

Халықаралық Ақпараттық Технологиялар Университеті

ӘОЖ: 004.451.7:004.7

Қолжазба құқығында

ИБРАЕВА ЖАНАР БАЗАРБЕКОВНА

Желілік трафикті талдау және болжау модельдерін жасау

6D070400 – Есептеу техникасы және бағдарламалық қамтамасыз ету

Философия докторы (PhD)
дәрежесін алу үшін дайындалған диссертация

Ғылыми кеңесшілер:
т.ғ.к., қауымд.проф. Бектемысова Г.У.,
PhD, қауымд.проф. Абдул Рахим Ахмад

Қазақстан Республикасы,
Алматы, 2023

МАЗМҰНЫ

БЕЛГІЛЕУЛЕР МЕН ҚЫСҚАРТУЛАР	4
КІРІСПЕ	5
1 МУЛЬТИСЕРВИСТІ ЖЕЛІДЕГІ ТРАФИКТІ ШОЛУ ЖӘНЕ МІНДЕТТЕРДІ ТҰЖЫРЫМДАУ.....	8
1.1 Қазақстандағы телекоммуникациялар желісінің негізгі ерекшеліктері.	8
1.2 Желілік трафикті болжау бойынша негізгі ғылыми еңбектерге шолу...	9
1.3 Уақыттық қатарларды талдаудағы негізгі түсініктер.....	11
1.4 Уақыттық қатарлардың модельдері.....	15
1.5 Уақыттық қатарды стационарлыққа тексеру.....	17
1.6 Уақыттық қатарларды болжау әдістері.....	18
1.7 Зерттеу мақсаты және міндеттерді тұжырымдау.....	20
1.8 Бірінші тарау бойынша қорытынды.....	21
2 ЖЕЛІЛІК ТРАФИКТІ ТАЛДАУ ЖӘНЕ СТАЦИОНАРЛЫҚТЫ АНЫҚТАУ БАҒДАРЛАМАСЫН ЖАСАУ	22
2.1 Өлшенген трафиктің сандық сипаттамаларын талдау.....	22
2.2 Трендтің бар болу гипотезасын тексеру.....	25
2.3 Уақыттық қатардың стационарлығын анықтайтын бағдарлама сипаттамасы.....	28
2.4 Уақыттық қатардың стационарлығын бағалау.....	29
2.5 Қатардың стационарлығын бірлік түбір (unit root) тесттері арқылы бағалау.....	31
2.6 Екінші тарау бойынша қорытындылар.....	33
3 УАҚЫТТЫҚ ҚАТАРДЫ УАҚЫТ-ЖИЛІК БОЙЫНША ТАЛДАУ	34
3.1 Спектрлік талдау әдістері.....	34
3.2 Сингулярлық спектрлі талдау.....	36
3.3 Сингулярлық спектрлі талдау әдісімен қатарды ыдырату бағдарламасының сипаттамасы.....	39
3.3.1 Сингулярлық спектрлі талдау әдісімен ыдыраған қатарды қалпына келтіру.....	44
3.4 Үшінші тарау бойынша қорытындылар.....	46
4 ЖЕЛІЛІК ТРАФИКТІ БОЛЖАУ МОДЕЛЬДЕРІН ЖАСАУ.....	47
4.1 Стационарлы емес уақыттық қатарларды болжаудың ерекшеліктері.....	47
4.2 ARIMA әдісімен болжау.....	50
4.3 Желілік трафикті зерттеу үшін нейрондық желіні қолдану.....	58
4.4 Жасанды нейронды желі негізіндегі желілік трафикті болжау.....	62
4.5 NARX ANN негізіндегі желілік трафикті болжау.....	65
4.6 Нейро-анық емес болжау модельдерін әзірлеу.....	74
4.6.1 Уақыттық қатарды болжауға нейро-анық емес Чен және Ченг алгоритмдерін қолдану.....	74
ҚОРЫТЫНДЫ.....	80
ПАЙДАЛАНЫЛҒАН ӘДЕБИЕТТЕР ТІЗІМІ	82

ҚОСЫМША А - Стационарлықты анықтайтын бағдарлама бөлігі.....	90
ҚОСЫМША Ә - ARIMA моделімен болжау кодының бөлігі	92
ҚОСЫМША Б - NARX (Nonlinear AutoRegressive exogenous Network) көмегімен болжау кодының бөлігі.....	93
ҚОСЫМША В - Нейро-анық емес болжау кодының бөлігі.....	96
ҚОСЫМША Г - Енгізу актісі.....	98
ҚОСЫМША Д - Авторлық куәлік №1.....	99
ҚОСЫМША Е - Авторлық куәлік №2.....	100

БЕЛГІЛЕУЛЕР МЕН ҚЫСҚАРТУЛАР

- АКТ – Ақпараттық коммуникациялық технологиялар
- АКФ – Автокорреляция функциясы
- ИИ – Искусственный интеллект (Жасанды интеллект)
- ИНС – Искусственная нейронная сеть (Жасанды нейрондық желі)
- ARIMA (Auto-Regressive Integrated Moving Average) – авторегрессивті интеграцияланған (айырма алынған) жылжымалы орташа
- IP (Internet Protocol) – желіаралық протокол
- NARX (Nonlinear AutoRegressive with exogenous Network) – сызықты емес экзогенді авторегрессивті желі
- ANFIS (Adaptive Neuro-Fuzzy Inference System) – бейімделетін нейро-анық емес қорытындылар жүйесі
- NGN (Next Generation Network) – келесі буын желілері
- MPLS (Multi-Protocol Label Switching) – көп протоколды ауыстыру
- MPEG (Moving Picture Experts Group) – қозғалыстағы сурет бойынша сарапшылар тобы
- MSE (Mean squared error) – орташа квадрат қатесі
- TDM (Time Division Multiplexing) – арналарды уақыт бойынша бөлу
- SDN/NFV (Software Defined Network/Network Functions Virtualization) - Бағдарламалық құралмен анықталған желі/желі функцияларын виртуалдандыру
- WT (Wavelet transform) – толқындық түрлендіру
- VoIP (Voice over Internet Protocol) – интернет арқылы дауыс беру протоколы
- UDP (User Datagram Protocol) – пайдаланушы датаграммасының протоколы
- IP VPN (Virtual private network) – виртуалды жеке желі
- IPTV (Internet protocol television) – Интернет протоколдық теледидар
- SSA (Singular Spectrum Analysis) – Сингулярлық спектрлі талдау
- RSME - Root mean square error (түбір орташа квадрат қатесі)
- MAE - Mean absolute error (Орташа абсолютті қате)

КІРІСПЕ

Дүниежүзілік байланыс желісінің пакеттік технология негізіндегі дамуы ақпарат ағындарымен байланысты деректер көлемінің күрт өсуіне себеп болды.

Қазақстанда 2007 жылдан бастап TDM (Time Division Multiplexing) арналарды ауыстыру (коммутация каналов) технологиясының IP (Internet Protocol) пакеттік технологияға ауысуы орасан зор өзгеріс алып келді. Ақпараттық коммуникациялық технологиялар (АКТ) саласында VoIP (Voice over Internet Protocol), IP VPN (Virtual private network), IPTV (Internet protocol television) қызметтерін көрсететін заманауи инфрақұрылым пайда болды. Бұл желі NGN (Next Generation Network) жаңа буын желісі болып табылады.

Қазіргі заманғы гетерогенді желі құрылымы өте қиын, біркелкі емес желілік трафик тудырады. Нақты өлшенген деректерді зерттеулер де желілік құрылғыларға қызмет көрсететін пакеттердің біркелкі емес қарқындылығын көрсетеді. Сонымен қатар уақыттық қатарлар негізінде құрылған модельдер стационарлық емес қасиетке ие, бұл олардың құрылымының көп компонентті екенін білдіреді. Демек, трафик көлемі ұлғайып қана қоймастан, оның құрылымы да өзгеруде, сондықтан желілік трафикті талдау әлі де өзекті мәселе болып табылады. Күрделі құрылымның құрамдас бөліктерін анықтау және сандық анықтау – трендтің, кезеңділіктің, кездейсоқ құрамдас бөліктің болуы/болмауы уақытша қатарларды талдаудың негізгі міндеті болып табылады.

Қазіргі заманғы зерттеулер желілік трафикті талдау және болжау трафикті басқарудағы ең маңызды міндет болып қала беретінін көрсетті. Желіні басқарудың міндеттері басқа нәрселермен қатар дұрыс шешім қабылдау үшін болжамды болашақ деректерге де негізделген.

Инфокоммуникациялық желілердегі гетерогенді трафик көлемінің өсуі ұсынылатын байланыс қызметтерінің сапасын қамтамасыз ету мәселелерін өзекті етеді, бұл өз кезегінде болжау моделдеріне жүгінуді талап етеді.

Болжамды деректер желідегі ақпарат ағындарын басқару мәселесін шешу үшін қажетті ақпаратты береді және басқару негізінде пакеттердің жоғалуын болдырмауға мүмкіндік береді.

Диссертациялық жұмыстың мақсаты нақты өлшенген желілік трафикті талдау және болжау үшін модельдерді әзірлеу болып табылады.

Қойылған мақсатқа жету үшін диссертация келесі **міндеттерді** шешеді:

1. Эмпирикалық деректердің уақыттық қатарының құрылымын зерттеу;
2. Уақыттық қатарларды болжаудың классикалық үлгілерімен эксперименталды зерттеулер жүргізу;
3. Желілік трафикті, оның гетерогенділігін ескере отырып, болжау үлгілерін әзірлеу;

Зерттеу объектісі. Зерттеу объектісі – әрбір секунд сайын (18000 нүкте) бес сағат бойы магистральдық желі арқылы берілетін MPEG хаттама пакеттерінің жиынтығын көрсететін уақыттық қатар.

Диссертациялық жұмыстың ғылыми жаңалығы келесі алынған нәтижелермен анықталады:

1. Нақты өлшенген желілік трафик құрылымына талдау жүргізудің нәтижесінде қатарды стационарлыққа тексеретін бағдарлама әзірленді;
2. Уақыттық қатарларды болжаудың классикалық ARIMA моделінің параметрлері анықталып, нәтижеде ARIMA(0,2,1) моделі құрылды;
3. NARX (Nonlinear AutoRegressive Network with exogenous inputs) негізінде болжау моделі құрылды;
4. Анық емес логика модельдері бағдарламалық іске асырылды.

Теориялық және практикалық құндылығы. Бұл жұмыстың теориялық маңыздылығы эмпирикалық деректердің стационарлы емес жағдайында қолдануға болатын болжау модельдерін анықтауда жатыр. Жұмыстың практикалық маңыздылығы ұсынылған болжау модельдерін желілік трафикті басқару жоспарын әзірлеуде және басқару кезінде дұрыс шешімдер қабылдауда жатыр. «Алматы Технологиялар Институтында» алынған нәтижелер сынақтан өткізілді.

Диссертация құрылымы кіріспеден, 4 бөлімнен, қорытындыдан, пайдалы дереккөздер мен қосымшалардан тұрады. Диссертация көлемі 100 бет, оның ішінде 41 сурет пен 6 кесте, 51 формула, 113 дереккөз бар.

Кіріспеде диссертацияның өзектілігін анықтайтын негіздемелер, ғылыми зерттеу жұмысының қалыптастырушы мақсаты, объектісі, міндеттері берілген. Жүргізілген зерттеу нәтижелеріне сипаттама беріліп, олардың ғылыми жаңалығы және практикалық маңыздылығы көрсетілген.

Диссертациялық жұмыстың **бірінші бөлімінде** мультисервистік желінің желілік трафигіне шолу және талдау, Қазақстандағы телекоммуникация желісінің негізгі ерекшеліктері, желілік трафикті болжау бойынша негізгі ғылыми еңбектерге шолу жасалды. Жалпы сипаттамалар мен зерттеу міндеттері берілген. Стохастикалық процестердің (қандай да бір жүйе күйінің ықтималдық заңдылықтарға сәйкес уақыт бойынша өзгеру процесі) классификациясы, деректердің уақытша қатарының стационарлық және стационарлық еместігін анықтау параметрлері, болжау әдістері анықталған.

Екінші бөлімде өлшенген трафиктің сандық сипаттамаларының талдауы жүргізіледі. Бастапқы деректер бес сағат бойы алынды, бұл әрбір секунд сайын бес сағат бойы магистральдық желі арқылы берілетін MPEG хаттама пакеттерінің жиынтығын көрсетеді.

Қатардың қалыптылығын статистикалық бағалау жүзеге асырылды, өйткені бұл желілік трафикті болжауда сәйкес әдістерді қолдану кезінде маңызды рөл атқарады. Зерттелетін қатардың таралу қалыптылығы туралы гипотеза тексерілді және келесі критерийлер жойылды: Колмогоров пен Смирновтың өзгертілген критерийлері; Крамер-Мизес; Андерсон-Дарлинг; Шапиро Франция; асимметрия коэффициенті; Жарк Бер; Гири және Д'Агостино. Жілілік бөлу формасының сандық сипаттамалары: асимметрия және куртозис зерттелетін қатардың қалыпты таралу заңына сәйкессіздігін көрсетті.

Зерттелетін уақыттық қатардың таралу заңы қалыпты болмағандықтан, трендтің бар екендігі туралы гипотезаны тексеру үшін параметрлік емес сынақтарды қолданып, нәтижесінде зерттелетін қатарда тренд бар екені

анықталды. Python бағдарламалық ортасында қатарды стационарлыққа тексеретін бағдарлама жасалды.

Үшінші бөлім уақыттық қатарлардың уақыттық-жиілік талдауына арналған. Зерттелетін қатар кездейсоқтық, стационарлық емес қасиеттерге ие болғандықтан, бұл қатарға сингулярлық спектрлі талдау SSA (Singular Spectrum Analysis) әдісі қолданылды. Меншікті векторлардың графиктері олардың құрамында жоғары жиілікті тербелістер бар екенін және шу құрамдастарына жатқызу керек екенін көрсетеді. Сингулярлы спектрлік талдау арқылы өңдеуден кейін қатар динамикасының қалпына келуі көрсетілген. Одан әрі қатар гармоникалық құрамдастарға зерттелді, ол төмен жиілікті компонентке (тренд) – бірінші негізгі компонент және жоғары жиілікті шумен қатар мерзімді құрамдастарға ыдырады. Уақыттық қатардың ыдырауы қатардың стационарлы емес екенін және оның тренд, гармоникалық және шу құрамдастарын қамтитынын көрсетті.

Төртінші бөлімде стационарлы емес уақыттық қатарларды болжау ерекшеліктерінің сипаттамасы берілген. Статистикалық тәсілдердің ішінде ARIMA (Авто-регрессивті айырымды жылжымалы орташа) әдісі бастапқы уақыттық қатардан кейбір реттік айырмаларды алу арқылы стационарлы қатарға келтірілетін стационарлы емес уақыттық қатарларды сипаттауға мүмкіндік береді. Нәтижесінде ARIMA(0,2,1) моделі болжау интервалы 95% барабар болжамды көрсетті.

ЖИ негізіндегі әдістердің ішінен сызықты емес жүйелерді модельдеуге өте қолайлы NARX (Nonlinear AutoRegressive exogenous Network) моделі және анық емес логикалық алгоритмдер (Fuzzy Logic) қолданылды.

NARX моделін бағалау нәтижесінде орташа квадраттық қатенің (MSE) нәтижесі 1,0627 тең болды.

Қорытындыда диссертацияның негізгі нәтижелері мен қорытындылары берілген.

Жарияланымдар. Диссертация тақырыбы бойынша негізгі нәтижелер 15 мақалада: Қазақстан Республикасы Білім және ғылым министрлігі, Ғылым және жоғары білім сапасын қамтамасыз ету комитеті ұсынған журналдарда - 7 мақала, Халықаралық Scopus деректер базасында – 1 мақала, Халықаралық конференцияларда - 7 мақала жарық көрді.

Жұмыстың нәтижесі бойынша 2 авторлық куәлік алынды:

1. Разработка модели прогнозирования с использованием статистического метода Auto-Regressive Integrated Moving Average. № 32481 от «9» февраля 2023 года.
2. Нечеткая модель прогнозирования временного ряда. № 35224 от «27» апреля 2023 года

1 МУЛЬТИСЕРВИСТІ ЖЕЛІДЕГІ ТРАФИКТІ ШОЛУ ЖӘНЕ МІНДЕТТЕРДІ ТҰЖЫРЫМДАУ

1.1 Қазақстандағы телекоммуникациялар желісінің негізгі ерекшеліктері

Қазіргі уақытта Қазақстанның барлық әкімшілік орталықтарында NGN конвергентті мультисервистік желісі жұмыс істеп тұр. Дүниежүзілік экономикалық форумның 2022ж. Қорытындысы негізінде Қазақстан Республикасы желіге дайындық индексі бойынша 130 елдің арасында 58-ші орынды иеленді [1].

Бүгінгі күні ақпараттық технологиялар саласындағы бизнес мультисервистік желілердің ауқымдылығына жоғары талаптар қояды. Мультисервистік желілердің даму тенденциялары:

- трафик көлемінің жылдам өсуі және оның құрылымының бейне тарату және бірыңғай байланыс UC (Unified communications) бағытында өзгеруі;
- мобильді пайдаланушыларға BYOD (Bring Your Own Device) және әлеуметтік желілерді қолдау қажеттілігі;
- Үлкен деректерді өңдеуге арналған жоғары өнімді кластерлер (BIG DATA);
- бұлттық қызметтерді ұсыну үшін виртуалдандыру (Cloud Bursting).

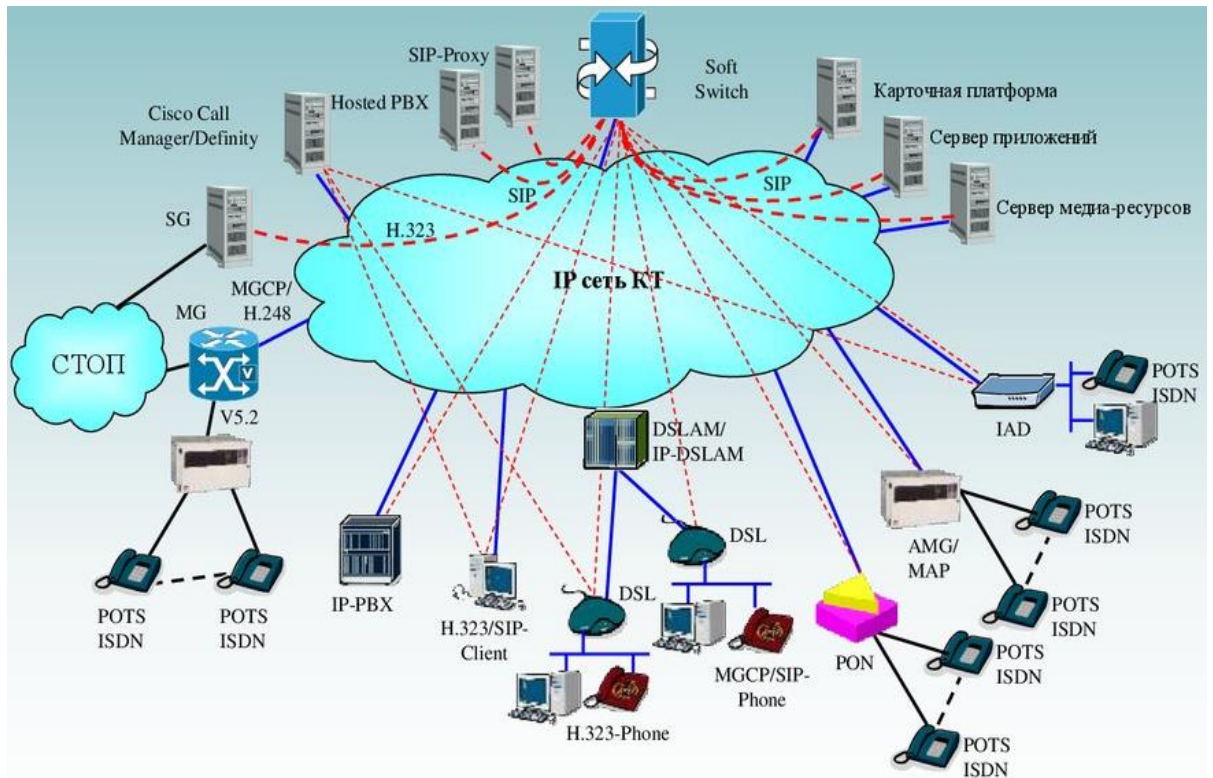
Мультисервистік желіде ұсынылатын инфокоммуникациялық қызметтер дәстүрлі байланыс желілерінің қызметтерінен ерекшеленеді:

- инфокоммуникациялық қызметтердің негізгі бөлігі клиент/сервер архитектурасы негізінде жүзеге асырылады. Бұл ретте қосымшаның клиенттік бөлігі пайдаланушы деректерінің терминалдық жабдығына жүктеледі, ал сервер бөлігі сәйкес бөлінген желілік қызмет түйіндеріне жүктеледі;
- әртүрлі қызметтерді дамыту ISO моделінің жоғарғы деңгейлерінде жүзеге асырылады;
- инфокоммуникациялық қызметтер өткізу қабілеттілігінің кеңдігімен сипатталатын мультимедиялық ақпаратты беруді жүзеге асырады.

Мультисервистік желі көп деңгейлі иерархиялық архитектураға ие (трафиктің жоғары жылдамдықты ауысуы бар ядро, трафикті біріктіру функциясын орындауға арналған тарату деңгейі және оператор желісіне қол жеткізу, сегменттеу функциялары, пайдаланушыларды логикалық топтарға біріктіру, аутентификация және т.б.)

Көрсетілетін қызмет түрлерінің ұлғаюымен осы қызметтерді көрсету сапасына қойылатын талаптар да өсуде. Сонымен қатар, мұндай ағындардың қарқындылығының жоғарылауымен өңдеу әдісін үнемі жетілдіріп отыру керек.

Мультисервисті желінің маңызды ерекшеліктерінің бірі байланыс желісінде қабылданатын хабарламалар құрылымының біркелкі еместігі болып табылады. Мультисервисті желі өнімділігі мен жылдамдығы жоғары IP коммутаторларын пайдаланады. 1.1 суретте қолданыстағы мультисервистік желі және қатынау жүйелерінің пакеттік желіге бейімделуі, NGN желісінің жалпы архитектурасы көрсетілген.



Сурет 1.1 - Қазақстан Республикасындағы NGN желісінің жалпы архитектурасы

1.2 Желілік трафикті болжау бойынша негізгі ғылыми еңбектерге шолу

Болжау мәселесін шешу өзекті болып табылады, өйткені болжау деректері желіні бақылау мәселесін шешуге қажетті ақпаратты береді.

[2] жұмыста кез келген уақыт аралығы үшін жоғары дәлдікпен желілік трафиктің өзіндік ұқсастық қасиеттерін ескере отырып болжау алгоритмдері әзірленді. [3] жұмыста қысқа мерзімді уақыттық қатарларды болжау және кептелісті анықтау үшін желілік трафикті модельдеу сызықты емес динамика әдістерін пайдалана отырып сипатталған. [4] жұмыста әзірленген трафикті болжау әдісі қолданудың қарапайымдылығымен және желі өлшеміне шектеу қажет етпейтін әмбебап әдіс болып табылады. Циклдік талдау негізінде трафикті болжаудың математикалық моделін жасаумен аномалияны анықтауға негізделген басқару әдісі [5] жұмыста қарастырылған. [6]-да әлеуметтік-экономикалық параметрлерді ескере отырып, аймақтық ақпараттық-коммуникациялық жүйелердің сипаттамаларын болжауға арналған тапсырмаларды автоматтандырылған орнату технологиясы әзірленді. Ұялы байланыста ресімделуі қиын процестер үшін ұқсас объектілер туралы ақпаратты пайдалана отырып болжау әдісі [7] жұмыста ең жақын аналогтарды іздеу алгоритмін әзірлеумен сипатталған, ал [8] жұмыста процестің логикалық-математикалық модельдері болмаған жағдайда қолданылатын спектрлік компоненттердің динамикасын талдауға негізделген объектілердің техникалық шарттарын болжаудың жаңа әдісі әзірленген. Нейрондық желіні болжау [9, 10]-

жұмыстарда оқыту және тест үлгілері бойынша болжау қателерін есептеу арқылы нейрондық желілердің әртүрлі конфигурациялары салынды. [11-13]-жұмыстарда болжау үшін экзогендік кірісі бар сызықты емес авторегрессивті қайталанатын нейрондық желі (NARX) пайдаланылды. Алынған нәтижелер әзірленген модельдердің күн радиациясын болжауға және оның нақты бағалаулар мен болжамдарды беруге қабілеттілігін көрсетті. Уақыттық қатарларды талдау негізінде компьютерлік желілердегі кептелісті болжау әдісі [14] жұмыста қарастырылды. Корпоративтік желінің техникалық жай-күйін болжау [15] жұмыста орындалды, ал [16] жұмыста айнымалы құрылымы бар уақыттық қатарларға негізделген болжау әдістері қарастырылған. [17] – де сипатталған жұмыс максималды ықтималдық үлгісіне негізделген уақыттық қатарларды болжау үлгісін қарастырады. [18, 19, 20] жұмыстарда авторегрессивті модельдер арқылы қозғалыс қарқындылығын болжау мәселесі қарастырылады. Интерактивті қызметтерді алдын ала сәйкестендіру және түрін анықтау арқылы телекоммуникация желісіндегі ақпарат ағынының әрекетін болжау әдісі болжау дәлдігін арттыру мақсатында әзірленді [21]. Жіберілетін трафик құрылымын өзгерту компьютерлік желі дамуының заманауи тенденциясы болып табылады және бұл трафик пакеттердің кешігу уақытын, буферлік жадыны анықтау дәлдігіне қойылатын талаптардың жоғарылауына байланысты мультимедиялық болды деп атап өтілген [22].

[23] жұмыстағы зерттеу кездейсоқ процестер теориясын уақыт қатарларын болжаудағы мәселені шешу үшін салмақты модельдерді ұсынады.

Көптеген зерттеулер көрсеткендей, желідегі қызмет көрсету стратегиялары классикалық әдістерден «болжау және алдын алу» әдістемесіне көшудің қажеттігін анықтайды. Жағдайды автоматты түрде бақылау және болжау әдістерін құру және енгізу экономикалық шығындарды азайтып, желіге қажетті компоненттерді уақтылы ауыстыруға және ақпарат ағынының тоқтап қалуына жол бермеу мәселелерін шешеді.

Жоғарыда шолу жасалған жұмыстар нәтижесі жалпы деректер жиындары үшін жақсы жұмыс істейді. Өкінішке орай, нақты деректерде олардың бірқатары өнімділік немесе тиімділік тұрғысынан жоғары мәндерге қол жеткізе алмайды. Жалпыға қолжетімді деректер жиынының көпшілігінің бірнеше кемшіліктері бар: шынайы емес аномалия, біркелкі емес және жетіспейтін деректердің болуы. Осы кемшіліктерге байланысты нақты деректер негізіндегі уақыттық қатардың құрылымын анықтау бойынша зерттеулер жасап, нәтижесінде таңдалатын модельдер анықталуы қажет.

Уақыттық қатар – бұл көптеген құрылғыларда қабылданған деректер тізбегі, олар уақыт бойынша ретті тәртіпте сақталады. Пайдалы статистикалық мәліметтер (тренд, аномалия және құбылмалылық) уақыттық қатарлардың деректерінен алынады. Оларға негізделіп құрылған модельдер болжау кезінде дұрыс шешім қабылдауға септігін тигізеді.

1.3 Уақыттық қатарларды талдаудағы негізгі түсініктер

Уақыттық қатар – тұрақты уақыт аралықтарында өлшенген айнымалының мүмкін мәндерінен (жол деңгейі, пакеттер саны) тұратын қатар.

Жалпы жағдайда уақыттық қатарлар кездейсоқ шамалардың бір параметрлі тобы болып табылады, олардың таралу заңы және сандық сипаттамалары t уақытына байланысты болуы мүмкін. Көрсеткіштің сандық мәндері уақыттық қатардың деңгейлері деп аталады.

Уақыттық қатарлар интервалды және сәттік (моментный) болуы мүмкін.

Интервалдық қатар белгілі бір уақыт аралығында айнымалының жинақталуы арқылы қалыптасады.

Сәттік қатар белгілі бір уақыт нүктелеріндегі құбылыстың өлшемін сипаттайды.

Кездейсоқ процестің сипаттамаларына мыналар жатады:

- математикалық күту (математическое ожидание) – кездейсоқ шаманың орташа мәнін сипаттайтын ықтималдықтар теориясының тұжырымдамасы. Кездейсоқ шаманың мәні әр уақытта әртүрлі болуы мүмкін эксперименттер қатарының нәтижесінде қандай орташа мәнді күтуге болатынын көрсетеді.

- дисперсия – кездейсоқ шама мәндерінің оның математикалық күтуіне қатысты таралуының көрсеткіші. Дисперсия кездейсоқ шаманың мәндері оның ортасынан қаншалықты алыс болуы мүмкін екенін өлшейді.

- коварианттық функция – екі кездейсоқ шама арасындағы уақыт немесе кеңістік бойынша тәуелділік дәрежесін өлшейтін функция. Ол екі кездейсоқ шаманың бірге қалай өзгеретінін сипаттау үшін қолданылады.

Уақыттық қатарларды зерттеу кезінде туындайтын міндеттер:

- құбылыстың кезеңнен кезеңге дейінгі даму қарқындылығын сипаттау;
- белгілі бір кезеңдегі уақыттық қатардың орташа көрсеткіштерін анықтау;
- жеке кезеңдердегі және жалпы қарастырылатын кезеңдегі зерттелетін құбылыстың динамикасының негізгі заңдылықтарын анықтау;

- зерттелетін объектінің уақыт бойынша өзгеруін бақылайтын факторларды анықтау;

- болашақта құбылыстың дамуын болжау.

Модельдеу кезінде уақыттық қатар кездейсоқ процесс (стохастикалық), ықтималдықтар теориясының заңдарына сәйкес уақыт бойынша дамиды статистикалық құбылыс ретінде қарастырылады. Уақыттық қатарлар детерминделген және кездейсоқ процестерге бөлінеді. Детерминделген процесс бірге тең ықтималдығы бар берілген мәнді қабылдайды, ал кездейсоқ процесс мәні бойынша кездейсоқ болып табылады.

Стохастикалық процестер стационарлық/стационарлық емес процестер болып бөлінеді. Стационарлық және стационарлық емес уақыттық қатарлар қатар деңгейлерін құрайтын факторлардың болуы/болмауымен ерекшеленеді [24, 25]. Уақыттық қатарларды зерттегенде олардың стационарлық немесе стационарлық еместігін білу маңызды, өйткені олар әртүрлі статистикалық сипаттамаларға ие және сәйкесінше әртүрлі тәсілдермен бағаланады [26].

Уақыттық қатар стационарлы болуы үшін келесі қағидалар орындалуы тиіс:

1. күтудің тұрақтылығы

$$E(y_1) = E(y_2) = E(y_3) = \dots \quad (1.1)$$

(яғни қатар ұлғаюға немесе азаюға бейім емес, бірақ кездейсоқ жоғары немесе кездейсоқ төмендей алады, бірақ орташа алғанда сол деңгейде қалады деген болжам):

$$E(y_t) = \mu = \text{const} \quad (1.2)$$

Бұл формула әрбір t уақытындағы y -тің орташа мәні (күтілетін мәні) μ тұрақты мәніне тең болатын уақыттық қатар моделін сипаттайды. Мұндай модель тұрақты деңгейлі модель (constant level model) деп аталады.

2. дисперсияның тұрақтылығы

$$\text{Var}(y_1) = \text{Var}(y_2) = \text{Var}(y_3) = \dots = \gamma_0 \quad (1.3)$$

(қатардың шашырауының (разброс ряда) уақыт бойынша ұлғаятын фактілерінің жоқтығы, яғни қатардың уақыт бойынша ауытқуы, мысалы, бір жыл бұрынғы мен бүгінгі күнгі ауытқулары шамамен бірдей деген болжам):

$$V(y_t) = \gamma_0 \quad (1.4)$$

Формула әрбір t уақытындағы y дисперсиясы γ_0 тұрақтысына тең екенін білдіреді. Бұл уақыттық қатар мәндері уақыт бойынша дисперсиясын өзгертпейтінін және уақыттың әрбір нүктесінде бірдей таралатынын білдіреді. Осы шартты қанағаттандыратын модель тұрақты дисперсия моделі немесе гомоскедастық модель деп аталады.

3. ковариация (сызықтық байланыс күші) яғни бір тәуліктік

$$\text{Cov}(y_1, y_2) = \text{Cov}(y_2, y_3) = \text{Cov}(y_3, y_4) = \dots = \gamma_1 \quad (1.5)$$

және (2 не одан да көп)

$$\text{Cov}(y_1, y_3) = \text{Cov}(y_2, y_4) = \text{Cov}(y_3, y_5) = \dots = \gamma_2 \quad (1.6)$$

арақашықтықтағы уақыт көрсеткіштері бірдей байланыс дәрежесіне ие.

k -ші ретті автоковариация коэффициенті:

$$Cov(y_t, y_{t-k}) = \gamma_k, \quad (1.7)$$

γ_k - y_t және y_{t-k} процестерінің арасындағы автоковариациялық функция. μ, γ_0, γ_k параметрлері уақыттан тәуелсіз.

1.8 формула екі уақыттық қатардың мәндері, y_t және y_{t-k} арасындағы ковариация функциясын анықтайды, мұндағы k — мәндер арасындағы уақыт кідірісін білдіретін бүтін сан.

Ковариация екі кездейсоқ процестің өзара қалай байланысатынын көрсетеді: егер олар бір бағытта қозғалса, онда ковариация оң болады, егер әртүрлі бағытта болса, онда теріс болады.

Коварианттық функция әртүрлі уақыт аралықтарында уақыттық қатарлардың мәндері арасындағы байланысты анықтайды және уақыттық қатарларды талдау және модельдеу үшін пайдаланылуы мүмкін [27].

k -ші ретті автокорреляция коэффициенті:

$$\rho_k = Corr(y_t, y_{t-k}) = \frac{Cov(y_t, y_{t-k})}{\sqrt{V(y_t) \cdot V(y_{t-k})}} = \frac{\gamma_k}{\sqrt{\gamma_0 \cdot \gamma_0}} = \frac{\gamma_k}{\gamma_0}, \quad (1.8)$$

ρ_k - автокорреляция функциясы (АКФ, АСФ).

Стационарлы қатарлардың келесі ерекшеліктері бар:

- ұзақ мерзімді кезеңде қатардың деңгейлері тұрақты орташа мәннің айналасында ауытқиды;
- уақыттық қатардың дисперсиясы уақытқа тәуелді емес;
- АСФ төмендеуі, яғни кешігу (лаг) ұзақтығы артқан сайын АСФ мәні төмендейді.

Уақыттық қатарлар стационарлы болса, қатарда:

- тренд;
- маусымдық компонент;
- интервенциялар;
- гетероскедастық;
- уақыттық қатардың құрамдас бөліктері арасындағы кез келген жүйелі түрде өзгертін өзара тәуелділік болмайды.

Стационарлық процестің ең қарапайым мысалы – ақ шу.

$$(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{t-1}, \varepsilon_t, \dots) \quad (1.9)$$

$$E(\varepsilon_t) = 0 \quad (1.10)$$

$$Var(\varepsilon_1) = \sigma^2 \quad (1.11)$$

$$Cov(\varepsilon_t, \varepsilon_{t-k}) = 0 \quad (1.12)$$

Ақ шу – барлық жиіліктерді бірдей мөлшерде қамтитын тұрақты спектрлік тығыздығы бар корреляцияланбаған кездейсоқ шамалардың тізбегі.

Егер уақыттық қатардың сипаттамалары: орташа мәні, дисперсиясы, АКФ уақытқа тәуелді болса, бұл қатар *стационарлы емес* деп аталады. Егер уақыттық қатардың орташа мәні уақытқа сызықты тәуелді болса, онда сызықтық тренд барын, егер уақыттық қатардың дисперсиясы уақыт бойынша өзгерсе, АКФ циклдік өзгерсе, онда бұл қатарда мерзімді құрамдас бөліктер бар екенін көрсетеді.

Стационарлы емес қатарлардың келесі ерекшеліктері бар:

- ұзақ мерзімді кезеңде қатар деңгейлері әртүрлі орташа мәндер төңірегінде топтасады;
- уақыттық қатардың дисперсиясы период бойынша өзгереді, яғни уақытқа байланысты;
- автокорреляция функциясы өте баяу төмендейді.

Тұрақтылықты анықтаудың келесі жолдары бар:

- график түрінде уақыттық қатарды ұсыну және трендті визуалды тексеру;
- автокорреляцияның бар-жоғын зерттеу;
- детерминделген трендтің болу тесттерін қолдану;
- стохастикалық трендтің болуына тесттерді қолдану (мысалы, бірлік түбірге арналған тесттер).

Детерминделген тренд – функционалдық форма немесе математикалық теңдеу арқылы сипатталуы мүмкін болатын уақыт бойынша айнымалының өзгеру формасы. Трендтің бұл түрі кездейсоқ немесе стохастикалық трендтен ерекшеленеді, оны нақты болжау мүмкін емес, өйткені оның нақты функционалдық формасы жоқ және кездейсоқ құрамдас болуы мүмкін.

Детерминделген тренд уақыт бойынша айнымалы мәnnің тұрақты өзгеруін көрсететін сызықтық немесе сызықтық емес функция ретінде ұсынылуы мүмкін. Мысалы, сызықтық детерминделген тренд $y(t) = a + bt$ теңдеуімен ұсынылуы мүмкін, мұндағы $y(t) - t$ уақытындағы айнымалының мәні, a – айнымалының бастапқы мәні, ал b – айнымалы мәnnің уақыт бойынша өзгертін жылдамдығы.

Стохастикалық тренд – кездейсоқ құрамдас бөлігінің болуымен сипатталатын уақыт бойынша айнымалының өзгеру формасы. Детерминделген трендтен айырмашылығы, стохастикалық трендті математикалық функция немесе теңдеу арқылы сипаттауға болмайды, өйткені оның белгілі бір формасы немесе өзгеру заңы жоқ.

