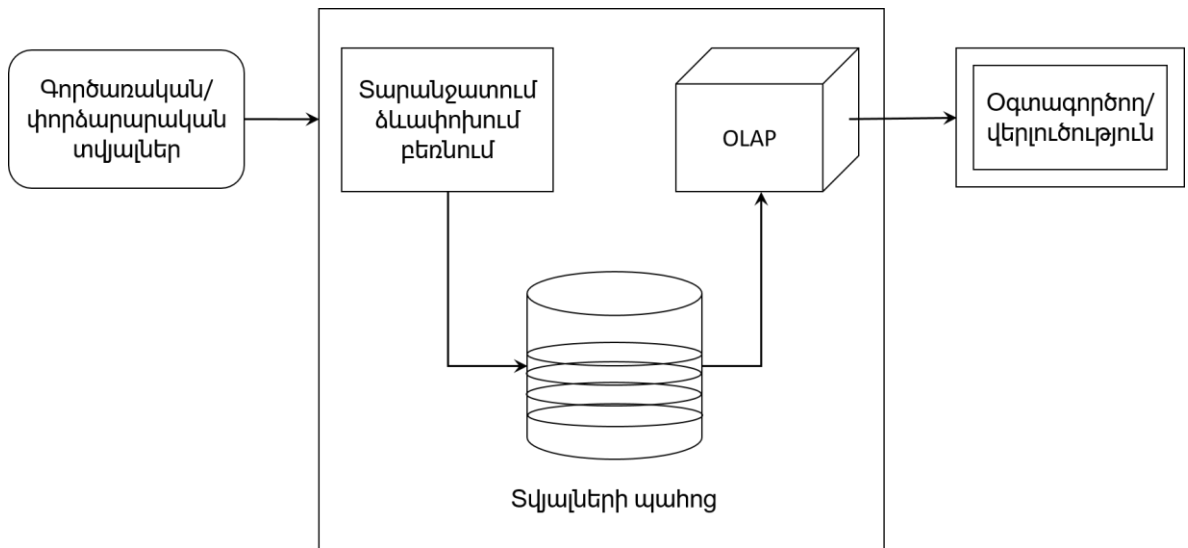
Ստրուկտուրավորված տվյալների պահպանման համար օգտագործվում է հարաբերական տվյալների բազայի կառավարման համակարգ (Relational Database Management System - RDBMS), որը տվյալների հարաբերական մոդելի հիման վրա

թվային բազա է: Այս մոդելում տվյալները պահվում են տողերով և սյուներով աղյուսակներում, որտեղ ամեն տող պարունակում է յուրահատուկ բանալի: Աղյուսակների միջև փոխգործակցության ստեղծման համար օգտագործվում են տրամաբանական կապեր հանդիսացող հարաբերություններ:

Մշակվել են մեթոդներ և ծրագրային միջավայրեր տվյալները բազայում պահելու համար: Տվյալների աղբյուրներից ստացված ինֆորմացիաները պահվում են iRODS-ում: Օգտագործելով Python ծրագրավորման լեզվով կառուցված iRODS Client-ը (python-irodscient [36]), ավտոմատ կերպով համակարգում նոր տվյալներին կցվում են մետատվյալներ, որոնց միջոցով հնարավորություն է տրվում արդյունավետ կատարել փնտրման հարցումներ: Մշակված միջանկյալ ծրագրային միջոցները՝ կախված տվյալի ձևաչափից և օգտագործելով iRODS-ի տրամադրած հրամանները, վերլուծում և պեղում են այդ տվյալներից անհրաժեշտ ինֆորմացիան և մուտքագրում հարաբերական բազայում, որտեղից և հետագայում օգտվում են վերլուծության կամ վիզուալիզացիայի բաղադրիչները:

3.4 Տվյալների վերլուծություն

Տվյալների ամենօրյա աճը և նրանց վերլուծության պահանջը նպաստել են վերլուծական միջոցների մշակմանը: OLAP-ը (Online analytical processing) հանդիսանում է տվյալների պահոցներում բազմաչափ տվյալների հասանելիության մոդել [37]: Տվյալների խորանարդները (Cubes) և սեսիաները հանդիսանում են OLAP-ի հասկացություններ: Տվյալների խորանարդը տվյալների պահոցի տվյալները կազմակերպող փաստերի և չափերի հավաքածու է՝ ըստ տարբեր վերլուծության առանցքների և միաձուլման (aggregation) չափումների: OLAP-ը տրամադրում է մի շարք գործողություններ (drill-down և slice-and-dice), որոնք փոխարկում են մեկ բազմաչափ հարցումը մյուսի, որն ապահովում է բարձր որակի որոնում (նկ. 3.4):



Նկար 3.4. Տվյալների հոսք:

Որպես հարցումների լեզու օգտագործվում են բազմաչափ արտահայտություններ (MDX – MultiDimensional eXpression) [38], որոնք հնարավորություն են տալիս փոխազդել և իրականացնել հարցումներ՝ օգտագործելով բազմաչափ տվյալների բազաներ (OLAP Cubes): OLAP հարցումները ձևավորվում են որպես հաջորդականություններ, որոնք կոչվում են OLAP սեսիաներ: Տարածական և աշխարհագրական բաղադրիչներով տվյալների վերլուծության համար օգտագործվում է Spatial OLAP (SOLAP) տեխնոլոգիան, որը հնարավորություն է տալիս արագ և հեշտ տեղաշարժվել տարածական տվյալների բազաներում և տրամադրում է ինֆորմացիայի բազմաբնույթության բազմաթիվ մակարդակներ, թեմաներ և սինխրոնացված և ոչ սինխրոնացված ցուցադրման ռեժիմներ՝ քարտեզներ, աղյուսակներ և դիագրամներ [39]: Այն թույլ է տալիս GIS և OLAP համակարգերի սերտ ինտեգրում: SOLAP համակարգը հնարավորություն է տալիս աշխատել 3 տեսակի տարածական չափերի հետ՝ ոչ երկրաչափական տարածական, երկրաչափական տարածական և խառը տարածական: OLAP սեսիայի ընթացքում օգտագործողը վերլուծում է հարցման արդյունքները և կախված կոնկրետ տվյալներից, կիրառում է գործողություն, որոշելու համար նոր հարցումը, որը կտա ինֆորմացիայի մասին ավելի լավ տեղեկատվություն:

Mondrian-ը հանդիսանում է OLAP սերվեր, մշակված Pentaho ընկերության կողմից, բաց աղբյուրի կոդերով, գրված JAVA ծրագրավորման լեզվով: Mondrian սերվերը

բաղկացած է երեք շերտերից՝ չափման, աստղային և պահպանման: Չափման շերտը վերլուծում, ստուգում և իրականացնում է MDX հարցումները: Աստղային շերտը պատասխանատու է համակցված քաշի պահպանման համար: Պահպանման շերտը ապահովում է ագրեգատների ձեռքբերման համար անհրաժեշտ հումքի պահեստավորում: Մշակված միջանկյալ ծրագրային միջոցը հնարավորություն է տալիս ստեղծել Cube-եր և MDX հարցումներ, որոնք իրականացվում են Mondrian սերվերում:

Գոյություն ունեն խնդիրներ, որտեղ հաջորդական հաշվարկները կամ արդյունքները ունենում են նախատիպեր, որոնք կարելի է տրամաբանորեն կապել միմյանց հետ: Օրինակ, եղանակի կանխատեսման տվյալները, որոնք ստացվում են որոշակի թվային մեթոդների միջոցով, հաճախ ունենում են բավականին մեծ շեղվածություն իրական տվյալներից, որը կախված է օգտագործված մոդելի կարգավորումներից տվյալ տարածաշրջանի համար: Սակայն, քանի որ յուրաքանչյուր տարածաշրջան ունի իրեն հատուկ ռելիեֆ և եղանակային պայմաններ, տարբեր տարիների եղանակային դիտարկումների միջև կարելի է տեսնել տրամաբանական կապ: Այս և նմանատիպ դեպքերի համար կարելի է կատարել տվյալների ուղղում և օպտիմիզացիա՝ կանխատեսման ենթարկվող ալգորիթմների մշակումը նեյրոնային ցանցերի միջոցով:

Google AI կազմակերպության Google Brain թիմի հետազոտողների և ինժեներների կողմից մշակված TensorFlow-ն հանդիսանում է բաց կոդով ծրագրային գրադարան, բարձր արդյունավետության թվային հաշվարկներ իրականացնելու համար: Գրադարանը կիրառվում է նաև նեյրոնային ցանցերի ուսուցում իրականացնելու համար: Այս բաղադրիչը հնարավորություն է տալիս անհրաժեշտության դեպքում կարգավորել և օգտագործել նեյրոնային ցանցերը կանխատեսման ենթարկվող ալգորիթմների մշակման համար: Մշակված մեթոդների միջոցով տվյալները կառավարման բաղադրիչից կարելի է ստանալ նեյրոնային ցանցի ուսուցման համար անհրաժեշտ ձևաչափում:

3.5 Տվյալների վիզուալիզացիա

Այս բաղադրիչը իրականացնում է տվյալների ինտերակտիվ վիզուալիզացիա, հիմնվելով տվյալների կառավարման կամ վերլուծության բաղադրիչների տվյալների վրա: Մշակված մեթոդները հնարավորություն են տալիս օգտագործելով վիզուալիզացիայի վեբ գրադարանները ստանալ տարբեր տեսակի արտացոլումներ սկսած պարզագույն գրաֆիկներից մինչև բարդ կառուցվածքով պատկերներ՝ 2D և 3D ձևաչափերում: Փորձերի ժամանակ օգտագործվել են D3.js և Vis.js վիզուալիզացիայի գրադարանները:

Լավ և ճշգրիտ վիզուալիզացիաների նախագծման գործընթացը կարելի է բաժանել տարբեր քայլերի: Տվյալ բաղադրիչում աշխատանքի հիմքում ընկած են [40]–ում նշված 6 քայլերը.

1. Ձևափոխում – Առաջին քայլն է, որը նշանակում է, թե ինչպես արտացոլել ինֆորմացիան կամ ինչպես կոդավորել այն տեսողական ձևի: Լավ ձևափոխումը ապահովում է ճշգրիտ տեսողական ներկայացում, որին կարելի է հասնել տվյալների օբյեկտների և տեսողական օբյեկտների միջև ճշգրիտ հարաբերությունների ներկայացման միջոցով:
2. Ընտրություն – Ընտրել անհրաժեշտ տվյալները տրված խնդրին կամ աշխատանքին համապատասխան հասանելի տվյալներից: Տվյալների ընտրությունը ուղղակիորեն կախված է սպասվող վիզուալիզացիայի տեսակներից: Այս քայլը ամենակարևորն է, քանի որ սխալ ընտրությունը կհանգեցնի սխալ որոշումների կայացմանը: Այդ իսկ պատճառով պետք է խուսափել ավելորդ տվյալների ներգրավումից:
3. Ներկայացում – Վիզուալիզացիայի տեսանկյունից նշանակում է արդյունավետ կերպով կառավարել, կազմակերպել ինֆորմացիայի արտացոլումը էկրանի հասանելի հատվածում: Ինտուիտիվ ձևափոխումից և տվյալների հստակ և ճշգրիտ ընտրությունից հետո շատ կարևոր է ներկայացնել տվյալների վիզուալիզացիան իմաստալից և հասկանալի ձևով:

4. Ինտերակտիվություն – Նշանակում է, թե ինչպիսի հատկություններ են տրամադրված վիզուալիզացիայի կազմակերպման, ուսումնասիրության և վերադասավորման համար: Ինտերակտիվությունը օգտատիրոջը հնարավորություն է տալիս լավագույնս ուսումնասիրել, հասկանալ և մեկնաբանել տվյալները, որոնք բարելավում են հետազոտական կարողությունները:
5. Մարդկային գործոն – Ներառում է երկու լայն կատեգորիա՝ օգտագործելիություն և մատչելիության: Տեսողական ընկալման գիտելիքները և ճանաչողականության ասպեկտները հեշտացնում են արդյունավետ արտացոլման նախագծումը:
6. Գնահատում – օգտագործելի տեսողական ինտերֆեյս ստեղծելուց հետո վերջին քայլն է հանդիսանում գնահատել ստեղծված արդյունքը: Այն անհրաժեշտ է պարզելու համար, թե արդյոք վիզուալիզացիայի մեթոդը ունի արդյունավետություն և հասել է իր նպատակին, թե՛ ոչ:

Մշակված API-ի միջոցով, օգտագործողի միջավայրը ստանում է տվյալները սերվերային հատվածից և կառուցում վիզուալիզացիաներ, օգտագործելով խնդրի պահանջներին բավարարող տեսակներ: Լրացուցիչ ստեղծված մեթոդները հնարավորություն են տալիս ստանալ տվյալները տարբեր ձևաչափերով, քանի որ յուրաքանչյուր գրադարան պահանջում է իր մուտքին տրամադրել տվյալները համապատասխան տեսքով:

3.6 Տվյալների տարածում

Տվյալները հիմք են հանդիսանում ճիշտ ու արդյունավետ գիտական որոշումների կայացման, կառավարման, ռեսուրսների օգտագործման համար: Գիտական տվյալների հսկայական չափի պատճառով լավ պրակտիկա է տարածել ոչ միայն տվյալները, այլ նաև մետատվյալները (տվյալների մասին տվյալները): Մետատվյալները ներկայացնում են տվյալի մասին բազային ինֆորմացիան, որը դյուրին է դարձնում անհրաժեշտ տվյալների փնտրումն ու աշխատանքը նրանց հետ: Այն ստրուկտուրավորված ինֆորմացիա է, որը հայտարարում, գտնում, բացատրում կամ, այլ կերպ ասած, հեշտացնում է տեղեկատվական ռեսուրսի ստացումը, օգտագործումն ու կառավարումը: Այն նկարագրում է տարբեր տեսակի ռեսուրսներ, ինչպես միավոր, այնպես էլ

հավաքածուներ կամ նույնիսկ ավելի մեծ ռեսուրսների հատվածներ: Բազային փաստաթղթի մետատվյալների օրինակներ են հեղինակը, փոփոխման ամսաթիվը և ֆայլի չափը: Մետատվյալների մեջ փնտրելու հնարավորությունը օգնում է հեշտորեն գտնել անհրաժեշտ տվյալները: Ինտենսիվ տվյալներով գիտությունների դեպի ավտոմատացված պրոցեսների շարժման փուլում, մշտական նույնացուցիչների (PID) օգտագործումը բոլոր տեսակի թվային օբյեկտների համար դառնում է կրիտիկական [41]: Ներկայումս գոյություն ունեցող լուծումները տրամադրում են տվյալների տարածման համար բավարար ծառայություններ, սակայն սովորաբար նրանք ունենում են բարդ կառուցվածքներ և գիտնականը, որը չունի բավական գիտելիքներ այդ տեսակի ծառայություններից օգտվելու համար, կարող է հանդիպել բարդությունների: Նաև այդ ծառայությունները կարող են տրամադրել մետատվյալների այնպիսի կառուցվածք, որը կարող է չհամապատասխանել տվյալ խնդրի արդյունքների կառուցվածքի պահանջներին:

FAIR տերմինը առաջին անգամ օգտագործվել է 2014 թ.-ին Լորենցի սեմինարի ժամանակ, որի արդյունքում ստացված FAIR սկզբունքները հրապարակվել են 2016 թվականին [42]: Այս սկզբունքների նպատակն է տվյալները դարձնել հայտնաբերվող (Findable), մատչելի (Accessible), համատեղելի (Interoperable) և վերամշակելի (Reusable).

- **Հայտնաբերվող (Findable)**

F1. (մետա) տվյալներին կցվում է գլոբալ եզակի և հավերժական մշտական նույնացուցիչներ:

F2. տվյալը նկարագրվում է հարուստ մետատվյալով:

F3. (մետա) տվյալները գրանցված են կամ ինդեքսավորվում են որոնելի ռեսուրսում:

F4. մետատվյալը որոշում է տվյալի նույնացուցիչը:

- **Մատչելի (Accessible)**

A1. (մետա)տվյալները ստացվում են իրենց նույնացուցիչով՝ ստանդարտացված հաղորդակցությունների արձանագրության միջոցով:

A1.1. արձանագրությունը բաց է, ազատ և համընդհանուր կիրառելի:

A1.2. արձանագրությունը անհրաժեշտության դեպքում թույլ է տալիս վավերացման գործընթաց:

A2. մետատվյալները հասանելի են, նույնիսկ եթե տվյալները այլևս հասանելի չեն:

- **Համատեղելի (Interoperable)**

I1. (մետա)տվյալները օգտագործում են ֆորմալ, մատչելի, համօգտագործելի և լայնորեն կիրառելի լեզու գիտելիքների ներկայացման համար:

I2. (մետա)տվյալները օգտագործում են բառապաշարներ, որոնք համապատասխանում են FAIR սկզբունքներին:

I3. (մետա)տվյալները ներառում են որակյալ հղումներ այլ (մետա)տվյալների վրա:

- **Վերամշակելի (Reusable)**

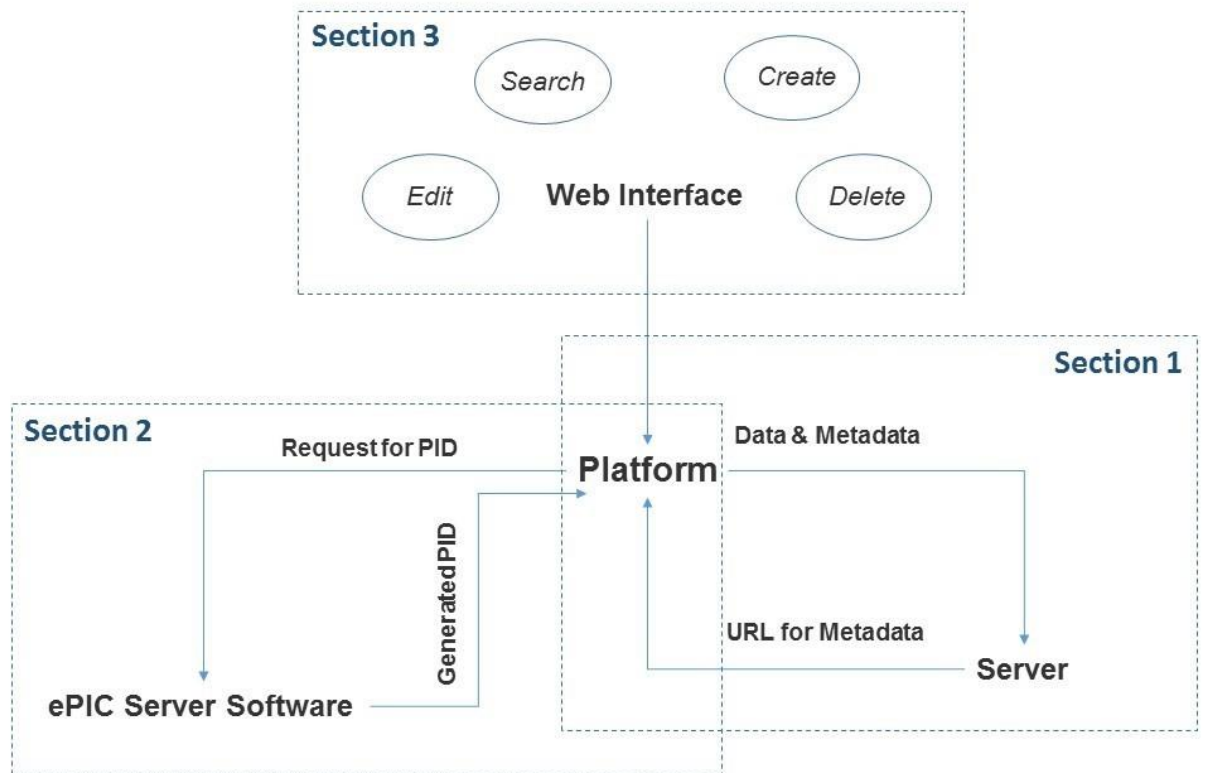
R1. մետա(տվյալները) ունեն ճշգրիտ և համապատասխան հատկանիշների բազմազանություն:

R1.1. (մետա)տվյալները թողարկվում են հստակ և հասանելի տվյալների օգտագործման արտոնագրով:

R1.2. (մետա)տվյալները կապվում են իրենց աղբյուրի հետ:

R1.3. (մետա)տվյալները համապատասխանում են համայնքի համապատասխան չափանիշներին:

Ստորև ներկայացված է FAIR սկզբունքներ ապահովող մշակված համակարգի սխեմատիկ ներկայացումը [43] (նկ. 3.5):



Նկար 3.5. Ենթակառուցվածքի մոդելը:

Ենթակառուցվածքի մոդելը բաղկացած է երեք բաժիններից.

Բաժին 1: Գրանցված համայնքները կարող են իրենց տվյալների համար ձևավորել մետատվյալներ և հղում տալ իրենց ռեսուրսների վրա: Մետատվյալները պահվում են լոկալ սերվերում և ստեղծվում է URL նրանց համար:

Բաժին 2: Օգտագործելով ePIC-ի կողմից տրամադրած API-ը, համակարգը հարցում է կատարում գեներացնելու համար նոր PID՝ տրամադրելով Բաժին 1-ում ստեղծված URL-ը: ePIC API-ը հնարավորություն է տալիս կառուցել PID՝ նշելով GUID-ի ընտրելի նախածանց և վերջածանց: Քանի որ այն աշխատում է օգտագործելով ePIC-ի կողմից տրամադրված միևնույն հաշիվը՝ ստացված PID-երի համար նախածանցը ունենում է նույն կառուցվածքը:

<Registered INST in ePIC system>-<GUID>-<Registered community INST in our system>

Բաժին 3: Ենթակառուցվածքը կառուցված PID-ը կցում է սերվերում պահված մետատվյալի հետ և օգտատերերը կարող են փոփոխել իրենց մետատվյալներն առանց փոփոխելու տվյալ PID-ը: Նրանք նաև կարող են ջնջել ունեցած մետատվյալները, որն

ավտոմատ կերպով կջնջի նաև PID-ը ePIC-ի համակարգից: Տրամադրված է նաև որոշման հնարավորություն ինչպես ePIC-ի PID-երի համակարգում, այնպես էլ տեղական սերվերում:

Մշակված տվյալների տարածման միջավայրը տրամադրում է API, որի օգտագործմամբ մեկ այլ համակարգ կարող է մետատվյալների հիման վրա ստեղծել մշտական նույնացուցիչներ:

3.7 Եզրակացություն

Մշակված տվյալների վերլուծության և վիզուալիզացիայի համալիր համակարգը նախատեսված է իրականացնելու բարձր արտադրողականությամբ հաշվարկների, տվյալների պահպանման, վերլուծության, վիզուալիզացիայի և տարածման գործառույթները մեկ համակարգում այն խնդիրների համար, որտեղ առկա են մեծ տվյալներ և բարձր արտադրողականությամբ հաշվարկներ: Համակարգում օգտագործվել են ժամանակակից ծրագրային միջոցները և հաշվի են առնվել միջազգային հիմնօրինակները: Համակարգի ճկուն կառուցվածքը հնարավորություն է տալիս հարմարեցնել այն տարատեսակ խնդիրների պահանջներին, որը հնարավորություն կտա հետազոտողներին խնայել ժամանակ և հարմարավետ միջավայրում կատարել դրված խնդրի համար անհրաժեշտ գործառույթները:

ԳԼՈՒԽ 4

ՄՇԱԿՎԱԾ ՀԱՄԱԼԻՐ ՀԱՄԱԿԱՐԳԻ ՓՈՐՁԱՐԿՈՒՄՆԵՐ

Այս գլխում ներկայացված են առաջարկվող համալիր համակարգի հենքի վրա մշակված ամպային ծառայությունները երկու տարբեր ոլորտներում՝ քվանտային ֆիզիկայում ֆոտոնիկ դիսիպատիվ համակարգերի թվային հաշվարկների ու վիզուալիզացիայի և եղանակի կանխատեսման տվյալների վիզուալիզացիայի և նրանց համապատասխանելիության վերլուծության համար: Նկարագրված են խնդիրները և ամպային ծառայությունների կառուցվածքները: Երկու դեպքում էլ բերված են կիրառության օրինակներ, մշակված ամպային ծառայությունների արդյունավետությունը և օգտակարությունը ընդգծելու համար:

4.1 Ամպային ծառայություն ֆոտոնիկ դիսիպատիվ համակարգերի թվային հաշվարկների և վիզուալիզացիայի համար

Ներկայումս քվանտային ֆիզիկան շատ կարևոր է մի շարք գիտական ծրագրերի համար, որտեղ այլևս հնարավոր չէ անտեսել շրջակա միջավայրի փոխազդեցությունը: Նման դեպքերում սովորաբար քվանտային համակարգերը դիտվում են որպես բաց համակարգեր և նրանց ժամանակային էվոլյուցիան նկարագրվում է հիմնական հավասարման շրջանակներում խտության մատրիցայով՝ Հիլբերտի տարածության վեկտորի և Շրեդինգերի հավասարման փոխարեն: Այս քվանտային համակարգերի վիզուալիզացիան թույլ է տալիս օգտագործողներին հաշվարկել և ուսումնասիրել այնպիսի պարամետրերի զգայունությունը, ինչպիսիք են՝ գրգռված ֆոտոնների թիվը, ֆոտոնների թվի բաշխման ֆունկցիան և Վիգների ֆունկցիաները: Ֆոտոնիկ դիսիպատիվ համակարգերի թվային հաշվարկների և վիզուալիզացիայի համար ներկայացված ամպային ծառայությունը թույլ է տալիս կատարել ցանկացած ժամանակային կախվածությամբ Համիլիտոնյանների լայն տեսականու թվային հաշվարկներ և վիզուալիզացիա, որոնք լայնորեն օգտագործվում են շատ ֆիզիկական ծրագրերում [44]: Մշակված ամպային ծառայությունը հնարավորություն է տալիս կառուցել գրաֆիկներ և դիագրամներ՝ բարդ համակարգերի փոխազդեցության համար, և ժամանակի էվոլյուցիայի ընթացքում կատարել թվային հաշվարկներ:

4.1.1 Նախաբան

Ընդհանուր առմամբ, հիմնական հավասարումը չի կարող անալիտիկորեն լուծվել շատ ֆիզիկական համակարգերի համար կամայական գնահատման ժամանակում, այդ իսկ պատճառով առաջանում է թվային մեթոդների օգտագործման անհրաժեշտություն: Այնուամենայնիվ, անգամ հիմնական հավասարման թվային լուծումը խնդիր է: Եթե վիճակի վեկտորը հիլբերտյան D տարածությունում ներկայացնելու համար պահանջվում է հիմնական վեկտոր, համապատասխան խտության օպերատորը պահանջում է իրական թվեր: Այս բազան պահանջում է խոշորամասշտաբ հաշվարկներ, հատկապես, երբ համակարգը ներառում է իր մեջ մեկից ավելի ազատության աստիճան: Այս խնդիրը կարող է լուծվել խտության օպերատորի էվոլյուցիան բաժանելով քվանտային հետազդերի և Quantum State Diffusion (QSD) մեթոդներով [45]: Քանի որ քվանտային հետազդերը ներկայացնում են համակարգը որպես վիճակի վեկտոր, այլ ոչ թե խտության օպերատոր, նրանք հաճախ ունեն թվային առավելություն հիմնական հավասարումը լուծելու նկատմամբ, անգամ, եթե հարկ է լինում միջինացնել շատ քվանտային հետազդեր՝ հիմնական հավասարման լուծումը վերականգնելու համար: Բաց քվանտային ρ համակարգերի համար առավել լայն տարածված մոտավորություններից մեկն է Մարկովյան դինամիկան, որը նկարագրվում է նվազեցված խտության մատրիցայի համար՝ Լինդբլադի հիմնական հավասարման տեսքով: Այս դեպքում ժամանակի էվոլյուցիան որոշվում է համակարգի Համիլտոնիան օպերատորով և Լինդբլադյան օպերատորներով: QSD մեթոդն իրականացվում է քառասային հավասարման հիման վրա, որը $|\Psi(t)\rangle$ վիճակի համար ներառում է ինչպես Համիլտոնիան, այնպես էլ Լինդբլադյան օպերատորները: Հաշվարկվում է խտության օպերատորը, որն օգտագործում է վիճակի $|\Psi\rangle$ վեկտորի ընդլայնումը հարմոնիկ օսցիլյատորի (ֆոտոնային վիճակ) Ֆոկի թվերի վիճակի սեղմված հիմքում:

Մշակված է ECAPT (Extendible C++ Application in Photonic Technologies) գրադարան և ինտերնետային կայք [46], որում ներկայացված է կոմպլեքս քվանտային համակարգերի թվային հաշվարկ և մոդելավորում՝ դեկոհերենցիայի առկայությամբ ֆոտոնիկայի մի շարք կիրառություններում: Ընդլայնվել է պորտալի ֆունկցիոնալությունը, տրամադրելով ամպային ծառայություն և վերլուծություն՝

գնահատելու համար կոմպլեքս քվանտային համակարգերի տարբեր ֆիզիկական արժեքներ: Ծառայությունը թույլ է տալիս օգտագործողներին, ներառելով գրգռված ֆոտոնների թիվը, հավանականությունը և Վիգների ֆունկցիաները, մոդելավորել և գրաֆիկորեն ղեկավարել արդյունքները տարբեր պարամետրերի և ժամանակի ինտերվալների համար:

4.1.2 Կեռի ոչ գծային ռեզոնատորը չիրպացված արտաքին դաշտի ազդեցությամբ

Այս բաժնում Կեռի ոչ գծային չիրպացված ռեզոնատորը անընդհատ ալիքի տարածման ռեժիմում ներկայացվում է հիմնական հավասարման բազայի վրա, ինքնառեզոնանսային անցումների և փուլային փականքի երևույթների ամբողջական քվանտային վերլուծությունում՝ դիսիպացիայի առկայության դեպքում: Այս մոտեցումը թույլ է տալիս հաշվի առնել ֆոտոնի թվի վիճակների արտադրության վերահսկումը, և ֆոտոնների արգելափակումը ուժեղացնել ռեզոնատորի արագության հաճախականությունը ինքնակարգավորելու միջոցով: Կեռի ոչ գծային տարրի ֆոտոն-ֆոտոն փոխհարաբերության վրա հիմնված ֆոտոնային արգելափակման մեջ (Photon Blockade - PB) մեկ ֆոտոնի օպտիկական արձագանքը մոդելավորվում է այլ ֆոտոնների ներկայությամբ կամ բացակայությամբ: Հիմնականում համակարգում մեկ ֆոտոնի ներգրավումը ազդում է երկրորդ ֆոտոնի ժամանման հավանականության վրա: Ֆոտոնային արգելափակման պարզ հետևանք է արտանետման ժամանակ ֆոտոնների անբավարարությունը՝ արտանետման երկմակարդակ ատոմի ռեզոնանսային ֆլուորեսցենայի նմանօրինակությամբ [47]: Ֆոտոնային արգելափակումը առաջին անգամ դիտարկվել է օպտիկական խոռոչում՝ զուգորդված մեկ արգելափակված ատոմով [48]: PB-ն կանխատեսված էր քվանտային էլեկտրադինամիկայի (QED) խոռոչում [49], իսկ վերջերս միկրոալիքային հաղորդման գծի ռեզոնատորի հետ կապված մեկ գերհաղորդիչ արհեստական ատոմով QED-ի շրջանում [50]: PB-ն նաև փորձնականորեն ցուցադրվել է ֆոտոնիկ բյուրեղյա խոռոչի հետ՝ ուժեղ միացված քվանտային կետով, ինչպես նաև կանխատեսվել է քվանտային օպտոմեխանիկական համակարգերում: Ֆոտոնային արգելափակման նմանատիպ երևույթ կանխատեսվել է

նանումեխանիկական ռեզոնատորին զուգակցված արհեստական գերհաղորդիչ ատոմի համար, ինչպես նաև բևեռատոն-բևեռատոնային փոխազդեցությունների միջոցով բևեռատոնի ազդեցությունը դիտարկվել է [51]-ում: Վերջերս PB-ն դիտարկվել է qubit դաշտի ցրված փոխազդեցություններում գերհաղորդիչ կոպլանար ալիքային գոտում և ժամանակի մոդուլացված մուտքով [52]: Մենք հստակեցնում ենք Կեռի ոչ գծային ռեզոնատորում ճառագայթային ազդեցությունները, հաշվի առնելով ֆոտոնային թվերի ազդեցությունները և վերլուծելով ռեզոնատորային ռեժիմի փուլային տարածքային հատկությունները: Այսպիսով, մենք կենտրոնանում ենք ֆոտոնների միջին թվի (գրգռվածության թիվ), ֆոտոնների բաշխվածության հավանականության և Վիգների ֆունկցիաների վերլուծության վրա ֆազային տարածությունում:

Մենք դիտարկում ենք չիրպացված Կեռի ոչ գծային ռեզոնատորը որպես բաց քվանտային համակարգ և ենթադրում, որ ժամանակի էվոլյուցիան նկարագրված է Մարկովյան դինամիկայով՝ օգտագործելով Լինդբլադի հիմնական հավասարումը՝ խտության մատրիցը կրճատելու համար: Փոխազդեցության պատկերում, որը համապատասխանում է $\rho \rightarrow e^{-i\omega a^\dagger a t} \rho e^{i\omega a^\dagger a t}$ վերափոխմանը, որտեղ $a^\dagger$ և a օսցիլյատորի ռեժիմի Bose-ի ստեղծման և ոչնչացման օպերատորներն են և գրգռման հաճախականությունը, հավասարումը Մարկովյան ձևում ստացվում է.

$$\frac{d\rho}{dt} = \frac{-i}{\hbar} [H, \rho] + \sum_{i=1,2} (L_i \rho L_i^\dagger - \frac{1}{2} L_i^\dagger L_i \rho - \frac{1}{2} \rho L_i^\dagger L_i) \quad \text{Հավասարում 1}$$

Պտտվող ալիքի մոտավորությունը իմպուլսային գրգռվածությամբ Կեռի ոչ գծային ռեզոնատորի համար Համիլտոնիանը տրվում է.

$$H = (\Delta - \delta t) a^\dagger a + \chi (a^\dagger)^2 a^2 + (\Omega a^\dagger + \Omega^* a) \quad \text{Հավասարում 2}$$

Այստեղ, ժամանակային կախվածությամբ Ω կապի հաստատունն է, որը համարժեք է գրգռվածության դաշտի ամպլիտուդային, մինչդեռ $a^\dagger$, a Bose-ի ստեղծման և ոչնչացման օպերատորներն են, χ -ն ոչ գծայնության ուժն է, $\Delta = \omega_0 - \omega$ գրգռվածության դաշտի միջին հաճախականության և օսցիլյատորի հաճախականության միջև շեղումն է, δ -ն չիրպի պարամետրն է: Դիսիպացիայի և դեկոհերենսի ազդեցությունները,

կորուստները և ջերմային աղմուկը ընդգրկված են այս հավասարման վերջին մասում, որտեղ L_i -ն Լինդբլադի օպերատորն է՝

$$L_1 = \sqrt{(N+1)\gamma}a, \quad L_2 = \sqrt{N\gamma}a^\dagger \quad \text{Հավասարում 3}$$

որտեղ γ -ն դիսիպացիոն գործընթացի ինքնական տրոհումն է, N -ը ջերմային խոռոչի քվանտների միջին թիվը:

Այս մոդելը փորձականորեն կիրառելի է և հնարավոր է իրականացնել մի քանի ֆիզիկական համակարգերում: Մասնավորապես, արդյունավետ Համիլտոնիանը (Հավասարում 2) նկարագրում է խոռոչի հետ ոչ ռեզոնանսային կապված քուբիթը: Իրականում հայտնի է, որ ցրված մոտեցման մեջ երկմակարդակ ատոմի և խոռոչի ռեժիմի փոխազդեցության Համիլտոնիանը կարող է կրճատվել մինչև արդյունավետ Համիլտոնիանի (Հավասարում 2): Այս մոդելը նկարագրում է նաև նանոմեխանիկական օսցիլյատորը $a^\dagger$ և a օպերատորներով՝ կապված քվանտային շարժման դիրքի և իմպուլսի օպերատորների հետ: Վերջինս ստացվել է Կեռի տիպի ռեզոնատորի կարևոր իրականացում գերհաղորդիչ սարքերի կոնտեքստում, որը հիմնված է ոչ գծային Ջոզեֆսոնի անցման վրա:

Հիմնական բանաձևը թվայնորեն վերլուծվել է քվանտային դիֆուզիոն մեթոդի միջոցով [53]: Այս մոտեցման համաձայն՝ նվազեցված խտության օպերատորը հաշվարկվում է որպես համույթի միջին արժեք:

$$\rho(t) = M(|\psi_\xi(t)\rangle\langle\psi_\xi(t)|) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{\xi}^N |\psi_\xi(t)\rangle\langle\psi_\xi(t)| \quad \text{Հավասարում 4}$$