Бірлік түбір – уақыттық қатарларды сипаттау үшін қолданылатын ұғым [28]. Бұл келесі формуламен берілген бірлік теңдеудің түбірі:

$$y(t) = a*y(t-1) + e(t) \quad (1.13)$$

мұндағы $y(t) - t$ уақытындағы қатардың мәні, $a - y(t)$ мәні бұрынғы $y(t-1)$ мәніне қаншалықты тәуелді екенін анықтайтын коэффициент, $e(t) -$ кездейсоқ шама, қате, бұл t уақытындағы уақыт қатарының мәніне әсер ететін кездейсоқ фактор.

Егер a коэффициенті 1 болса, онда теңдеу келесідей болады:

$$y(t) = y(t-1) + e(t) \quad (1.14)$$

Мұндай теңдеу бірлік түбірі бар уақыттық қатарды сипаттайды, бұл қатардың стационарлы емес екенін және оның орташа мәні уақыт бойынша тұрақты емес екенін білдіреді. Бұл, мысалы, қатарда уақыт өте келе оның орташа мәнінің тұрақты өзгеруіне әкелетін тренд болған жағдайда орын алуы мүмкін.

Бірлік түбір уақыттық қатарларды талдауда маңызды ұғым болып табылады, өйткені ол стационарлық емес қатарлар үшін модельдеу және болжау әдістерінің қажеттілігін көрсетуі мүмкін.

Статистикалық тестілеу уақыттық қатарларды талдаудың маңызды қадамы болып табылатынын атап өту маңызды, өйткені көптеген статистикалық әдістер мен модельдер уақыттық қатарлардың стационарлық болуын талап етеді. Егер қатар стационарлы емес болса, стационарлық қатарды алу үшін сәйкес деректерді түрлендіруді, мысалы, дифференциалдау немесе маусымдық түзетуді жүргізу қажет [29].

1.4 Уақыттық қатарлардың модельдері

Мультисервистік желінің процестерін талдаудың жалпы үлгісі уақыттық қатарларға негізделген модель болып табылады. Уақыттық қатарларды талдаудың негізгі мақсаты:

- қатардың мінез-құлқындағы ішкі заңдылықтарды анықтау;
- болжау.

Уақыттық қатардың жалпы статистикалық моделі:

$$y_t = f(x_t, a) + \varepsilon_t. \quad (1.15)$$

$y_t - f(x_t, a)$ компонентінің қосындысы ретінде қарастырылатын бақылаулар қатары; a - параметр. $f(x_t, a)$ компоненті сызықты және келесілерге ыдыраған болуы мүмкін: тренд, кезеңділік, маусымдық, және p кешігу параметріне жылжитын авторегрессивті құрам. Жалпы уақыттық қатарларды статистикалық талдаудың өзі – детерминделген компоненттерді моделдеу және қалдықтарды стационарлық қалыпқа келтіру. D_t детерминделген компонентіне әдетте тренд, маусымдық және циклдік компоненттерді жатқызады.

Уақыттық қатардың тренді – бұл әсері бірте-бірте әсер ететін ұзақ мерзімді факторлардың таза әсерін сипаттайтын бірқалыпты өзгеретін, циклдік емес компонент $T(t)$ функциясы арқылы сипатталады.

Маусымдық компонент талданатын белгідегі өте ұзақ емес кезеңдегі ауытқуларды білдіреді. Маусымдық кезең маусымдық ұзақтығына еселік болатын $S(t)$ функциясымен сипатталады.

Циклдік компонент ұзақ мерзімді циклдердің әсерінен талданатын белгінің өзгеруін білдіреді. Циклдік компонент $C(t)$ функциясымен сипатталады. Бұл өзгерістер трендке де, маусымдық мерзімді компонентке де жатпайтындай анық

байқалу керек. Маусымдық және циклдік компоненттер бірігіп, периодтық компонент деп аталады, өйткені оларды тек тербеліс периодының шамасы ғана ажыратады.

Кездейсоқ компонент – зерттелетін белгінің стохастикалық сипатын білдіретін қысқа мерзімді және болжауға келмейтін факторлар әсерінің нәтижесі.

Әртүрлі жағдайларда уақыттық қатардағы аталған компоненттердің әрқайсысы жұмыс істей алады, бірақ кездейсоқ компоненттің болуы міндетті болып саналады.

Кез келген модель y_t мәндеріндегі жеке компоненттердің үлесін бағалауды қамтиды. Бұл процедура уақыттық қатар деңгейлерінің ыдырауы деп аталады. y_t құрамындағы әрбір компоненттің үлесіне қарай оларды бағалауға және сәйкесінше y_t болашақ мәндерін болжауға болады.

Стационарлық уақыттық қатарлардың негізгі сызықтық үлгілеріне мыналар жатады:

- авторегрессиялық модельдер;
- жылжымалы орташа модельдер;
- авторегрессивті және жылжымалы орташа модельдер.

Қалдықтар қатарының стационарлығын тексеру және оның дисперсиясын бағалау үшін:

- таңдамалы автокорреляция функциясы (коррелограмма);
- жеке автокорреляция функциялары пайдаланылады.

Егер уақыттық қатар стационарлы болмаса, онда бұл уақыт қатарында не сызықтық тренд, не қалдықтардың гетероскедастикасы, не қалдықтардың автокорреляциясы, не периодтық компонент немесе барлығы да бар екенін білдіреді. Стационарлы емес уақыттық қатарлардың үлгілері:

- уақыттық қатарлардың периодтық компоненттерінің моделі;
- уақыттық қатарлардың гетероскедастикалық моделі;
- қалдықтардың автокорреляция моделі.

Тренд анық көрінбесе, процестің АКФ және жеке автокорреляция (ЧАКФ) функцияларын қарастыру керек.

Уақыт өте өзгертін стационарлық емес процесс келесі процестермен көрсетіледі:

- детерминделген тренды бар процесс (TS (trend stationary));
- стохастикалық тренді бар қатарлар DS (difference stationary). Бұл кездейсоқ жүру процесі (случайное блуждание - random walk).

Стационарлық емес кездейсоқ процесс (random walk process) келесі формуламен анықталады:

$$X_t = \mu + X_{t-1} + \varepsilon_t \quad (1.16)$$

Кездейсоқ жүру – стационарлы емес процесс, бірақ егер оған реттік айырманы алу операторы қолданылса, онда стационарлы қатар пайда болады.

Осы тұста қатарлардың екі түрі де кездейсоқ бөлікпен ерекшеленетінін атап өткен жөн, TS-қатарда (ағымдық шок, бұзылулар) уақыт өткен сайын

бұрынғы соққылар ыдырауының әсері әлсірейді, ал DS сериясында (барлық алдыңғы соққылардан жинақталған бұзылулар) мұндай әлсіреу болмайды және әрбір жеке соққы қатардың барлық келесі мәндеріне бірдей күшпен әсер етеді.

Ұзақ мерзімді регрессиялық модельдерді дұрыс құру үшін қарастырылатын қатардың TS/DS кластарының қайсысына жататынын шешу қажет.

Кестеде TS және DS модельдерінің ерекшеліктері сипатталған.

Кесте 1 - TS және DS модельдерінің ерекшеліктері

TS процесі	DS процесі
Тұрақты емес трендке байланысты стационарлық емес	Тұрақты емес дисперсияға байланысты стационарлық емес
Соққылар туралы жады шекті, яғни ол алдыңғы кадамдағы қатені бірден ұмытады. Егер ақ шудың орнына жалпылама процесс ARMA(p,q) қолданылса, онда соққылар біраз уақыт әсер етеді, бірақ уақыт өте келе олардың әсері әлсірейді.	Барлық алдыңғы ε_t қосындысы болған соң, соққылар үнемі есте қалады. Соққылар туралы жады шексіз.

1.5 Уақыттық қатарды стационарлыққа тексеру

Уақыттық қатарды стационарлыққа тексеру үшін математикалық модельді құрудың негізгі қадамдары мыналар екендігі [30]-да сипатталған:

- қатардың стационарлығын тексеру;
- сынақ моделін анықтау;
- модель параметрлерін бағалау;
- модельдің дұрыстығын тексеру.

Тәжірибеде қатарлардың стационарлық гипотезасын тексеру үшін жоғары ықтималдықпен ақиқат болжамды қабылдауды және жалған гипотезаны жоққа шығаруды қамтамасыз ететін критерийлер (тесттер) қолданылады [31]. Олар параметрлік және параметрлік емес болып бөлінеді [32]:

- параметрлік тесттер есептеу формуласына таралу параметрлерін (орташа және дисперсия) қосады, яғни талданатын барлық сипаттамалар қалыпты түрде таралғанда қолданылады;

- параметрлік емес тесттер таралу параметрлерін есептеу формуласына қоспайды және жиіліктер мен рангтарға негізделеді.

Төменде сипатталған әдістер қатардың стационарлығын анықтау үшін қолданылады [33]:

- графикті визуалды зерттеу (тренд пен периодтық құрамдастың болуы, дисперсияның тұрақтылығы, т.б.);
- детерминделген трендтің барын тексеретін тесттер:
- статистикалық сипаттамалардың тұрақтылығына арналған тесттер;
- стохастикалық трендтің барын тексеретін тесттер (бірлік түбірге арналған тесттер).

Уақыттық қатар алдымен деректердің таралуының қалыптылығына тексеріледі, яғни қалыпты таралу жағдайында параметрлік тест, әйтпесе параметрлік емес тест қолданылады.

Уақыттық қатардың стационарлығын келесі параметрлік тесттер арқылы тексеруге болады:

- Стьюдент критерийі бойынша математикалық күтудің тұрақтылығы туралы;
 - Фишер үлестірімі бойынша дисперсияның тұрақтылығы туралы.
- Параметрлік емес критерийлер:
- графикалық әдіс;
 - математикалық күтудің тұрақтылығын тексеру үшін Манн-Уитни тесті;
 - математикалық күтудің тұрақтылығын тексеру үшін Кокс-Стюарт трендінің белгі критерийі;
 - өсу және кему қатарлары критерийі;
 - математикалық күтудің тұрақтылығын тексеру үшін Фостер-Стюарт критерийі;
 - дисперсияның тұрақтылығын тексеру үшін Фостер-Стюарт критерийі;
 - автокоррекциялық талдау;
 - TS және DS стационарлық емес қатарларының түрлерін ажырату үшін бірлік түбір сынағы (Дикки-Фуллер), т.б..

1.6 Уақыттық қатарларды болжау әдістері

Болжау – құбылыстың немесе процестің дамуының жай-күйі мен ықтимал жолдарын ғылыми негізделген, қалыптасқан себеп-салдарлық байланыстар мен заңдылықтар жүйесіне негізделіп отырып анықтау [34].

Уақыттық қатарды болжау кезінде қатардың өткен мәндері оның болашақтағы әрекеті туралы ақпаратты қамтитыны ескеріледі. Болашақ өткенге қаншалықты тәуелді болса, соғұрлым жақсы болжам жасауға болады – бұл негізгі қасиет.

Болжау әдістері сан алуан:

- эвристикалық болжау әдістері;
- уақытша экстраполяция әдістері;
- кеңістіктік экстраполяция;
- болжамды экстраполяция әдістері;
- экспоненциалды тегістеу модельдері;
- максималды ұқсастық үлгісіне негізделген модель;
- Марков тізбектеріндегі модель;
- жасанды интеллекттің логикалық әдістері;
- регрессия әдістері;
- авторегрессивті болжау модельдері.

Қазіргі кезде кең таралған әдістер: авторегрессивті болжау модельдерімен қиылысатын уақытша экстраполяциялық әдістер, экспоненциалды тегістеу модельдері, нейрондық желілік болжау әдістері және т.б..

Сызықты емес модельдердің коэффициенттерін бағалау үшін мыналар қолданылады [35]:

- модельді сызықтандыру (бастапқы айнымалылардың түрлендірулер);
- коэффициенттердің бағалауын табу үшін сызықты емес оңтайландырудың сандық әдістері қолданылады.

[36]- жұмыста уақыттық қатарларды болжамға дейін алдын ала талдау кезеңдері сипатталған:

- бастапқы мәліметтердің графигін құру;
- негізгі статистикалық көрсеткіштерді (орташа мән, дисперсия, және т.б.) есептеу және талдау;
- деңгейлерді бір-бірімен салыстыру нәтижесінде алынған статистикалық көрсеткіштердің көмегімен құбылыстың жылдамдығы мен қарқындылығын уақыт бойынша талдау.

Алдын ала зерттеулерден кейін келесі қадамдар:

- уақыттық қатардың әрекетін графикалық бейнелеу және сипаттау;
- уақытқа байланысты уақыттық қатардың тұрақты компоненттерін таңдау және жою (тенденция, маусымдық, циклдік компоненттер);
- процестің төменгі немесе жоғары жиілікті компоненттерін таңдау және жою (сүзу);
- жоғарыда аталған компоненттерді алып тастағаннан кейін қалған уақыттық қатардың кездейсоқ құрамдас бөлігін зерттеу;
- кездейсоқ компонентті сипаттаудың математикалық моделін құру және оның сәйкестігін тексеру;
- уақыттық қатармен ұсынылған процестің болашақ дамуын болжау;
- әртүрлі уақыттық қатарлардың өзара әрекеттесуін зерттеу.

Жоғарыда аталған қадамдарды шешу үшін уақыттық қатарларды талдаудың келесі әдістері қолданылады:

- бір процесс ішіндегі корреляциялық талдау (автокорреляция);
- уақыттық қатардың периодтық және квазипериодтық компоненттерін табуға мүмкіндік беретін спектрлік талдау;
- жоғары жиілікті немесе маусымдық ауытқуларды жою үшін уақыттық қатарларды тегістеу және сүзу;
- орташа мәннің айналасында біркелкі ауытқуларды көрсететін процестерді сипаттау және болжау үшін пайдалы авторегрессивті және жылжымалы орташа модельдер;
- уақыттық қатарлардың мінез-құлқының таңдалған моделіне сүйене отырып, болашақта оның мәндерін болжауға мүмкіндік беретін экстраполяция және болжау.

Уақыттық қатарларды талдаудың міндеті – детерминделген компонентті пайдалана отырып, уақыттық қатардың болжамдық мәнін болжау және кездейсоқ құраушыны пайдалана отырып, мүмкін болатын ауытқудың шамасын және мұндай ауытқудың ықтималдығын болжау.

Болжау әдістері – классикалық декомпозиция, эконометриялық модельдер, көп өлшемді регрессия.

1.7 Негізгі міндеттерді тұжырымдау

Әдебиеттерді шолу, желіні талдау зерттеу мақсатын, осы диссертациялық жұмыс аясында шешілетін негізгі міндеттерді тұжырымдауға мүмкіндік берді:

- өлшенген мультисервистік трафик құрылымын стационарлыққа зерттеу;
- өлшенген трафикті болжау.

Уақыттық қатарлар стационарлы және стационарлы емес болуы мүмкін. Тәжірибеде қатардың стационарлық гипотезасын тексеру үшін параметрлік және параметрлік емес критерийлер қолданылады. Стационарлық және стационарлық емес уақыт қатарларының статистикалық қасиеттері айтарлықтай ерекшеленеді және оларды модельдеу үшін әртүрлі әдістерді қолдану қажет.

Демек қатарды стационарлыққа тексеру міндеті, стационарлы болмаған жағдайда стационарлық жағдайға келтіруді жүзеге асыру үшін трендті таңдау, маусымдық компонент немесе интеграция дәрежесін анықтау қажеттігі туады.

Болжау міндеті – T болжау сәтінде барлық белгілі ақпаратқа тәуелді болатын функцияны f_T табу қажет, бұл функция және каншалықты ілгері болжау қажеттігін көрсететін қосымша N параметрін қабылдай алады:

$$y_{T+h} \approx f_T(y_T, \dots, y_1, h) \equiv \hat{y}_{T+h|T}, \quad (1.17)$$

мұндағы N – болжау көкжиегі, $h \in \{1, 2, \dots, N\}$, y_{T+h} – $T+h$ сәтіндегі қатардың мәні, $\hat{y}_{T+h|T}$ – $T+h$ сәтіндегі болжам мәні.

Осылайша, h болжам көкжиегі мен T уақытына дейінгі қолжетімді ақпаратты ескере отырып, y тарихи мәндерін және f_T болжау функциясын пайдаланып $T+h$ уақытындағы y -тің болашақ мәнін болжау қажет.

1.8 Бірінші тарау бойынша қорытынды

1. Деректерді талдаудың классикалық мәселелерінде бақылаулар тәуелсіз, ал болжау кезінде қатардың алдыңғы мәндері болашақ мәндер туралы ақпаратты береді деп күтіледі.

Кейінгі әсері жоқ ақпарат ағынының (поток без последействия) тарихы болашақты болжауда ешқандай рөл ойнамайды.

2. Уақыттық қатарлар стационарлы және стационарлы емес болуы мүмкін. Тәжірибеде қатардың стационарлық гипотезасын тексеру үшін параметрлік және параметрлік емес критерийлер қолданылады.

3. Параметрлік критерийлер талданатын деректердің қалыпты таралуын талап етеді.

4. Параметрлік емес критерийлер есептеу формуласында таралу параметрлерін талап етпейді және операциялық жиіліктерге немесе рангтарға негізделеді.

5. Стационарлық және стационарлық емес уақыт қатарларының статистикалық қасиеттері айтарлықтай ерекшеленеді және оларды модельдеу үшін әртүрлі әдістерді қолдану қажет.

6. Стационарлық уақыт қатарларының моделі уақыттың тұрақты орташа мәнімен, дисперсиясымен және автокорреляциясымен сипатталады.

7. Стационарлық емес уақыт қатарларының моделі көпкомпонентті: трендті, маусымдық және кездейсоқ құрамдастарды қамтиды.

8. Стационарлық емес қатарлар TS детерминирленген трендті және DS стохастикалық тренді бар қатарларға бөлінеді, бұл мұндай қатарлармен жұмыс істеу кезінде стандартты асимптотикалық теорияның қолданылмайтындығын көрсетеді.

9. TS процесі тұрақты емес тенденцияға байланысты стационарлық емес. Бұл процесс шексіз жадымен емес.

10. DS процесі тұрақты емес дисперсияға байланысты стационарлық емес. Бұл процесс шексіз жадымен.

11. Стационарлық жағдайға келтіру мынадай операциялармен жүзеге асырылады: трендті таңдау, маусымдық компонент немесе интеграция дәрежесі.

12. Дикки-Фуллер тесті қатарды DS (стохастикалық тренд) немесе TS (детерминирленген тренд) бар-жоғын тексереді.

13. Түрлендірулер көмегімен сызықтық функцияға келтіруге болатын жұптық (бір факторлы) және көп факторлы (көп факторлы) регрессия, сондай-ақ сызықтық және сызықтық емес регрессия болады.

2 ТРАФИКТІ ТАЛДАУ МОДЕЛЬДЕРІ ЖӘНЕ СТАЦИОНАРЛЫҚТЫ АНЫҚТАУ

2.1 Өлшенген трафиктің сандық сипаттамаларын талдау

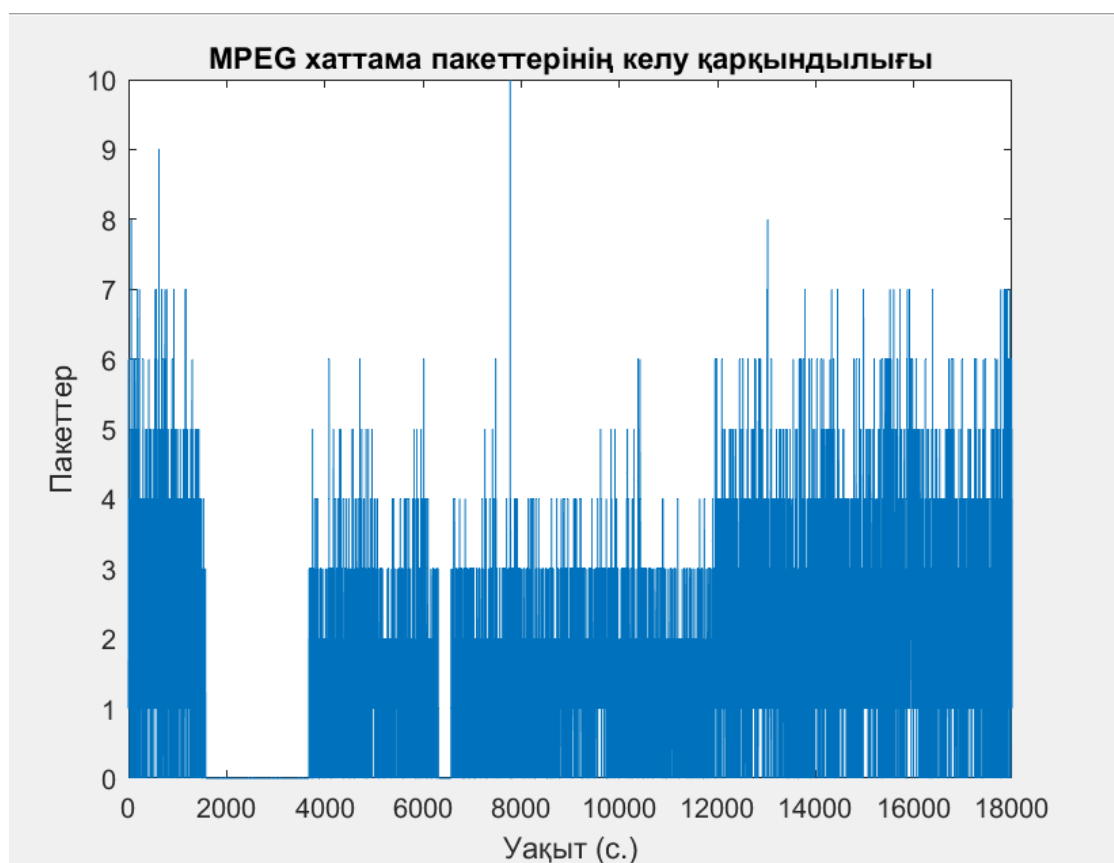
Бастапқы деректер Wireshark sniffer бағдарламасы арқылы алынды. 5 сағат ішінде 278 557 пакет бақыланды: оның ішінде 158 TCP (Transmission Control Protocol - Таратуды басқару хаттамасы) пакеттері; 493 ARP (Address Resolution Protocol - Адресі шешу хаттамасы) пакеттері; 25733 MPEG (Moving Picture Experts Group - Қозғалмалы сурет бойынша сарапшылар тобы), 250241 UDP (User Datagram Protocol - Пайдаланушы деректерінің хаттамасы) және т.б. UDP пакеттерінің үлесі 89,83% құрады. Бақылау трафигінің қалған бөлігі TCP, ARP, DHCP (Динамикалық хост конфигурациясының хаттамасы), DNS (Domain Name System - домендік атаулар жүйесі), MPEG (Moving Picture Experts Group), IGMP (Internet Group Management Protocol - Интернет тобын басқару хаттамасы), STP (Spanning Tree Protocol – байланыс ағаш хаттамасы) және т.б. арасында таратылды.

```
1 0.000000 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
2 0.002104 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
3 0.004069 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
4 0.006304 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
5 0.008388 192.168.172.30 239.2.4.4 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
6 0.008391 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
7 0.010750 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
8 0.010753 192.168.172.30 239.2.4.4 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
9 0.012718 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
10 0.013497 192.168.172.30 239.2.4.4 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
11 0.014810 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
12 0.016488 192.168.172.30 239.2.4.4 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
13 0.016931 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
14 0.018195 192.168.172.30 239.2.4.4 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
15 0.019525 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
16 0.020687 DTS_14321.272788888 PTS_14321.312788888 MPEG_PES_video-stream 1358 bytes
17 0.021821 192.168.172.20 239.2.2.40 UDP Source_port: bre Destination_port: cisco-sccp 1358 bytes
```

Сурет 2.1 - Өлшенген желілік трафиктің бөлігі

Уақыттық қатардың стационарлығын анықтау үшін әртүрлі әдістерді қолдануға болады: уақыттық қатардың графикалық көрінісін визуалды талдау, жиілік таралу формасын анықтау, тренд пен мерзімді компоненттің болуын тексеру, автокорреляцияның болуы және т.б.

Өлшенген MPEG қатары 2.2 суретте көрсетілген, ол 5 сағат ішіндегі әрбір секунд сайын магистральдық желі арқылы берілетін пакеттер санын көрсетеді. Уақыттық қатарда 18000 нүкте бар. Графиктен қатардың біркелкі емес қарқындылыққа ие екендігін бақылауға болады (бақылаулардың шашырауы уақыт өткен сайын артады және азаяды), пакеттер бір жерлерде тығыз шоғырланса, кейбір жерлерде аз не мүлдем жоқ болатын бөлімдер бар.



Сурет 2.2 - MPEG хаттама пакеттерінің келіп түсу қарқындылығы

Қатарды статистикалық бағалау желілік трафикті болжауда сәйкес әдістерді қолдану кезінде маңызды рөл атқарады.

Жиілік таралу формасының *сандық сипаттамаларына* мыналар жатады: асимметрия және эксцесс (куртозис).

Асимметрия таралу қалыптылығына тән симметриялы таралудан ауытқу дәрежесін көрсетеді, ол мына формуламен есептеледі:

$$Ax = \frac{1}{(n-1)Sx^3} \sum_{i=1}^n x_i - Mx \quad (2.1)$$

және -3-тен +3-ке дейінгі мәндерді қабылдайды.

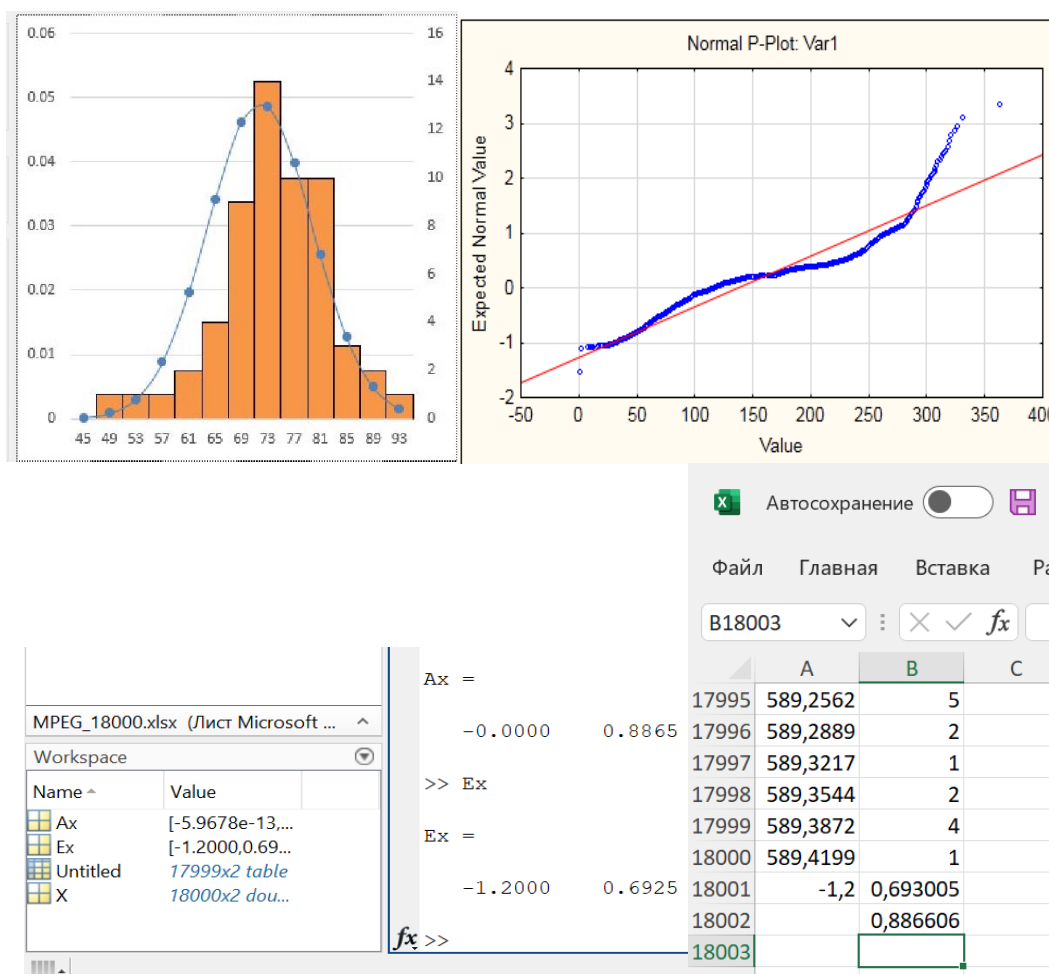
$Ax=0$ болғанда таралу симметриялы болады, $Ax<0$ болғанда – сол жақты асимметрия, $Ax>0$ болғанда оң жақты асимметрия болады.

Эксцесс (куртозис) таралу қисығының шыңын анықтайды, мына формуламен есептеледі:

$$Ex = \frac{1}{(n-1)Sx^4} \sum_{i=1}^n (x_i - Mx)^4 - 3 \quad (2.2)$$

және -3-тен +3-ке дейінгі мәндерді қабылдайды. $Ex=0$ кезінде таралу қалыпты, $Ex<0$ кезінде төбесі тегіс шыңды, $Ex>0$ кезінде таралу үшкір шыңды болады. $Ax=0,886606>0$ – оң жақ асимметриясының таралуы. Куртозис $Ex=0,693005>0$ (2.3 сурет). Екі мән де зерттелетін қатар мен қалыпты таралу заңы арасындағы сәйкессіздікті көрсетеді.

Бастапқы деректердің таралу қалыптылығын графикалық тексеру нәтижесі 2.3 суретте көрсетілген. Егер деректер қалыпты болса, онда гистограмма қоңырау тәріздес болып, Normal P-Plot экранында график нүктелері қызыл сызық бойымен орналасар еді. Нәтижелер бастапқы MPEG қатарының қалыпты таралу заңының жиынтық тобына (генеральная совокупность нормального закона распределения) жатпайтынын көрсетеді.



Сурет 2.3 - Таралу формасын сандық және графикалық тексеру нәтижелері

2.4 суретте қалыпты таралу сынақтарын жасайтын критерийлер нәтижелері көрсетілген. Олардың барлығы дерлік қатардың қалыпты таралуын жоққа шығарды.

	A	B	C	D	E	F	G	H	I
1	0,032746	1		Проверка нормальности распределения					
2	0,065491	2		Выдача обычно включает:					
3	0,098237	6		Статистика, Р-значение двустороннее, вывод					
4	0,130982	1		Выбранное пороговое значение					
5	0,163728	2		0,05					
6	0,196473	3		Численность выборки					
7	0,229219	3		18000					
8	0,261964	3		Модифицированный критерий Колмогорова					
9	0,29471	3		0,200817	0	Гипотеза о нормальности отклоняется			
10	0,327456	2		Модифицированный критерий Смирнова					
11	0,360201	3		0,200817	0	Гипотеза о нормальности отклоняется			
12	0,392947	3		Критерий Крамера-Мизеса					
13	0,425692	2		115,6663	0	Гипотеза о нормальности отклоняется			
14	0,458438	1		Критерий Андерсона-Дарлинга					
15	0,491183	3		722,6446	0	Гипотеза о нормальности отклоняется			
16	0,523929	2		Критерий Шапиро-Уилка					
17	0,556674	4		Много данных. Расчет не выполнен.					
18	0,58942	1		Критерий Шапиро-Франсиа					
19	0,622165	3		0,876696	0,367667	Гипотеза о нормальности не отклоняется			
20	0,654911	3		Критерий коэффициента асимметрии					
21	0,687657	3		0,886606	0	Гипотеза о нормальности отклоняется			
22	0,720402	3		Критерий эксцесса					
23	0,753148	0		0,693005	0	Гипотеза о нормальности отклоняется			
24	0,785893	1		Критерий Жарка-Бера					
25	0,818639	3		2717,461	0	Гипотеза о нормальности отклоняется			

Сурет 2.4 - Қалыпты таралу критерийлерінің нәтижелері

2.2 Трендтің бар болу гипотезасын тексеру

Зерттелетін қатар қалыпты заң бойынша таралмағандықтан, қатардың стационарлық немесе стационарлық еместігін анықтау әдістері үшін параметрлік емес критерийлер қолданылады:

- қатар критерийі;
- «өсу және кему» қатарының критерийі;
- Уоллис-Мур фазалық жиілік критерийі;
- Кокс-Стюарт трендінің белгі критерийі.

Қатар критерийі трендті көрсетеді, бұл жағдайда әрбір бақылау (+) немесе (-) ретінде белгіленетін бір-бірін жоққа шығаратын екі тізбектің біріне тағайындалады. Қатар – бір типті бақылаулар тізбегі, оның алдында және одан кейін қарама-қарсы типтегі бақылаулар бар [37].

Кез келген қажетті маңыздылық деңгейі α болатын гипотезаны тексеру үшін алынған қатарлар санының мәнін гипотезаны қабылдау аймағының шекараларымен салыстыру қажет.

Егер нәтижелердің саны осы аймақтан тыс болса, онда гипотеза α маңыздылық деңгейімен қабылданбауы керек. Әйтпесе, гипотезаны қабылдауға болады.

«Өсу және кему» қатарының критерийі, оған сәйкес тренд келесі алгоритм бойынша анықталады:

$$\delta_i = \begin{cases} +, & \text{егер } y_{t+1} - y_t > 0 \\ -, & \text{егер } y_{t+1} - y_t < 0 \end{cases} \quad (2.3)$$

яғни, *Өсу қатары* - әрбір келесі мән алдыңғысынан үлкен болатын мәндер тізбегі, ал *Кему қатары* - әрбір келесі мән алдыңғысынан аз болатын мәндер тізбегі. Бұл критерий бойынша, егер келесі мән алдыңғысына тең болса, онда бір ғана мән есепке алынады. «Өсу және кему» қатарының критерийі деректерде тренд бар-жоғын бағалау үшін пайдаланылатын параметрлік емес статистикалық тест болып табылады. Ол деректер тізбегіндегі өсу және кему қатарларының санын санауға негізделген.

Әрі қарай өсу және кему қатарларының әрқайсысының саны - $v(n)$ есептеледі. Серия - қатар орналасқан плюстер немесе минустар тізбегі, бір плюс не бір минус та серия болып саналады. $l_{\max}(n)$ - ең ұзын серия ұзындығы.

Егер келесі теңсіздіктердің кем дегенде біреуі бұзылса, онда трендтің жоқтығы туралы гипотеза 0,95 сенімділік ықтималдығымен жоққа шығарылады:

$$\left\{ \begin{cases} v(n) > \frac{1}{3}(2n-1) - u_t \sqrt{\frac{16n-29}{90}}, \\ l_{\max}(n) < t_{кр} \end{cases} \right. \quad (2.4)$$

u_t - деңгейдің қалыпты таралу квантилі $(1-\alpha)/2$;

Уоллис-Мур фазалық жиілік критерийі деректерде маусымдық немесе мерзімділік бар-жоғын анықтау үшін қолданылатын параметрлік емес статистикалық сынақ болып табылады. Ол деректердегі қайталанатын оқиғалардың немесе құбылыстардың фазалары мен жиіліктерін талдауға негізделген.

Фаза - қайталанатын мәндер тізбегіндегі реттік нөмір, ал жиілік - қайталаулар саны. Уоллис-Мур фазалық жиілік критерийі деректердегі фазалық жиілік құрылымының кездейсоқ немесе маңызды екенін бағалауға мүмкіндік береді.

Ол критерийдің мәнін шығарады және деректер құрылымының альфа=0,05 мәнділік деңгейінде кездейсоқ реттіліктен айтарлықтай айырмашылығы бар-жоғын анықтайды. Уоллис-Мур критерийі оң және теріс фазалардың жиілігін - белгілердің бірдей реттілігін қарастырады және айырмашылықтарды, сондай-ақ

фазалардың санын L есептеу кезінде бастапқы және соңғы фазаларды есепке алмайды. Фазаның есептелген мәні - жиілік айырмасының критерийі формуласымен анықталады

$$\hat{z} = \frac{|L - \frac{2 \cdot n - 7}{3}| - 0,5}{\sqrt{\frac{16 \cdot n - 29}{90}}}. \quad (2.5)$$

\hat{z} – фазалық-жиілік айырмашылық критерийі;
 L – фазалар саны;
 n – деңгейлер саны.

Кокс-Стюарт әдісі n мәннен тұратын зерттелетін қатарды 3 бірдей бөлікке x_i^1 , x_i^2 және x_i^3 бөлуді талап етеді және бірінші және үшінші топтардың мәндер айырмасын $x_i^3 - x_i^1$ қарастырады. Яғни, осы деректер негізінде плюс белгісі өсу тенденциясымен (s^+), минус – кему тенденциясымен (s^-), нөл – тренд жоқ болғанда (s^0) болады, нәтижеде

$$s^+ + s^- + s^0 = n/3 \quad (2.6)$$

$$\hat{z} = \frac{|s - \frac{n}{6}| - 0,5}{\sqrt{\frac{n}{12}}} \quad (2.7)$$

$$S = \max\{s^+, s^-\} \quad (2.8)$$

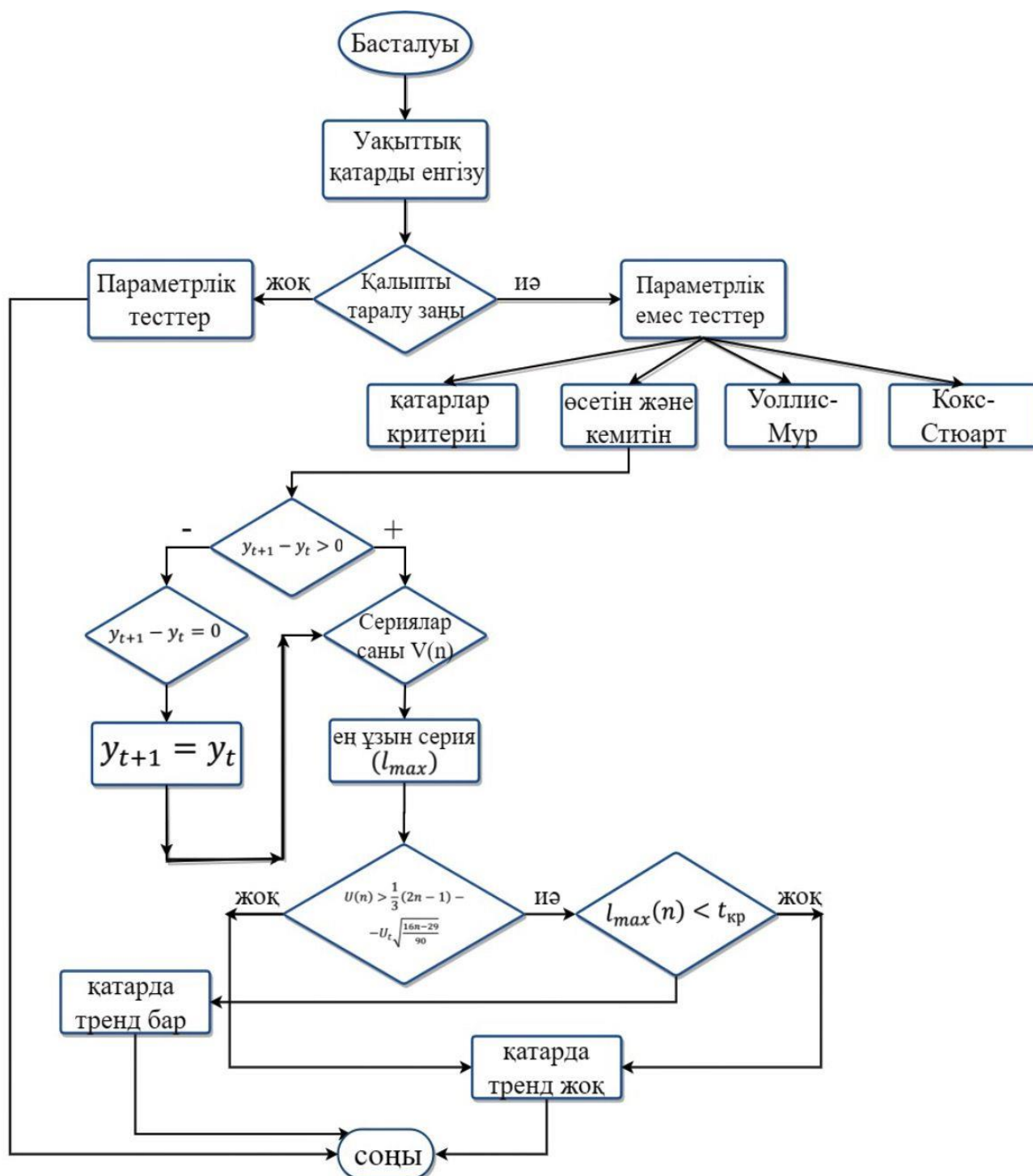
Кокс-Стюарт сынағы деректердің дәйекті топтарындағы бағыттың өзгеруін (жоғары немесе төмен) бағалайды. Ол деректердегі трендтің кездейсоқ немесе маңызды екенін $\alpha=0,05$ маңыздылық деңгейінде анықтауға мүмкіндік береді.

Бұл критерийлер маңызды тренд құрылымдарын анықтау үшін уақыттық қатарларды және деректер ретін талдау үшін пайдалы болуы мүмкін.

Осы критерийлер негізінде диссертациялық жұмыста зерттелетін қатарда трендтың бар-жоғын анықтайтын бағдарлама жасап шығарылды. Алдымен зерттелетін нақты өлшенген 18000 мәннен тұратын бастапқы уақыттық қатардан әрбір 10 с. сайынғы пакеттер санын жинақтау арқылы 1800 мәнге ие қатар алынды. Бағдарламаның алгоритмі 2.5 суретте көрсетілген. Егер қатарда тренд болса, ол қатар стационарлы емес деп саналады. Бағдарлама жалпы мақсаттағы Python бағдарламалау тілі негізінде жүзеге асты.

Бағдарламаның кодтық бөлігі 1-ші қосымшада толық берілген.

2.3 Уақыттық қатардың стационарлығын анықтайтын бағдарлама сипаттамасы



Сурет.2.5 - Қатардың стационарлығын тексеретін бағдарлама алгоритмі

2.4 Уақыттық қатардың стационарлығын бағалау

Уақыттық қатардың стационарлығын бағалау үшін корреляция – қатардың стационарлық дәрежесін бағалауға мүмкіндік беретін уақыттық қатарларды талдау әдістерінің бірі. Стационарлық қатардың статистикалық қасиеттері уақыт өте келе өзгермейтінін білдіреді, соның ішінде орташа және дисперсия.

Кездейсоқ процестің математикалық күтуі кездейсоқ шамалардың осы процесін құрайтын математикалық күтулердің траекториясы болып табылады:

$$m(t) = MX(t). \quad (2.8)$$

Кездейсоқ процестің дисперсиясы кездейсоқ шамалардың осы процесін құрайтын дисперсиялардың траекториясы деп аталады:

$$\sigma^2(t) = DX(t) \quad (2.9)$$

Стационарлықтың бір белгісі деректерде жүйелі тәуелділіктің (тенденцияның) болмауы және автокорреляция сияқты қатардың статистикалық қасиеттерінің уақыт бойынша өзгерісінің болмауы.

«Корреляция» терминін алғаш рет француз палеонтологы Ж.Кювье қолданған, ал бұл терминді статистикаға 1886 жылы ағылшын биологы және статистигі Фрэнсис Гальтон енгізген. Корреляция коэффициентін есептеу формулаларын математик және биолог Карл Пирсон жасаған.

Корреляция кездейсоқ шамалардың тәуелділігінің қаншалықты күшті екенін және басқа бір көрсеткіштің мәнін біле отырып, бір көрсеткіштің мүмкін мәндерін болжауға болатындығын көрсетеді.

Егер уақыттық қатарда тренд және циклдік компонент болса, онда қатардың келесі деңгейінің мәндері алдыңғыларға тәуелді [38]. Бұл тәуелділік қатар деңгейлерінің автокорреляциясы болып табылады. Оның сандық мәнін өлшеу бастапқы уақыттық қатарлардың деңгейлері мен уақыт бойынша бірнеше кадамға жылжыған осы қатардың деңгейлері арасындағы корреляциялық көрсеткішті қолдану арқылы жүзеге асырылады.

x және y үлгілері арасындағы Пирсон корреляция коэффициентінің формуласы:

$$r_{xy} = \frac{\overline{\Sigma(x_j - \bar{x}) \cdot (y_j - \bar{y})}}{\sqrt{\Sigma(x_j - \bar{x})^2 \cdot \Sigma(y_j - \bar{y})^2}} \quad (2.10)$$

мұнда:

r_{xy} – Пирсон корреляция коэффициенті

x_j, y_j - сәйкесінше x және y үлгілерінің элементтері

\bar{x}, \bar{y} - сәйкесінше x және y үлгілерінің орташа мәндері

2.10 формула үлгілер арасындағы ковариациясының үлгілердің стандартты ауытқуларының көбейтіндісіне қатынасын білдіреді. Пирсон корреляция коэффициенті -1-ден 1-ге дейінгі мәндерді қабылдайды, мұнда:

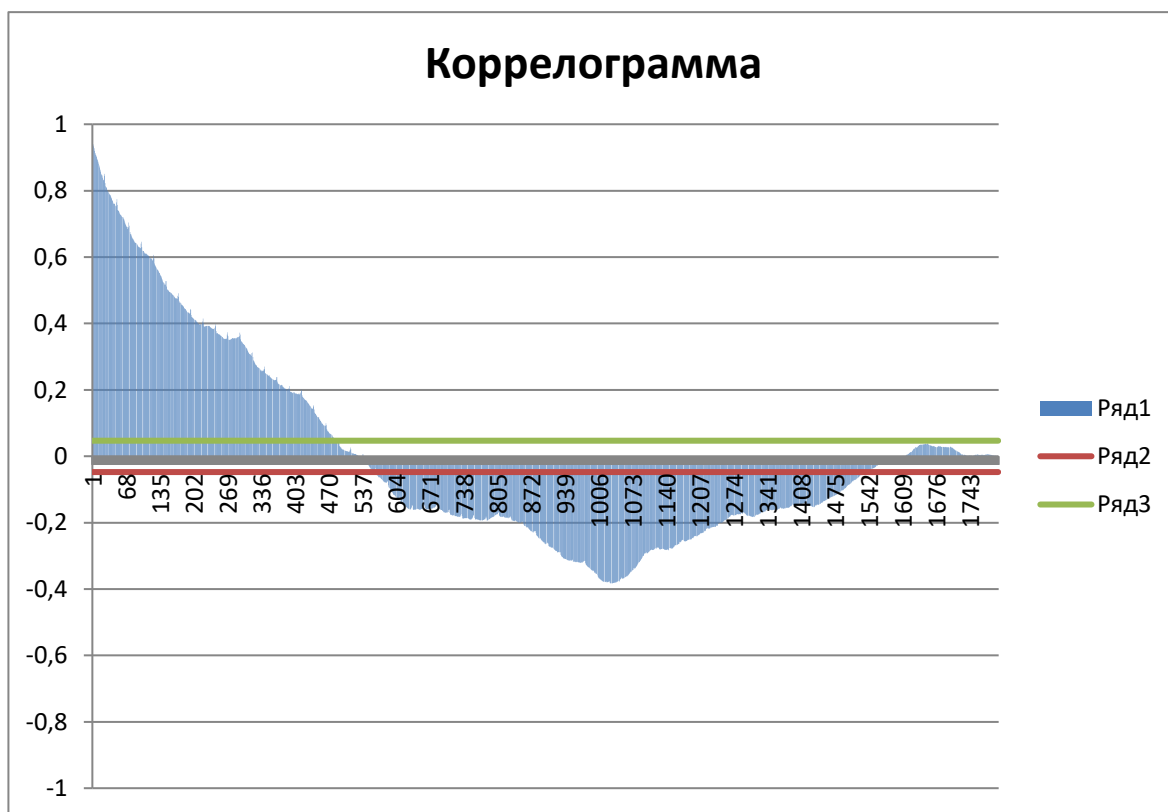
1 мәні үлгілер арасындағы толық тікелей сызықтық корреляцияны білдіреді

-1 мәні үлгілер арасындағы толық кері сызықтық корреляцияны білдіреді

0 мәні үлгілер арасындағы сызықтық корреляцияның жоқтығын білдіреді.

Уақыттық қатар элементтерінің автокорреляциясы - уақыттық қатардың бірізді элементтері арасындағы корреляциялық тәуелділік. Кешігу (лаг) – қатар элементтерінің жұптары арасындағы автокорреляция коэффициенті есептелетін кезеңдердің саны. Уақыттық қатардың автокорреляциялық функциясы 1, 2, 3... тең кешігуі бар автокорреляция коэффициенттерінің тізбегі болып табылады. Функция мәндерінің лаг мәніне тәуелділік графигі коррелограмма деп аталады.

Төмендегі 2.6 суреттегі коррелограмманы пайдалана отырып, ең жоғары кешігуді (лагты) анықтауға болады. Біздің жағдайда автокорреляцияның ең жоғары деңгейінің сандық мәні 1-ге тең лагпен 0,962896, сондықтан бастапқы қатар тек трендті қамтиды. Әрі қарай автокорреляция коэффициенттерінің деңгейлері баяу төмендейді және артта қалған сайын нөлге ұмтылмайды, демек қатарда тренд бар, яғни қатардың стационарлы емес екендігі байқалады [39].



Сурет 2.6 - Қатардың коррелограммасы

2.5 Бірлік түбір тесті арқылы стационарлықты бағалау

«Бірлік түбір» түсінігі жүйедегі тербелістердің сипатын анықтайтын көрсеткіш екендігі [40] жұмыста сипатталған және оның есептелу формулалары 1-тараудағы 1.8 және 1.9 формулаларда берілген.

Ықтималдықтар теориясы мен статистикада бірлік түбір кейбір стохастикалық процестердің (мысалы, кездейсоқ жүрістердің) сипаттамасы болып табылады. Бірлік түбір сынағы уақыттық қатардың стационарлық еместігін және уақыттық қатарларды талдауда бірлік түбірден тұратынын тексереді.

N-ші ретті сызықтық айырымдық теңдеулер жүйесінің N түбірі бар. Олардың кез келгенінің абсолютті мәні 1-ден үлкен болса, жүйе «жарылысқа» жақындайды, егер барлық түбірлер 1-ден аз болса, кез келген уақытша ауытқулардан кейін жүйе сөзсіз өзінің бастапқы тепе-теңдігіне ұмтылады. Абсолютті мәнің 1-ге тең түбірі немесе бірлік түбір жүйенің тұрақты ығысуын тудырады, ал бірқатар бұзушылықтар бастапқы позициядан шексіз ауытқуды тудыруы мүмкін.

Аналитикалық түрде бірлік түбір сынағы келесі түрде ұсынылуы мүмкін:

$$y_t = D_t + z_t + \varepsilon_t \quad (2.11)$$

мұндағы y_t – t уақытындағы қатардың деңгейі;

D_t – детерминирленген компонент;

z_t – стохастикалық компонент;

ε_t – стационарлық қателік процесс.

Бірлік түбірді тексерудің негізгі тұжырымдамасы z_t (стохастикалық құрамдас) бірлік түбірден тұратынын немесе жоқтығын анықтау болып табылады.

Диссертациялық жұмыста уақыттық қатарға екі бірлік түбір сынағы қолданылды:

- Augmented Dickey-Fuller Test (ADF-test), оны 1979 жылы американдық ғалымдар Дэвид Алан Дики мен Уэйн Артур Фуллер әзірлеген;

- Филлипс-Перрон сынағы (PP-тест) 1988 жылы жасалған. [41,42] жұмыстарда Дикки-Фуллер тесті қатардың стационарлық/стационарлық емес түрін анықтау үшін бірлік түбірдің (единичный корень) болуын тексеретіні сипатталған. Бұл жағдайда H_0 нөлдік гипотеза қатардың DS (Difference stationary) типті, ал балама TS (Trend stationary) типті қатарға сәйкес келетіні айтылған. Сынақтың авторлары – америкалық статистиктер Дэвид Алан Дики мен Уэйн Артур Фуллер. Бұл сынақты олар 1979 жылы ұсынған болатын.

Келесі гипотеза қолданылады:

- H_0 : – қатар стационарлы емес: ол бірлік түбірді қамтиды және кездейсоқ жүру (случайное блуждание) процесімен сипатталады.

- H_1 : – қатар стационарлы - онда бірлік түбір болмайды және 1-ші ретті стационарлық авторегрессивті процесс арқылы сипатталады. Matlab бағдарламасындағы Econometric Modeler құралы арқылы тексерілген тест

нәтижесі 2.7 суретте көрсетілген, онда зерттелетін қатар бірлік түбірге ие екенін көруге болады.

Augmented Dickey-Fuller Test(X2)									
Null Hypothesis: X2 contains a unit root									
Results									
	Select	Null Rejected	P-Value	Test Statistic	Critical Value	Lags	Model	Test Statistic	Significance Level
1	<input type="checkbox"/>	true	1.0000e-03	-57.8492	-1.9416	0	AR	t1	0.0500

Сурет 2.7 - Дикки-Фуллер тестінің нәтижесі

Гипотезаны растау немесе теріске шығару үшін маңыздылық деңгейінің ең кіші мәні р-мәні қолданылады. Егер р-Value 5%-дан жоғары болса, онда нөлдік гипотеза жоққа шығарылады, демек, қатар стационарлық болып табылады [43].

2.8 сурет көрсеткендей, р-мәні 0,001, ал маңыздылық деңгейінің мәні 0,05 екенін көруге болады. р-мәні маңыздылық деңгейінің мәнінен аз болғандықтан, зерттелетін уақыттық қатар кездейсоқ жүріс сипатына ие, яғни зерттелетін қатар стационарлық емес деген гипотезаны жоққа шығару мүмкін емес.

Таблица 1.1. Test Parameters

	Lags	Model	Test Statistic	Significance Level
1	0	AR	t1	0.05

Таблица 1.2. Test Results

	Null Rejected	P-Value	Test Statistic	Critical Value
1	true	0.001	-57.8492	-1.9416

Сурет 2.8 - ADF-тест параметрлерінің мәндері

Егер тексеру кезінде уақыттық қатардың авторегрессивті моделінде бірлік түбір алынса, онда бұл уақыттық қатардың айырмасын білдіреді. Нәтижесінде қатар стационарлы емес деген болжам расталды.

[44]-те қатарды стационарлық немесе стационарлы емес деп талдау үшін бір емес, бірнеше тесттерді қолдану орынды екені айтылған. Сондықтан басқа да тесттер, атап айтқанда 1988 ж. жасалған Филлипс-Перрон тесті (PP-test) қолданылды. PP тесті нөлдік гипотеза айнымалы шамада бірлік түбір бар, ал балама гипотеза айнымалы стационарлық процесс екенін көрсетеді. 2.9 суреттен зерттелетін қатар бірлік түбірге ие, яғни қатар стационарлы емес екенін көруге болады.

Phillips-Perron Test for One Unit Root(X2)									
Null Hypothesis: X2 contains a unit root									
Results									
	Select	Null Rejected	P-Value	Test Statistic	Critical Value	Lags	Model	Test Statistic	Significance Level
1	<input type="checkbox"/>	true	1.0000e-03	-57.8492	-1.9416	0	AR	t1	0.0500

Сурет 2.9 – Филлипс - Перрон тестінің нәтижесі

Екі тест те уақыттық қатарды стационарлық емес деп бағалағандықтан, ол шын мәнінде стационарлық емес деуге болады.

2.6 Екінші тарау бойынша қорытындылар

1. Уақыттық қатардың стационарлығын анықтау үшін әртүрлі әдістерді қолдануға болады: уақыттық қатардың графикалық көрінісін визуалды талдау, жиілік таралу формасын анықтау, тренд пен мерзімді компоненттің болуын тексеру, автокорреляцияның болуы және т.б.;
2. Зерттелетін қатар қалыпты қатарға жатпайды. Куртозис $E_x=0,693005>0$ (қалыпты жағдайда $E_x=0$). Normal P-Plot графигінде қатар мәндері қызыл сызықтың бойымен емес (қалыпты жағдайда қызыл сызықтың бойымен болады);
3. Колмогоров және Смирнов; Крамер-Мизес; Андерсон-Дарлинг; Шапиро-Франсиа; эксцесс; Жарк-Бер, т.б. критерийлер қатардың қалыпты емес екенін дәлелдеді;
4. Бірнеше параметрлік емес критерийлер негізінде әзірленген бағдарлама қатарда тренд барын растады яғни қатардың стационарлы емес екендігі байқалды;
5. Корреляция деңгейлері баяу төмендеді және артта қалған сайын нөлге ұмтылмады, демек қатарда тренд бар, яғни қатардың стационарлы емес екендігі байқалды;
6. Matlab бағдарламасында бірнеше тесттер (ADF-test, PP-test) қатарда бірлік түбірдің барын растады;
7. Қатарда бірлік түбір болса, ол стационарлы емес болып саналады.

3 УАҚЫТТЫҚ ҚАТАРДЫ УАҚЫТ-ЖИІЛІК БОЙЫНША ТАЛДАУ

3.1 Спектралды талдау әдістері

Кез келген кездейсоқ процесте уақыт облысы да, жиілік көрінісі де болады. Кейбір сипаттар үшін уақытты зерттеу, кейбірі үшін жиілік қолайлырақ болуы мүмкін. Спектрлік талдаудың классикалық әдістеріне сингулярлық спектрлік талдау (SSA), тәуелсіз компоненттік талдау (ICA) және толқындық талдау (WA) әдістерін қосуға болады.

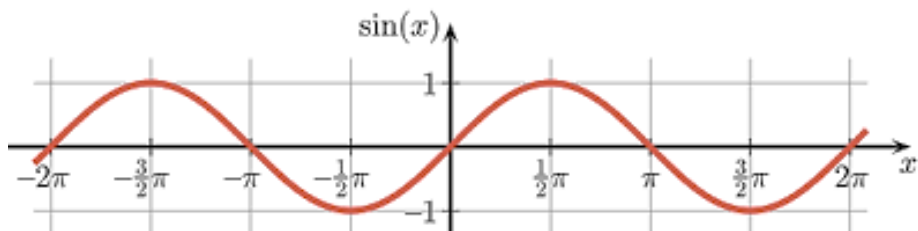
Спектрлік әдісті қолдану әртүрлі уақыт шкалаларында қатардың құрылымы мен циклдік мінез-құлқын түсінуді жақсарты алады.

Спектрлік талдау қатардың жиілік құрамын сипаттайды. Толқындық теңдеудің шешімін 1807 жылы Фурье қатарын енгізу арқылы алғаш ашқан барон Жан Батист Джозеф де Фурье болды. Фурье теориясын 1836 жылы Штурм мен Лиувилл ерікті ортогоналды функциялар жағдайына кеңейтті. Штурм-Лиувилл теориясы 1925 және 1926 жылдары Гейзенберг пен Шредингердің кванттық механиканы тұжырымдауына байланысты ең сәтті эмпирикалық спектрлік талдауға айналды. 1929 жылы Джон фон Нейман атомның спектрлік теориясын Гильберт кеңістігінде спектрлік бейнелеу теоремасы арқылы берік математикалық негізге қойды. Осы уақытта Винер 1923 жылы броундық қозғалыстың математикалық теориясын жасады, ал 1930 жылы стационарлық кездейсоқ процестің спектрлік бейнесінің жалпылама гармоникалық талдауын енгізді.

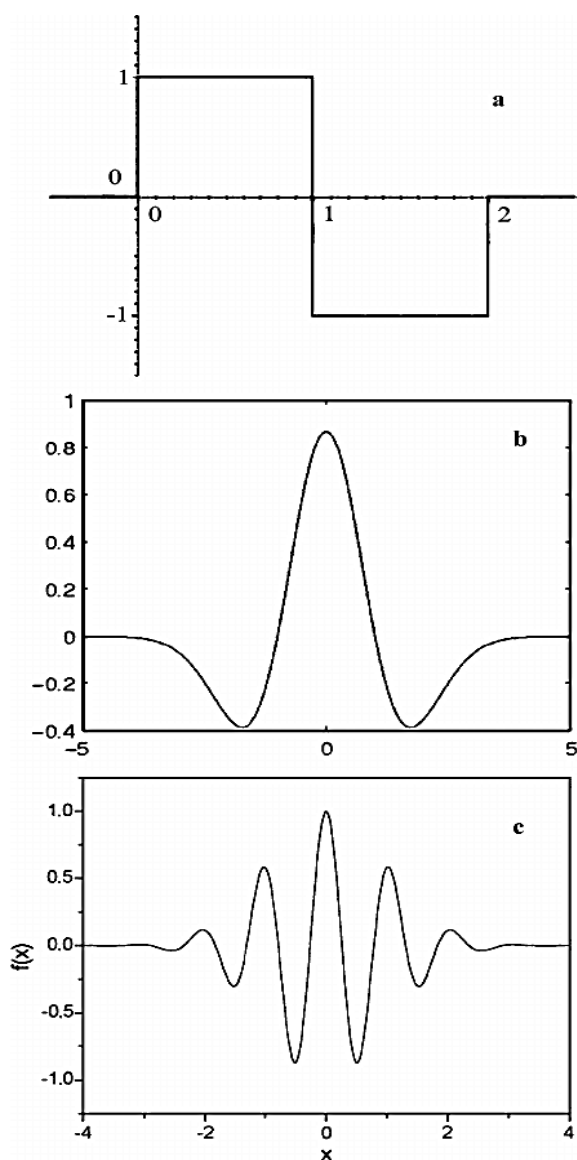
Спектрлік талдаудың қазіргі тарихы 1949 жылы Дж. В.Тьюкидің спектрлік бағалауды ашуынан басталады, бұл Фурье ашқан жаңалығының статистикалық баламасы [45]. Дегенмен, спектрлік талдау есептеу үшін қымбат болды. Ол 1965 жылы Дж.С.Кули мен Дж. В.Тьюкидің FFT (Fast Furier Transform - Жылдам Фурье түрлендіруі) алгоритмін жариялаумен қол жетімді болды. Кули-Тьюки әдісі тәжірибеде сигналдарды уақыт және жиілік облыстарында өңдеуге мүмкіндік берді, бұл үздіксіз жүйелерде бұрын мүмкін болмаған. Фурье түрлендіруі тек теориялық сипаттама ғана емес, сонымен қатар құрал болды. ФТ дамуымен спектрлік талдаудың эмпирикалық өрісінің маңызы арта түсуде. Одан кейінгі үлестері: 1967 жылы Джон Бургтың максималды энтропиялық спектрлік талдауды енгізуі; 1950 жылдары Эммануэль Парзен және т.б. ның спектрлік терезелерді жасауы және т.б. [46]

Фурье түрлендіруі аналитикалық түрде сипатталмайтын және талдауы салыстырмалы түрде қиын ерікті форманың бастапқы периодты емес функциясын әртүрлі жиіліктері мен амплитудалары бар синустар немесе косинустар жиынтығы ретінде көрсетеді. Яғни, күрделі функция Фурье түрлендіруінің нәтижесінде қарапайым синусоидтардың (немесе косинус толқындарының) жиынтығына айналады. Бұл Фурье спектрін құрайтын спектрлік компоненттер (гармониялар) [47]. Графикалық түрде әрбір спектрлік құрамдас көлденеңі оның жиілігіне, ал биіктігі оның амплитудасына сәйкес келетіндей көрсетіледі.

Толқынды түрлендіру (Вейвлет-преобразование) немесе вейвлеттер сигналдарды интерпретациялаудың жаңа әдісін ашты. Ол толқындар шексіз таралатын синус немесе косинус толқынын пайдаланатын Фурье қатарымен салыстырғанда шағын толқындарды пайдаланады (3.1 сурет). Вейвлеттер тез ыдырайтын мини-толқын түріндегі толқынның қысқа жарылыстары (3.2 сурет).



Сурет. 3.1 - Шексіз синусоида



Сурет. 3.2 - Вейвлеттер: Haar (a), Mexican Hat (b), Morlet (c)

Жоғарыда келтірілген суреттер синус немесе косинус толқындарының уақыт бойынша шектелмейтінін, ал вейвлет толқындарының уақыт бойынша шектелгенін көрсетеді. Вейвлет толқындары уақыт пен жиілікте шектеулі болғандықтан пайдалы. Төмен және жоғары жиіліктерді өңдеу кезінде вейвлеттер әртүрлі әрекет етеді. Жоғары жиіліктер үшін синусоида қысылады, ал төмен жиіліктер үшін ол созылады.

Классикалық Фурье талдауы уақыттық қатар деректеріндегі стационарлы сигналдарды анықтаудың теңдесі жоқ құралы болып табылады және стационарлы емес сигналдарды зерттеуде қауқарсыз [48]. Бұл ретте олардың амплитудалық және жиілік сипаттамалары бағаланады.

Толқындық талдау – көптеген уақыт шкалаларында сигнал қуатының өзгеруін көрсететін уақыт бойынша стационарлы емес немесе біркелкі емес уақыт қатарын талдауға арналған қуатты математикалық құрал. Яғни, талдау нәтижелері сигналдың жалпы жиілік реакциясын ғана емес, сонымен қатар сигналдың жиілік құрамдас бөліктерін де қамтиды [49,50].

3.2 Сингулярлық спектрлі талдау

Зерттелетін қатарға сингулярлық спектрлі талдау - SSA (Singular Spectrum Analysis - Сингулярлы Спектрлік Талдау) әдісі қолданылды. Әдіс уақыттық қатардың құрылымдық компоненттерін талдаудың заманауи құралдарының бірі болып табылады. Уақыттық қатарларды ыдырату әдісі (SSA) қатарды тренд, мерзімділік және шу ретінде түсіндірілетін жиынтық құрамдас бөліктерге ыдыратады және бұл әдіс негізгі құрамдастардың бөліну мүмкіндігіне баса назар аударады және әртүрлі уақыт шкалаларында орын алатын кезеңділіктерді, тіпті өте шулы қатарларда оңай ажырата алады. Бастапқы уақыттық қатар өңдеу нәтижесінде алынған оның барлық құрамдастарын қосу арқылы қалпына келтіріледі. Сонымен қатар, SSA әдісі оны пайдалану үшін уақыттық қатарлардың табиғаты туралы белгілі бір шарттарды талап етпейді. SSA әдісі параметрлік емес және зерттелетін бір өлшемді қатардың сәйкес таралу заңдарына жатуын, қатарды алдын ала тұрақтандыруды талап етпейді және зерттелетін қатардың стационарлық немесе стационарлы еместігіне байланысты емес, және оның бір ғана реттелетін L параметрі бар.

SSA әдісі бір өлшемді уақыт қатарын бір параметрді жылжыту арқылы көп өлшемдіге түрлендіреді. Алынған траектория матрицасының қасиеттерін пайдалана отырып, ол сингулярлық мәннің ыдырауын қолданады және негізгі компоненттерді таңдау арқылы өңделген деректерді қалпына келтіреді (жуықтатады) [51]. SSA әдісі көп функциялы және негізгі құрамдас талдауды (PCA, Principal Component Analysis), кластерлік талдауды және жеке векторлардың графикалық талдауын орындайды.

Бұл әдіс $X = X_N = (X_1, \dots, X_N)$ нақты уақыттық қатарды қарастырады. $N > 2$ және X нөлдік емес қатар, яғни X қатарында $f(i) \neq 0$ болатындай i бар деп ұйғарылады. Негізгі параметр L ($1 < L < N$) терезенің ұзындығы деп аталатын қандай да бір бүтін сан (N -нің жартысынан аспауы керек). L өлшемді $K = N - L + 1$ ендіру векторлары жасалады [52]:

$$X_i = (f_{i-1}, \dots, f_{i+L-2})^T, 1 \leq i \leq K. \quad (3.1)$$

Әрі қарай, бір өлшемді f қатары L -траекториялық матрицасына түрленеді, ол бағандар ретінде кірістіру векторларынан тұрады:

$$X = [X_1 : \dots : X_K]. \quad (3.2)$$

Траекторлы матрица:

$$X = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} f_0 & f_1 & f_2 & \cdot & \cdot & \cdot & f_{K-1} \\ f_1 & f_2 & f_3 & \cdot & \cdot & \cdot & f_K \\ f_2 & f_3 & f_4 & \cdot & \cdot & \cdot & f_{K+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ f_{L-1} & f_L & f_{L+1} & \cdot & \cdot & \cdot & f_{N-1} \end{pmatrix}. \quad (3.3)$$

Траектория матрицасы жылжымалы терезе арқылы қалыптасады, бұл кезде белгілі бір қадамы бар жолдың элементтері матрицаның бағандарына қосылып, бағандар саны терезенің ұзындығымен анықталады. Осылайша, матрицаның әрбір бағаны бастапқы уақыттық қатардың векторы болып табылады.

K параметрі траектория матрицасындағы бағандар санын білдіреді. X бағандары L -баяу векторлар деп аталады, ал жолдар K -баяу векторлар деп аталады. Бұл жағдайда коварианттық матрицаның негізгі компоненттерін талдау процесі жүзеге асады. Нәтижесінде сәйкес негізгі компоненттердің таңдамалы дисперсиялары болып табылатын матрицаның меншікті мәндері (собственные вектора) анықталады [53, 54]. Яғни, матрицаның сингулярлық мәнінің ыдырауы болады.

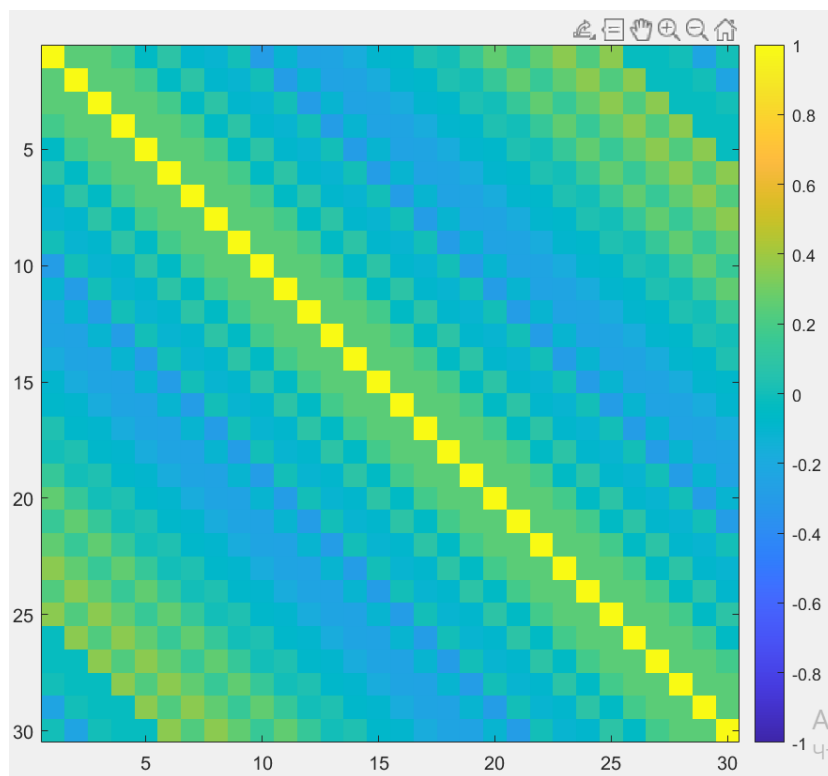
Матрицаның меншікті мәндері компоненттердің уақыттық қатарға әсер ету дәрежесін сипаттайды, ал меншікті векторлар уақыт функциясы ретінде түсіндіріледі.

[55] жұмыста алынған үлгілік коварианттық матрицаның спектрінің, яғни оның меншікті мәндерінің жиындары SSA сызықты емес спектрлік талдаудың қазіргі әдісінде маңыздылығы сипатталған. Бұл жағдайда уақыттық қатардың коварианттық матрицасының спектрі мен оның спектрлік қуаты арасындағы байланыстың болуы сипатталады. Сондай-ақ, коварианттық матрицаның меншікті мәндерінің спектрін сигналдар санын бағалау үшін пайдалануға болатыны атап өтілген.

Ганкель матрицасы уақыт қатарын блоктарға бөлу және олардан матрица құру арқылы қалыптасады, мұнда бірінші бағанда блоктардың бірінші

элементтері, екіншісінде - блоктардың екінші элементтері және т.б. Матрица бағандарының саны блок ұзындығына тең.

X матрицасы диагональ бойынша элементтері бірдей $i + j = const$, бұл Ганкель матрицасы.



Сурет. 3.3 - Ганкель матрицасы

Диссертациялық жұмыста SSA әдісі арқылы бір өлшемді уақыттық қатардың бір параметрін жылжыту арқылы көп өлшемдіге түрлендіру бағдарламасы жасалды. Алынған траектория матрицасының қасиеттерін пайдалана отырып, ол сингулярлық мәннің ыдырауын қолданады және негізгі компоненттерді таңдау арқылы өңделген деректерді қалпына келтіреді (жуықтатады-аппроксимация).

Меншікті мәндердің қосындысы m -ге тең болатыны [56]-да сипатталған және келесі формула сәйкес негізгі компоненттердің дисперсиясының пайыздарын анықтау үшін пайдаланылуы мүмкін:

$$c_i = \frac{\lambda_i}{m} \cdot 100\% \quad (3.4)$$

мұнда, $c_i = \sum_{k=1}^i c_k$ - жиналған пайыздар, олар i -нші негізгі компоненттің үлесін көрсетеді.

λ_i - i -ші компоненттің меншікті мәні (дисперсия мөлшерінің өлшемі),
 m - құрамдас бөліктердің жалпы саны.

Осылайша, формула деректердің жалпы дисперсиясына әрбір негізгі компоненттің үлесін пайызбен бағалауға мүмкіндік береді.

Меншікті векторлардың уақытқа тәуелділігінің графиктері бойынша уақыттық қатардың үш компонентіне қандай негізгі құрамдас бөліктер жатқызылатыны анықталады:

- тренд (баяу өзгертін төмен жиіліктегі ауытқулар);
- периодтық компоненттер;
- шу (жоғары жиілікті құрамдас бөліктер).

3.3 Сингулярлық спектрлі талдау әдісімен қатарды ыдырату бағдарламасының сипаттамасы

Бұл тарауда 3.2 тараудағы көрсетілген математикалық процестердің бағдарламалық код түріндегі сипаттамасы беріледі.