Ստոխաստիկ մաքուր վիճակների վրա քվանտային հետազոտվ էվոյուցիան նկարագրող $|\psi_\xi(t)\rangle$ վիճակի ստոխաստիկ հավասարումը ներառում է և Համիլտոնիանը (Հավասարում 2), և Լինբլանդի օպերատորները (Հավասարում 3): Մենք հաշվարկում ենք խտության օպերատորը՝ օգտագործելով $|\psi_\xi\rangle$ վիճակի վեկտորի ընդլայնումը՝ Ֆոկի ներդաշնակ օսցիլյատորի թվային վիճակների կտրված հիմքում:

$$|\psi_{\xi}(t)\rangle = \sum_n a_n^{\xi}(t)|n\rangle \quad \text{Հավասարում 5}$$

Ֆոտոնների թվի բաշխումը հաշվարկվում է որպես ֆոտոն-թվերի $|n\rangle$ վիճակների անկյունագծային տարրեր.

$$P(n) = \langle n|\rho|n\rangle \quad \text{Հավասարում 6}$$

մինչդեռ Վիգների ֆունկցիայի հաշվարկները կատարվում են օգտագործելով իր ստանդարտ ձևը Ֆոկի տարածությունում.

$$W(\rho, \theta) = \sum_{m,n} \rho_{mn} W_{mn}(\rho, \theta) \quad \text{Հավասարում 7}$$

Այստեղ ρ, θ -ն բարդ ֆազային տարածությունում բևեռային կոորդինատներն են՝ $x = \frac{a+a^+}{\sqrt{2}}, y = \frac{(a-a^+)}{\sqrt{2}i}$, իսկ $W_{mn}(\rho, \theta)$ գործակիցները՝ Վիգների բնորոշ ֆունկցիայի մատրիցային տարրերի Ֆուրյեի ձևափոխություններն են:

4.1.3 Ամպային ծառայության ենթակառուցվածք

Ներկայացված վեբ հիմքով բազմաօգտատեր թվային հաշվարկների իրականացման և արդյունքների վիզուալիզացիայի միջավայրը կառուցվել է օգտագործելով մշակված DAVS համակարգը: Հաշվողական ենթակառուցվածքը ապահովում է HPC-ն և տվյալների ռեսուրսները, որոնք հաշվարկները իրականացնում են կլաստերային համակարգում: Օգտագործվել են Հայաստանի էլեկտրոնային ենթակառուցվածքի ռեսուրսները, որոնք հանդիսանում են բարդ ազգային SS ենթակառուցվածք, որը բաղկացած է հաղորդակցման և տարաբաշխված հաշվարկային ենթակառուցվածքներից: Ամպային API-ի օգտագործման հետ կապված բոլոր բարդությունները, վիզուալիզացիայի գրադարանները և MPI ծրագրավորումը, տեսանելի չեն օգտագործողին, որը հանդիսանում է աշխատանքի ընդհանուր մոտեցման հիմնական նպատակը (նկ. 4.1):

Հաշվողական շերտը՝ հաշվի առնելով խնդրի հաշվարկային բարդությունը, իրականացնում է հաշվարկը տեղային կամ HPC ռեսուրսներում:

4.1.4 Տվյալների կառավարում

Յուրաքանչյուր հաշվարկի արդյունքի վիզուալիզացիայի համար օգտագործվում են երկու տեսակի տվյալների հավաքածուներ ASCII ֆայլերում: Առաջինը ներկայացնում է գրգռված ֆոտոնների թվի և ֆոտոնների բաշխվածության հավանականության երկչափ ֆունկցիաներ: Յուրաքանչյուր տողում առկա են երկու իրական թիվ (x, y) , միմյանցից առանձնացված դատարկ սիմվոլով, որոնցից առաջինը ժամանակն է, իսկ երկրորդը՝ համապատասխան արժեքը: Երկրորդ տեսակը իրենից ներկայացնում է եռաչափ Վիգների ֆունկցիայի արդյունքներ, որի յուրաքանչյուր տողը բաղկացած է երեք իրական արժեքներից (x, y, z) ՝ բաժանված դատարկ սիմվոլներով:

Օգտագործողների և հաշվարկների պատմության պահպանման համար օգտագործվում է MySQL տվյալների պահոցը, որի միջոցով նաև ապահովվում է iRODS-ի հետ սինխրոնիզացիան: “results” անունով աղյուսակը պարունակում է ինֆորմացիա յուրաքանչյուր հաշվարկի մասին: Յուրաքանչյուրի համար այն պահում է նախնական մուտքային պարամետրերը ECAPT ծրագրի համար և հաշվարկի կարգավիճակը: Յուրաքանչյուր հաշվարկի համար iRODS-ում տրամադրվում է աշխատանքային կատալոգ, որում հետագայում տեղակայվում են արդյունքները:

Ծառայության աշխատանքի ընթացքում cron job-ը տվյալների պահոցից կարդում է նոր ներկայացված հաշվարկները և յուրաքանչյուրի համար՝ կախված հաշվարկային բարդությունից, աշխատեցնում է ECAPT ծրագիրը տեղային միջավայրում կամ ուղարկում է խնդիրը կլաստերային համակարգ կատարման համար: Հաշվարկների ընթացքում կազմավորվում են գրգռված ֆոտոնների թիվը, բաշխվածության հավանականության և Վիգների ֆունկցիաները, որոնք պահվում են տեքստային ֆայլերում: Վիզուալիզացիայի համար ծրագիրը կարդում է այդ ֆայլերի պարունակությունը և ձևափոխում այն վիզուալիզացիայի համար անհրաժեշտ ձևաչափի:

4.1.5 Օգտագործողի միջավայրի նկարագրություն

Ամպային ծառայությունը հիմնված է օգտագործողի վավերացման հնարավորությունների վրա, տրամադրելով հարմարավետ միջավայր հաշվարկների և վերլուծությունների իրականացման համար: Ծառայությունը հնարավորություն է տալիս հաշվարկել գրգռված ֆոտոնների թիվը, բաշխվածության հավանականության և Վիգների ֆունկցիաները տարբեր մուտքային պարամետրերի համար: Օգտատերերը կարող են գրանցվել և մուտք գործել համակարգ լրացնելով համակարգում առկա համապատասխան դաշտերը (նկ. 4.2):

The image displays two side-by-side forms for user authentication. The 'SIGN IN' form on the left has two input fields: 'Email Address' and 'Password', followed by a teal 'SIGN IN' button. Below it is a link that says 'or if you do not have an account then Sign Up'. The 'SIGN UP' form on the right has four input fields: 'Company Name', 'Email Address', 'Password', and 'Retype Password', followed by a teal 'SIGN UP' button.

Նկար 4.2. Համակարգ մուտք գործելու դաշտեր:

Համակարգ մուտք գործելուց հետո, օգտատերը տեղափոխվում է գլխավոր վահանակի էջ: Վերջինս բաղկացած է երկու մասից: Առաջին մասում տեղակայված են անհրաժեշտ պարամետրերի մուտքագրման դաշտերը (նկ. 4.3): Այն լռությամբ լրացված է վերջին հաշվարկների մուտքային պարամետրերով:

The image shows a grid of input fields for simulation parameters. The first row contains 'Detuning (Δ)' with value -100, 'Nonlinearity strength (χ)' with value 25, 'Driving field amplitude (Ω)' with value 10, and 'Time' with value 20. The second row contains 'Time step (dt)' with value 0,00001, 'State vector (Ψ)' with value 5, and 'Chirp parameter (δ)' with value 5. A teal 'SUBMIT' button is located at the bottom right.

Նկար 4.3. Հաշվարկի համար անհրաժեշտ մուտքային պարամետրերի լրացման դաշտեր:

Կեռի ոչ գծային ռեզոնատորում իմպուլսային գրգռվածության և շրջադարձային ակիբի պարագայում Համիլտոնիանի համար մուտքային պարամետրերն են (Հավասարում 2).

- Ω - Ժամանակային կախվածությամբ կապի հաստատուն,
- χ - ոչ գծայնության ուժ,
- Δ - գրգռվածության դաշտի միջին հաճախականության և օսցիլյատորի հաճախականության միջև շեղում $\Delta = \omega_0 - \omega$,
- δ – չիրպի գործակից,
- t - Ժամանակային ինտերվալ 0-ից t միջակայքում,
- dt - հաշվարկի ժամանակային քայլը,
- Ψ – Ֆոկի վիճակների կրճատված բազիսի վիճակի վեկտորի չափ:

Երկրորդ մասում առկա է աղյուսակ, որը պարունակում է հաշվարկների վիճակի մասին ինֆորմացիա (նկ. 4.4):

Id	Status	Δ	χ	Ω	Time	dt	Ψ	δ	Photon mean number	P(N)	Wigner
9443750711	done	-100	10	10	10	0.000001	5	5	View	View	View
7364215202	new	-100	25	10	20	0.00001	5	5	View	View	View
6412115643	done	-100	10	10	20	0.00001	5	5	View	View	View
5843331632	done	-100	10	20	20	0.000001	5	5	View	View	View
1459294119	done	-100	25	27	20	0.00001	7	2	View	View	View
1136859602	done	-100	25	10	20	0.00001	5	5	View	View	View

« Previous **1** Next »

Նկար 4.4. Ընթացիկ և ավարտված հաշվարկների աղյուսակ:

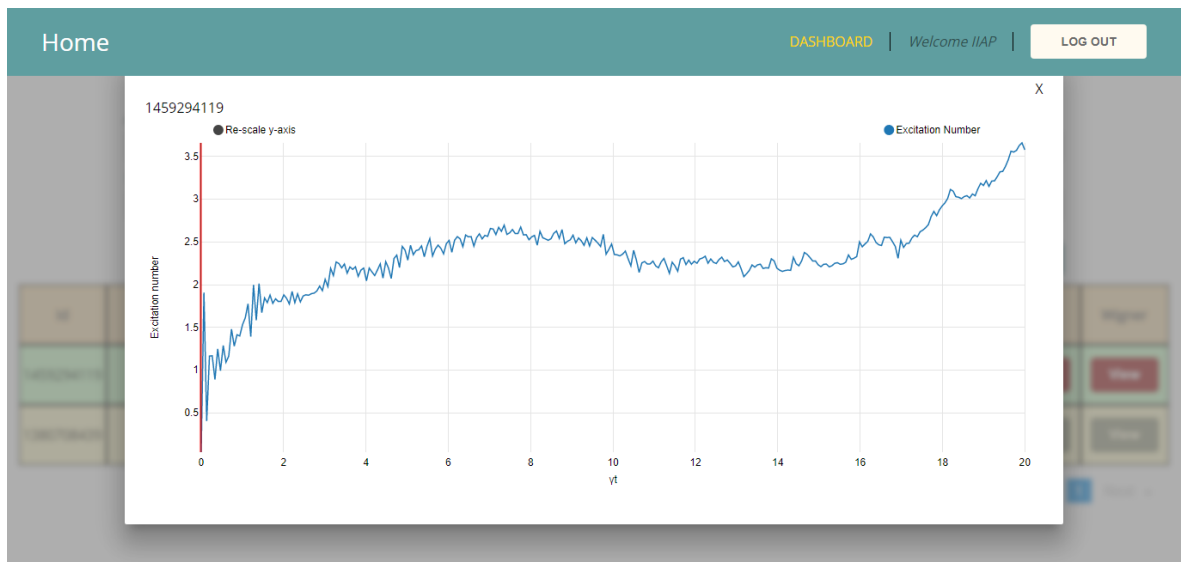
Աղյուսակը պարունակում է հաշվարկի մասին տեղեկատվություն՝ հետագայում հարմար դարձնելով արդյունքների վերստին ուսումնասիրությունը: Այն ցուցադրում է տվյալների բազայում նույնականության համարը (id), հաշվարկի համար տրված մուտքային պարամետրերը և ներկայիս կարգավիճակը (done, new and fail): Վերջին երեք սյուներում իրենց համապատասխան կոճակներով ներկայացված են հնարավոր վիզուալիզացիաները (Photon mean number, P(N) and Wigner):

Նոր հաշվարկների հարցման համար անհրաժեշտ է առաջին մասում լրացնել անհրաժեշտ մուտքային պարամետրերը, այնուհետև սեղմելով “Submit” կոճակը՝ ինտերֆեյսը հրաման է ուղարկում REST API-ի միջոցով և աղյուսակում ավելանում է համապատասխան տող “new” վիճակի արժեքով: Հաշվարկի ավարտից հետո, եթե հաշվարկը հաջողությամբ է ավարտվել, վիճակը փոխվում է “done”-ի և ակտիվանում են վերջին երեք սյունակների կոճակները, որոնց միջոցով հնարավոր է դառնում տեսնել այդ արդյունքի վիզուալ ներկայացումը: Անհաջող ավարտի դեպքում, տողն ունենում է “fail” վիճակ, և անհրաժեշտ է կրկնել նույն հաշվարկը կամ կատարել որոշակի պարամետրերի փոփոխություն: Գծային գրաֆիկների համար օգտագործվում է D3.js [54] գրադարանը, իսկ 3D ֆունկցիայի վիզուալիզացիայի համար՝ Vis.js [28]: Օգտագործողի միջավայրը իրականացված է Nodejs և Angular JS framework-երի միջոցով:

Համակարգում յուրաքանչյուր հաշվարկի համար առկա է երեք տեսակի վիզուալիզացիա.

- Photon mean number – գրգռված ֆոտոնների միջին թիվ (նկ. 4.5)
- $P(N)$ - ֆոտոնների թվի բաշխման ֆունկցիաներ (նկ. 4.6)
- Wigner - Վիգների ֆունկցիայի էվոլյուցիան տրված t պահին (նկ. 4.7)

Գծային գրաֆիկական վիզուալիզացիաների դեպքում (նկ. 4.5, 4.6) վահանակի էջում բացվում է լողացող պատուհան, որտեղ ներկայացված է համապատասխան վիզուալ ներկայացումը: Հորիզոնական առանցքում տրված է ժամանակային պարամետրը, իսկ ուղղահայաց առանցքում՝ համապատասխան արժեքը:



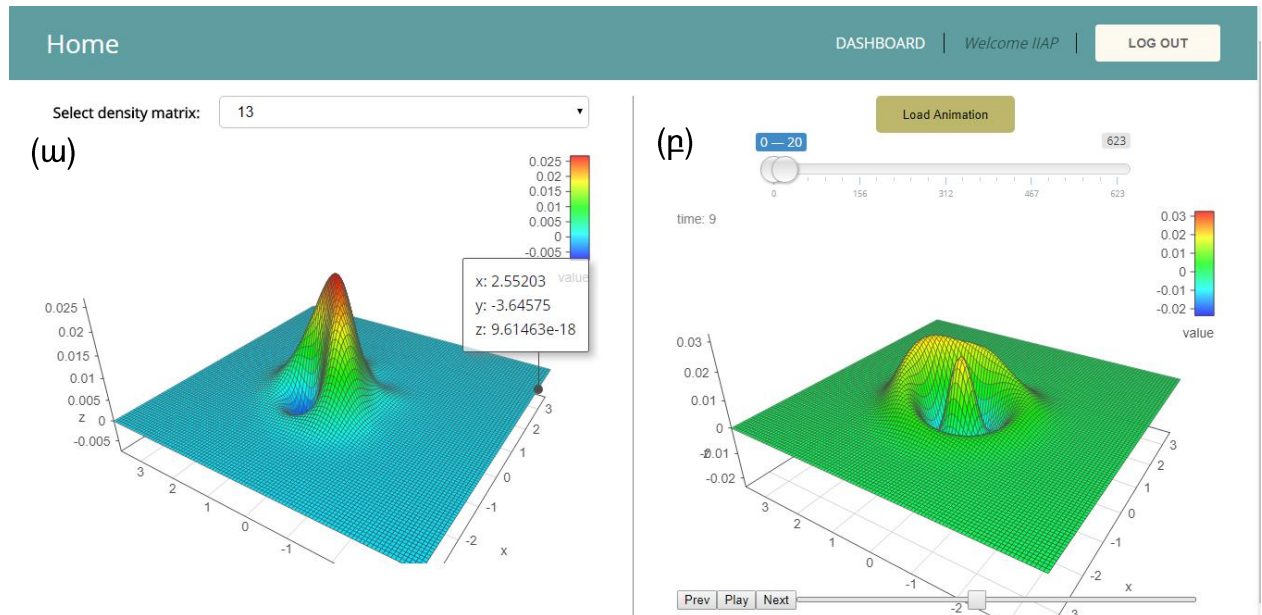
Նկար 4.5. Գրգռված ֆոտոնների թվի վիզուալիզացիայի գրաֆիկ:



Նկար 4.6. Ֆոտոնների թվի բաշխման ֆունկցիաների վիզուալիզացիայի գրաֆիկ:

Վիզների ֆունկցիայի վիզուալիզացիայի դեպքում բացվում է նոր էջ, որը բաղկացած է երկու մասից՝ (ա) և (բ): Ձախ (ա) մասում ցույց է տրվում մեկ 3D Վիզների ֆունկցիայի արտացոլում, որը համապատասխանում է մեկ ժամանակային պահին: Վերին հատվածում գտնվող “Select density matrix” անունով դաշտի միջոցով օգտագործողը կարող է ընտրել ժամանակային պահը, որի համար ցանկանում է տեսնել վիզուալ արտապատկերումը: Աջ (բ) մասում ներկայացվում է ժամանակային հատվածի համար Վիզների ֆունկցիաների փոփոխությունների էվոլյուցիան շատ գեղեցիկ անիմացիայի միջոցով: Օգտագործողը նշելով իրեն հետաքրքրող ժամանակային

հատվածը և սեղմելով “Load Animation” կոճակը, կարող է տեսնել վիզուալիզացիայի արդյունքը: Ներքևում առկա վահանակը հնարավորություն է տալիս տեսնելու եռաչափ ֆունկցիայի վիզուալիզացիան ոչ միայն անիմացիոն ձևով, այլ նաև քայլ առ քայլ: Վիզուալիզացիաները ինտերակտիվ են, ինչը հնարավորություն է տալիս օգտագործողին շրջել, մոտեցնել/հեռացնել և տեսնել ֆունկցիայի յուրաքանչյուր կետի (x, y, z) արժեքները՝ այդ կետի վրա սեղմելով:

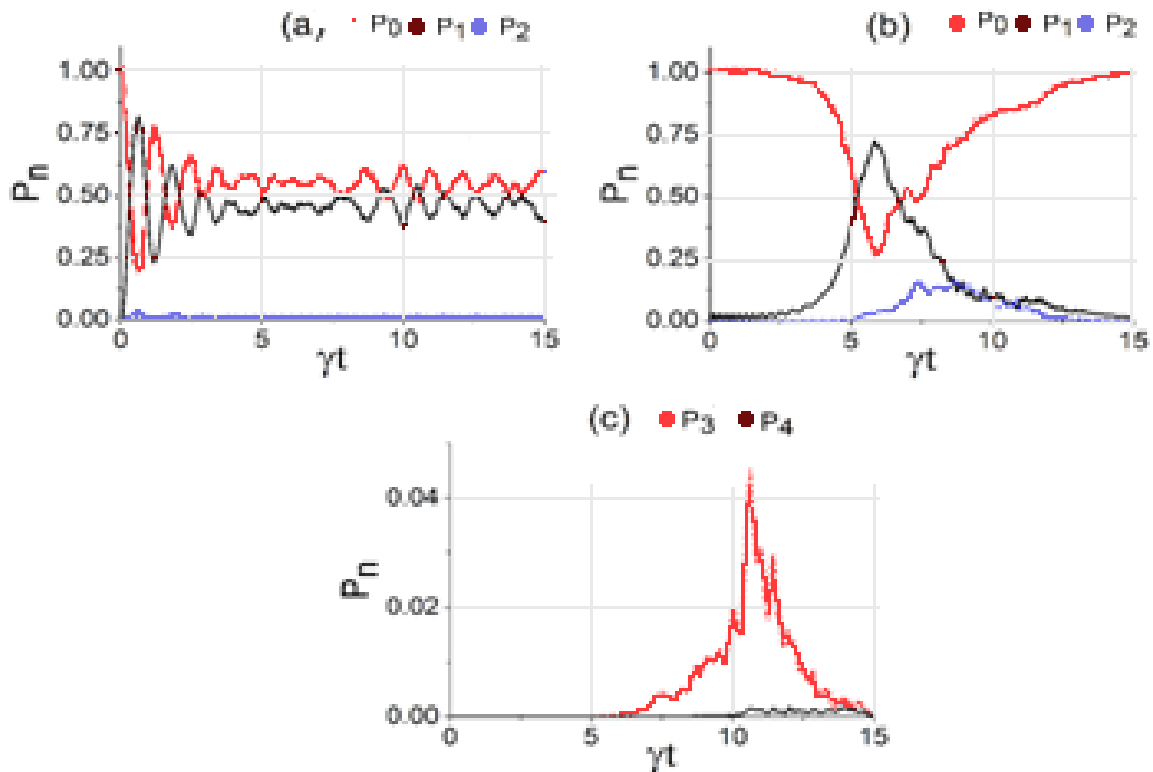


Նկար 4.7. Վիզուալիզացիայի ֆունկցիայի վիզուալիզացիայի էջ:

4.1.6 Կիրառություն

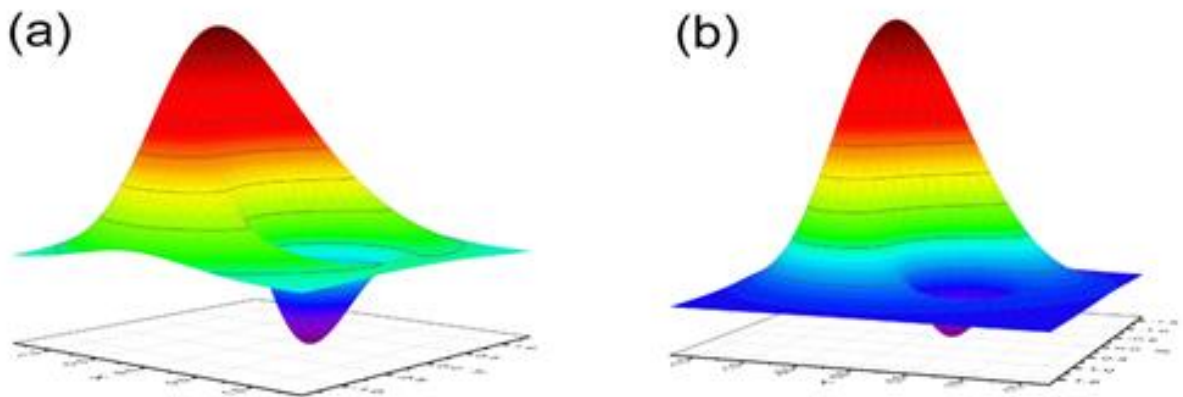
Ներկայումս զգալի հետաքրքրություն է ներկայացնում ֆոտոնային արգելափակման իրականացման հնարավորությունը՝ չիրպացված դիսիպատիվ դաշտի օգտագործմամբ: Նկարագրված համակարգը հետազոտելու համար իրականացվել են տարբեր հաշվարկներ մշակված ամպային ծառայության կիրառմամբ: Նկար 4.8-ում ցուցադրված է ֆոտոնի վիճակի հավանականության ժամանակային դինամիկան: Արտաքին դաշտի դինամիկան՝ առանց ճառագայթի պատկերված է նկար 4.8a-ում, իսկ նկար 4.8b-ում ներկայացված է դինամիկան՝ ճառագայթի կիրառմամբ: Նկար 4.8b-ում կարելի է տեսնել, որ փոքր ոչ գծայնությունների ($\chi / \gamma = 3$) և դաշտային ընդարձակման ($\Omega / \gamma = 2.5$) համար ստեղծվում է զգալի հավանականություն ունեցող վիճակ:

Այժմ դիտարկենք նկար 4.8b-ում ներկայացված վիճակը, երբ տարբեր ժամանակային պահերի P_0 և P_1 վիճակները ունեն միևնույն 0,5 հավանականային արժեքը, և, հետևաբար, այս ժամանակային պահերի համար հնարավոր է արտադրել սուպերպոզիցիոն վիճակներ: Մենք ստուգում ենք սուպերպոզիցիայի վիճակների գոյությունը իսկության օգտագործմամբ, որը ցույց է տալիս երկու քվանտային վիճակների մոտիկությունը: Այսպիսով, մենք հաշվարկում ենք այն $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ սուպերպոզիցիոն վիճակի և հատման կետերին համապատասխան վիճակների համար: Հաշվարկները ցույց են տվել, որ այդ պահերի իսկությունը հավասար է 0,78-ի, ինչը ցույց է տալիս, որ ձեռք բերված պայմանը շատ մոտ է մաքուր սուպերպոզիցիոն վիճակին: Այս երկու վիճակների նմանությունը կարող է դրսևորվել նաև Վիգների ֆունկցիաների օգտագործմամբ: Վերջինս $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ վիճակի համար և տրված ժամանակային պահի համար համապատասխանաբար ներկայացված է նկար 4.9a և 4.9b-ում:



Նկար 4.8. Վիգների ֆունկցիաները արտաքին չիրացված դաշտի պարագայում՝ (a) $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ վիճակի համար, (b) չիրացված դաշտի համար. Պարամետրերը հետևյալն են՝ (a), $\chi/\gamma = 5$, $\Omega/\gamma = 2,5$, $\alpha/\gamma^2 = 0$, (b),(c) $\chi/\gamma = 5$, $\Omega/\gamma = 2,5$, $\alpha/\gamma^2 = 7$:

Ինչպես երևում է նկարից՝ Վիզների ֆունկցիաները շատ մոտ են, սակայն ակնհայտ տարբերություններ կարելի է տեսնել Վիզների ֆունկցիաների բացասական հատվածներում: Ի տարբերություն մաքուր վիճակի՝ ստեղծված վիճակի բացասական մասը ավելի փոքր է: Հայտնի է, որ Վիզների ֆունկցիայի բացասական հատվածը նշանակում է, թե արդյոք տրված վիճակը մաքուր քվանտային վիճակ է, թե՛ ոչ: Ֆունկցիայի բացասական հատվածի փոքր բացարձակ արժեքը կարող է բացատրվել անջատման և դիսֆունկցիայի առկայությամբ, քանի որ այդ ազդեցությունները վերացնում են քվանտային վիճակը:



Նկար 4.9. Վիզների ֆունկցիաները արտաքին չիրապցված դաշտի պարագայում՝ (a) $\frac{1}{\sqrt{2}} (|0\rangle - |1\rangle)$ վիճակի համար, (b) չիրապցված դաշտի համար. Պարամետրերը հետևյալն են՝ (a), $\chi/\gamma = 5$, $\Omega/\gamma = 2,5$, $\alpha/\gamma^2 = 0$, (b) $\chi/\gamma = 5$, $\Omega/\gamma = 2,5$, $\alpha/\gamma^2 = 7$:

Նկար 4.9a-ում ցուցադրվածը լավ ուսումնասիրված դեպք է, երբ արտաքին դաշտը, որը ազդում է քվանտային ոչ գծային օսցիլյատորի վրա, մոնոխրոմատիկ է, և հաշվի է առնված դիսպենսատի առկայությունը: Ինչպես տեսնում ենք, մեկնարկային կետում կան մի քանի գրգռվածություններ, որոնք ժամանակի հետ նեղանում են և, ի վերջո, համակարգը գալիս է կայուն վիճակի: Ինչպես երևում է նկար 4.9b-ում, իրավիճակը ամբողջովին տարբեր է չիրաի կիրառման դեպքում: Սկզբնական կետում գոյություն ունեն որոշակի փոքր տատանումներ և էներգիան տեղայնացվում է որոշակի ժամանակային պահի: Այդ պահին ստացվում է միաֆոտոնային վիճակ:

4.1.7 Եզրակացություն

Արդյունքների վիզուալիզացիան հնարավորություն է տալիս հետազոտողներին վերլուծելու և ուսումնասիրելու համակարգերը առավել պրակտիկ և լիովին տարբեր ժամանակային պարբերությունների և ժամանակային կետերի համար, օրինակ, վերլուծել Համիլտոնիան համակարգի (Հավասարում 2) ելքային արդյունքները տարբեր մուտքային պարամետրերի և միևնույն Համիլտոնիանի տարբեր արժեքները տարբեր ժամանակների դեպքում: Բացի տարբեր հաշվարկների համար տրամադրված պահոցից, մշակված միջավայրը դարձնում է տեսական հետազոտությունը ավելի փորձարարական և վիզուալ, որը խնայում է հետազոտության ժամանակը և կիրառվող ռեսուրսները: Հիմնական նպատակն է եղել մշակել մեկ մուտքով բազմաշերտ ծառայությունների միջավայր, որը կտրամադրի հեշտ հասանելիություն քվանտային հաշվարկների և ալգորիթմների տարբեր տեսակների համար՝ իրականացված հիբրիդային տարաբաշխված հաշվողական ենթակառուցվածքներում՝ օգտագործելով կլաստերի, գրիդի և ամպային տեխնոլոգիաների առավելությունները:

4.2 Ամպային ծառայություն եղանակային տվյալների վերլուծության և վիզուալիզացիայի համար

Հայաստանը զբաղեցնում է Հայկական լեռնաշղթայի հյուսիս-արևելյան մասը և Փոքր Կովկասի կենտրոնական հատվածը (լայնություն - $38^{\circ}51'$ -ից մինչև $41^{\circ}18'$ հյուսիս, երկայնություն - $43^{\circ}29'$ -ից մինչև $46^{\circ}37'$ արևելք) մոտ 30000 քմ տարածքով: Հայաստանի աշխարհագրական դիրքը և բարդ լեռնային ռելիեֆը հանգեցրել են երկրի ողջ բնական պայմանների բազմազանությանը: Հայաստանը գտնվում է ենթատարրային գոտու հյուսիսային եզրին, լայնություններում բնութագրվում են չոր և կոնտինենտալ կլիմայով: Լեռնային ռելիեֆի շնորհիվ առկա են տարբեր կլիմայական գոտիներ և եղանակը կարող է ունենալ բարձր տարածական գրադիենտ: Հայկական կլիմային բնորոշ են տարեկան և ցերեկային ջերմաստիճանի բարձր տատանումները: Հայաստանի հիդրոոդերևութաբանության և մոնիթորինգի ծառայության համար եղանակային կանխատեսումների և կլիմայի կանխատեսման համար լրացուցիչ մարտահրավերներ են առաջացնում չոր ցիտրատիկ գոտիների չոր մերձարևադարձայինից մինչև խիստ բարձր լեռնային և հավերժացող ձնագնդերի և տաք խոնավ ենթոդպիկական անտառների և խոնավ կիսաանապատային դաշտերի առկայությունը: Եղանակային տվյալների ուսումնասիրության, կանխատեսումների ճշգրտության վերլուծության համար կառուցվել է վեբ միջավայրում աշխատող ամպային ծառայություն [55] [56], որը տրամադրում է վերլուծական և ինտերակտիվ վիզուալիզացիայի գործիքներ:

4.2.1 Նախաբան

Օդերևութաբանական տվյալները ստացվում են 47 օդերևութաբանական (ներառյալ 3 մասնագիտացված) կայաններից, որոնցից ընդամենը չորս կայանն է տրամադրում պատմական տվյալներ և ամսական թարմացումներ Գլոբալ Կլիմայի Դիտարկումների Համակարգի Մակերևույթի Ցանցին և երեք օդերևութաբանական կայաններ տրամադրում են սինոպտիկ տվյալներ, որոնք օգտագործվում են ցանցային վերլուծության տվյալների հավաքածուում, որը հանդիսանում է որպես մուտք գլոբալ մթնոլորտային մոդելների համար՝ համաշխարհային մասշտաբով եղանակի կանխատեսումների մշակման համար: Դիտարկման տվյալները ծառայում են որպես

մուտքագրման տվյալներ՝ ջերմաստիճան, տեղումներ և այլ մթնոլորտային տարրեր, եղանակի կանխատեսման թվային մոդելներում արդյունքներ ստանալու համար [57] [58]:

Եռաչափ վարիացիոն (3DVAR) տվյալների յուրացման մեթոդը օգտագործվում է տվյալ ժամանակահատվածում մթնոլորտային վիճակի վերաբերյալ առկա բոլոր տեղեկությունները համադրելու համար, տվյալ վերլուծության ժամանակահատվածում մթնոլորտային պայմանների գնահատականի ստեղծման համար [59]: Վերլուծության համար օգտագործվող տեղեկատվության աղբյուրները ներառում են օդերևութաբանական կայանների դիտարկումներ, նախնական կանխատեսումներ (ֆոնային կամ առաջին գուշակման վիճակ) և արբանյակային պատկերներ:

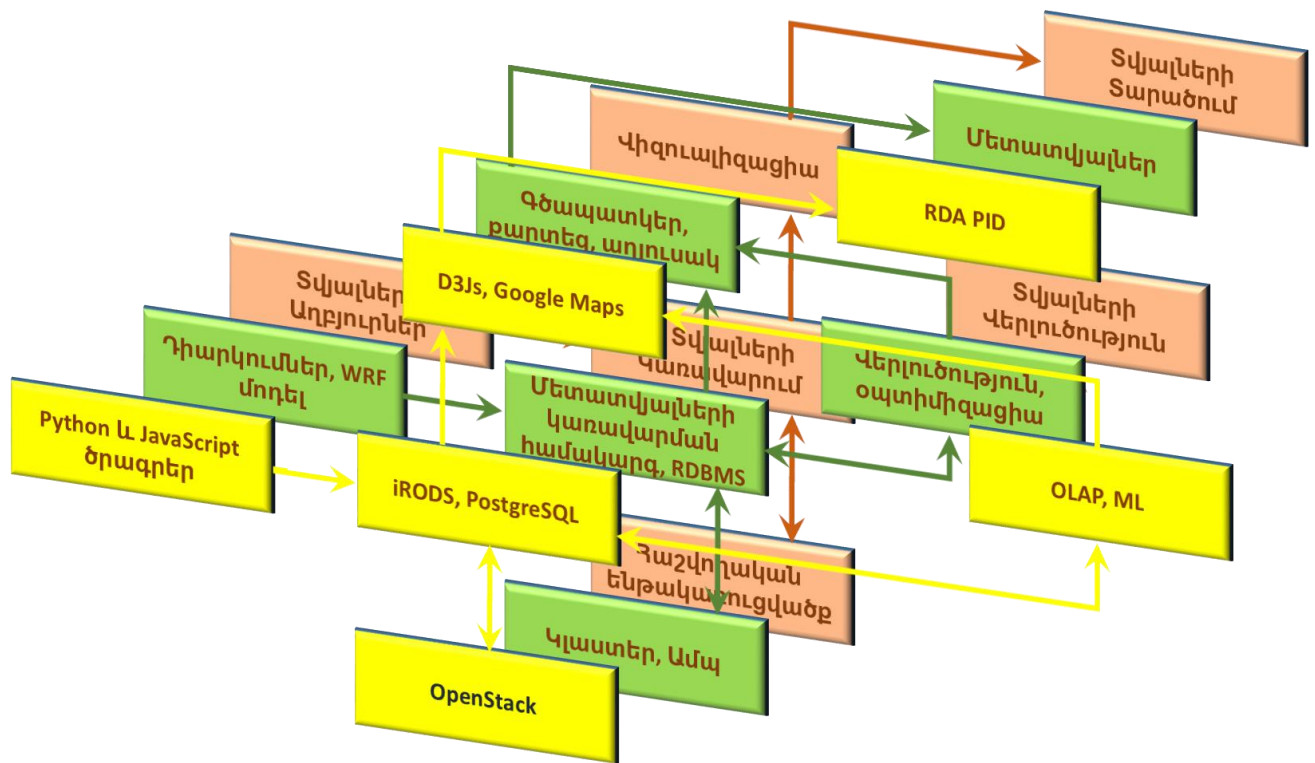
Ներկայումս արբանյակային պատկերները օգտագործվում են ապագա փորձերի համար, քանի որ նաև շատ կարևոր է հաշվի առնել այն տարածքները, որտեղ չկան օդերևութաբանական կայաններ: Հայաստանի էլեկտրոնային ենթակառուցվածքի բարձրակարգ հաշվարկային ռեսուրսները օգտագործվում են ավելի լավ վերլուծելու եղանակի մթնոլորտային իրադարձությունները՝ կարճ շրջանում բավականաչափ ճշգրիտ կանխատեսումներ տրամադրելու համար [60] [61].