Зерттелетін нақты өлшенген 18000 мәннен тұратын бастапқы уақыттық қатардан алдымен әрбір 10 с. сайынғы пакеттер санын жинақтау арқылы 1800 мәнге ие қатар алынды. Ол жинақтаудың бағдарламалық кодының бөлігі төменде көрсетілген:

```
i=1;x=VarName2;  
for j=1:(length(x))/10  
    y(j)=x(i)+x(i+1)+x(i+2);  
    i=i+10;  
end
```

Әрі қарай негізгі параметр L ($1 < L < 1800$) терезенің ұзындығы деп аталатын қандай да бір бүтін сан (N -нің жартысынан аспауы керек).

```
--  
>> N=length (X);  
>> L=900;  
>> t=(1:N);  
>> figure;  
plot (t,X);
```

Әрі қарай L өлшемді $K=N-L+1$ ендіру векторлары жасалады, нәтижеде траекторлы матрица қалыптасады:

Траектория матрицасы жылжымалы терезе арқылы қалыптасады, бұл кезде белгілі бір қадамы бар жолдың элементтері матрицаның бағандарына қосылып, бағандар саны терезенің ұзындығымен анықталады. Осылайша, матрицаның әрбір бағаны бастапқы уақыттық қатардың векторы болып табылады.

```

>> %Construction of trajectory matrix of the time series
Y=zeros(L,N-L+1); %trajectory matrix
for m=1:N-L+1
    Y(:,m) = X((1:L)+m-1);
end;
figure ;
set(gcf,'name','Trajectory matrix of the time series');
clf;
imagesc(Y); %trajectory matrix
title('Trajectory matrix of time series')
axis square
colorbar

```

Әрі қарай негізгі компоненттердің таңдамалы дисперсиялары болып табылатын матрицаның меншікті мәндері (собственные вектора) анықталады. Яғни, матрицаның сингулярлық мәнінің ыдырауы болады.

```

>> S=Y*Y'/(N-L+1) ;%calculation of matrix S
[V,LAMBDA] = eig(S);
LAMBDA = diag(LAMBDA);
[LAMBDA,ind]=sort(LAMBDA,'descend');
V = V(:,ind);
>> figure;
>> set(gcf,'name','Eigenvalues of the matrix S')
clf;
plot(LAMBDA,'o-');|
title('Eigenvalues of the the matrix S')
>> |

```

Әрі қарай қатарды негізгі компоненттерге ыдырату қажет, бағдарламаның бұл бөліктері төменде көрсетілген:

```

ylabel(sprintf('PC %d',900));
>> %Construction of Reconstructed Components
R=zeros(N,L);
for m=1:L
    inp=PC(:,m)*V(:,m)'; %invert projection
    inp=inp(end:-1:1,:);
    for n=1:N % anti-diagonal averaging
        R(n,m)=mean(diag(inp,-(N-L+1)+n));
    end
end;

```

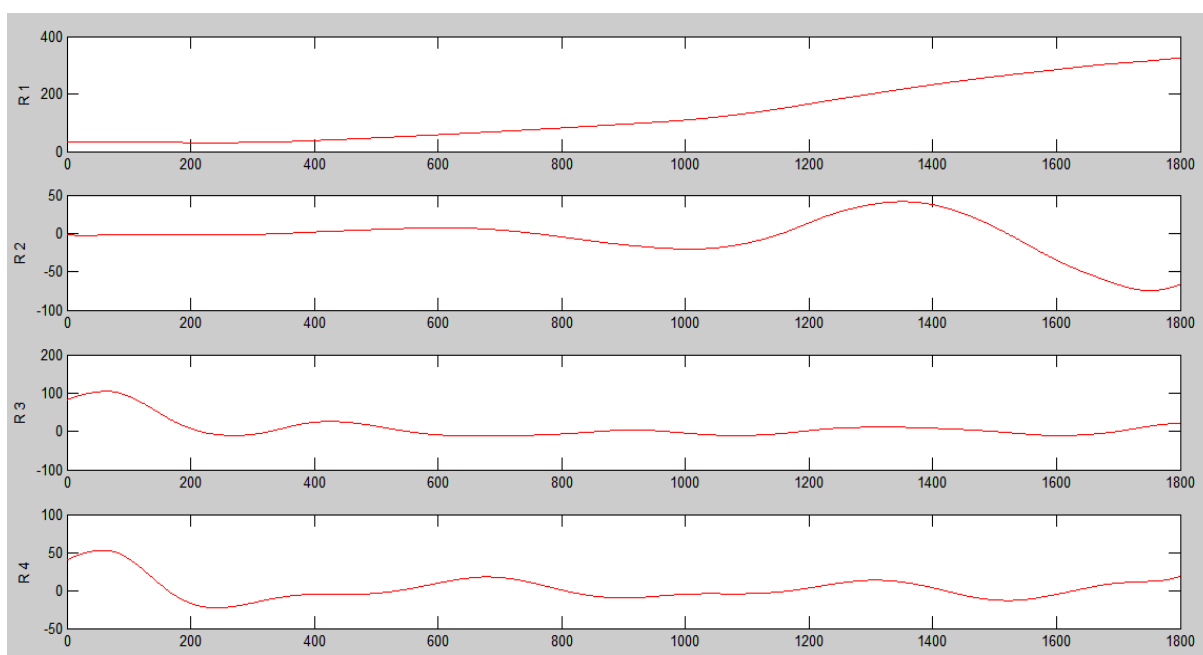
Әрі қарай алғашқы төрт меншікті векторды көру үшін график құру процедурасы орындалады:


```

>> %Plot first four reconstructed components
figure;
set(gcf, 'name', 'Reconstructed components')
clf;
subplot(4,1,1);
plot(t,R(1:N,1), 'r-');
ylabel(sprintf('R %d',1));
subplot(4,1,2);
plot(t,R(1:N,2), 'r-');
ylabel(sprintf('R %d',2));
subplot(4,1,3);
plot(t,R(1:N,3), 'r-');
ylabel(sprintf('R %d',3));
subplot(4,1,4);
plot(t,R(1:N,4), 'r-');
ylabel(sprintf('R %d',4));
fx >> |

```

Осылайша алынған графиктерді визуалды талдау арқылы сәйкес негізгі компоненттердің гармонияға немесе шуылға не трендке тиесілігін топтастыру процедурасы орындалады.

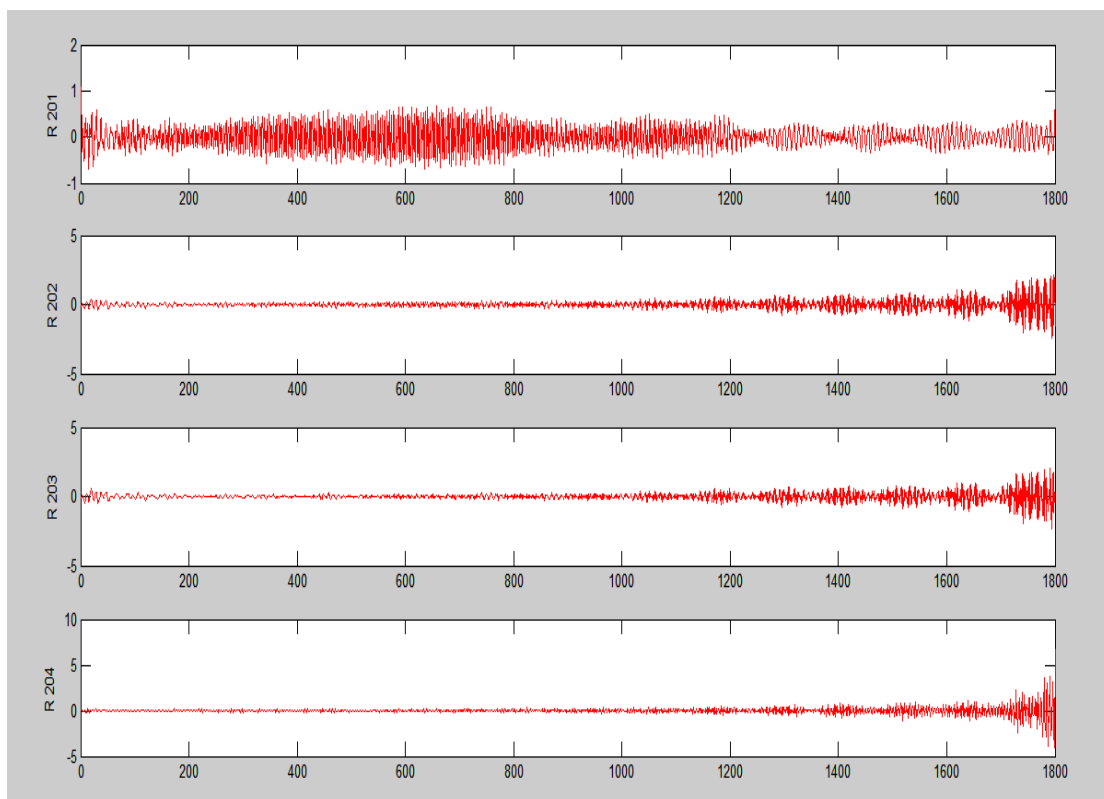


Сурет. 3.4 - 1-4 меншікті векторлар

Шуды ажырату үшін бірнеше ескертулерді қолдануға болады: сингулярлы векторлардың ретсіз әрекеті олардың шу құрамдас бөлігі тудырған жиынға жататынын көрсетуі мүмкін. Бұл сонымен қатар баяу, іс жүзінде секірусіз, меншікті мәндердің белгілі бір саннан төмендеуімен дәлелденуі мүмкін.

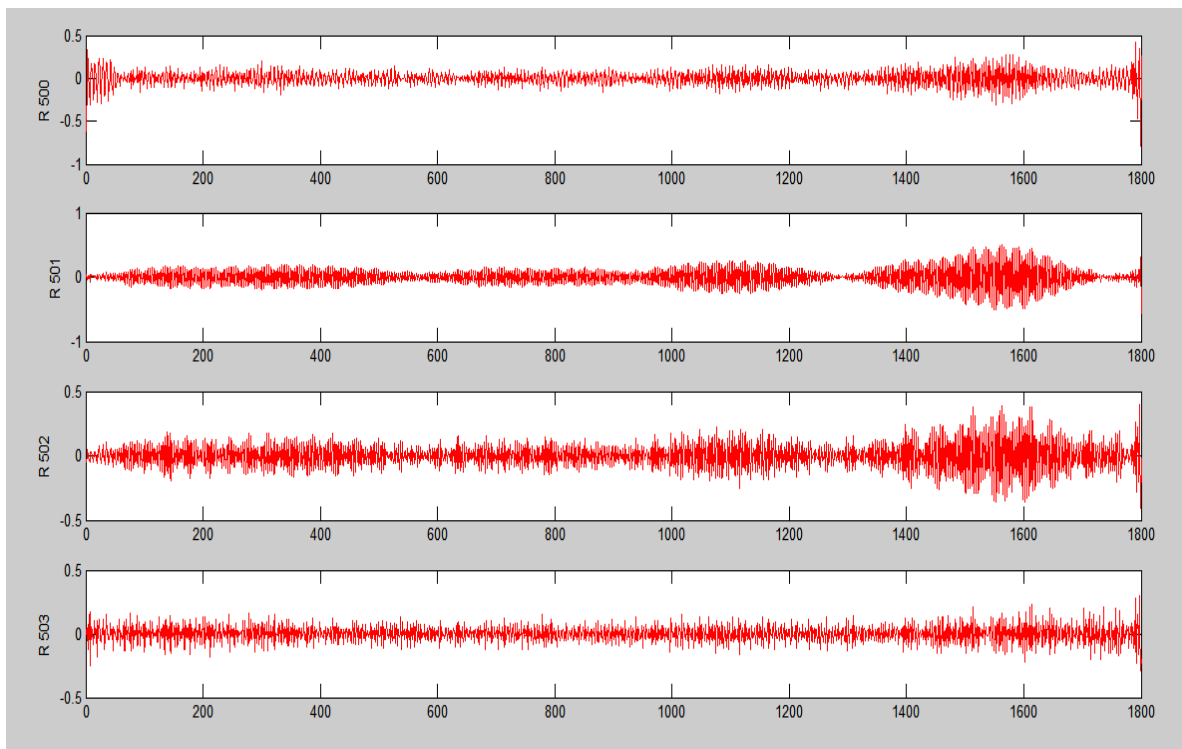
Меншікті векторлардың 3.4 суреттегі графикте бірінші жолдың графигі салыстырмалы түрде тегіс екенін көрсетеді, оны тренд компонентіне жатқызуға болады. Қалған үш графикке келетін болсақ, оларды мерзімді құрамдас бөліктерге жатқызуға болады.

3.5 сурет 201-ден 204-ке дейінгі негізгі компоненттерінің (векторларының) графиктерін көрсетеді. Бұл жағдайда векторларды мерзімді құраушыларға жатқызуға болады.



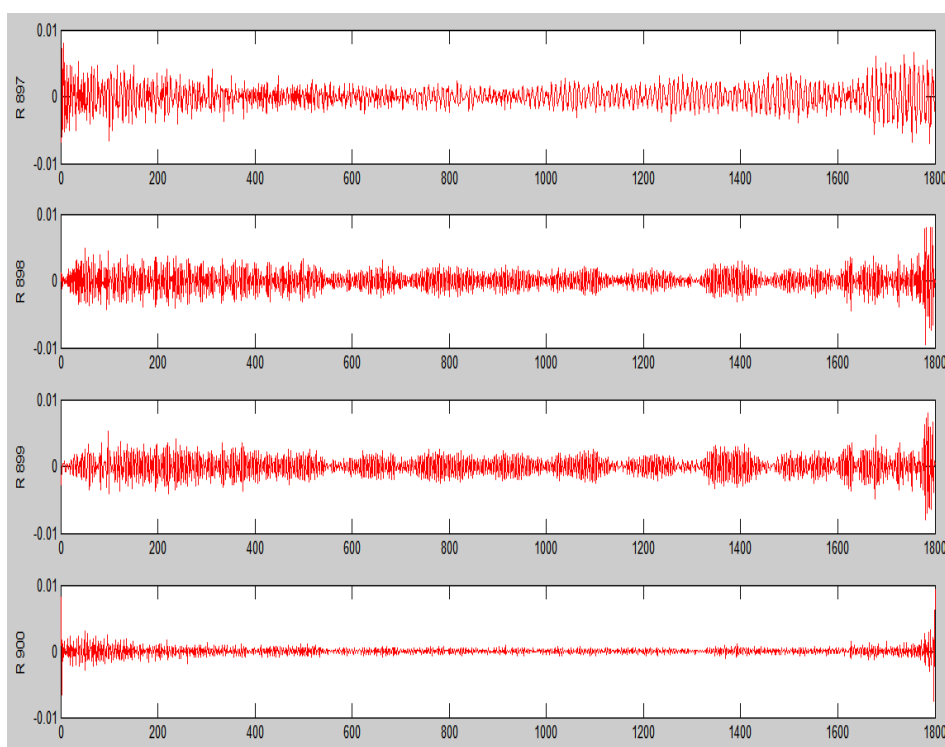
Сурет. 3.5 - 201–204 меншікті векторлар

3.6 суретте тиісінше 500-ден 503-ке дейінгі негізгі құрамдастардың графиктері көрсетілген, оларды мерзімді құраушыларға жатқызуға болады.



Сурет. 3.6 - 500–503 меншікті векторлар

3.7 суретте 897-900 аралығындағы меншікті векторлардың графиктері олардың құрамында жоғары жиілікті тербелістер бар екенін және шу құрамдарына жатқызу керек екенін көрсетеді.



Сурет 3.7 - 897-900 меншікті векторлар

3.3.1 Сингулярлық спектрлі талдау әдісімен ыдыраған қатарды қалпына келтіру

Қатарды қалпына келтіру негізгі компоненттердің әртүрлі комбинациялары арқылы жүзеге асырылады [57].

3.8 сурет сингулярлық спектрлік талдау арқылы өңдеуден кейін қатардың динамикасын қалпына келтіруге арналған бағдарлама терезесін көрсетеді.

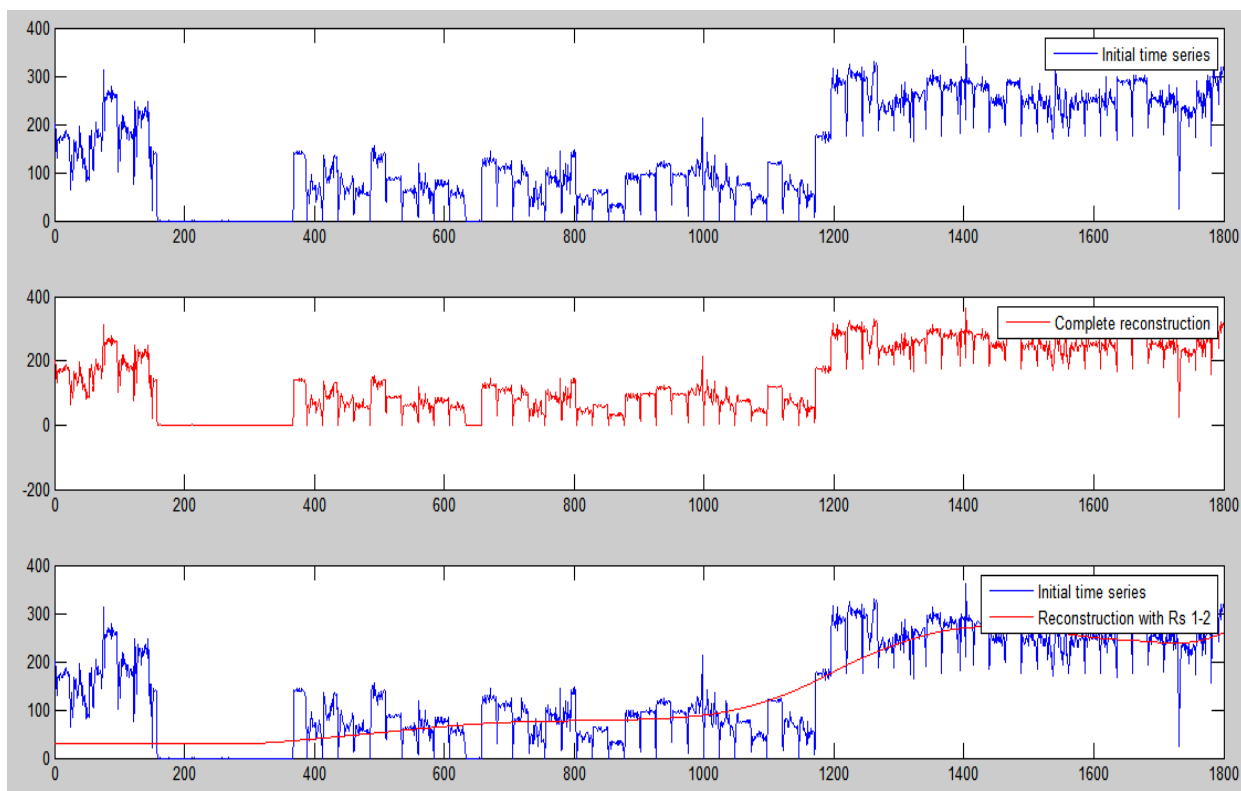
```
subplot(3,1,1)
plot(t(1:N),X,'b-');
legend('Initial time series');
subplot(3,1,2)
plot(t,sum(R(:,:),2),'r-');
legend('Complete reconstruction'); %sum of all reconstructed
components
subplot(3,1,3)
plot(t(1:N),X,'b-',t,sum(R(:,1:2),2),'r-');
legend('Initial time series','Reconstruction with Rs 1-2');
%comparison time series with sum of first two reconstructed components
Undefined function or variable 'components'.

>> %% Compare initial time series and reconstructed components
figure;
set(gcf,'name','Initial time serie X and Reconstruction R')
clf;
subplot(3,1,1)
plot(t(1:N),X,'b-');
legend('Initial time series');
subplot(3,1,2)
plot(t,sum(R(:,:),2),'r-');
legend('Complete reconstruction'); %sum of all reconstructed components
subplot(3,1,3)
plot(t(1:N),X,'b-',t,sum(R(:,1:2),2),'r-');
legend('Initial time series','Reconstruction with Rs 1-2');
%comparison time series with sum of first two reconstructed components|
fx >>
```

Сурет. 3.8 - Қатардың динамикасын қалпына келтіруге арналған бағдарлама

3.9 сурет үш уақыттық қатарды көрсетеді: көк түспен бастапқы қатар, өңдеуден кейін қалпына келтірілген қатар қызыл түспен және жуықтау функциясы бар бастапқы қатар сызылған.

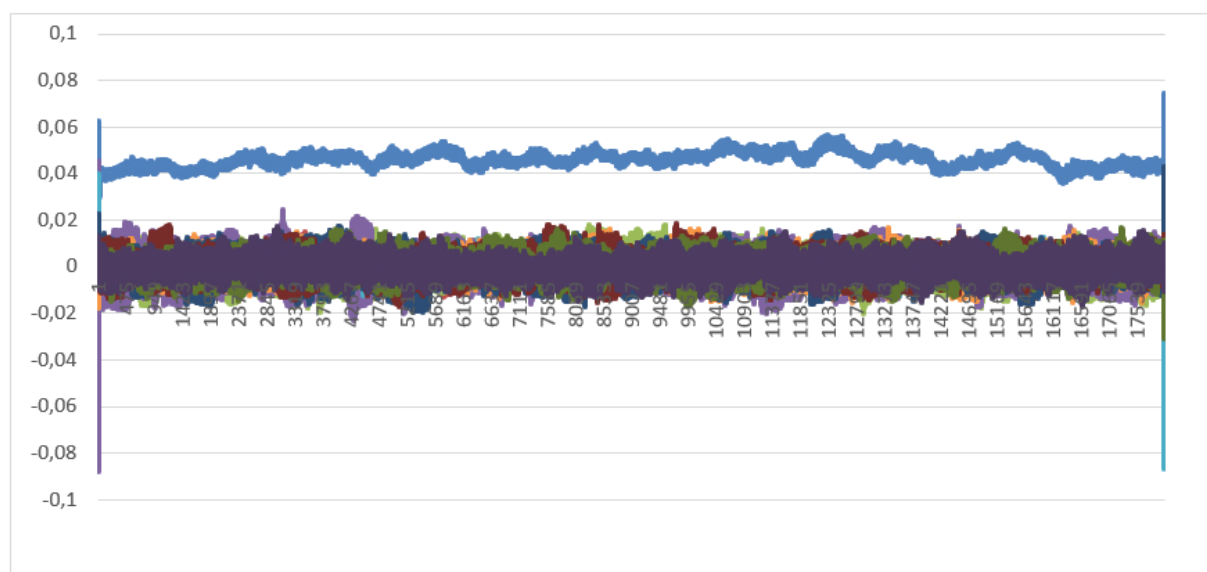
Қайта қалпына келтірілген қатарды болжамдық модельдерді әзірлеу үшін пайдалануға болады, себебі ол болашақта бастапқы қатардың әрекетіне әсер етуі мүмкін әртүрлі құрамдас бөліктер туралы ақпаратты қамтиды. Мысалы, шуға немесе кездейсоқ ауытқуларға қатысты құрамдастарды жою арқылы қалпына келтірілген қатарды бастапқы қатардың тегістелген нұсқасын алу үшін пайдалануға болады.



Сурет 3.9 - Бастапқы қатар, қайта құрылған қатар және жуықтау функциясы бар бастапқы қатар

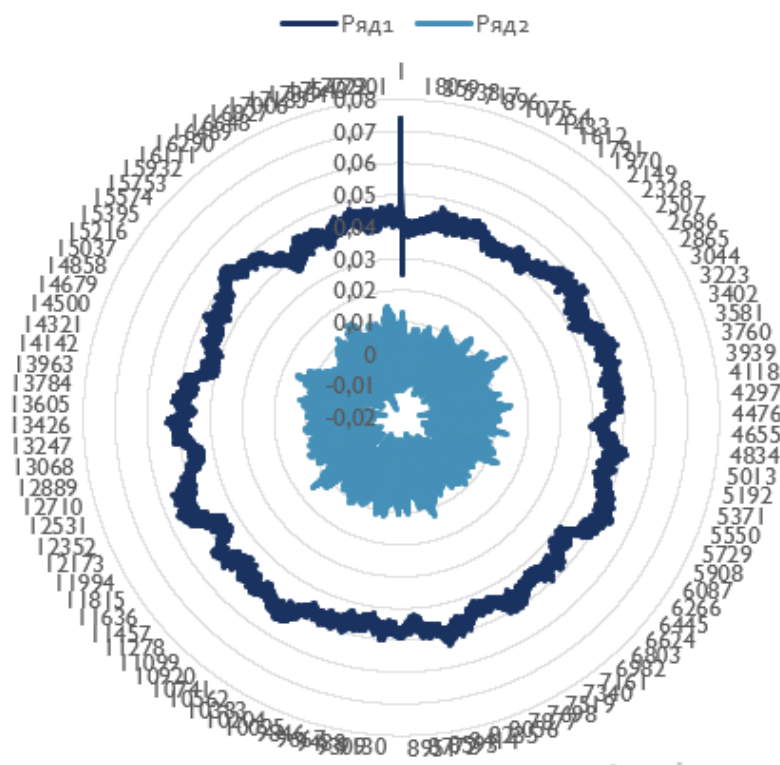
Әрі қарай, әрбір 100 секунд сайын 100 деңгейі бар жинақталған бастапқы уақыттық қатар MS Excel электрондық кесте процессорының қосымшасы ретінде жасалған AtteStat [58] бағдарламасында қарастырылды. Сингулярлық спектрлік талдауға параметрлер келесідей орнатылды - терезе ені 50 және гармоника саны да 50.

3.10, 3.11 суреттерде уақыттық қатардың құрамдастарға ыдырауы көрсетілген.



Сурет 3.10 - Уақыттық қатардың ыдырауы

Уақыттық қатардың ыдырауы оның стационарлық емес екенін көрсетеді. Онда бірінші негізгі компонент - тренд (1-қатар), гармоника және шу бар.



Сурет 3.11 - Уақыттық қатардың ыдырауы

3.4 Үшінші тарау бойынша қорытындылар

1. Классикалық Фурье талдауы уақыттық қатар деректеріндегі стационарлы сигналдарды анықтаудың теңдесі жоқ құралы болып табылады және стационарлы емес сигналдарды зерттеуде қауқарсыз;
2. Толқынды түрлендіру (Вейвлет-преобразование) немесе вейвлеттер сигналдарды интерпретациялауда маңызды, Вейвлеттер тез ыдырайтын минитолқын түріндегі толқынның қысқа жарылыстары;
3. Толқындық талдау – көптеген уақыт шкалаларында сигнал қуатының өзгеруін көрсететін біркелкі емес уақыт қатарын талдауға арналған қуатты математикалық құрал;
4. Уақыттық қатарды ыдырату әдісі (SSA) қатарды тренд, мерзімділік және шу ретінде түсіндіретін жиынтық құрамдас бөліктерге ыдыратады;
5. Алынған құрамдас бөліктер талданады және деректерде болжау, шуды сүзу және аномалияны анықтау мәселелерін шешу үшін пайдаланылады;
6. Уақыттық жиілікті деректерді талдау (немесе спектрлік талдау) уақыттық қатардың жиілік құрамдастарын ерекшелеуге және олардың жалпы деректер құрылымына қосқан үлесін анықтауға мүмкіндік береді.
7. Уақыттық жиілікті деректерді талдау қатардың болашақ мәндерін болжау немесе болжамдық интервалдың оңтайлы ұзақтығын анықтау үшін пайдаланылуы мүмкін.

4 ЖЕЛІЛІК ТРАФИКТІ БОЛЖАУ МОДЕЛЬДЕРІН ЖАСАУ

4.1 Стационарлы емес уақыттық қатарларды болжаудың ерекшеліктері

Бүгінгі күні телекоммуникацияны дамытудың негізгі тенденциясы конвергентті келесі буын желілеріне (NGN) көшу болып табылады. Алдыңғы қатарда бірыңғай телекоммуникациялық инфрақұрылым шеңберінде әртүрлі сипаттамалар мен талаптардағы қызметтерге қызмет көрсету сапасын қамтамасыз ету мәселесі тұр. Әртүрлі деңгейдегі телекоммуникациялық желілердің жұмыс істеу сапасын арттыру мәселесіне жүйелі көзқараспен трафикті басқарудың тиімді алгоритмдерінсіз мұны істеу мүмкін емес. Осыған байланысты желідегі трафик тәртібінің адекватты моделін әзірлеу маңызды міндет болып табылады. Көптеген зерттеулердің нәтижелері [59-62] қазіргі телекоммуникациялық желілердің стационарлық қасиеттерге және, атап айтқанда, біркелкі емес қасиетіне ие екенін көрсетті. Белгілі болғандай, мұндай процестерді болжауға заманауи тәсілдер мүмкіндік береді [63-67].

Стационарлық емес уақыт қатарын болжаудың негізгі әдістеріне мыналар жатады:

- статистикалық әдістер: ARIMA (Авто-регрессивті айырымды жылжымалы орташа), Экспоненциалды тегістеу әдісі, SSA (Сингулярлы спектрлі талдау).
- жасанды интеллектке (ЖИ) негізделген жаңа әдістер:
 - ANN (Artificial Neural Network) – жасанды нейрондық желілер;
 - ANFIS (Adaptive Neuro-Fuzzy Inference System) – анық емес қорытындылар жүйесіне негізделген адаптивті желі;
 - Fuzzy Logic – анық емес логика;
 - SVM (Support Vector Machine) – машиналық оқыту әдісі - тірек векторлық машина;
 - Neuro-Fuzzy Network – нейро-анық емес желілер;
 - Evolutionary Optimization Algorithm - эволюциялық оңтайландыру алгоритмдері;
 - кластерлеу әдістері.

ARIMA (Авто-регрессивті айырымды жылжымалы орташа) әдісі бастапқы уақыттық қатардан кейбір реттік айырмашылықтарды алу арқылы стационарлық қатарға келтірілетін стационарлық емес уақыт қатарларын сипаттауға мүмкіндік береді.

Экспоненциалды тегістеу болашақ мәндерді болжау үшін қатардың алдыңғы мәндерін пайдаланатын әдіс. Ол деректерді тегістеу және алдыңғы мәндер үшін салмақ орнату арқылы жұмыс істейді. Бұл әдіс әсіресе трендтері мен маусымдылығы бар деректер үшін пайдалы.

SSA (Сингулярлы спектрлі талдау) – уақыттық қатарды оның құрамдас бөліктеріне бөлетін және оларды болашақ мәндерді болжау үшін пайдаланатын

әдіс. Ол қатардың негізгі құрамдастарын шығару үшін сингулярлық мәннің ыдырауын пайдаланады, содан кейін оны болжау үшін пайдалануға болады.

Нейрондық желілер – уақыттық қатарларды болжау үшін жасанды нейрондық желілерді пайдаланатын әдіс. Нейрондық желілер деректердің үлкен көлемін өңдей алады және олардың арасындағы күрделі қатынастарды таба алады, бұл *нейрондық желілер* әдісін күрделі стационарлық емес қатарларды болжау үшін пайдалы етеді.

ANFIS – бұл жасанды нейрондық желілер мен анық емес логиканың артықшылықтарын біріктіретін гибриді жүйе. Ол қолда бар деректер негізінде күрделі жүйелерді модельдеу және болжау үшін қолданылады.

ANFIS негізгі функциялар мен ережелерді дайындаудан басталады, содан кейін олар анық емес ережелерді құру үшін пайдаланылады. Одан соң ол нақты болжау нәтижесін шығару үшін қолда бар деректер негізінде ережелердің коэффициенттерін оңтайландыратын оқыту алгоритмін пайдаланады.

Болжау кезінде *ANFIS* кірістер мен шығыстар арасындағы қатынастарды анықтау үшін анық емес ережелерді пайдаланады. Бұл ережелерді қол жетімді деректер негізінде болашақ мәндерді болжай алатын анық емес модель жасау үшін пайдалануға болады.

Осылайша, *ANFIS* жасанды нейрондық желілер мен анық емес логиканың артықшылықтарын біріктіретін гибриді жүйе болып табылады және қол жетімді деректер негізінде күрделі жүйелерді модельдеу және болжау үшін пайдаланылуы мүмкін.

Анық емес логика анық емес деректерді талдау және анық емес шешімдер қабылдау үшін қолданылады. Ол ақиқат және жалған арасында қандай да бір аралық деңгей бар деген болжамға негізделген, оны «ішінара ақиқат», «жартылай жалған», «өте дұрыс», «төмен шындық» сияқты терминдер арқылы сипаттауға болады. Мәлімдемелері тек ақиқат немесе жалған болуы мүмкін классикалық логикаға қарағанда, анық емес логика мәндердің кең ауқымымен жұмыс істеуге мүмкіндік береді.

Анық емес логика *мүшелік функциялары* арқылы анықталған анық емес жиындарды пайдаланады. Мүшелік функциясы элементтің берілген жиынға қалай жататынын, яғни қандай дәрежеде жататынын анықтайды. Әдетте мүшелік функциялар қисық функция немесе 0-ден 1-ге дейін бірқалыпты өзгеретін функциялар түрінде болады.

Анық емес деректермен жұмыс істеу үшін *Fuzzy Logic* бірнеше кадамдарды қамтитын анық емес қорытынды жасау процесін пайдаланады:

1. Анық емес кіріс мәні мүшелік функциясының көмегімен анық емес жиынға түрлендіріледі;
2. Берілген анық емес жиын үшін ережелер анықталған болса, онда шығыс анық емес жиын осы ережелер негізінде құрастырылады;
3. Шығарылатын анық емес жиын шығыс функциясының көмегімен сандық мәнге түрлендіріледі.

Машиналық оқыту әдістерін: градиентті арттыру және кездейсоқ орман сияқты әдістерді стационарлық емес уақыт қатарын болжау үшін қолдануға болады. Олар деректер арасындағы тәуелділіктерді модельдеу және осы тәуелділіктер негізінде болжамды модельдерді құру үшін пайдаланылады.

Машиналық оқыту әдісі *SVM (тірек векторлық машина)* – желілік трафикті болжау үшін тарихи трафик мәндері мен сәйкес уақыт белгілерінен тұратын деректер жинағын қажет етеді. Содан кейін деректер оқу және сынақ жинақтарына бөлініп, болжау үшін бір класты жіктеу әдісін пайдаланады. Бұл SVM трафиктің тарихи деректерінен үйренетінін және ағымдағы трафик мәнінің аномальды екенін анықтай алатын модель құрастыратынын білдіреді. Ол үшін тірек векторлық машина қалыпты трафик мәндерін қалыпты емес мәндерден мүмкіндігінше бөлетін гипержазықтық қалыптастырады.

Желілік трафикті болжауға SVM-ді пайдалану үшін ядро параметрлері және терезе ені сияқты модель параметрлерін дұрыс таңдау маңызды болып табылады. Бұл параметрлер болжау дәлдігіне айтарлықтай әсер етуі мүмкін, сондықтан модельдің өнімділігіне қол жеткізу үшін мұқият талдау қажет.

Neuro-Fuzzy Network – нейро-анық емес желілер жасанды нейрондық желілер мен анық емес логиканың қосындысы болып табылады. Ол деректерді болжау және жіктеу (классификация) үшін қолданылады. Негізгі идея нейрондық желілердің және анық емес логиканың артықшылықтарын болжау мен жіктеудің күрделі мәселелерін шешу үшін қолданылатын бір жүйеге біріктіру болып табылады.

Нейро - анық емес желінің жұмыс істеу процесі анық емес ережелерді қолдану арқылы енгізілген деректерді өңдеуден басталады. Содан кейін кірістер мен шығыстар арасындағы байланыстарды анықтайтын жүйені оқыту үшін нейрондық желілер қолданылады. Оқыту нәтижесінде оқыту процесінде пайдаланылмаған жаңа деректерді болжау немесе жіктеу үшін жүйелерді пайдалануға болады.

Чен және Ченг алгоритмдері нейро - анық емес желілерді оқыту әдістеріне жатады және болжау модельдерін құру үшін қолданылады. Екі алгоритм де модель параметрлерін жаңартудың итерациялық процесіне негізделген. Чен алгоритмі ANFIS сияқты нейро-анық емес желілерді үйретуге арналған.

Evolutionary Optimization Algorithm – (Эволюциялық оңтайландыру алгоритмі EOA) - оңтайландыру мәселелерін шешу үшін эволюциялық биология тәсілін қолданатын оңтайландыру алгоритмдерінің класы.

Уақыттық қатарларды болжауда эволюциялық алгоритмдерді нейрондық желілер немесе ARIMA модельдері үшін оңтайлы гиперпараметр мәндері сияқты модель параметрлерін таңдау үшін пайдалануға болады.

Осындай болжамдарды пайдалана отырып, трафикті басқарудың тиімді алгоритмдерін жасауға болады.

Тәжірибелік тұрғыдан алғанда, қызмет көрсету сапасын қамтамасыз ету үшін тиісті шараларды қабылдау үшін мүмкін болатын ең жоғары жүктеменің шамасын және оның пайда болу уақытын болжау маңызды.

4.2 ARIMA әдісімен болжау

Стационарлы емес уақыт қатарларын модельдеу үшін жиі қолданылатын әдіс параметрлік бағалау болып табылады. Бұл жағдайда тренд компоненті үшін қандай да бір функционалдық тәуелділіктің параметрлері таңдалады, оларды алып тастағаннан кейін стационарлық қатар қалады. Осы мақсатта әртүрлі стационарлық сынақтар (Дикки-Фуллер тесті (ADF), Филлипс-Перрон тесті (PP-test)) қолданылады, олар әдетте белгілі функционалдық тәуелділіктерді қолдануға арналған [68-70].

Жалпы статистикалық уақыттық қатарлардың моделі:

$$y_t = f(x_t, a) + \varepsilon_t \quad (4.1)$$

Стационарлық уақыттық қатарлардың негізгі сызықтық үлгілеріне мыналар жатады:

- авторегрессиялық модельдер;
- жылжымалы орташа модельдер;
- авторегрессивті және жылжымалы орташа модельдер.

Қалдық қатарының стационарлығын тексеру және оның дисперсиясын бағалау үшін келесілер пайдаланылады:

- таңдамалы автокорреляция функциясы (АКФ);
- жеке автокорреляция функциясы (ЧАКФ).

Егер уақыттық қатар стационарлы болмаса, онда бұл уақыт қатарында не сызықтық тренд, не қалдықтардың гетероскедастикасы, не қалдықтардың автокорреляциясы, не периодтық құраушы немесе барлығы да бірдей бар екенін білдіреді.

Дикки-Фуллер тестінде (нөлдік) гипотеза зерттелетін қатар DS класына жатады, ал альтернативті гипотеза зерттелетін қатар TS класына жататыны тексеріледі. «Интеграция» термині стационарлық емес уақыт қатарын стационарлы жасауды білдіреді. Стационарлықтың орындалуы $I(d)$ белгіленеді. $I(0)$ – стационарлы қатар. Негізінде интеграция процедурасы айырмашылықтарды алу арқылы жүзеге асырылады [71-73].

Интеграцияның негізгі әдістеріне мыналар жатады:

- шекті айырмашылықтарды алу. Бастапқы қатар y_1, y_2, \dots, y_t стационарлы емес. Онда жаңа x_1, x_2, \dots, x_{n-1} қатар құрамыз, мұндағы $x_t = y_t - y_{t-1}$. Егер бұл қатар стационарлық шарттарды қанағаттандырса, онда бастапқы қатар $I(1)$ деп белгіленеді және бастапқы заң сызықтыққа жақын деген қорытынды жасалады. Әйтпесе, келесі жолға ораламыз: Жаңа қатар құрылады z_1, z_2, \dots, z_{n-1} , мұнда $z_t = x_t - x_{t-1} = y_t - 2y_{t-1} + y_{t-2}$. Бастапқы қатар $I(2)$ деп белгіленіп, бастапқы заң квадраттыққа жақын деген қорытынды жасалады. Демек, интегралдау дәрежесі зерттелетін уақыт қатарының стационарға айналуындағы айырмашылық ретін көрсетеді. $\Delta^d y_t$ жазуы айырмашылық операторы қалай қолданылатынын оқиды, мысалы $\Delta^2 y_t$;

ARMA (Авто-регрессивті жылжымалы орташа) моделі стационарлық қатарларға арналған.

[74]-те өткен ғасырдың ортасында Бокс-Дженкинс (BJ) жасаған модельдің негізі болып табылатын стационарлық емес уақыттық қатарларды талдау әдісі сипатталған. Бұл әдістің мәні мынада: сызықтық әдістерді қолдану арқылы стационарлық емес қатарды стационарлық қатарға келтіруге болады. Бұл модель ARIMA (Авто-регрессивті айырмалы жылжымалы орташа) моделі деп аталады. у уақыттық қатарды қарастырайық:

$$(y_1, y_2, \dots, y_n) \quad (4.2)$$

Айырмалардан тұратын екінші қатарды құрастырамыз:

$$(y_2 - y_1, y_3 - y_2, \dots, y_n - y_{n-1}) \quad (4.3)$$

$$\text{жалпы түрі: } \Delta y_t = y_t - y_{t-1} \quad (4.4)$$

Мұндағы Δ – айырма операторы.

Шекті айырмалар есебі (Исчисление конечных разностей) формула бойынша туындыға (дифференцирование) ұқсас. Ал бастапқы қатарға келу үшін келесідей тәсілді қолданамыз:

$$(y_1 + y_2 - y_1, y_2 + y_3 - y_2, \dots, y_{n-1} + y_n - y_{n-1}) \quad (4.5)$$

Шекті айырмалар есебіндегі осы кері процесс айырмалау (интегрирование) деп аталады. ARIMA сөзіндегі I параметрі осы айырма алу, яғни инеграциялауды білдіреді.

$$y_t = x_t + a + bt \quad (4.6)$$

мұндағы x_t - стационарлы қатар, ал $a + bt$ – сызықты тренд. Осы қатарға айырма алу процедурасын қолданайық.

$$y_t - y_{t-1} = x_t - x_{t-1} + (a + bt) - (a + b(t-1)) = x_t - x_{t-1} + b \quad (4.7)$$

Туындыларды алу сияқты, бірінші айырманы алу қатарды стационарлы етті, себебі b – тұрақты сан.

Егер тренд квадраттық болса:

$$y_t = x_t + at^2 + bt + c \quad (4.8)$$

$$\Delta^2 y_t = (1 - L)^2 y_t = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2} \quad (4.9)$$

мұндағы L – уақыт қатарының мәндерін бір период артқа жылжытатын *лаг* операторы. Екінші айырманы алу квадраттық трендті стационарлы етеді және жоғары дәрежелі көпмүшеліктер үшін осылай жалғаса береді.

Демек айырмалар алу әдісін (бір немесе бірнеше, әдетте 2-ден аспайды) қолдану арқылы стационарлық емес қатарды стационарлық қатарға келтіруге болады.

Бұл ARIMA әдісін ойлап тапқан Бокс және Дженкинс ұсынған классикалық әдіс болып табылады.

Бокс-Дженкинс (BJ) немесе ARIMA(p,d,q) болжау әдісі модельдің үш бөлігінен тұрады [75];

- AR - авторрегрессияны сипаттайтын уақыттық қатарлар моделінің бөлігі, онда уақыттық қатардың қазіргі мәндері сол қатардың алдыңғы мәндерінің сызықтық комбинациясы және ақ шу қасиетіне ие кездейсоқ қате ретінде көрсетілуі мүмкін (p параметрі);

- I – қатардың айырма ретін сипаттайтын уақыттық қатар моделінің бөлігі (d параметрі);

- MA – қатардың ағымдағы мәнін сипаттайтын және оның қасиеттерінде «ақ шуға» сәйкес келетін ағымдағы және өткен қате мәндерінің сызықтық комбинациясы ретінде ұсынылатын уақыттық қатар моделінің бөлігі (q параметрі).

Айырма алудан бөлек трендті жою әдісі де бар. Қайсысы жақсы деген сұрақ - нақты тапсырмаға және бастапқы уақыттық қатардың қасиеттеріне байланысты. Бастапқы қатарда детерминделген тенденция болса, онда трендті жою әдісін қолданған дұрыс. Егер бастапқы қатарда кездейсоқ құрамдас болса, онда айырмашылықтарды алуды пайдалану тиімдірек болуы мүмкін. Бұл жағдайда біз бастапқы қатардағы автокорреляциядан құтыла аламыз, бұл бізге стационарлық талдау жүргізуге де мүмкіндік береді.

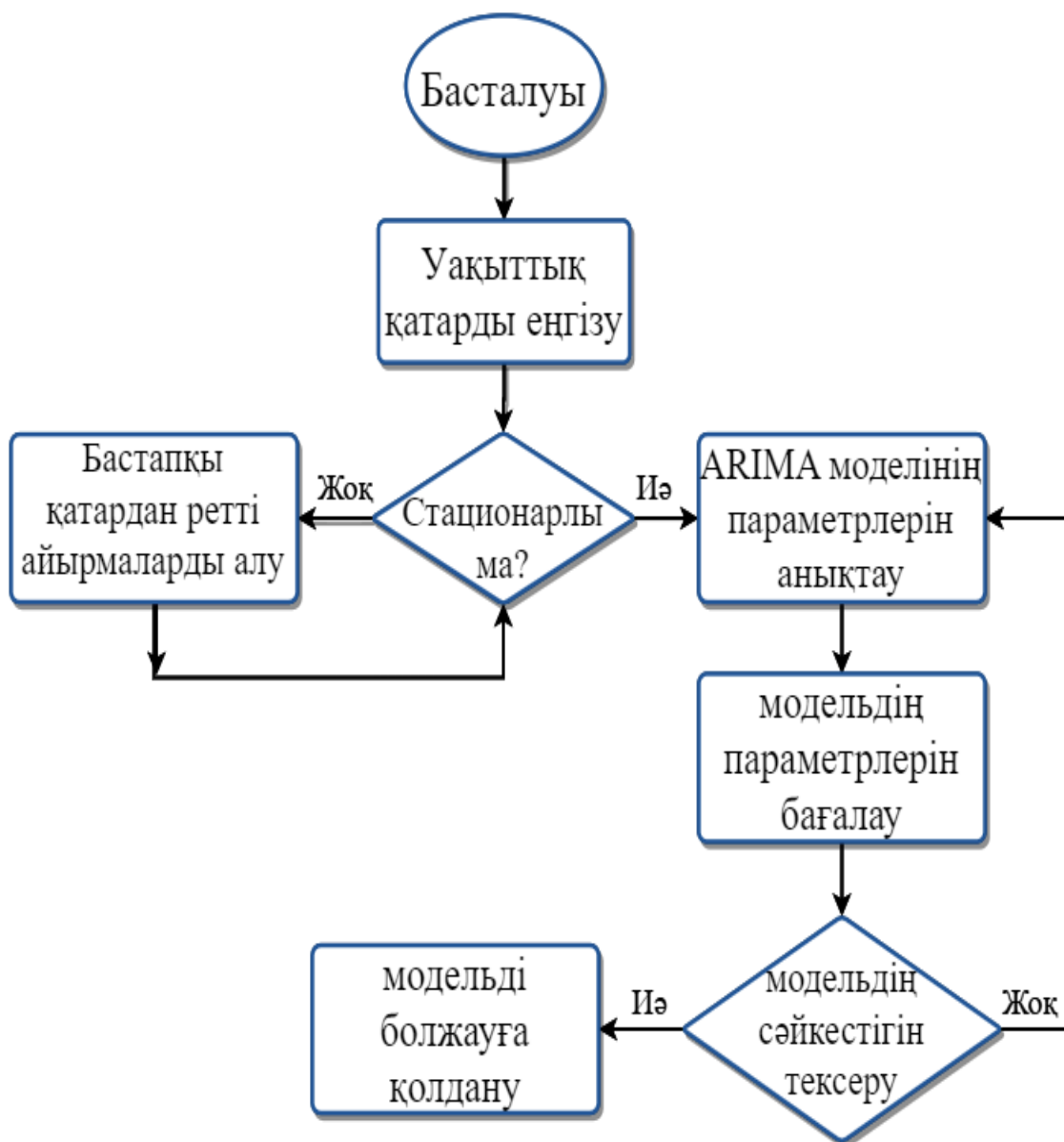
Жалпы әдісті таңдау уақыттық қатарлардың нақты қасиеттеріне және талдау мақсаттарына байланысты.

Моделдерді құрудың бірінші қадамы - деректерден стационарлы болғанша айырма алу.

I – қатардың айырма ретін сипаттайтын уақыттық қатар моделінің бөлігі, жоғарыда аталғандай d параметрі 2-ге тең.

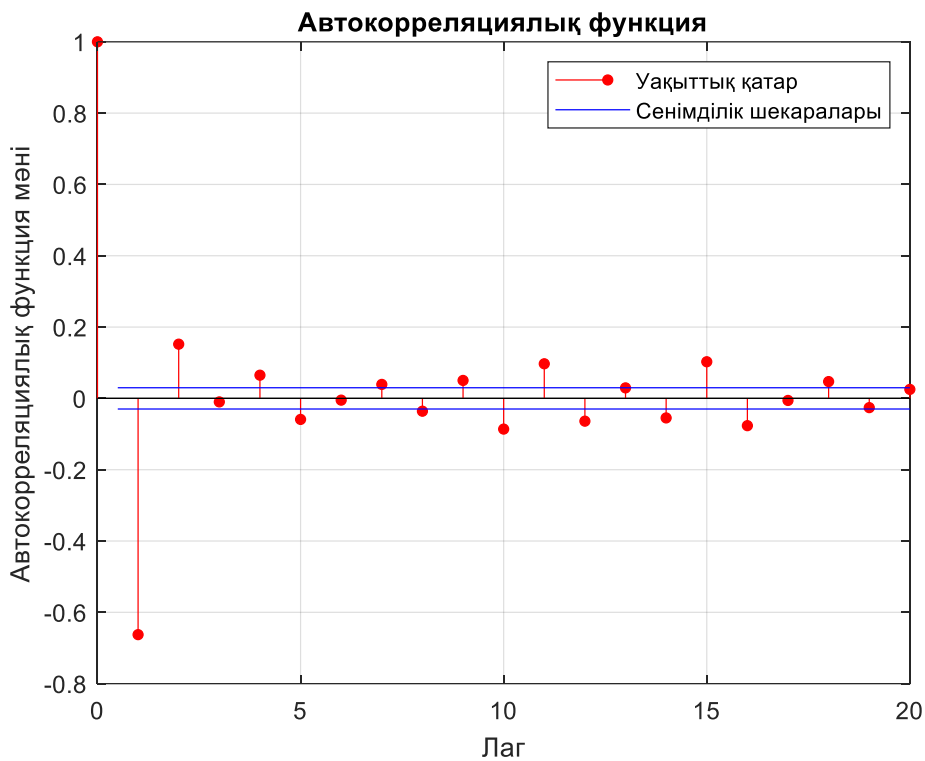
Сондай ақ бұл параметрді Matlab бағдарламасының Econometric Toolbox қосымшасы арқылы алынған графиктер де растайды. Бұл қосымша деректерді түрлендіруді, деректерді визуализациялауды, гипотезаны сынауды, стационарлық тестілеуді, ARIMA үлгілерін жасауды және т.б. қоса, уақыттық қатарларды талдау мен модельдеуді интерактивті түрде орындайды.

ARIMA моделін құру алгоритмі 4.1 суретте берілген.

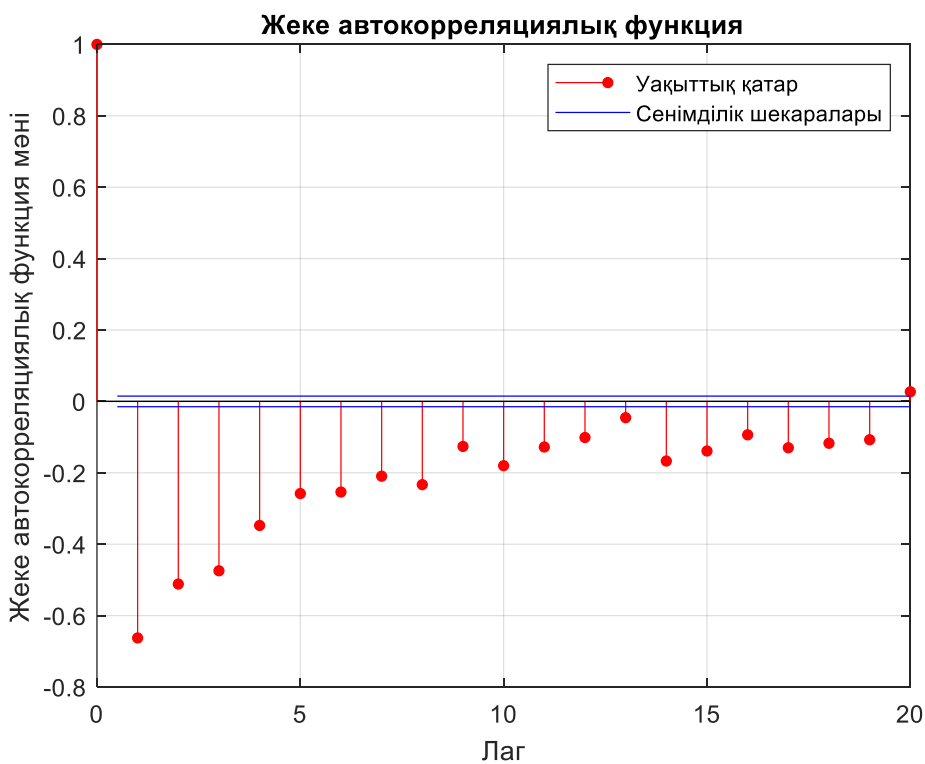


Сурет 4.1 - ARIMA моделін құру алгоритмі

Әрі қарай АКФ және ЖАКФ графиктерінің көмегімен ARIMA моделінің параметрлерін анықтауға болады.



Сурет 4.2 - АКФ графигі



Сурет 4.3 - Жеке автокорреляция графигі

Екінші ретті айырманың ACF және PACF графиктері бойынша d параметрі екіге тең (екінші айырма). Авторегрессия (AR) процесі бойынша ACF деңгейлері бірден кеміді, ал PACF деңгейлері біртіндеп кеміді. Авторегрессивті процесс болған жағдайда ACF функциясы баяу кемитін еді, сондықтан p параметрі нөлге тең (Авторегрессивті процесс жоқ).

Бұл автокорреляциялық функция (ACF) сызбасындағы авторегрессивті (AR) процесс үшін әртүрлі кешігулердегі қатар мәндерінің арасындағы корреляциялық мәндер тез төмендейді және нөлге дейін ыдырайды, ал жеке автокорреляциялық функцияда (PACF) олар баяу төмендейді және бірнеше лагтар үшін маңызды болып қалуы мүмкін.

Автокорреляциялық функция модулінің максимумдары лагтардың, яғни кездейсоқ шамалардың жасырын тәуелділігі көрінетін уақыт аралықтары болуын көрсетеді.

Жылжымалы орташа мән (MA) процесі бойынша ACF бірінші кідірістен (лаг) кейін күрт кемиді (соңғы маңызды кідіріс q параметрін көрсетеді). Демек, q параметрі бірге тең.

Модельді орнату процедурасы күрделі моделдерде сәйкестікті азайтуға көмектесетін Akaike ақпараттық критерийін (AIC) және Байес ақпарат критерийін (BIC) азайтатын параметрлерді табуға негізделген.

Нәтижесінде ARIMA(0,2,1) моделі (ARIMA_X2DiffDiff) алынды.

Кесте 2 - Бағалау нәтижелері

Parameter	Value	Standard Error	TStatistic	PValue
Constant	0	0		
MA{1}	-0,3452	0,016298	-21,1803	1,4509e-99
MA{2}	-0,32033	0,018528	-17,2888	5,7169e-67
Variance	698,8574	10,5872	66,0097	0

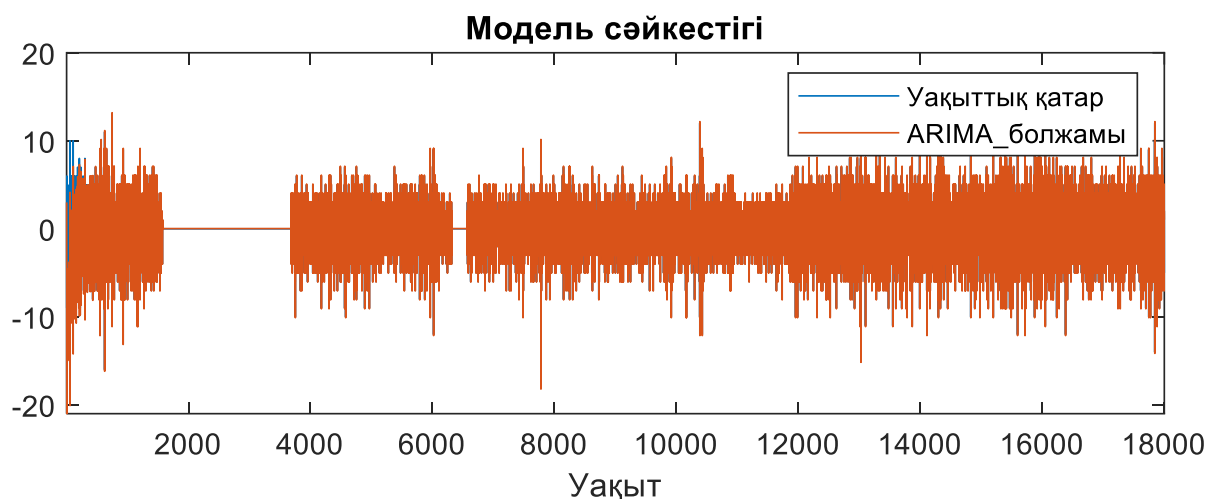
Жылжымалы орташаның MA{1} құрамдас бөлігінің PValue мәні 0,05 ($0 < 0,05$) мәнділік деңгейінен аз, сондықтан жылжымалы орташа мәннің коэффициенті статистикалық маңызды параметр деп қорытынды жасауға болады.

Кесте 3 - Ақпараттық критерийлер

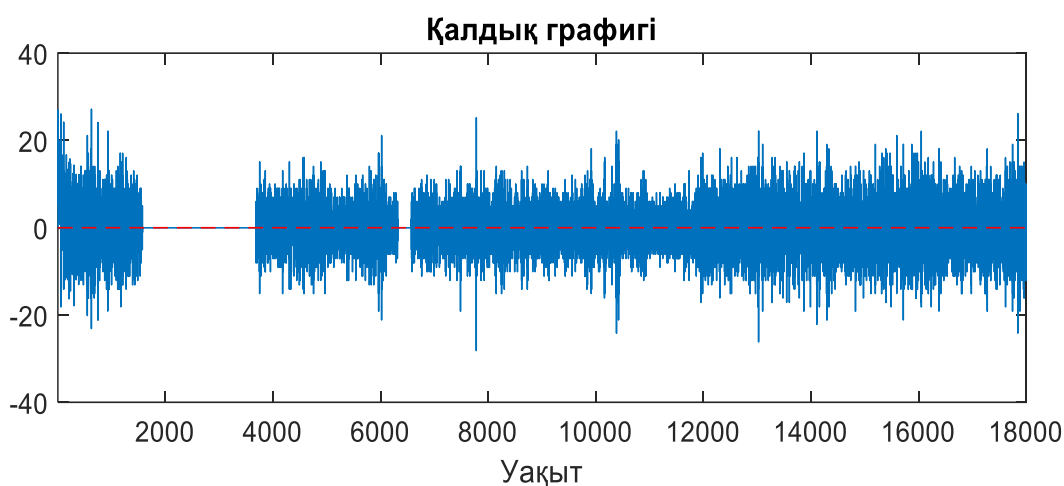
AIC	14715,9358
BIC	14732,0047

AIC және BIC критерийлері моделдің дәлдігі мен күрделілігі арасындағы кейбір айырбасқа модельдің сәйкестік дәрежесін көрсетеді.

Бастапқы қатар мен болжамды деректердің біріктірілген графигі 4.4 суретте көрсетілген және модельдің дұрыс таңдалғанын көрсетеді.



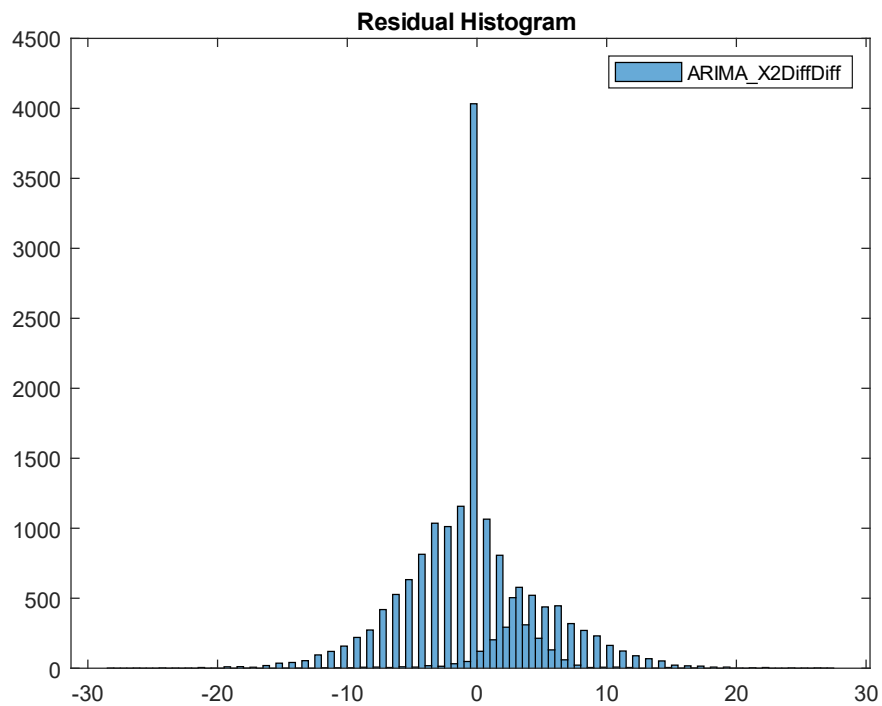
Сурет 4.4 – Бастапқы қатар мен болжамды деректердің сәйкестік графигі



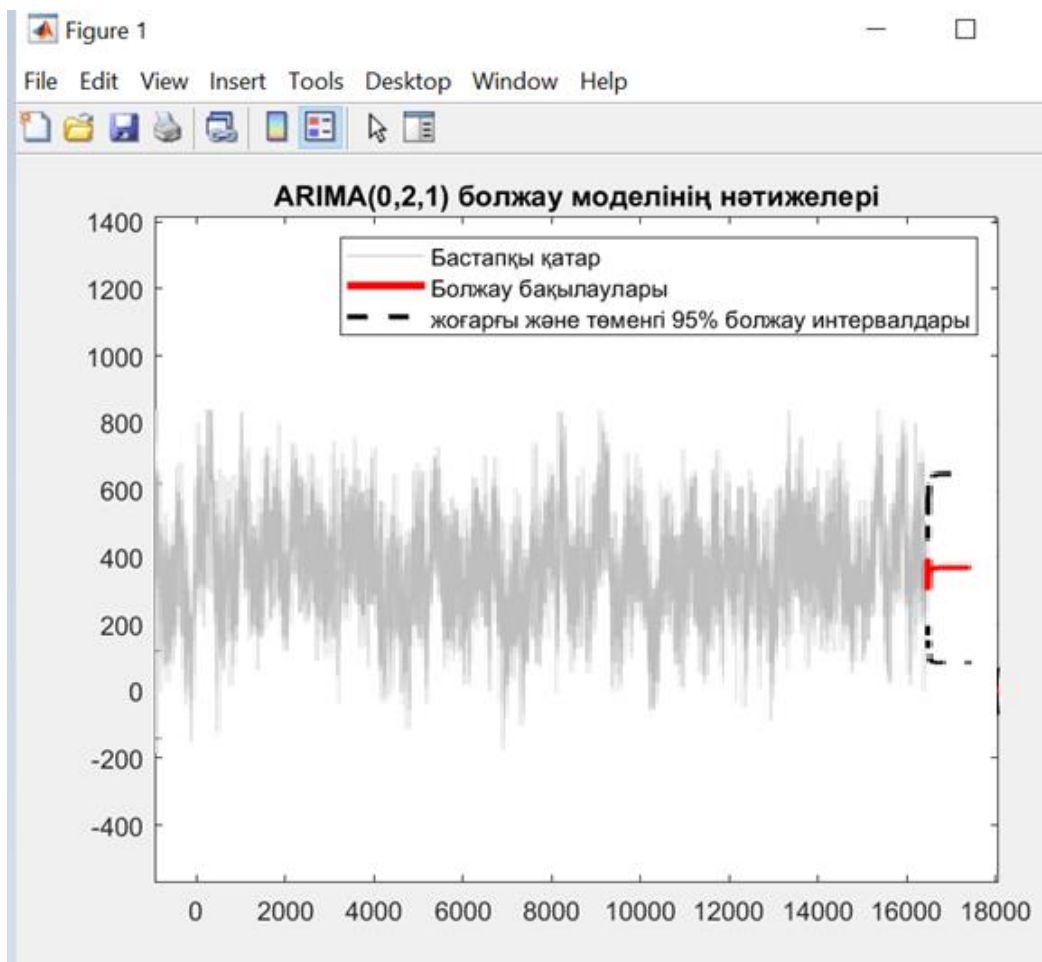
Сурет 4.5 – Қалдықтың графигі

4.5 суретте қатардың қалдық бөлімі көрсетілген. Графиктерді талдай отырып, қатардың қалдықтары орташа мәні нөлге жақын қалыпты таралуға ие екенін атап өтуге болады.

Жоғарыда айтылғандардың барлығын ескере отырып, нәтижесінде алынған ARIMA болжамдық моделі (сенімділік интервалдарын ескере отырып) барабар (адекватный) деп айта аламыз.



Сурет 4.6 – Қалдықтың гистограммасы



Сурет 4.7 –ARIMA(0,2,1) моделі көмегімен болжау нәтижесі

Нәтижелерді талдай отырып, келесі қорытындылар жасалды:

- Уақыттық қатарлардың графигін визуалды талдау қатардың біркелкі емес қарқындылығы бар екенін көрсетеді. Сондықтан қатарды стационарлы емес деп санауға болады;
- Қатарда бірлік түбір бар екені расталды, сондықтан қатар стационарлы емес.
- ARIMA моделінің сәйкестік графигі (4.4 сурет) дұрыс параметрлері мен барабар болжамдары бар дұрыс модель таңдалғанын көрсетеді. Графиктің көк бөлігі әрең көрінеді, бұл тамаша болжамды көрсетеді.
- Қалдықтар графиктері (4.5, 4.6 суреттер) қатар қалыпты таралуға ие екендігін көрсетеді. ARIMA әдісіне сәйкес, моделдерді құрудағы қадамдардың бірі бастапқы уақыт қатары стационарлы болғанша айырмасы болып табылады. Қажет болғандай бұл графиктер нөлге жақын орташа мәнмен стационарлы болып табылады.
- ARIMA(0,2,1) моделі (4.7 сурет) болжау интервалы 95% барабар болжамды көрсетті.

4.3 Желілік трафикті зерттеу үшін нейрондық желіні қолдану

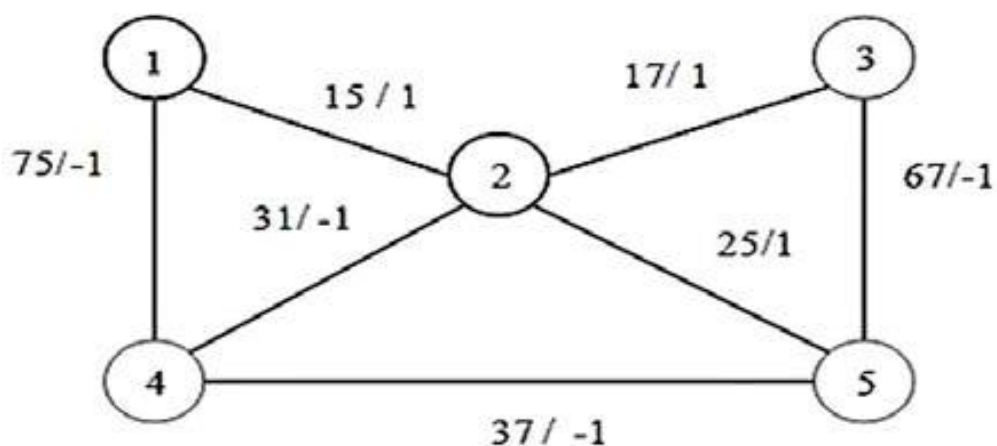
Нейрондық желілердің математикалық модельдері келесі маңызды сипаттамалары бар ыңғайлы есептеу әдісі болып табылады:

- оқуға қабілеттілік;
- бейімділік;
- есептеулердің жоғары параллелділігі мен жылдамдығы;
- болжамды есептеу сенімділігі, яғни дәл емес кіріс деректермен шешімдерді дұрыс анықтау мүмкіндігі.

Нейрондық желінің алгоритмдері маршруттау тапсырмасының нейрондық баптауымен (дәлірек айтқанда, оның нейрондық интерпретациясы) және берілген ортадағы (баптау) есептерді шешетін сәйкес (қолайлы) нейрондық желі моделімен анықталады.

Нейрондық компьютерлерді құру және адаптивті нейрондық желілерді модельдеу жасанды интеллекттің көптеген мәселелерін шешудің ең перспективалы бағыты ретінде танылды. Маршрутизациялау алгоритмінде қолданылатын нейрондық желі параллельді есептеудің математикалық моделі болып табылады, оның құрамында қарапайым өзара әрекеттесетін өңдеу элементтері – жасанды нейрондар бар. Нейрондық желілердің дәстүрлі алгоритмдерден артықшылығы – олардың үйрену қабілеті.

Жергілікті компьютерлік желінің құрылымы граф түрінде берілген деп есептейік (4.8 сурет).



Сурет 4.8 - Белгілері бар график

Белгілер графиктің шеттерінде келесі атрибуттарға сәйкес орнатылады: маршруттың ең аз ұштарына 1 мәні, ал басқа ұштарға мән (-1) мәні тағайындалады. Мысалы, 1-ші түйіннен 2-ші түйінге дейінгі жол 15 болып табылады және 1 ретінде кодталған.

$$\beta_{12} = 15 \rightarrow 1. \quad (4.10)$$

немесе, мысалы, 1 және 3 түйіндер арасындағы ықтимал маршруттарды келесідей кодтауға болады:

$$\mu_{13} = \beta_{12} \rightarrow \beta_{23} = 1 \rightarrow 1. \quad (4.11)$$

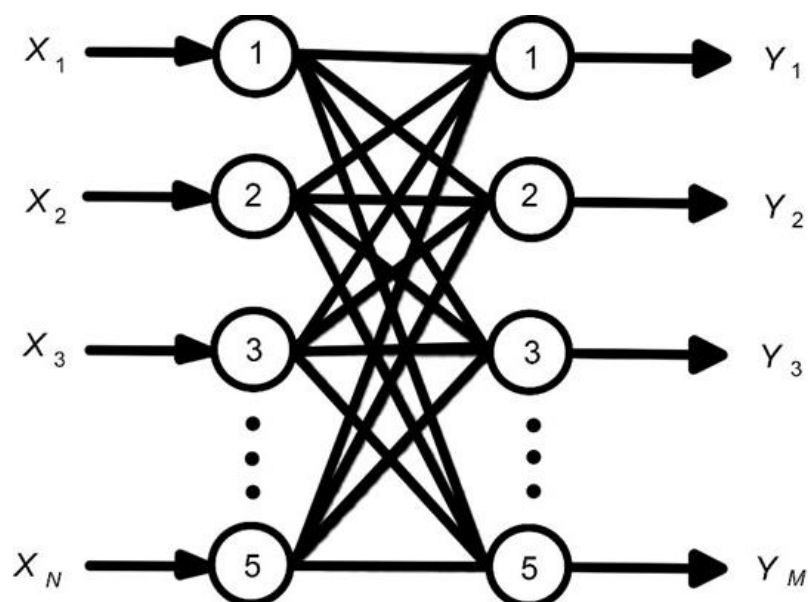
$$\mu_{13} = \beta_{14} \rightarrow \beta_{42} \rightarrow \beta_{23} = -1 \rightarrow -1 \rightarrow -1. \quad (4.12)$$

$$\mu_{13} = \beta_{14} \rightarrow \beta_{45} \rightarrow \beta_{53} = -1 \rightarrow 1 \rightarrow -1. \quad (4.13)$$

мұнда μ_{ij} – код формасы түрінде ұсынылған желі түйіндері арасындағы жалпы қашықтық болып табылады

β_{ij} – желі түйіндері арасындағы кодталған қашықтық (желі әкімшісінің шарттарына сәйкес бұл параметр сонымен қатар деректерді беру жылдамдығын, кешіктіруді және т.б. көрсете алады).

Нейрондық желі оң жолдар тізбегімен кодталған маршрутты таңдайды, мұндағы (1) дұрыс бағыт және (-1) дұрыс емес бағыт. Маршрут графигін көрсету үшін нейрондық желі құрылымы таңдалады, мұнда X_i нейрондық желінің кіріс деректері, Y_j нейрондық желінің шығыс деректері, мұнда $i = 1..N$; $j = 1..M$; $N = M = 5$, x_i кіріс сигналының ағымдағы мәні болып табылады, ал y_j шығыстың ағымдағы мәні болып табылады (4.9 сурет).



Сурет 4.9 - Нейрондық желінің құрылымы

Берілген нейрондық желі үшін оңтайлы маршруттардың кодталған кестесін қарастырайық (4.3 кесте).

Кесте 4 - Нейрондық желі үшін оңтайлы графикалық жолдар

	y_1	y_2	y_3	y_4	y_5
x_1	-1	1	-1	-1	-1
x_2	1	-1	1	-1	1
x_3	-1	1	-1	-1	-1
x_4	-1	1	-1	-1	-1
x_5	-1	1	-1	1	1

SGN қолтаңбаны белсендіру функциясы бар нейрондық желіні қарастырайық (с функцией активации сигнатур SGN) [].

$$Y_i = SGN(\sum_{i=1}^N W_{ij}X_i) = \begin{cases} 1 \text{ нпу } \sum_{i=1}^N W_{ij}X_i \geq 0 \\ -1 \text{ нпу } \sum_{i=1}^N W_{ij}X_i < 0 \end{cases} \quad (4.14)$$

мұнда W_{ij} – нейрондық желінің салмақ факторы.

Нейрондық желінің кірісі маршруттар болып табылатын кодталған сигналдарды қабылдайды делік:

Бейнелер (суреттер) үшін нейрондық желінің салмақ факторы матрицалары келесі формула арқылы есептеледі:

$$W_i = x_i x_i^T. \quad (4.15)$$

$$W_1 = x_1 x_1^T = \begin{pmatrix} 1 & -1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 \end{pmatrix};$$

$$W_2 = x_2 x_2^T = \begin{pmatrix} 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 \end{pmatrix};$$

$$W_3 = x_3 x_3^T = \begin{pmatrix} 1 & -1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 \end{pmatrix};$$

$$W_4 = x_4 x_4^T = \begin{pmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & -1 & 1 \end{pmatrix};$$

$$W_5 = x_5 x_5^T = \begin{pmatrix} 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 \end{pmatrix}$$

Соңғы желі салмағы матрицасы W келесі формула бойынша есептеледі:

$$W = \sum_1^5 W_i = \begin{pmatrix} 5 & -5 & 5 & 1 & 3 \\ -5 & 5 & -5 & -1 & -3 \\ 5 & -5 & 5 & 1 & 3 \\ 1 & -1 & 1 & 5 & -1 \\ 3 & -3 & 3 & -1 & 5 \end{pmatrix}$$

Жадта сақталған анықтамалық кескінді анықтау үшін көбейтіндіні есептеу қажет

$$WX_i = Y_i (i = 1, 2, 3, 4, 5) \quad (4.16)$$

Мысалы, x_3 үшін

$$WX_3 = \begin{pmatrix} 5 & -5 & 5 & 1 & 3 \\ -5 & 5 & -5 & -1 & -3 \\ 5 & -5 & 5 & 1 & 3 \\ 1 & -1 & 1 & 5 & -1 \\ 3 & -3 & 3 & -1 & 5 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -19 \\ 19 \\ -19 \\ -7 \\ -13 \end{pmatrix}.$$

Іске қосу функциясын ескере отырып, біз шығыс сигналын аламыз $(-1 \ 1 \ -1 \ -1 \ -1)$ яғни нейрондық желі сақталған кескінді дұрыс қалпына келтірді. Басқа суреттер үшін де осындай нәтижелер алынды.

Осылайша, ұсынылған тәсіл компьютерлік желі маршрутизаторының бағдарламалық қамтамасыз етуінде сәтті қолданылуы мүмкін.

4.4 Жасанды нейронды желі негізіндегі желілік трафикті болжау

Мультисервистік пакеттік желіде генерацияланған уақыт қатарының жоғарыда келтірілген зерттеулері пакеттер санының біркелкі емес интенсивтілігін көрсетті, яғни қатар стационарлы емес болып шықты. Мұндай қатарларды болжау үшін классикалық болжау әдістері жарамайды.

Нейрондық желіні болжау арқылы сызықты емес сәйкестендіру мүмкін болады. Сонымен қатар, нейрондық желілер мен терең оқыту мұндай тапсырмалармен жақсы жұмыс істейді [76,77].