Միջավայրը մշակվել է Հայաստանում եղանակային տվյալների համար, երեք հարթակների՝ օդերևութաբանական կայանների դիտարկումների, եղանակի թվային կանխատեսման և արբանյակային նկարների մշակման ինտեգրման միջոցով: Միջավայրը AHMS (Armenian State Hydrometeorological and Monitoring Service)-ին տալիս է վերլուծական կարողություններ՝ վերլուծելու համար տվյալ ժամանակահատվածում յուրաքանչյուր կայանի և տարածաշրջանի տարածքի, մոդելի և արբանյակային պատկերների շեղումները:

4.2.1 Ամպային ծառայության ենթակառուցվածքը

Ներկայացված վեբ հիմքով ինտերակտիվ վիզուալիզացիայի և վերլուծության միջավայրը կառուցվել է օգտագործելով մշակված DAVS համակարգը: Հաշվողական ենթակառուցվածքը ապահովում է HPC-ն և տվյալների ռեսուրսները, որոնք հատկապես կարևոր են թվային մոդելների և արբանյակային նկարների մշակման համար [62]:

Օգտագործվել են Հայաստանի էլեկտրոնային ենթակառուցվածքի ռեսուրսները, որոնք հանդիսանում են բարդ ազգային SS ենթակառուցվածք, որը բաղկացած է հաղորդակցման և տարաբաշխված հաշվարկային ենթակառուցվածքներից (նկ. 4.10):



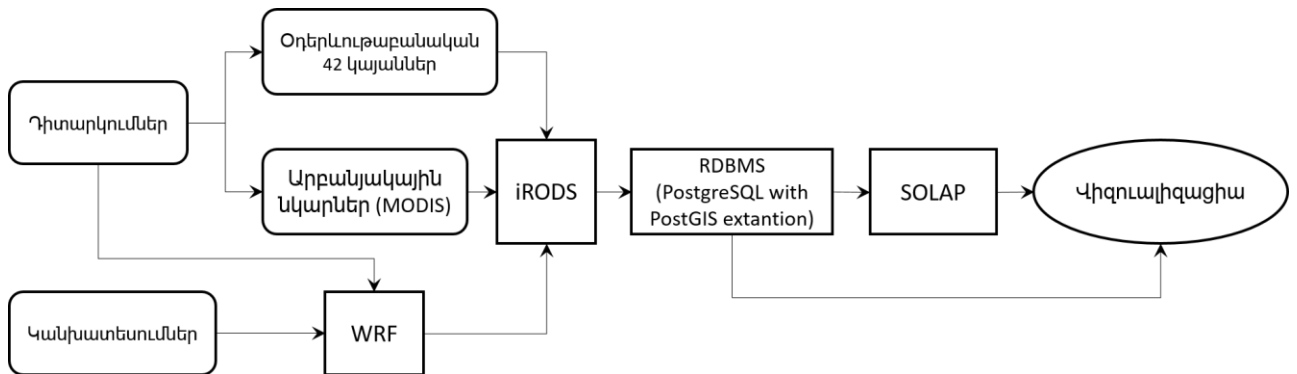
Նկար 4.10. Ամպային ծառայության ենթակառուցվածքը եղանակային տվյալների վերլուծության և վիզուալիզացիայի համար¹⁵:

Որպես տվյալների աղբյուրներ հանդես են գալիս երեք տեսակի տվյալների հարթակներ՝ հետագա վերլուծությունների իրականացման համար.

- Մոդելի արդյունք – եղանակի կանխատեսման մոդելների արդյունք,
- Արբանյակային նկարներ – Հայաստանի տարածքը ներառող բազմատեսակ արբանյակային նկարներ,
- Օդերևութաբանական կայանների դիտարկումներ – հիմք են հանդիսանում մոդելի արդյունքների և արբանյակային նկարների շեղումների վերլուծությունների համար:

¹⁵ Ամպային ծառայություն. <http://meteo.grid.am/>

Տվյալների կառավարման բաղադրիչը տրամադրում է խելացի գործիքներ չմշակված տվյալները տվյալների վերլուծության շերտ տեղափոխելու համար: iRODS-ը հանդիսանում է միջնորդ համակարգ, որն ապահովում է ֆիզիկական տվյալների պահպանման համակարգերի և օգտագործողի ինտերֆեյսի միջև կապը (նկ. 4.11):



Նկար 4.11. Ամպային ծառայությանում տվյալների հոսք:

Երբ տվյալը հասնում է վերլուծության շերտ, այն մշակվում է և միայն որոշ ցուցանիշներ են մնում նախնական չմշակված տվյալների ահռելի քանակությունից: Վերջապես ամենավերևի և վերջին շերտը արդեն մշակված տվյալների վիզուալիզացիայի շերտն է, որտեղ արդյունքները ձևափոխվում են օգտագործողի համար ավելի հասկանալի գրաֆիկների և աղյուսակների: Ավելին, Google Maps-ի առավելությունները օգտագործվում են աշխարհագրական տվյալները քարտեզի վրա պատկերման համար:

4.2.2 Տվյալների տարրեր, գործիքներ և մոդելներ

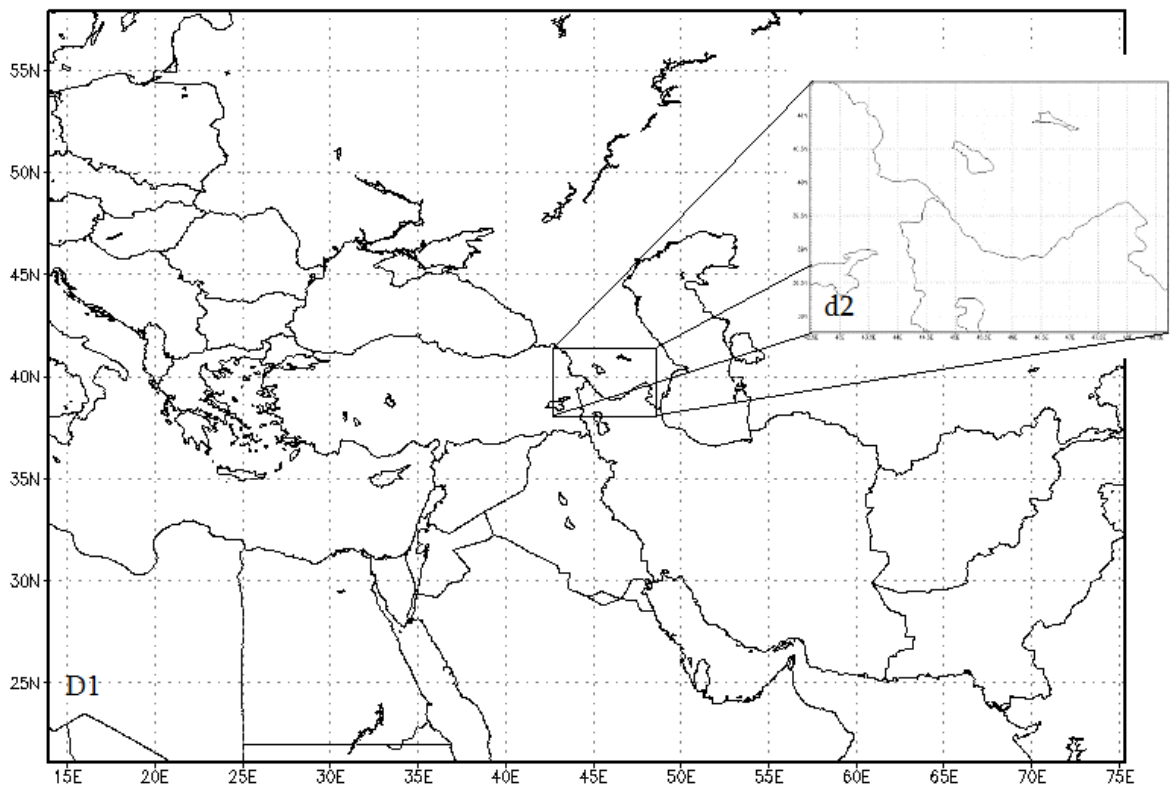
Ռիտարկման տվյալների բազաները ապահովում են SYNOP (մակերևութային սինոպտիկ դիտարկումներ) հաղորդագրություններից ստացված տարբեր վայրերից դիտարկված եղանակային տվյալները՝ տրամադրված պաշտոնական օդերևութաբանական կայանների կողմից: SYNOP-ի հաշվետվությունները սովորաբար ուղարկվում են յուրաքանչյուր երեք ժամը մեկ, որը բաղկացած է եղանակային կայանների ընդհանուր եղանակի մասին տեղեկությունները նկարագրող թվերի խմբերից, ինչպիսիք են ջերմաստիճանը, բարոմետրիկ ճնշումը և տեսանելիությունը եղանակային կայանում:

Երկրի մակերևույթի ջերմաստիճանը, ներառյալ հողի մակերևույթի ջերմաստիճանը (LST – Land Surface Temperature), կարևոր պարամետրեր են, որոնք արտացոլում են երկրի մակերեսային միջավայրը և լայնորեն օգտագործվում են կլիմայի փոփոխության ու եղանակի կանխատեսման համար: Օգտագործվում է Landsat պատկերը, որն ապահովում է բարձր հստակության տեսանելի և ինֆրակարմիր պատկերներ, ջերմային և panchromatic պատկերներով, որը հասանելի է նաև ETM + (Enhanced Thematic Mapper Plus) սենսորից [63]: LST-ի որոնման համար օգտագործվում է մեկ ալիքային մեթոդ: Այս մեթոդը կիրառում է արբանյակային պատկերների միայն մեկ ջերմային խողովակ: Այս մեթոդը համապատասխանում է այն սենսորներին, որոնք ունեն միայն մեկ ջերմային խողովակ, ինչպիսիք են՝ Landsat TM / ETM+: Տվյալների ձեռքբերման ամսաթվերը ունեցել են հստակ պարզ մթնոլորտային պայմաններ և պատկերները ձեռք են բերվել United States Geological Survey Earth Explorer Data Center-ից, որն ուղղել է պատկերների ռադիոմետրիկ և երկրաչափական աղավաղումները մինչև 1G-ի որակի մակարդակ: Պատկերները հասանելի են GeoTiff ձևաչափով, թույլ տալով վերափոխել, վերամշակել և իրականացնել գործառնություններ հեշտ ճանապարհով: Հայաստանի տիրույթը ներառող արբանյակային նկարները հասանելի են օրվա 2 հատվածների համար՝ ցերեկվա և գիշերվա:

Որպես աշխարհագրական տեղեկատվական համակարգ (GIS – Geographic Information System) ծրագրային համալիր, աշխարհագրական ռեսուրսների վերլուծության աջակցության համակարգը (GRASS GIS) օգտագործվում է արբանյակային պատկերների վերամշակման, գրաֆիկայի և քարտեզների ստեղծման, տարածքային և ժամանակային մոդելավորման ու արտացոլման համար [64]: GRASS GIS-ը ներկայումս օգտագործվում է աշխարհով մեկ ակադեմիական և կոմերցիոն համատեքստերում, ինչպես նաև բազմաթիվ պետական մարմիններում ու բնապահպանական խորհրդատվական ընկերություններում: Այն կարողանում է մշակել ռաստերային, տոպոլոգիական վեկտորի, նկարների մշակման և գրաֆիկական տվյալները: GRASS GIS-ը պարունակում է ավելի քան 350 մոդուլներ մոնիթորի և թղթի վրա քարտեզների ու պատկերների մատուցման համար. մանիպուլացնել ռաստեր և վեկտորի տվյալներ, ներառյալ վեկտորային ցանցերը, մշակել բազմատեսակ

պատկերային տվյալներ և մշակել, կառավարել ու պահել տարածական տվյալները: GRASS GIS-ի հետևյալ մոդուլները օգտագործվել են back-end-ում արբանյակային պատկերների մշակման համար: LST-ը հաշվարկելու համար օգտագործվում են մի քանի Grass մոդուլներ: Նախ և առաջ, յուրաքանչյուր սեսիայի համար ստեղծվում է աշխատանքային տարածք: Երբ բոլոր պատկերները լինում են միևնույն GeoTiff տեսակի, աշխատանքային տարածքները դառնում են համանման և օգտագործվում են միևնույն EPSG geodetic պարամետրերի տվյալների տարրերը: Աշխատանքային տարածքի ստեղծումից հետո օգտագործվում է "r.in.gdal" հրամանը՝ ռաստեր նկարների բեռնման համար: Այնուհետև օգտագործելով r.mapcalc հրամանը հաշվարկվում է LST ցուցիչը քարտեզի յուրաքանչյուր կետի համար և օգտագործելով "r.what" մոդուլը ստացվում է LST-ն տրված կոորդինատների համար: Այս պարզ քայլերի իրականացման արդյունքում ստացվում են LST արժեքները որպես արդյունք:

Հայաստանի տարածքի համար հարմարեցված WRF (Weather Research and Forecasting) մոդելը օգտագործվում է եղանակի օպերատիվ կանխատեսման համար: Ուսումնասիրելու համար, թե որքանով է լավ աշխատում մոդելը Հայաստանի տարածքի համար, օգտագործվել է 2 տիրույթ (նկ. 4.12): Առաջինը՝ D1-ը, ներառում է Հայաստանի ամբողջ տարածքը և կենտրոնը գտնվում է 44,7 երկայնության և 40,0 լայնության կետում, հորիզոնական տարածությունը 18 կմ է և ունի 202x202 ցանցային կետ: Երկրորդ տիրույթը՝ D2-ը, ներդրված է առաջինում, ունի 6 մ հորիզոնական չափ և 97x70 ցանցային կետ: Եղանակի կանխատեսումը կատարվում է ամեն օր՝ օգտագործելով նշված մեկ մակարդակի ներդրված մարտավարությունը: Մոդելը օգտագործում է ուղղահայաց 31 eta_level-ներ, իսկ աշխարհագրական տվյալների չափը 30 վայրկյան է: Մոդելը սկզբնավորվել է Global Forecast System (GFS)-ի նախնական և սահմանային պայմաններով՝ 0000 UTC (տեղական ժամանակով 04:00) 31 օրվա ընթացքում, մասնավորապես, 1-31 հունվարի 2016թ.:



Նկար 4.12. Մոդելում օգտագործվող ծնող և ներդրված տիրույթները:

Կիրառված WRF-ARW մոդելը ներառում է WRF «Single Moment 6» դասի սխեմա ամպային միկրոֆիզիկայի համար, որը ներառում է սառույցը, ձյունը և graupel (ձյան հատիկ) պրոցեսները: Խորը և մակերեսային կոնվեկցիայի ենթատարածքային պարամետրերը հիմնված են Kain-Fritsch սխեմայի վրա, որը կիրառվում է զանգվածային հոսքի մոտեցման համար ներքևում և CAPE հեռացման ժամանակահատվածում: Այս ուսումնասիրության մեջ օգտագործվել է Mellor-Yamada-Janjic (MYJ) PBL համակարգը Eta-ի նմանության մակերեսային շերտով: Արագ ռադիացիոն փոխանցման մոդելի (RRTM) սխեման օգտագործվել է երկարատև ճառագայթման համար, իսկ կարճ ալիքային ճառագայթման գործընթացները ներկայացված է Դուդայի սխեմայով:

4.2.3 Եղանակի կանխատեսման WRF թվային մոդելը

Եղանակի հետազոտման և կանխատեսման (WRF) մոդելը [65] [66] եղանակի թվային կանխատեսման (NWP – Numerical Weather Prediction) համակարգ է, որը ներառում է գործիքներ և ծրագրային ապահովումներ, որոնք նախատեսված են ինչպես

մթնոլորտային հետազոտությունների, այնպես էլ օպերատիվ կանխատեսման ծրագրերի համար: WRF-ն ունի 2 դինամիկ (հաշվողական) միջուկներ, տվյալների ասիմիլյացիայի համակարգ և ծրագրային ապահովման ճարտարապետություն, որը թույլ է տալիս կատարել զուգահեռ հաշվարկներ և ունենալ համակարգի հասանելիություն: Մոդելը կիրառվում է լայնածավալ օդերևութաբանական ծրագրերում՝ տասնյակ մետրից մինչև հազարավոր կիլոմետրեր: Այն ներառում է հետևյալ հատկանիշները.

- Օդերևութաբանական ուսումնասիրություններ,
- NWP իրական ժամանակում,
- Իդեալականացված,
- Հաշվարկներ,
- Տվյալների ասիմիլյացիա,
- Երկրի համակարգի մոդելի կապակցում,
- Մոդելի ուսուցում և կրթական աջակցություն:

Շրջակա միջավայրի կանխատեսման ազգային կենտրոնում և այլ ազգային օդերևութաբանական կենտրոններում WRF-ը ներկայումս կիրառվում է եղանակի կանխատեսման և մոդելավորման համար: Շատ լաբորատորիաներում, համալսարաններում և մասնավոր ընկերություններում WRF մոդելը օգտագործվում է իրական ժամանակում կանխատեսման կարգավորումներում:

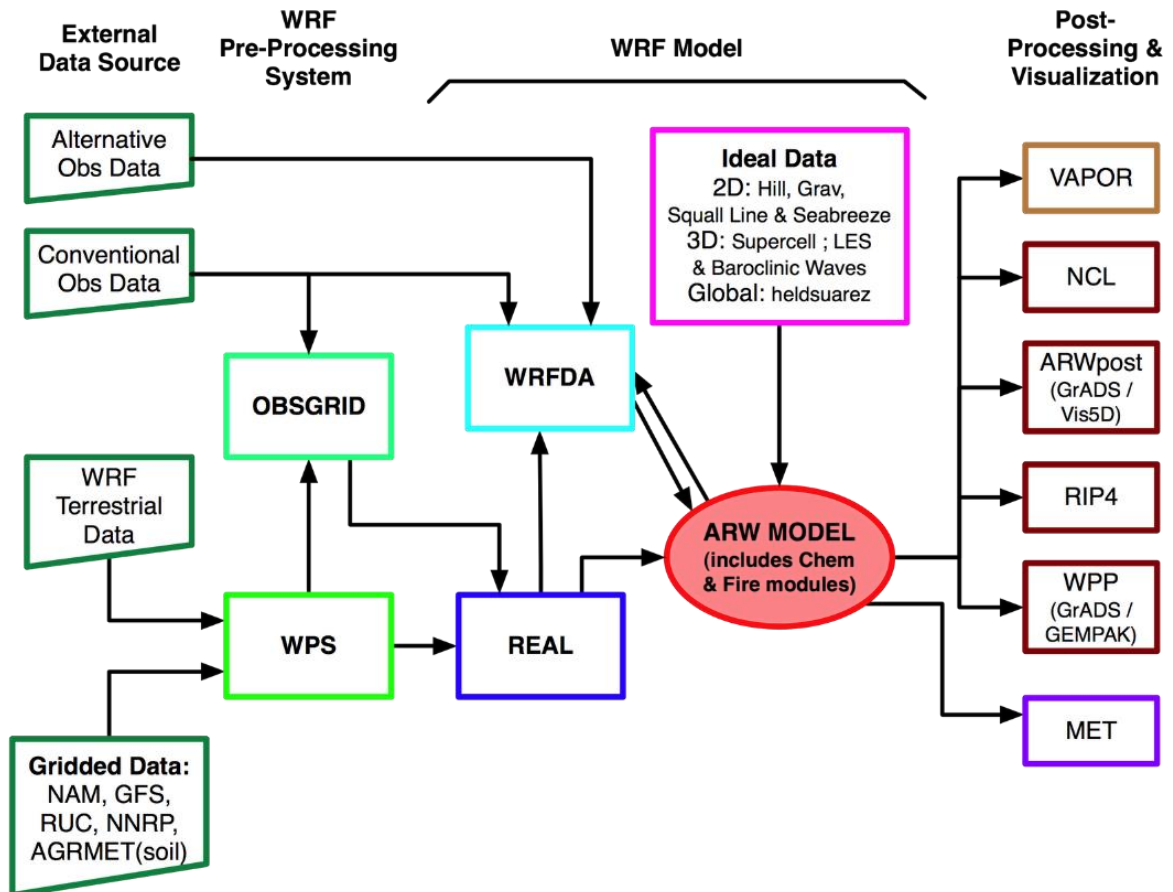
WRF-ը բաղկացած է հետևյալ հիմնական մասերից.