Бұл жұмыста олардың әмбебап жуықтау және болжау мүмкіндіктерімен байланысты ЖНЖ (жасанды нейрондық желі) негізінде болжам жасалған. 1943 жылы шетелдік ғалымдар У.Маккаллох пен У.Питтс «Логическое исчисление идей, относящихся к нервной активности» мақаласында сипатталғандай, ЖНЖ түсінігін енгізіп, жасанды жүйенің формальды моделін ұсынған. Одан әрі 1949 жылы Д.Хебб желінің негізгі принциптерін сипаттап, нейрондық оқытудың бірінші алгоритмін жасады. Ал бүгінгі күні ЖНЖ көпөлшемді функцияларды жуықтау, процестерді болжау және диагностикалау, ассоциациялар бойынша іздеу және деректер массивтеріндегі заңдылықтарды іздеу, адаптивті басқару және статистикалық талдау, заңдылықтарды анықтау және тану және т.б. үшін модель және құрал ретінде пайдаланылады [78-80].

ЖНЖ - бұл тірі организмнің жүйке жасушаларының желілері - биологиялық нейрондық желілерді ұйымдастыру және жұмыс істеу принципіне негізделген математикалық модельдер (бағдарламалық немесе аппараттық жүзеге асыру).

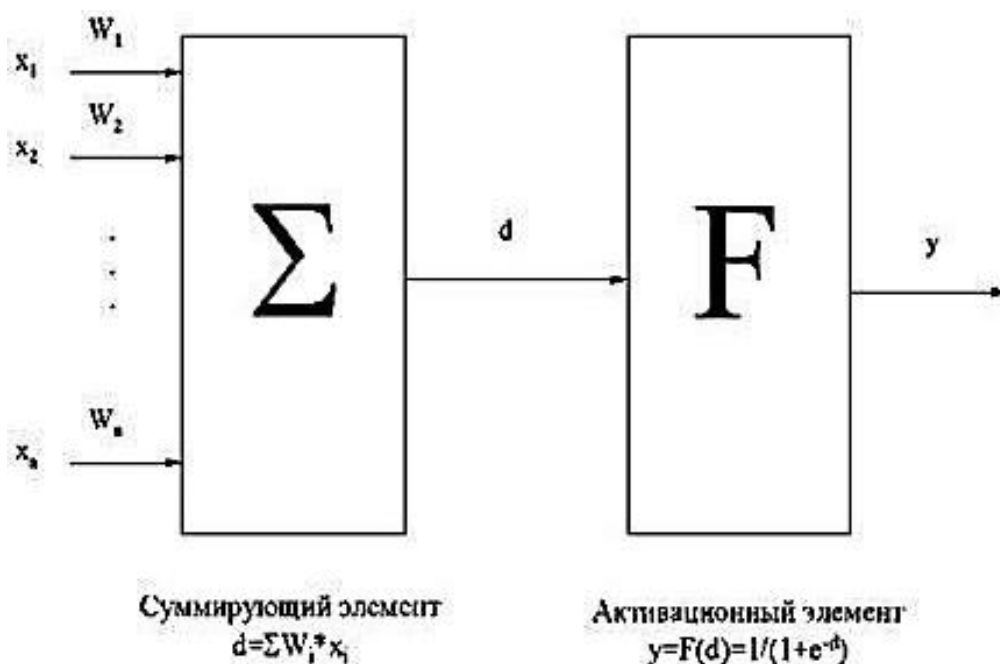
ЖНЖ – өзара байланысқан жасанды нейрондардың жиынтығы.

Нейрондық желілер параллель жұмыс істейтін қарапайым элементтерден тұрады. Бұл элементтердің атауы биологиялық жүйке жүйесінен шыққан. Табиғаттағы сияқты элементтер арасындағы байланыстар желінің қызметімен жақсы анықталған. Жасанды нейрондық желілердің мүмкіндіктері нейрондық желі мәліметтерін өңдеу технологиясына сәйкес алгоритмдердің соңғы

параллелизациясын қамтамасыз ететін жіктеу, болжау және басқару мәселелерін шешу үшін сәтті қолданылады [81,82].

Артқа таралу алгоритмі (Алгоритм обратного распространения - Back propagation) - бұл алгоритмді бұрынғы қарапайым алгоритмдерден ерекшелендіретін пайдалы жаңа мүмкіндіктерді жасауға қабілетті көпқабатты перцептрондарды оқытуға арналған бақыланатын оқыту алгоритмі. Сонымен қатар көпқабатты перцептронның мәні кері таралу әдісінің перцептронның бірнеше қабаттары бар болғандықтан емес, ал барлық қабаттарды меңгеруінде жатыр. Back propagation нейрондық желіні оқытудың мәні болып табылады. Бұл алдыңғы дәуірде (яғни итерация) алынған қателік жылдамдығына (яғни жоғалту) негізделген нейрондық желінің салмақтарын дәл баптау тәжірибесі. Салмақтарды дұрыс баптау қателіктердің төмен деңгейін қамтамасыз етеді, бұл оның жалпылауын арттыру арқылы модельді сенімді етеді. Оқыту деп нейрондар арасындағы байланыстардың салмақтық коэффициенттерін өзгерту арқылы берілген анықтамалық үлгілерге желіні бейімдеу түсініледі (4.10 сурет). Нейрондардың қабаттық ұйымы мидың белгілі бір бөліктерінің қабатты құрылымдарына еліктейді [83,84]. Көпқабатты нейрондық желі кіріс векторларын шығысқа кез келген бейнелеуге қабілетті.

Нейрондық желіні оқытушымен (кіріс деректерінің де, қажетті шығыс сигналдарының да мәндері ұсынылған және ол өзінің синаптикалық қосылыстарының салмағын кейбір алгоритмге сәйкес реттейді) және оқытушысыз (салмақтар сәйкес өзгереді) оқытуға болады. Оқытушымен бірге оқытылатын нейрондық желі көпқабатты және сызықты емес авторегрессия мәселесін шешуге қолайлы.



Сурет 4.10 - Жасанды нейрон

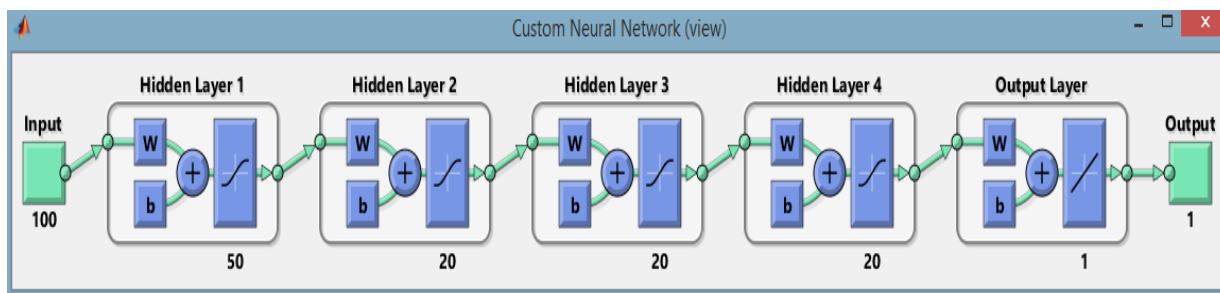
Стационарлық емес уақыт қатарын болжауды жүзеге асыру үшін Matlab ядросының бақылауымен жұмыс істейтін Math Works Neural Network Toolbox (NNT) компаниясының қолданбалы пакеті (APP) таңдалды.

Бұл пакетке келесі оқыту функциялары енгізілген:

- traindx - MORO импульсі және бейімделуі;
- traincdr - градиенттер (Поллак-Рибейра);
- trainlm – Левенберг-Марквардт әдісі.

Болжаудың дәлдігі оның архитектурасына, яғни жасырын қабаттардың санына және әрбір қабаттағы нейрондардың санына байланысты. Көптеген қабаттары бар перцептронның есептеу қабілеті жоғары және бір қабаты бар перцептронға қарағанда үлкен сызықтық емес көп өлшемді қатынастарды жақсы сипаттайды. Қолданбалы пакеттің артықшылықтарына уақыттық қатарды болжау, көпқабатты перцептрондарды құру және пайдалану және т.б. үшін әртүрлі модельдер мен әдістердің үлкен саны кіреді.

Бастапқы желілік трафикті болжау үшін біз ЖНЖ пайдаланамыз [85-87]. Тренинг үшін пакетті жіберу қарқындылығының уақыттық қатары ұсынылды. Шығару – болжам деректері. Бес қабаты бар желі, оның ішінде төрт жасырын қабат пайдаланылды. Бірінші жасырын қабат екіншіге, екіншісі үшіншіге, үшінші төртіншіге байланысты. Бірінші жасырын қабатта 50 нейрон, екіншісінде - 20, үшіншіде - 20, төртіншіде - 20, бесіншіде - 1 болады (4.11 сурет).



Сурет 4.11 - ЖНЖ құрылымы

ЖНЖ іске қосу және оқыту нәтижесінде ақпараттық графиктер алынды:

- Train, жаттығу параметрлерін сипаттау.

Бұл ретте дәуірдің (циклдің) мәні 1-ге тең, ол жаңа реттеу үлгісін енгізу арқылы ұсынылған желіні көрсетеді (қатенің кері таралуы) тікелей белсендіру ағынымен бірге, түбір-орта квадрат қатесі бар салмақтарын түзету кезінде орын алды. Дәуірдің мәнінің одан әрі жоғарылауы орташа квадраттық қателік MSE мәнінің төмендеуіне әкеледі, яғни оқыту 4 дәуірге созылды;

- Валидация – 1-ге тең дәуір мәнінен кейін зерттелетін деректер оның мәндерін өзгертпейді, яғни ең жақсы өнімділік валидация қатесі ең аз дәуірден алынады (дөңгелек нүкте);

- Тест – оқудың ең жақсы көрсеткіштерін көрсетеді;
- Үздік – ең жақсы өнімділік.

[88] жұмыста статистикалық және құрылымдық болжау үлгілері талданады, мұнда қысқа мерзімді болжау үшін статистикалық үлгілерді және орта мерзімді және ұзақ мерзімді болжау үшін құрылымдық үлгілерді қолдану мүмкіндігі туралы қорытынды жасалады.

Желілік трафикті болжау ресурстарды басқаруды ұйымдастыруда және қазіргі заманғы телекоммуникация желілерін дамытуды жоспарлауда шешуші рөл атқарады. Дәстүрлі түрде трафикті болжау ағымдағы трафик көлемі мен оның болашақтағы мүмкін мәндері арасындағы корреляцияны белгілейтін статистикалық авторегрессивті модельдер негізінде құрылды. Дегенмен, бұл модельдердің айтарлықтай шектеулері бар, атап айтқанда, олардың көпшілігі стационарлық статистикалық деректерге ғана қатысты, бұл олардың қолданылу аясын уақыт бойынша тұтынылатын трафиктің тұрақты көлемдері бар қарапайым құрылымдағы желілерге дейін тарылтады. Қазіргі заманғы желілердің көпшілігі күрделі құрылымға ие және трафик ағындарының жылдам өзгеруіне ұшырайды. Дәстүрлі үлгілердің шектеулерін еңсеру және болжауды жақсарту үшін соңғы уақытта машиналық оқыту технологияларын қолдану бойынша күшті талпыныс жасалды.

4.5 NARX желісі көмегімен желілік трафикті болжау

[89,90] жұмыстарда NARX модельдері сызықты емес жүйелерді модельдеу үшін өте қолайлы көп сатылы болжау үшін қуатты модельдер класы екендігі сипатталған.

ЖНЖ архитектурасында сигналдардың таралуына байланысты нейрондық желілер 2 түрге бөлінеді:

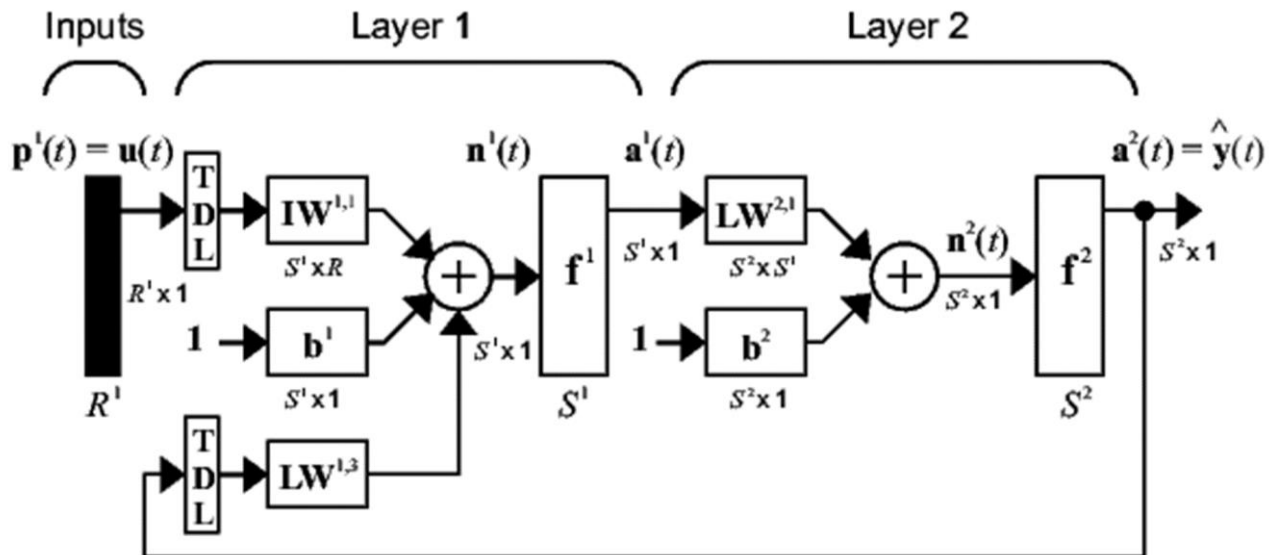
1. Тікелей таралу желілері
2. Қайталанатын желілер

Бұл жұмыста – сигналдың тікелей таралуы бар желі – NARX сызықты емес авторегрессивті желісі (Nonlinear AutoRegressive with exogenous Network) қолданылды. Онда нейрондық желілерді пайдаланып динамикалық сызықтық емес жүйелерді жүзеге асыру үшін айырмашылық авторегрессивті теңдеуіне сызықтық емес коэффициенттер енгізіледі. Сонымен қатар, бұл типтегі желі сигналды тікелей жіберу және кері байланысы бар көп қабатты, ал кіріс сигналы уақыт кідірістерінің векторы арқылы өтеді.

NARX желілері екі нұсқада келеді:

- параллель;
- реттік-параллель (последовательно-параллельный).

Бұл жұмыста желінің параллель нұсқасы қолданылды (4.12 сурет).



Сурет 4.12 - NARX желісінің архитектурасы

Зерттелетін стационарлы емес уақыттық қатарды болжауды жүзеге асыру үшін Matlab жүйесіндегі Neural Network Toolbox (NNT) қолданбалы пакеті (APP) таңдалды.

Сыртқы кірістері бар NARX [91-94] моделінде жасанды нейронды желіге түсетін бастапқы уақыттық қатар салмақ коэффициенттеріне сәйкес өңделеді, содан кейін шығыс деректері қайтадан желілік кіріске енеді, осылайша қатені кері тарату процедурасы жүзеге асады. Бұл желіні қайта даярлау кезінде бастапқы дайындықтан кейін алынған салмақ коэффициенттерін есепке алуға мүмкіндік береді және бұл өз кезегінде оның жұмысының дәлдігін арттырады.

Цифрлық сигналдарды өңдеу мәселелерінде динамикалық нейрондар кеңінен қолданылады, олар айырмашылық немесе дифференциалдық теңдеулер арқылы сипатталады, олардың мінез-құлқы олардың алдыңғы тарихымен анықталады. Ең қарапайым динамикалық нейрондардың бірі Хопфилд нейроны болып табылады. NARX сызықты емес авторегрессивті желісіне келетін болсақ, онда күйі әлдеқайда алыс тарихпен анықталатын модификацияланған Хопфилд нейроны қолданылады. NARX моделінің динамикасы келесідей сипатталған

$$y(n + 1) = F(y(n), \dots, y(n - q + 1), u(n), \dots, u(n - q + 1)) \quad (4.17)$$

мұндағы F – жаттығу кезінде жуықталатын сызықтық емес функция,
 q – кідіріс (задержка).

NARX желісі - тікелей сигнал беру және шығыс бойынша кері байланысы бар көп деңгейлі желі, оның шығысы уақыт бойынша кідіріс векторы арқылы өтеді. Сигма тәрізді функция белсендіру функциясы ретінде қолданылады.

Желінің әрбір қабаты белгілерді енгізу кеңістігін басқа өлшеммен басқа кеңістікке түрлендіреді. Мұндай сызықты емес түрлендіру шығыс қабатының сызықты түрде бөлінетін нейрондары болғанша жалғасады. Кіріс пен шығыстан

басқа, ЖНЖ барлық деңгейлері; желіге сызықтық емес құбылыстарды моделдеу мүмкіндігін береді.

Нейрондық желіні оқыту үшін Левенберг-Маквардт алгоритмі қолданылды, ал нейрондық желінің өнімділігін бағалау үшін орташа квадраттық қате (MSE) қолданылады [95]. Левенберг-Маквардт алгоритмі сызықты емес модельдер үшін оңтайландыру және жуықтау есептерін шешу үшін қолданылатын қайталанатын ең кіші квадраттар әдісі. Ол Гаусс-Ньютон әдісінің модификациясы болып табылады және градиенттік түсу идеяларын және Ньютон әдісін біріктіреді. Левенберг-Маквардт алгоритмі әдетте бақыланатын мәндер мен болжанатын мәндер арасындағы квадраттық айырмашылықтардың қосындысын азайту үшін модель параметрлерін реттеу үшін қолданылады. Ол әсіресе нейрондық желіні оқыту немесе сызықты емес оңтайландыру мәселелері үшін пайдалы.

Кіріс (бастапқы деректер) және мақсатты деректер векторлары кездейсоқ түрде келесі үш жинаққа бөлінді:

- оқыту үшін векторлардың 70% пайдаланылды;
- 15% векторлар - нәтижелердің сенімділігін тексеру және желіні қайта жаттықтыруды болдырмау үшін пайдаланылды;
- 15% желіні тәуелсіз тестілеуге пайдаланылды.

Зерттелетін уақыттық қатардың бірнеше қадамға ілгері келіп түсетін пакеттер қарқындылығын болжау бағдарламасы әзірленді (4.13 сурет).

```
clc
clear all
close all
Data_Inputs=xlsread('Колич.ряд_MPEG_18000.xlsx'); % Import file

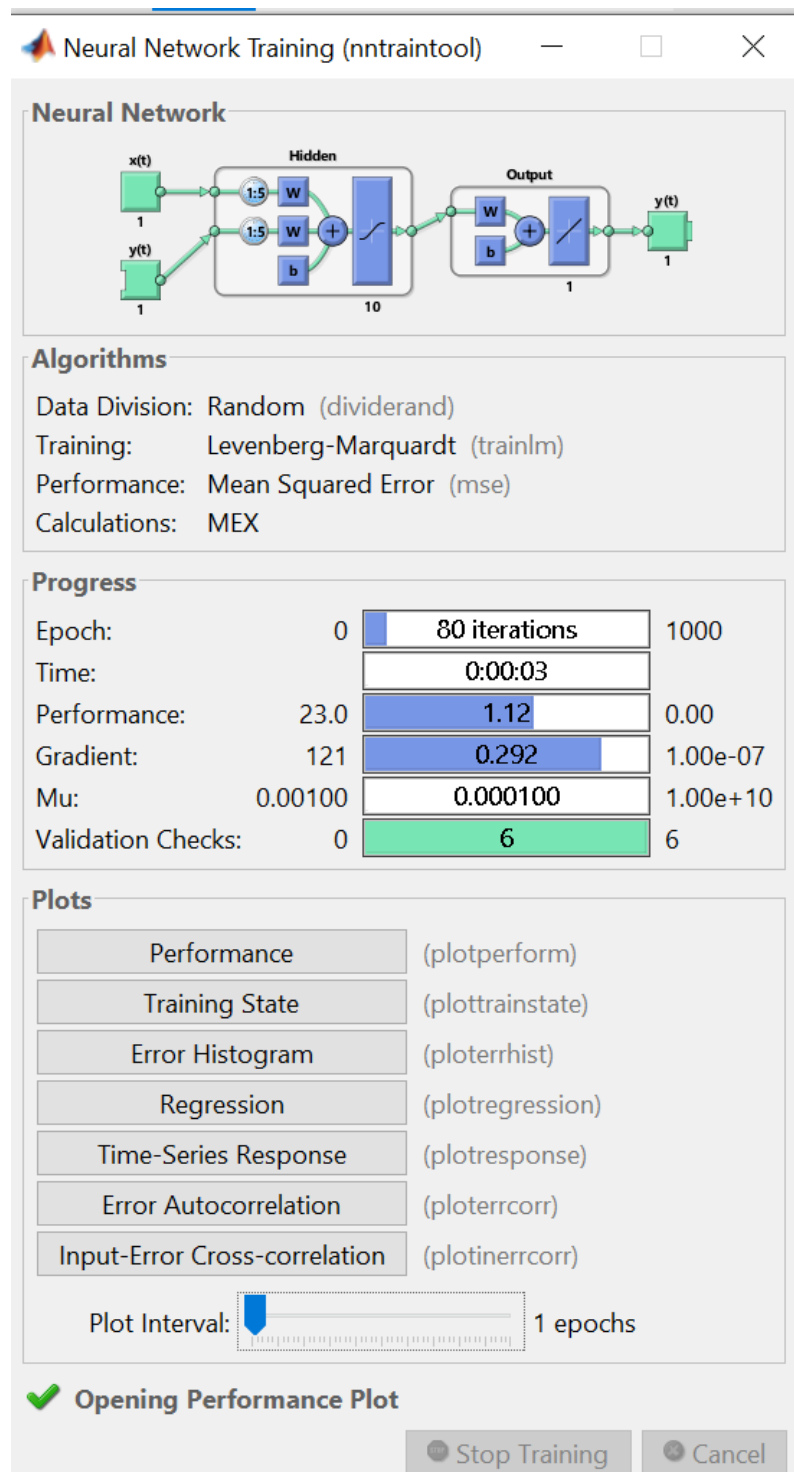
Shuffling_Inputs=Data_Inputs(randperm(end),1:2); % integers (training sample)
Training_Set=Data_Inputs(1:end,1);% specific training set
Target_Set=Data_Inputs(1:end,2); % specific target set
Input=Training_Set'; % Convert to row
Target=Target_Set'; % Convert to row
X = con2seq(Input); % Convert to cell
T = con2seq(Target); % Convert to cell

%% 2. Data preparation
N = 600; % Multi-step ahead prediction
..
```

Сурет 4.13 - Пакеттер қарқындылығын болжауға арналған бағдарлама кодының фрагменті

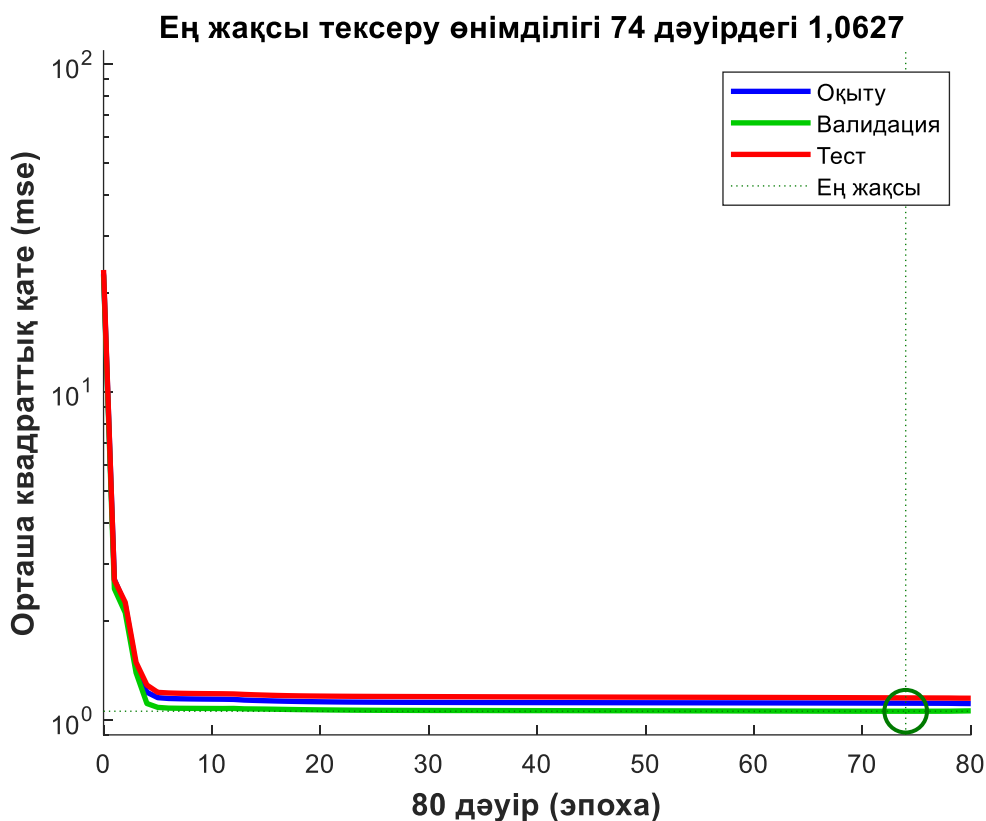
4.14 суретте NARX желісін оқыту процесінің басында ашылатын терезе көрсетілген.

Мұндай оқыту процестерінің принципі - қайталанатын нейрондық желі «уақыт өте қайта ашылады» және көпқабатты перцептрон ретінде ұсынылады, әрбір қабат кейбір өткен циклге сәйкес келеді.



Сурет 4.14 - NARX желісін оқытудың шығыс нәтижелері

Оқу процесі мақсатты дәлдікке жеткенде аяқталды, ал зерттелетін желі 74 итерацияда оқытылып бітті. «Performance» батырмасын басқанда итерациялар санына тәуелді орташа квадраттық қатенің (MSE) графигі алынады (4.15 сурет). Тексерудің (валидация) ең кіші қатесі 1,0627 тең.



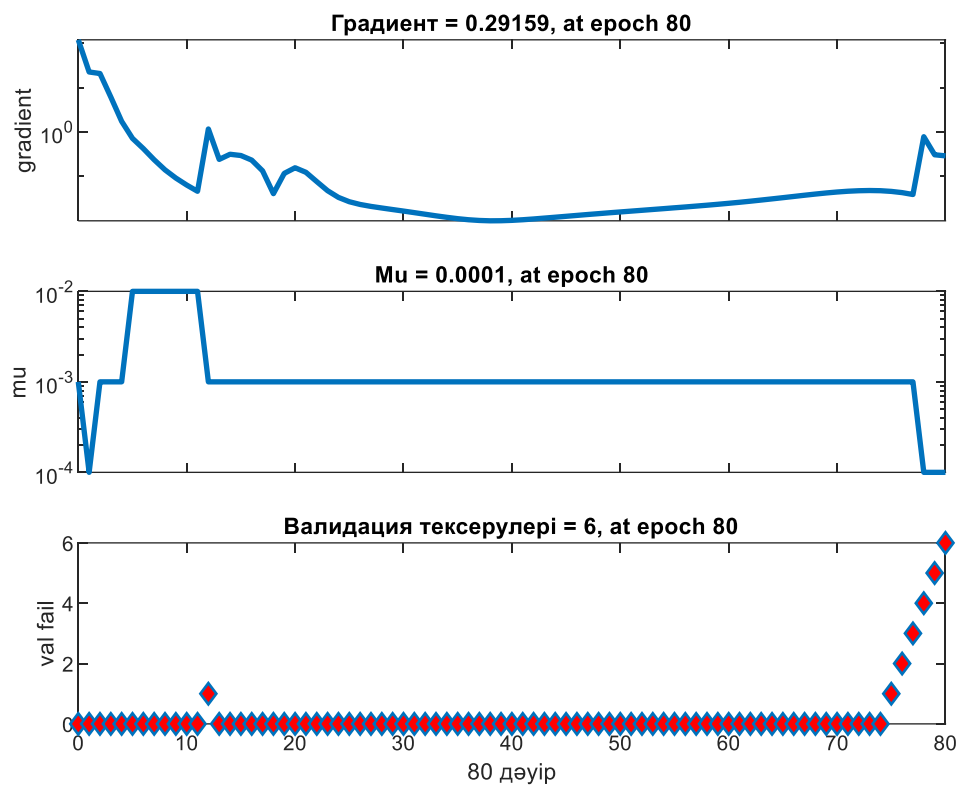
Сурет 4.15 - Орташа квадрат қатесінің итерациялар санына тәуелділігі

4.16 суретте ЖНЖ қосымша оқыту құралдары: әртүрлі болжау барысы параметрлері көрсетілген.

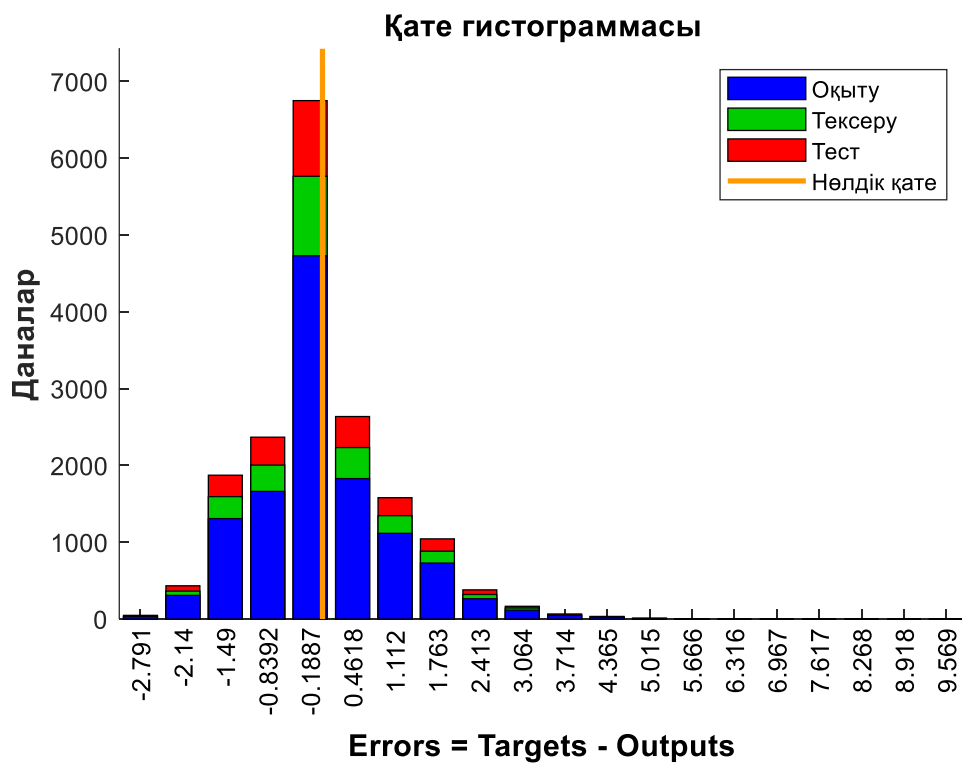
Сызықты емес авторегрессивті желінің шығыс қателерінің гистограммасы 4.17 суретте көрсетілген.

Жуықтаудың сапасын толық талдау үшін желілік нәтижелердің регрессиялық талдауы алынды, онда желілік оқыту нәтижелерінің сызықтық регрессиясы қарастырылған үш ішкі жиынға және бүкіл оқыту жиынына құрылады. Әрбір нәтиже үшін R корреляция коэффициенті есептеліп, графиктер салынды (4.18 сурет). Регрессиялық талдау - бұл тәуелді айнымалы (мақсатты айнымалы) және бір немесе бірнеше тәуелсіз айнымалылар (факторлар) арасындағы байланысты зерттеу үшін қолданылатын статистикалық әдіс. Ол тәуелсіз айнымалылардың тәуелді айнымалыға әсерін зерттеуге және тәуелсіз айнымалылардың мәндеріне негізделген тәуелді айнымалының мәндерін болжай алатын модель құруға мүмкіндік береді.

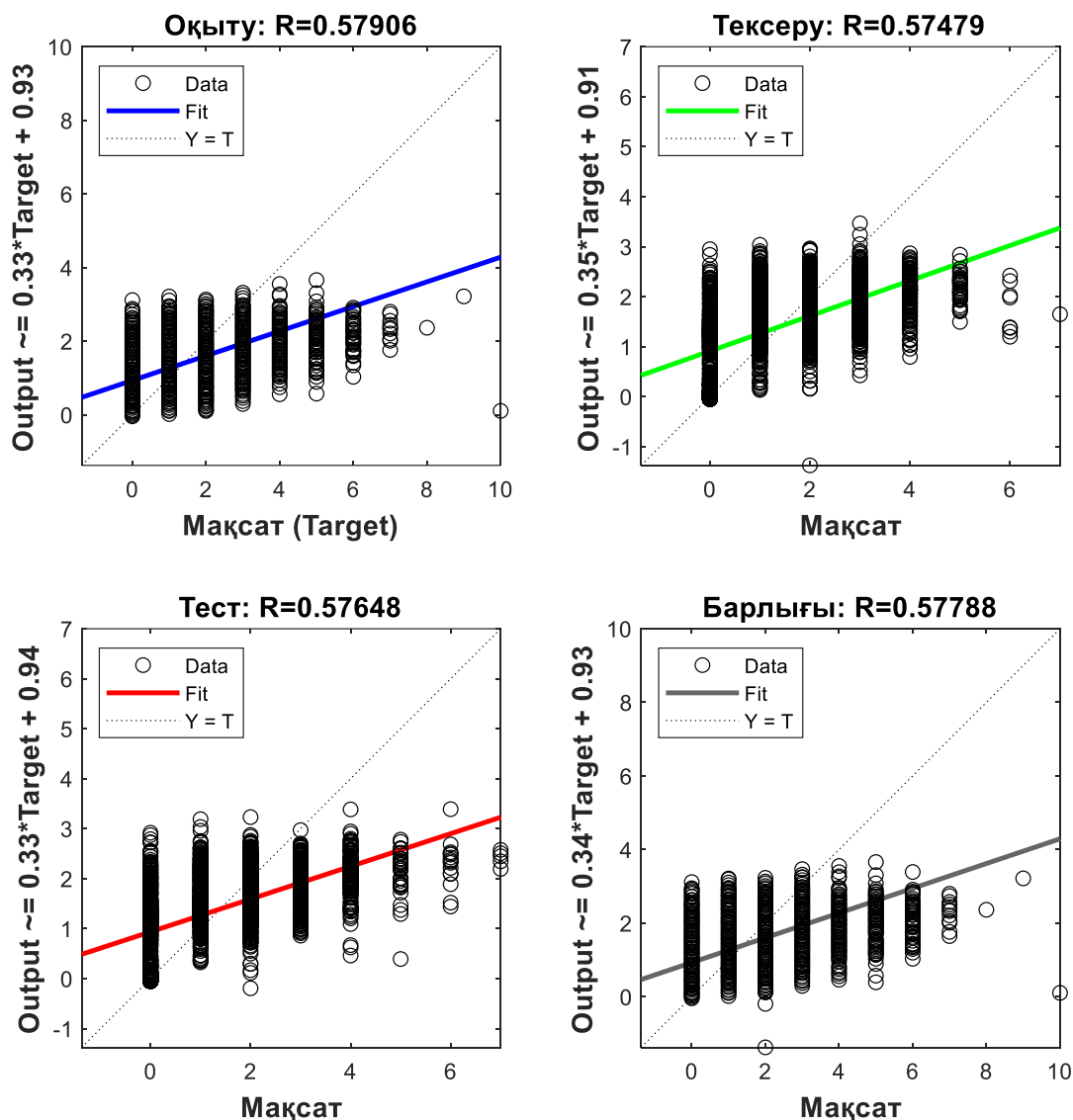
АКФ сызбасында (4.19 сурет) тек бір ғана нөлдік емес мән бар, ол нөлдік кідіріс кезінде орын алды (rms қатесі). Бұл болжам үлгісінің салыстырмалы түрде мінсіз екенін көрсетеді (болжау қателері бір-бірімен корреляцияланбайды, яғни олар ақ шу). Қалған АКФ мәндері сенімділік шегінен төмен және нөлге жуық шамамен 95% сенімділік шегіне түседі. Әйтпесе, болжамды жақсарту үшін желіні қайта оқыту қажет болады [96-98].