- WRF-ի նախամշակման համակարգ (WPS),
- WRFDA,
- WRF մոդել (ARW մոդել, NMM մոդել),
- Վիզուալիզացիայի գործիքներ:

Նկար 4.13-ում պատկերված և նշված բոլոր բաղադրիչները հանդիսանում են առանձին ծրագրեր և, որպեսզի աշխատեն որպես ամբողջական համակարգ հարկավոր է կատարել անհրաժեշտ կարգավորումներ: Յուրաքանչյուր բաղադրիչ ունի իր կարգավորման ֆայլը:

WPS-ի ֆունկցիան է սահմանել կանխատեսման տիրույթները (domain), ինտերպոլացնել երկրային տվյալները (տեղանք, հողի տեսակ և այլն) և այլ մոդելից ստացված օդերևութաբանական տվյալները կանխատեսվող տիրույթին: WPS-ը բաղկացած է 3 ծրագրերից՝ geogrid, ungrib, metgrid:

WRF Modeling System Flow Chart



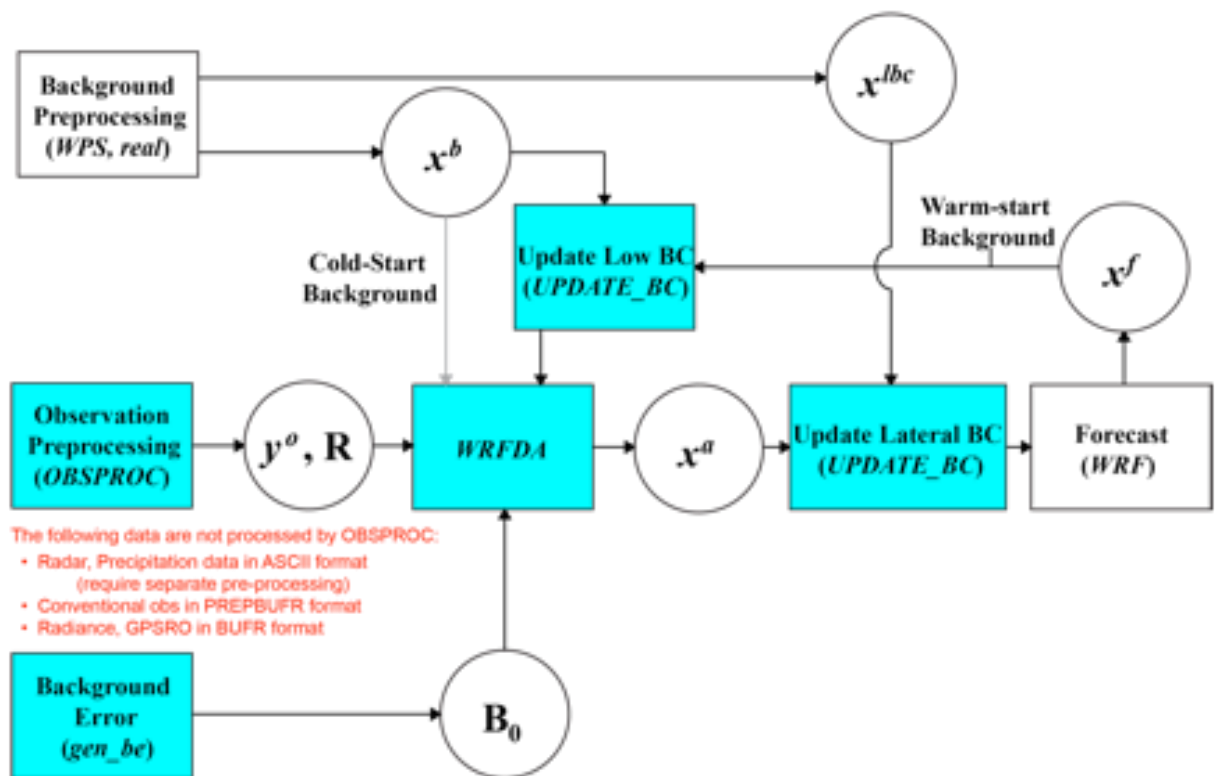
Նկար 4.13. WRF մոդելի բաղադրիչներ¹⁶:

Տվյալների ասիմիլյացիան հնարավորություն է տալիս դիտարկված տվյալները համակցել NWP արդյունքներին (նախնական կանխատեսում) և նրանց համապատասխան սխալների վիճակագրությանը՝ կարճաժամկետ կանխատեսումներում փոքր ուղղումներ կատարելու համար, որն ունի դրական ազդեցություն մոդելի վերլուծության վրա: Վարիացիոն (Var) տվյալների ասիմիլյացիան

¹⁶ <http://www2.mmm.ucar.edu/wrf/users/model.html>

հասնում է արդյունքի՝ օգտագործելով տրված արժեքի իտերատիվ նվազեցման եղանակը:

WRFDA համակարգը ստեղծված է որպես ճկուն, ժամանակակից մթնոլորտային տվյալների ասիմիլյացիայի համակարգ, որը հանդիսանում է տեղափոխելի և արդյունավետ հասանելի զուգահեռ հաշվարկների հարթակներում (նկ. 4.14) [67]: Այն հարմար է լայն շրջանակներում կիրառման համար, տարածաշրջանային մոդելների համար կիլոմետրերից սկսած մինչև մեգոմասշտաբ մոդելավորման հազարավոր կիլոմետրերի համար: Համակարգի OBSPROC ծրագիրը օգտագործում է դիտարկված տվյալները LITTLE_R ֆորմատով ֆայլերի միջոցով: Տարբեր տեսակի դիտարկման արդյունքներ համախմբվում են մեկ ընդհանուր LITTLE_R ֆորմատով ֆայլում, մշակվում, ձևափոխվում և գեներացվում են ASCII ֆորմատի ֆայլեր, որը համապատասխանում է WRFDA համակարգի մուտքի համար՝ 3DVAR, FGAT (First Guess at Appropriate Time), 4DVAR:



Նկար 4.14. WRFDA համակարգ¹⁷:

¹⁷ http://www2.mmm.ucar.edu/wrf/users/wrfda/Docs/user_guide_V3.9.1/WRFDA_Users_Guide.pdf

Ընթացիկ սցենարը հիմնականում օգտագործվում է 3DVAR համակարգի համար, որի հիմնական նպատակն է վերլուծության ժամանակ փնտրել իրական մթնոլորտային վիճակի օպտիմալ գնահատական:

LITTLE_R

LITTLE_R-ը ASCII հիմքով դիտարկման ֆայլային ֆորմատ է: Քանի որ գոյություն ունեն դիտարկման տվյալների ֆայլերի տարբեր հնարավոր ֆորմատներ, ինչպիսիք են՝ ASCII, BUFR, PREBUFR, MADIS և HDF, LITTLE-ը ստեղծված է որպես միջանկյալ ֆորմատ, այնպես որ WRFDA-ն կարող է համակողմանի կերպով յուրացնել հնարավորինս շատ դիտարկման տեսակներ [5]: Ջեկույցի վրա հիմնված ֆայլի ձևաչափը թույլ է տալիս բոլոր տեսակի դիտարկման ինֆորմացիաները հեշտությամբ համատեղել մեկ ընդհանուր հեշտ խմբագրվող տեքստային ֆայլում:

Այն բաղկացած է զեկույցներից, որոնք կազմված են 3 բաժիններից՝ header, data և երեք վերջավորության թվեր: header-ը մեկ տող է տեղակայված յուրաքանչյուր դիտարկման առաջին տողում և պարունակում է ինֆորմացիա դիտարկման տեղի, տեսակի և այլ պարամետրերի մասին: Data բաժինը հաջորդում է header բաժնին և մարմնում կարող է պարունակել մեկից ավել տող ինֆորմացիա, որը կախված դիտարկման տեսակից, համապատասխանում է մեկ դիտարկման բազմակի ուղղահայաց մակարդակների տվյալներին: Հայաստանի դիտարկման տվյալների դեպքում, որոնք FM-12 տեսակի են, Data բաժնի մարմինը բաղկացած է մեկ տողից: Բոլոր դիտարկման տեսակների դեպքում Data բաժնի վերջում առկա է վերջնական մեկ տող, որում OBSPROC ծրագրի համար անհրաժեշտ է բարձրության և ճնշման դաշտերը լրացնել ֆիքսված “-777777.00000” արժեքով: Վերջավորության թվերը օգտագործվում են OBSPROC ծրագրի կողմից, որոշելու համար, թե որ դիտարկումը պահպանել, եթե առկա են երկու կամ ավելի միանման դիտարկման տեսակներ միևնույն տեղի և ժամանակի համար:

LITTLE_R-ը օգտագործվում է ոչ միայն OBSPROC/WRFDA համակարգերում, այդ իսկ պատճառով շատ դաշտեր չեն օգտագործվում, և քանի որ անհրաժեշտ է պահպանել ընդհանուր կառուցվածքը, նրանք լրացվում են “-888888.00000” բաց թողնված դրոշի արժեքով, նշելու համար, որ այդ դաշտի արժեքը բացակայում է:

SYNOP կող

SYNOP (մակերեսային սինոպտիկ դիտարկումներ)-ը թվային կող է WMO (World Meteorological Organization)-ի կողմից անվանված FM-12, որը օգտագործվում է ավտոմատացված կայանների կողմից եղանակային դիտարկումները զեկուցելու համար: Զեկույցը բաղկացած է թվերի խմբից, որը նկարագրում է ընդհանուր եղանակային ինֆորմացիան, ինչպիսիք են ջերմաստիճանը, բարոմետրային ճնշումը և եղանակային կայանի տեսանելիությունը: SYNOP զեկույցները սովորաբար հաղորդվում են 3 ժամը մեկ:

Ավտոմատացման ծրագիր

Ծրագիրը մշակված է Python ծրագրավորման լեզվով [68]: Ծրագրի կառուցվածքը նախագծված է այնպես, որպեսզի համապատասխանի WRF մոդելի ծրագրերի աշխատեցման ձևին: Աշխատանքի համար անհրաժեշտ արգումենտները տրվում են կարգավորման “namelist” կոչվող ֆայլում, որն օգտագործվում է գլխավոր աշխատեցվող ծրագրի կողմից: Այն պարունակում է 2 բաժին՝ input և output: Առաջին բաժինը պարունակում է “stations_file” (կայանների CSV ֆայլի ճանապարհը) և “synop_file” (SYNOP պարունակող ֆայլի ճանապարհը) արգումենտները: Կայանների CSV ֆայլը պետք է ունենա հետևյալ կառուցվածքը՝ [station WMO number], [name], [elevation], [latitude], [longitude] (նկ. 4.15).

```
37609;Ashotsk;2009;41,029444;43,876389
37618;Tashir;1502;41,123333;44,28
37626;Bagratashen;453;41,25;44,819722
37627;Odzun;1200;41,059167;44,613611
37682;Amasia;1876;40,96;43,78
```

Նկար 4.15. Կայանների CSV ֆայլի օրինակ:

SYNOP կողային ֆայլերը պետք է պարունակեն բաժիններ, որոնցից յուրաքանչյուրը պետք է ունենա ժամանակը առաջին տողում և, որին հաջորդի կողի ինֆորմացիան յուրաքանչյուր կայանի համար (նկ. 4.16):

00:00 June 01, 16

```
37875 326// 70301 10133 20092 38988 48482 51001 86930 555 10013=  
37792 32998 00000 10047 20033 38224 48488 54000=  
37682 32597 33402 10069 20064 38106 48487 53002 81940=  
37686 32999 00000 10066 20038 38455 48484 52001=  
37801 325// 53501 10154 20042 38010 48472 54000 85300=  
37950 323// 80701 10072 20045 38153 48494 52003 885//=  
37895 42996 62001 10123 20099 39173 48114 52004 86080=
```

Նկար 4.16. SYNOP կողային ֆայլի օրինակ:

Երկրորդ output բաժինը պարունակում է “prefix” պարամետրը, որն օգտագործվում է LITTLE_R ֆայլերի անվանումներում:

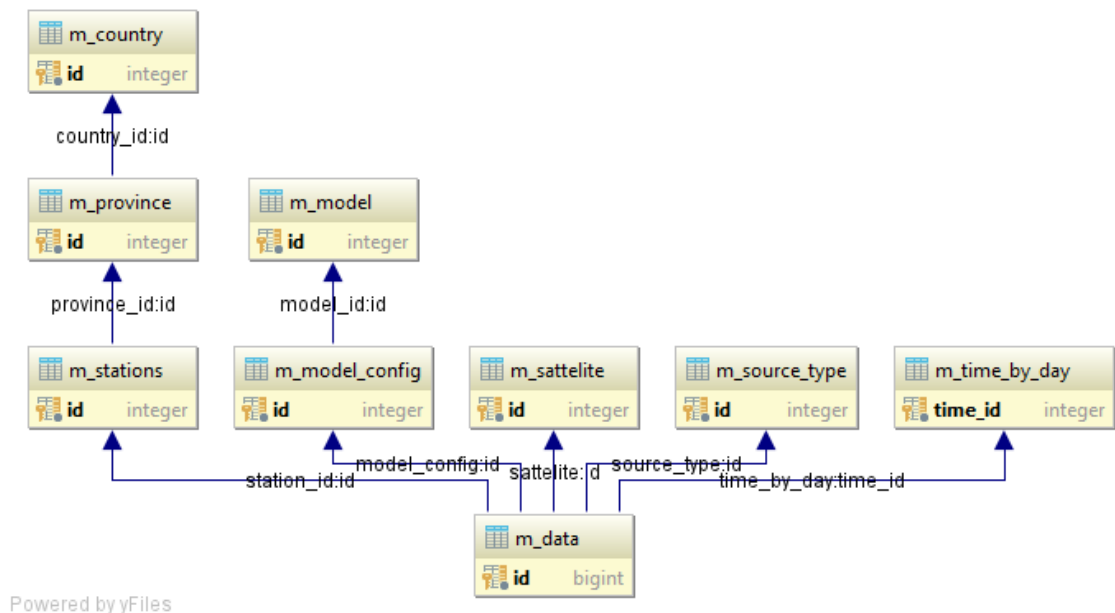
Ծրագիրը օգտագործելու համար անհրաժեշտ է լրացնել “namelist” կարգավորման ֆայլի արգումենտները համապատասխան արժեքներով, և դրանից հետո աշխատեցնելով գլխավոր “run.py” ֆայլը՝ արդյունքում կգեներացվեն համապատասխան LITTLE_R ֆայլերը:

Օդերևութաբանական 47 կայաններից դիտարկումները հաղորդվում են օրվա ընթացքում 8 անգամ (00, 03, 06, 09, 12, 15, 18, 21 ժամերի համար) և համախմբվում են մեկ ընդհանուր զեկույցի մեջ: Յուրաքանչյուր խումբ պարունակում է դիտարկումը ֆիքսված ժամանակի համար և այս ժամանակների միջև ինտերվալը կազմում է 3 ժամ: Յուրաքանչյուր LITTLE_R ֆայլ պարունակում է մեկ կամ ավել կայանների դիտարկումներ, բայց միևնույն ժամանակի համար: Այսպիսով, ամեն օրական զեկույցի համար ծրագիրը կառուցում է 8 LITTLE_R ֆորմատի ֆայլեր:

4.2.4 Տվյալների կառավարում և վերլուծություն

Միջավայրի համար օգտագործվող տվյալների պահպանման համար օգտագործվել է PostgreSQL տվյալների բազայի կառավարման համակարգը՝ մշակված POSTGRES 4.2-ի հիման վրա Կալիֆորնիայի համալսարանի Berkeley Computer Science Department-ում: Աշխարհագրական տվյալների պահպանման և կիրառման համար օգտագործվել է PostgreSQL-ի PostGIS ընդլայնումը, որն ավելացնում է լրացուցիչ ֆունկցիոնալություններ, ընդլայնում է PostgreSQL-ի հնարավորությունները տարածական տվյալների պահպանման, հարցման և նրանց ղեկավարման տեսանկյունից:

Տվյալների պահպանման համար տվյալների բազայում առկա են հիմնական 9 աղյուսակներ (նկ. 4.17):



Նկար 4.17. Տվյալների բազայում առկա հիմնական աղյուսակները:

Տարածաշրջանի տարածական տվյալների պահպանման համար օգտագործվում են m_country, m_province, m_stations աղյուսակները համապատասխանաբար երկրի, մարզերի և կայանների համար: Կառուցված հիերարխիկ կառուցվածքը ընդլայնման հնարավորություն է տալիս: Այս աղյուսակում պահվում են երկրի, մարզերի աշխարհագրական ծածկույթը, իսկ կայանների համար իրենց դիրքը: Եղանակի կանխատեսման մոդելների և նրանց կարգավորումների մասին ինֆորմացիան պահվում է m_model և m_model_config աղյուսակներում: Ներկա պահին օգտագործվում է միայն