Сурет 4.16 - Болжау барысы параметрлері



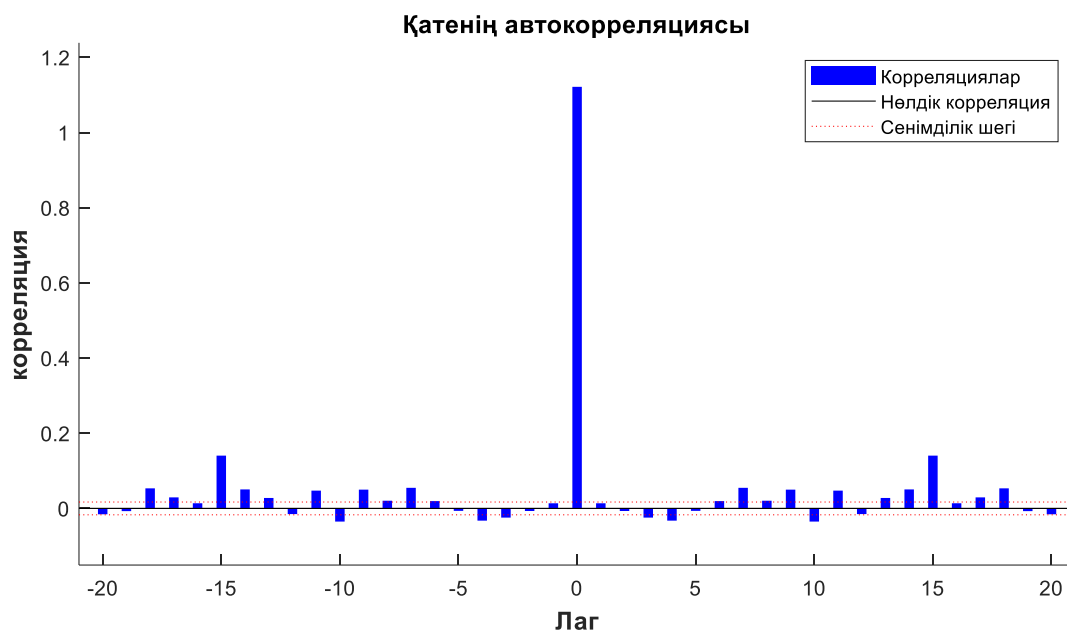
Сурет 4.17- СЫЗЫҚТЫ емес авторегрессивті желінің ШЫҒЫС қателерінің гистограммасы



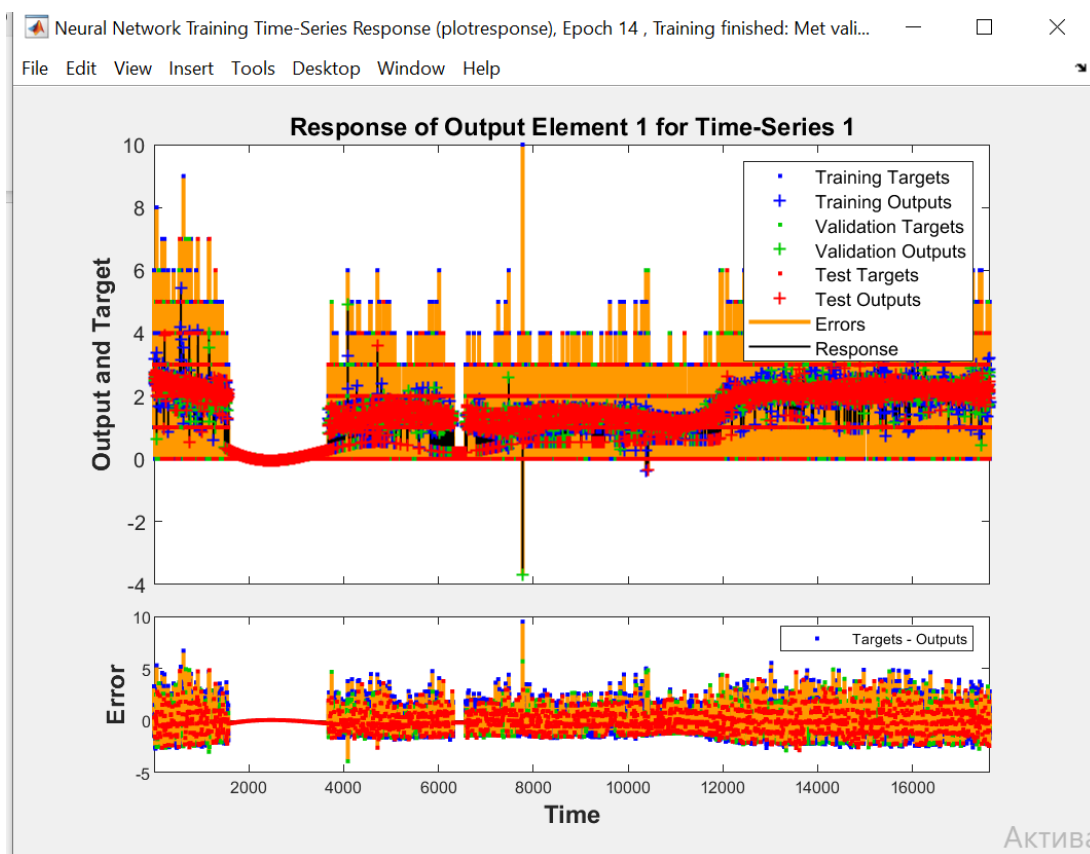
Сурет 4.18 - R корреляция коэффициенттері

Қарапайым сызықтық регрессия, көп сызықтық регрессия, логистикалық регрессия, көпмүшелік регрессия және т.б. қоса алғанда, бірнеше түрлі регрессия талдау әдістері бар. Белгілі бір әдісті таңдау деректердің сипаттамаларына және зерттелетін мәселенің сипатына байланысты.

Регрессиялық талдау әртүрлі салаларда кеңінен қолданылады. Регрессиялық талдау нәтижесінде тәуелді және тәуелсіз айнымалылар арасындағы байланысты сипаттайтын регрессиялық модель құрастырылады. Бұл модельге сүйене отырып, тәуелсіз айнымалылардың мәндерін ескере отырып, тәуелді айнымалының мәндерін болжауға болады.



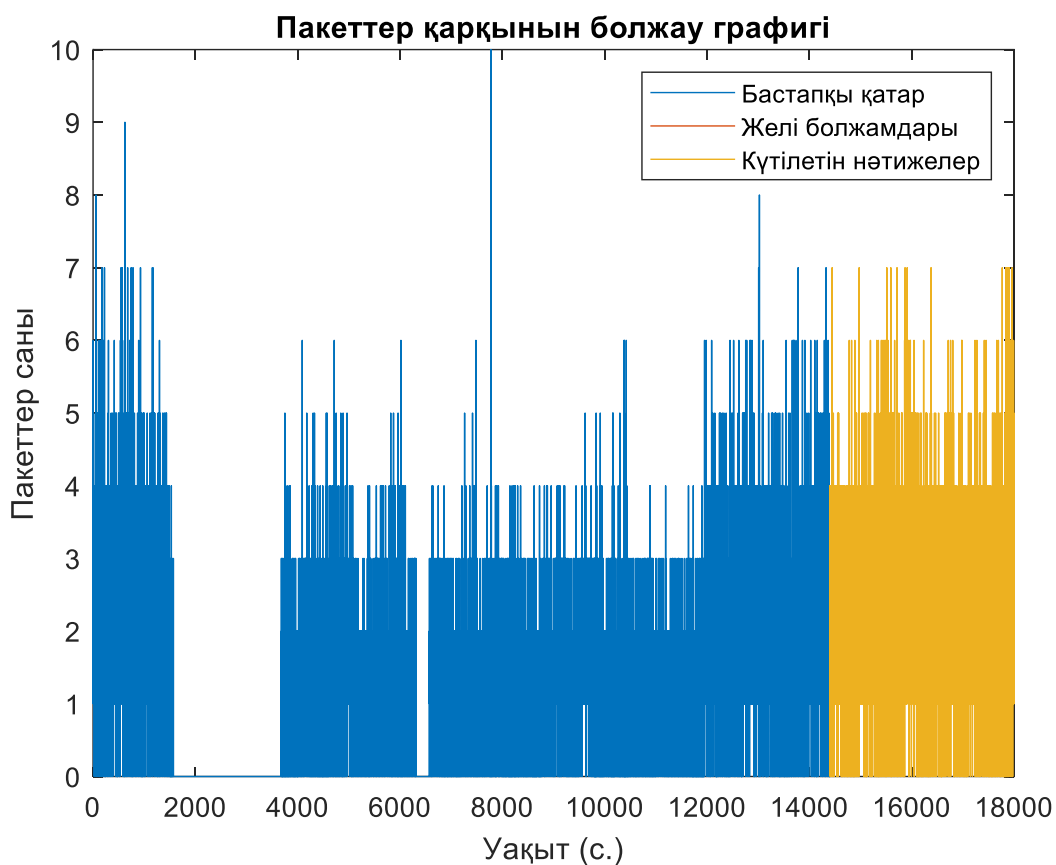
Сурет 4.19 - Қатенің автокорреляциясы



Сурет 4.20 - Жуықтау функциясы (функция аппроксимации)

Тәуелділік жуықтау есептерін шешу кезінде желі жұмысының сапасы функцияның белгілі мәндері және желімен жуықталған мәндер арасындағы айырмашылықтармен анықталады. 4.20 сурет Neural Network Toolbox Matlab жүйесіндегі жуықтау функциясын көрсетеді.

4.21 суретте пакеттер қарқындылығын болжау графигі көрсетілген, ол экзогендік кірістері бар NARX желісі болжамының қанағаттанарлық сапасын көрсетіп тұр.



Сурет 4.21 - Пакеттер қарқындылығын болжау

Кесте 5 - Тұрақты және нейрондық желілер үшін трафикті болжау алгоритмінің сипаттамаларының салыстырмалы талдауы

Критерий	RIP алгоритм	NARX
Бейімделу қабілеті	+	+
Оңтайлы маршрутты бағалау уақыты, с	5	2
Бұрмаланған сигналды түзету	-	+
Орындалу уақыты, с	58	11,9
Іске асыру жеңілдігі	+	+
Желі параметрлерін болжау мүмкіндігі	-	+
Пакетті жеткізу уақыты 150 кбит (с)	15	7

4.6 Нейро-анық емес болжау модельдерін әзірлеу

«Анық емес жиын» термині, сонымен қатар 1965 жылы құрылған анық емес жиындар теориясындағы алгоритмдерді сипаттаудың көптеген әдістері профессор Лотфи Задеге тиесілі.

Әртүрлі процестерді модельдеу кезінде нейрондық желілердің мүмкіндіктері ақпаратты өңдеу технологиялары анық емес жиындар мен анық емес қорытындыға негізделсе, күшеюі мүмкін. Гибридті желі – бұл анық емес логикалық әдістер мен жасанды нейрондық желі әдістерін біріктіретін желі. Бұл жағдайда мүшелік функцияны анықтау үшін анық емес жиындар пайдаланылады, ал ЖНЖ алгоритмдері дефаззификатордың функцияларын орындайды. Бұл осы әдістердің артықшылықтарына сүйене отырып, білім қорын тиімдірек орындауға мүмкіндік береді.

4.6.1 Уақыттық қатарды болжауғанейро-анық емес Чен және Ченг алгоритмдерін қолдану

1993 жылы шетелдік ғалымдар Сонг пен Чиссом анық емес уақыт қатары тұжырымдамасын ұсынды. Олар бірінші болып FTS (Fuzzy Time Series - АЕУҚ) анық емес уақыт қатарларының әдістемесін ұсынды. Бұл ретте уақыттық қатарлардың екі анық емес моделі қарастырылды: уақытқа тәуелсіз және уақытқа тәуелді. 1994 жылы Салливан мен Вудолл лингвистикалық белгіні пайдалана отырып, Сонг және Чиссом әдістерін уақыт бойынша өзгермейтін Марков үлгісімен салыстырды. 1996 жылы ғалым Чен Сон және Чиссом ғалымдарының алгоритмдерімен жасалған жеңілдетілген арифметикалық амалдарға негізделген әдісті кеңейтуді жүзеге асырды. Ғалым Ченнің идеясы бұрын жасалған модельдің есептеу күрделілігін азайту үшін топтық қатынастардың анық емес логикалық кестелеріне негізделген [99,100]. Ғалым Ченнің алгоритміне келсек, ол өз кезегінде ғалым Чен жасаған модельді кеңейтіп, тренд бойынша өлшенген АЕУҚ (FTS) моделін енгізді. Сонымен бірге болжау үшін ол анық емес қатынастарды бөлуге тиісті салмақтар берді [101]. 2001 жылы Хуарнг Чен моделін біріктіру арқылы эвристикалық модельді енгізді. 2002 жылы Чен жоғары ретті АЕУҚ ұсынды.

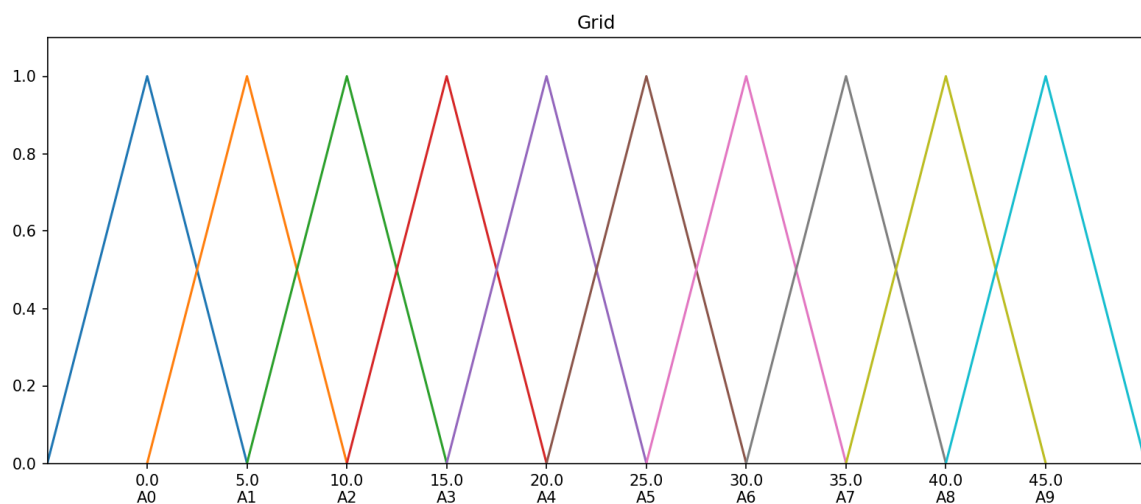
2006 жылы Хван стационарлық емес АЕУҚ - мен жұмыс істей бастады. Жалпы идея – уақыттық қатарлардың үлгілерінен ережелерді алу арқылы әрбір аймақ өзін қалай әрекет ететінін білу үшін уақыттық қатарлар дискурсының ғаламын интервалдарға/бөлімдерге бөлу. Бұл модельдердің ережелері мәндер бір жерден екінші жерге ауысқанда уақыт өте келе бөлімдердің бір-бірімен байланысын көрсетеді. Басқаша айтқанда, біз сандық уақыт қатарын көрсету үшін тілдік айнымалыны жасаймыз және бұл аймақтар айнымалының лингвистикалық терминдері болады.

[102] жұмыста динамикалық процестердің кең класы үшін оның анық емес мәндері негізінде уақыт қатарларын болжауға болатыны сипатталған. [103] жұмыста FAT (Fuzzy Approximation Theorem) теоремасына сәйкес анық емес логиканы пайдалана отырып, табиғи тілдің «егер-онда» формалды

мәлімдемелері бар ерікті қатынасты дәл жақындатуға болатыны атап өтілген. Яғни, анық емес уақыт қатары – тарихи деректер лингвистикалық тұрғыдан қалыптасқан кезде болжау үшін қолданылатын ұғым. Сонымен қатар, уақыттық қатардың ортақ қасиеті бар және оның элементтері берілген жиынға әртүрлі дәрежеде жатады.

Python бағдарламасында Chen және Cheng әдістерін бағдарламалық түрде енгізу үшін интервалдарға бөлетін `pyFTS.partition` модулі пайдаланылды.

Модель деректерінің уақыттық қатары 10 интервалға бөлінді және әрбір интервал үшін анық емес жиын құрылды.



Сурет 4.22 - айқындау (фаззификация) үшін қолданылатын 10 интервалға мүшелік функциялардың графигі

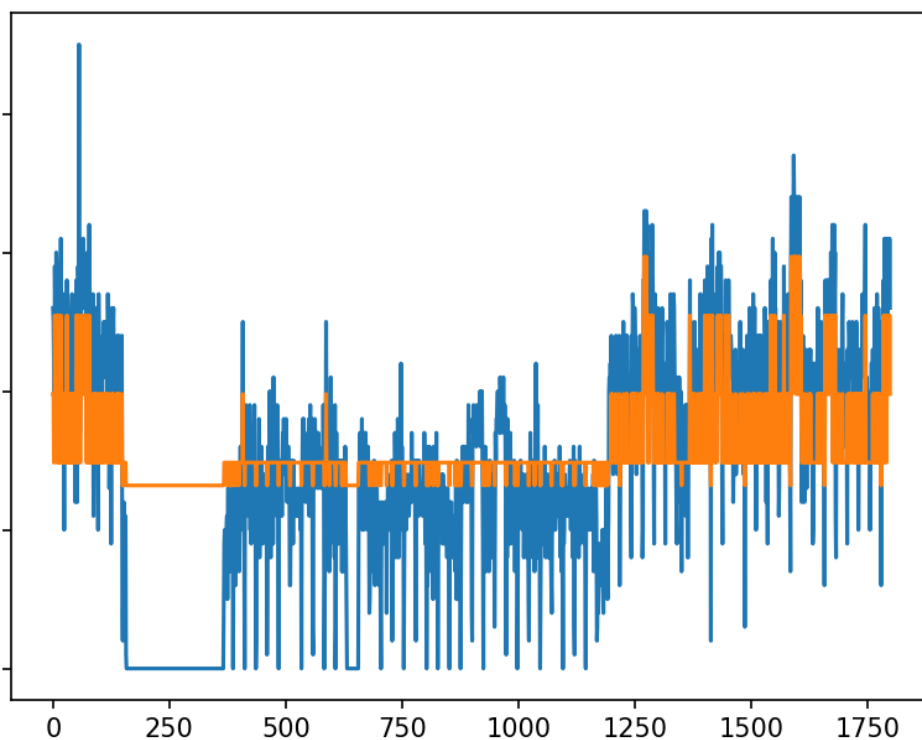
4.23 суретте бастапқы деректер мен болжам деректерінен тұратын біріктірілген график көрсетілген, ол Чен әдісі болжам деректерінің дәлдігі төмен екенін көрсетеді, сонымен қатар, MSE сандық мәні үлкен, 9,078161087779954, ал Ченг әдісінің дәлдігі жоғары, $MSE=1,2359176594533703$ (4.24 сурет).

MSE (Орташа квадраттық қате) – регрессия үлгілеріндегі орташа квадраттық қатені өлшеу үшін пайдаланылатын метрика. Ол нақты мәндер мен болжамды мәндер арасындағы айырмашылықтың орташа квадратын өлшейді.

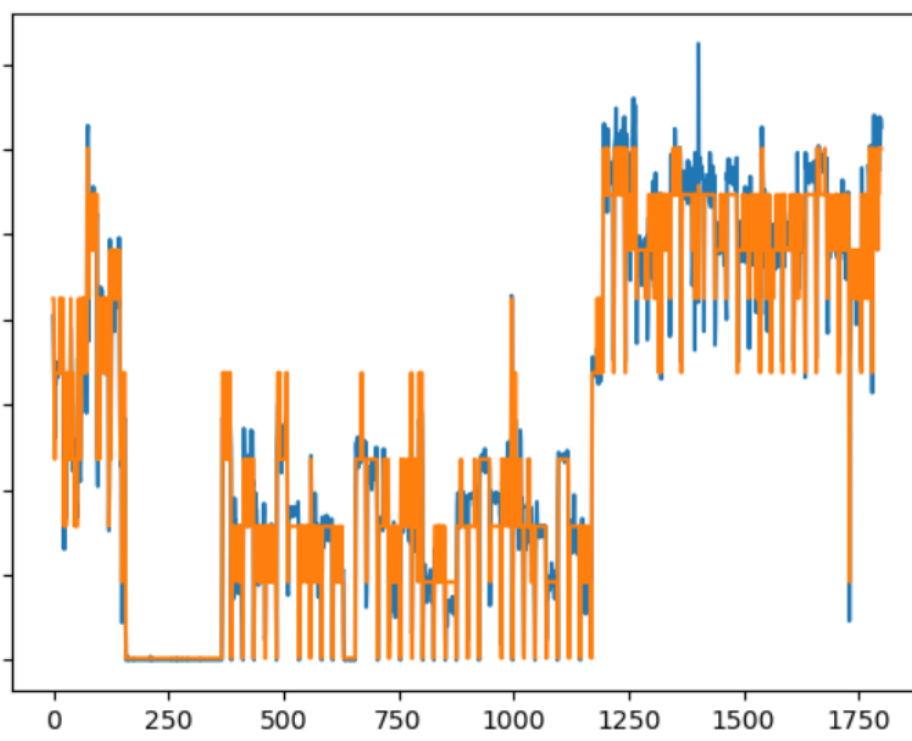
Ресми түрде MSE нақты мәндер (y) мен болжамды мәндер (\hat{y}) арасындағы квадраттық айырмашылықтардың қосындысын алып, содан кейін бақылаулар санына (n) бөлу арқылы есептеледі:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4.18)$$

MSE мәні неғұрлым аз болса, модель деректерге соғұрлым жақсы сәйкес келеді. MSE регрессия үлгісінің дәлдігі мен сапасын бағалауға мүмкіндік береді, өйткені ол қатенің шамасын да, олардың таралуын да ескереді.



Сурет 4.23 - Чен алгоритмі көмегімен бастапқы және болжамды деректердің біріктірілген графигі



Сурет 4.24 - Ченг алгоритмі көмегімен бастапқы және болжамды деректердің біріктірілген графигі

Бұл нәтиже Ченгтің анық емес уақыт қатарлары әдісінің уақыттық қатарларды болжауда жақсы нәтиже беретінін көрсетеді. Бұл әдістердің айырмашылығы анық емес жиындардың қалыптасу кезеңдерінде және анық емес қатынастардың әрбір тобында салмақтар болатынында жатыр.

R-квадрат сынағы аралас корреляцияның квадраты болып табылады, ол квадраттардың регрессия сомасының (SSR) квадраттардың жалпы сомасына (SST) қатынасы ретінде есептеледі:

$$SSR = \sum_{k=1}^n w_k (\hat{y}_k - \bar{y})^2 \quad (4.19)$$

$$SST = \sum_{k=1}^n w_k (y_k - \bar{y})^2 \quad (4.20)$$

$$R_{\text{квадрат}} = \frac{SSR}{SST} \quad (4.21)$$

R-квадрат метрикасындағы R (немесе анықтау коэффициенті) регрессия үлгісіндегі нақты мәндер мен болжамды мәндер арасындағы корреляцияны немесе сәйкестікті білдіреді. R-квадрат – регрессия үлгісінің деректерге қаншалықты сәйкес келетінін өлшейтін статистикалық көрсеткіш. Ол 0-ден 1-ге дейінгі мәнді қабылдайды, мұнда 0 модельдің күші жоқ дегенді білдіреді және 1 модель деректерге тамаша сәйкес келетінін білдіреді. R-квадрат мәні 1-ге неғұрлым жақын болса, модель нақты мәндерді болжауда соғұрлым жақсы болады.

$$\text{Adjusted Rsquare} = 1 - \frac{SSE(n-1)}{SST(n-m)} \quad (4.22)$$

Түзетілген R-квадрат - модельдегі айнымалылар санын ескеретін және артық бағалауды болдырмайтын R-квадраттың түзетілген түрі. Түзетілген R-квадрат – регрессия үлгісінің модельдің күрделілігін ескере отырып, деректерге қаншалықты сәйкес келетінін көрсететін статистикалық көрсеткіш. Жаңа айнымалылар қосылған сайын өсетін (тіпті олар статистикалық маңызды болмаса да) тұрақты R-квадраттан айырмашылығы, түзетілген R-квадрат айнымалылардың санын ескереді және түсіндіруге ықпал етпейтін айнымалыларды шамадан тыс пайдалануға тыйым салады. Түзетілген R-квадратты пайдаланудың артықшылығы – ол модельдің күрделілігін ескере отырып, болжамдық күшін объективті түрде бағалайды.

Кесте 6 - Бағалау метрикаларының салыстырмалы талдауы

Модель	MSE	R ²	R _{adj}
Нейро-анық емес Чен алгоритмі	9,0781	0,3967	0,2876
Нейро-анық емес Ченг алгоритмі	1,2041	0,9194	0,8772

Кластерлеу - бұл деректерді өндіру (Data Mining) әдістерінің бірі. Кластерлік талдау әдістері зерттелетін объектілер жиынтығын кластерлер деп аталатын «ұқсас» объектілер тобына бөлуге мүмкіндік береді.

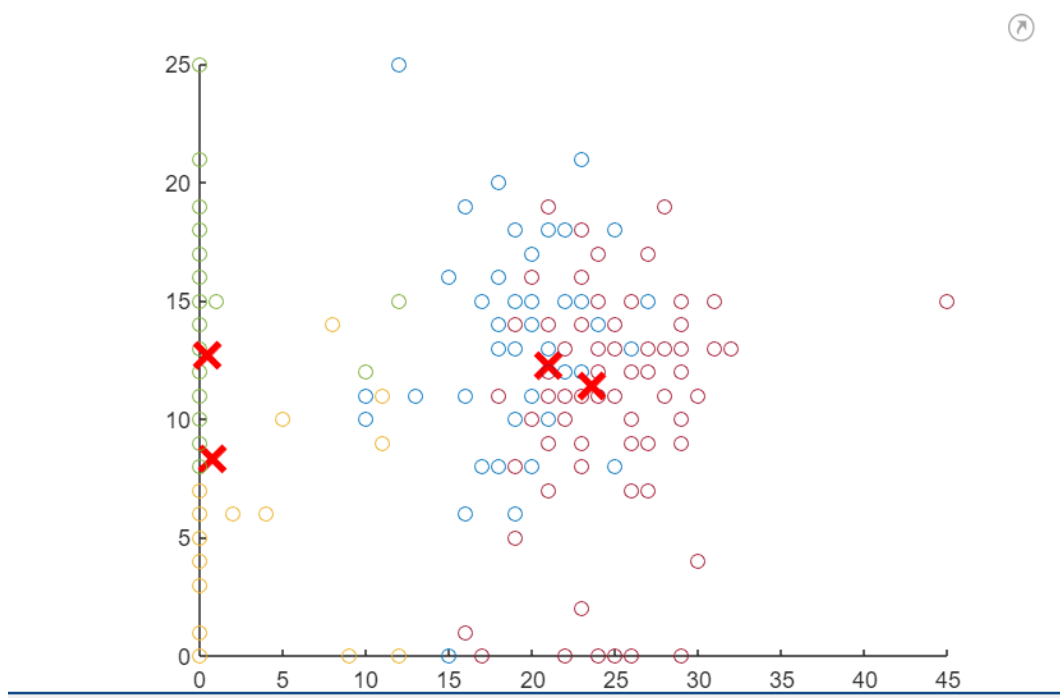
Кластерлеу алгоритмдерінің көпшілігі дәстүрлі статистикалық болжамдарға сүйенбейді; олар деректердің таралу заңдарына тәуелсіз. Кластерлеудің көптеген әдістері бар, оларды анық және анық емес деп жіктеуге болады. Анық кластерлеу әдістері қатардың бастапқы жинағын бірнеше қабаттаспайтын ішкі жиындарға бөледі. Сонымен қатар, X-тің кез келген нысаны тек бір кластерге жатады. Анық емес кластерлеу әдістері бір объектінің бір уақытта бірнеше (тіпті барлық) кластерге, бірақ әртүрлі дәрежеде тиесілі болуына мүмкіндік береді.

C-орташалардың анық емес кластеризациясы (Fuzzy C-Means Clustering - FCM) мүшелік дәрежесімен анықталатын (яғни әр нүкте белгілі бір дәрежеде кластерге жататын) деректерді кластерлеу әдісі. fcm функциясы әрбір кластердің орташа орнын көрсетуге арналған кластерлердің орталықтары туралы бастапқы болжамнан басталады. Бұл кластерлік орталықтар туралы бастапқы болжам қате болуы мүмкін.

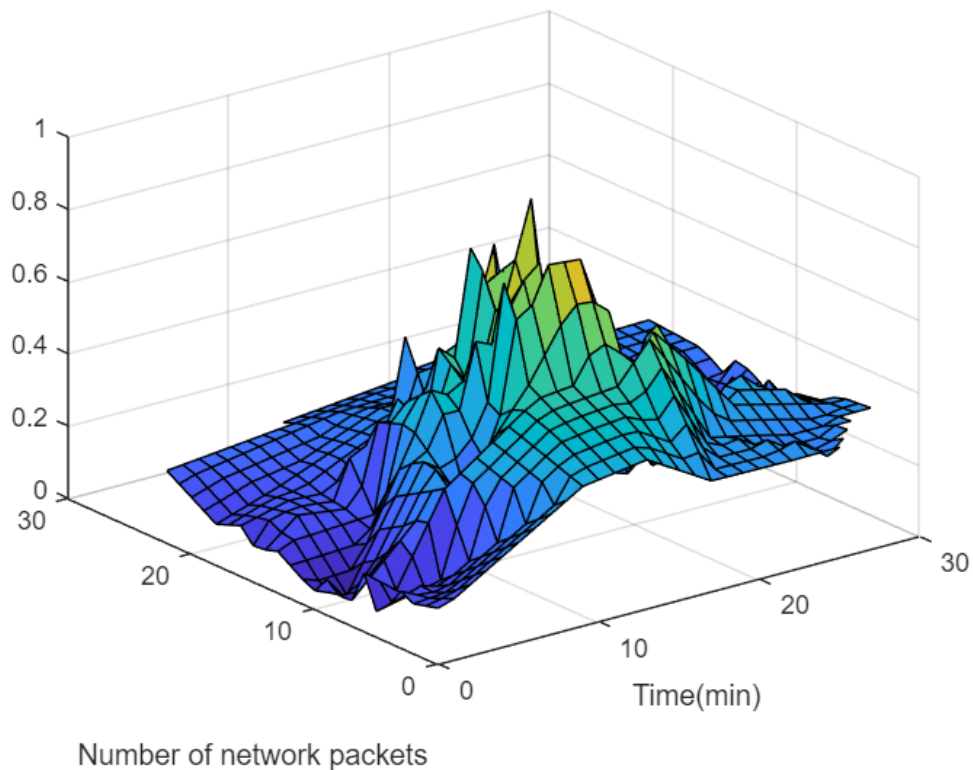
Оған қоса, fcm әрбір деректер нүктесіне әрбір кластер үшін мүшелік класын тағайындайды. Әрбір деректер нүктесі үшін кластер орталықтарын және мүшелік ұпайларын итеративті түрде жаңарту арқылы, fcm кластер орталықтарын деректер жиынындағы дұрыс орынға жылжытады. Бұл итерация мақсаттық функцияны азайтуға негізделген, ол кез келген берілген деректер нүктесінен кластердің орталығына дейінгі қашықтық, сол деректер нүктесінің мүшелігімен өлшенеді (4.25, 4.26 суреттер).

Зерттелетін қатар 4 бөлікке бөлінеді:

dataset = [X(1:450) X(451:900) X(901:1350) X(1351:1800)];



Сурет 4.25 - Кластерлеу нәтижелері (2-өлшемді)



Сурет 4.26 - Кластерлеу нәтижелері (3-өлшемді)

Гибридті нейро-анық емес желі - бұл арнайы құрылымның көп қабатты нейрондық желісі. Тәжірибеде ең кең тарағандары нейро-анық емес шығыстың адаптивті жүйелері түрінде жүзеге асырылатын гибридті ANFIS желілері. MATLAB-та олар Fuzzy Logic Toolbox пакетінде жүзеге асырылады. Гибридті желілермен жұмыс графикалық ANFIS редакторының көмегімен интерактивті түрде жүзеге асырылады. ANFIS редакторы нейро-анық тұжырымның адаптивті жүйесінің нақты моделін құруға, оны жаттықтыруға, құрылған құрылымды қарауға, оның параметрлерін өзгертуге және реттеуге, содан пайда болған нейрондық желіні кейін анық емес шығыс нәтижелерін алу үшін пайдалануға мүмкіндік береді.

ANFIS көмегімен анық емес жүйені жаттықтыру үшін Fuzzy Logic Toolbox бағдарламалық құралы кері таралу алгоритмін жеке немесе ең кіші квадраттар алгоритмімен бірге пайдаланады. Бұл оқу процесі FIS мүшелік функциясының параметрлерін жүйе кіріс/шығыс деректерін моделдейтіндей етіп реттейді.

ҚОРЫТЫНДЫ

Қазіргі заманғы зерттеулер желілік трафикті талдау және болжау трафикті басқарудағы ең маңызды міндет болып қала беретінін көрсетті. Желіні басқарудың міндеттері басқа тәсілдермен қатар дұрыс шешім қабылдау үшін болжамды болашақ деректерге де негізделген.

Инфокоммуникациялық желілердегі гетерогенді трафик көлемінің өсуі ұсынылатын байланыс қызметтерінің сапасын қамтамасыз ету мәселелерін өзекті етеді, бұл өз кезегінде болжау моделдеріне жүгінуді талап етеді.

Болжамды деректер желідегі ақпарат ағындарын басқару мәселесін шешу үшін қажетті ақпаратты береді және басқару негізінде пакеттердің жоғалуын болдырмауға мүмкіндік береді.

Бұл диссертациялық жұмыста ең алдымен желілік пакеттердің келіп түсу қарқындылығын сипаттайтын уақыттық қатар зерттелді. Зерттеу визуалды түрде және көптеген критерийлер, тесттер көмегімен, сондай ақ математикалық және программалық-кодтық түрде жүзеге асты. Нәтижелер көрсеткендей, зерттелетін қатар стационарлық емес деген тұжырым орын алды.

Стационарлық және стационарлық емес уақыт қатарларының статистикалық қасиеттері айтарлықтай ерекшеленеді және оларды модельдеу үшін әртүрлі әдістерді қолдану қажет [104]. Стационарлық уақыт қатарларының моделі уақыттың тұрақты орташа мәнімен, дисперсиясымен және автокорреляциясымен сипатталады. Стационарлық емес уақыт қатарларының моделі көпкомпонентті: трендті, маусымдық және кездейсоқ құрамдастарды қамтиды.

Уақыттық қатарды ыдырату әдісі (SSA) қатарды тренд, мерзімділік және шу ретінде түсіндіретін жиынтық құрамдас бөліктерге ыдыратады. Алынған құрамдас бөліктер талданып, деректерде болжау, шуды сүзу және аномалияны анықтау мәселелерін шешу үшін пайдаланылады. Уақыттық жиілікті спектрлік талдау уақыттық қатардың жиілік құрамдастарын ерекшелеуге және олардың жалпы деректер құрылымына қосқан үлесін анықтауға мүмкіндік берді.

Уақыттық жиілікті деректерді талдау қатардың болашақ мәндерін болжау немесе болжамдық интервалдың оңтайлы ұзақтығын анықтау үшін пайдаланылуы мүмкін.

Түрлендірулер көмегімен стационарлық емес уақыт қатарларын стационарлық жағдайға келтіру интеграциялау арқылы ARIMA (Autoregressive Integrated Moving Average) моделінің көмегімен жүзеге асты.

ARIMA әдісіне сәйкес, моделдерді құрудағы қадамдардың бірі бастапқы уақыт қатар стационарлы болғанша айырма алу болып табылады. ARIMA моделінің дұрыс параметрлерін таңдау арқылы ARIMA(0,2,1) моделі таңдалды. Ол болжау интервалы 95% барабар болжамды көрсетті.

ARIMA моделі желілік трафикті болжауды қоса алғанда, уақыттық қатарларды болжау әдісінде кеңінен қолданылады. Дегенмен, оның кейбір кемшіліктері де бар:

1. Параметрлерді таңдауға тәуелділік: ARIMA моделі үш негізгі параметрді

баптауды талап етеді: авторегрессивті параметр (p), айырмашылық параметрі (d) және жылжымалы орташа параметрі (q). Бұл параметрлер үшін оңтайлы мәндерді таңдау күрделі болуы мүмкін және оларды автоматты түрде таңдау сияқты басқа әдістерді пайдалануды талап етеді.

2. Бір айнымалы деректерге шектеу: ARIMA моделі желі трафигінде болуы мүмкін көп айнымалы қатынастарды немесе әртүрлі айнымалылар арасындағы қатынастарды модельдеу мүмкіндігін шектейтін бір айнымалы уақыт қатарларына ғана қолданылады.
3. Деректердің үлкен көлемін өңдеудің қиындығы [105-107]: егер желілік трафикте деректердің үлкен көлемі болса, ARIMA моделі есептеу тұрғысынан күрделі болуы мүмкін және оқыту мен болжау үшін үлкен көлемдегі есептеу ресурстары мен уақытты қажет етеді.

Желілік трафикті болжау үшін күрделі тәуелділіктерді тиімдірек есепке алатын [108] және деректердің үлкен көлемін [109] өңдейтін терең оқыту модельдері, мысалы қайталанатын нейрондық желілер, жасанды нейронды экзогенді кірісі бар желі (NARX желісі) сияқты балама әдістер қолайлы болуы мүмкін.

NARX желісі - тікелей сигнал беру және шығыс бойынша кері байланысы бар көп деңгейлі желі, оның шығысы уақыт бойынша кідіріс векторы арқылы өтеді. Сигма тәрізді функция белсендіру функциясы (функция активации) ретінде қолданылады.

Желінің әрбір қабаты белгілерді енгізу кеңістігін басқа өлшеммен басқа кеңістікке түрлендіреді. Мұндай сызықты емес түрлендіру шығыс қабатының сызықты түрде бөлінетін нейрондары болғанша жалғасады. Кіріс пен шығыстан басқа, ЖНЖ барлық деңгейлері; желіге сызықтық емес құбылыстарды моделдеу мүмкіндігін береді [110,111].

Нейрондық желіні оқыту үшін Левенберг-Маквардт алгоритмі қолданылды, ал нейрондық желінің өнімділігін бағалау үшін орташа квадраттық қате (MSE) қолданылады.

MSE-ге шектен тыс көрсеткіштер әсер етуі мүмкін және бұл метрика деректер масштабына сезімтал болуы мүмкін екенін есте ұстаған жөн. Сондықтан модельдерді салыстыру немесе бағалау үшін MSE пайдалану кезінде оны басқа көрсеткіштермен бірге қарастыру және зерттелетін деректердің белгілі бір ерекшеліктерін ескеру қажеттігі айқындалды.

Гибридті желі – бұл анық емес логикалық әдістер мен жасанды нейрондық желі әдістерін біріктіретін желі [112].

Әртүрлі процестерді модельдеу [113] кезінде нейрондық желілердің мүмкіндіктері ақпаратты өңдеу технологиялары анық емес жиындар мен анық емес қорытындыға негізделсе, күшеюі мүмкін.

Қорытындылай келе, диссертацияның тақырыбы өте өзекті және қазіргі ақпарат ағынының қарқынды өсуімен оның өзектілігі күн сайын арта түспек. Әдістер мен модельдер жан-жақты қарастырылып, зерттелетін нысанға сәйкес ерекшеліктерді ескере отырып, дұрыс модель таңдау қажеттілігі айқындалды.

ПАЙДАЛАНЫЛҒАН ӘДЕБИЕТТЕР ТІЗІМІ

- 1 Рейтинг стран мира по индексу сетевой готовности. – URL: <https://gtmarket.ru/ratings/networked-readiness-index>. 21.12.2022.
- 2 Иванов А. В. Разработка и исследование алгоритмов прогнозирования и управления очередями в компьютерных сетях: дис. ... канд. тех. наук: 05.13.01 – Санкт-Петербург, 2001. – 147 с.
- 3 Колесников А. В. Моделирование сетевого трафика и алгоритмы борьбы с перегрузками на основе методов нелинейной динамики и краткосрочного прогнозирования временных рядов: автореф. ... канд. тех. наук: 05.13.15 – Москва, 2015. – 22 с.
- 4 Сан В.А. Исследование и разработка метода прогнозирования трафика при проектировании компьютерной сети: автореф. ... канд. тех. наук: 05.13.15 – Москва, 2012. – 20 с.
- 5 Марьенков А. Н. Управление трафиком вычислительной сети на основе идентификации аномалий: дис. ... канд. тех. наук: 05.13.01 – Астрахань, 2012. – 196 с.
- 6 Датьев И. О. Метод и имитационная модель прогнозирования характеристик региональных информационно-коммуникационных систем: дис. ... канд. тех. наук: 05.13.18 – Москва, 2011. – 190 с.
- 7 Афанасьева К. Е. Метод и имитационная модель прогнозирования характеристик региональных информационно-коммуникационных систем: дис. ... канд. тех. наук: 05.13.01 – Челябинск, 2008. – 161 с.
- 8 Ляхманов Д. А. Прогнозирование состояний технических объектов на основе спектрального анализа: автореф. ... канд. тех. наук: 05.13.01 – Нижний Новгород, 2013. – 21 с.
- 9 Станкевич А. А. Разработка дисциплины обслуживания на основе нейросетевого прогноза трафика дифференцированных услуг: дис. ... канд. тех. наук: 05.12.13 – Санкт-Петербург, 2006. – 122 с.
- 10 Покровская М.А. Метод прогнозирования изменения трафика с использованием нейросетевой модели // Т-Comm. - 2012. – № 6. – С. 27-29.
- 11 Yazeed A., Al-Sbou and Khaled M. Alawasa. Nonlinear Autoregressive Recurrent Neural Network Model For Solar Radiation Prediction // International Journal of Applied Engineering. – 2017. – Vol.12, № 14. – P. 4518-4527.
- 12 Mohammed L. B., Hamdan M. A., Abdelhafez E. A. and Shaheen W. Hourly Solar Radiation Prediction Based on Nonlinear Autoregressive Exogenous (NARX) Neural Network // Jordan Journal of Mechanical and Industrial Engineering (JJMIE). – 2013. – Vol.7, № 1. – P. 11-18.
- 13 Piazza D., Di Piazza M. C. and Vitale G. Solar and wind forecasting by NARX neural networks // Renew. Energy Environ. Sustain. – 2016. – Vol.1, №39. – P. 1-5.
- 14 Соколов А. С. Метод прогнозирования перегрузок в компьютерных сетях на основе анализа временных рядов: автореф. ... канд. тех. наук: 05.13.19 – Санкт-Петербург, 2011. – 21 с.

- 15 Нарыжная Н. Ю. Разработка методического обеспечения для прогнозирования технического состояния корпоративной сети: автореф. ... канд. тех. наук: 05.13.01 – Краснодар, 2009. – 21 с.
- 16 Сизов А. А. Модели, способы и программные средства поддержки принятия решений на основе прогнозирования временных рядов с переменной структурой: дис. ... канд. тех. наук: 05.13.17 – Смоленск, 2014. – 139 с.
- 17 Чучуева И. А. Модель прогнозирования временных рядов по выборке максимального правдоподобия: дис. ... канд. тех. наук: 05.13.18 – Москва, 2012. – 146 с.
- 18 Овчинников К.А. Прогнозирование сетевого трафика при помощи авторегрессионных моделей // Материалы первой МНПК «Проблемы инфокоммуникаций. Наука и технологии» - Харьков. 2013. – С.177-179.
- 19 Крюков Ю.А. ARIMA – модель прогнозирования значений трафика: // Информационные технологии и вычислительные системы. - 2011. – Т.16, № 2. – С. 41-49.
- 20 Игнатенко Е.Г. Методика краткосрочного прогнозирования трафика телекоммуникационных сетей // Збірник наукових праць ДонІЗТ. – Д., 2011. – № 28. – С. 102-108.
- 21 Босько В.В. Разработка метода прогнозирования поведения информационного потока в телекоммуникационной сети // Харківського національного університету Повітряних Сил. - 2010. – Т.4, № 3. – С.126-130.
- 22 Коннов А. Л. Методы расчета показателей производительности сетей ЭВМ с неоднородным трафиком / Оренбургский гос. ун-т. – Оренбург: ОГУ, 2013. – 139 с.
- 23 Кирьянова Л.В. Теория случайных процессов / под ред. Нац. исследоват. Моск. гос. строит. ун-т. – М.: Изд-во Моск. гос. строит. ун-та. - Мир, 2016. – 96 с.
- 24 Лоскутов А.Ю. Анализ временных рядов. – М: Физический факультет МГУ, 2012. – 113с.
- 25 Канторович Г.Г. Анализ временных рядов // Экономический журнал ВШЭ, 2002. - №3. – С. 192-201.
- 26 Афанасьев В.Н. Анализ временных рядов и прогнозирование: учебник. – М.: Финансы и статистика, 2001. – 228 с.
- 27 Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов: учебное пособие для вузов. – М.: Финансы и статистика, 2003. – 416 с.
- 28 Rath B. N., & Akram V. (2021). Popularity of Unit Root Tests: A Review // Asian Economics Letters. – 2021. – Vol.4, №2. – P. 305-324.
- 29 Драница Ю.П. Быстрый алгоритм построения нестационарной векторной линейной авторегрессии. – Мурманск: МГТУ, 2011. – 137 с.
- 30 Бокс Дж. Анализ временных рядов, прогноз и управление: пер. с англ. Дж. Бокс, Дженкинс Г. / под ред. В.Ф. Писаренко. – М.: Мир, 1974. – 406 с.

- 31 Любушин А.А. Анализ структуры временных рядов многолетних колебаний речного стока // Труды Британско-Российской конференции – Барнаул: Пять плюс, 2009. – С. 91-102.
- 32 Яковлев П. В. Анализ пространственно-временных особенностей временных рядов. – Москва, 2016. – 115 с.
- 33 Носко В.П. Эконометрика. – М.: «Дело» РАНХиГС, 2011. – 672 с.
- 34 Афанасьева Т.В. Прогнозирование временных рядов нечеткие модели. – Ульяновск: УлГТУ, 2014. – 145 с.
- 35 Ярушкина Н.Г. Интеллектуальный анализ временных рядов: учебное пособие. – Ульяновск: УлГТУ, 2010. – 320 с.
- 36 Бендат Дж. Прикладной анализ случайных данных: Пер. с англ. – М.: Мир, 1989. – 540 с.
- 37 Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Основы моделирования и первичная обработка данных. – М.: Финансы и статистика, 1983. – 470 с.
- 38 Вершинина М.С. Анализ предположений о стационарности некоторых временных рядов // Сборник трудов Всероссийской конференции по математике с международным участием "МАК-2018" / АлтГУ и др. – Барнаул: Изд-во АлтГУ, 2018. – С. 172-176.
- 39 Бектемысова Г.У., Ибраева Ж.Б. Исследование временного ряда на стационарность // Вестник Национальной инженерной академии РК. – Алматы, 2022. – №4(86). – С. 20-27.
- 40 Вес F., & Alain G. A simple unit root test consistent against any stationary alternative // Working Papers. – 2020. – №28, – P. 105-114.
- 41 Бектемысова Г.У., Ибраева Ж.Б., Кулакаева А.Е., Кожакметова Б.А. Анализ измеренного сетевого трафика на стационарность // Вестник Казахской Академии Транспорта и Коммуникации им. М.Тынышпаева. – Алматы, 2022. – № 3 (122). – С. 302-308.
- 42 ADF test - interpreting the results. URL: <https://stats.stackexchange.com/questions/265603/adf-test-interpreting-the-results> 16.04.2022.
- 43 Amsler C. and Lee J. An LM Test for a Unit Root in the Presence of a Structural Change // Econometric Theory. –1995. – Vol.11, – P. 359-368.
- 44 Кратович П. В. Предпрогнозный анализ временных рядов финансовых данных на основе методов фрактального анализа // Молодой ученый. – 2010. – Т. 1, №2. – С. 11-18.
- 45 Поршнева С.В. Исследование особенностей применения метода сингулярного спектрального анализа в задаче анализа и прогнозирования временных рядов: Монография: – Ульяновск: Зебра, 2016. – 167с.
- 46 Голяндина Н.Э. Метод «Гусеница»-SSA анализ временных рядов: Учебное пособие / – СПб., 2004. – 76 с.
- 47 Han H., Zha H., Giles C.L. Name disambiguation in author citations using a K-way spectral clustering method // Digital Libraries - 2005. JCDL '05. Proceedings of the 5th ACM/IEEE-CS Joint Conference. - IEEE, 2007. – P. 334-343.

- 48 Матвеев М.Г. Применение метода сингулярно-спектрального анализа для идентификации сигналов электрической активности мозга // Вестник ВГУ, Серия: Системный анализ и информационные технологии, - 2012, – №2. – С. 70-78.
- 49 Василишин В.И. Оценивание числа гармонических компонент сигнала с использованием технологии суррогатных данных // Прикладная радиоэлектроника. – 2013. – Т.12, №4. – С. 542-552.
- 50 Фаткиева Р. Р., Левоневский Д.К. Детектирование атак методом сингулярного спектрального разложения // Труды СПИИРАН. – 2013. № 2 (25). – С. 135-147.
- 51 Арутюнов А.Г. Прогнозирование емкости рынка летательных аппаратов с использованием методов «Гусеница» SSA и «Анализ кривой выживаемости» // Journal of applied informatics. – 2017. – Vol.12, №2. – С. 118-128.
- 52 Decomposing Time Series Data with Singular-Spectrum Analysis. URL: <https://www.kaggle.com/jdarcy/introducing-ssa-for-time-series-decomposition>. 25.09.22.
- 53 Поршнева С.В. Исследование особенностей применения метода сингулярного спектрального анализа в задаче анализа и прогнозирования временных рядов. – Ульяновск: Зебра. 2016. – 167 с.
- 54 Голяндина Н.Э. Метод «Гусеница»-SSA: анализ временных рядов: Учеб.пособие. – СПб., 2004. – 76 с.
- 55 Безрукавный Д.С. О возможности применения алгоритма «Гусеница» к задаче контроля сетевого трафика // Научные разработки: евразийский регион: материалы международной научной конференции теоретических и прикладных разработок. – Москва: Издательство Инфинити, 2019. – С. 236-240.
- 56 Матвеев М.Г. и др. Применение метода сингулярно-спектрального анализа для идентификации сигналов электрической активности мозга // Вестник ВГУ, Серия «Системный анализ и информационные технологии». – 2012. № 2. С. 42–47.
- 57 Василишин В.И. Оценивание числа гармонических компонент сигнала с использованием технологии суррогатных данных // Прикладная радиоэлектроника. – 2013. – Т.12., №4. – С. 542-552.
- 58 Гайдышев И.П. Моделирование стохастических и детерминированных систем: Руководство пользователя программы AtteStat / – Курган, 2013. – 490с.
- 59 Groth A., and M. Ghil, 2015: Monte Carlo Singular Spectrum Analysis (SSA) revisited. Detecting oscillator clusters in multivariate datasets // Journal of Climate. – 2015. – Vol. 2, №8. – P. 7873-7893.
- 60 Метод расчета критерия Колмогорова-Смирнова для оценки нормальности распределения. URL: <http://statyx.ru/metod-rascheta-kriteriya-kolmogorova-smirnova-dlya-otsenki-normalnosti-raspredeleniya> 25.11.2021
- 61 Dai M. Analysis and modeling of MPEG-4 and H.264 multilayer video traffic // Proceedings – IEEE. – INFOCOM, 2005. – Vol.4, – P. 2257-2267.

- 62 Мараев В.С. Инструменты визуализации временных рядов в космических исследованиях // Исследования наукограда. – 2017. –Т. 1, № 4. – С. 200-207.
- 63 Мефферт Б. Инструменты обработки сигналов – основы, примеры применения и задачи. Авторы перевода Г.Безель, В.Баербах. – 2018. – 320с.
- 64 Носко В.П. Эконометрика. Введение в регрессионный анализ временных рядов. – Москва: НФПК, 2002. – 271 с.
- 65 Сычев В.Н. Применение аппарата нелинейного анализа динамических систем для обработки экспериментальных данных // Вестник КРСУ. – 2011. Т. 11, № 4. – С. 91-96.
- 66 Schreiber T., Kantz H. Noise in chaotic data: Diagnosis and treatment // CHAOS. – 1995. – Vol. 5. – P. 133-142.
- 67 Data Science. Анализ временных рядов URL: <https://ppt-online.org/431804>. 30.04.2022.
- 68 Bektemyssova G., Abdul R., Mirzakulova Sh., Ibraeva Zh. Time series forecasting by the Arima method // Scientific Journal of Astana IT University. – Almaty, 2022. – Vol.11, – P. 14-23.
- 69 Бестугин А. Р. Контроль и диагностирование телекоммуникационных сетей. – СПб: Политехника, 2003. – 174с.
- 70 Ибраева Ж.Б., Мирзакулова Ш.А. Analysis of a one-dimensional time series for a trend // Материалы международной научной конференции молодых ученых, ИМА-2022. – Суми-Нур-Султан, 2022, – С. 258-259.
- 71 Ибраева Ж.Б., Мирзакулова Ш.А. Network traffic analysis using Leybourne-McCabe test // Материалы международной научной конференции молодых ученых, ИМА-2022, – Суми-Нур-Султан, 2022, – С. 259-261.
- 72 KPSS test for stationarity – MATLAB. URL: <https://www.mathworks.com/help/econ/kpsstest.html> 19.01.2022.
- 73 Branitsky A. A., Kotenko I. V. Analysis and Classification of Network Attack Detection Methods // Proc. of SPIIRAS. – 2016. – №2(45). – P. 207-244.
- 74 Бокс Дж. Анализ временных рядов, прогноз и управление: Пер. с англ. Дж. Бокс, Дженкинс Г. / Под ред. В.Ф. Писаренко. – М.: Мир, 1974, кн. 1. – 406 с.
- 75 Rutka G. Network Traffic Prediction using ARIMA and Neural Networks Models // Elektronika Ir Elektrotechnika. – 2008. – Vol.84, №4. – С. 53-58.
- 76 Шарипбай А.А. Нейронные сети: учебное пособие. – Алматы: Эверо, 2017. – 278 с.
- 77 Николаева С.Г. Нейронные сети. Реализация в Matlab. Учебное пособие. – Казань: Казанский государственный энергетический университет, 2015. – 92 с.
- 78 Белов М.П. Разработка рекуррентных нейронных сетей в задачах идентификации систем управления следящими электроприводами крупного радиотелескопа // Известия СПбГЭТУ «ЛЭТИ». – 2018. – №8. – С. 270-278.

- 79 Boussaada Z. A Nonlinear Autoregressive Exogenous (NARX) Neural Network Model for the Prediction of the Daily Direct Solar Radiation // *Energies*. – 2018. – № 11. – P. 620-628.
- 80 Бодянский Е.В. Искусственные нейронные сети: архитектуры, обучение, применения. – Харьков: ТЕЛТЕХ, 2004. – 369 с.
- 81 Rehak M. et al. Adaptive Multiagent System for Network Traffic Monitoring // *IEEE Intelligent Systems*. – 2009. – Vol.24, №3. – P. 16-25.
- 82 Anu Gowsalya R. S., Miruna Joe Amali S. SVM Based Network Traffic Classification Using Correlation Information // *International Journal of Research in Electronics and Communication Technology* – 2014. – №3(35). – P. 107-114.
- 83 Singh J., Nene M. J. A Survey on Machine Learning Techniques for Intrusion Detection Systems // *Intern. Journal of Advanced Research in Computer and Communication Engineering*. – 2013. – Vol.2, №11. –P. 4349-4355.
- 84 Abraham S., Nair S. Cyber Security Analytics: A Stochastic Model for Security Quantification Using Absorbing Markov Chains // *Journal of Communications*. – 2014. – Vol.9, №12. –P. 899-907.
- 85 Szabo G. et al. Multi-Level Machine Learning Traffic Classification System // *The Eleventh International Conference on Networks*. - 2012. – P. 69-77.
- 86 Sun R. et al. Traffic Classification Using Probabilistic Neural Networks // *Sixth Intern. Conf. on Natural Computation (ICNC 2010)*. – 2010. – P. 1914-1919.
- 88 Прогнозирование производственной программы выпуска продукции с учетом предпрогнозного анализа временных рядов URL: <https://http://www.konspekt.biz/index.php?text=481>
30.05.2022.
- 89 Капитонова О.В. Прогнозирование социально-экономических процессов: Учебно-методическое пособие. – Нижний Новгород: Нижегородский университет, 2016. – 74 с.
- 90 Коннова Н. С. Методы и алгоритмы обработки, анализа и визуализации данных ультразвукового доплеровского измерителя скорости кровотока: дис. ... канд. тех. наук: 05.13.01 – Москва, 2017. – 205 с.
- 91 Serikov T., Zhetpisbayeva A., Mirzakulova S., Zhetpisbayev K., Ibraeva Z., Soboleva L., Tolegenova A., & Zhumazhanov B. Application of the NARX neural network for predicting a one-dimensional time series // *Eastern-European Journal of Enterprise Technologies*. – 2021. Vol. 5., №4 (113). – P. 12-19.
- 92 Li X., Yang C. (2014). Research and Application of Data Mining and NARX Neural Networks in Load Forecasting // *International Journal of Database Theory and Application*. –2014. Vol.7, №2. – P. 13-24.
- 93 Chandra R. Competition and Collaboration in Cooperative Coevolution of Elman Recurrent Neural Networks for TimeSeries Prediction // *IEEE Transactions on Neural Networks and Learning Systems*. – 2015. – №26 (12), – P. 3123-3136.
- 94 Havaluddin Alfred, R. Performance of modeling time series using nonlinear autoregressive with eXogenous input (NARX) in the network traffic forecasting // *International Conference on Science in Information Technology (ICSITech)*. – 2015. – P. 87-94.

- 95 Boussaada Z., Curea O., Remaci A., Camblong H., Mrabet Bellaaj N. (2018). A Nonlinear Autoregressive Exogenous (NARX) Neural Network Model for the Prediction of the Daily Direct Solar Radiation // *Energies*, – 2018. – Vol.11, №3. – P. 620-628.
- 96 Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. – М: Горячая линия – Телеком, 2006. – 452 с.
- 97 Воронина В.В. Теория и практика машинного обучения: учебное пособие. – Ульяновск: УЛГТУ, 2017. – 290 с.
- 98 Малинецкий Г.Г. Нелинейная динамика и проблемы прогноза // *Вестник Российской Академии Наук*. – 2001. – Т. 71. № 3. – С. 210-232.
- 99 Chen S.M. Forecasting enrollment based on high-order fuzzy time series // *Cybernetic Systems*. – 2002. – Vol.33, № 1. – P.1-16.
- 100 Yu Hui-Kuang. Weighted fuzzy time series models for TAIEX forecasting // *Physica A: Statistical Mechanics and its Applications*. – 2005. – Vol. 349, № 3. – P. 609-624.
- 101 Flavio Manenti, Francesco Rossi, Alexey G. Goryunov, Valeriy F. Dyadik, Kirill A. Kozin M., Igor S. Nadezhdin, Sergei S. Mikhalevich. Fuzzy adaptive control system of a non-stationary plant with closed-loop passive identifier // *Resource-Efficient Technologies*. – 2015. № 1. – P.10-18.
- 102 M. El-Bardini, A.M. El-Nagar. Interval type-2 fuzzy PID controller for uncertain nonlinear inverted pendulum system // *ISA Trans*. – 2014. Vol.53, №3. – P. 732-743.
- 103 Пегат А. Нечеткое моделирование и управление, пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2009. – 312 с.
- 104 Бектемысова Г.У., Ибраева Ж.Б. Возможности применения искусственного интеллекта в строительстве // *Вестник КазГАСА*. – Алматы, 2018. – Т. 68, № 2. – С. 205-212.
- 105 Бектемысова Г.У., Ибраева Ж.Б., Луганская С.П., Миркасымова Т.Ш. инструментов МАТЛАВ для анализа больших данных по энергоэффективности зданий // *Вестник КБТУ*. – Алматы, 2019. – Т. 16, №3. – С. 324-328.
- 106 Аймагамбетова З.Т., Ибраева Ж.Б. Аспекты формирования комфортной городской среды // *Вестник КазГАСА*. – Алматы, 2021. – Т. 81, №3. С. 15-20.
- 107 Aimagambetova Z.T, Bektemyssova G.U, Ibraeva Zh.B. Buildings energy consumption modeling methods // *STCCE-2020 IOP Conf. Series: Materials Science and Engineering (Scopus)*. – Kazan, 2020. Vol. 890, – P.14-23.
- 108 Ибраева Ж.Б., Айтжанов Д., Каттабек А. Анализ и оптимизация сетевого трафика // *Международный журнал информационных и коммуникационных технологий*, – Алматы, 2022. – С. 166-170.
- 109 Ибраева Ж.Б., Мирзакулова Ш.А. Analysis of a one-dimensional time series for a trend // *Материалы международной научной конференции молодых ученых, ИМА-2022*, – Суми-Нур-Султан, 2022, – С. 258-259.

110 Ибраева Ж.Б., Мирзакулова Ш.А. Network traffic analysis using Leybourne-McCabe test // Материалы международной научной конференции молодых ученых, IMA-2022. – Суми-Нур-Султан, 2022, – С. 259-261.

111 Ибраева Ж.Б., Миркасимова Т.Ш. Мәліметтерді басқарудың ертеңі мен бүгіні // Сборник материалов международной научно-методической конференции «Современные концепции науки и образования». – Алматы, 2017. – С. 56-59.

112 Mirzakulova Sh., Ibraeva Zh. Clustering Time Series Data // Материалы международной научной конференции молодых ученых, IMA-2023, – Суми-Нур-Султан, 2023, – С. 397-398.

113 Доскен Б., Ибраева Ж.Б. Компьютерное моделирование усилителя сигналов // Материалы международной научной конференции молодых ученых, IMA-2023. – Суми-Нур-Султан, 2023, – С. 399-400.

ҚОСЫМША А

Стационарлықты анықтайтын бағдарлама бөлігі

```
import openpyxl
def get_sign_sequence(numbers):
    sequence = []
    for i in range(len(numbers) - 1):
        diff = numbers[i+1] - numbers[i]
        if diff > 0:
            sequence.append('+')
        elif diff < 0:
            sequence.append('-')
        else:
            if numbers[i] > 0:
                sequence.append('+')
            elif numbers[i] < 0:
                sequence.append('-')
            else:
                # Если уi == 0, онда таңба қоспаймыз
                pass
    # Соңғы элементке таңба қосамыз
    if numbers[-1] > 0:
        sequence.append('+')
    elif numbers[-1] < 0:
        sequence.append('-')
    return sequence
# файлды оқу
workbook = openpyxl.load_workbook('18000.xlsx')
worksheet = workbook.active
for col in worksheet.iter_cols(values_only=True):
    numbers = [int(x) for x in col]
    sequence = get_sign_sequence(numbers)
    print(numbers)
    print(sequence)
def get_series_info(numbers):
    series_count = 0
    current_series_length = 0
    max_series_length = 0
    last_sign = '+' if numbers[0] > 0 else '-'
    for i in range(1, len(numbers)):
        sign = '+' if numbers[i] - numbers[i-1] > 0 else '-' if numbers[i] - numbers[i-1] <
0 else last_sign
        if sign != last_sign:
            series_count += 1
```

```

    current_series_length = 0
else:
    current_series_length += 1
    max_series_length = max(max_series_length, current_series_length)
last_sign = sign
return series_count, max_series_length

for col in worksheet.iter_cols(values_only=True):
    numbers = [int(x) for x in col]
    series_count, max_series_length = get_series_info(numbers)
    print(f"Сериялар саны: {series_count}, Ең ұзын серияның саны: {max_series_length}")
# өсетін және кемитін қатарларды анықтау
num_up_series = 0
num_down_series = 0
prev_val = df.iloc[0, 0]
for i in range(1, len(df)):
    curr_val = df.iloc[i, 0]
    if curr_val > prev_val:
        num_up_series += 1
    elif curr_val < prev_val:
        num_down_series += 1
    prev_val = curr_val
# күтілетін мән мен дисперсияны анықтау
n = len(df)
expected_up = (n * (n + 1)) / 4
expected_down = (n * (n + 1)) / 4
var_up = (n * (n + 1) * (2 * n + 1)) / 24
var_down = (n * (n + 1) * (2 * n + 1)) / 24

# z-статистиканы анықтау
z_up = (num_up_series - expected_up) / (var_up ** 0.5)
z_down = (num_down_series - expected_down) / (var_down ** 0.5)

# p-мәндерді анықтау
p_up = 2 * norm.sf(abs(z_up))
p_down = 2 * norm.sf(abs(z_down))

# сынақ нәтижелерін шығару
if p_up > 0.05 and p_down > 0.05:
    print('Уақыттық қатарда тренд жоқ, қатар стационарлы')
else:
    print('Уақыттық қатарда тренд бар, қатар стационарлы емес')

```

ҚОСЫМША Ә

ARIMA моделімен болжау кодының бөлігі

```
Mdl = ARIMA_VarName21DiffDiff;
idxpre = 1:17500;
idxest = 17500:18000;
EstMdl = estimate(Mdl,VarName2(idxest),...
'Y0',VarName2(idxpre));
yf0 = VarName2(idxest(end - 1:end));
yf = forecast(EstMdl,500,yf0);
figure
h1 = plot(VarName2(1:17500));
hold on
h2 = plot(VarName2(17501:18000),yf,'r');
legend([h1 h2],"Бастапқы қатар","Болжам",...
'Location','NorthWest')
title("DataTable.VarName2")
xlabel("Уақыт (s)")
ylabel("Пакеттер")
hold off

rng('default')
figure

plot(X2DiffDiff,'Color',[.75,.75,.75])
hold on
title('ARIMA(0,2,1) болжау моделінің нәтижелері')
[Yf,YMSE] = forecast(Mdl,1000,X2);
upper = Yf + 1.96*sqrt(YMSE);
lower = Yf - 1.96*sqrt(YMSE);

y2 = 18001:19000;

plot(18001:19000,Yf,'r','LineWidth',3)
plot(18001:19000,[upper,lower],'k--','LineWidth',2)
xlim([0,19000]);
ylim([45,75]);
legend({'Бастапқы қатар','Болжау бақылаулары','жоғарғы және төменгі 95% бол-
жау интервалдары'})

Mdl1 = estimate(Mdl,X2);
pmse = mean((y2-yF).^2);
```

ҚОСЫМША Б

NARX (Nonlinear AutoRegressive exogenous Network) көмегімен болжау
кодының бөлігі

```
clc
clear all
close all

Data_Inputs=xlsread('18000.xlsx'); % Import file
% The training data sample are randomized by using the function'randperm'
Shuffling_Inputs=Data_Inputs(randperm(end),1:2); % integers (training sample)
Training_Set=Data_Inputs(1:end,1);% specific training set
Target_Set=Data_Inputs(1:end,2); % specific target set
Input=Training_Set'; % Convert to row
Target=Target_Set'; % Convert to row
X = con2seq(Input); % Convert to cell
T = con2seq(Target); % Convert to cell
%% 2. Data preparation
N = 365; % Multi-step ahead prediction
% Input and target series are divided in two groups of data:
% 1st group: used to train the network
inputSeries = X(1:end-N);
targetSeries = T(1:end-N);
inputSeriesVal = X(end-N+1:end);
targetSeriesVal = T(end-N+1: end);
% Create a Nonlinear Autoregressive Network with External Input
delay = 2;
inputDelays = 1:2;
feedbackDelays = 1:2;
hiddenLayerSize = 100;
net = narxnet(inputDelays,feedbackDelays,hiddenLayerSize);
% Prepare the Data for Training and Simulation
% The function PREPARETS prepares timeseries data for a particular network,
% shifting time by the minimum amount to fill input states and layer states.
% Using PREPARETS allows you to keep your original time series data un-
changed, while
% easily customizing it for networks with differing numbers of delays, with
% open loop or closed loop feedback modes.
[inputs,inputStates,layerStates,targets] = preparets(net,inputSeries,{ },tar-
getSeries);
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
```

```

% Train the Network
[net,tr] = train(net,inputs,targets,inputStates,layerStates);
% Test the Network
outputs = net(inputs,inputStates,layerStates);
errors = gsubtract(targets,outputs);
performance = perform(net,targets,outputs)
% View the Network
view(net)
% Plots
% Uncomment these lines to enable various plots.
% figure, plotperform(tr)
% figure, plottrainstate(tr)
% figure, plotregression(targets,outputs)
% figure, plotresponse(targets,outputs)
% figure, ploterrcorr(errors)
% figure, plotinerrcorr(inputs,errors)
% Closed Loop Network
% Use this network to do multi-step prediction.
% The function CLOSELOOP replaces the feedback input with a direct
% connection from the outout layer.
netc = closeloop(net);
netc.name = [net.name ' - Closed Loop'];
view(netc)
[xc,xic,aic,tc] = preparets(netc,inputSeries,{ },targetSeries);
yc = netc(xc,xic,aic);
closedLoopPerformance = perform(netc,tc,yc)
% Early Prediction Network
% For some applications it helps to get the prediction a timestep early.
% The original network returns predicted  $y(t+1)$  at the same time it is given  $y$ 
(t+1).
% For some applications such as decision making, it would help to have predicted
%  $y(t+1)$  once  $y(t)$  is available, but before the actual  $y(t+1)$  occurs.
% The network can be made to return its output a timestep early by removing one
delay
% so that its minimal tap delay is now 0 instead of 1. The new network returns the
% same outputs as the original network, but outputs are shifted left one timestep.
nets = removedelay(net);
nets.name = [net.name ' - Predict One Step Ahead'];
view(nets)
[xs,xis,ais,ts] = preparets(nets,inputSeries,{ },targetSeries);
ys = nets(xs,xis,ais);
earlyPredictPerformance = perform(nets,ts,ys)
%% 5. Multi-step ahead prediction
inputSeriesPred = [inputSeries(end-delay+1:end),inputSeriesVal];

```

```
targetSeriesPred = [targetSeries(end-delay+1:end), con2seq(nan(1,N))];  
[Xs,Xi,Ai,Ts] = preparets(netc,inputSeriesPred,{ },targetSeriesPred);  
yPred = netc(Xs,Xi,Ai);  
perf = perform(net,yPred,targetSeriesVal);  
figure;  
plot([cell2mat(targetSeries),nan(1,N);  
      nan(1,length(targetSeries)),cell2mat(yPred);  
      nan(1,length(targetSeries)),cell2mat(targetSeriesVal)]')  
legend('Original Targets','Network Predictions','Expected Outputs');
```

ҚОСЫМША В

Нейро-анық емес болжау кодының бөлігі

```
importing libraries into the current environment

import dill
from scipy.spatial import KDTree

# The "pyFTS.partition" module to split the time series into non-overlapping intervals
from pyFTS.data import Enrollments
from pyFTS.partitions import Grid

from pyFTSmodels import chen, cheng
import matplotlib.pyplot as plt
from openpyxl import load_workbook

wb = load_workbook('C:/Users/Жанар/Desktop/Докторантура/МША/РядМРЕG.xlsx')
sheet = wb.get_sheet_by_name('Лист1')

test = []

for row in range(1, 18001):
    test.append(sheet.cell(row=row, column=2).value)

train = test[:]

#train = Enrollments.get_data()

#test = Enrollments.get_data()

count_intervals = 10

#Universe of Discourse Partitioner
#A class containing all the entities referred to in a discourse or argument.
partitioner = Grid.GridPartitioner(data=train, npart=count_intervals)

fig, ax = plt.subplots(nrows=1, ncols=1, figsize=[15,5])

partitioner.plot(ax)
plt.show()

# Create an empty model using the Chen method
model = chen.TrendWeightedFTS(partitioner=partitioner)
```



```

# The training procedure is performed by the method fit
model.fit(train)

# Print the model rules
print('Term expert: ',model)

# The forecasting procedure is performed by the method predict
forecasts = model.predict(test)
print('forecasts_Chen')
print(forecasts)
print()
print('test')
print(test)
print()
MSE = sum([(forecasts[i]-test[i])**2 for i in range(len(test))])/len(test)

print('U=[',min(test),', ',max(test),']')
print('MSE=', MSE)

#plot
plt.plot(test)

```

ҚОСЫМША Г

Енгізу актісі

Утверждаю
Проректор
По глобальному партнерству и
дополнительному образованию

Е.А. Дайнеко
2023 г.

Утверждаю
Директор ТОО
«Алматынский Институт Технологий»

Н.Б. Богуслаев
2023 г.

АКТ ВНЕДРЕНИЯ

РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ ДИССЕРТАЦИОННОЙ РАБОТЫ

Настоящий акт составлен об использовании и внедрения результатов исследований диссертационной работы Ибраевой Жанар Базарбековны, обучающейся в PhD докторантуре АО «Международный университет информационных технологий» по специальности «Вычислительная техника и программное обеспечение».

Темой диссертационного исследования является «Разработка моделей анализа сетевого трафика и прогнозирование».

Разработанные модели анализа сетевого трафика и прогнозирования используются в работе ТОО «Алматынский Институт Технологий» с января 2023 года.


Полученные модели представляют собой систему анализа, включающую модель прогнозирования с использованием статистического метода ARIMA, модель прогнозирования сетевого трафика на основе ИНС NARX (Nonlinear AutoRegressive Network with exogenous inputs), модель гибридных сетей ANFIS (Adaptive Neuro-Fuzzy Inference System), систем нейро-нечеткого вывода, прогностические модели нечеткой логики.


Данная система анализа предоставляет необходимые сведения для решения задачи управления информационными потоками в сети и *позволят на основе управления предотвратить потери пакетов.*

Модели вполне подходят для интеграции в сетевые устройства как аналитическая программа, работающая на агрегационных данных, что позволит в реальном времени предсказывать поведение объема трафика на коротком промежутке времени и позволяют обосновать перспективные требования к объемам памяти узлового оборудования инфокоммуникационной сети.

Электронная версия описания объекта внедрения и руководство пользователя прилагаются и являются неотъемлемой частью акта.

Представители АИТ:
Директор  Богуслаев Н.Б.

Представители учебной части:
Зав. кафедрой
«Радиотехника, электроника и
телекоммуникации»  Бахтиярова Е.А.

Научный
консультант  Бектемысова Г.У.

ҚОСЫМША Д
Авторлық куәлік №1

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ

РЕСПУБЛИКА КАЗАХСТАН

АВТОРЛЫҚ ҚҰҚЫҚПЕН ҚОРҒАЛАТЫН ОБЪЕКТІЛЕРГЕ ҚҰҚЫҚТАРДЫҢ
МЕМЛЕКЕТТІК ТІЗІЛІМГЕ МӘЛІМЕТТЕРДІ ЕНГІЗУ ТУРАЛЫ

КУӘЛІК

2023 жылғы «9» ақпан № 32481

Автордың (лардың) жөні, аты, әкесінің аты (егер ол жеке басын куәландыратын құжатта көрсетілсе):
БЕКТЕМЫСОВА СУЛНАРА УМИТКУЛОВНА, Ибраева Жанар Базарбековна

Авторлық құқық объектісі: **ЭЕМ-ге арналған бағдарлама**

Объектінің атауы: **Разработка модели прогнозирования с использованием статистического метода Auto-
Regressive Integrated Moving Average**

Объектіні жасаған күні: **07.02.2023**



Құжаттың түпнұсқасын <http://www.kazpatent.kz/ru/avtorlyk>
"Авторлық құқық" бөлімінде тегеурінге бірізді <https://copyright.kazpatent.kz>

Подлинность документа возможно проверить на сайте [kazpatent.kz](http://www.kazpatent.kz)
в разделе «Авторские права» <https://copyright.kazpatent.kz>

ЭЦҚ қол қойылды

Н. Абулкаиров

ҚОСЫМША Е
Авторлық куәлік №2

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ

РЕСПУБЛИКА КАЗАХСТАН

АВТОРЛЫҚ ҚҰҚЫҚПЕН ҚОРҒАЛАТЫН ОБЪЕКТІЛЕРГЕ ҚҰҚЫҚТАРДЫҢ
МЕМЛЕКЕТТІК ТІЗІЛІМГЕ МӘЛІМЕТТЕРДІ ЕНГІЗУ ТУРАЛЫ

ҚУӘЛІК

2023 жылғы «27» сәуір № 35224

Автордың (ардың) жөні, аты, әнестінің аты (егер ол жеке баспан куәлімдерімен қаржыға көрсетілсе):
БЕКТЕМЫСОВА ГҮЛНАРА УМИТКУЛОВНА, Ибраһим Жанар Базарбекқызы

Авторлық құқық объектісі: **ЭЕМ-ге арналған бағдарлама**

Объектінің атауы: **Незеткая модель прогнозирования временного ряда**

Объектіні жасаған күні: **24.04.2023**



Құжат тіркелді/талғанын <http://www.kazpatent.kz> сайтымен
"Авторлық құқық" Бөлімшесінің веб-сайты арқылы <http://copyright.kazpatent.kz>
Полнота документа свободно проверяется на сайте www.kazpatent.kz
и разделе «Авторской правы» <http://copyright.kazpatent.kz>

ЭЦҚ қол қойылды

Е. Оспанов