WRF մոդելը, բայց կառուցվածքը թույլ է տալիս օգտագործել նաև այլ մոդելների արդյունքները: Արբանյակային նկարների աղբյուրների մասին ինֆորմացիան պահվում է `m_sattelite` աղյուսակում: Տվյալի աղբյուրի մասին ինֆորմացիան պահվում է `m_source_type` աղյուսակում, որը ներկա պահին ընդունում է 3 արժեք. 0 – օդերևութաբանական կայանի դիտարկում, 1 – մոդելի արդյունք, 2 – արբանյակային պատկերի վերլուծություն: Ժամանակային տվյալները պահվում են `m_time_by_day` աղյուսակում, պարունակելով մանրամասն ինֆորմացիա տվյալ ժամանակային պահի մասին (ժամ, օր, շաբաթ, ամիս, եռամսյակ և տարի):

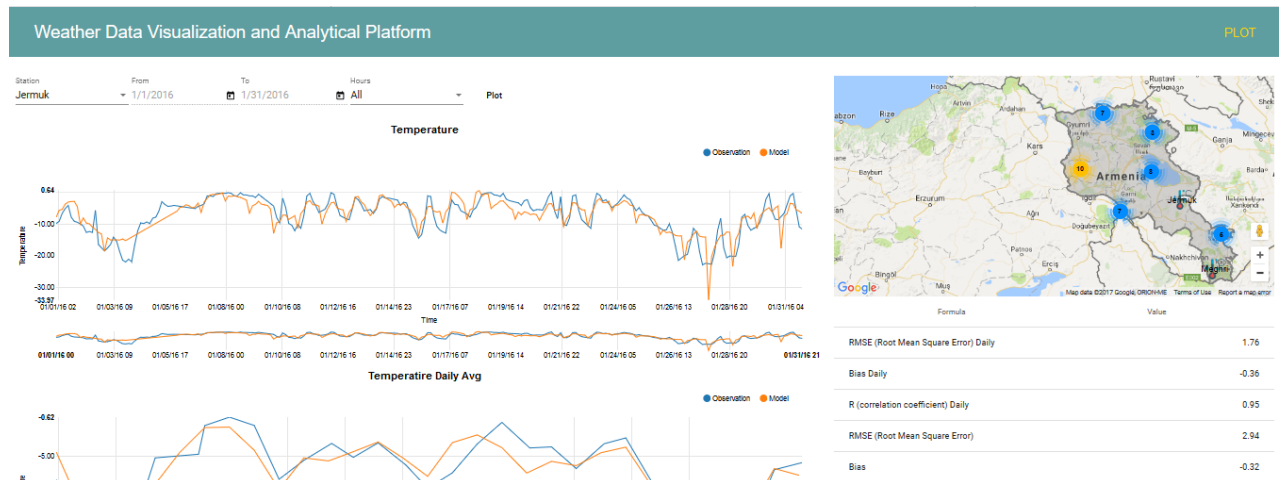
Բոլոր աղբյուրների տվյալները պահվում են `m_data` աղյուսակում, որը պարունակում է հղումներ նկար 4.17-ում նշված աղյուսակներին: Այն պարունակում է եղանակի վիճակի մասին տեղեկություններ, ինչպիսիք են ջերմաստիճան, քամու արագություն և ուղղություն, նվազագույն և առավելագույն ջերմաստիճան, հարաբերական խոնավություն, տեղումների քանակ և այլն:

Եղանակի կանխատեսման և ուսումնասիրության համար անհրաժեշտ տվյալների ծավալի առկայությունը առաջ է բերում նոր տեխնոլոգիաների կիրառման: Ներկա միջավայրում այդ խնդրի լուծման համար կիրառվել է SOLAP տեխնոլոգիան: Հայաստանի տարածքի համար օգտագործվում է 42 օդերևութաբանական կայաններից հաղորդած դիտարկումները: Այդ տվյալներին ավելանում են նույն քանակի մոդելի արդյունքները և մոտ 4 անգամ քիչ արբանյակային տվյալները: Գումարելով այս բոլորը և հաշվի առնելով այն հանգամանքը, որ որքան շատ է պատմական ինֆորմացիան, այնքան ավելի լավ հետազոտությունների արդյունք կարելի է ստանալ՝ SOLAP-ը լավ հնարավորություն է ընձեռում հաղթահարելու այսքան ահռելի քանակի ինֆորմացիայի հետ աշխատելու բարդությունները:

4.2.5 Օգտագործողի միջավայր և վիզուալիզացիա

Օգտագործողի միջավայրը տրամադրում է տարատեսակ վիզուալիզացիաներ հիմնված տվյալների պահոցում հավաքագրված տվյալների վրա: Օգտագործելով հարմարավետ և ինտերակտիվ վիզուալիզացիաների գրաֆիկները, որոնք ցույց են տալիս օդերևութաբանական կայանների դիտարկումները և WRF մոդելի

կանխատեսման արդյունքները, կարելի է կատարել համեմատական վերլուծություններ (նկ. 4.18):



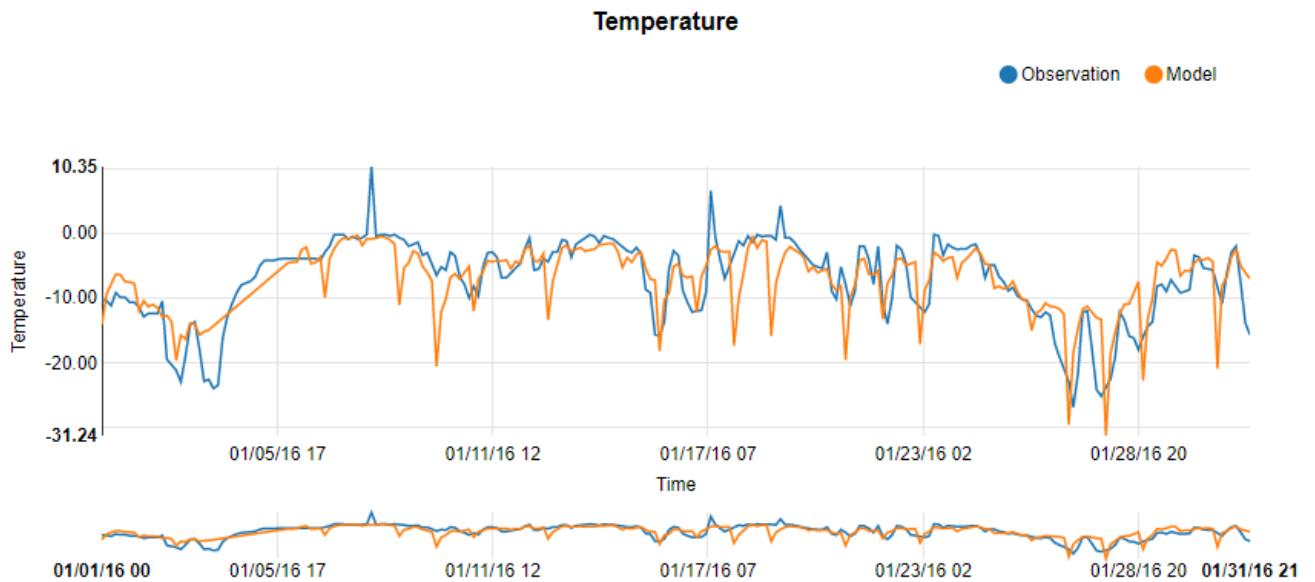
Նկար 4.18. Վեր հիմքով վիզուալիզացիայի և վերլուծության միջավայր:

Համակարգում մուտքագրված տվյալների հիման վրա վիզուալիզացիոն արդյունք տեսնելու համար անհրաժեշտ է նշել համապատասխան արժեքները հարցման դաշտերում (նկ. 4.19):

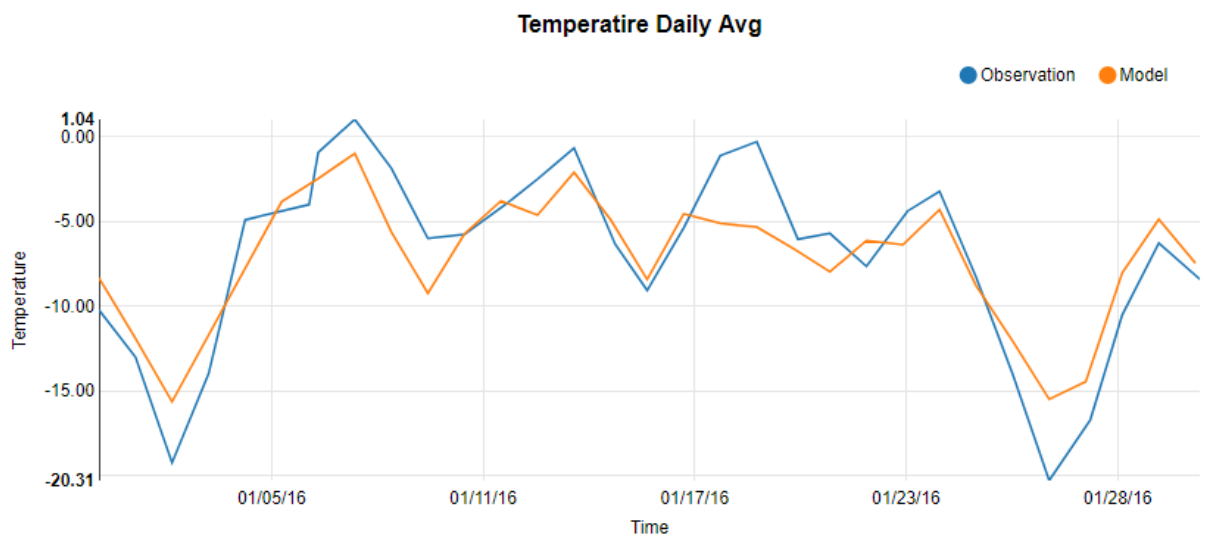
Station: **Aparan** From: **1/1/2016** To: **1/31/2016** Hours: **Hours** Plot

Նկար 4.19. Հարցման համար անհրաժեշտ տվյալների դաշտերը:

“Station” դաշտում անհրաժեշտ է ընտրել համապատասխան օդերևութաբանական կայանը, “From” և “To” դաշտերում նշել համապատասխան դիտարկվող ժամանակահատվածի սկիզբը և ավարտը, իսկ “Hours” դաշտում ընտրել օրվա դիտարկման ժամը (0, 3, 6 ... 21, չնշելու դեպքում օգտագործվում են բոլոր ժամերը): Դաշտերը լրացնելուց հետո սեղմելով “Plot” կոճակը կարելի է տեսնել վիզուալիզացիայի արդյունքները (նկ. 4.20, 4.21):



Նկար 4.20. Փաստացի և մոդելի կանխատեսած ջերմաստիճանային տվյալների գրաֆիկ. Օր.՝ Ապարանի օդերևութաբանական կայան:



Նկար 4.21. Փաստացի և մոդելի կանխատեսած օրական միջին ջերմաստիճանային տվյալների գրաֆիկ. Օր.՝ Ապարանի օդերևութաբանական կայան:

Բացի տվյալների գրաֆիկական ներկայացումներից հաշվարկվում և ցուցադրվում են նաև վերլուծությունների համար անհրաժեշտ գործակիցները (նկ. 4.22).

RMSE (Root Mean Square Error) Daily	2.54
Bias Daily Root Mean Square Error	0.31
R (correlation coefficient) Daily	0.94
RMSE (Root Mean Square Error)	4.17
Bias	-0.05

Նկար 4.22. Հարցված տվյալների հիման վրա հաշվարկված գործակիցները:

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (\hat{y}_t - y_t)^2}{n}}, \quad Bias = \frac{\sum_{t=1}^T (\hat{y}_t - y_t)}{n},$$

$$R = \frac{\sum_{t=1}^T (\hat{y} - m_{\hat{y}})(y - m_y)}{\sqrt{\sum_{t=1}^T (\hat{y} - m_{\hat{y}})^2 \sum_{t=1}^T (y - m_y)^2}}$$

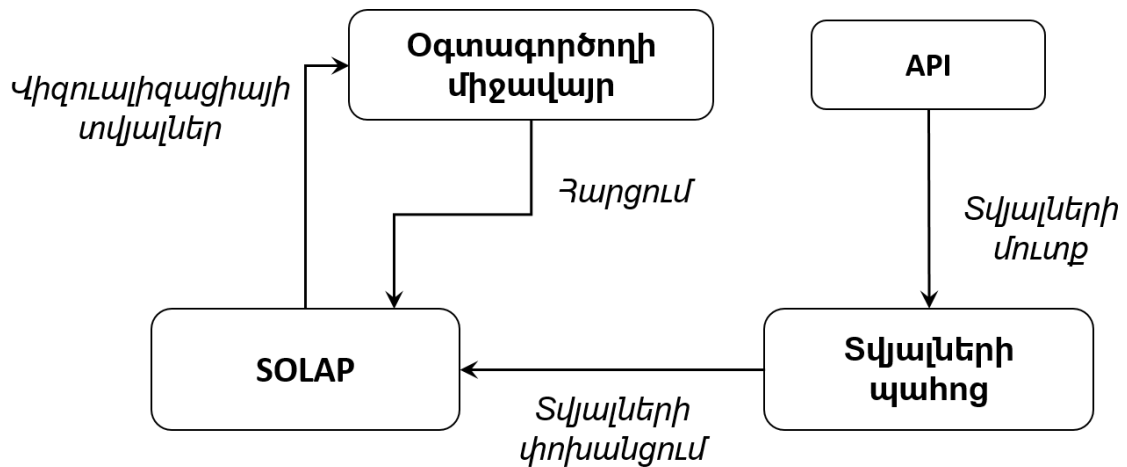
որտեղ $\hat{y}_t$ -ն t պահին փաստացի դիտարկումն է, y_t -ն մոդելի կանխատեսած արժեքը, իսկ $m_{\hat{y}}$ և m_y $[1, T]$ միջակայքի համապատասխան միջին արժեքներն են: Մոդելի կանխատեսման և դիտարկման արժեքների տարբերության միջին տատանումը չափվում է միջին քառակուսային սխալով (RMSE), որը բացասական կողմնորոշված մեծություն է և նրա արժեքներն ընկած են $[0, \infty)$ միջակայքում: Նախընտրելի է այն մոդելը, որի համար հաշվարկված միջին քառակուսային սխալը ստացվում է առավել փոքր, այսինքն՝ տվյալ մոդելն օժտված է կանխատեսումային առավել ունակություններով: Bias-ը նկարագրում է փաստացի և կանխատեսված տվյալների միջև միջին շեղումը: R հարաբերակցության գործակիցը օգտագործվում է գտնելու համար, թե որքան ամուր է տվյալների միջև հարաբերությունը: R-ի բանաձևը վերադարձնում է արժեք $[-1, 1]$ միջակայքից, որտեղ 1 արժեքը նշանակում է ամուր դրական հարաբերություն (մեկ փոփոխականի յուրաքանչյուր դրական աճի դեպքում մյուսում դիտվում է ֆիքսված համամասնության դրական աճ), -1-ը՝ ամուր բացասական

հարաբերություն (մեկ փոփոխականի յուրաքանչյուր դրական աճի դեպքում մյուսում դիտվում է ֆիքսված համամասնության բացասական անկում), իսկ O-ի դեպքում տվյալների միջև հարաբերությունը բացակայում է:

Նկար 4.23. Google Maps քարտեզ՝ նշված ուսումնասիրվող տարածքը և
օդերևութաբանական կայանները:

Օգտագործողի միջավայրը կարող է օգտագործվել տվյալների բազայում նոր տվյալներ վերբեռնելու համար՝ տրամադրելով կայանների ինֆորմացիան CSV ֆորմատի համապատասխան կառուցվածքով և WRF մոդելի կանխատեսման արդյունքների ինֆորմացիան netCDF ֆորմատով: Python ծրագրավորման լեզվով կառուցված ծրագիրը գտնում է կայաններին համապատասխանող ինֆորմացիան և API-ի միջոցով լրացնում է այն տվյալների բազայում: Դիտարկման տվյալների համար մշակվել է ձևափոխիչ

ծրագիր JavaScript ծրագրավորման լեզվով՝ SYNOP տեսակի ֆայլերը կարդալու համար (նկ. 4.24):



Նկար 4.24. Տվյալների հոսք:

4.2.6 Կիրառություն

Որպես օրինակ 2016թ.-ի հունվար ամսվա համար դիտարկումների և WRF կանխատեսման մոդելի տվյալների օգտագործմամբ հետազոտվել է երկրի մակերևույթից 2 մետր բարձրության ջերմաստիճանը: Օպերատիվ կանխատեսումների ստուգման վիճակագրությունը հաշվարկվել է Հայաստանի տարածքով տարածված 42 դիտարկման կետերում: Յուրաքանչյուր կանխատեսման ժամի՝ դիտարկման և մոդելի կանխատեսած տվյալների համար հաշվարկվել են RMSE, BIAS և R հարաբերակցության գործակիցները:

Աղյուսակ 1-ում ներկայացված են որոշ օդերևութաբանական կայանների համար ջերմաստիճանի կանխատեսումների ստուգման միջին ցուցանիշները: Կանխատեսման և դիտարկումների միջև միջին տարբերությունը 4-5°C է: Ուսումնասիրված բոլոր կայանների համար ձեռք են բերվել ընդունելի կոռելտության գործակիցներ, որոնք թույլ են տալիս դատել մոդելի համապատասխանության մասին: Վերլուծությունները, որոնք կատարվել են 42 կայանների համար, ցույց են տալիս, որ գրեթե բոլոր կայանների համար մոդելի արդյունքում ստացված տվյալները տվել են փաստացի տվյալներից դեպի վերև արժեքներ:

Աղյուսակ 1. Որոշ օդերևութաբանական կայանների համար ջերմաստիճանի կանխատեսումների ստուգման միջին ցուցանիշներ:

Տարածաշրջան	Կայան	Բարձր ծ.մ., մ	RMSE	BIAS	R
Արարատ	Արմավիր	870	3,97	-3,53	0,82
	Արարատ	818	5,65	-4,83	0,62
	Արտաշատ	829	5,69	-5,10	0,73
	Մերձավան	942	4,71	-4,11	0,79
Երևան	Զվարթնոց	853	4,44	-3,91	0,82
	Արաբկիր	1113	3,24	-1,92	0,77
Սյունիք	Մեղրի	627	9,89	-7,37	-0,70
	Կապան	705	9,69	-7,81	-0,43
Տավուշ	Իջևան	732	7,61	-4,47	-0,40
	Բագրատաշեն	453	7,56	-6,02	-0,31
Շիրակ	Աշոտցք	2012	2,52	-1,10	0,91
	Գյումրի	1513	3,91	-3,04	0,82
Գեղարքունիք	Մասրիկ	1934	5,13	-4,14	0,80
	Գավառ	1960	2,74	-1,61	0,89
	Սևանա լիճ	1917	2,59	1,74	0,90

RMSE-ի արժեքները մոտավորապես 1,6-2,5°C են՝ 1500-2000մ բարձրությունների վրա գտնվող կայանների համար, 4,6-9,8°C՝ 1000 մետրից պակաս բարձրության վրա գտնվող կայանների համար: Դա նշանակում է, որ մոդելը լավ է կանխատեսում երկրի մակերևույթից 2մ բարձրության ջերմաստիճանը 1500մ-ից բարձր գտնվող կայանների համար և անբավարար արդյունքներ է տալիս 1500մ-ից ցածրերի համար: Ամենավատ արդյունքները ստացվում են 1000մ-ից պակաս բարձրության վրա գտնվող կայանների համար, ինչպիսիք են՝ Մեղրիում (627մ), Կապանում (705մ), Իջևանում (732մ), Բագրատաշենում (453մ) գտնվող կայանները: 2016թ.-ի հունվարին ջերմաստիճանը Արարատյան դաշտավայրի և Երևանի համար նույնպես անբավարար է: Արարատյան դաշտավայրում RMSE-ն եղել է 5,6-6,4°C, որն ամենավատ արդյունքն է ստացված

Արարատի կայանից: Երևանում լավ արդյունք է ստացվել Երևան-Արաբկիր կայանի համար, որը տեղակայված է 1113մ բարձրությունում, իսկ ամենավատ արդյունքը Երևան-Զվարթնոցի համար՝ 853մ: Լեռնային և ցածրադիր տարածքների համար համադրման գործակիցը տատանվում է 0,80-0,97, հովիտներում՝ 0,25-0,7 և -0,7-ից -0,31:

Վերը ներկայացվածից կարելի է եզրակացնել, որ WRF մոդելը, նկարագրված կարգավորումներով, ցուրտ ժամանակաշրջանում երկրի մակերևույթից 2մ բարձրության փոփոխական օդի ջերմաստիճանի համար տալիս է դրական կանխատեսում հանրապետության լեռնային և հարավային շրջաններում (Շիրակ, Կոտայք, Գեղարքունիք, Լոռի և Արագածոտնի լեռնային և հովտային շրջաններ), միևնույն ժամանակ անբավարար արդյունք Սյունիքի և Արարատի դաշտավայրերում, Տավուշում և Երևանում:

Վերլուծելով ջերմաստիճանը կազմավորող բոլոր գործոնները, պարզ է դառնում, որ ընտրված ժամանակահատվածում ձորերում առկա էր մառախուղ, ինչը ցածր ջերմաստիճանի առաջացման պատճառ էր: Հետևաբար, կարելի է եզրակացնել, որ WRF մոդելը, ինչպես նաև այլ կանխատեսող մոդելները, չեն կանխատեսում մառախուղի առկայության պայմաններում, որը շատ մեծ ազդեցություն ունի ընդհանուր ջերմային ֆոնի վրա:

4.2.7 Եզրակացություն

Առաջարկվող միջավայրը հնարավորություն է տալիս ինտեգրել արդեն հասանելի դիտարկման, մոդելի կանխատեսման և բազմատեսակ արբանյակային պատկերների արդյունքները և օգտագործել այդ տվյալները վեբ միջավայրում ուսումնասիրությունների և վերլուծությունների համար: Երկրի մակերևույթից 2մ բարձրության ջերմաստիճանի ինտերակտիվ համեմատական գրաֆիկները թույլ են տալիս իրականացնել տեսողական վերլուծություն և ստանալ ինֆորմացիա կանխատեսման մոդելի ճշգրտության մասին: Ծառայությունը հնարավորություն է տալիս ճշգրտելու կանխատեսումների արդյունքները լրացուցիչ մեթոդներով, վիճակագրական վերլուծությունների իրականացման միջոցով և բավականին բարձր արդյունք է ապահովում այն դեպքերում, երբ մոդելի զգայունությունը ցածր է:

Հետագա աշխատանքների նպատակն է բարելավել միջավայրը՝ ավելացնելով տարբեր տեսակի նոր վիզուալիզացիաներ՝ վերլուծելու և համեմատելու համար մակերևույթին մոտ այլ մթնոլորտային տարրեր: Վտանգավոր հիդրոոդերետութաբանական երևույթների ազդանշանային համակարգի մշակման համար նախատեսվում է իրականացնել արհեստական բանականության և արբանյակային պատկերների օգտագործման վրա հիմնված տարբեր մեթոդաբանություններ:

Աշխատանքի հիմնական արդյունքները

1. Մշակվել է տվյալների վերլուծության և ինտերակտիվ վիզուալիզացիայի՝ տվյալների ստացման, մշակման, ձևափոխման, պահպանման, բարձր արտադրողականությամբ հաշվարկների իրականացման, օպտիմիզացիայի և տարածման մեթոդներ պարունակող, համալիր համակարգ [43] [44] [55] [56] [68]:
2. Մշակվել է տվյալների լոկալ պահպանման հնարավորությամբ միջավայր՝ խնդրին համապատասխան մետատվյալների ստեղծման և նրանց տարածման համար, որում հաշվի են առնված միջազգային հիմնօրինակները [43]:
3. Մշակվել է բազմաօգտատեր ամպային ծառայություն քվանտային ֆիզիկայում ֆոտոնիկ դիսիպատիվ համակարգերում թվային հաշվարկների կատարման, արդյունքների վերլուծության և ինտերակտիվ վիզուալիզացիայի իրականացման համար [44]:
4. Մշակվել է ամպային ծառայություն Հայաստանում եղանակի ուսումնասիրության և կանխատեսման համար օգտագործվող թվային մեթոդների արդյունքների վերլուծության և ինտերակտիվ վիզուալիզացիայի համար [55] [56]:

Օգտագործված գրականության ցանկ

- [1] V. Dhar, "Data science and prediction" *Communications of the ACM*, vol. 56, no. 12, pp. 64-73, 2013.
- [2] S. R. Sukumar, M. A. Matheson, R. Kannan and S.-H. Lim, "Mini-apps for high performance data analysis" in *IEEE*, Washington, DC, USA, 2016.
- [3] E. T. d'Amico and G. Guignard, "Special lattice computation for the CERN Compact Linear Collider" *PHYSICAL REVIEW SPECIAL TOPICS - ACCELERATORS AND BEAMS*, vol. 4, no. 021002, 2001.
- [4] F. Nielsen, Introduction to HPC with MPI for Data Science, Springer International Publishing, 2016.
- [5] S.-J. Chen, G.-H. Lin, P.-A. Hsiung and Y.-H. Hu, Hardware Software Co-Design of a Multimedia SOC Platform, Springer Netherlands, 2009.
- [6] J. Mathew and D. Vijayakumar, "The Performance of Parallel Algorithms by Amdahl's Law, Gustafson's Trend" *International Journal of Computer Science and Information Technologies*, vol. 2, no. 6, pp. 2796-2799, 2011.
- [7] OpenMP Architecture Review Board, "OpenMP Application Programming Interface" <https://www.openmp.org>, 2015.
- [8] J. Bruck, D. Dolev, C.-T. Ho, M.-C. Roşu and R. Strong, "Efficient Message Passing Interface (MPI) for Parallel Computing on Clusters of Workstations" *Journal of Parallel and Distributed Computing*, vol. 40, no. 1, pp. 19-34, 1997.
- [9] R. Buyya, High Performance Cluster Computing: Architectures and Systems, Melbourne: Prentice Hall, 1999.
- [10] N. M. Gonzalez, T. C. M. d. B. Carvalho and C. C. Miers, "Cloud resource management: towards efficient execution of large-scale scientific applications and workflows on complex infrastructures" *Journal of Cloud Computing Advances, Systems and Applications*, vol. 6, no. 13, 2017.
- [11] U. Fayyad, G. Piatetsky-Shapiro and P. Smyth, "From Data Mining to Knowledge Discovery in Databases" *AI MAGAZINE*, pp. 37-54, 1996.

- [12] E. Curry, *The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches*, Springer, Cham, 2016.
- [13] D. Laney, "3-D Data Management: Controlling Data Volume, Velocity and Variety" META Group Research Note, 2001.
- [14] W. Fan and A. Bifet, "Mining big data: current status, and forecast to the future" *ACM SIGKDD Explorations Newsletter*, vol. 14, no. 2, pp. 1-5, 2012.
- [15] D. Zhang, "Inconsistencies in big data" in *12th IEEE International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)*, New York, NY, USA, 2013.
- [16] Wikipedia, "Machine Learning" https://en.wikipedia.org/wiki/Machine_learning, 2018.
- [17] Z. Zhang, "A gentle introduction to artificial neural networks" *Ann Transl Med*, vol. 4, no. 19, p. 370, 2016.
- [18] E. Zudilova-Seinstra, T. Adriaansen и R. v. Liere, *Trends in Interactive Visualization*, Springer-Verlag London, 2009.
- [19] S. Few, "Eenie, Meenie, Minie, Moe: Selecting the Right Graph for Your Message" Perceptual Edge, 2004.
- [20] R. Webb, «Stella: Polyhedron navigator» *Symmetry: Culture and Science*, т. 11, № 1-4, pp. 231-268, 2003.
- [21] Lidong Wang, Guanghui Wang, Cheryl Ann Alexander, "Big Data and Visualization: Methods, Challenges and Technology Progress" *Digital Technologies*, vol. 1, no. 1, pp. 33-38, 2015.
- [22] M. Khan and S. S. Khan, "Data and Information Visualization Methods, and Interactive Mechanisms: A Survey" *International Journal of Computer Applications*, vol. 34, no. 1, 2011.
- [23] J. E. M. III, *Political Analysis Using R*, Springer International Publishing, 2015.
- [24] OriginLab Corporation, "Origin" <https://www.originlab.com/>, 2018.
- [25] K. Inc, "ParaView" <https://www.paraview.org/>, 2017.
- [26] M. Bostock, "D3js Visualization Library" <https://d3js.org/>, 2017.

- [27] Mozilla and individual contributors, "Document Object Model (DOM)" https://developer.mozilla.org/enUS/docs/Web/API/Document_Object_Model/Introduction, 2005-2018.
- [28] A. B.V., "Vis.js" <http://visjs.org/>, 2017.
- [29] V. Agafonkin, "LeafletJs" <http://leafletjs.com/>, 2017.
- [30] Bestiario_, "Quadrigram," <http://www.quadrigram.com>, 2017.
- [31] D. Crockford, "The application/json Media Type for JavaScript Object Notation (JSON)" <https://www.rfc-editor.org/rfc/pdf/rfc4627.txt.pdf>.
- [32] "Google Fusion Tables" https://developers.google.com/fusiontables/docs/v2/getting_started, 2016.
- [33] Datawrapper GmbH, "Datawrapper," <https://www.datawrapper.de/>, 2017.
- [34] TABLEAU SOFTWARE, "Tableau," <https://www.tableau.com/>, 2018.
- [35] H. T. Blanke и A. Hasan, «Rule-based curation and preservation of data: A data Grid approach using iRODS » *Future Gener. Comput. Syst.*, т. 25, № 4, pp. 446-452, 2009.
- [36] iRODS Consortium, "iRODS Client," <https://github.com/irods/python-irodsclient>.
- [37] S. Chaudhuri and U. Dayal, "An overview of data warehousing and OLAP technology" *ACM SIGMOD Record*, vol. 26, no. 1, pp. 65-74, 1997.
- [38] M. Whitehorn, R. Zare and M. Pasumansky, *Fast Track to MDX*, London: Springer-Verlag London, 2006.
- [39] S. Aissi, M. S. Gouider, T. Sboui and L. B. Said, "Enhancing Spatial Data Warehouse Exploitation: A SOLAP Recommendation Approach" *Computer and Information Science*, vol. 656, pp. 131-147, 2016.
- [40] L. Chittaro, "Visualizing Information on Mobile Devices" *ACM Computer*, vol. 39, no. 3, pp. 40-45, 2006.
- [41] G. Berg-Cross, R. Ritz and P. Wittenburg, "RDA DFT Core Terms and Model" <http://b2share.eudat.eu>, 2016.
- [42] Mark D. Wilkinson et al, "The FAIR Guiding Principles for scientific data management and stewardship" *Scientific Data*, vol. 3, 2016.

- [43] H. Grigoryan, "Increasing the Visibility of Scientific Data in Armenia Using Persistent Identifiers" *Mathematical Problems of Computer Science*, vol. 46, pp. 81-86, 2016.
- [44] H. Grigoryan, H. Astsatryan, T. Gevorgyan and V. Manukyan, "Cloud Service for Numerical Calculations and Visualizations of Photonic Dissipative Systems" *Cybernetics and Information Technologies*, vol. 17, no. 5, pp. 89-100, 2017.
- [45] N. Gisin and L. C. Percival, "The quantum-state diffusion model applied to open systems" *Cambridge University Press*, vol. A25, p. 5677, 2000.
- [46] H. Astsatryan, T. Gevorgyan and A. Shahinyan, "Web Portal for Photonic Technologies Using Grid Infrastructures" *Journal of Software Engineering and Applications*, vol. 5, no. 11, pp. 864-869, 2012.
- [47] H. J. Carmichael, P. Drummond, P. Meystre and D. F. Walls, "Intensity correlations in resonance fluorescence with atomic number fluctuations" *Journal of Physics A: Mathematical and General*, vol. 11, no. 5.
- [48] K. M. Birnbaum, A. Boca, R. Miller, A. D. Boozer, T. E. Northup and H. J. Kimble, "Photon blockade in an optical cavity with one trapped atom" *Nature*, vol. 436, pp. 87-89, 2005.
- [49] L. Tian and H. J. Carmichael, "Quantum trajectory simulations of two-state behavior in an optical cavity containing one atom" *Phys. Rev. A*, vol. 46, no. 11, 1992.
- [50] G. Aad et al. (ATLAS Collaboration), "Search for dilepton resonances in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector" *Phys. Rev. Lett.*, vol. 107, no. 27, 2011.
- [51] A. Verger, C. Ciuti and I. Carusotto, "Polariton quantum blockade in a photonic dot" *Phys. Rev. B*, vol. 73, no. 19, 2006.
- [52] A. Faraon, A. Majumdar and J. Vučković, "Generation of nonclassical states of light via photon blockade in optical nanocavities" *Phys. Rev. A*, vol. 81, no. 3, 2010.
- [53] N. Gisin and I. C. Percival, "The quantum-state diffusion model applied to open systems," *Journal of Physics A: Mathematical and General*, vol. 25, no. 21, 1992.
- [54] A. Jain, B. Hills and S. SD, "Data visualization with the D3.JS Javascript library" *Journal of Computing Sciences in Colleges*, vol. 30, no. 2, pp. 139-141, 2014.

- [55] H. Grigoryan and R. Abrahamyan, "Cloud Service for Analysis and Interactive Visualization of Weather Data in Armenia" *Mathematical Problems of Computer Science*, vol. 49, pp. 49-57, 2018.
- [56] H. Astsatryan, H. Grigoryan, E. Gyulgyulyan, A. Hakobyan, A. Kocharyan, W. Narsisian, V. Sahakyan, Y. Shoukourian, A. Mkoyan, R. Abrahamyan, Z. Petrosian and J. Aligon, "Weather Data Visualization and Analytical Platform" *Scalable Computing: Practice and Experience*, vol. 19, no. 2, pp. 79-86, 2018.
- [57] A. Gevorgyan, "Summertime wind climate in Yerecan: valley wind systems" *Climate Dynamics*, vol. 48, no. 56, pp. 1827-1840, 2017.
- [58] A. Gevorgyan, H. Melkonyan, R. Abrahamyan, Z. Petrosyan, A. Shahnazaryan, H. Astsatryan, V. Sahakyan and Y. Shoukourian, "A Persistent Surface Inversion Event in Armenia as Simulated by WRF Model" in *CSIT*, Yerevan, 2015.
- [59] C. Faccini, D. Cimini, R. Ferratti and A. Taramasso, "3DVAR assimilation of SSM/I data over the sea for the IOP2b MAP case" *Adv Geosci*, vol. 2, pp. 229-235, 2005.
- [60] H. Astsatryan, V. Sahakyan, Y. Shoukourian, P.-H. Cross, M. Dayde, J. Dongarra and P. Oster, "Strengthening Compute and Data intensive Capacities of Armenia" in *IEEE Proceedings of 14th RoEduNet International Conference - Networking in Education and Research, NER*, 2015.
- [61] Y. Shoukourian, V. Sahakyan and H. Astsatryan, "E-Infrastructures in Armenia: Virtual research environments" in *IEEE Proceedings CSIT - 9th International Conference on Computer Science and Information Technologies*, Yerevan, 2013.
- [62] H. Astsatryan, Y. Shoukourian and V. Sahakyan, "The ArmCluster Project: Brief Introduction" in *Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications*, 2004.
- [63] J. A. Sobrino, J. C. Jimenez-Munoz and L. Paolini, "Land surface temperature retrieval from Landsat TM 5" *Remote Sensing of Environment*, vol. 90, pp. 434-440, 2004.
- [64] M. Netelera, M. H. Bowman, M. Landac and M. Metza, "GRASS GIS: A multi-purpose open source GIS" *Environmental Modelling & Software*, vol. 31, pp. 124-130, 2012.
- [65] W. C. Skamarock and J. B. Klemp, "A time-split nonhydrostatic atmospheric model for weather research and forecasting applications" *Journal of Computational Physics*, vol. 227, no. 7, pp. 3465-3485, 2008.

- [66] J. G. Powers, J. B. Klemp, W. C. Skamarock, C. A. Davis, J. Dudhia, D. O. Gill, J. L. Coen и D. J. Gochis, «The Weather Research and Forecasting Model: Overview, System Efforts, and Future Directions» *Bulletin of the American Meteorological Society*, т. 98, № 8, pp. 1717-1737, 2017.
- [67] J. Ban, Z. Liu, X. Zhang, X.-Y. Huang and H. Wang, "Precipitation data assimilation in WRFDA 4D-Var: implementation and application to convection-permitting forecasts over United States" *Tellus A: Dynamic Meteorology and Oceanography*, vol. 69, no. 1, 2017.
- [68] H. Grigoryan, "Automation Script for WRF Model Data Assimilation System" *Mathematical Problems of Computer Science*, vol. 48, pp. 57-63, 2018.